Ashish Kumar Luhach · Dharm Singh Jat ·
Kamarul Bin Ghazali Hawari · Xiao-Zhi Gao ·
Pawan Lingras (Eds.)

Communications in Computer and Information Science 1075

Advanced Informatics for Computing Research

Third International Conference, ICAICR 2019
Shimla, India, June 15–16, 2019
Revised Selected Papers, Part I

iCAICR

Part 1

COMPUTER SOCIETY OF INDIA
ESTD. 1965
सर्वे भवन्तु सुखिनः

Springer

Communications in Computer and Information Science 1075

More information about this series at http://www.springer.com/series/7899

Ashish Kumar Luhach · Dharm Singh Jat ·
Kamarul Bin Ghazali Hawari ·
Xiao-Zhi Gao · Pawan Lingras (Eds.)

Advanced Informatics for Computing Research

Third International Conference, ICAICR 2019
Shimla, India, June 15–16, 2019
Revised Selected Papers, Part I

Springer

Editors

Ashish Kumar Luhach
Papua New Guinea University
of Technology
Lae, Papua New Guinea

Kamarul Bin Ghazali Hawari
Universiti Malaysia Pahang
Pekan, Pahang, Malaysia

Pawan Lingras
Department of Mathematics
and Computing Science
Saint Mary's University
Halifax, Canada

Dharm Singh Jat
Computer Science Department
Namibia University of Science
and Technology
Windhoek, Namibia

Xiao-Zhi Gao
School of Computing
University of Eastern Finland
Kuopio, Finland

ISSN 1865-0929 ISSN 1865-0937 (electronic)
Communications in Computer and Information Science
ISBN 978-981-15-0107-4 ISBN 978-981-15-0108-1 (eBook)
https://doi.org/10.1007/978-981-15-0108-1

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721,
Singapore

Preface

The Third International Conference on Advanced Informatics for Computing Research (ICAICR 2019) targeted state-of-the-art as well as emerging topics pertaining to advanced informatics for computing research and its implementation for engineering applications. The objective of this international conference is to provide opportunities for researchers, academics, industry professionals, and students to interact and exchange ideas, experience, and expertise in the current trends and strategies in information and communication technologies. Moreover, participants were informed about current and emerging technological developments in the field of advanced informatics and its applications, which were thoroughly explored and discussed.

ICAICR 2019 was held during June 15–16 in Shimla, India in association with Namibia University of Science and Technology and technically sponsored by the CSI Jaipur Chapter, MRK Institute of Engineering and Technology, Haryana, India, and Leafra Research Pvt. Ltd., Haryana, India.

We are very thankful to our valuable authors for their contribution and our Technical Program Committee for their immense support and motivation for making the first edition of ICAICR 2019 a success. We are also grateful to our keynote speakers for sharing their precious work and enlightening the delegates of the conference. We express our sincere gratitude to our publication partner, Springer, for believing in us.

June 2019

Ashish Kumar Luhach
Dharm Singh Jat
Kamarul Bin Ghazali Hawari
Xiao-Zhi Gao
Pawan Lingras

Organization

Conference Chairs

Kamarul Hawari bin Ghazal	Universiti Malaysia Pahang, Malaysia
Dharm Singh	Namibia University of Science and Technology, Namibia

Conference Co-chair

Ashish Kr. Luhach	The PNG University of Technology, Papua New Guinea

Publicity Chair

Aditya Khamparia	Lovely Professional University (Punjab), India

Technical Program Committee

Pawan Lingras (Chair)	Saint Mary's University, Canada
Xiao-Zhi Gao	University of Eastern Finland, Finland
Xin-She Yang	Middlesex University, UK

Program Committee

K. T. Arasu	Wright State University Dayton, USA
Mohammad Ayoub Khan	Taibah University, Saudi Arabia
Rumyantsev Konstantin	Southern Federal University, Russia
Wen-Juan Hou	National Taiwan Normal University, Taiwan
Syed Akhat Hossain	Daffodil University (Dhaka), Bangladesh
Zoran Bojkovic	University of Belgrade, Serbia
Sophia Rahaman	Manipal University, UAE
Thippeswamy Mn	University of KwaZulu-Natal, South Africa
Lavneet Singh	University of Canberra, Australia
Pao-ann Hsiung	National Chung Cheng University, Taiwan
Wei Wang	Xi'an Jiaotong-Liverpool University, China
Mohd. Helmey Abd Wahab	Universiti Tun Hussein Onn, Malaysia
Andrew Ware	University of South Wales, UK
Shireen Panchoo	University of Technology, Mauritius
Sumathy Ayyausamy	Manipal University, UAE
Kamarul Hawari bin Ghazal	Universiti Malaysia Pahang, Malaysia
Dharm Singh	Namibia University of Science and Technology, Namibia

Almir Pereira Guimaraes	Federal University of Alagoas, Brazil
Fabrice Labeau	McGill University, Canada
Abbas Karimi	Islamic Azad University of Arak, Iran
Kaiyu Wan	Xi'an Jiaotong-Liverpool University, China
Pao-Ann Hsiung	National Chung Cheng University, Taiwan
Paul Macharia	Data Manager, Kenya
Yong Zhao	University of Electronic Science and Technology of China, China
Upasana G. Singh	University of KwaZulu-Natal, South Africa
Basheer Al-Duwairi	Jordan University of Science and Technology, Jordan
M. Najam-ul-Islam	Bahria University, Pakistan
Ritesh Chugh	CQ University Sydney, Australia
Yao-Hua Ho	National Taiwan Normal University, Taiwan
Pawan Lingras	Saint Mary's University, Canada
Poonam Dhaka	University of Namibia (UNAM), Namibia
Amirrudin Kamsin	University of Malaya, Malaysia
Ashish Kr. Luhach	The PNG University of Technology, Papua New Guinea
Pelin Angin	Purdue University, USA
Indra Seher	CQ University Sydney, Australia
Adel Elmaghraby	University of Louisville, USA
Sung-Bae Cho	Yonsei University, South Korea
Dong Fang	Southeast University, China
Huy Quan Vu	Victoria University, Australia
Basheer Al-Duwairi	JUST, Jordan
Sugam Sharma	Iowa State University, USA
Yong Wang	University of Electronic Science and Technology of China, China
T. G. K. Vasista	King Saud University, Saudi Arabia
Nalin Asanka Gamagedara Arachchilage	The University of New South Wales, Australia
Durgesh Samadhiya	National Applied Research Laboratories, Taiwan
Akhtar Kalam	Victoria University, Australia
Ajith Abraham	Director at MIR Labs, USA
Runyao Duan	Tsinghua University, China
Miroslav Skoric	IEEE Section, Austria
Al-Sakib Khan Pathan	IIU, Malaysia
Arunita Jaekal	Windsor University, Canada
Pei Feng	Southeast University, China
Ioan-cosmin Mihai	A.I. Cuza Police Academy, Romania
Abhijit Sen	Kwantlen Polytechnic University, Canada
R. B. Mishra	Indian Institute of Technology (IIT-BHU), India
Bhaskar Bisawas	Indian Institute of Technology (IIT-BHU), India

Contents – Part I

Computing Methodologies

Contents – Part II

Hardware

Information Systems

Networks

Software and Its Engineering

Computing Methodologies

Naïve Bayes Model Based Improved K-Nearest Neighbor Classifier for Breast Cancer Prediction

Sonia Goyal and Maheshwar(✉)

Government Boys Senior Secondary School, Holambi Kalan, Delhi, India
sonia.ymca88@gmail.com, maheshwar1524@gmail.com

Abstract. Breast cancer is one of the major cancers that is common to women all over the world. Though, the cancer is curable and can be prevented if it is detected in early stages. In medical science, lots of different strategies have been developed to detect and diagnose the cancer patients. Data mining techniques are no far behind and are widely used to extract information from large databases of the cancer patients to discover some patterns making decisions. Classification is one of the data mining techniques that can be used to classify the data in two stages i.e. benign or malignant. This paper presents the Naïve Bayes improved K-Nearest Neighbor method (NBKNN) for breast cancer prediction and compares the results with traditional classifiers like traditional K-nearest Neighbor and naïve Bayes. In the experiments, the standard dataset used is taken from UCI repository. Sensitivity and specificity have been used as accuracy measures for comparing the results. Experimental results show that proposed classifier is better than traditional classifiers.

Keywords: KNN · Naïve Bayes · NBKNN · Sensitivity · Specificity

1 Introduction

Breast cancer is most common cancer in women worldwide and has overtaken other cancers. In India, 144,937 women were newly detected with breast cancer and 70, 218 women died of breast cancer in 2012 [1]. So, in India, for every two women newly diagnosed with breast cancer, one woman is dying of it. We are now seeing the more number of patients with breast cancer at younger age (in thirties and forties). Figure shows the breast cancer at different age group in India 25 years back and at present [1]. The trend reflects that presently, there is an increasing number of patients in the 25 to 40 years of age which is very disturbing (Fig. 1).

This study shows that how it becomes crucial to detect and diagnose the breast cancer at early stage to reduce the deaths. Breast Cancer Estimation (BSE) and clinical breast examination (CBE) are two common methods used in medical science to screen the breast cancer [2]. Mammography, CAD, Ultrasonography, MRI are some other techniques used in medical science for diagnosing breast cancer [3].

Data mining is the process to discover the relation between items in a dataset and extract the hidden information to make decisions. It has been now used widely in health

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 3–11, 2019.
https://doi.org/10.1007/978-981-15-0108-1_1

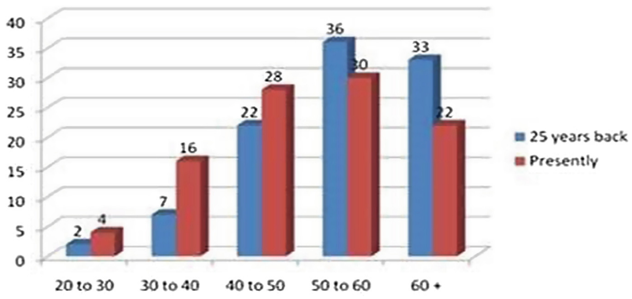


Fig.1 Breast Cancer trend in India.

care industries [4]. Various data mining techniques like classification, clustering are now used to detect diseases and getting the better results than the traditional methods for curing the diseases. Among these techniques, classification is one of the most widely used method in health sector. Various classification techniques like decision tree, KNN, SVM and Naïve Bayes have been used for detection of breast cancer. This paper is focused on using Naïve Bayes model based improved K-Nearest Neighbor method (NBKNN) to classify breast cancer in either benign or malignant categories and compares the results with other classifiers (Table 1).

The whole paper is organized as follows: Sect. 1 discussed the breast cancer severity, trends in India, different methods used in medical science to diagnose it. Section 2 shows the related work done and various approaches used by researchers. Section 3 discusses the proposed approach following with the result analysis and comparison in Sect. 4. Finally, Sect. 5 concludes the paper with some future directions to the work done in this paper.

2 Literature Survey

Paper [5], compares the different data mining method for predicting the breast cancer survivability. In this papers authors have used two popular data mining methods: Artificial Neural Network (ANN) and decision tree. 10-fold cross validation method was used to measure the biasing of the used algorithms. Different data mining classification techniques have been used in [6] for breast cancer diagnosis and prognosis. This paper is actually a review paper and summarizes various review and technical articles on breast cancer. In paper [7], authors have used decision tree classifiers for the diagnosis of breast tumour in medical ultrasonic images. The authors have described a novel computer-aided diagnosis (CADx) system that uses decision tree technique for classification to provide the immediate opinion for the physician. Paper [8] presents a diagnosis system for detecting breast cancer based on RepTree, RBF Network and Simple Logistic. This study uses the simple logistic for reducing the dimension of feature space and Rep Tree and RBF Network model to obtain fast automatic diagnostic system for diseases. Different decision tree algorithms' performance analysis for breast cancer classification is done by authors in [9]. Various classification algorithms like j48, CART, AD tree and BF tree are used on cancer data set for classification and prediction. In Paper [10], Kharya et al., have used the Naïve Bayes classification model for breast cancer detection. The aim of their work is to provide a GUI to enter patient record and probability of breast cancer in women.

In paper [11], breast cancer risk prediction is done by using data mining techniques. In this paper, the authors have used Naïve Bayes classifiers to predict the breast cancer risk in Nigerian patients. In paper [12], Survey of different data mining approaches for healthcare is discussed. In this paper, authors have discussed the various challenges and issues of data mining in healthcare sector. In paper [13], authors have used KNN classifier along with PCA for feature extraction. The authors used the traditional breast cancer dataset for performing the result analysis. Nearest Neighbor classifier has also been used in [14] for prediction of human cancer from imbalanced data. Support Vector Machine (SVM) classifiers have also been used for breast cancer recurrence in [15]. In paper [15], authors have used decision tree, support vector machine (SVM) and artificial neural network (ANN) to develop predictive models. Sensitivity and specificity are used to measure accuracy. In Paper [16] authors evaluate KNN, SVM, logistic regression classifiers using resampling on breast cancer dataset. Predicting breast cancer by combining two or more data mining techniques is also quite prevalent. In paper [20], breast cancer prediction is done using support vector machine and K-Nearest Neighbors. K-Nearest Neighbor classifier is most common and widely used machine learning technique that is used by many researchers [23–27]. A comparison of different machine learning techniques for predicting the breast cancer is done in [21, 22]. The main aim of the authors is to find the best classifier which grants the maximum accuracy and revealed that quadratic support vector machine provides largest accuracy.

3 Proposed Approach

Classification is the data mining technique in which the model is trained using the predefined class labels of each instance and then is used to predict the class label of unseen data. Classification comes under the category of supervised learning. Different classification models have been proposed and well used in different areas like fraud detection, credit card risk evaluation, recommender systems design. The traditional models for classification are decision tree, Naïve Bayes and K-nearest Neighbor. These model as discussed in related work section have been used by different authors in healthcare sector.

Naïve Bayes classifiers are statistical classifiers and use the concept of probability to predict the class of unseen data. Naïve Bayes classifiers assumes the attributes of the dataset independent and use Bayes theorem. An input data, X, will belong to the class for which it is having highest probability. The probability that X belong to a class Ci is derived from Bayes' theorem.

$$P(C_i|\mathbf{X}) = \frac{P(\mathbf{X}|C_i)P(C_i)}{P(\mathbf{X})} \tag{1}$$

Since P(X) is constant for all classes. So main focus is to maximize the

$$P(Ci|\mathbf{X}) = P(\mathbf{X}|Ci)P(Ci) \tag{2}$$

Naïve Bayes classifiers get trapped if the probability of any attribute comes out to be zero. This is handled by using Laplace estimator or Laplace correction.

KNN comes under the category of lazy learners [17]. It keeps on storing the data until it is actually having the input data whose class or label is to be predicted. KNN classifier uses some distance measure to predict the closeness of the unknown tuple to the K training tuples. The most common distance metric is Euclidean distance given below:

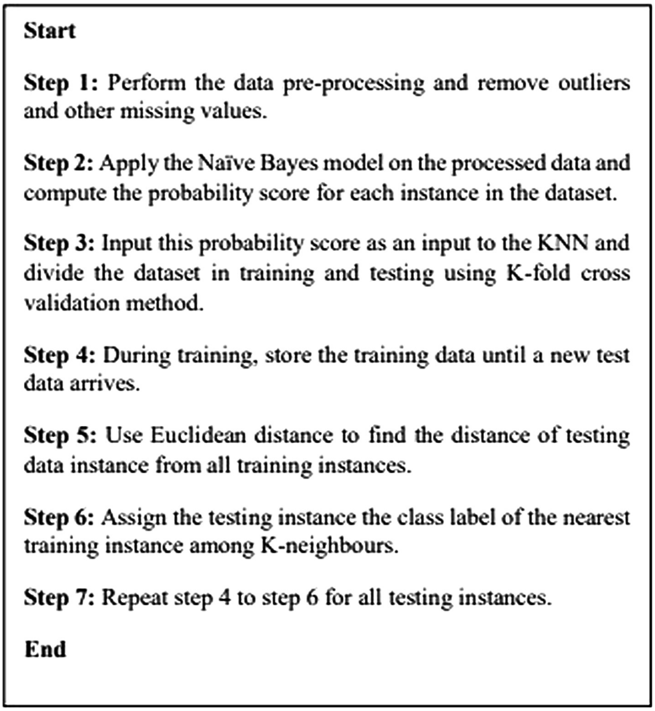
$$D = \sqrt{(x_1 - x_2)^2 + (y_{1-}y_2)^2} \tag{3}$$

KNN is conceptually simple, yet able to solve complex problems. It can work with relatively little information.

In the proposed NBKNN method, both Naïve Bayes and K-Nearest Neighbor (KNN) models are used. First, the Naïve Bayes model is applied on the initial database and find the predictive probability for each instance in the dataset. This predictive probability of each instance is used as an input to the KNN classifier. The distance function used to find the distance of an instance from the other instances is same as given in Eq. (3). The algorithm given below shows the steps involved in the proposed approach.

The instances in the dataset are represented as multidimensional vectors with class labels. During training phase, the model stores the feature vectors for each instance along with the class labels. When a new testing data is arrived, the model compares the testing instance with each of the training instance stored using Euclidean distance.

A 10-fold cross validation technique is used for evaluating the accuracy of each classifier. At each fold only 10% instances are used for testing purpose and rest are used training purpose. All these partitions are mutual exclusive. Before apply model, data cleaning is also done so as to remove outliers and missing values. After preprocessing the dataset, the models are applied one by one (Fig. 2).



Start

Step 1: Perform the data pre-processing and remove outliers and other missing values.

Step 2: Apply the Naïve Bayes model on the processed data and compute the probability score for each instance in the dataset.

Step 3: Input this probability score as an input to the KNN and divide the dataset in training and testing using K-fold cross validation method.

Step 4: During training, store the training data until a new test data arrives.

Step 5: Use Euclidean distance to find the distance of testing data instance from all training instances.

Step 6: Assign the testing instance the class label of the nearest training instance among K-neighbours.

Step 7: Repeat step 4 to step 6 for all testing instances.

End

Fig. 2 Proposed NBKNN Method

4 Result and Discussion

The breast cancer dataset has been taken from UCI machine learning repository [18, 19] as input data. The dataset contains 699 instances and 10 attributes with two classes i.e. either benign or malignant.

Table 1. Dataset details

Dataset details	Number of instances	Number of attributes	Number of classes
Breast Cancer Wisconsin	699	10	2

The study of different classification models provided the basis for comparison among them. Confusion matrix for each classifier gives the effectiveness of each classifier. Sensitivity and specificity are used to calculate the accuracy of the model. Sensitivity and specificity are defined as below:

Sensitivity: proportion of the positive instances that are correctly classified as such.

$$Sensitivity = true\,positives\,/\,(true\,positives\, + false\,negatives) \tag{4}$$

Specificity: proportion of the negative instances that are correctly classified as such

$$Specificity = true\,negatives\,/\,(true\,negative\, + false\,positives) \tag{5}$$

Table 2 shows the comparison of sensitivity and specificity of different classifiers. The comparison study shows that proposed method gives highest sensitivity with 97.4%.

Table 2. Comparing sensitivity and specificity

Algorithm	Sensitivity	Specificity
KNN	96.6	97.5
Naïve Bayes	95.2	97.5
NBKNN	97.4	97.5

The accuracy of each classifiers is calculated using accuracy formula given in Eq. 6.

$$Accuracy\ = TP + TN\,/\,(TP + TN + FP + FN) \tag{6}$$

Where,

TF = True positive, TN = True Negative, FP = False Positive, FN = False Negative

Table 3 below shows the accuracy comparison of KNN, Naïve Bayes and proposed NBKNN method. The study shows that NBKNN method outperforms when compared to others classifiers in terms of accuracy giving an accuracy of 97.5%.

Table 3. Classification Accuracy Comparison

Algorithm	Accuracy
KNN	96.7
Naïve Bayes	95.9
NBKNN	97.5

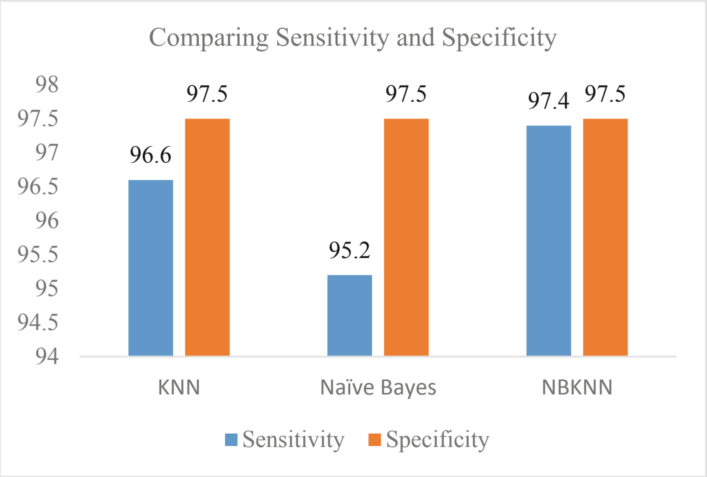


Fig. 3 Comparison of Sensitivity and Specificity for KNN, Naïve Bayes and NBKNN.

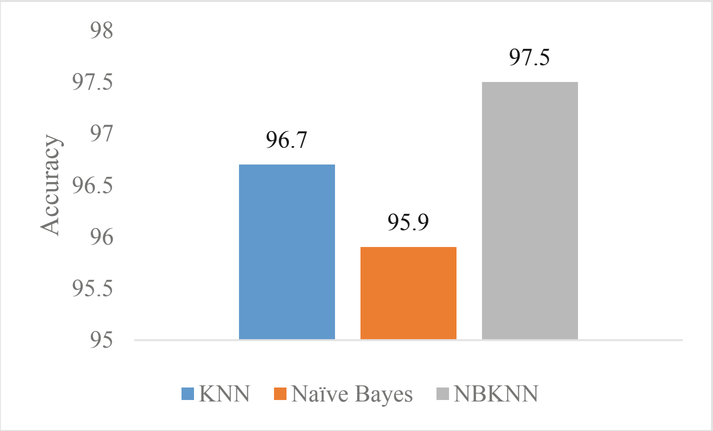


Fig. 4 Comparing Accuracy for KNN, Naïve Bayes and NBKNN

Figures 3 and 4 show the graphical representation of the result to have a more clear view of the result. So, overall result and analysis of different classification models shows that proposed method is giving best result when compared with traditional Naïve Bayes and KNN methods.

5 Conclusion and Future Work

Comparing to all other cancers, breast cancer is one of the major cause of death in women. Early detection of it can significantly reduce the life losses. Various early detection techniques are being used for breast cancer cell detection. This paper also shows the use of machine learning techniques namely KNN, Naïve Bayes along with the proposed method. Results and comparison study shows that proposed NBKNN method gives high accuracy when compared to other classifiers.

This paper can be further extended by considering other classification models like using Support Vector Machine (SVM). Using different methods for features selection and finding the best model among all the classifiers to predict breast cancer is a really challenging task. Using classification models for large dataset and analysing how each classifier behaves can be a future direction to this paper. Moreover, different bio-inspired techniques like Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO) can be used to optimize the result.

References

1. www.breastcancerindia.net/statistics/
2. Ratanachaikanont, T.: Clinical breast examination and its relevance to diagnosis of palpable breast lesion. J. Med. Assoc. Thailand **88**(4), 505–507 (2005)
3. Nover, A.B., Jagtap, S., Anjum, W., et al.: Modern breast cancer detection: a technological review. Int. J. Biomed. Imaging, **2009**, 14 p. (2009). Article ID 902326, https://doi.org/10.1155/2009/902326
4. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques, 2nd edn. Morgan Kaufmann, San Fransisco (2005)
5. Delen, D., Walker, G., Kadam, A.: Predicting breast cancer survivability: a comparison of three data mining methods. Artif. Intell. Med. **34**(2), 113–127 (2005)
6. Gupta, S., Kumar, D., Sharma, A.: Data mining classification techniques applied for breast cancer diagnosis and prognosis. Indian J. Comput. Sci. Eng. (IJCSE) **2**(2), 188–195 (2011)
7. Kuo, W.-J., Chang, R.-F., Chen, D.-R., Lee, C.C.: Data mining with decision trees for diagnosis of breast tumor in medical ultrasonic images. Breast Cancer Res. Treat. **66**(1), 51–57 (2001)
8. Chaurasia, V., Pal, S.: Data mining techniques: to predict and resolve breast cancer survivability. Int. J. Comput. Sci. Mob. Comput. **3**(1), 10–22 (2014)
9. Venkatesan, E., Velmurugan, T.: Performance analysis of decision tree algorithms for breast cancer classification. Indian J. Sci. Technol. **8**(29), 1–8 (2015)
10. Kharya, S., Agrawal, S., Soni, S.: Naive Bayes classifiers: a probabilistic detection model for breast cancer. Int. J. Comput. Appl. **92**(10), 0975–8887 (2014)
11. Williams, K., Idowu, P.A., Balogun, J.A., Oluwaranti, A.I.: Breast cancer risk prediction using data mining classification techniques. Trans. Netw. Commun. **3**(2), 01–11 (2015)
12. Tomar, D., Agarwal, S.: A survey on data mining approaches for healthcare. Int. J. Bio-Sci. Bio-Technol. **5**(5), 241–266 (2013)
13. Ramadevi, G.N., Rani, K.U., Lavanya, D.: Importance of feature extraction for classification of breast cancer datasets—a study. Int. J. Sci. Innovative Math. Res. **3**(2), 368–763 (2015)

14. Majid, A., Ali, S., Iqbal, M., Kausar, N.: Prediction of human breast and colon cancers from imbalanced data using nearest neighbor and support vector machines. Comput. Meth. Programs Biomed. **113**(3), 792–808 (2014)
15. Ahmad, L.G., Eshlaghy, A.T., Poorebrahimi, A., Ebrahimi, M., Razavi, A.R.: Using three machine learning techniques for predicting breast cancer recurrence. J. Health Med. Inform. **4**(124), 3 (2013)
16. Ramadevi, G.N., Rani, K.U., Lavanya, D.: Evaluation of classifiers performance using resampling on breast cancer data. Int. J. Sci. Eng. Res. **6**(2) (2015)
17. Kotsiantis, S.B., Zaharakis, I., Pintelas, P.: Supervised machine learning: a review of classification techniques. Emerg. Artif. Intell. Appl. Comput. Eng. **160**, 3–24 (2007)
18. Mangasarian, O.L., Wolberg, W.H.: Cancer diagnosis via linear programming. SIAM News **23**(5), 1–18 (1990)
19. Wolberg, W.H., Mangasarian, O.L.: Multisurface method of pattern separation for medical diagnosis applied to breast cytology. In: Proceedings of the National Academy of Sciences, U.S.A., vol. 87, pp. 9193–9196, December 1990
20. Islam, Md.M., Iqbal, H., Haque, Md.R., Hasan, Md.K.: Prediction of breast cancer using support vector machine and K-nearest neighbors. In: 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), pp. 226–229. IEEE (2017)
21. Obaid, O.I., Mohammed, M.A., Ghani, M.K.A., Mostafa, S.A., Taha, F.: Evaluating the performance of machine learning techniques in the classification of Wisconsin Breast Cancer. Int. J. Eng. Technol. **7**(4.36), 160–166 (2018)
22. Yue, W., Wang, Z., Chen, H., Payne, A., Liu, X.: Machine learning with applications in breast cancer diagnosis and prognosis. Designs **2**, 13 (2018). https://doi.org/10.3390/designs2020013
23. Gupta, A., Kaushik, B.N.: Feature selection from biological database for breast cancer prediction and detection using machine learning classifier. J. Artif. Intell. **11**, 55–64 (2018)
24. Chawla, S., Kumar, R., Aggarwal, E., Swain, S.: Breast cancer detection using K-nearest neighbor algorithm. Int. J. Comput. Intell. IoT **2**(4) (2018)
25. Cherif, W.: Optimization of K-NN algorithm by clustering and reliability coefficients: application to breast-cancer diagnosis. Proc. Comput. Sci. **127**, 293–299 (2018)
26. Sun, J., et al.: Predicting medical conditions using k-nearest neighbors. University of Nevada, Las Vegas (2017)
27. Alarabeyyat, A., Alhanahnah, M.: Breast cancer detection using k nearest neighbor machine learning algorithm. In: 2016 9th International Conference on Developments in eSystems Engineering (DeSE), pp. 35–39. IEEE (2016)

Evaluation of Model Using J-48 and Other Classifier on Kddcup99 Through Performance Metrics

Saby Singhal(✉) and Pradeep Yadav

Institute of Technology and Management Group of Institutions, Gwalior, India
saby.singhal048@gmail.com,
er.pradeepyadav0610@gmail.com

Abstract. Intrusion detection system (IDS) is one of the complete solution against harmful attacks, as well as the attackers always keep changeable their tools and techniques. However, implementing an approved intrusion detection system is also a challenging task. In this paper, we have taken the dataset of KDDCUP99. KDD cup99 is the most widely used data set for the evaluation of the system in anomaly based detection. This paper we have used twelve attributes from the KDD 99 dataset and weka tool for simulation. In this paper J-48 with other classifier shows the better results in terms of precision and recall metrics. It achieves to compute several performance metrics are available for the measurement in order to evaluate the selected classifiers.

Keywords: KDDCUP99 · IDS · Weka tool · J-48

1 Introduction

Enormous Data is the information that are hard to store, oversee, and break down utilizing customary database and programming strategies. Enormous Data incorporates high volume and speed, and furthermore assortment of information that requirements for new procedures to manage it. Interruption recognition framework (IDS) is equipment or programming screen that dissects information [1] to distinguish any assault toward a framework or a system. Conventional interruption discovery framework methods make the framework increasingly mind boggling and less effective when managing Big Data, since its examination properties process is perplexing and take quite a while.

In Today's reality, a few associations store their information in a few different ways. These associations's solitary necessity is to protect their private and authority information from the intruder and outside, inner intruder. It might likewise be conceivable that a few approved clients may release the information of the association for any reason. Continuously, it is trying to perceive the aggressor since copy IP and assault bundles can make. Methods utilized before like firewall, and IDS was not ready to distinguish the ongoing aggressors which happened without the administrator without his insight [2]. A PC organize is the mix of a lot of equipment and programming. The two segments have their dangers, vulnerabilities and security issues.

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 12–19, 2019.
https://doi.org/10.1007/978-981-15-0108-1_2

The assault in the programming makes the information powerless. The ones who know programming and frameworks can without much of a stretch discover the different exercises performed on the frameworks utilizing log records.

As there exists different kinds of expert system, some of them are rule based expert system, reasoning based expert system and framework expert system [3]. There is another kind of method that is known as pattern matching which is used for intrusion detections.

The IDS has three methods for detecting attacks. First one is signature based detection, anomaly based detection and hybrid based detection. The signature based detection is used to detect known attacks by using signature of those attacks [4]. Anomaly based detection compares current user activities against pre-defined profiles, that is used to detect abnormal behaviors [5, 6].

They can help in guaranteeing security. The issue arrives when individuals try not to have any basic information of programming, and their framework gets assaulted by the gatecrashers, and they can't discover out the issue. There are different kinds of attacks. Be that as it may, the most difficult one is to discover the insider/interior attacks. The system security is a territory where each client needs his frameworks to shield from all the harmful attacks (inward or outer assaults). The outer attacks by the interlopers can be recognized by IDS, and IIDS can distinguish the inner unauthorized person.

2 Literature Survey

2.1 In This Section Literature Review Is Being Discussed

Jianguo and pei [7] in his researched work asserted that designation of expert system that is intrusion detection system which is used to check and supervised equipment of underground railway transit system and all about its environment. He also focused on the designing of engine of intrusion detection system which is based on expert system.

Akash and prachi [8] asserted that in their researched work they provided a security strategy for endeavoring to recognize different types of attacks. He checked on grunt as intrusion misuse detection system framework just as NETPAD that is based on statistical algorithm of anomaly.

Arkadiusz and Grzegorz [9] asserted that intrusion detection system plays an crucial role in the field of cyber security area. They asserted about the intrusion detection system that is anomaly based. In the network area anomaly based intrusion detection refers to the finding of problem about irregular events that do not adjust to the normal ordinary patterns. For detection of anomalies various types of security system performs clustering and classification algorithm.

Liwei Kuang asserted that in his researched work the steadfastness of an Intrusion Detection System (IDS) depends on two components capacity to identify interruptions [10] and survivability in antagonistic conditions. He proposed a Dependable Network Intrusion Detection System which is based on the Isolation measure K-Nearest Neighbours (CSI-KNN) algorithm.

3 Type of Attacks

3.1 This Section Consists of Following Types of Attacks in an IDS Network

A. **Active attack:** Active attack involves creation of false statement including some changes of the data stream. Active attack refers to the exploitation in the network security in which the attacker making an attempt to break down into the network or system. At the time of active attack, the intruder will create changing in data potentially. Masquerades, distributed DOS, replay session are the types of active attacks.

B. **Passive Attack:** The passive attacks does not making an effect on the system resources and it is a kind of network attack in which system is sometimes scanned for open ports and monitors vulnerabilities without any interaction. The types of passive attacks are encryption, scanning and tapping. An attack can be created and caused by the insider and outsider of the organization.

C. **Insider Attack**: An attack can be created and caused by the insider and outsider of the organization. An insider attack is a type of destructive attack that is carried out by a person with authorized system access into the system and computer network. The types of insider attacks are UBS and PaineWebber.

D. **Outsider Attack**: An outsider attack is initiated by the illegal use of the computer system. The outsiders attack do not have directly access to any one of the authorized nodes in the network. The examples of the outsider attacks are spoofing, spin and spam.

E. **Denial of service (DOS)**: Denial of service is a such type of cyber attack in which the malefactor seeks to make a resources of network and machine unavailable to its intended users. The denial of service attack is mainly characterized by the attackers for the prevention of recognized use of a service. The Dos attacks are mainly in two forms- one is those that flood services and other is those that crash services. The most common examples of Dos attacks are flooding in the network, prevention in the access of individuals, disrupting in the connections.

4 Types of Dataset

4.1 Datasets and Tools

Here in this work we have taken the dataset of KDDCUP99. .From the last decade the KDD CUP 99 has been become the point of interesting for many researchers in the field of intrusion detection. KDD cup99 is the most widely used data set for the evaluation of the system in anomaly based detection. Investigation on the anomaly based detection we have used KDD data set. To apply KDD there are two types of common approaches. In the first approach KDD 99 training portion is employed for sampling both the trained sets and test sets. However, in the second approach, the training samples are randomly collected from the KDD train set, while the samples for testing are arbitrarily selected from the KDD test set. It consists of mainly 42 attributes but here class attribute should not be considered as an attribute, because it shows whether the given instance is a normal one or an attack. Therefore here only 41 attributes should be considered.

In this Table 1 we have used twelve attributes from the KDD 99 dataset. Here in the below table all twelve attributes are given.

Table 1. Twelve attributes of the KDD99 Dataset

Attributes name	Description
Duration	It describes time about the connection
Protocol type	It demonstrates the protocol that is used in the connection
Src BYTES	Total no. of data types transferred from the destination to source in single connection
Dst BYTES	Total no. of data types transferred from destination to source in single connection
HOT	It shows the no. of hot indicators in the content such as creating and executing programs
Num Failed Logins	It shows the total failed login attempts
Logged in	It shows login status: it shows 1 if successful logged in otherwise it shows 0
Count	It shows the total no. of connections to the same destination host as the current connections as the past in the past 2 s
Urgent	It shows the number of urgent packets
Class	Normally or anomaly
Srv count	Total no.of connection to the same service no. or port no. as the current connection in the past two seconds
Wrong_fragment	It shows the number of wrong fragments

5 Performance Metrics

There are different types of metrics are available for the measurement or to calculate the performance. Some of them are

5.1 Precision Metric

$$Precision = \frac{L}{L+S} \tag{1}$$

where L represents True Positive and S represents False Positive value respectively

5.2 Recall Metric

$$\text{Re}\mathit{call} = \frac{L}{L+N} \tag{2}$$

Here L represents True positive and N represents false negative value respectively.

Table 2. Analysis between different types of classifier

Researcher	Type of classifier	Database	Performance metric								
			Accuracy	Precision	Recall	F-measure	Roc	Sensitivity	Training time (s)	Prediction time	Specificity
Mustapha and Salah [13]	SVM	UNSW-NB15	92.28	NA	NA	NA	NA	92.13	38.91	0.20	91.15
Salah and Idhammad [13]	Naïve Bayes	UNSW–NB15	74.19	NA	NA	NA	NA	92.16	2.25	0.18	67.82
Sahilpreet and Meenakshi [14]	Multi Layer Perceptron	NSLKDD	NA	NA	NA	NA	NA	NA	26.72	NA	NA
Zahangir and Tarek [15]	Deep Neural Network	KDD99	90.12	NA	NA	NA	NA	NA	NA	NA	NA
Vimalkumar and Radhika [16]	DNN	KDD99	79.86	NA	60.62	NA	NA	NA	NA	40.83	89.82
Vimal and Radhika [16]	SVM	Syncrophasor Dataset	75.92	NA	73.26	NA	NA	NA	NA	15.69	76.01
Rohit and Suman [17]	J48	Balance Scale	NA	73.0	77.0	0.75	0.81	NA	NA	NA	NA
Rohit and Suman [17]	J48	Diabetes	NA	74.0	74.0	0.74	0.75	NA	NA	NA	NA
Kanupriya and Ritu [18]	DNN	KDD99	98	98.2	98	98	99.0	NA	NA	NA	NA
Suad and Fadl [1]	SVM	NA	NA	94.36	94.36	NA	96.80	NA	25.5	1.37	NA
Suad and Amal [1]	Chi-SVM	KDD99	NA	96.24	96.24	NA	99.55	NA	10.79	1.21	NA
Govind and Manish [19]	SVM	KDD99	78.84	NA	NA	NA	NA	90.53	479.124	10.08	30.48
Govind and Manish [19]	Logistic Regression	KDD99	91.56	NA	NA	NA	NA	89.91	289.105	12.9	98.4
Manish and Priyanka [20]	Logistic Regression	KDD99	91.64	NA	NA	NA	NA	90.02	341.66	21.23	98.31
Raushan and Govind [20]	SVM	KDD99	92.13	NA	NA	NA	NA	92.17	561.04	26.36	91.18
Proposed Method	J-48	KDD99	NA	98.4	98.3	98.4	99.9	NA	6.15	NA	NA

5.3 F-Measure

$$F_{Measure} = 2 \times \frac{P \times R}{P + R} \tag{3}$$

Where, P = Precision R = Recall.

5.4 Accuracy

$$Accuracy = \frac{L + T}{L + T + S + N} \tag{4}$$

Where L represents true positive, T represents true negative,
S represents false positive and N represents false negative.

6 Comparisons Made by Different Researchers

6.1 Comparisons

In the above comparison table we have compared the different types of performance metrics by various researchers. In this Table 2 different classifiers are classified on the basis of real time dataset (Unsw-nb15, Kdd99, NslKdd, Synchrophasor, BalanceScale, Diabetes) Here different types of performance metric have been taken such as accuracy, sensitivity, specificity, precision recall and F-measure. Training time and Prediction time are also compared on the basis of given parameters. Where NA represents not available in Table 2. In this the evaluated value of these parameters of performance metric are calculated in the percentage form.

7 Proposed Work

7.1 Work

In this paper we have taken twelve attributes of Kddcup dataset. Attributes are selected according to the requirement in the IDS network. Here in the Table 4, three classifiers are used to create the models. Multilayer perceptron, J48 and Logistic regression are some of the classifiers that we have taken. In multilayer perceptron we have taken the learning rate at 0.3 and 0.2 respectively. The results that are obtained using learning rate 0.2 shows the better precision and recall value as compared to the other (Table 3).

Table 3. Outcome of multilayer perceptron classifier

Classifier (Multilayer Perceptron)	Precision	Recall	ROC	Time taken (sec)
Learning Rate (0.3)	94.9	94.9	99.1	392.42
Learning Rate (0.2)	95.3	95.2	99.1	394.31

The precision and recall values that are obtained using j-48 classifier in Table 4 shows the better results as compared to the other classifiers. Also the time taken to build the model is less than the other classifiers.

Table 4. Comparision among proposed and existing classifier

Classifier	Precision	Recall	ROC	Time taken (sec)
Multilayer Perceptron	95.3	95.2	99.1	394.31
J-48	98.4	98.3	99.9	6.15
Logistic Regression	93.1	92.9	97.7	11

8 Conclusion

Here the comparisons between the previous work done by the different researchers the proposed work is shown. The proposed method shows the best results among the existing methods. Therefore the J-48 classifier shows the promising results as compared to the others methods.

References

1. Othman, S.M., Ba-Alwi, F.M., Alsohybe, N.T., Al-Hashida, A.Y.: Intrusion detection model using machine learning algorithm on Big Data environment. J. Big Data **5**(1), 34 (2018)
2. Borkar, A., Donode, A., Kumari, A.: A survey on intrusion detection system (IDS) and internal intrusion detection and protection system (IIDPS). In: 2017 International Conference on Inventive Computing and Informatics (ICICI), pp. 949–953. IEEE (2017)
3. Zhao, W.: The study and implement of the high-speed EMU's operation and maintenance decision knowledgebase, Jiaotong University, Beijing (2016)
4. Khanna, R., Liu, H.: System approach to intrusion detection using hidden markov model. In: Proceedings of the 2006 International Conference on Wireless Computations and Mobile Computing, pp. 349–354. ACM (2006)
5. Zolotukhin M, Hämäläinen T, Kokkonen T, et al. Analysis of HTTP requests for anomaly detection of web attacks. In: 2014 IEEE 12th International Conference on Dependable, Autonomic and Secure Computing (DASC), pp. 406–411. IEEE (2014)
6. Eason, G., Noble, B., Sneddon, I.N.: On certain integrals of Lipschitz-Hankel type involving products of Bessel functions. Phil. Trans. Roy. Soc. Lond. **A247**, 529–551 (1955)
7. Maxwell, J.C.: A Treatise on Electricity and Magnetism, 3rd edn. vol. 2, pp. 68–73, Clarendon, Oxford (1892)
8. Jacobs, I.S., Bean, C.P.: Fine particles, thin films and exchange anisotropy. In: Rado, G.T., Suhl, H., (eds.) Magnetism, vol. III, pp. 271–350. Academic, New York (1963)
9. Elissa, K.: Title of paper if known (unpublished)
10. Nicole, R.: Title of paper with only first word capitalized. J. Name Stand. Abbrev. (in press)
11. Yorozu, T., Hirano, M., Oka, K., Tagawa, Y.: Electron spectroscopy studies on magneto-optical media and plastic substrate interface. IEEE Trans. J. Magn. Jpn. **2**, 740–741 (1987). Digests 9th Annual Conference on Magnetics Japan, p. 301 (1982)
12. Young, M.: The Technical Writer's Handbook. University Science, Mill Valley (1989)

13. Belouch, M., El Hadaj, S., Idhammad, M.: Performance evaluation of intrusion detection based on machine learning using Apache Spark. Proc. Comput. Sci. **127**, 1–6 (2018)
14. Singh, S., Bansal, M.: Improvement of intrusion detection system in data mining using neural network. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(9) (2013). kkyuu
15. Alom, Md.Z., Taha, T.M.: Network intrusion detection for cyber security on neuromorphic computing system. In: 2017 International Joint Conference on Neural Networks (IJCNN). IEEE (2017)
16. Vimalkumar, K., Radhika, N.: A big data framework for intrusion detection in smart grids using Apache Spark. In: 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE (2017)
17. Arora, R.: Comparative analysis of classification algorithms on different datasets using WEKA. Int. J. Comput. Appl. **54**(13) (2012)
18. Arora, K., Chauhan, R.: Improvement in the performance of deep neural network model using learning rate. In: 2017 Innovations in Power and Advanced Computing Technologies (i- PACT). IEEE (2017)
19. Gupta, G.P., Kulariya, M.: A framework for fast and efficient cyber security network intrusion detection using apache spark. Proc. Comput. Sci. **93**, 824–831 (2016)
20. Kulariya, M., et al.: Performance analysis of network intrusion detection schemes using Apache Spark. In: 2016 International Conference on Communication and Signal Processing (ICCSP). IEEE (2016)

Malaria Detection Using Custom Convolutional Neural Network Model on Blood Smear Slide Images

Rahul Kumar, Sanjay Kumar Singh(✉), and Aditya Khamparia

Department of Computer Science and Engineering,
Lovely Professional University, Jalandhar, Punjab, India
rahul.cse.ccu@gmail.com, sanjayksingh.012@gmail.com,
aditya.khamparia88@gmail.com

Abstract. Malaria is a life-threatening disease and is a concern of global health threat. The standard way of diagnosing the malaria is by visually examining them under microscope and is very lengthy and tedious task. In this paper, the authors has purposed custom Convolutional Neural Network model for detection of malaria on blood smear slide images. The images are available on website of U.S. National Library of Medicine. The proposed model uses various deep learning layers like convolution layer, max pooling layer, batch normalization layer and fully connected layer. The model achieves 99.71% accuracy in training and 98.23% accuracy on the test data. The study purposes a robust CNN models for detecting infected cell. The training and testing were performed on the 27,558 single cell images.

Keywords: Convolutional neural network · Deep learning · Malarial machine learning

1 Introduction

Malaria is a very serious life threatening, mosquito borne blood disease affecting both humans and animals. The Anopheles, a mosquito of a genus which is most prominent in the warmer countries, transmits the malaria parasite of the genus Plasmodium to the human. According to the WHO report in 2000, the number of malaria cases are estimated to be 262 million globally and which in turn counts to 8,39,000 deaths. And in 2017, the malaria cases had been decreased to 219 million, shrinking down the death count to 4,35,000. Majority of these deaths had been in warmer countries and in Sub-Saharan Africa regions. This can be concluded from the fact that the environmental conditions facilitate the growth of the malaria and also due to the poor socio-economic conditions of country which inhibits the access to the better healthcare program and resources [1]. Malaria being caused by a Plasmodium parasite infection, there are many types of Plasmodium parasites out of which only five infects human, namely, Plasmodium malariae, Plasmodium falciparum, Plasmodium knowlesi, Plasmodium ovale, and Plasmodium vivax. Among all, the one being lethal is Plasmodium falciparum [2].

Dealing with malaria, it becomes extensively difficult to identify the uninfected cells and infected cells. The mortality rate for the malaria can be increased by

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 20–28, 2019.
https://doi.org/10.1007/978-981-15-0108-1_3

increasing the diagnosis accuracy or by eradicating malaria. The latter is still on long road and there are countries trying to eliminate malaria by devising better malaria surveillance system [3] but for our concern, increasing the diagnosis accuracy will suffice the goal of this research paper.

The common traditional way of detecting malaria is to take the blood sample under the microscope observe the color and morphological changes in erythrocytes and determine if it is infected or not. The traditional process is very tedious and requires substantial amount of pathological lab experience. This being perceptual problem is prone to misdiagnosis. Using the microscope imagery techniques, it requires expertise in the domain and if person lacking the required training, it will certainly yield out irrelevant or ambiguous information, which in turn will certainly misdiagnose the malaria.

There are several studies pertaining to the malaria diagnosis and methodology [4–8]. Pattanaik et al. [7] used the object detection technique with different level of procedure, they are Kernel-based detection and Kalman filtering process to figure out the infected malaria cell. Hendrawan et al. [8] used the color image segmentation using cascading method with various colour normalization process, edge enhancement, fuzzy c-means clustering and malaria infection for four plasmodium types from 574 images.

2 Related Work

Saraswat et al. [9], proposed a system automating the manual work done by a technician in order to cut down the human error and increased the accuracy of the malaria diagnosis. They are using the HSV segmentation technique. With this technique they have accomplish a 98.5% sensitivity, 75% specificity, 95% accuracy and 95.7% ppv. Liang et al. [15], used the custom designed CNN model and from Focus-stack of blood smear Images acquired using Custom-built slide scanner, they have acquired accuracy in terms of sensitivity as 97.06% and 98.50% in terms of specificity. An evaluation paper [10] gives a great insight while comparing the well-known Convolution neural network models, namely the LeNet, AlexNet and GoogLeNet. The simulation accuracy for all the models have reached an accuracy above 95%. The details for each CNN model with their accuracy is given Table 1.

Table 1. LetNet-5, AlexNet and GoogLeNet comparison with their accuracy

CNN	LeNet-5	AlexNet	GoogLeNet
Number of layers	4	8	22
Number of convolutional layers	3	5	21
Kernel size	5	11, 5, 3	7, 1, 3, 5
Number of fully connected layer	1	3	1
Number of parameters	3628072	20176258	5975602
Dropout	No	Yes	Yes
Data augmentation	No	Yes	Yes
Inception	No	No	Yes
Local response normalization	No	Yes	Yes
Accuracy	96.18%	95.79%	98.13%

3 Proposed Methodology

In this section, proposed method for detecting the infected and uninfected cells is discussed. We have used the deep learning, especially the CNN models. The custom CNN models were designed and the simulation was performed on system running Ubuntu 19.04 with following configuration: Nvidia GTX 1050 4 GB GPU, 8 GB RAM and Intel Core i5-8300H processor. Deep Learning, a machine learning subfield is a learning technique based on data representations. There are three types of learning: supervised, semi-supervised and the unsupervised. For feature extraction and transformations, it uses the cascade of layers comprising the nonlinear processing units. The, output of the layer serve as the input for the successive layer. As the number of layers increases, the ability to learn the abstract feature increases and this is carried upon only by the better representational layer. A hierarchy of concepts can be concluded from the different levels. The number of successive layers in the model counts for the "deep" in the deep learning and it also exhibits how the data is transformed.

3.1 Convolutional Neural Networks

Convolutional neural network is an artificial neural network which are inspired by the animal visual system [11]. CNN architecture mainly comprises three main types of layer, they are namely the Convolutional Layer, the Pooling Layer and the Fully Connected Layer. CNNs has this ability to extract unique feature without losing any spatial correlations. Each and every neuron in one layer is connected to the every neuron in the another layer. And because of this "fully-connectedness", CNN is very prone to overfitting. CNN, because of its architecture, requires less preprocessing time than other Image classification algorithm. CNN consists of input, multiple hidden layers and output layer. The hidden layers in CNN generally comprises of Convolution layers, Activation Layer (generally ReLU), pooling layers, fully connected layers and normalization layers [12]. The Fig. 1 shows diagrammatic representation of the CNN model [13].

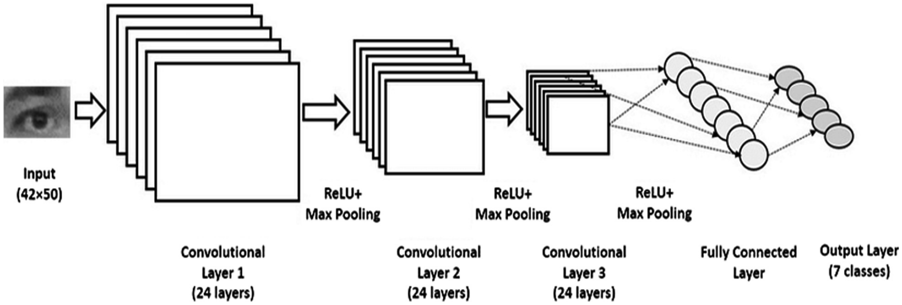


Fig. 1. CNN model

The CNN's using filters is able to capture minute image features as possible. As, the architecture of the CNN model get complex, this demands the need of more powerful computational resources. Use of high-performance GPUs, cluster computing and the High- Speed Computing significantly reduces the training time. Since, the CNN models are data hungry, a deficit in data will not be suffice to train a large deep learning model with many parameters. Transfer learning can be an alternative where the time for training can cut off cost of performance.

3.2 Configuration of CNN Model

The basic modus operandi of the deep learning is to apply a series of transformation from the input layer to the output layer through the multiple hidden layers in between. This optimal mapping of the input data to the output data through the multiple hidden layer is carried upon by a process called the back-propagation. The partial derivative of the output layer supplies the parameter for the partial derivative of the input layer. This in turn helps recursively compute the changes in one layer, by analyzing the changes made to the layer connected to it. Two inverse computation takes place in CNN: The feed-forward and the back-propagation [14]. The feedforward propagation is used to propagate in the forward direction of the model and compute the output for each layer and from the preceding layer, all the weighted sum of the input is aggregated and then the non-linear activation function is applied [15]. The activation function, generally Rectified Linear Unit (ReLU) is mainly used for the hidden layer, but hyperbolic tangent (tanh), logistic function etc. is also used. Back to fine tuning each layer, back propagation is applied which works by optimizing the weight of each layer. By applying above propagation mechanism, the CNN architecture feeds in data and propagates it through the entire network.

3.3 CNN Architecture for Malaria Image Classification

Our custom CNN model comprises of 12 layers. Each layer allows enhanced learning and allow better representation. The proposed CNN model takes in images with $50 \times 50 \times 3$ as input shape. The first convolution layer has 32 filters, with 3×3 kernel size and "ReLU" as activation function followed by Max Pooling layer with 2×2 pooling size and finally the Batch Normalization layer which reduces the Covariance shift, which in turn helps determine the value shift in hidden layer. Also, the Batch Normalization allows each layer to learn a little bit more independently than other layers. With every block, the number of filters are increasing, this helps extract high-level abstract features. This is in accordance to the VGG16 architecture but to reduce the cost of computation, extra parameters had been cut off. The proposed CNN model has 11,44,258 parameters compare to actual VGG16 model which has 14,36,65,448 parameters.

4 Results and Discussion

We archived the Image dataset of the malaria from the U.S. National Library of Medicine website [16, 17]. It comprises of the uninfected and parasitized images of 27,558 cells. The ratio to the uninfected and parasitized images is 1:1. The images in the dataset has width and height as 121 and 106 pixels respectively. All the images are normalize, with height and width reduced to 50 × 50 pixels for both the training and testing, with all three colour channels.

4.1 Data Preprocessing

The images in the dataset is resized to the 50 × 50 pixels to generalize the input. Before feeding the data to the model, we have applied normalization, gamma correction and logarithmic correction to improve brightness and adjust contrast as shown in Fig. 2.

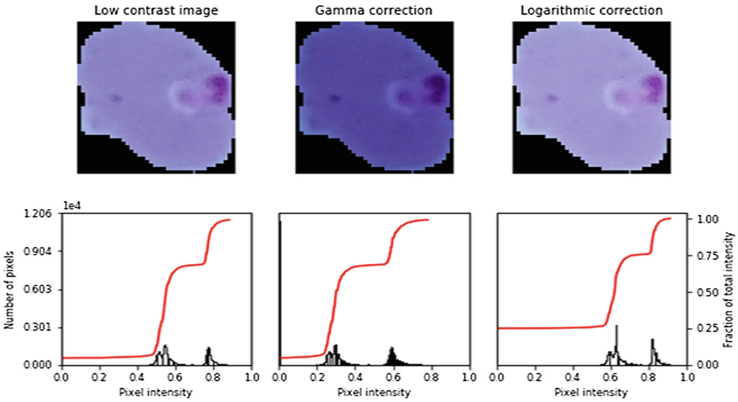


Fig. 2. Image correction

We have used two techniques first without using data augmentation and second by using the augmentation. The latter showed increased in accuracy and image augmentation parameter like horizontal flip, vertical flip, width shift, height shift, fill mode, zoom range, rotational range has been applied. The Fig. 3 demonstrates the proposed CNN architecture for the proposed methodology.

Image outputs have been visualized from different layers as shown in Figs. 4, 5, and 6 respectively. The model outperforms the accuracy achieved by [15] and with less layers compared to the proposed model as shown in Fig. 7. The model uses 1144258 parameters which is relatively very less compared to the well-known model like the LeNet and the AlexNet and still manages to achieve higher accuracy as shown in Table 2. This being comprised of less parameter, cut off the computational cost and makes the algorithm faster as compared to the other known model.

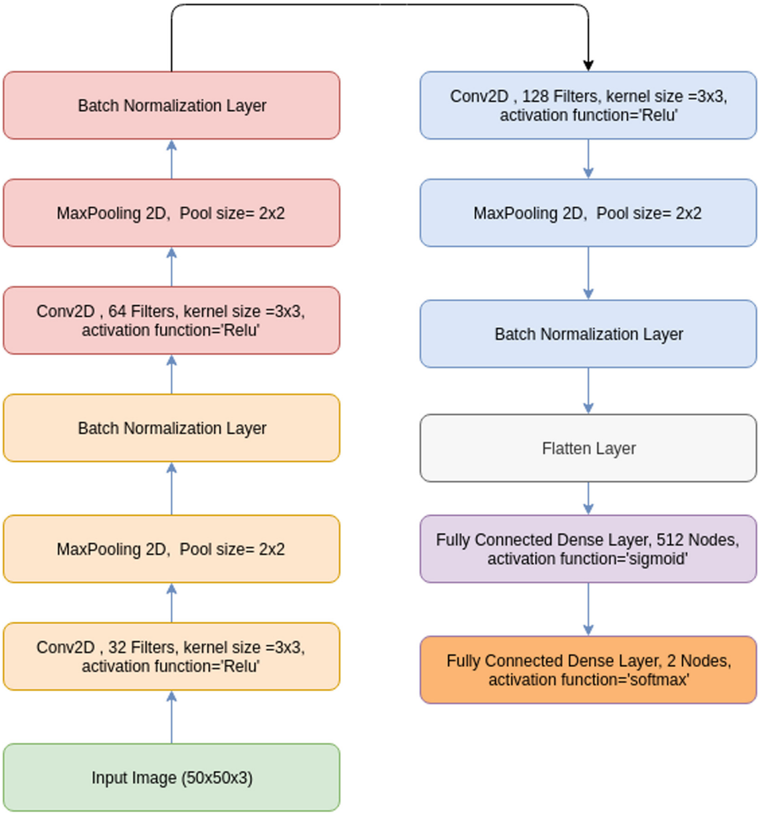

Batch Normalization Layer
MaxPooling 2D, Pool size= 2x2
Conv2D , 64 Filters, kernel size =3x3, activation function='Relu'
Batch Normalization Layer
MaxPooling 2D, Pool size= 2x2
Conv2D , 32 Filters, kernel size =3x3, activation function='Relu'
Input Image (50x50x3)
Conv2D , 128 Filters, kernel size =3x3, activation function='Relu'
MaxPooling 2D, Pool size= 2x2
Batch Normalization Layer
Flatten Layer
Fully Connected Dense Layer, 512 Nodes, activation function='sigmoid'
Fully Connected Dense Layer, 2 Nodes, activation function='softmax'

Fig. 3. Proposed architecture

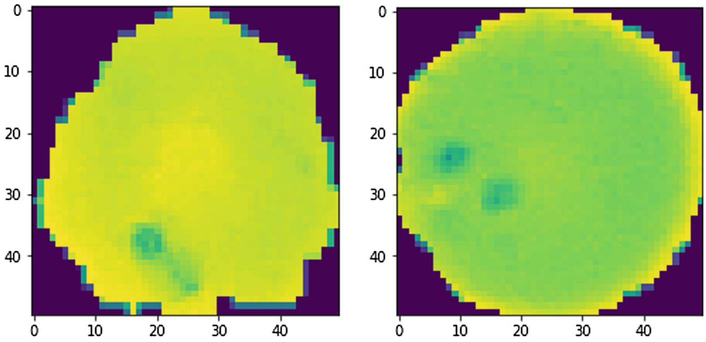


Fig. 4. Image output from first layer

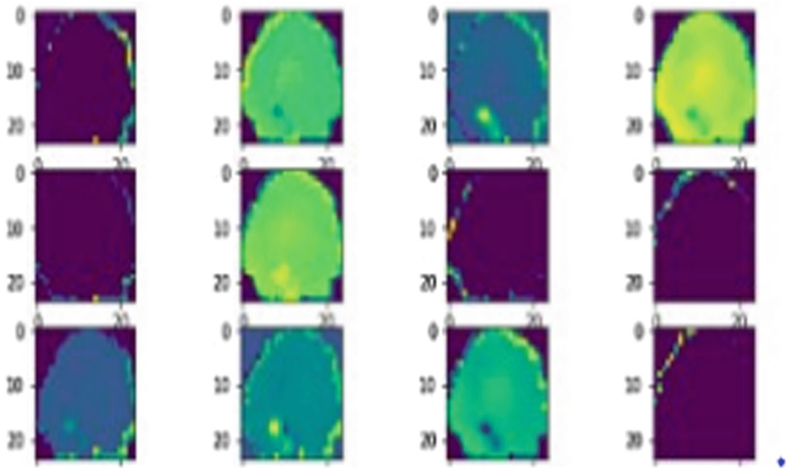


Fig. 5. Image output from the second layer

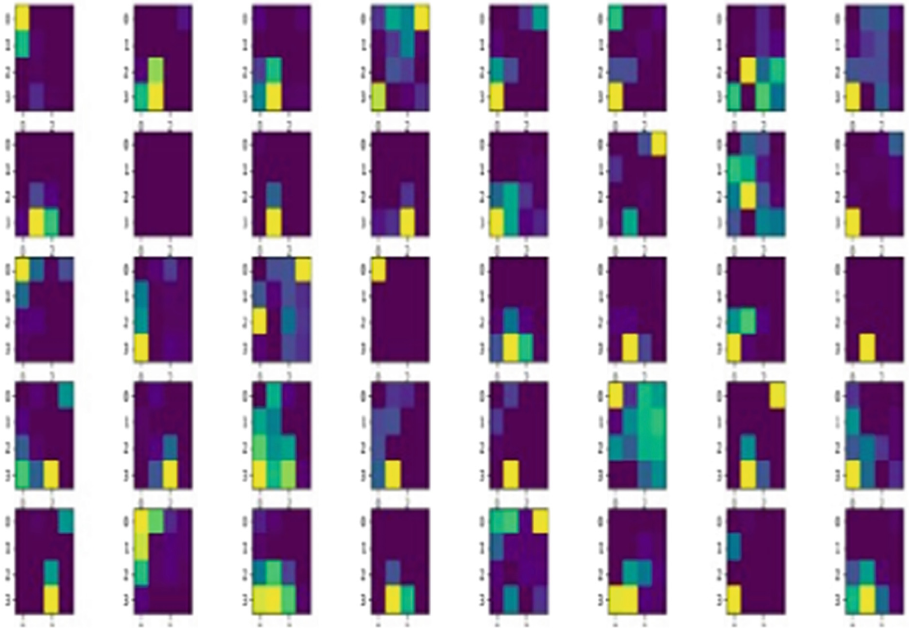


Fig. 6. Cell Image output from the eight layers

The performance of the model was concluded from the following evaluation metrics

1. Accuracy = (TP + FP)/(TP + FP + TN + FN)
2. Sensitivity = TP/(TP + FN)

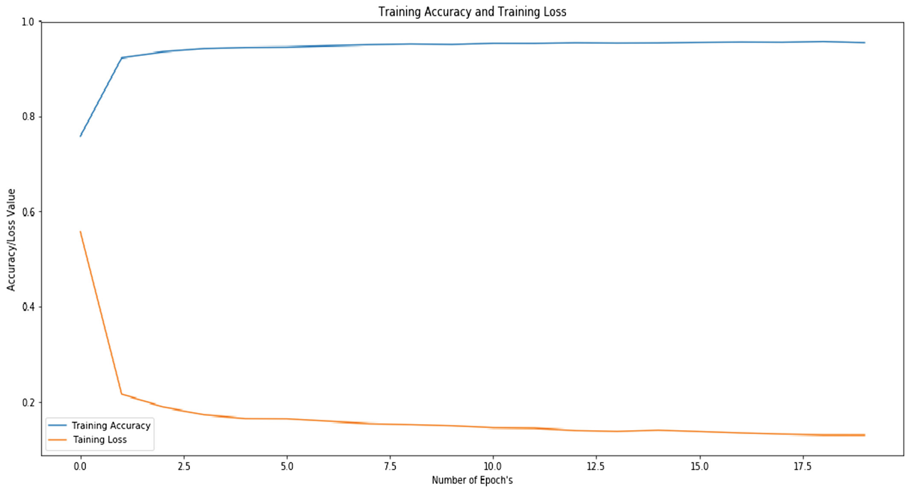


Fig. 7. Model training accuracy and loss

3. Specificity = TN/(TN + FP)
4. Precision = TP/(TP + FP)
5. F1 Score = 2 × Precision × Recall/(Precision + Recall)

where TP = True Positive, FP = False Positive, TN = True Negative, FN = False Negative

Table 2. Accuracy measure of proposed model

Measure	CNN model
Accuracy	98.23%
Sensitivity	96.44%
Specificity	99.99%
Precision	99.09%
F1 Score	97.74%

5 Conclusion

The proposed custom designed CNN models correctly classify the infected and uninfected cells, with relatively less parameter and higher accuracy, it does a better classification with 22558 images as training data and 5000 images as training data. Therefore, it makes it computationally less expensive without making any trade off. As with the deep learning advent, the perceptual problem is significantly solved and with the global health threat as malaria is imposing, diagnosing malaria will be very precise and fast and avoid misdiagnosis.

References

1. World Health Organization: World Malaria Report (2018)
2. Hisaeda, H., Yasutomo, K., Himeno, K.: Malaria: immune evasion by parasites. Int. J. Biochem. Cell Biol. **37**(4), 700–706 (2005)
3. Global Technical Strategy for Malaria 2016–2030, WHO'S E-2020 initiative and malaria elimination
4. Di Ruberto, C., Dempster, A., Khan, S., Jarra, B.: Analysis of infected blood cell images using morphological operators. Image Vis. Comput. **20**, 141–144 (2002)
5. Ross, N.E., Pritchard, C.J., Rubin, D.M., Duse, A.G.: Automated image processing method for the diagnosis and classification of malaria on thin blood smears. Med. Biol. Eng. Comput. **44**, 427–436 (2006)
6. Mitiku, K., Mengistu, G., Gelaw, B.: The reliability of blood film examination for malaria at The Peripheral health unit. Ethiop. J. Health Dev. **7**, 97–204 (2003)
7. Pattanaik, P.A., Swarnkar, T., Sheet, D.: Object detection technique for malaria parasite in thin blood smear images. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (2007)
8. Hendrawan, Y.F., Angkoso, C.V., Wahyuningrum, R.T.: Colour image segmentation for malaria parasites detection using cascading method. In: International Conference on SIET (2017)
9. Gopakumar, G., Swetha, M., Siva, G.S., Subrahmanyam, G.R.K.: CNN based malaria diagnosis from focus-stack of blood smear images acquired using custom-built slide scanner, Online Wiley Library (2018)
10. Dong, Y., et al.: Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. IEEE (2017)
11. Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. J. Physiol. **195**(1), 215–243 (1968)
12. Karpathy, A.: CS231n Convolutional Neural Networks for Visual Recognition (2018)
13. George, A., Routray, A.: Real-time eye gaze direction classification using convolutional neural network. In: International Conference on Signal Processing and Communications (SPCOM) (2016)
14. Jarrett, K., Kavukcuoglu, K., Ranzato, M.A., LeCun, Y.: What is the best multi-stage architecture for object recognition? In: International Conference on Computer Vision, pp. 2146–2153. IEEE (2009)
15. Liang, Z., et al.: CNN-based image analysis for malaria diagnosis. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (2016)
16. Rajaraman, S., et al.: Pre-trained convolutional neural networks as feature extractors toward improved Malaria parasite detection in thin blood smear images. PeerJ **6**, e4568 (2018). https://doi.org/10.7717/peerj.4568
17. National Library of Medicine website, U.S. (2018). https://ceb.nlm.nih.gov/repositories/malaria-datasets/

Rank Based Multi Path Job Execution Sequencing for Multi Cluster Environment to Find Shortest Path

Jasleen Kaur(✉), Anil Kumar, and Dhanpreet Singh Dhingra

CSE Department, GNDU, Amritsar, India
jasleenahuja8@gmail.com, anil.gndu@gmail.com,
dhanpreet.cits@gndu.ac.in

Abstract. Job scheduling is major problem in advanced computing system. To tackle the issue of larger Makespan and Flowtime, several techniques are being researched over. This paper works towards creation of job scheduling policy where burst time is considered for arranging the jobs in clusters. Proposed system is categorised into two phases: first phase arranges the jobs by following smaller burst time first scheduling. The queue thus formed is presented to round robin scheduler with time quantum that varies depending upon the burst time of job. Jobs are arranged in batches of 10% of total jobs in queue and arranged according to Ranks. On basis of requirement of resources the ranks are given to jobs. SJF scheduler considered is non premptive where RRS scheduler is premptive. Second phase executes the jobs by looking at the resource clusters. Multi-source shortest path dynamic algorithm is used for selecting jobs from cluster of resources assigned. Once job execution is complete Ranks are assigned which will be from 0–10. Higher the Rank more proficient is the result. The application of proposed system is done and optimum result is obtained in terms of Makespan and flowtime. Simulation is conducted in MATLAB showing improvement of 6% in overall result.

Keywords: Job scheduling · SJF · RRS · PBS · Rank based multi source shortest path · Rank

1 Introduction

Advanced computing provides resources on as you pay as use to the users. As the user of cloud increases, resources availability degrades. Resource utilization thus becomes critical within advanced computing like cloud. Resource allocation and management within distinct fields of advanced computing has been researched over. The in depth study of resource selection and allocation policies are provided in this section.

1.1 Resource Allocation Within Cloud

It is critically based on cost. Resources can be served as service in advance computing over the internet. Dedicated connection is rare in chance of cloud.

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 29–38, 2019.
https://doi.org/10.1007/978-981-15-0108-1_4

In other words shared resources are used among customary clients of cloud. Buyers not required to buy costly resources rather they can get distributed computing administrations on the basis of as you pay as you use (Mirashe and Kalyankar 2010). The equipment support is given the assistance of information centers(DCs). As it were physical resources are given through DCs. The DCs are additionally divided into virtual machines. Abnormal state versatility, spryness and accessibility is guaranteed by DCs. As indicated by (Armbrust et al. 2010) the flexibility of resource pool is remarkable system given to assist IT industry without paying an overwhelming add up to specialist organizations. This allows reduction in cost during maintenance and deployment. Demand of resources from the cloud is enhanced over the past decades. The availability associated with resources becoming an issue which is tackled using the techniques suggested by (Singh et al. 2012; Ranganathan et al. 2002). The growing demands of resources through single cloud provider are lacking scalability and hence availability decreases. To resolve the issue, public and private cloud is merged to form federation of cloud.

Cloud computing infrastructure weather associated with single or multiple clouds, is a complex multitude resource cluster (Kaur 2014). The resources that are computable required to be managed to resolve the issue of starvation or performance degradation (Kaur et al. 2014). Cloud resource are significantly managed the cost affect, availability and performance issues efficiently. Resource management in cloud distinct layers are unique from one another in many aspects. Main differing aspects include elasticity, workload complexity and availability. In case fluctuation in workload is present, booking of resources policies comes into picture. In case unplanned spikes in workload auto scaling of workload can be done to balance the load out. Auto scaling of work load is provided by PaaS providers. Spikes in workload are unpredictable and hence centralised resource control by single cloud provider is inefficient.

1.2 Resource Allocation in Grid

It generally depends of level of different resource requirement along with booking of resources scheme. (Depoorter et al. 2014) booking of resources in grid computing is critical since it provide simultaneous access to resources hence increasing performance according to the condition of resource availability and speed (Li et al. 2014). Booking of resources (AR) for global grids turns into a critical research zone as it enables clients to increase simultaneous access for their applications to be executed in parallel, and ensures the accessibility of resources at indicated future circumstances (Yousif et al. 2011; Kaur 2014). Booking of resources (BR) is a methodology of requesting Resources for use at a specific time later on. Normal Resources whose utilization can be held or requested are CPUs, memory, circle space and framework information transmission. BR for a framework Resource handles the above issue by empowering customers to increment concurrent access to adequate Resources for applications to be executed. BR in like manner guarantees the openness of Resources to customers and applications at the required conditions. Surveying distinctive BR circumstances can't conceivably be finished on a certified matrix condition on account of its dynamic nature. During booking of resources in Grid, there exists life cycle accompanied with states. These states are elaborated as under:

- Requested: it is an initial state when resources are first requested.
- Rejected: The booking of resources process is not successful or existing reservation is expired.
- Accepted: Booking of resources is successful. The reservation process yield resources for the process and it can be successfully completed.
- Committed: Resource reservation is committed before expiration of policy. Hence resource is assigned to the process and process can execute.
- Active: Process is currently executing since resources are successfully allotted to it.
- Cancelled: User does not require the resources anymore and request cancellation of allotted resources.
- Completed: Booked end time is reached.
- Termination: The user ends the reservation before end time is reached.

Resource allocation in federated Cloud depends on the distance of virtual machine from the jobs. Need for federated cloud originate as the resource deprecate from single cloud platform. As the distance increases so does the cost associated with the resources. In order to tackle the issue, nearest neighbour is identified and selected for allocation to virtual machines within federated cloud.

Rest of the paper is organised as under: Sect. 2 gives the literature survey consisting of existing approaches used to optimise the resource allocation process, Sect. 3 gives the flow of proposed system with experimental details, Sect. 4 gives the performance analysis and result, Sect. 5 gives conclusion and future scope and last section gives the references.

2 Literature Survey

Ideal resource allocation to execute job inside cloud condition is basic to ration Makespan and Flowtime. (Horng and Lin 2015) proposed a joint effort of insect settlement and ordinal advancement to take care of the issue of job scheduling. Job plan is situated through the said system and after that executed. Result got is diminished execution time. Calendar to be pursued is spoken to with insect for this situation. Change portrayal is utilized to demonstrate activity inside the timetable. Ordinal improvement component is utilized so as to dispose of replication if any inside the calendar to advance the spending allocation process. (Barbosa & Monteiro 2008) proposed an instrument to allocate jobs started from numerous clients to the bunch. Fixed arrangement of resources are allocated to every client in this work. The execution of jobs relies on the quantity of jobs inside the bunch.

Makespan and Flowtime increments as number of jobs increments. (Xhafa et al. 2011) proposed hybridization of hereditary and tabu scan instrument for job allocation and execution. (Rodger 2016; Elghirani et al. 2008) To execute the jobs hereditary methodology is pursued and to find the resource tabu inquiry is utilized. Wellness work is characterized as far as expense. The wellness work consequently must be limited and is accomplished through said writing. (Switalski and Seredynski 2014) proposed a summed up external improvement (GEO) which is upgrade of hereditary methodology. The examined methodology comprises of two stages. In the primary stage, ideal virtual machine out of the accessible machines is chosen. In the second stage, bunches are booked to execute on chosen virtual machine. (Kliazovich et al. 2013) proposed an energy mindful job scheduling inside the server farms. Energy efficiency and system mindfulness is being introduced in this writing for accomplishing improvement as far as Makespan and Flowtime.

Writing overview of multi heuristic methodologies recommend that there is absence of essential scheduling approaches alongside premption. This could prompt starvation. To yield ideal outcome, Premptive scheduling and be converged alongside non premptive scheduling and credit based component for choosing most ideal answer for executing jobs out of accessible arrangements. Next segment presents stream of proposed framework.

3 Proposed System

The proposed system consists of two phases used to optimize job scheduling process. Phase 1 consists of hybridization of three algorithms: Non premptive SJF, Premptive Round robin and Process batch scheduling mechanism to form batches to be submitted to the next phase. Next phase consist of multi source shortest path to identify the jobs which are related to previous jobs for execution.

Environment Considered for Experiment in the Proposed System is Public Cloud: The cloud environment is divided into data centres and virtual machines. Virtual machines are selected for allocation of resources.

The flow of proposed system is given in Fig. 1.

According to Fig. 1, initially job list is fetched from source. Source in proposed literature is initialised randomly. The burst time is obtained through random function in MATLAB environment. The algorithm in 3.1 considers sorting of jobs according to their burst time. Algorithm of data structure known as bubble sort is used for this purpose. The obtained result is sorted jobs which are maintained within the queue. The obtained jobs are fed into shortest job first scheduler.

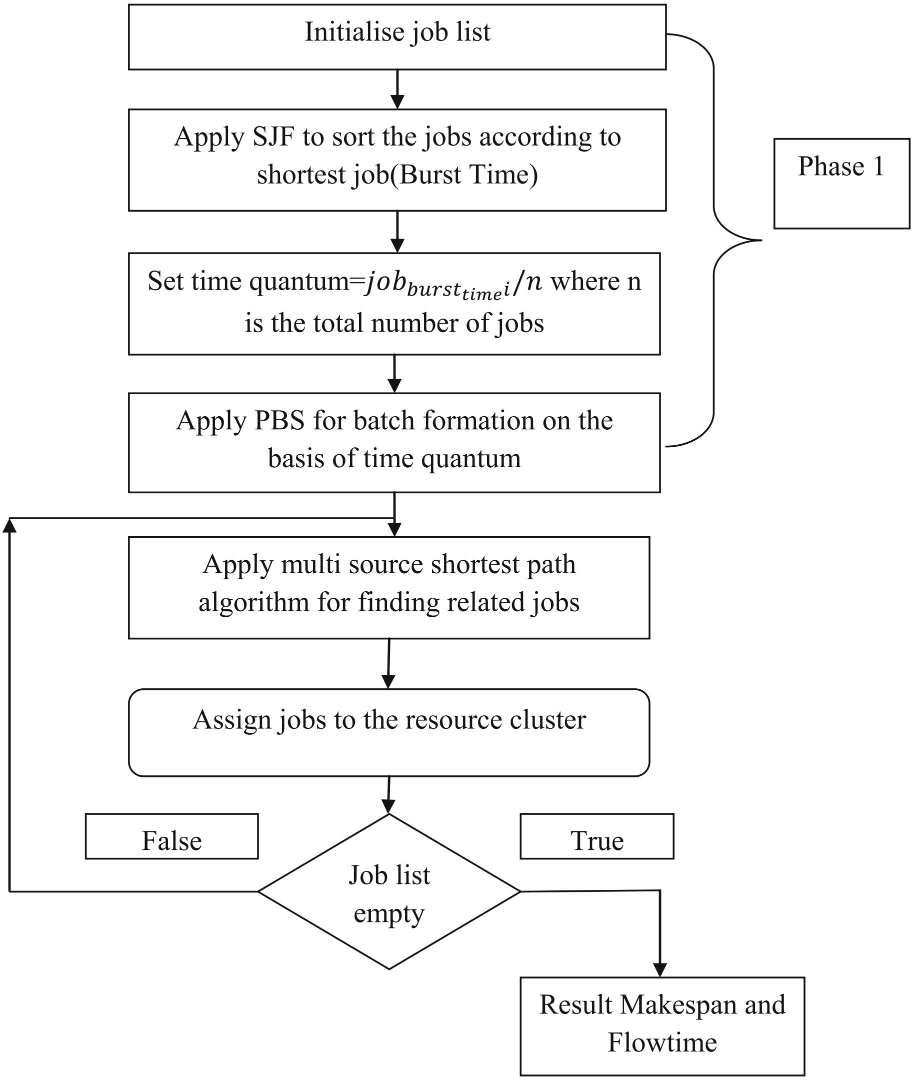

Initialise job list
Apply SJF to sort the jobs according to shortest job(Burst Time)
Set time quantum=$job_{burst_{time}i}/n$ where n is the total number of jobs
Apply PBS for batch formation on the basis of time quantum
Phase 1
Apply multi source shortest path algorithm for finding related jobs
Assign jobs to the resource cluster
False
True
Job list empty
Result Makespan and Flowtime

Fig. 1. Flow of hybrid rank based system.

3.1 Working of Shortest Job First Scheduler

The shortest job first scheduler (Suri and Rani 2017) get the burst time of each job and arrange the jobs in ascending order to burst time. Pseudo code for the SJF is given as under

```
For i=1:N
For j=1:N-i
If($Job\_bursttime_j$>$Job\_bursttime_{j+1}$)
Temp= $Job\_bursttime_j$
$Job\_bursttime_{j=}$ $Job\_bursttime_{j+1}$
$Job\_bursttime_{j+1}$=Temp
End of for
End of for
```

Algorithm 1: Shortest Job First Algorithm.

3.2 Working of Round Robin Scheduler

As in Fig. 1, SJF scheduler once finished operation, round robin scheduler works to achieve premption. This mechanism ensures that starvation problem can be resolved. The starvation could occurs if resource is occupied by job for long periods of time. Time quantum is evaluated using the equation

$$Time_{Quantum} = \frac{Job_{Burst_i}}{N}$$

Equation 1: Time quantum evaluation.

3.3 Working of PBS

As in Fig. 1: The main task of PBS scheduler is to form batches of jobs that can be allocated to the virtual machine clusters. Tasks that are present within PBS are analysed for relatedness. Since jobs are partitioned according to time quantum, hence tasks from same jobs may lie in distinct clusters. Relatedness thus becomes critical to join the processes for evaluation of parameters like Makespan and flow time.

3.4 Working of Multi Source Shortest Path Algorithm

As in Fig. 1 This algorithm is used to identify the relatedness from PBS system. Multi source indicates that there are multiple clusters and tasks from multiple jobs may lie in distinct clusters. Shortest path indicates the closeness of tasks which are initiated from same job. The pseudo code for the multi source shortest path algorithm is given as under

```
Receive batches of jobs
Buffer$_i$=batches
For(i=1:N) where N is the total number of tasks within each batch
If(batch_job$_i$_Pid== batch_job$_{i+1}$_Pid
  Distance$_i$=0
Else
Distance$_i$=1
End of if
End of for
```

Algorithm 2: Multi source shortest path algorithm.

The distance factor if 1 then job is from similar source. After finding the similarity jobs are merged and executed.

Result is obtained in terms of Makespan and Flowtime. Makespan is total time taken to execute each and every job. Flowtime on the other hand is obtained by determining execution of individual job. The steps are repeated and credits are assigned at each step of job execution. The simulation finishes when specified number of iterations are completed. Iteration with the highest credit value gives desired Makespan and Flowtime. Pseudo code for entire proposed system is given as under

```
Phase 1
    • Index all the jobs using SJF-RRS-PBS with time quantum 4 and arrange
      them into temporary batch.
    • Repeat the following until all jobs in cluster are processed
    • Check whether resources are available in cluster or not, If available  than
        ➢ Allocate the resources and obtain the make span, flow time.
    • Otherwise
        ➢ Choose the new job from batch using Dynamic programming
          algorithm  and update the jobs using SJF-RRS-PBS
      End if
      End Loop
    • Go to Phase 2
Phase 2
    • Shuffle the jobs in the clusters.
    • Repeat the following steps until all jobs are processed
        ➢ Apply dynamic programming to obtain optimal solution for jobs in
          cluster
        ➢ Calculate the number of jobs for which optimal solution is obtained
        ➢ Obtain the normalized function
    • End Loop
    • Apply credit depending upon  the values of make span and flow time
    • Go to phase 1 until required number of generations met.
```

The algorithm collaborate all the steps from 3.1 to 3.4 to determine optimal value of Makespan and Flowtime. Next section gives the results obtained and performance analysis.

4 Performance Analysis and Results

Job scheduling within cloud environment is simulated through this research. The job relatedness is the main criteria along with the rank assigned to each iteration of batch execution. The process continues until specified number of times, simulation executed. Highest rank value assigned to iteration gives desired Makespan and Flowtime. Obtained results are given in this section.

First of all results are given in terms of Makespan (Table 1)

Table 1. Result in terms of makespan.

Number of jobs	Without rank based system	With rank based system
10	43	39
20	58	40
30	61	52
40	75	65
50	87	79

Plot for the Makespan is given as under (Fig. 2):

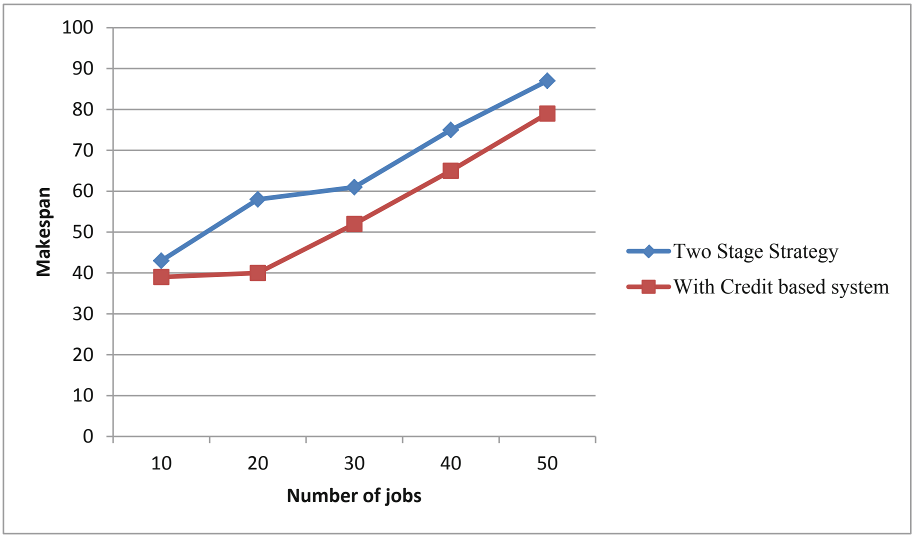


Fig. 2. Makespan plot.

Flowtime is generally smaller than the Makespan. The result obtained through proposed system and existing system is given as under (Fig. 3):

Number of jobs	Without rank based system	With rank based system
10	25	10
20	32	16
30	42	20
40	49	25
50	58	28

Fig. 3. Result in terms of Flowtime.

In terms of plot Flowtime is given as under (Fig. 4):

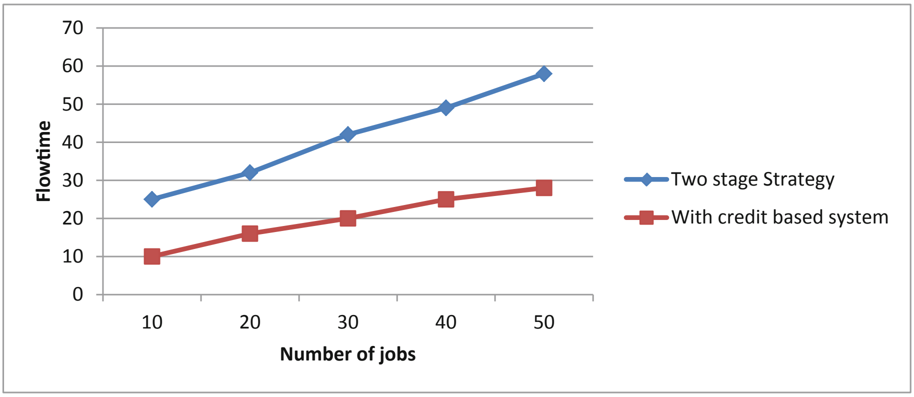


Fig. 4. Flowtime plot.

Results obtained from proposed system in terms of Makespan and Flowtime is better hence proving worth of the study.

5 Conclusion and Future Scope

Job scheduling becomes critical in order to utilize the resources efficiently. To this end, proposed system uses hybridization of multiple scalar schedulers along with PBS for efficiently forming a batch. Job execution is associated with the rank values. Schedule formed are entered into the execution mechanism and relatedness is identified by the application of multi source shortest path algorithm. Obtained results are in terms of Makespan and Flowtime. Result improvement of nearly 10% is observed.

In future, relatedness calculation mechanism can be further improvised by incorporating multiheuristic algorithm like genetic approach to improve Makespan and Flowtime.

References

Armbrust, M., et al.: A view of cloud computing. Commun. ACM **53**(4), 50 (2010). http://portal.acm.org/citation.cfm?doid=1721654.1721672

Barbosa, J., Monteiro, A.P.: A list scheduling algorithm for scheduling multi-user jobs on clusters. In: Palma, J.M.L.M., Amestoy, P.R., Daydé, M., Mattoso, M., Lopes, J.C. (eds.) VECPAR 2008. LNCS, vol. 5336, pp. 123–136. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-92859-1_13

Vanmechelen, D.: Booking of resources and cooallocation in parallel computing for reducing makespan and flowtime. FGCS **41**, 1–15 (2014). http://dx.doi.org/10.1016/j.future.2014.07.004

Elghirani, A., et al.: Using genetic algorithm to enhance the performance of parallel grid computing. In: 6th IEEE/ACS International Conference on Computer Systems and Applications, AICCSA 2008, pp. 436–443 (2008)

Horng, S.C., Lin, S.S.: Solving the problem of makespan and flowtime by integrating ant colony optimization with ordinal clustering in parallel system. In: Proceedings - International Conference on Intelligent Systems, Modelling and Simulation, ISMS, October 2015, pp. 70–75 (2015)

Kaur, M., Sharma, S., Kaur, R.: Improving scheduling mecahnsim in advanced computing. IJARCSC **4**(7), 2277–128 (2014)

Kliazovich, D., Bouvry, P., Khan, S.U.: DENS: data center energy-efficient network-aware scheduling. Cluster Comput. **16**(1), 65–75 (2013)

Li, B., et al.: Resource availability-aware booking of resources for parallel jobs with deadlines, pp. 798–819 (2014)

Mirashe, S.P., Kalyankar, N.V.: Cloud computing. In: Antonopoulos, N., Gillam, L. (eds.) Communications of the ACM, vol. 51, no. 7, p. 9 (2010). http://arxiv.org/abs/1003.4074

Rajvir Kaur, S.K.: Review of distinct job scheduling algorithm used in advanced computing. IJCTT **2**(March), 12–22 (2014)

Ranganathan, K., Iamnitchi, A., Foster, I.: Improving data availability through dynamic model-driven replication in large peer-to-peer communities. In: 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid, CCGrid 2002, pp. 0–5 (2002)

Rodger, J.A.: Information discovery and scheduling within advanced computing using the application of multiheuristic algorithms, pp. 17–26 (2015). http://dx.doi.org/10.1016/j.imu.2016.01.002

Singh, D., Singh, J., Chhabra, A.: Increasing reliability of cloud using the checkpointing algorithm with mean time between failure enhancement. In: CSNT 2012, pp. 698–703 (2012)

Suri, P.K., Rani, S.: Design of task scheduling model for cloud applications in multi cloud environment. In: Kaushik, S., Gupta, D., Kharb, L., Chahal, D. (eds.) ICICCT 2017. CCIS, vol. 750, pp. 11–24. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-6544-6_2

Switalski, P., Seredynski, F.: Scheduling parallel batch jobs in grids with evolutionary metaheuristics. J. Sched. **18**(4), 345–357 (2014). https://doi.org/10.1007/s10951-014-0382-0

Xhafa, F., et al.: A GA+TS hybrid algorithm for independent batch scheduling in computational grids. In: Proceedings - 2011 International Conference on Network-Based Information Systems, NBiS 2011, pp. 229–235 (2011)

Yousif, A., Abdullah, A.H., Nor, S.M.: Scheduling jobs on grid computing using. IEEE Access **33**(2), 155–164 (2011)

Optimization of a Real Time Web Enabled Mixed Model Stochastic Assembly Line to Reduce Production Time

Rangith Baby Kuriakose(✉) and Hermanus Jacobus Vermaak

Department of Electrical, Electronics and Computer Systems,
Central University of Technology, Bloemfontein 9301, Free State, South Africa
{rkuriako,hvermak}@cut.ac.za

Abstract. The role of assembly lines has never been more critical as it is now with the world entering the 4th Industrial Revolution, commonly referred to as Industry 4.0. If the focus of the previous industrial revolution was on mass production, the focus of Industry 4.0 is on mass customization. One of the major changes mass customization brings about to an assembly line is the need for them to be autonomous. An autonomous assembly line needs to have the following key features; ability to provide a ubiquitous input, the ability to optimize the model in real time and achieve product variety. Product variety, in this context, refers to different variants of the same product as determined by the user. Assembly lines that make provision for introducing product variety are termed as mixed-model assembly lines. Mixed-model assembly lines become stochastic in nature when the inputs are customized as time cannot be predetermined in a stochastic process. The challenge, as it stands, is that there are limited discussions on real-time optimization of mixed model stochastic assembly lines. This paper aims to highlight this challenge by considering the case study of optimizing a mixed model assembly line in the form of a water bottling plant. The water bottling plant, which needs to produce two variants of the bottled water, 500 ml, and 750 ml, takes customer inputs through a web interface linked to the model, thereby making it stochastic in nature. The paper initially details how the model replicating the functioning of the water bottling plant was developed in MATLAB. Then, it proceeds to show how the model was optimized in real time with respect to certain constraints. The key results of the study, among others, showcase how the optimization of the model is able to significantly reduce production time.

Keywords: Real Time Optimization · Cloud manufacturing · Mixed model · Assembly lines · Stochastic processes · Industry 4.0

1 Introduction

Assembly lines are critical to the operation of any successful manufacturing plant [1]. This holds value even as the world is embracing the fourth Industrial Revolution [2] commonly termed as Industry 4.0 (I.4.0). The industrial revolutions of the past focused on mass production [3, 4] and reducing production time.

However, with the advent of I.4.0, the focus has shifted from mass production to mass customization [5]. Mass customization is key to introducing product variety [6]

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 39–50, 2019.
https://doi.org/10.1007/978-981-15-0108-1_5

into manufacturing. With product variety, customers are given the opportunity to make variations in the final product.

Product variety is achieved by using Multi/Mixed model assembly lines [7, 8] with stochastic task times [9] and as a result poses problems in assembly line balancing. Assembly Line Balancing Problems (ALBP) are a study of the different types of problems [10] encountered in balancing assembly lines.

There have been many classifications of ALBP's, but the most referenced include that by Ghosh and Gagnon in 1989 [11], Scholl and Becker in 2006 [12] and finally by Sivasankaran and Shahubdeen in 2014 [13]. The later study has split Multi/Mixed model Stochastic (MMS) assembly lines in accordance to the line layout to be either S-type or U-type.

The complexity of a MMS assembly line is exacerbated if one considers real-time inputs as this would need Real-Time Optimization (RTO) [14, 15]. As it stands there is very limited research done on MMS assembly line balancing and even less literature on Real-Time Optimization of MMS assembly lines.

Therefore, in order to garner fresh perspectives in this research niche area, this paper considers optimization of a real-time mixed model assembly line. A case study of a water bottling plant that can manufacture 500 ml and 750 ml with real time inputs from a web server is considered for this purpose.

The paper initially considers the design criteria for the water bottling plant. Then it looks at how the plant was modelled in MATLAB and real-time inputs were provided through a Web Server. Thirdly, the paper focuses on the problem formulation and optimization of the model and finally the results thereof are discussed.

2 Water Bottling Plant as a Case Study

As stated in the introduction, in-order to design a multi mixed assembly line, a case study needs to be selected. In this study, a business plan [16] to realize a water bottling plant at the Bloemfontein campus of the Central University of Technology, was chosen as the case study. A three-dimensional CAD model of the proposed model is shown in Fig. 1.

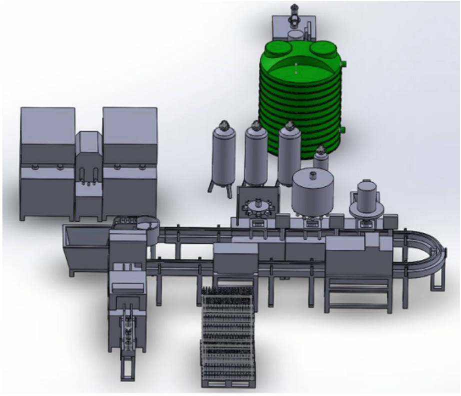


Fig. 1. 3-dimensional CAD model of the plant

The CAD model was split into three main subsystems for creating a Simulink model. The three subsystems with their specific tasks are defined as follows;

- Subsystem A – Source and Storage tank
- Subsystem B – Bottle manufacturing and storage
- Subsystem C – Water filling

The Simulink model is depicted in Fig. 2 with the different subsystems. As seen in Fig. 2, the input is provided through a web server. The input is fed to the water filling substation. Based on the inputs, the water filling subsystem triggers water and bottles from subsystem A and B respectively.

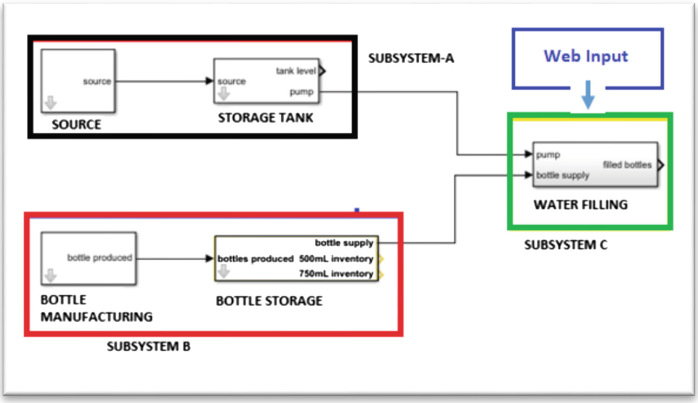


Fig. 2. Simulink model of the plant with three subsystems

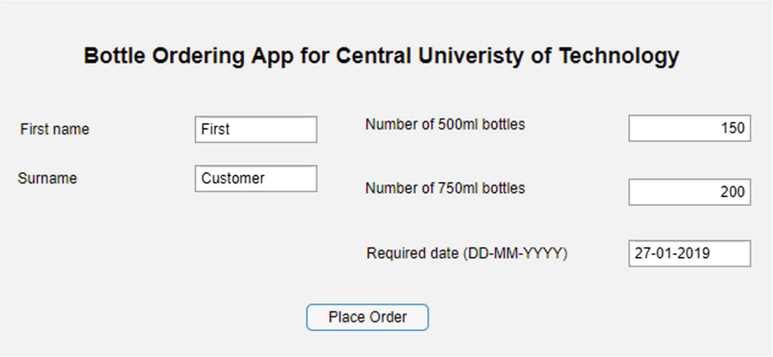
3 Model Overview

This section elaborates on the Simulink model described in Fig. 2. This is done by initially focusing on the web input and how it is setup. Secondly the focus will be on the three subsystems with an aim to explain the functioning of each block in the respective subsystems. This will form the base for discussion on the problem formulation and optimization, which is discussed in Sect. 4.

3.1 Customer Inputs Using Web Apps

Web apps are MATLAB [17] applications that can run in a web browser. An interactive Graphic User Interface (GUI) is created using the App Designer and packaged using a Web App Compiler and hosted on the MATLAB Web Server App. A unique URL is created for each web app and can be accessed from a browser using HTTP or HTTPS protocols.

In this paper, the web app is designed to first collect information from customers on the number of 500 ml and 750 ml bottles of water that needs to be ordered and the required date of delivery. This information is captured and saved on the. A program written in MATLAB is used to access this saved file and process it so as to start the water filling process. A picture depicting the GUI is shown in Fig. 3.



Bottle Ordering App for Central Univeristy of Technology

First name: First

Surname: Customer

Number of 500ml bottles: 150

Number of 750ml bottles: 200

Required date (DD-MM-YYYY): 27-01-2019

Fig. 3. GUI for bottle ordering app

3.2 Source and Storage Subsystem-Subsystem A

As depicted in Fig. 3, subsystem A acts as one of the inputs to subsystem C. Subsystem A consists of two subsystems being the water source and the storage tank. The source subsystem is designed to provide purified water to the plant. For design purposes, the source subsystem is a masked meaning that parameters such as the flow rate and upper limit of water from the source system are user defined in the mask.

The storage tank subsystem needs to take purified water from the source subsystem and pump it into the water filling subsystem. These conditions make it a continuous state subsystem. Continuous states, in the context of Simulink, refers to a variable whose value is determined through numerical integration of its derivative with respect to time.

As with the source subsystem, the storage tank subsystem is also masked, with values like the upper limit of the tank, maximum capacity and pump flow rate defined under the mask. For optimization purposes, the pump flow rate is defined as ‘x’ in the mask as it acts as the handle and needs to be varied with respect to the constraints.

3.3 Bottle Manufacturing Subsystem-Subsystem B

Subsystem B consists of the bottle manufacturing subsystem and the bottle storage subsystem. Subsystem B is independent of subsystem A and the output of the bottle storage subsystem is fed to subsystem C.

The bottle manufacturing subsystem, in keeping with the design parameters, require that bottles be manufactured in two sizes, 500 ml and 750 ml. The bottles are manufactured according to the order from the web input through the bottle ordering app depicted in Fig. 3.

The manufacturing and storage is done by retrieving the orders, saved on the server, into a one dimensional look-up table. The data in the look-up table can be transposed and flattened to appear as a row of information to the table. A relational table with a memory block can be used to check if the number of bottles in each row has been reached.

3.4 Water Filling Subsystem-Subsystem C

As explained previously, subsystem C takes inputs from subsystem A and B. Subsystem A provides the water and subsystem B provides the bottles. The data from the one dimensional lookup table described in subsystem B is used for filling the bottles.

As soon as the number of bottles in the first row has been achieved, a trigger element can be set up to first index the data and move to the next row in the look-up table. This can be continued till the last row in the look-up table has been read into the model and defined in the index.

The distinction between 500 ml and 750 ml bottles can be made by calling a function in the model which checks the index where the data has been read from. By default, the 500 ml bottles will be indexed in the odd rows while the 750 ml will be indexed in the even rows. The output of the bottle storage unit is then provided as input to subsystem C.

4 Problem Formulation and Optimization

The model that has been designed in this study has customized inputs as the inputs are fed through a web interface. This would mean that model will have real inputs. Therefore, the optimization problem needs to be formulated [20] as a Real Time Optimization (RTO) problem.

The formulation can be split into five steps described as follows;

- Step 1: Determine process variables – The plant model has the following variables;
  - Water stored in the tank in the Source subsystem
  - Flow rate of water from the pump in the storage tank subsystem
  - Initial number of 500 ml bottles in the bottle manufacturing subsystem
  - Initial number of 750 ml bottles in the bottle manufacturing subsystem
  - Expected date of delivery of customer orders
- Step 2: Defining the objective function – The objective function of the plant model is to reduce the production time for completing the customer orders. The hypothesis is that with optimization the production time can be significantly improved.
- Step 3: Development of process models - The model has considered three constraints being firstly the water level of the tank and secondly the number of 500 ml and 750 ml bottles available in storage. The water level in the tank should never go below 25%. The number of bottles in storage should never go below zero as this would result in the system crashing in a physical setup.

- Step 4: Simplify the process model – In order to simplify the process, the initial number of bottles is kept above zero. The pump flow rate from the storage tank subsystem acts as the handle which can be varied to meet the constraints.
- Step 5: Apply a suitable optimization technique – On analyzing the objective function, process variables and the constraints, it is noted that they exhibit a non-linear relationship. Since the modelling is done with Simulink, the optimization can be done using MATLAB.

The Optimization ToolBox™ in MATLAB [21] offers variety of functions to solve the different types of optimization problems. Step 5 described in the problem formulation localizes the problem in this paper to focus on analyzing constrained nonlinear optimization techniques.

There are mainly three functions available in MATLAB for nonlinear constrained optimization. They along with their specific purpose is listed as follows;

- *fminbnd* – Finding minimum variable function on fixed intervals
- *fmincon* – Find minimum of constrained nonlinear multivariable function
- *fseminf* – Find minimum of semi-infinitely constrained multivariable nonlinear function

It is evident that in the listed MATLAB functions, *fmincon*, is best suited for the plant model optimization as it accommodates the minimizing of nonlinear objective equations with constraints and multivariable functions. The *fmincon* function has the following syntax;

$$x = fmincon(fun, x0, A, b, Aeq, Beq, lb, ub, nonlcon, opt)$$

Where,

x0 = starting point of minimization

fun = function to minimize

b and *beq* are linear constraints

A and *Aeq* are nonlinear constraints

lb = lower boundary of constraint

ub = upper boundary of constraint

Here, the syntax needs *x0*, a starting point for the minimization. This is kept at 0.1. Next it needs a lower and upper boundary for the pump flow rate, which is the handle. This will ensure that the solution is always within the range of $lb \leq x \leq ub$. Here, *lb* = 0 and the *ub* = 1. As there are no inequality constraints, *Aeq* = [] and *beq* = [].

5 Results and Discussion

This section presents the results of the work described. This is done by comparing the non-optimized and optimized pump flow rate against the constraints defined in the problem formulation. Firstly, the customer requirements as obtained from the cloud server for this set of tests are shown in Table 1.

Table 1. Customer requirements table as per the input from the cloud server.

Customer number	Number of 500 ml bottles required	Number of 500 ml bottles required	Required date and time of delivery
1	100	100	16-Jan-2019 15:00
2	85	95	16-Jan-2019 15:00
3	120	120	16-Jan-2019 15:00
4	100	150	17-Jan-2019 15:00
5	120	150	17-Jan-2019 15:00
6	79	69	17-Jan-2019 15:00

Secondly, the pump flow rate, which acts as the handle, is varied. A set of three tests are done on the inputs shown in Table 1. The first two being with manual, non-optimized pump flow rates and the third one a model generated, optimized flow rate. The flow rate can be varied between 0 m/s to 1.0 m/s.

As a first test, the model is provided with a manual pump flow rate of 0.6 m/s. The assumption here is that with a high flow rate the orders will be completed at a fast rate. The result of this test is given Table 2.

Table 2. Customer requirements table with non-optimized delivery date with pump flow rate at 0.6 m/s

Customer number	Number of 500 ml bottles required	Number of 500 ml bottles required	Required date and time of delivery	Non-optimized date and time of delivery
1	100	100	16-Jan-2019 15:00	16-Jan-2019 02:54
2	85	95	16-Jan-2019 15:00	16-Jan-2019 03:14
3	120	120	16-Jan-2019 15:00	16-Jan-2019 03:42
4	100	150	17-Jan-2019 15:00	17-Jan-2019 12:07
5	120	150	17-Jan-2019 15:00	**17-Jan-2019 16:37**
6	79	69	17-Jan-2019 15:00	**17-Jan-2019 17:01**

It can be deduced from Table 2 that first four orders (Customer's 1, 2, 3 and 4) are completed before the required date of delivery. However, the last two orders (Customer's 5 and 6) are not completed on time. This is owing to the fact that the high water flow rate has resulted in the water in the tank being depleted. Therefore, the assembly line process has to be halted to allow for the tank to be replenished.

After the tank has been replenished, the assembly line process continues, but the time lost during the replenishing cannot be made up and hence the customer orders 5 and 6 fall behind of schedule. A GUI depicting the status of the constraints using gauges is shown in Fig. 4. It can be seen here that the gauge showing tank level has gone below 25%. This is a fail as in the problem formulation the first condition was that the tank level should not go below 25%.

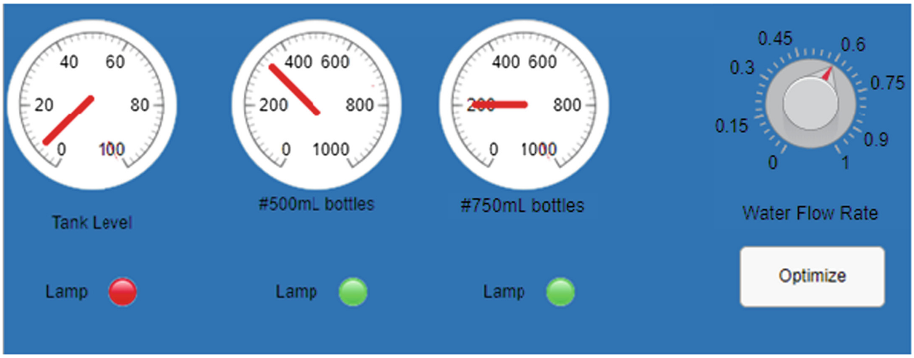


Fig. 4. GUI showing the status of the constraints at 0.6 m/s pump flow rate

The second tests had the model being provided with a manual pump flow rate of 0.1 m/s. The assumption here is that the slower pump flow rate would result in the constraints being met and therefore the orders will be completed in time. The result of this test is given Table 3.

Table 3. Customer requirements table with non-optimized delivery date with pump flow rate at 0.1 m/s

Customer number	Number of 500 ml bottles required	Number of 500 ml bottles required	Required date and time of delivery	Non-optimized date and time of delivery
1	100	100	16-Jan-2019 15:00	16-Jan-2019 12:50
2	85	95	16-Jan-2019 15:00	16-Jan-2019 13:23
3	120	120	16-Jan-2019 15:00	**16-Jan-2019 21:25**
4	100	150	17-Jan-2019 15:00	17-Jan-2019 07:14
5	120	150	17-Jan-2019 15:00	17-Jan-2019 13:26
6	79	69	17-Jan-2019 15:00	**17-Jan-2019 15:53**

It can be seen from Table 3 that the rate of completing the orders is much slower in this instance. This is evident in that customer orders 3 and 6 are not completed within the required time. However, unlike with the previous test, all constraints are met in this test. This is depicted in Fig. 5.

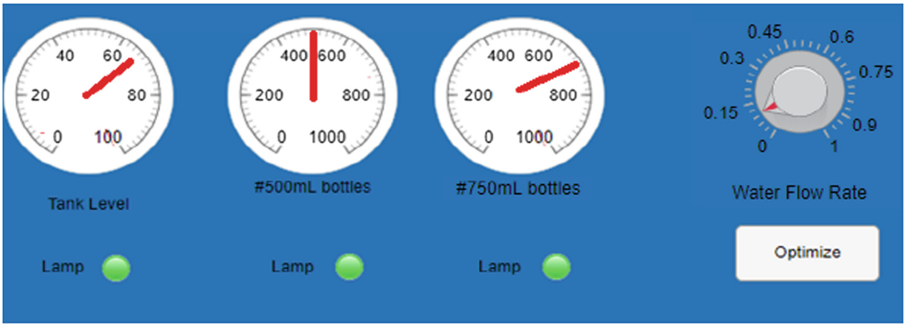


Fig. 5. GUI showing the status of the constraints at 0.1 m/s pump flow rate

The instances where production time exceeds the allotted time is termed as positive drift [3]. These tests and results put a strong emphasis for the need for a robust optimization function for this model. The optimization should be able to meet all constraints while ensuring that the required delivery dates are met.

As described in the problem formulation in Sect. 4, the MATLAB optimization function, *fmincon*, is applied on the model. The syntax defined in the problem formulation is expanded out as follows;

$$[xOpt, TTMOpt] = fmincon(fun, [0.1; 0.1], [], [], [], [], [0; 0], [1; 1], funConstr, opt)$$

Here, the starting point of the pump flow rate and the lower and upper boundaries of the constraint are defined. The starting point is defined as 0.1 and the lower boundary is 0, while the upper boundary is 1. *XOpt* is the optimized pump flow rate which will meet the defined constraints and *TTMOpt* is the 'Time To Manufacture' with optimized pump flow rate.

After completing the optimization, the *fmincon* function does a further check to test the robustness of the model by adding 0.001 to the optimized pump flow rate. The check should meet the constraints as before to ensure the model is robust and not prone to even the minutest of changes. The result of the optimization is summarized in Table 4.

Table 4. Customer requirements table with optimized time and date of delivery

Customer number	Number of 500 ml bottles required	Number of 500 ml bottles required	Required date and time of delivery	Optimized date and time of delivery
1	100	100	16-Jan-2019 15:00	16-Jan-2019 02:54
2	85	95	16-Jan-2019 15:00	16-Jan-2019 03:30
3	120	120	16-Jan-2019 15:00	16-Jan-2019 04:26
4	100	150	17-Jan-2019 15:00	16-Jan-2019 05:17
5	120	150	17-Jan-2019 15:00	16-Jan-2019 06:42
6	79	69	17-Jan-2019 15:00	16-Jan-2019 08:47

The optimized pump flow rate for this specific set of inputs was 0.36. It can be seen here that at the optimized pump flow rate, all the customer orders are met well before the required date and time of delivery.

On further analysis, the constraints such as the level of water in the tank and the number of bottles available in the inventory are met on completion of the orders. The status of the constraints is depicted in Fig. 6.

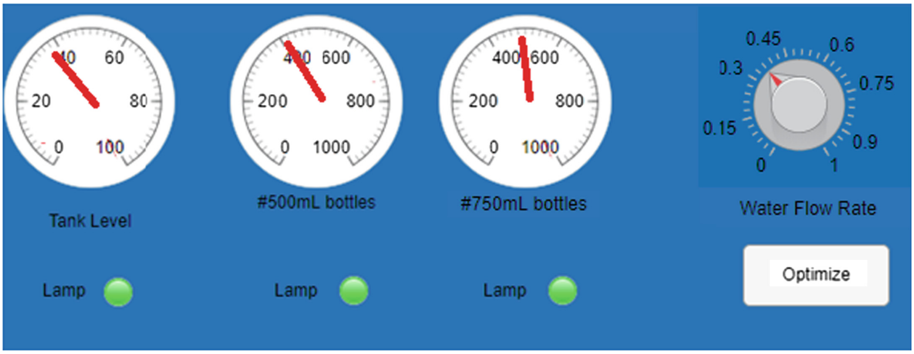


Fig. 6. GUI showing the status of the constraints at 0.1 m/s pump flow rate

6 Conclusion and Future Work

This study was necessitated due to the limited research done on Real Time Optimization of Multi Mixed Stochastic assembly lines. In order to facilitate this study, a case study was chosen.

The case study was on a water bottling plant that can manufacture 500 ml and 750 ml with real time customer inputs provided through a web server. The paper initially describes how the plant was modelled on Simulink. Secondly, the paper shows how a web app was developed on MATLAB and finally how the optimization problem was formulated.

The results of the paper initially discuss how a set of orders received through the web server is completed using a manual pump flow rate. Two sets of pump flow rates were chosen for test purposes. The first pump flow rate was chosen to be 0.6 m/s and the second pump flow rate was chosen as 0.1 m/s.

The first test instance showed that the first few orders were completed at a rapid pace. However, this resulted in the water in the tank being depleted at a fast rate. This meant that after the first few orders had been completed, the process had to be halted for the water in the tank to be replenished. This eventually resulted in two orders not meeting the required delivery date and time. These results are summarized in Table 2 and Fig. 4.

The second test, which used a slower pump flow rate, resulted in all constraints being met as depicted in Fig. 6. However, like in the first instance, a few orders were not met as per the required date of delivery and time. This is evident from the data shown in Table 3.

The required date of delivery was not met in both instances and necessitated the need for optimization. The MATLAB optimization function *fmincon* was used to optimize the production time based on the customer inputs and the required date of delivery. It was seen that the required delivery date and the constraints were met upon optimization. This is depicted in Table 4 and Fig. 6.

This proves that optimization technique used in this study was successfully able to reduce the production time for a mixed model stochastic assembly line with real time inputs. As part of future work, to establish the veracity and robustness of the model, numerous sets of inputs with various scenarios as far as the required date of delivery and number of orders needs to be applied to the model. The number of constraints, three in this study, could also be increased to see the stress it creates on the model.

The results of the suggested tests, as part of the future work, can add valuable insight into the field of mixed model stochastic assembly line balancing. The results could also determine if this model can be tested on an existing manufacturing plant and form part of creating an industry standard that can be used in future to combat problems related to positive drift.

Acknowledgements:. The authors would like to thank the RGEMS group of Central University of Technology, Free State for providing the lab resources to pursue this project. The authors would also like to thank MerSETA for financially contributing to the successful completion of the project.

References

1. Tun, U., Onn, H., Sulaiman, S., Ismail, N.: Assembly Line and Balancing Assembly Line, January 2015
2. Baldassarre, F., Ricciardi, F., Campo, R.: The Advent of Industry 4.0 in manufacturing industry: Literature review and growth opportunities. In: DIEM: Dubrovnik International Economic Meeting, pp. 632–643 (2017)
3. Hu, S.J., et al.: Assembly system design and operations for product variety. CIRP Ann. Manuf. Technol. **60**(2), 715–733 (2011)
4. Kuriakose, R.B., Vermaak, H.J.: A review of the literature on assembly line balancing problems, the methods used to meet these challenges and the future scope of study. Adv. Sci. Lett. **24**(11), 8846–8850 (2018)
5. Um, J., Lyons, A., Lam, H.K.S., Cheng, T.C.E., Dominguez-pery, C.: Product variety management and supply chain performance: a capability perspective on their relationships and competitiveness implications. Int. J. Prod. Econ. **187**, 15–26 (2017)
6. Wang, H., Hu, S.: Manufacturing complexity in assembly systems with hybrid configurations and its impact on throughput. CIRP Ann. Manuf. Technol. **59**(1), 53–56 (2010)
7. Kuriakose, R.B., Vermaak, H.J.: Optimization of a customized mixed model assembly line using MATLAB/Simulink. J. Phys.: Conf. Ser. **1201**, 012017 (2019)
8. Reginato, G., Anzanello, M.J., Kahmann, A., Schmidt, L.: Mixed assembly line balancing method in scenarios with different mix of products. Gestão Produção **23**(2), 294–307 (2016)
9. Baykasoglu, A., Ozbakir, L.: Stochastic U-line balancing using genetic algorithms. Int. J. Adv. Manuf. Technol. **32**, 139–147 (2007)

10. Kumar, N., Mahto, D.: Assembly line balancing: a review of developments and trends in approach to industrial application. Glob. J. Res. Eng.: Ind. Eng. **13**(2) (2013)
11. Ghosh, S., Gagnon, R.: A comprehensive literature review and analysis of the design, balancing and scheduling of assembly systems. Int. J. Prod. Res. **27**(4), 637–670 (1989)
12. Becker, C., Scholl, A.: A survey on problems and methods in generalized assembly line balancing. Eur. J. Oper. Res. **168**(3), 694–715 (2006)
13. Sivasankaran, P., Shahabudeen, P.: Literature review of assembly line balancing problems. Int. J. Adv. Manuf. Technol. **73**(9–12), 1665–1694 (2014)
14. Bonvin, D.: Preface to Real-Time Optimization Processes, Special edition, pp. 1–5 (2017)
15. Francois, G., Bonvin, D.: Real-time optimization: optimizing the operation of energy systems in the presence of uncertainty and disturbances. In: 13th International Conference on Sustainable Energy technologies (SET2014), pp. 1–12 (2014)
16. Kuriakose, R.B., Vermaak, H.J.: Customized mixed model stochastic assembly line modelling using Simulink. Int. J. Simul. Syst. Sci. Technol. **20**(1) (2019). ISSN 1473-804X
17. MATLAB: MATLAB Web Apps, MATLAB Documentation (2018). https://www.mathworks.com/help/compiler/web-apps.html. Accessed 28 Apr 2019
18. Kumar, D.N.: Introduction and basic concepts - classification of optimization problems, National Program on Technology Enhanced Learning. http://nptel.ac.in/courses/Webcourse-contents/IISc-ANG/OPTIMIZATION%20METHODS/pdf/Module_1/M1L3slides.pdf. Accessed Apr 2019
19. Kumar, D.N.: Optimization problem and Model formulation, Optimization Methods. http://msulaiman.org/onewebmedia/Lecture2mphil.pdf. Accessed 28 Mar 2018
20. Edgar, T., Himmelblau, D., Ladson, L.: Optimization of Chemical Processes, 2nd edn. McGraw-Hill Higher Education, New York (2001)
21. MATLAB: MATLAB Optimization Toolbox: User's Guide, MATLAB Documentation (2018). https://www.mathworks.com/help/pdf_doc/optim/optim_tb.pdf. Accessed 02 Jan 2019

Recommendation System of Dietary Patterns for Cardiovascular Disease Patients Using Decision Making Approach

Garima Rai and Sanjay Kumar Dubey(✉)

Department of Computer Science and Engineering,
Amity University Uttar Pradesh, Sec-125, Noida, India
Garima.rai003@gmail.com, sanjukundan@gmail.com

Abstract. Cardio vascular disease can be described as any disorder that constrains the normal functioning of heart or blood vessels. Diseases such as heart disease, Angina, Myocardial Infraction, Arrhythmia all can be termed as cardio vascular disease. One of the most common and frightening term is heart failure. It does not mean "failure" of heart or heart has stopped working but the heart does not pump as well as it should. This retains salt and water which result in swelling and shortness of breath. Cardio Vascular Disease is primarily caused by Building-up-of fatty deposits inside the arteries or damage to arteries in brain, kidneys and eyes. Increased level of blood pressure, cholesterol, diabetes and weights leads to cardio vascular diseases in human body. In this paper work, our interest is to discover ideal diet recommendation plan for people with cardio vascular disease and to obtain this we are implementing Analytic Hierarchy Process (AHP). Ideal plan of diet recommendation for people with CVDs is recently discovered. The outcomes reveal ideal diet recommendation among the currently available ones for people with CVDs ailment. The diet consists of food items to be taken up at breakfast, lunch and dinner by CVDs patients. In order to justify our work we are implementing fuzzy method which uses preference order similar to absolute solutions. The outcomes justify our work with results using Analytic Hierarchy Process.

Keywords: Recommendation system · CVD · AHP · Entropy · Model

1 Introduction

Cardio vascular disease, mainly cause due to overweight, poor nutrition, smoking and other factors such as high blood pressure, physical inactivity, lack of sleep, diets with high fat combined with carbohydrates, excessive drinking cause symptoms such as pains in the chest, slurred speech, cramping pain in leg. The symptoms vary from person to person and are specific for condition of an individual. Cardio Vascular diseases (CVDs) are one of the leading contributors to the burden of deaths in developing countries. In 2016, 17.9 million deaths recorded from CVDs. Year 2015 recorded 17 million premature deaths from Cardio Vascular Disease. Recent studies indicate that cardiovascular diseases are significantly present among the males from both rural and urban population. In India, there are annually more than 10 million

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 51–59, 2019.
https://doi.org/10.1007/978-981-15-0108-1_6

deaths due to diseases of circulatory systems with 40 percent as women. These diseases are the considerable cause of death among women, in middle age, in rural and in urban women. In western world risk factors for cardiovascular are enlarging terribly. One fearful aspect with cardiovascular disease includes that it evokes numerous significant ailments eventually giving rise to innumerable body problems. This global pervasiveness of CVDs is due to unhealthy eating habits, physical inactivity, social isolation and financial issues leading to irregular medical checkup. Earlier, works have discovered deal diet recommendation by using several techniques and methodologies. Diet-Right an ideal diet recommendation system has been developed. It is based on cloud and use users' pathology reports for selecting an ideal diet to fulfill one's requirements for nutrition. Diet-Right, a system has been developed for ideal diet recommendation. It uses pathology reports of users and is based on cloud. It resolves the problem of selecting an ideal diet to fulfill nutrition requirement of CVDs patient. The system use an algorithm named ant colony to determine the optimal list of foods and thus prescribes appropriate diets in accordance with values mention in the pathology report of the patients [1]. The limitation with the system includes recommendations for three meals at different times of a day i.e. breakfast, lunch and dinner with nutrition amount in different food items as per timings and daily needs of the patients. In 1970s, Thomas L. Saaty developed a structural technique for categorizing and examining complicated outcomes, based on mathematics and psychology termed as Analytic Hierarchy Process [2]. Hwang and Yoon proposed a fuzzy method which uses preference order similar to absolute solutions for solving MCDM problems. The technique involves that the selected alternative should be with minimum distance to Positive Ideal Solution (PIS) and with long distance to Negative Ideal Solution (NIS) [3]. The previous works on diet recommendation emphasize more on exterior aspects including necessity, methods, trend, opportunity etc. instead of internal aspects like the nutrients to be present in the diet.In our work nutrients plays a vital role in determining the meals for CVDs patients instead of the other exterior aspects. In the proposed work we have implemented Analytic Hierarchy Process for recommending the diet containing nutrients with due importance. It involves AHP for determining an ideal diet plan comprising of three meals intake for the CVDs patients in order to preserve nutrients in body all through the day. As, AHP formally implement selecting the best substitute comprising of all essential elements our paper aim this in order to determine the ideal diet plan for CVDs patients [4].

2 Methodology

Multi Criteria Decision Making (MCDM) technique is applied as a piece of circumstances when the results depends on the basis and the influence of those models on the outcomes. AHP is a MCDM technique which was introduced and arranged by Thomas L. Satyr of in 1980. It is a to an incredible degree significant system, one which incorporates both numerical instruments and mental methodology [11]. It settles on complex essential initiative a basic endeavor by making pairwise connections of all the fundamental elements in view of requirements. It in like manner settles on pairwise relationships of the decisions in view of all components. The Eigen regards, Eigen vector, Consistency rundown and Consistency extent assist in deciding the outcomes. Consistency extent with regard under 10% presents an anticipated result. AHP uses the use of examination grids (Table 1). So as to play out the analysis, Satyr proposed a primary size of inside and out numbers which is showed up in the table underneath (Table 2). This scale suits the numbers which can be used to present a connection among establishments and the choices in perspective of each measure [8]. Usage of AHP is positive as the eating regimen elective ends up being really easy to do as it melds all the fundamental supplements in the proposed devour less calories in numerous respectable ways. The very first step in diet selection process is to establish the factors to be used to evaluate the diet. In order to meet the factors and their importance for the selection process diet the data is collected by taking reviews from the peoples itself. Based on the above mention decision parameters, each parameter is passed through and given the rank on the scale of four points (i) Less important (1 to 3) (ii) Little important (4 to 5) (iii) Somewhat more Important (6 to 7) (iv)Very important (8 to 9) [4]. A relationship is established between the diets (D1, D2 and D3) and the selection factors (C1, C2, C3 and C4). AHP method is used. Analytic Hierarch Process (AHP) is a very powerful tool for decision making. Analytical Hierarchy Process has its own ability to make rank of the choices, in order to make an effective result out of conflicting objects. Another favored outlook of AHP is that it examines the uniformity of a Decision Maker's responses likewise taking out any biasness in the essential initiative procedure. For these purposes of intrigue and the limit of AHP to assist settle on correct decisions we have used this system in our paper. Entropy technique is one of the assorted systems for discovering weights. This strategy is used for assurance of weight of surveying file in fuzzy mathematics, which joins the opinions of the expert with fuzzy analysis survey, realizing a specific level to sort. It is a blend strategy of subjective examination and quantitative examination [9]. The entropy weight is dictated by the framework developed and, its structure of the choice execution cross section is demonstrated in lattice beneath (Fig. 1).

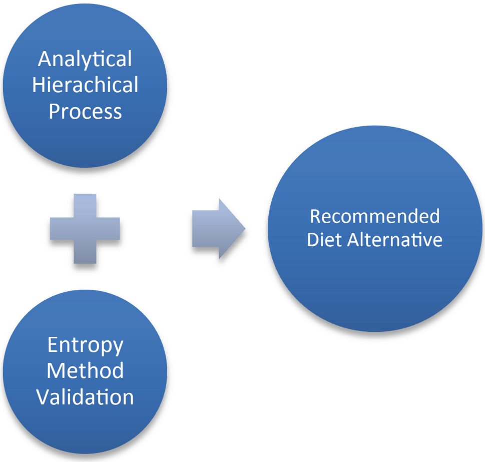

Analytical Hierachical Process
Entropy Method Validation
Recommended Diet Alternative

Fig. 1. Methodology adopted

3 Experimental Work

The Diet outline of cardiovascular disease relies upon a few essential segments. These factors were settled with the help of an investigation driven by the creators. Various famous dieticians and specialists gave their feedbacks. Essential, four parts/measures were settled which are mentioned in the table exhibited as takes after (Table 1).

Table 1. Factors and their importance in the diet

Factors	Importance of the Factors in the Diet [5]
Coenzyme Q10	The human heart beats 86,400 times on an average. Therefore it requires tremendous amount of energy. CoQ10 recharges mitochondria of the cells which are the factories for energy production
Omega-3S	It consists of fish oils which decrease the risk of heart attack or strokes. It also enhances insulin sensitivity
Niacin	It decreases Lp (a) an autonomous danger aspect for heart ailments. It also helps in lowering the cholesterol
Vitamin K	It comes as K1 and K2 flavours. K1 is responsible for clotting and K2 is responsible for regulating the amount of calcium in bones

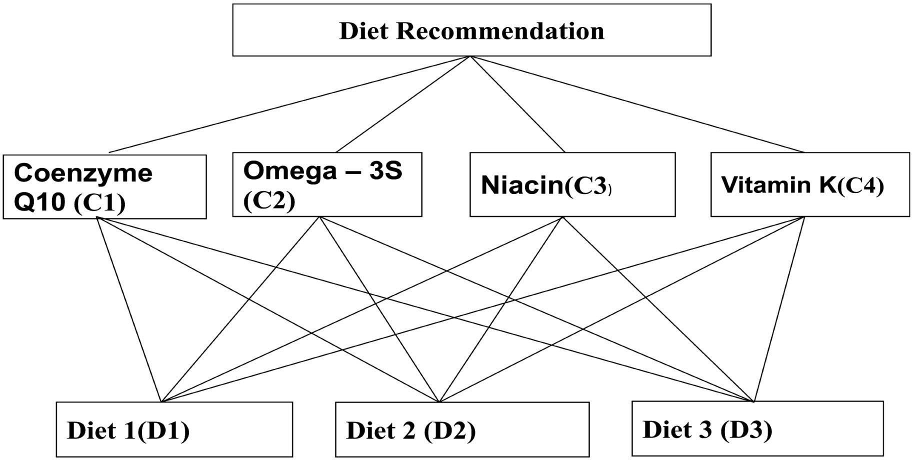


Fig. 2. Proposed hierarchical model

In light of the above parts we are presently showing the elective eating routine outlines which were given to us by esteemed Doctors, who are contemplated pros in medication of Anemia and have enormous inclusion of this area [12]. These eating regimen choices are readied remembering the above parts (Fig. 2).

i. By using AHP method, one best diet is to be selected based on the priority vector out of three alternatives.
ii. First compare all four types of Decision factors w.r.t each other (C1 to C4). Then diet is compared with each decision parameter to show their relative importance.
iii. Calculate the consistency index (C.I) and consistency ratio (C.R) for each matrix and ranked three diets using AHP Methodology.
iv. Validate our work using Entropy Method

The nutritive estimation of the feast is moreover remembered to give the required proportion of supplement required by a particular patient in perspective of the ordinary admission of the supplements. Each eating regimen elective includes three fundamental dinners of the day to be particular: Breakfast, Lunch and Dinner. The components of diets are mentioned in the tabubar underneath (Tables 2, 3, 4, 5, 6, 7 and 8) (Fig. 3).

Table 2. Alternatives of diet and their components

Diet Alternatives	Components of the Diet [6–8]
Diet 1 (D1)	**Breakfast:** One cup Scottish oats with fruits and walnuts garnish, one cup fat free milk, one banana **Lunch:** Oven-bake chicken with Multi-grain bread, sliced avocado and tomato, black beans combined with vinegar and olive oil **Dinner:** Leafy vegetables, Half bowl brown rice, Few blackberries
Diet 2 (D2)	**Breakfast:** Few slices of whole-wheat bread, Six or eight ounces of one percent, yogurt along with blueberries and seeds of sunflower. Three-fourth cup of calcium-fortified juice of orange **Lunch:** Half-bowl brown rice and black beans, Half-bowl spinach or cauliflower with roasted fish or roasted chicken toppings **Dinner:** Three ounces stir-fry chicken with basil, One bowl brown rice with one tbsp diced apricots, One bowl steamed cauliflower, One red wine
Diet 3 (D3)	**Breakfast:** Omelet of one egg and one egg white with one sliced orange and yogurt **Lunch:** Beans or chilies soup low in sodium with toppings of avocado slices **Dinner:** Cucumber raita, Spiced bhindi, one bowl cooked green beans
Diet 4 (D4)	**Breakfast**: One cup Kellogg's all bran wheat flake, 1 cup fat free milk, Half cup strawberries **Lunch:** Tuna salad of half smashed avocado with diced walnuts and grapes **Dinner:** Two bowls baked tofu, one roasted potato, one cup reduced fat milk.

Table 3. Pairwise comparison matrix of criterions

	Coenzyme Q10	Omega - 3S	Niacin	Vitamin K	Eigen Vector
Coenzyme Q10	1.0000	3.0000	7.0000	7.0000	0.5654
Omega - 3S	0.3333	1.0000	5.0000	7.0000	0.3001
Niacin	0.1429	0.2000	1.0000	3.0000	0.0879
Vitamin K	0.1429	0.1429	0.3333	1.0000	0.0466

$\lambda_{avg(max)} = 4.2278$, C.I = 0.0759, C. R = 0.0844

Table 4. D1, D2 and D3 w.r.t Coenzyme Q10

	DIET 1	DIET 2	DIET 3	Eigen Vector
DIET 1	1.0000	0.3333	5.0000	0.2790
DIET 2	3.0000	1.0000	7.0000	0.6491
DIET 3	0.2000	0.1429	1.0000	0.0719

$\lambda_{avg(max)}$ = 3.0649, C.I = 0.0324, C. R = 0.0559

Table 5. D1, D2 and D3 w.r.t Omega - 3S

	DIET 1	DIET 2	DIET 3	Eigen Vector
DIET 1	1.0000	3.0000	7.0000	0.6694
DIET 2	0.3333	1.0000	3.0000	0.2426
DIET 3	0.1429	0.3333	1.0000	0.0879

$\lambda_{avg(max)}$ = 3.0070, C.I = 0.0035, C. R = 0.0061

Table 6. D1, D2 and D3 w.r.t Niacin

	DIET 1	DIET 2	DIET 3	Eigen Vector
DIET 1	1.0000	7.0000	5.0000	0.7306
DIET 2	0.1429	1.0000	0.3333	0.0810
DIET 3	0.2000	3.0000	1.0000	0.1884

$\lambda_{avg(max)}$ = 3.0649, C.I = 0.0324, C. R = 0.0559

Table 7. D1, D2 and D3 w.r.t Vitamin K

	DIET 1	DIET 2	DIET 3	Eigen Vector
DIET 1	1.0000	7.0000	9.0000	0.7854
DIET 2	0.1429	1.0000	3.0000	0.1488
DIET 3	0.1111	0.3333	1.0000	0.0658

$\lambda_{avg(max)}$ = 3.0803, C.I = 0.0401, C. R = 0.0692

Table 8. Reliability index for diet alternatives

	Results	Rank
DIET 1	0.459435	1
DIET 2	0.453883	2
DIET 3	0.086681	3

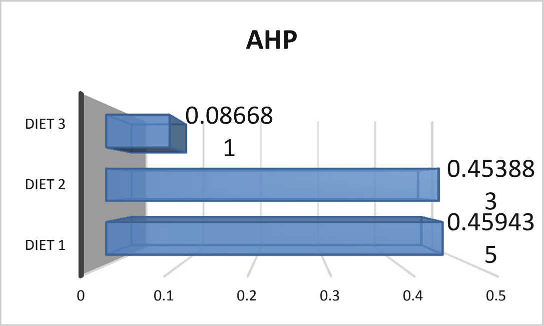


Fig. 3. AHP results

4 Validation of the Work

Entropy method is one of the different frameworks for finding weights. This system is utilized for affirmation of weight of studying document in fluffy arithmetic, which joins the assessments of the master with fluffy investigation study, understanding a particular level to sort. It is a mix procedure of abstract examination and quantitative examination [9]. The entropy weight is managed (Tables 9, 10 and 11).

Table 9. Diets vs Nutrients matrix

	DIET 1	DIET 2	DIET 3
Coenzyme Q10	0.2790	0.6491	0.0719
Omega - 3S	0.6694	0.2426	0.0879
Niacin	0.7306	0.0810	0.1884
Vitamin K	0.7854	0.1488	0.0658

Table 10. Entropy calculations

	Coenzyme Q10	Omega - 3S	Niacin	Vitamin K
D	0.7518	0.7519	0.6802	0.5937
E	0.2482	0.2481	0.3198	0.4063
W	0.203	0.2029	0.2616	0.3324

Table 11. Ranking

	Diet 1	Diet 2	Diet 3
R3I	0.6447	0.2517	0.1036
Rank	1	2	3

5 Results and Discussions

Investigative Hierarchy Process is used to choose the most fitting and reasonable eating routine among the three options recommended or endorsed by master dieticians and exceptionally experienced specialists.

i. By AHP method we observe that the Diet type 1 has the rank I among all other available types of diet.
ii. Obtained rank is same to the rank obtained by Entropy.
iii. Outcomes obtained are validated and thus diet D1 is the best substitute amongst all types of diets.

The results eating regimen among the others (Fig. 4).

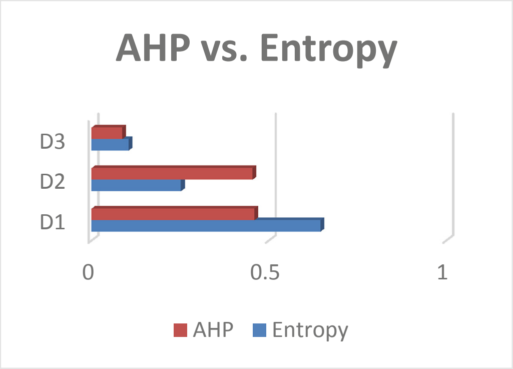


Fig. 4. AHP vs entropy results.

6 Conclusions

India is seen as the growing capital of the world in terms of cardiovascular disease. As shown by recurring pattern measures, India will before long have the most raised number of cardiovascular disease cases. This paper presented three eating regimen choices which are most appropriate to fix and additionally forestall cardiovascular disease. These eating methodologies were figured after a study which included very experienced specialists and master dieticians. These eating routine options comprise the majority of the four most essential/required supplements by a cardiovascular ailment understanding. Along these lines logical procedure, which is AHP a Multi Criteria Decision Making strategy, was utilized to associate those supplement and eating regimen options. We collected the data from esteemed dieticians to figure out the essentials factors in the CVDs patients' diet. Validated the data collected (Calculated consistency). Selecting various diet factors on the basis of the responses given by various dieticians. Diet plans finalised on the basis of selected diet factors. Then we obtained the ranking among the finalised diet plan. Analytical Hierarchical Process (AHP for ranking). Entropy by weights Method(For validation) Thus, it was seen that

Diet elective 1 was the most great and eating regimen 2 was second to it and eating routine 3 was third. This outcome or perception was then approved by utilizing the strategy for Entropy. The Entropy approved the outcomes and gave indistinguishable outcomes from AHP. Henceforth, it very well may be inferred that the relationship of the eating regimen options and supplement, and the choice, that eating routine elective 1 is the best, is substantial and deductively just.

References

1. Rehman, F., et al.: Diet-right: a smart food recommendation system. KSII Trans. Internet Inf. Syst. **11**, 2910–2925 (2017). https://doi.org/10.3837/tiis.2017.06.006
2. https://en.wikipedia.org/wiki/Analytic_hierarchy_process
3. Nădăban, S., Dzitac, S., Dzitac, I.: Fuzzy TOPSIS: a general view. Proc. Comput. Sci. **91**, 823–831 (2016)
4. Saaty, T.: The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation. Mcgraw- Hill, Texas (1980)
5. https://www.betternutrition.com/features-dept/supplements-for-heart-health
6. https://www.mayoclinic.org/diseases-conditions/heart-disease/in-depth/hearthealthy-diet/art-20046702
7. https://www.cookinglight.com/eating-smart/heart-healthy-foods-fuel-cardiac-diet
8. https://www.bhf.org.uk/informationsupport/heart-mattersmagazine/nutrition/cooking-skills/10-heart-healthy-meals-in-less-than-10-minutes
9. Singh, M.P., Dubey, S.K.: Recommendation of diet to anaemia patient on the basis of nutrients using AHP and fuzzy TOPSIS approach. Int. J. Intell. Eng. Syst. **10**(4) 100À108 (2017)
10. Cheng, T., Zhang, C.X.: Application of fuzzy AHP based on entropy weight to site selection of solid sanitary landfill. Environ. Sanit. Eng. **12**(2), 64–67 (2003)
11. Sharma, C., Dubey, S.K.: Reliability evaluation of software system using AHP and fuzzy TOPSIS approach. In: Pant, M., Deep, K., Bansal, J.C., Nagar, A., Das, K.N. (eds.) Proceedings of Fifth International Conference on Soft Computing for Problem Solving. AISC, vol. 437, pp. 81–92. Springer, Singapore (2016). https://doi.org/10.1007/978-981-10-0451-3_9
12. Chi, Y., Chen, T., Tsai, W.: A chronic disease dietary consultation system using OWL-based ontologies and semantic rules. J. Biomed. Inf. **53**, 208–219 (2015)

A Model for Classification of Breast Cell Density Using ANN and Shift Invariance Wavelet Transform ConvNet

Sulaiman Sadiya(✉), and C. A. Hafsath

Department of CSE, ICET, KTU University, Thiruvananthapuram, Kerala, India
sadiya.nader7@gmail.com, hafsath.ca@gmail.com

Abstract. Breast cancer is among world's second most happening cancer in a wide range of cancer. Early location of cancer followed by the best possible treatment can decrease the danger of passings. AI can assist restorative experts with diagnosing the illness with more precision. Where deep learning or neural networks is one of the strategies which can be utilized for the characterization of ordinary and strange breast detection. This exploration presents a double tree complex valued discrete wavelet transform constructed in ConvNet and ANN to conduct breast cell density grouping out of mammography. For mammogram image characterization assignments, customary Convolutional Neural Networks (ConvNet) are: (1) slanted to disregard significant surface data of the image because of the constraints of pooling methodologies, and (2) inadequately vigorous to commotion. To defeat the hindrances, an other area transformation procedure is embraced in ConvNet. In ConvNet (WConvNet) the convolution layer is connected with double-tree complex valued wavelet Through complex valued discrete wavelet transformation, the picture winds up flexible dimensional way, enabling exact characterization. The WConvNet deteriorates the image into various wavelet subbands, and lessens boisterous information. The exhibition of WConvNet is tried on MIAS datasets, and connected for early conclusion of cancer. Contrasted with the conventional ConvNets utilizing maximum valued pooling, trial outcome show that the WConvNet method acquires remarkable evenness and exactness. The study is finished utilizing 322 images of the MIAS database and has brought about characterization achievement rates running from 90% to 94.00% for various breast cell density category (Scattered fibroglandular density, Heterogeneously dense, Extremely dense).

Keywords: Double-tree complex valued discrete wavelet transform · MIAS database · Breast cell density · Double tree complex valued discrete wavelet transform ConvNets (WConvNet)

1 Introduction

Cancer in breast is one of the most common malignancies in women around the world. Subsequently, numerous investigations of programmed analyze and characterization framework have been realized in the writing that can be useful for radiologists and specialists. Inside this extension, mini-MIAS databases of mammogram [1] and Digital

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 60–69, 2019.
https://doi.org/10.1007/978-981-15-0108-1_7

Database for screening Mammogram [2] individually are the commonly used databases. Breast cell density is a significant factor in analyze of cancer from mammographic imaging [3–5]. In this paper, the breast cell density is arranged into three categories, which are Scattered fibroglandular density, Heterogeneously dense, Extremely dense. The feature extraction utilized in this paper is double-tree complex valued discrete wavelet transformation (DTCWT) implemented in ConNet and classified using ANN. Models are trained on an enormous number of datasets and ConvNet structures containing numerous layers. The ConvNet actualize a stratified description of input by sliding to enable extremity and maximum pooling to consolidate significant comparable characteristic [6]. Despite that, maximum pooling down-samples the features that have been extracted, together with the fast decrease in dimensional size, can actuate a corruption of ConNet learning execution [7]. Among different spectral investigation strategies, double-tree complex valued discrete wavelet transformation (DTCDWT) is a wave signal handling strategy that provides translational invariance [8]. The usefulness of calculations can be debased in cancer analysis when the images are in low resolution because of inadequate data open to the ROIs. The utilization of complex valued discrete wavelet transformation to ConNet allows high-recurrence loud information to be removed. To characterize the breast cell density, a precise and parameter-disposed of Double-Tree Complex valued discrete Wavelet ConvNet (WConvNet) is presented. Inside this examination, 322 digital mammogram from the repository was used. Here, fundamentally complex valued discrete wavelet transformation that is coordinated with ConvNet was connected to the digital mammogram [13]. The subset of elements acquired is isolated to two separate DWT decompositions (tree *a* and tree *b*) attribute factor. Through these two separate DWT decompositions image attribute factor some analytical information is determined. Then, 6 digital mammogram attribute vectors altogether including 3 vectors for every part were developed, for which these analytical information is connected to every one of the tree *a* and tree *b* part. The digital mammogram utilized in this investigation is characterized independently utilizing 6 image attribute vector, ConvNet and NeuralNetwork. The outcomes got in the most recent phase of the investigation, by 6 arrangement gathering is consolidated. The outcome identified with the framework was achieved in the last procedure.

2 Related Work

Sun [9] in 2017 introduced an intelligent gear fault diagnosis methodology using a complex wavelet enhanced convolutional neural network. Habibzadeh in 2019 [10] introduced analysis of white blood cell differential counts using dual-tree complex wavelet transform and support vector machine classifier. Lu [12] in 2018 displayed a DTCWT based CNN model for Human Thyroid Medical Image Segmentation. Yaşar [13] in 2018 displayed a System Using Neural Networks and Complex valued discrete Wavelet Transform for tissue Density classification in Digital Mammogram. This paper proposes Double tree complex valued Discrete wavelet Transform implemented in ConvNet models to perform Digital mammogram characterization to decide if its is cancerous or not and This work provides assessment on the performance of characterization of breast cell density.

3 Proposed System

3.1 Database Used in the Study

This examination profited by 322 images from mini-MIAS repository of mammogram, It is the most easily accessed repository and therefore the most commonly used. The repository involves 322 digital mammogram of which 106 has Scattered fibroglandular, 104 Heterogeneously dense and 112 Extremely dense cell density. The first size of digital mammogram from repository is1024 × 1024.

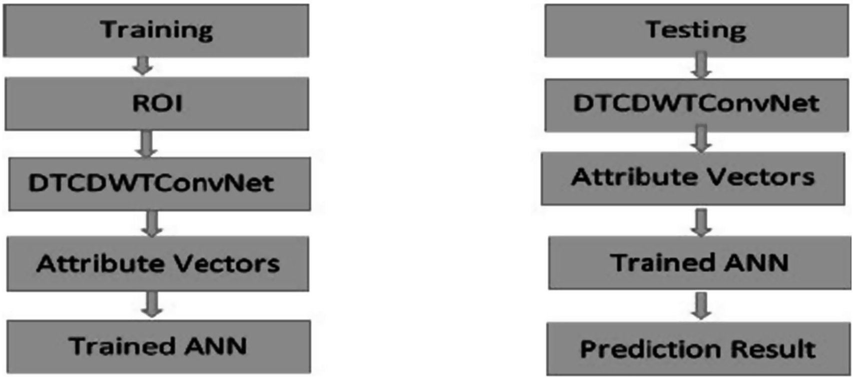


Fig. 1. Training and testing procedure for classification model

3.2 Double Tree Complex Valued Discrete Wavelet Transform

This calculates the complex transform of a signal using two separate DWT decomposition (tree *a* and tree *b*). If the filters used in one are specifically designed different from those in the other it is possible for one DWT to produce the real attribute factor and the imaginary attribute factor. These attribute part are determined 6 distinctive way ($\pm 15^0$, $\pm 45^0$, $\pm 75^0$) for the two DWT decomposition. The block diagram is shown in Fig. 2. In the Fig. 2, h0 is tree *a* high-cut filter factor, *h1* is tree *a* low-cut filter factor, *g0* is reverse tree *b* high-cut filter factor, *g1* is reverse tree *b* low-cut filter factor; separately [13]. In the examination, the real and imaginary argument picture sub-set of dimension 32 × 32 was acquired, Implementing DTCDWT to the picture territories of dimension 256 × 256 in level 3. The Low pass filter LLLL sub band of the picture sub-set was isolated into two separate DWT decomposition and is utilized in acquiring the picture attribute vectors.

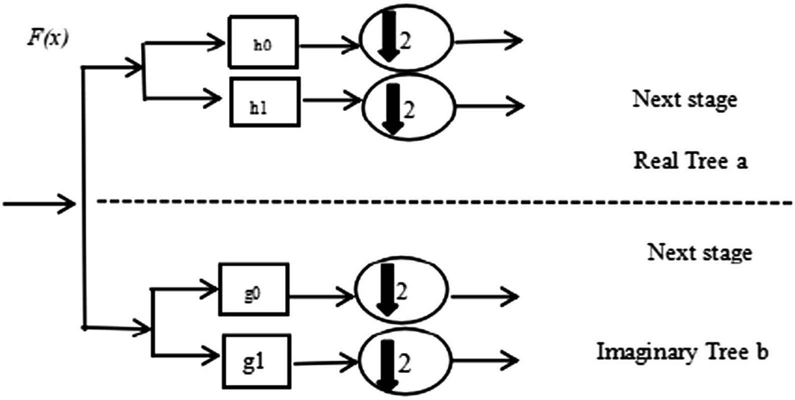


Fig. 2. Double Tree Complex valued Discrete wavelet transformation tree structure

3.3 Digital Mammogram Attribute Vectors

The Low pass filter sub band of the picture sub set was isolated into two separate DWT decomposition (tree *a* and tree *b*). Then for both the parts correlation, contrast, homogeneity, entropy and energies is determined independently. Utilizing these measurable qualities, 3 picture attribute vectors for the two separate DWT decomposition (tree *a* and tree *b*) parts is built [13].

3.4 Convolutional Neural Networks

The ConvNet same as neural network that includes convolution layers/pooling layers that associate with each neuron connected to only subset of input images [12]. ConvNet helps to reduce the number of parameters in the system and makes the computation efficient, ConvNet consists of filter that slide over the complete image and along the way take the dot product between the filter and chunks of the input image [12].

All the picture of consecutive chunks are denoted as *a* = *{a0, a1, ..., an},* when n image tests are given and the yields are *z* = *{z0, z1, ..., zn}*. The convolution layer consists of set of filters. Filtering is independently performed in convolutional layer. In the *j*-th convolution layer, assume that *i*-th filter is M_i, the bias factor is B_i, the output after convolution from small chunk a_y is [12]:

$$C_i^j = f\left(\sum\nolimits_y a_m^{j-1} . M_i^j + B_i\right) \tag{1}$$

After convolution is performed pooling is done to pack the image. Normally most commonly used is max pooling. In max pooling, for the *j*-th layer the yield from small chunk a_y is:

$$p_i^j = MAX\{a_{y_{\pi\varphi}}\}. \tag{2}$$

The scale in pooling is denoted by π and ϕ. When picture data is properly packed after convolving and pooling the ConvNet is ready to learn [11]. In final layer the data is fed to a Neural network for classification of images. Training was done on 70% of the dataset from Mammograms MIAS from an aggregate of 322 images.

3.5 Dual-Tree Wavelet Based Convolutional Neural Networks

A WConvNet architecture is planned in Fig. 3 with DTCDWT link. There are 5 convolution levels, DTCDWT pooling layer (WT1and WT2) linked to Con 3 and Con 4. Con1 and Con2 layers have kernels whose sizes are 1 × 1. Con1 and Con2 layers have similar dimension of input with padding. Con3, Con4 and Con5 is of dimension 2 × 2 and the stride is 2 × 2 by not using padding. WT1 is linked to Con3 and WT2 layer is linked to Con4 for feature extraction [12]. ANN is connected to yield layer and prediction is performed.

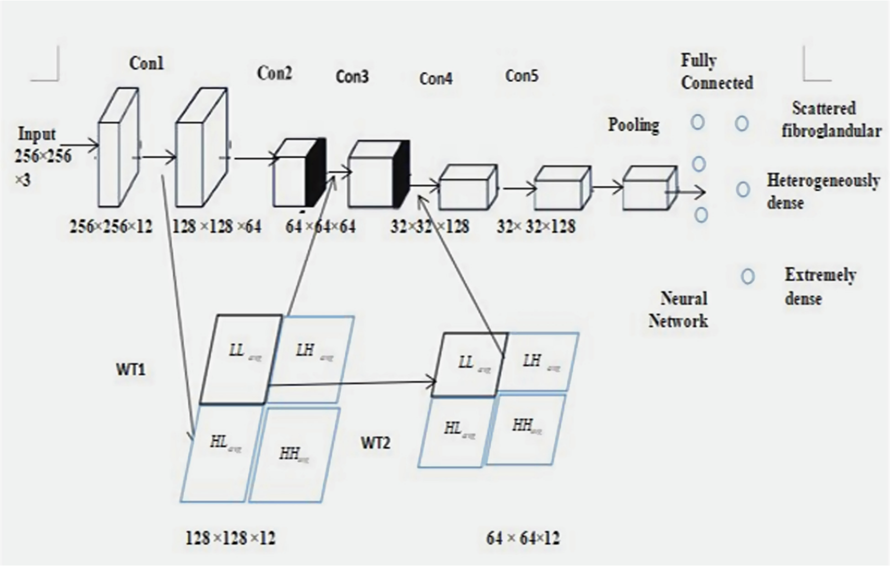


Fig. 3. A WConvNet model. The size of input pictures is 256 × 256. Con: Convolution; WT: DTCDWT in ConvNet

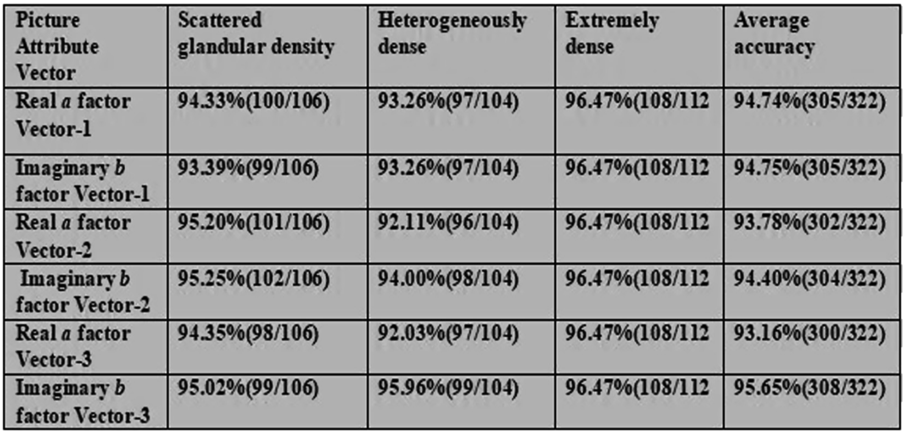
3.6 Artificial Neural Networks (ANN)

In this paper, a lot of ANN are utilized to classify the mammograms in various levels. In the training stage, wavelets coefficients are utilized as a network input. The training procedure proceeds until an acceptable order rate is achieved. In the test stage, the module utilizes a lot of mammogram images to test the framework finding assessment. The number of the output nodes in the neural system module relies upon the classification levels of the framework.

4 Result Analysis and Discussion

In the examination, the outcomes acquired from WConvNet and ANN and the attribute vector group for the three breast cell density gatherings, is presented in Table 1. The value for three breast density groups is 92.03% to 98.21%. The proportion is 93.16% and 95.65%, for mean genuine grouping. Six classification classes set up at the most recent phase of this framework is joined and compared and the characterization outcome is created. The outcomes of this joined framework is shown in Table 2. In Table 2, it is concluded that the framework presented delivered "sure" outcomes of 305 from 322 digital pictures. Here, the framework that creates "sure" outcome is limited to 94.75%. From these outcomes, 280 outcomes was accomplished a genuine characterization, and 25 outcomes was characterized false. Here, 91.8% was grouped as genuine for pictures with "sure" outcome. Also, the genuine breast cell density group of "unsure" picture of this framework is 98.24%. The information acquired from this framework is gathered and found in Table 3. Some of information got utilizing the normal system of CNN is appeared Tables 4 and 5.

Table 1. Characterization results acquired by attribute vectors for various breast cell densities in WConNet and ANN



Picture Attribute Vector	Scattered glandular density	Heterogeneously dense	Extremely dense	Average accuracy
Real *a* factor Vector-1	**94.33%(100/106)**	**93.26%(97/104)**	**96.47%(108/112**	**94.74%(305/322)**
Imaginary *b* factor Vector-1	**93.39%(99/106)**	**93.26%(97/104)**	**96.47%(108/112**	**94.75%(305/322)**
Real *a* factor Vector-2	**95.20%(101/106)**	**92.11%(96/104)**	**96.47%(108/112**	**93.78%(302/322)**
Imaginary *b* factor Vector-2	**95.25%(102/106)**	**94.00%(98/104)**	**96.47%(108/112**	**94.40%(304/322)**
Real *a* factor Vector-3	**94.35%(98/106)**	**92.03%(97/104)**	**96.47%(108/112**	**93.16%(300/322)**
Imaginary *b* factor Vector-3	**95.02%(99/106)**	**95.96%(99/104)**	**96.47%(108/112**	**95.65%(308/322)**

Table 2. Results of proposed system (WConvNet and ANN)

	"Unsure"	"Sure"		Total
		True	False	
Scattered fibroglandulardensity	6	5	95	106
Heterogeneously dense	7	15	82	104
Extremely dense	4	5	103	112
		25	280	
	17	305		322

Table 3. Informations of the proposed system (WConvNet and ANN)

"Sure" result generation capacity	Accuracy of "sure" results			Reduction capacity to two groups of "unsure" results	True result capacity of "unsure" results
305/322	SG	HD	ED	17/17	16/17
	95/100	82/97	103/108		
	95.09%	84%	95.37%		
94.00%	91.8%			100%	98.2%

Table 4. Results of system using ConvNet

	"Unsure"	"Sure"		Total
		True	False	
SG	5	5	96	106
HD	29	15	60	104
ED	13	12	87	112
		32	243	
	47	275		322

Table 5. Informations of the system using ConvNet

"Sure" result generation capacity	Accuracy of "sure" results			Reduction capacity to two groups of "unsure" results	True result capacity of "unsure" results
275/322	SG	HD	ED	47/47	46/47
	96/101	60/75	87/99		
	95.04%	80%	87.87%		
85.40%	88.36%			100%	98.24%

4.1 Test on Datasets

The WConvNet design in Fig. 3 is tried on the MIAS dataset. To comprehend the WConvNet accomplishment, the ordinary ConvNet structure is connected for correlation by expelling the W T 1 and W T 2 layers in Fig. 3 and keeping up a similar number of convolutional filters and the precision is determined for each model [11].

4.2 Results of Datasets

The WConvNet model exhibited the adequacy with a 94.00% precision on the MIAS dataset, beating the customary ConvNet design.

4.3 Digital Mammogram Pictures Classification Outcomes

The assessment results are in Figs. 4 and 5. Contrasting among WConvNet and ConvNet in Fig. 4 WConvNet demonstrates higher precision than ConvNet alone by 10–12%.

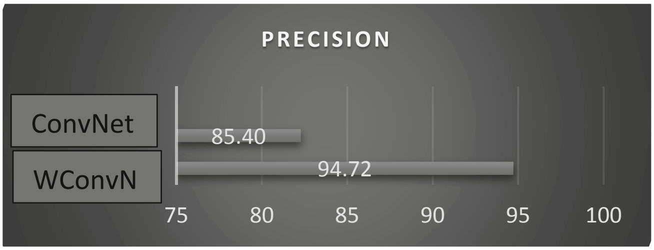


Fig. 4. Comparison of precision between ConvNet and WConvNet models

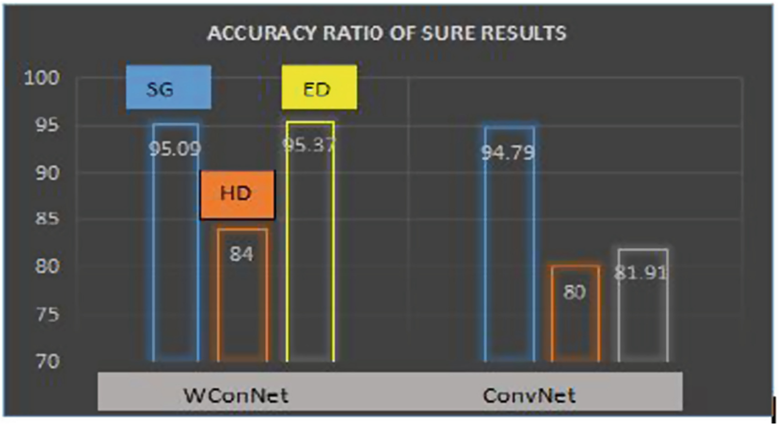


Fig. 5. Comparison of precision ratio of "sure" results for each class between ConvN et and WConvNet model

5 Conclusion and Future Work

The last outcomes got were inspected and reasoned that that the framework capacities with a high limit of 94.72% in getting "sure" results, and of these "sure" results 91.8% are delegated "genuine". Also, the characterization outcome is turn down to two gatherings for the outcomes marked "unsure". In this framework 98.24% of genuine breast cell density group of the picture named "unsure" is accomplished. At the point when "sure" and "unsure" results are thought about, it is seen that out of the 322 images 305 give genuine data to the radiologist who should execute the pertinent classification. This research proposes a WCNN and ANN strategy for characterization of breast cell density. The CWT with prevalence in shift in-variance guarantees viable image pressure execution that is incorporated to the customary ConvNet engineering during the image dimension contraction procedure, and grouped utilizing ANN. Here, WConvNet model is overhauled, It also assessed the DTCDWT based ConvNet (WConvNet) strategy on freely open datafile. From correlation of exactness between WConvNet and customary ConvNet, WConvNet produces understandable picture side, that enhance the precision and stability. The outcomes can be improved by utilizing genetic algorithm. Thus a completely programmed framework is able to create, that separate the digital mammogram pictures named "unsure" as per the outcomes acquired from the ANN training sets.

References

1. Suckling, J., et al.: The mammographic image analysis society digital mammogram database. Exerpta Medica. Int. Congr. Ser. **1069**, 375–378 (1994)
2. Heath, M., Bowyer, K., Kopans, D., Moore, R., Kegelmeyer, P.: The digital database for screening mammography. In: Digital Mammography, pp. 431–434 (2000)
3. Byng, J.W., Boyd, N.F., Fishell, E., Jong, R.A., Yaffe, M.J.: The quantitative analysis of mammographic densities. Phys. Med. Biol. **39**(10), 1629 (1994)
4. Li, H., Giger, M.L., Olopade, O.I., Margolis, A., Lan, L., Bonta, I.: Computerized texture analysis of mammographic parenchymal patterns of digitized mammograms. Int. Congr. Ser. **1268**, 878–881 (2004)
5. Yaşar, H., Ceylan, M.: A novel approach for reduction of breast tissue density effects on normal and abnormal masses classification. J. Med. Imaging Health Inform. **6**(3), 710–717 (2016)
6. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–445 (2015)
7. Graham, B.: Fractional max-pooling, arXiv preprint, arXiv:1412.6071, December 2014
8. Kushwaha, A., Khare, A., Prakash, O., Song, J.I., Jeon, M.: 3D medical image fusion using dual tree complex wavelet transform. In: 2015 Interna-tional Conference on Control, Automation and Information Sciences (ICCAIS), pp. 251–256. IEEE, October 2015
9. Sun, W., Zeng, N.: An ıntelligent gear fault diagnosis methodology using a complex wavelet enchanced convolutional neural network. Material **10**, 790 (2017)

10. Habibzadeh, M., Krzyżak, A., Fevens, T.: Analysis of white blood cell differential counts using dual-tree complex wavelet transform and support vector machine classifier. In: Bolc, L., Tadeusiewicz, R., Chmielewski, Leszek J., Wojciechowski, K. (eds.) ICCVG 2012. LNCS, vol. 7594, pp. 414–422. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-33564-8_50
11. Duan, Y., Liu, F., Jiao, L., Zhao, P., Zhang, L.: SAR Image segmenta-tion based on convolutional-wavelet neural network and markov random field. Pattern Recognit. **64**, 255–267 (2017)
12. Lu, H., Wang, H., Zhang, Q.: A dual-tree complex wavelet transform based convolutional neural network for human thyroid medical ımage segmentation. IEEE (2018)
13. Yaşar, H., Kutbay, U.: A new combined system using ANN and complex wavelet transform for tissue density classification in mammography ımages. IEEE (2018)

Social Data Sentiment Analysis of a Multilingual Dataset: A Case Study with Malayalam and English

Deepa Mary Mathews[1,2](✉) and Sajimon Abraham[3]

[1] School of Computer Sciences, Mahatma Gandhi University,
Kottayam, Kerala, India
deepamarymathews@gmail.com
[2] Federal Institue of Science and Technology, Angamaly, Kerala, India
[3] School of Management and Business Studies, Mahatma Gandhi University,
Kottayam, Kerala, India
sajimabraham@rediffmail.com

Abstract. Opinion Analysis is an articulate methodology that is indivisibly crypt to the sector of Emotional Sciences which concern the individual supposition, emotions or thoughts and thereby identifies the personal deeds. Natural language processing techniques presume that in common most of the user annotations are written in English dialect, but as focus shifts onto processing comments from internet sources such as microblogging services, this becomes progressively harder to guarantee. Conveying the empathy in their own language can be well thought-out as the homey means for expressing the exact opinion and it leads to the generation of multilinguistic societal data. So the challenge is to analyze the formal textual content along with the informal and mixed linguistic nature of social data. This prompts the need of Sentiment Analysis in multilingual dialects. This article surveys the methodologies used for analyzing the sentiment of multilingual data and proposed a model built using Long Short Term Memory approach.

Keywords: Deep learning · Multilingual · Malayalam · Opinion mining · Sentiment analysis · Long Short Term Memory

1 Introduction

People have always had a curiosity in what individuals assume, or what their opinion is. With web-based social networking channels such as Facebook, Twitter, and LinkedIn, it is becoming viable to automate and judge what well-liked opinion is on a given subject, news thread, item, or brand. Viewpoints that are mined from such feeds can be helpful for decision making. Reviews that are accumulated can be examined and displayed in such a way that it turns out to be easy to recognize if the frame of mind is optimistic or off-putting. This enables the individuals or business to be proactive instead of reactive when a

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 70–78, 2019.
https://doi.org/10.1007/978-981-15-0108-1_8

pessimistic conversational thread is budding. Instead, positive response can be known thus permitting the identification of product advocates or to see which parts of a business tactic are working. When applied to social media channels, it can be used to mark the prickles in the sentiment, thereby enabling to identify for instance the potential product promoters. So the opinion characterization is nowadays a potential and intense research theme.

1.1 Motivation

Albeit the fact that English is the universal language, according to the statistics [1], just 28.6% of the Internet clients used to commune in English. Dialect-switching ensue when someone uses two or more languages in a single piece of text. These days, an immense part of the population is able to handle two or more dialects. English as being the universal language is highly possible to be dialect-switched with the local tongue to form multilingual yet, linguistically right and imperative sentences which necessitates the need of analysis in dialects apart to English [2].

In India, among the roughly thirty official dialects, more than thirty five million individuals spreading on the districts of Kerala, Pondicherry and Lakshadweep are having Malayalam as their local tongue. The rules and methods are different for the Malayalam dialect compared to the English language. A quick search shows that less works have been done on the Malayalam dataset [3]. Additionally it is noticed that the majority of the clients while communicating their suppositions in Malayalam, they used to utilize some common words in English like beautiful, good and so forth rather than the Malayalam term. So in this article multilingual which comprises of a blend of Malayalam and English tokens is consider for analysis.

The rest of the article is structured as follows. Brief descriptions about major contributions on Sentiment Analysis in Malayalam dialect and various approaches to multilingual datasets are summarized in the next section. In Sect. 3, the proposed methodologies used to address the research questions and the implementation are depicted followed by the Sect. 4 which sketches and discusses the experimentation results and proposed solutions for dealing with this type of multilingual dataset. The conclusion and the future scope of the work are delineated in the Sect. 5.

2 Related Works

As dialect changing nature of the Internet users' increases, the growing need to toil on blend of languages arises. The main issue in performing the multilingual sentiment analysis is the lack of resources in the dialects [4]. A common approach to handle the mulitlingual dataset is to translate them to the unilingual dataset where the resources are available [5]. Languages that have been considered for the study includes Chinese, Japanese, German, French, Italian, Spanish, Swedish, Arabic, Romanian etc, but only less works considered multilingual dataset. The resources constructed for a dialect is not applicable to other dialects as each having its own unique constructs [6] Igor Mozetic, et al. applied different supervised algorithms to the datasets in the selected thirteen dialects and the role of human annotators in multilingual sentiment classification is depicted [7]. Opinion mining generally reliant on built lexicons found in dictionary or corpora while dealing with informal and scarce resource languages [8]. The methodologies commonly used for the multilingual studies are lexicon, corpus or translator based and the approaches based on concepts and sentics [9]. The contributions using deep learning algorithms in other formal languages such as Spanish, Japanese, Chinese and German are portrayed respectively in [10–13]. More focus is required to improve results in the multilingual sentiment analysis discipline [14].

The study also depicted that less works considers Indian languages especially Malayalam. Analyzing the opinions marked in Indian languages micro text using recurrent neural network is depicted in [15] and [16]. The major articles that mentions and explains the sentiment analysis on Malayalam dialect are summarized in [3]. The deep learning frame work used for Malayalam sentiment analysis is depicted in [17]. In this article, the multilingual data which consists of Malayalam and English dialects is considered for the experimentation and an LSTM based model is proposed to perform their sentiment analysis.

3 Proposed Methodology and Implementation

The ratio of each linguistic phrase in the dataset considered is vital and so necessitates concern while analyzing the opinion marked in the review. As no standard dataset is available for the Malayalam dialect, human annotation is required for labeling the dataset considered for the experimentation. The lexicons in varied language need to be handled differently. The YouTube reviews are extracted using YouTube Comment Scraper tool which is a free web-based tool that can be used to scrape the comments from a YouTube video and save them in either a JSON or CSV format. Here the user comments marked in that video information for the search query "Sabarimala Women Entry Issue" is scraped and extracted around 16580 comments. The proposed algorithm for the multilingual social data sentiment analysis is depicted in Algorithm 1. The impact of data preprocessing phase in sentiment analysis of Malayalam reviews is depicted in [18].

3.1 Data Processing Phase

The reviews are tokenized into fitting units to build a representation of the data which then undergone the process of cleansing to remove the noises and punctuation marks. As every language has its own specific character patterns and frequencies, they need to be handled differently. An off-the-shelf language identification algorithm called langid is used to detect the language of the tokens. In

Algorithm 1. Multilingual_SentiAnal

```
ReviewsList ← multilingual_user_reviews;
enum noisy = {punctuations, URLs,digits, hashtags, extraspaces};
languages = langid.setLanguages{"ml","en" };
foreach row in tempReviewsList do
    ReviewsList ← ReviewsList.apply(word_tokenize);
    if ReviewsList.contains noisy then
        ReviewsList.remove(noisy);
    landet ← ReviewsList.apply(langid.classify);
    if ReviewsList.langdet is én´ then
        ReviewsList ← ReviewsList.remove(nltk.stopwords);
    if ReviewsList.lang is ḿl´ then
        ReviewsList ← ReviewsList.remove(mlstopwords);
Create a unique vocabulary of tokens and an indices vector;
Remove if any outlier reviews;
Pad all the reviews to a specific sequence length;
Split the dataset into training set, test set and validation set;
Define the hyperparameters for building the bidirectional LSTM model;
Define the layers of the model and classify the reviews;
Test the bidirectional LSTM classifier model;
Fine tune the accuracy of the model by modifying the hyperparameters;
```

view of the language detected, the dataset is splitted into unilingual datasets and afterward experienced various text processing phases. The percentage of unilingual lexicons in the multilingual dataset used for experimentation is delineated in Table 1. The tokens in English get converted to lower case and the stopwords are removed. While annotating the dataset, the main challenge is that the annotator should be ease in handling the languages used in the review. So an interpretation stage is additionally required to convert the multilingual to unilingual dataset.

Table 1. Percenatage of unilingual lexicons

Language identified	Percentage of lexicons
Malayalam	37172
English	72680

Table 2. Feature data

Type	Feature shapes
TrainSet	(12976,200)
ValidationSet	(1622,200)
TestSet	(1622,200)

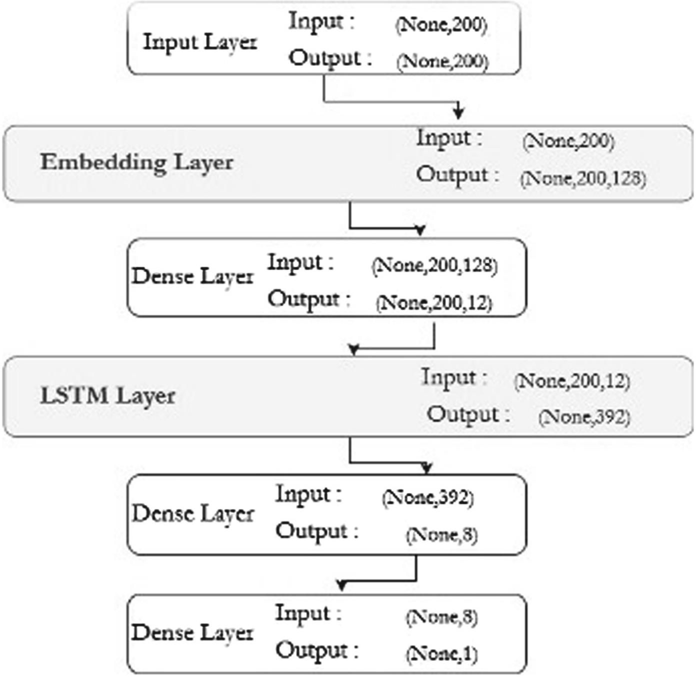


Fig. 1. Proposed LSTM model for multilingual social data sentiment analysis

Google Translate is a free multilingual machine translation service that quickly deciphers words, phrases, etc between English and over 100 other languages.

3.2 Data Encoding Phase

The audits are stored as individual list elements like [review_1, review_2, .., review_n]. The vocabulary of tokens and the respective indices vector of every audit are created. In index mapping dictionary, the recurrently occurring words are assigned with lower indexes. Utilizing this dictionary, the audits are encoded by replacing the individual terms in audits with integers. Now each individual review is a list of integer values and all of them are stored in one huge list.

3.3 Removal of Outliers

By analyzing the length of the audit, it is found that some entries, nearly 360, are of zero length. These were pulled out from the dataset. Likewise it is found that there are quite a few reviews that are extremely long. They are manually examined to check whether to include or exclude them from our analysis. To manage the extremely long or short audits, padding/truncating have been done to make all to a particular length. This sequence length is same as number of time steps for LSTM layer. Here the length of the vector is set as 200 and is fed to the model.

3.4 Training, Validation and Test Dataset Split

Once the data is encoded and outliers are removed, the dataset is parted to three sets – the training_set, validation_set and the test_set. The ratio used for parting is taken as 80% for the training_set and the rest for validation and test set. The shape of the resultant feature data is shown in Table 2.

3.5 LSTM Model Creation

The meaning of a word in the audit relies upon the context of the previous text and accordingly a memory is required to store the word dependencies. The LSTM network consists of a memory cell which is able to handle the dependencies between the terms. A memory cell in LSTM is composed of four key elements: an input_gate, a neuron with a self connection, a forget_gate and an output_gate. The neurons ensure that the condition of a memory cell can stay constant from one time step to another. At each time step t, the values of the input_gate I_t, the forget_gate F_t, the output_gate O_t and the memory cell C_t is used to calculate the output of the hidden layer H_t. The forget_state figures out what to discard from the cell state. The inputs to the forget_state are H_{t-1} and X_t where H_{t-1} is the output from the previous step. To generate the output values to be in between 0 and 1, a sigmoid activation function (σ) is used, where 0 means nothing to remember and 1 to remember all. The equation used to calculate the value of forget_state can be given as

$$F_t = \sigma(W_f.X_t + U_f.H_{t-1} + b_f) \tag{1}$$

Accordingly, what data will be forwarded to the cell is resolved. The inputs are H_{t-1} and X_t. The input layer initially applies a sigmoid layer over the input to figure out which portion of the memory need to be refreshed and a tanh layer is then applied to produce C_t. These two values are then consolidated to refresh the cell state Ct. The state C_{t-1} is multiplied by F_t, to wipe out the terms that are not needed further and the new information is forwarded to the memory. Equations to calculate the values can be given as

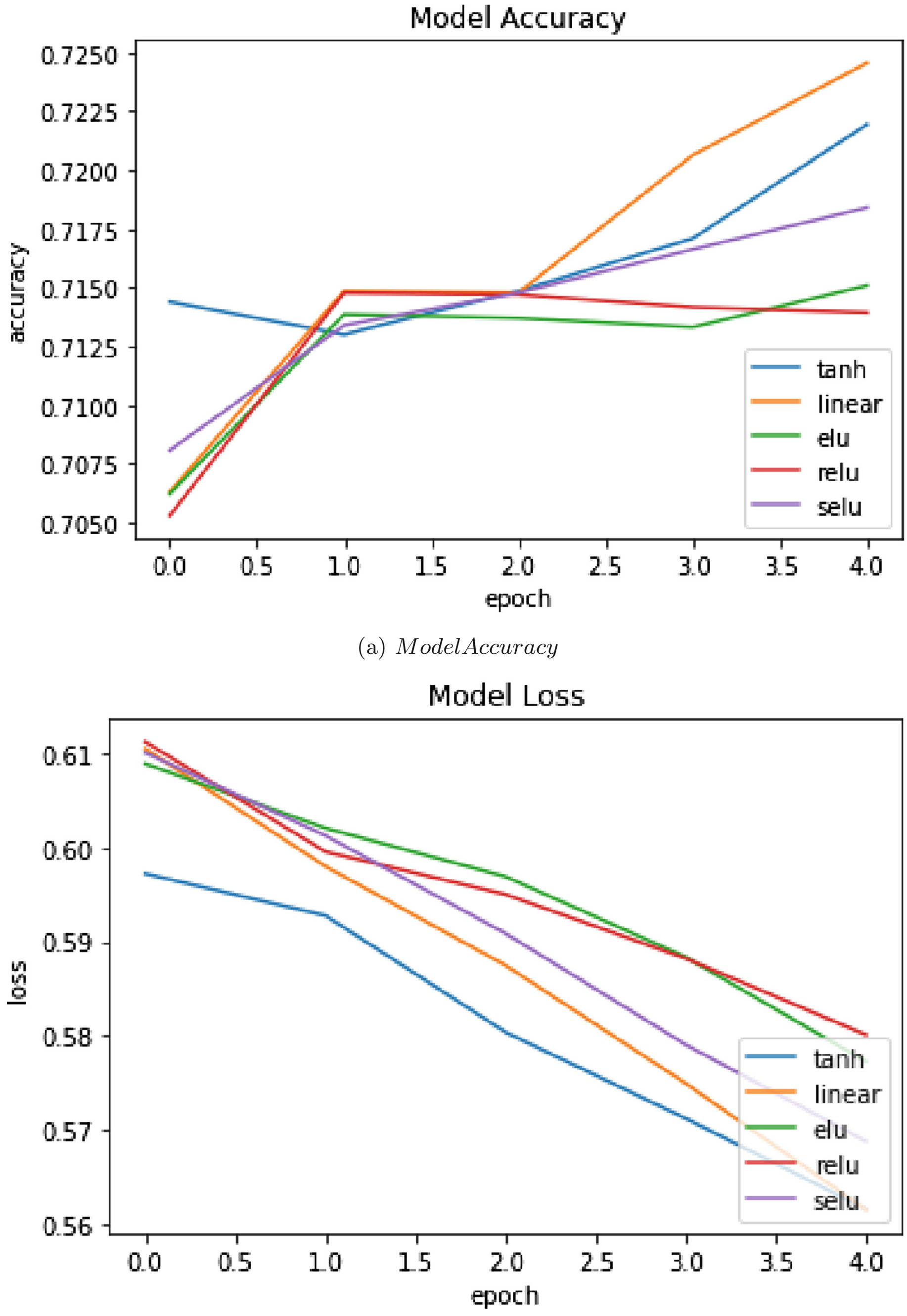
$$I_t = \sigma(W_i.X_t + U_i.H_{t-1} + b_i) \tag{2}$$

$$C_t = (I_t.(tanh((W_c.X_t + U_c.H_{t-1} + b_c) + F_t.C_{t-1}) \tag{3}$$

The value of the output_gate O_t can be determined by applying a sigmoid layer to the H_{t-1} and X_t as given in Eq. (4). C_t is then altered by a tanh function and these changed cell values are multiplied by O_t to create H_t Eq. (5). This output will be pipelined to the next stage in the network.

$$O_t = \sigma(W_o.X_t + U_o.H_{t-1} + b_o) \tag{4}$$

$$H_t = (O_t.tanh(C_t) \tag{5}$$



(a) $ModelAccuracy$

(b) $ModelLoss$

Fig. 2. Model accuracy and loss using various activation functions

3.6 Classification Phase

As depicted in Fig. 1, in the proposed model there are four layers. The input is the set of integers where each integer represents a term in the review. The term is fed into the next layer (embedding_layer) where the term is changed to a vector. An LSTM layer is then added where each feature is multiplied by a weight. The final layer is the output layer fed with a neuron where the outputs from the LSTM layer is taken and using the sigmoid function, a value of 0 or 1 is generated.

Dropout is used to avoid the over-fitting issue and 0.5 is fixed as the dropout parameter for the experimentation. Adam is used as the optimizer for the model. Various activation functions like Relu, tanh, Selu, linear and Elu are used for the experimentation.

4 Results and Discussions

Here 572,245 parameters are considered for the model to be trained. As we need to classify the reviews only into two classes, the loss function used by the model is binary crossentropy. The Fig. 2 depicts the accuracy (2a) and the loss (2b) values for each epoch on training and validation set. Various activation functions like relu, tanh, linear, selu and elu are used for the experimentation. The embedded dimension used for experimentation is 128. Both the input (I_t) and the output (O_t) gates outputs a value between 0 and 1. The tanh function converges faster and the gradient computation is less expensive. Relu seems more successful in computer vision applications.

The diagrams (Fig. 2) depicts that the classifier using linear activation function outperforms the other models. The accuracy of the classifier model using various activation functions is in the range of 70 to 75% which is a significant result in this line.

5 Conclusion and Future Scope

The article depicts the analysis of reviews in multilingual dataset which consists of Malayalam and English words using Bidirectional Long Short Term Memory deep learning approaches. The reviews in Malayalam dialect are extracted and physically annotated. The accuracy of the classifier model using various activation functions is in the range of 70 to 75% which is a significant result in this line. The increase in the accuracy of the classifier can be obtained by fine tuning the model and is a scope for future research.

References

1. https://unbabel.com/blog/top-languages-of-the-internet/. Accessed 4 Feb 2019
2. Boiy, E., Moens, M.F.: A machine learning approach to sentiment analysis in multilingual Web texts. Inf. Retr. **12**(5), 526–558 (2009)

3. Mathews, D.M., Abraham, S.: Sentiment analysis on Malayalam language: a survey. Int. J. Manag. Technol. Eng. **8**(XII), 1046–1054 (2018)
4. Balahur, A., Turchi, M.: Multilingual sentiment analysis using machine translation? In: Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis. Association for Computational Linguistics, pp. 52–60 (2012)
5. Denecke, K.: Using SentiWordNet for multilingual sentiment analysis. In: IEEE 24th International Data Engineering Workshop, ICDEW 2008, pp. 507–512. IEEE (2008)
6. Mirchev, U., Last, M.: Multi-document summarization by extended graph text representation and importance refinement. In: Innovative Document Summarization Techniques: Revolutionizing Knowledge Understanding, pp. 28–53. IGI Global (2014)
7. Mozetic, I., Grcar, M., Smailovic, J.: Multilingual Twitter sentiment classification: the role of human annotators. PLoS ONE **11**(5), e0155036 (2016)
8. Lo, S.L., et al.: Multilingual sentiment analysis: from formal to informal and scarce resource languages. Artif. Intell. Rev. **48**(4), 499–527 (2017)
9. Riloff, E., Wiebe, J.:: Learning extraction patterns for subjective expressions. In: Conference on Empirical Methods in Natural Language Processing, pp. 105–112 (2003)
10. Paredes-Valverde, M., Colomo-Palacios, R., Salas Zarate, M., Valencia-Garcia, R.: Sentiment analysis in Spanish for improvement of products and services: a deep learning approach. Sci. Program. 1–6 (2017). https://doi.org/10.1155/2017/1329281
11. Kanayama, H., Nasukawa, T., Watanabe, H.: Deeper sentiment analysis using machine translation technology. In: Proceedings of the 20th International Conference on Computational Linguistics, COLING 2004 (2004)
12. Chen, H., et al.: Fine-grained sentiment analysis of Chinese reviews using LSTM network. J. Eng. Sci. Technol. Rev. **11**(1), 174–179 (2018)
13. Cieliebak, M., et al.: A Twitter corpus and benchmark resources for German sentiment analysis. In: Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media (2017)
14. Evans, D.K., Ku, L.-W., Seki, Y., Chen, H.-H., Kando, N.: Opinion analysis across languages: an overview of and observations from the NTCIR6 opinion analysis pilot task. In: Masulli, F., Mitra, S., Pasi, G. (eds.) WILF 2007. LNCS (LNAI), vol. 4578, pp. 456–463. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73400-0_57
15. Seshadri, S., Madasamy, A.K., Padannayil, S.K.: Analyzing sentiment in Indian languages micro text using recurrent neural network. IIOAB J. **7**, 313–318 (2016)
16. Bhargava, R., Arora, S., Sharma, Y.: Neural network-based architecture for sentiment analysis in Indian languages. J. Intell. Syst. **28**, 361–375 (2018)
17. Kumar, S.S., Kumar, M.A., Soman, K.P.: Sentiment analysis of tweets in malayalam using long short-term memory units and convolutional neural nets. In: Ghosh, A., Pal, R., Prasath, R. (eds.) MIKE 2017. LNCS (LNAI), vol. 10682, pp. 320–334. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-71928-3_31
18. Mathews, D.M., Abraham, S.: Effects of pre-processing phases in sentiment analysis for Malayalam language. Int. J. Comput. Sci. Eng. **6**(7), 361–366 (2018)

Trajectory Data Publication Through Individualized Sensitive Stay Location Anonymization

N. Rajesh[1](✉), Sajimon Abraham[1], and Shyni S. Das[2]

[1] Mahatma Gandhi University, Kottayam, India
rajeshshyni2000@gmail.com, sajimabraham@rediffmail.com
[2] SAS SNDP Yogam College, Konni, India
shynirajesh2000@gmail.com

Abstract. The abundance of GPS embedded devices accumulates trajectories in an excessive scale and it is enriched with personal information. While publishing traces for research activities, we must ensure to publish the anonymized trajectory in order to prevent the disclosure of individual privacy. During anonymization, we need to anonymize the stay locations where the user considered it as most sensitive instead of anonymizing all locations. For finding and extracting the most sensitive stay locations, we adopt a new method by considering the individual spatial and temporal factors using SSLF function. This combines the stay points within a threshold distance to form stay locations and also anonymize these locations in a stay zone using the generalization SSLA approach. The proposed model is tested with a real-world dataset and it guarantees a better trade-off between privacy and utility compared with other models of same nature.

Keywords: Sensitive stay locations · Spatio-temporal · Trajectory anonymization

1 Introduction

In recent years, the use of mobile phones and other embedded GPS devices like navigation systems collects and stores the mobility traces of individuals in an unprecedented manner. These huge mobility traces are called spatio-temporal trajectories, because it contains the information like latitude and longitude about a location and the time that the user has visited. The collection, storing and management of these movement traces is little bit a cumbersome. The mobility traces are very important for the mobility management people, since it is essential for the efficient traffic monitoring and also for the smooth running of smarter cities. But publishing the spatio-temporal trajectory database in its raw form definitely raises the issue of individual privacy. Because it contains many sensitive parameters like religious habits, socio-economic preferences, sexual preferences, health details etc. Even though the spatio-temporal trajectory is a must for the research activities, without proper anonymization of these could definitely results in the leakage of personal privacy details.

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 79–90, 2019.
https://doi.org/10.1007/978-981-15-0108-1_9

For the protection of sensitive data, simply removing external identifiers is not adequate enough for the trajectory protection. The one work which was carried out earlier, deals only with significant stay point protection [1]. The succeeding one work considers the location frequency of the vehicles parked and anonymizing higher frequency parking locations are enough for the trajectory protection [2]. The authors in [3] suggests that above methods are not sufficient against moving attacks and a k-correlation model along with the mix of location frequency function and inverse user frequency function is better for the protection of location trajectory data. But all these have resulted in unnecessary protection in certain parts of trajectory data. These methods were not considered the temporal factors well.

In this paper, we propose a new model which considers the location frequency within the stay location and also considers the frequency of the days in which the user has stayed. A sensitive stay location finder function considers these factors and suggests only one or few stay locations which are most sensitive to the user. This avoids the over protection problem that are mostly found in the earlier works.

The rest of the paper is organized as follows. Section 2 discusses various privacy attacks and existing models. In Sect. 3, some of the basic and essential problem definitions which needed for this work are given. Section 4 states the proposed structure and methods which we adopted for the work. The experimental setup and evaluations are described in Sect. 5. Finally, Sect. 6 is the concluding comments.

2 Related Work

In this section describes the various privacy attacks on location and trajectories and also the trajectory protection models in a briefly manner.

2.1 Privacy Attacks

In order to prevent privacy attacks, we must know the capacity of the adversary whom the attack is planning. An adversary is characterized based on the knowledge that he/she has acquired from the history trajectories and also the outline of the attack that he/she is going to target. Based on these, there are two types of adversaries: (a) weak adversary and (b) strong adversary. The weak adversary is having a little knowledge about spatio-temporal components and where as a strong adversary has the wider knowledge about the user movements by analyzing QID’s (Quasi-Identifiers) and also linking with related known databases.

Some of the attacks that commonly associated with location are location oriented attacks, multiple query requisition attacks, spatial knowledge attack, maximum moving area attack etc. *Location oriented attacks* are occurred when the adversary is inferring the user movements based on continuous queries or snapshot queries. There are two kinds of attacks may be possible when issuing snapshot query: (a) query sampling attack [4] and (b) location linking attacks [5]. The *multiple query requisition attacks* were happened when the adversaries are trying to know the actual location of the user requester with the help of multiple spatial queries. The idea [6] of extracting exact locations of the service requester by analyzing various cloaked region, since these requests are from various cloaked regions. The *spatial knowledge attack* [7] is happened when the user requests a LBS (Location Based System) query from a location and also he doesn't want to disclose this location to anyone. For that the user issues an obfuscated region instead of actual location to the LBS. But a strong adversary is having semantics about this location and inferred easily. This type of attacks occurred when the user is unknowingly ignores the real semantics of the location. The *maximum moving area attack* occurs when the adversary has the knowledge about the user movements and the maximum boundary area that he may cover in two succeeding position coordinate update. The authors in [8] developed a method with spatio-temporal transformations to handle this type of attacks.

During trajectory data publication, simply removing external identifiers doesn't guarantee the privacy attacks. Inversion attack is one of the trajectory privacy attacks, which is happened when an adversary is aware about the identity of *k* potential users from their request. A solution to this type of attack is the reciprocity which was suggested by the authors in [9]. Another type of trajectory attack is the inference attack, which states that unlawfully gaining knowledge about a subject by the analysis of data. This can be handled by different obscuration methods [10] like spatial cloaking, noise addition, rounding etc. Query tracking is another type of attack on trajectory and these are possible when a user is cloaked with different users at various instances during the lifetime of query. By the usage of memorization property we can prevent query tracking attacks [4].

2.2 Privacy Models

The base of all privacy protected anonymity models either in location or in trajectory publishing is *k*-anonymity [11]. The authors in [12] extended the concept of *k*-anonymity with (*k*, δ) anonymity by considering the imprecision errors during the data acquisition by the GPS devices and δ being the possible location imprecision. The work in [13] suggests a micro aggregation approach for the privacy protection in trajectory data publishing with the use of micro-data anonymization. They clustered the trajectories based on the similarities with a measure of at least *k*-size and replaced these trajectories with synthetic data that are accumulated in the actual trajectories or in the visited locations.

Some of the works are based on the power of QID's. The authors in [14] presented a model called LKC- privacy model, where the parameter L is the maximum background knowledge held by the adversaries. Another work [15] proposes k^{m}-anonymity model for the location trajectory data, and while publishing the trajectory details it doesn't require the vast details about the QID's or the distinction between sensitive and non-sensitive data. In [2], the authors proposed to prevent de-anonymization attacks by modeling an individual mobility based function using a mobility Markov chain. An attack based semi supervised learning approach was presented by the authors in [16] to infer the riding trajectory of the user. A scalable sanitization method (SafePath) for publishing differentially-private trajectories by considering trajectories as a noisy prefix tree and which was proposed by the authors in [19]. The authors in [20] introduced a privacy model $k^{t,\epsilon}$ – anonymity, which is a recent refinement of k-anonymity using the generalization approach and it gives a good solution to probabilistic and linkage attacks against mobile user's trajectories.

3 Problem Definitions

The traces left by the moving objects under LBS are trajectories, is a sequence of spatio-temporal points. This section contains some important basic definitions that are needed to illustrate the trajectory privacy concepts used in our work.

Definition 1 (Trajectory). *A trajectory* Trj *is a sequence of time-stamped spatial points represented as* Trj = *{U_{id}, p_0, p_1, …………., p_n}, where p_i = {la_i, lo_i, t_i}, in which (la_i, lo_i) is representing latitude and longitude and t_i represents the timestamp at this point. Also the U_{id} is the trajectory-id of the user U.*

We also use TR(U) to represent the whole trajectories belongs to U in the historical dataset.

Definition 2 (Stay point). *A stay point* $\mathrm{Sp_i}$ *is a five tuple represented as {S_{id}, la_i, lo_i, arr_{ti}, dep_{ti}}, where S_{id} is the stay point identifier, (la_i, lo_i) is stay point's spatial coordinate and also arr_{ti} and dep_{ti} are the arrival and departure time of the user from point.*

A stay point is formed when the user wishes to stay at somewhere for over a user specified threshold time. The stay points may be different even if the person made multiple visits in the same building or room. This may happened because of the person may have moved a little or due to the imprecision of the GPS (Global Positioning System). So in order to tackle this we have to define another notion called *stay location*.

Definition 3 (Stay location). *A stay location* $\mathrm{STl_i}$ *is the collection of stay points within the user specified distance threshold, µd. i.e.,* $\mathrm{STl_i}$ = *{SL_{id}, Sp, µd}, where SL_{id} is the stay location identifier and Sp is the collection of all stay points which is located within the distance threshold.*

These stay locations may be the real-world places like restaurants, shopping mall, cinema theatre, parks, religious places etc. These stay locations are made available to all the users who may wish to stay or visit.

Definition 4 (Sensitive stay location). *A sensitive stay location* $\mathrm{SSl_i}$ *is the stay location, which is identified when the following conditions were satisfied:*

(1) A stay location where the user U_i has stayed for long time than the other stay locations
(2) In this stay location, the user has to visit more than once
(3) The no. of days visited to this stay location also more than once.

This stay location is considered to be the most sensitive stay location and these are the locations which have to be anonymized before the trajectory data publication in order to preserve privacy of the individual. Otherwise the published trajectory details may contain the sensitive information and publishing without proper anonymization mechanism will certainly results in the disclosure of vital information's like religious customs, health details, social or sexual preferences etc. into the hands of adversaries.

Definition 5 (Stay zone). *A stay zone* SZ_i *is an area which is formed by combining the sensitive stay location along with a k no. of adjacent grids containing stay locations during the anonymization process for maintaining privacy.*

During the publication of trajectory details to the public for research, we gave only the stay zone's bottom right corner coordinate and upper left coordinate of the zone instead of actual sensitive stay location coordinate and the other coordinates outside the zone will be published as such without any alteration.

4 Proposed Solutions

The proposed solution consisting of three major phases, namely sensitive stay location extraction, anonymization of sensitive stay location and the utility analysis. These phases are described as follows.

4.1 Overview of the Solution

The main aim of this work is to protect the sensitive stay location information from the hands of adversaries at the time and after the trajectory data publication. We assume that, the adversaries are capable of extracting the information from the published database by linking with other known databases. So our aim is to protect at least the sensitive stay location information from the published trajectories. Our main task is to anonymize the input historical trajectory database ITrD to a publishable version OTrD. The major and minor phases are shown through the Fig. 1 and also described as follows.

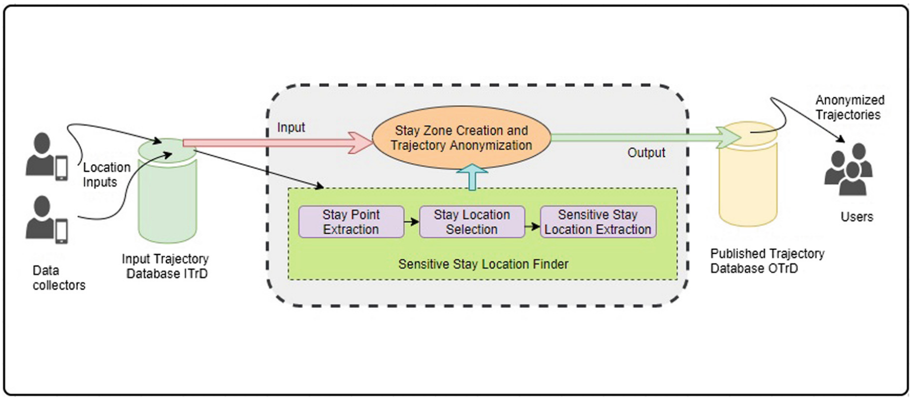

Location Inputs
Data collectors
Input Trajectory Database ITrD
Input
Stay Zone Creation and Trajectory Anonymization
Output
Stay Point Extraction
Stay Location Selection
Sensitive Stay Location Extraction
Sensitive Stay Location Finder
Anonymized Trajectories
Users
Published Trajectory Database OTrD

Fig. 1. Proposed approach – Sensitive Stay Location Anonymization (SSLA)

4.2 Stay Point Extraction

The stay points are extracted using the methods as adopted in the work [1] with minor modifications. Stay points are the location points where the user has arrived at a time arr_t and departed from the location on dep_t. We have to find stay duration by finding the difference between $\mathrm{dep_t}$ and $\mathrm{arr_t}$. i.e., | $\mathrm{dep_t} - \mathrm{arr_t}$ | > = δt, the user specified time threshold.

4.3 Selection of Stay Location

The stay location STl, is the collection of adjacent stay points spread over a user specified threshold distance, µd. Here we uses the Haversine distance measure instead of Euclidean measure, which is very suitable for the spatial distance calculations since the former considers the spherical nature of the earth very well.

4.4 Extraction of Sensitive Stay Location

The extraction of sensitive stay location is done through three steps. First step is to find the stay location frequency. For this purpose we use the following equation.

$$\mathrm{SLF}\left(\mathrm{U_i}, \mathrm{STl_j}\right) = \frac{\mathrm{fr}\left(\mathrm{U_i}, \mathrm{STl_j}\right)}{\sum_{STlp \in L} \mathrm{fr}\left(\mathrm{U_i}, \mathrm{STl_p}\right)} \quad (1)$$

Here SLF is stay location frequency, $\mathrm{U_i}$ be a user, $\mathrm{STl_j}$ be a stay location, $\mathrm{fr(U_i, STl_j)}$ means the no. of times the user has visited at the stay location and the denominator in the (1) is the sum of all total frequency of visits made by the user at various stay locations.

The second step is to find the day frequency. For that we use the following equation.

$$\mathrm{DF}\left(\mathrm{U_i}, \mathrm{STl_j}\right) = \frac{\mathrm{dr}\left(\mathrm{U_i}, \mathrm{STl_j}\right)}{\sum_{STlq \in D} \mathrm{dr}\left(\mathrm{U_i}, \mathrm{STl_q}\right)} \quad (2)$$

Here DF is the day frequency of the user $\mathrm{U_i}$ at the stay location $\mathrm{STl_j}$, $\mathrm{dr(U_i, STl_j)}$ is the no. of days he/she visited at the stay location and denominator in (2) means the total no. of days the user had went through the same route. The $\mathrm{dr(U_i, STl_j)}$ must be more than one day.

The third is the sensitive stay location finder function. For this we use the following equation.

$$\mathrm{SSLF}\left(\mathrm{U_i}, \mathrm{STl_j}\right) = \mathrm{SLF}\left(\mathrm{U_i}, \mathrm{STl_j}\right) \times \mathrm{DF}\left(\mathrm{U_i}, \mathrm{STl_j}\right) \quad (3)$$

This function gives as a numerical value and which predicts which stay location is most sensitive to the user. The higher SSLF value among the various SSLF values is the sensitive stay location of the user. We need to anonymize this sensitive stay location in order to achieve trajectory privacy. This will surely help as to prevent the unnecessary or overprotected anonymization.

4.5 Anonymization and Publication of Trajectories

For the anonymization purpose we use the generalization approach as used in [1] and we published the trajectory details in its anonymized form to the OTrD database. This model guarantees the unnecessary anonymization of more stay locations; it anonymizes only the stay location which is calculated as most sensitive. This will reduce the information loss considerably.

The proposed algorithm describes the trajectory anonymization approach based on the SSLF method.

Algorithm 1: Sensitive Stay Location Anonymization (SSLA)

Input : Historical trajectory dataset ITrD
Output : Publishable trajectory dataset OTrD (Anonymized trajectories)

```
/* Preprocessing / Sanitization */
for each Ui from 0 to n
    Read trajectories of users from ITrD to HTrj_tab
end for
/* Extracting stay point Sp */
Initialize threshold time $\delta t$
for i = 0 to m-1
    Read each trajectory coordinate from HTrj_tab
    Calculate $tm = | dept_i - arrt_i |$
    if $tm >= \delta t$ then put $Sp_i$ to Stp_tab
end for
/* Extracting stay location STl and finding sensitive stay location*/
Initialize spatial distance $\mu d$
Collect all ($STp_i \in$ Stp_tab) >= $\mu d$ and store into STl_tab
Calculate SLF, DF and SSLF
Set max(SSLF) as the sensitive stay location for each user in HTrj_tab
/* Creation of area zone */
Select neighbor $STl_k$ with minimum SSLF value and merge to form stay zone $SZ_j$
Replace $STl_j$ by the stay zone $SZ_m = \{ STl_j, STl_k \}$
Update SSLF
/ * Generalization of trajectories */
for each $Trj_i$ in HTrj_tab
if any $Sp_i \in SZ_m$ then
    Replace $Sp_i$ with brc and ulc coordinate values of $SZ_m$
otherwise preserve the same as in OTrD
Return OTrD
```

Here we anonymized the most sensitive stay location in a stay zone by selecting and combining other neighboring stay locations having the lesser SSLF value. Then update the SSLF value and stay zone coordinates. During the publication of trajectories, the anonymized stay point coordinates were updated with the coordinates of the bottom right corner and upper left corner of the stay zone and the non-anonymized stay points were published as such without any alteration.

4.6 Utility Analysis

Each time the anonymization aims to minimize the information and maximize the privacy. So each work has to measure the information loss occurred during the anonymization process. If the information loss is lesser but the privacy will be inversely affected. Everybody is aimed to achieve a better tradeoff between information loss and privacy gain. The following formulae (4) [17] is used to measure the average information loss occurred during the anonymization.

$$InL = \left(\sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{m} (1 - 1/area\_size\_of\_stay\ zone(SZ_i, t_j)) + \sum\nolimits_{x=1}^{k} L_x\right)/(N \times M) \tag{4}$$

Where *InL* represents average information loss occurred at the corresponding stay zone of SZ_i at time t_j when SZ_i stayed. L_x is the sensitive stay point that we are going to delete. N × M represents the no. of locations that anonymized in the published trajectory database. The value obtained from this analysis is certainly between 0 and 1.

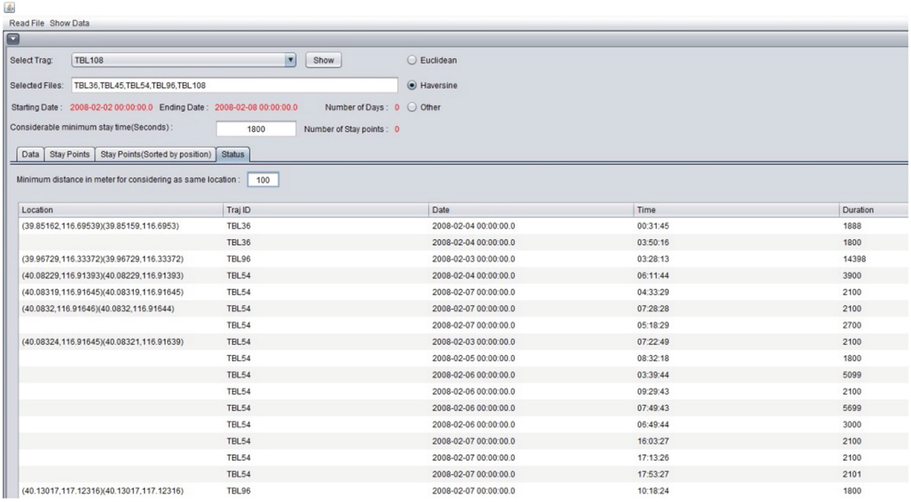
5 Experiments and Result Analysis

The proposed work's effectiveness is measured by the experimental evaluation and also through various analyses.

5.1 Experimental Setup and Outputs

For the experiment here we used T- Drive trajectory dataset, real-word dataset from Microsoft [18]. This dataset contains one week trajectories of 10,357 taxis run in China and it contain about 15 million trajectories and the taxis were covered a distance about 9 million kilometers. Each observation was taken with 10 min interval. The experiments were run on an Intel's 8^{th} generation core-i5 processor with up to 4 GHz speed and the machine is equipped with Windows 10 operating system and having 8 GB of RAM.

Here we set a time threshold as 1800 s and distance threshold as 100 m. We randomly took 50 users trajectory samples for our experiments. The Fig. 2 shows the selection and identification of stay locations for 5 users with their duration of stay at the location from 4044 location samples. Only one stay location contains more stay points of multiple days. So from the sample of five users only one user has a sensitive stay location (User no. 54), which satisfies our method SSLA.



Select Trag: TBL108

Selected Files: TBL36,TBL45,TBL54,TBL96,TBL108

○ Euclidean
● Haversine
○ Other

Starting Date : 2008-02-02 00:00:00.0 Ending Date : 2008-02-08 00:00:00.0 Number of Days : 0

Considerable minimum stay time(Seconds) : 1800 Number of Stay points : 0

Minimum distance in meter for considering as same location : 100

Location	Traj ID	Date	Time	Duration
(39.85162,116.69539)(39.85159,116.6953)	TBL36	2008-02-04 00:00:00.0	00:31:45	1888
	TBL36	2008-02-04 00:00:00.0	03:50:16	1800
(39.96729,116.33372)(39.96729,116.33372)	TBL96	2008-02-03 00:00:00.0	03:28:13	14398
(40.08229,116.91393)(40.08229,116.91393)	TBL54	2008-02-04 00:00:00.0	06:11:44	3900
(40.08319,116.91645)(40.08319,116.91645)	TBL54	2008-02-07 00:00:00.0	04:33:29	2100
(40.0832,116.91646)(40.0832,116.91644)	TBL54	2008-02-07 00:00:00.0	07:28:28	2100
	TBL54	2008-02-07 00:00:00.0	05:18:29	2700
(40.08324,116.91645)(40.08321,116.91639)	TBL54	2008-02-03 00:00:00.0	07:22:49	2100
	TBL54	2008-02-05 00:00:00.0	08:32:18	1800
	TBL54	2008-02-06 00:00:00.0	03:39:44	5099
	TBL54	2008-02-06 00:00:00.0	09:29:43	2100
	TBL54	2008-02-06 00:00:00.0	07:49:43	5699
	TBL54	2008-02-06 00:00:00.0	06:49:44	3000
	TBL54	2008-02-07 00:00:00.0	16:03:27	2100
	TBL54	2008-02-07 00:00:00.0	17:13:26	2100
	TBL54	2008-02-07 00:00:00.0	17:53:27	2101
(40.13017,117.12316)(40.13017,117.12316)	TBL96	2008-02-07 00:00:00.0	10:18:24	1800

Fig. 2. Output of the selection and identification of stay locations

The Table 1 shows the result of the extraction process of sensitive stay location for the user no. 54 (TBL 54).

Table 1. Extraction of sensitive stay location

User No. and location	No. of visits at the stay location	Total visits at stay locations	No. of days visited at stay location	SLF	DF	SSLF
54, STl_3	9	13	4	0.692	0.571	0.395

Here the other stay locations of user 54 don't satisfy the condition of total number of days visited at the stay location. So the user 54 has only one sensitive stay location and it is STl_3 and the highest is also the same.

5.2 Measure of Efficiency

Here we conducted two analyses and one algorithmic extraction. The Fig. 3 (a) shows the measure of information loss against the privacy parameter k, and this analysis shows that our approach SSLA is better than the approaches like Grid partition [1] and TRAMP [3] in terms of less information loss. The less information loss means greater privacy and so here the individual privacy protection comparatively well than the existing approaches. The Grid partition method anonymizes all higher stay points and TRAMP anonymizes the user's moving preference locations. None of them failed to protect the sensitive locations. But here we succeeded to protect this type of location with less information loss.

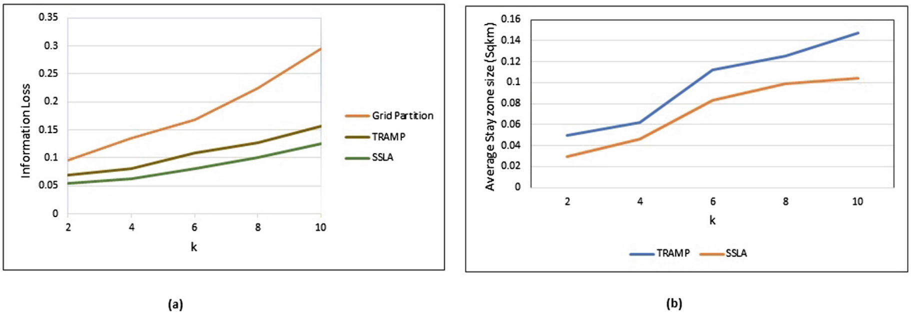


Fig. 3. Evaluation of SSLA. (a) information loss measure (b) average stay zone size

The analysis Fig. 3 (b) shows the average stay zone size during the anonymization. As we compared it with TRAMP method, our approach SSLA accommodates less stay zone size for the various values of privacy parameter k. So, unnecessary generalization of location points is less in our approach.

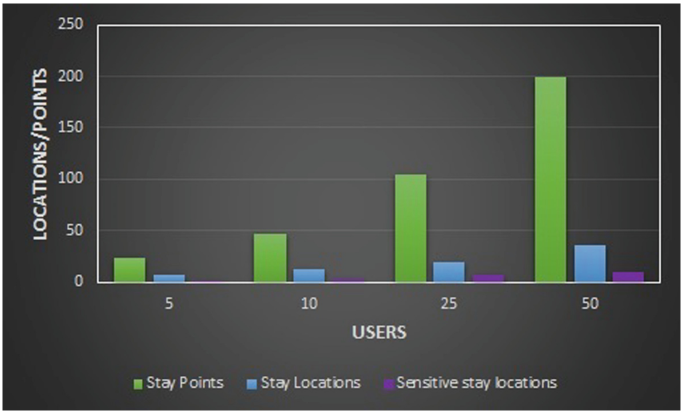


Fig. 4. Extraction of various parameters of SSLA

The Fig. 4 shows the extraction of various parameters of SSLA like stay point, Stay location and sensitive locations from a sample of randomly chosen 50 samples. The result shows that some of the users have sensitive stay locations and most of the taxi users have no sensitive stay locations. They were seasonal passengers and like travelling to various spots and don't have frequent visit spots or important locations.

6 Conclusion

The advancements in the field of GPS embedded systems has resulted a huge revolution in data collection and publication of these details are essential for the technological developments. The publication of mobility trajectories itself definitely raises the issue of individual privacy. The existing models are hard to find a better solution to reduce the information loss. In this work, we put forwarded a comparable solution for shielding individual's sensitive information through the anonymization of sensitive stay location by considering the temporal factors using SSLF and SSLA approaches. The proposed approach safeguards the most sensitive stay locations in stay zones and reduces the over protection problems. This model avoided the unnecessary anonymization of location traces, since only few individualized sensitive locations in a trajectory were anonymized and also provided a better solution against adversary attacks. We tested this approach with real-world dataset and result showed a positive effectiveness. However this research area is still challenging, because the reduction of information loss and maximizing the privacy is a major problem. In future, we are planning to apply more sophisticated approaches in the areas like frequent sub-trajectory mining, protection of customized sensitive locations in dynamic trajectories etc.

References

1. Huo, Z., Meng, X., Hu, H., Huang, Y.: *Y*ou *C*an *W*alk *A*lone: trajectory privacy-preserving through significant stays protection. In: Lee, S., Peng, Z., Zhou, X., Moon, Y.-S., Unland, R., Yoo, J. (eds.) DASFAA 2012. LNCS, vol. 7238, pp. 351–366. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-29038-1_26
2. Gambs, S., Killijian, M.O., Cotrez, M.N.P.: De-anonymization attack on geo-located data. J. Comput. Syst. Sci. **80**(8), 1597–1614 (2014)
3. Sui, P., Li, X., Bai, Y.: A study of enhancing privacy for intelligent transportation systems: k-correlation privacy model against moving preference attacks for location trajectory data. Special section on Internet-of-Things (IOT) Big Data Trust Management. IEEE Access **5**, 24555–24567 (2017). https://doi.org/10.1109/Access.2017.2767641
4. Chow, C.-Y., Mokbel, M.F.: Enabling private continuous queries for revealed user locations. In: Papadias, D., Zhang, D., Kollios, G. (eds.) SSTD 2007. LNCS, vol. 4605, pp. 258–275. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73540-3_15
5. Gruteser, M., Grunwald, D.: Anonymous usage of location-based services through spatial and temporal cloaking. In: Proceedings of Mobisys 2003, pp. 31–42 (2003)
6. Talukder, N., Ahamed, S.I.: Preventing multi-query attack in location-based services. In: Proceedings of 3rd ACM Conference on Wireless Network Security. ACM (2010)
7. Lee, B., Oh, J., Yu, H., Kim, J.: Protecting location privacy using location semantics. In: KDD 2011, San Diego, California, USA, pp. 1289–1297 (2011). https://doi.org/10.1145/2020408.2020602
8. Ghinita, G., Kalnis, P., Skiadopoulos, S.: Prive: anonymous location based queries in distributed mobile systems. Proc. WWW **2007**, 371–380 (2007). https://doi.org/10.1145/1242572.1242623
9. Kalnis, P., Ghinita, G., Mouratidis, K., Papadias, D.: Preventing location based identity inference in anonymous spatial queries. TKDE **19**(12), 1719–1733 (2007)
10. Krumm, J.: Inference attacks on location tracks. In: LaMarca, A., Langheinrich, M., Truong, K.N. (eds.) Pervasive 2007. LNCS, vol. 4480, pp. 127–143. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72037-9_8
11. Sweeny, L.: k-anonymity: a model for protecting privacy. Int. J. Uncertain. Fuzziness Knowl.-Based Syst. **10**(5), 557–570 (2002)

12. Abul, O., Bonchi, F., Nanni, M.: Never walk alone: uncertainty for anonymity in moving object databases. In: Proceedings of 24th International Conference on Data Engineering, pp. 376–385 (2008)
13. Domingo-Ferror, J., Trujillo-Rasua, R.: Micro-aggregation and permutation based anonymization of movement data. Inf. Sci. J. **208**(21), 55–80 (2012)
14. Mohammed, N., Fung, B.C.M., Debbabi, M.: walking in the crowd: anonymizing trajectory data for pattern analysis. In: Proceedings of 18th International Conference on Information and Knowledge Management, pp. 1441–1444 (2009)
15. Poulis, G., Skiadopoulos, S., Loukides, G., Gkoulalas-Divanis, A.: Apriori-based algorithms for k^m-anonymizing trajectory data. Trans. Data Priv. **7**(2), 165–194 (2014)
16. Hua, J., Shen, Z., Zhong, S.: We can track you if you take the metro: tracking metro riders using accelerometers on smartphones. IEEE Trans. Inf. Forensics Secur. **12**(2), 286–297 (2017)
17. Yarovoy, R., Bonchi, F., Lakshmanan, S., Wang, W. H.: Anonymizing moving objects: how to hide MOB in a crowd? In: 12th Int. Conf. Extending Database Tech. 72–83. ACM Press, New York (2009)
18. Yuan, J., Zheng, Y., Xie, X., Sun, G.: Driving with knowledge from the physical world. In: The 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2011. ACM, New York (2011)
19. Hussaeni, K.A., Fung, B.C.M., Iqbal, F., Dagher, G.G., Park, E.G.: SafePath: differentially-public publishing of passenger trajectories in transportation systems. Comput. Netw. **143**, 126–139 (2018)
20. Gramaglia, M., Fiore, M., Tarable, A., Banchs, A.: $k^{\tau, \varepsilon}$-anonymity: Towards Privacy-Preserving Publishing of Spatiotemporal Trajectory Data. CoRR, abs/1701.02243 (2017)

Medical Image Classification Based on Machine Learning Techniques

Naziya Pathan[1](✉) and Mukti E. Jadhav[2]

[1] Department of Management Science and Computer Studies,
Maulana Azad College of Arts, Science & Commerce, Aurangabad, India
nazi.pathan@yahoo.com
[2] Department of Computer Science and IT, MIT College, Aurangabad, India

Abstract. The usage of ultrasound has changed the district of therapeutic fetal examinations. The chance of distinguishing innate variations from the norm at an early dimension of the pregnancy is incredibly basic. To augmenting the odds of revising the ailment before it moves toward becoming dangerous. The issues identified with the standard strategy are its intricacy and the way that it requires many seeing around fetal life systems. Because of the deficiency of instruction among birthing assistants, explicitly in considerably less created nations, the outcomes of the examinations are as often as possible limited. Furthermore, the descent of the ultrasound gadget is frequently limited. These boundaries propose the requirement for an institutionalized method for the test to bring down the amount of time required, just as a programmed methodology for introducing the expectation of the embryo. Identification and grouping of medicinal pictures in the field of AI are a standout amongst the most testing issue. For restorative picture investigation, picture characterization is fundamental. In this paper, we have proposed a technique for removed ultrasound pictures. In view of the idea of standard view planes, a posting of predefined pictures is gotten of the embryo at some phase in the typical ultrasonography. In this sets perceptible pictures to contain with inherent irregularities. For the investigation of the medicinal pictures, the information of therapeutic pictures and the unusual ultrasound pictures are basic. A database is made to store the pictures of the embryo to perceive the separated picture is the ordinary or strange baby organ. This paper thinks about the order result by the utilization of Neural Network, K-Nearest Neighbor and Support Vector machine classifiers. It has been outlined from the result that the help Vector machine classifier outflanks a superior execution than Neural Network and K-Nearest Neighbor classifier.

Keywords: Ultrasound · Support Vector machine classifiers · K-Nearest Neighbor classifier

1 Introduction

- Advanced ultrasound (US) imaging is a broadly utilized logical anticipation strategy, owing to its safe noninvasive nature, low esteem, the usefulness of creating ongoing pictures, and continuing on with improvement in picture high caliber. It is anticipated that one out of every four logical analytic picture investigate inside the

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 91–101, 2019.
https://doi.org/10.1007/978-981-15-0108-1_10

worldwide includes US systems. Ultrasound imaging is utilized for picturing bulk and a few inward organs and therefore uncovering positive neurotic variations from the norm utilizing real-time picture. It is utilized for imagining the hatchling eventually of repeating and crisis pre-birth care. Obstetric sonography is typically utilized over the span of pregnancy. It has no respected long-term period feature results and only from time to time reasons any soreness to the patient. As it does not utilize ionizing beams, ultrasound involves no threats to the patient. It offers stay photographs from which the administrator can choose the extreme gainful portion, subsequently encouraging brief conclusions. The benefit of imaging is reduced by utilizing signal-subordinate commotion called spot clamor. This is multiplicative clamor that debases photograph extraordinary; therefore, it diminishes the capacity of an onlooker to segregate quality data inside the analytic test and renders treatment basic leadership hard. Edge security and clamor markdown are imperative for precise analysis Speckle decrease is the strategy of getting rid of dot commotion from the US. Dot commotion diminishes the skill of more pictures preparing including edge location. De-noising systems ought to diminish dot without obscuring or changing over the region of the edges, which are those focuses at which the glowing force modifications pointedly and commonly reflect fundamental changes inside the homes of the picture. Viewpoint identification is exceedingly basic in making sense of and understanding the total picture. Edge identification is, for the most part, the measurement and recognition of dim change, and in ultrasound depictions, its miles a troublesome test because of the reality the related calculations might be delicate to clamor. Dot commotion additionally debases the speed and exactness of US photo handling activities, comprising of division and enrollment Thus, the upgrade of photograph palatable is the subject of an imperative and exasperating examination, and this takes a gander at desire at smothering spot clamor through saving edges in ultrasound pictures. Images segmentation is a necessary part in the field of sample recognition and computer vision. The image segmentation algorithms are rising. The Snake model has been proposed by Kass et al. in 1987 [1]. Moreover, Xu and Prince have proposed the GVF-snake model in 1998 [2]. The major thought of dynamic structure present is fundamental. Right off the bat, a figured shape with point or curve is manufactured. At that point, pictures include are composed through the versatile twisting of the shape. A framework for restricting an essentialness work is coordinated. Finally, the division is finished. The dynamic structure model can be orchestrated into two headings as shown by the reliant qualities of pictures of the vitality work, for example edge based unique structures and district based powerful forms. It is difficult to achieve flawless division results with those techniques that simply rely upon the local or point of confinement information. The characteristic dot clamor and powerless edge of ultrasonic pictures make the division progressively troublesome. At present, it has transformed into an example to join region information and point of confinement information in a working shape show.

- Support vector machines (SVM) are directed learning models with related learning calculations that investigate information and perceive designs, utilized for order and reversion investigation. Given an arrangement of preparing precedents, each set apart as having a place with one of two classifications, a SVM preparing calculation manufactures a model that allocates new precedents into one class or the other, making it a non-probabilistic paired straight classifier. A SVM demonstrate is a depiction of the models as focuses in space, mapped with the goal that the precedents of the different classifications are isolated by a reasonable hole that is as wide as could be expected under the conditions. New models are mapped into that equivalent space and anticipated to have a place with a classification dependent on which side of the hole they fall on. Group learning has likewise been connected as an answer for preparing SVMs with imbalanced datasets. In these strategies; the common share class dataset is isolated into different associate datasets to such a level that every one of these sub-datasets has a comparable number of models as the minority class. This should be possible by arbitrary testing with or without substitution, or through grouping strategies. At that point an arrangement of SVM classifier development so everyone is prepared with a similar normal data and abnormal data. Finally, the choices made by the classifier gathering are consolidated by utilizing a strategy [3]. Neural network (NN) are related in numerous logical fields, for example, medication, computer science and engineering, neuroscience and engineering. It can be also connected in data preparing structure to take care of example examination issues or feature vectors classification for instance [4]. This paper describes pre-processing stages before the utilization of neural network that slot in (i) separating tools to reject noise, (ii) images segmentation to partition the images zone into isolated regions of value, and (iii) the utilization of mathematical strategy for extraction of feature vectors from each images segment. K nearest neighbor algorithm is used which is based on distance between items. This algorithm uttered in conditions of the similarity measure among the pair of data in N-dimension, the number of nearest data that are trained for classification, and vector of training data applied by the classifiers [5].

2 Related Work

Amid the most recent couple of decades, the utilization of ultrasonography for the identification of fetal variations from the norm has end up being tremendous in many industrialized nations. This brought about a move in time of the examination of acquired variations from the norm in infants from the neonatal length to the pre-birth time. This has foremost implications for both clinicians and the couples worried. In the case of ultrasound prognosis of fetal anomaly, there are numerous alternatives for the obstetric organization, range from widespread care to nonaggressive care and termination of pregnancy [6]. This paper investigates the system of together logical and parental decision making after ultrasound forecast of a fetal variation from the norm, with significance on the dutch situation. While ordinary discoveries at ultrasound examination have vigorous helpful mental outcomes at the pregnant female and her friend, the couple is frequently badly sorted out for horrendous news roughly the wellness of their unborn little child on account of unusual discoveries. This is, especially, genuine in settings where ultrasonography for the identification of fetal anomalies is offered as a basic piece of antenatal consideration without proper directing. A basic inquiry is to what degree the couple should be bolstered in decision making when a fetal variation from the norm is perceived. Upon the discovery of serious fetal variations from the norm, they will be looked with decisions on whether to hold the pregnancy. This exposition investigates the set of basic leadership after ultrasound conclusion of a fetal variation from the norm, with accentuation at the Dutch situation. On account of fetal irregularities, the conviction of decision among pregnant women shifts extraordinarily, troublesome the advising technique. These [7–12] researches use Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifiers. Support Vector Machine (SVM multiclass) shown a better classification performance as compared with other classification techniques in the paper the SVM multiclass will be used to classify many different classes at the same time while the old versions of (SVM) used for binary classification. This three classifier is the finest in training and testing time as its build N (N–1) classifiers one classifier to distinguish each pair of classes.

3 Materials and Methods

3.1 Image Collection

The images were acquired from Dr. Mishrikotkar Prasanna, Shanti Imagining Center at Aurangabad and Dr. Sachin Rathod, Rathod Diagnosis Center Buldhana, in India and appear Table 1.

Table 1. Types of Images

Sr.no	Typeof images	Normal	Abnormal
1	Heart	86	45
2	Kidney	55	63
3	Spine	100	51
4	Limbs	99	54

3.2 Method

Fetus study through ultrasound images are less contrast due to that proper guideline is not given to the parents to know about the abnormalities present in their babies due to that they can lost their child. The recognition and characterization of the fetus from images are a difficult undertaking in light of the fact that an image more often than not contains numerous groups and covering objects. The proposed framework helps through image processing techniques the Physician to characterize both ordinary and anomalous images precisely. The square chart for the proposed framework is appeared in Fig 1. The different stages associated with the proposed strategy incorporate image acquisition Set of test images are gathered from a ultrasound imaging framework. The second one is an images pre-preparing area, apply pre- processing techniques for the removal of inherent noises. Enhancement of images, segmentation of background cells, features extraction, and finally the classification.

3.2.1 Image Acquisition

Ultrasonography has developed as a helpful system for imaging interior organs and delicate tissue structures in the human body. It is noninvasive, compact, adaptable, and general ease. Ultrasonography can identify numerous fetal auxiliary and utilitarian variations from the norm. Ultrasound works by utilizing sound to create a picture of the baby. The principal arrange is the picture securing stage. High-recurrence imaging utilizes high recurrence sound waves and their echoes to create pictures that can show organ development progressively.

3.2.2 Image Pre-processing

Pre-Processing is to improve the interpretability of the data present in image for human. Image Enhancement is utilized for better-quality picture and it very well may be finished by either smothering the clamor or expanding the picture differentiate.

3.2.3 EdgeMap

In a picture, between two areas, many associated pixels are seen. Such a pixel bunch is called an edge. The identification of the edge is the most widely recognized strategy for distinguishing significant discontinuities, particularly in the therapeutic imaging field. It is used for starting the division. It is utilized to discover the edges of the images. It is used for initial segmentation. It is used to find out the edges of the images. Most of the edge detectors work on measuring the intensity gradient at a point in the image [13].

3.2.4 Image Segmentation

For segmentation, two unique methods are utilized that is, wind miss happenings utilizing GVF and Gabor. These strategies have effectively sectioned the kidney, heart, appendages, and Spine. The segmented fetus images are utilized to GVF; it is processed as dispersal of the angle vectors of an edge map subordinate from typical and strange pictures.

3.2.5 Image Classification

In proposed system, three classifiers are used. The SVM classifiers are used in Ultra Sound images of anomalies of fetal images (abnormal) dataset. The neural network feed forward back Propagation is used to identify normal and abnormal fetus image. NN is made out of three layers, information, yield and shrouded layer. Each layer can have a number of axes and axis from information layer is related with the hidden layer. The important Back Propagation (BP) Algorithm is proposed for feed forward NN calculations. The Back-propagation (BP) preparing calculation is an important agent of all iterative leaning throws calculations utilized for directed learning in neural systems. K nearest neighbor algorithm is used which is based on distance between items. K-NN was the parameter free method, which uses a Euclidean distance measure.

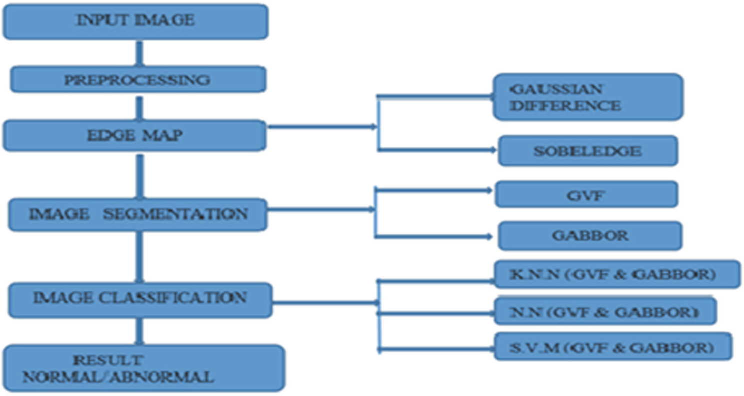


Fig. 1. Block diagram of proposed system

3.2.6 Training and Testing

The classification of fetus images of various part i.e. Kidney, heart, limbs and Spine images are used to training data set and testing data set. The collected data were grouped into two classes, namely normal fetus and abnormal fetus. The feature extraction is applied on the training image dataset and features are stored as knowledge base. The features extracted in the feature extraction phase are stored for classification analysis. An equal number of normal fetus and abnormal fetus samples were randomly picked from the total data set. In training, the normal fetus image is classified as normal as true positive (TP) and normal fetus image is classified as abnormal as false negative (FN).

4 Result Analysis and Discussion

In Analysis, we are comparing the performance level of three distinct classifier i.e KNN, SVM and NN.

4.1 KNN

KNN algorithm depends on the separation between things. It is the easiest strategy, which gives reasonable arrangement exactness. The K-NN has higher exactness and security for pictures data. The K-NN guideline comprises of a preparation part and a testing part. The exhibition of K-NN classifier with Gabor and GVF is appeared Table 2. K-NN algorithm involves following stages:

Algorithm no 1: K-Nearest Neighbor classifier Algorithm

```
Input: Set of US images with Segmentation.
Output: Identify the fetus images as normal or abnormal.
Step 1: Begin
Step2: Load data set of fetus segmented US images.
Step3: Initialize the value of K.
```

```
Step4: To find out the classified class, iteration can be done from 1 to total no of training dataset.
Step5: Calculate the Euclidean distance between the points.
Step6: On the bases of distance value sort the calculated distance in non-decreasing order.
Step7: Through the sorted array derived top K row.
Step8: Select the highest class through these rows.
Step9: Predict the accuracy of the classifier, which can identify the fetus images, are normal or abnormal.
Step10: End
```

Table 2. KNN classifier results with Gabor and GVF

	Gabor KNN	GVF KNN
RCALL	0.8987	0.9431
Precision	0.9034	0.9402
Specificity	0.9861	0.9917
F-score	0.8998	0.9411

4.2 SVM

The essential idea of SVM classifier is to build up a hyper plane classifier that confines the typical and irregular points of reference while boosting the smallest edge. The presentation of SVM classifier with Gabor and GVF appears Table 3. The SVM calculation includes the following stages

Algorithm no 2: The SVM classifier Algorithm

```
Input: Set of US images with Segmentation.
Output: Identify the fetus images as normal or abnormal.
Step 1: Start
Step 2: Fetus features of US image
Step 3: Determine optimal separating margin in hyper planes using
Step 4: Compute linear discriminate function with optimal separating margin using
Step 5: Classify US images as normal fetus or abnormal fetus
Step 6: End.
```

Table 3. SVM classifier results with Gabor and GVF

	Gabor SVM	GVFSVM
RCALL	0.9127	0.9553
Precision	0.9156	0.9605
Specificity	0.9883	0.9942
F-score	0.9135	0.9576

4.3 Neural Network

The information is fed in the form of GVF-Snake. The calculation of ultrasound images is done by feed forward and back propagation. The structure of neural system depends on the error which is registered from distinction of the objective output and real output. The error esteems are utilized to refresh the weight lattice of the neurons of the system. The informational indexes used to prepare the neural system can be parceled in to the autonomous weights. The performance of NN classifier with Gabor and GVF is shown in Table 4.

Table 4. NN classifier results with Gabor and GVF

	Gabor NN	GVF NN
RCALL	0.8681	0.9283
Precision	0.8762	0.9351
Specificity	0.9823	0.9903
F-score	0.8712	0.9302

Algorithm no 3: The NN classifier Algorithm

```
Step 1: numeral of input and output unknown neuron
Step 2: initial load and bias to chance value
Step 3: Do again until termination criteria satisfied present training and propagate it through network
Step 4: input useful, multiplied by weight
Step5: output passed to each neuron in next layer and working backwards.
```

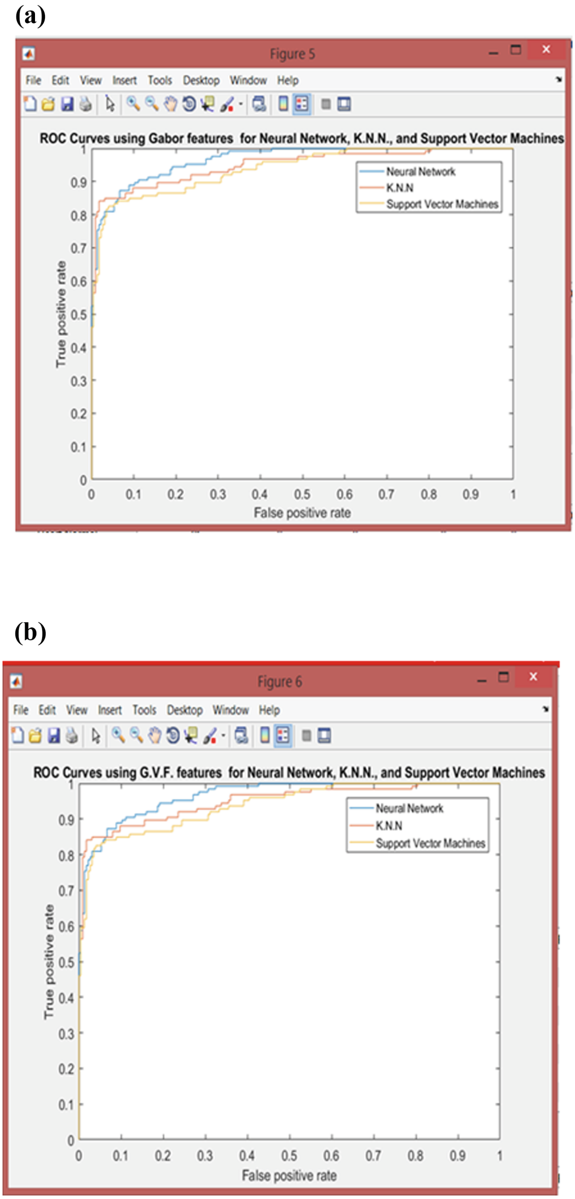
5 Comparison Between KNN, SVM and NN

KNN classifies data based on the gap metric while SVM need a right segment of training. Due to the most useful nature of SVM, its miles assured that the separated information would be optimally separated. KNN is used as multi-elegance classifiers whereas popular SVM separate binary statistics belonging to both of one elegance. For a multiclass SVM, One-vs-One and One-vs-All method is used. In One-vs-one

Table 5. Comparison between NN, KNN and SVM classifier results with Gabor and GVF

	GABOR	GVF	GABOR	GVF	Gabor	GVF
	NN	NN	KNN	KNN	SVM	SVM
RCALL	0.8681	0.9283	0.8987	0.9431	0.9127	0.9553
Precision	0.8762	0.9351	0.9034	0.9402	0.9156	0.9605
Specificity	0.9823	0.9903	0.9861	0.9917	0.9883	0.9942
F-score	0.8712	0.9302	0.8998	0.9411	0.9135	0.9576

technique, we need to educate n * (n − 1)/2 SVMs: for each pair of instructions, one SVM. To feed the sample, which is unknown to the entity and the very last verdict on the kind of records, is decided with the aid of the general public result among all consequences of all SVMs. This technique is used frequently utilized in multiclass



Graph 1. (a, b) ROC curves using Gabor features for Neural Network, KNN and Support Vector machine.

classification. When it comes to One-vs-All approach, to train as many SVMs as there are training of unlabeled data. As in the different technique, the unknown pattern to the system and the result is given to the SVM with the largest decision price. In KNN, the space metric is calculated on every occasion a across fixed recent unlabeled information. SVMs have primary instances in which classes is probably linearly separable or non-linearly separable. When the instructions are non-linearly separable, we use kernel characteristic such as Gaussian basis characteristic or polynomials. Hence, we best have to set the K parameter and select the space metric suitable for classification in KNN while in SVMs we have to select the R parameter and Additionally the parameters for kernel if the instructions are not linearly separable. The Comparison between NN, KNN and SVM classifier results with Gabor and GVF is shown in Table 5 and Graph 1(a, b).

6 Conclusion

The outcomes and perceptions show that SVMs are increasingly trustworthy more noteworthy classifiers. In any case, KNN is substantially less computationally serious than SVM. Since KNN is easy to put in power, the class of Multi-style records should be accomplished with KNN. The calculation that guarantees trustworthy identification in capricious conditions depends upon the realities. In the event that the data focuses are heterogeneously dispensed, both need to work appropriately. On the off chance that the data is homogenous to watch, one is most likely fit for grouping higher by means of introducing a portion into the SVM. For most extreme practical issues, KNN is a terrible want as it scales severely - if there are 535 marked models, it may set aside a long effort to find K closest neighbors.

Acknowledgments. We sincerely acknowledge the support provided by Dr. Mishrikotkar Prasanna, Shanti Imagining Center at Aurangabad and Dr. Sachin Rathod, Rathod Diagnosis Center Buldhana, in India for carrying out this work.

References

1. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. Int. J. Comput. Vis. **4**, 321–331 (1987)
2. Xu, C., Prince, P.L.: Snakes, shapes, and gradient vector flow. IEEETrans. Image Process. **7**, 359–369 (1998)
3. Qianqian, L., Qingyi, L., Lei, L., Yingxia, F., Peirui, B.: An ımproved method based on CV and snake model for ultrasound image segmentation. In: 2013 Seventh International Conference on Image and Graphics (2013)
4. Sheeja, S., Lavanya, V.S.: Empirical model decomposition based SVM classifier for abnormality detection in fetus us images. Int. Sci. Press **10**(12) (2017). ISSN 0974-5572
5. Mangayarkarasi, T., Jamal, D.N.: PNN-based analysis system to classify renal pathologies in kidney ultrasound ımages. IEEE (2017)

6. Bijma, H.H., van der Heide, A., Wildschut, H.I.J.: Decision-making after ultrasound diagnosis of fetal abnormality. Eur. Clin. Obstet. Gynaecol. **3**, 89–95 (2007). https://doi.org/10.1007/s11296-007-0070-0
7. Manning, C.D., Raghavan, P., Schütze, H.: Support vector machines and machine learning on documents, multiclass SVMs. In: Introduction to Information Retrieval. Cambridge University Press (2008). http://nlp.stanford.edu/IR-book/html/htmledition/irbook.html
8. Avni, U., et al.: X-ray categorization and retrieval on the organ and pathology level, using patch-based visual words. IEEE Trans. Med. Imaging **30**(3), 733–746 (2011)
9. Joachims, T.: Multi-Class Support Vector Machine. Cornell University, Department of Computer Science, Version: 2.20, 14 August 2008
10. Rifkin, R.: Multiclass Classification. 9.520 Class 06, 25 February 2008
11. Pal, M.: Multiclass Approaches for Support Vector Machine Based Land Cover, Classification
12. Khademi, et al.: A review of methods for the automatic annotation and retrieval of medical images. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **4**(7), 898–902 (2014)
13. Jayaraman, S., Esakkirajan, S., Veerakumar, T.: Digital Image Processing. Tata McGraw Hill Education Private Limited,New Delhi (2010)

Optimal Sizing and Placement of Micro-generators Using Particle Swarm Optimization Algorithm

R. Kiran(✉), B. R. Lakshmikantha, and Shanmukha Sundar

Electrical and Electronics Engineering,
Dayananda Sagar Academy of Technology and Management, Bangalore, India
Kiran.Bengaluru@gmail.com

Abstract. A reliable electrical supply plays a vital role in present era. The global electrification drives the industrial sector towards advancements & automation. To achieve this, Power system should be operated with safe margins & in stable state. It mainly depends on condition of several elements, environmental factors, operating conditions, safe limits etc. Placing the micro-generators in a power system & sizing helps in operating the power system in stable & reliable condition. The purpose of this research paper is to incorporate a micro-generator in the power system to have a reliable, stable electrical supply & to reduce the losses in the system. Optimal Placement & sizing of Micro-generators was implemented on IEEE-14 Bus system, producing efficient results.

Keywords: Optimal location · Micro-generator · Stability · Monitoring

1 Introduction

The regularly developing populace and the utilization of electrical power have set off the interest for reliable electrical supply. The power system network plays a vital job in transmitting power from the generators to the buyers. To forestall events, for example, loss of power, power network providers must perceive the quality and stability of different parts of the power transmission system through observing and measuring gear.

The advent of global electrification drives the world advancement in technology in every industrial sector. With thousands of generating units all interconnected by a complex network consisting of ac, transformers, high voltage direct current (dc) system and other complex equipment, the daily operation delivering safe, reliable and high quality electric power energy to the user is a very complicated task. It requires continuous planning work on top of the constant monitor and action of the operator [1].

Increasing numbers of recent wide-area system blackouts have been a constant reminder of the fact that the prevalent modus operandi on delivering power to customers is not sufficient to maintain power system security during operation.

Modern electrical system is a complex system, operates under complex interconnections, extreme conditions which is very close to the limits. Further, with expanded computerization and utilization of electronic gear, the quality has increased most

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 102–112, 2019.
https://doi.org/10.1007/978-981-15-0108-1_11

significantly, moving spotlight on to the ideas of voltage stability, oscillations, frequency stability and so forth. The power system is very non-linear and dynamic framework, with working parameters ceaselessly changing. Stability is henceforth an element of the underlying working condition and the idea of unsettling disturbance [1].

The worldwide call for energy especially in developing global areas has been seeing at colossal development because of fast populace increase, financial increase, enhancing social presence, creating business divisions and urbanization. The development of generation ability is the answer to fulfill the consistently expanding.

Need for energy. Enlarging the existing generation plants or incorporating in other locations helps in increasing the power generation [2]. Expansion of new generators to a present scenario has impacts on the network and the performance of the power stations. Therefore, the new generator location is an absolutely important part for planning of the power system.

Other than the technical and money related factors of finding the optimal location; another important inconvenience is the environmental effect. Environmental effect is produced from non-inexhaustible resources inclusive of coal, gas and oil. Those non-renewable energy sources produce contaminating vaporous like sulphur dioxide (SO_2), oxides of nitrogen (NO_X) and carbondioxide (CO_2). These emission gases can't be tolerated. Growing global conscious of environmental safety has pressurized the utilities enforcing the usage of renewable energy sources [3]. For the above reasons, an optimal location of new generator has to be decided based on technical, money & environmental factors. PSO Algorithm is utilized for solving this placement problem.

This paper discusses an optimization method limiting complete propellant price with negligible effects on nature susceptable to physical and mechanical limitations with the aid of the utilization of PSO technique has been proposed to choose the perfect region for new generation plant. The proposed strategy/technique suggested in this paper can be helpful/useful for power system utilities with a preparation for enhancement of generator capacity.

2 Voltage Stability

Voltage stability is engaged with "the capability of the energy system to sustain tolerable voltage at all bus in the system under usual circumstances and after being exposed to disruption". Voltage adherence depends upon maintaining equilibrium between supply & demand.

Progressive & uncontrollable decrease in voltage, could lead to cascaded outages & the framework goes into a phase of voltage vulnerability when a convulsion occurs. This cascaded outages can lead to loss of synchronism of some generators which leads to system blackout [4].

Voltage instability occurs when you are trying to consume more power beyond the capacity of the network. The reactive power and voltage stability are dependent to each other. Increase of reactive power leads to voltage raise while shortage of reactive power reduces the voltage in the system.

Voltage stability is sorted into following 2 sub-classes:

(1) Small Disturbance Voltage Stability:
It alludes to the system's capacity to keep up the enduring voltages when exposed to small changes in load.
(2) Large Disturbance Voltage Stability:
It alludes to the system's capacity to keep up the enduring voltages when exposed to extensive unsettling disturbances. It requires calculation of non-linear reaction of the power system to incorporate collaboration between different gadgets like motors, transformer tap changes and field current limiters.

Since power system devices are designed to be operated within specified limits, most pieces of system are protected via automatic tool which can cause device to be switched out of the system if those limits are violated. If the device violates those limits (line limits, voltage limits and many others.) continuously, and if there is no alternate action taken, then the whole device or massively a part of it may collapse & it is called as SYSTEM BLACKOUT [4].

Voltage Collapse
Voltage collapse is a mechanism where a collection of activities following voltage vulnerability causing objectionable voltage profile in a large scale. It could be demonstrated in numerous kinds of approaches. "Voltage collapses are identified as under:

i. The commencing occasion can be due to the following motives: Progressive system adjustments like escalation of load or huge immediate disturbances which include lack of generation or a loss of massively loaded line.
ii. The bottom line of the problem is the incapability of the system to satisfy its susceptible power requirements. Whilst shipping of susceptible power from nearby locations is tough, any exchange which calls for additive reactive power strength help might also ultimately cause voltage exhaustion.
iii. The voltage collapse, apart typically, reveals itself as a gradual decrease in voltage. It's a outcome of more than one procedure together with the actions and synergy of devices, control system as well as protection structures.
iv. Reactive power compensation can be formed handiest through an apt preference of a combination of shunt capacitors, static var compensator and probably synchronous condensers" [5].

Strategies of Enhancing Voltage Stability
"Voltage adherence of the system could be enhanced through these strategies:

(1) Reinforcing the localized reactive power strength (SVC, STATCOM) is efficient and comparatively cheap.
(2) Reduction in Net reactance and increase in power flow is obtained by balancing the line length.
(3) Extra transmission lines can be erected. It promotes reliability.
(4) Intensify incitement of generator, system voltage enhances and Q is endowed to the system.
(5) There is a increase in voltage as the reactive power burden reduces due to restoration of diplomatic load shedding [5].

3 Problem Formulation

In locating the optimal generation place, a non uniform multiobjective optimization situation which involves propellant price and discharge function of all alternators in the framework is being designed.

A system having N number of generating units is considered. The Input-output curve of the units are represented as $C_1(P_1)$, $C_2(P_2)$….$C_n(P_n)$. The objective is to find the optimal location of microgenerator to reduce the losses, minimizing the cost function

$$\text{Minimize Fuel cost(F(Pg)), Emission Cost(E(Pg))} \tag{1}$$

$$\text{Subject to: } g(Pg) = 0 \tag{2}$$

$$h(Pg) \ < 0 \tag{3}$$

here, g defines the power balance, equality restraint and *h* defines the power system security, inequality restraint.

The net propellant price of the alternators can be expressed through a quadratic characteristic equation as follows:

$$F(Pg) = \sum\nolimits_{i=1}^{Ng} a_i P_{gi}^2 + b_i P_{gi} + c_i \tag{4}$$

$$E(Pg) = \sum\nolimits_{i=1}^{Ng} d_i P_{gi}^2 + e_i P_{gi} + f_i \tag{5}$$

where,

a_i, b_i, c_i cost coefficients of generating unit
d_i, e_i, f_i emission coefficients of generating unit i
P_{gi} real power generation at bus i
N_g number of generators in the system

The multiobjective optimization situation is reformed to one optimization situation by means of proposing a rate retribution component *h* ($/kg) as given below:

$$\text{Minimize } f = F(Pg) \ + \ h\, E(Pg) \tag{6}$$

Where *h* is the ratio between the fuel price and

This is the proportion between the propellent price and the discharge of each alternator at its peak capacity. *h* can be determined as:

$$h_i = \frac{F_i(Pgimax)}{E_{i(Pgimax)}} \qquad i = 1\ldots\ldots\ldots..,Ng \tag{7}$$

The insertion of the price penalty element h in the function has defined that, the total operating price of the system is the price of fuel plus the implied value of emission. The most price penalty thing has been selected for combining value of gas.

Plus the implied price of emission as it gives a completely appropriate solution for emission limited fewer price situations [6].

This objective function is subjected to the following constraints.

A. Equality constraints
Power balance equations: it's far important to Make sure the output of generator serve the entire load and total losses in transmission lines.

The inclusion of the price retribution element h in the equation has shown that, the net working price of the unit equals the price of propellent added to the latent value of discharge.

The most cost retribution thing is tabbed for compounding value of gas added to the intended price of discharge since it gives a completely appropriate result for discharge limited price situations [6].

This equation is exposed to these restraints.

A. Equality restraint
Power balance eqs: it's far more important to make sure the result of generator deliver the entire heap and complete loss in transmission lines.

$$\sum_{i=1}^{N_g} P_{Gi} - P_D - P_{Loss} = 0 \tag{8}$$

The computation of the system failure may be finished by distinct techniques such As B-coefficients technique or general loss system. Here, power flow technique is used for calculating the loss compatible with real and reactive power.

$$P_{Gi} - P_{Di} - V_i \sum_{j=1}^{N_b} V_i \left[G_{ij} \cos\left(\delta_i - \delta_j\right) + B_{ij} \sin\left(\delta_i - \delta_j\right) \right] = 0 \tag{9}$$

$$Q_{Gi} - Q_{Di} - V_i \sum_{j=1}^{N_b} V_i \left[G_{ij} \cos\left(\delta_i - \delta_j\right) - B_{ij} \sin\left(\delta_i - \delta_j\right) \right] = 0 \tag{10}$$

B. Inequality constraints

(1) *Power generation limit:*

$$P_{Gi}^{min} \leq P_{Gi} \leq P_{Gi}^{max}, \quad i = 1, \ldots, N_g \tag{11}$$

(2) Voltage limit:

$$V_i^{min} \leq V_i \leq V_i^{max}, \quad i = 1, \ldots, N_b \tag{12}$$

(3) Line flow limit:

$$|P_l| \leq P_l^{max}, \quad l = 1, \ldots, N_l \tag{13}$$

4 PSO Algorithm

PSO is a biologically inspired algorithm which works on individual improvement and population criterion. This algorithm relies on simulation of social behaviors like bird flocking, fish schooling and Swarm theory [7]. This algorithm is motivated from social behavior. It is a mathematically driven model which is simple and cost effective in terms of space, time taken. The particles possess quick occurrence to native and/or global best location(s) considering a few ranges of generations. Swarm intelligence in PSO contains different type of particles. Each represent a different solution for optimization job. Each particle corresponds to a candidate solution. Every particle migrates to a new location with current speed based on previous best and global best solutions. The most effectual solution is used; every particle uses the path of local value however additionally uses the global premier location [8]. The swarm size is denoted as follows. In generic, 3 attributes are present, the particles' current position, current speed, and past best position, for particles within the search space. After every iteration the swarm particles are updated with the values obtained. In Shi [9] a new parameter i.e. inertia weight is accommodated in traditional PSO is used. The algorithm is depicted as given below.

Figure 1 depicts the working of a PSO alg. At some stage in the PSO method, every capability answer is depicted as a particle with a role vector **x**, known as phase weight factor **b** and a motion velocity delineated as **v**. The position and velocity of the ith particle For a **K** dimensional optimization may be represented as $\mathbf{b}_i = (b_{i,1}, b_{i,2}, \ldots, b_{i,K})$ and $\mathbf{v}_i = (v_{i,1}, v_{i,2}, \ldots, v_{i,K})$, respectively. Every particle has its own best position related to the objective gain of each particle derived so far for a period t, indicated as *pbest*. The global best (*gbest*) is shown by $\mathbf{b}^G = (b_{g,1}, b_{g,2}, \ldots, b_{g,K})$, symbolizes the prime value obtained so far at time t in the total swarm. The current momentum $\mathbf{v}_i(t+1)$ for particle i is modified as

$$v_i(t+1) = wv_i(t) + c_1 r_1 \left(b_i^P(t) - b_i(t)\right) + c_2 r_2 \left(b^G(t) - b_i(t)\right), \tag{14}$$

here w is the *inertia weight*, $v_i(t)$ is the previous momentum of the particle *i at time t*.

Allegedly in this, the current momentum is equated to the old momentum estimated by using w and additionally related to the location of the particle and the global best by means of acceleration constants c_1 and c_2. The acceleration constants in Eq. (14) modify the quantity of tension in PSO. Minimum value permit particles to move from required locations earlier than being returned, whilst excessive values bring about unexpected motion closer to, or beyond, target regions. The acceleration constants c_1 and c_2 are hence called as the cognitive and social rates. Due to the fact they constitute the weighting of the acceleration phrases that haul the character particle in the direction

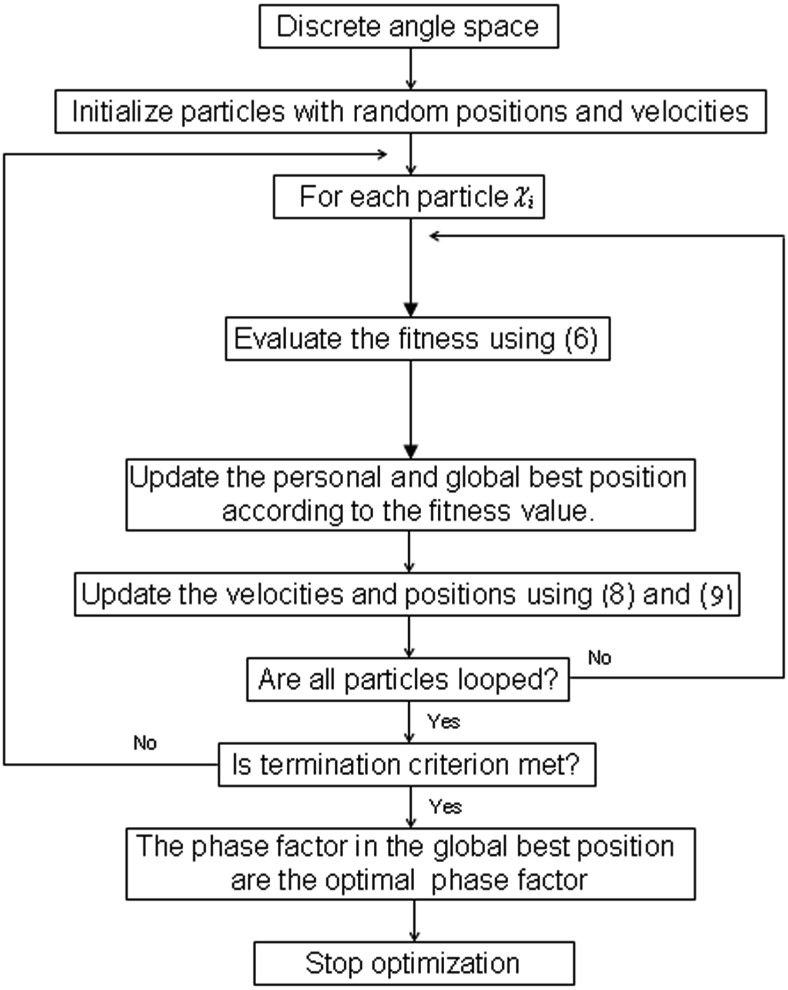

Discrete angle space
Initialize particles with random positions and velocities
For each particle x_i
Evaluate the fitness using (6)
Update the personal and global best position according to the fitness value.
Update the velocities and positions using (8) and (9)
Are all particles looped?
No
Yes
Is termination criterion met?
No
Yes
The phase factor in the global best position are the optimal phase factor
Stop optimization

Fig. 1. Shows the flowchart of the PSO algorithm

of the local best and global best locations. The momentum of the particles are given by $[v_{\min}, v_{\max}]$. If a momentum value exceeds threshold $v_{\min}$ and $v_{\max}$, it is replaced by the analogous threshold.

The inertia weight w in (14) is engaged to influence the effect of the past values of momentums at the present momentum. A massive inertia weight helps seeking new place while a minimal inertia weight ease great-looking inside the cutting-edge seek place. Appropriate choice of the inertia weight furnish a equity between worldwide expedition as well as neighborhood exploitation, and outcomes to less iterations on common to discover a adequately precise result. Observational results advise to set the inertia weight to a large value, giving preference to global expedition of the lookup area, presice lowering $w(t)$, in order to reap suitable solution [7–12].

In the direction of aiming to provoke the mild uncertain element of natural swarm behavior, two random functions r_1 and r_2 are carried out independently to offer systematic allotted numbers in the range [0, 1] to hypothetically alter the corresponding pull of the personal and global best particles. Based on the updated velocities, new position for particle i is calculated in accordance with the equation given below.

$$b_i(t+1) = b_i(t) + v_i(t+1) \tag{15}$$

The populations of particle are then changed in keeping with the new velocities and locations measured by using (14) and (15), and have a tendency to clump together from specific instructions. For this reason the assessment of each related health of the new populace of particles starts off evolved again. The set of rules runs thru those techniques iteratively until it stops. These days, Clerc [13] made use of one more parameter referred as constriction factor ψ, which additionally assist make sure concurrence. The narrowing model depicts the way of selecting w, c_1 and c_2 which helps in guaranteed concurrence. By selecting the numbers properly, the momentum of all the particles are hindered in the scope of $[v_{\min}, v_{\max}]$.

5 Proposed Algorithm

The planned PSO algorithm is defined as underneath:

1. Create a distributed framework with N range of DGs and related parameters. Make use of the Distributed Power System as the particles with random position & velocities.
2. Develop the cost matrix
3. Specify the fitness rule to decrease the collective cost over the disbursed power system
4. Fitness is evaluated
5. The local & global best values is updated according to fitness rule.
6. Update the positions & velocities
7. Stop when the criterion met.
8. This will be the global best factor.

6 Case Study and Simulation Results

This paper proposes a method examined on IEEE 14 bus system. Programming using MATLAB is developed to check the projected method.

In short, the system includes 14 Buses, five generators and eleven loads. In this example, every load is consistently improved by 10% with the additional load of 336.7 MW. It is presumed that the 10% accelerated load was not met by the extremum capableness of the present generators. In these state of affairs, increase or additional installation is essential. Consequently, a new Generator with 100 MW ability may be introduced to the framework to fulfill the predicted load development. The acquired fuel value, overall discharge and machine failure with the new generator at each bus inside the system are depicted In Table 1. Figure 2 suggests the assessment of outcomes acquired for every bus location graphically. For every gadget, a top-quality vicinity can be observed immediately.

Table 1. Outcome of new generator placed at each bus

Bus Number	Fuel cost	Emission	System loss
1	15393	314	7
2	15353	313	5
3	15294	313	3
4	15297	311	4
5	15322	313	4
6	15347	313	5
7	15298	313	3
8	15300	313	4
9	15300	313	3
10	15312	313	4
11	15343	313	5
12	15377	314	6
13	15330	313	5
14	15311	313	4

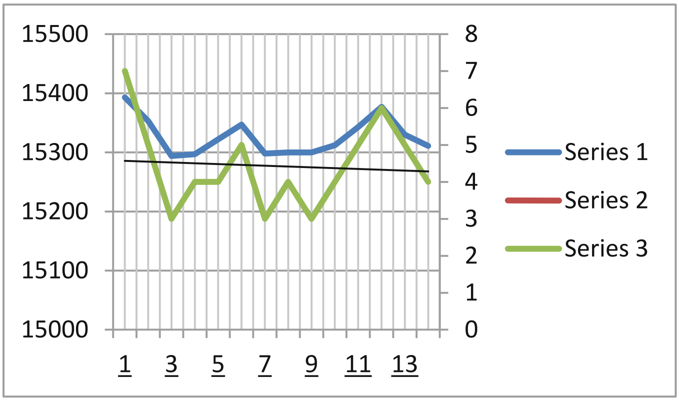


Fig. 2. Shows the variation of results for IEEE 14-bus system.

Based totally on the outcome in Table 1, it's far discovered that exceptional fuel price, overall emission and machine loss sample are acquired whilst finding the new Generator at special buses. It is able to determine that finding it at bus 3 has the minimum fuel price, general discharge and system failure with respect to all different bus locations. Therefore, it's selected because the maximum appropriate position for mounting the new generation plant. Also, this region is Exceptional fit as there'll not be any trouble of Transmitting congestion, without even including Extra lines. Alternatively, it is also observed that bus 1 and 12 are especially terrible positions for the new era unit. Finding the new generator at those buses return maximum fuel value, overall discharge and system failure.

7 Conclusion

The optimal plant location has been determined with the use of a PSO technique, for a proposed method. The method has been developed with fuel value as well as emission goals and is dependent to a number of restriction. The most appropriate generation Plant position is chosen primarily appertaining to minimal fuel Value, overall discharge and overall line losses of the plant. This approach is executed on IEEE 14-bus method taking into account further additional load. The structure of fuel cost, gross emission and system losses were acquired with newer generators placed at every bus. The modeling consequences ensures that the optimum plant positioning can salvage a fuel price in addition to cut down the overall discharge of generators and failure in the network. Except, the planned approach guarantee the increase of new generator can be completed without any trouble of transmitting line over-crowding. These rescue the finance and operation cost of network elaboration. Consequently, for improved extended time period operational welfare of the system, it's far important to determine the new plant at a Premiere position.

The optimal plant area has been resolved with the utilization of a Swarm based methodology, for the proposed framework. Problems figured are with fuel value and emission goals and is exposed to various restrictions. The most suitable Plant position is picked essentially dependent on negligible fuel Value, by and large emissions and generally line losses in the system. This methodology is executed on IEEE 14-bus system thinking about forthcoming extra load. The examples of fuel cost, in general discharge and losses were acquired from recent generator situated at all bus. The recreation outcome demonstrate that an ideal plant area can spare a fuel cost an incentive not withstanding decrease the emission of generators and leaks in the system. But, this methodology guarantees that expanding further generator is done with no inconvenience of transmission line blockage. This will spare the venture and task cost of system extension. Thus, for more effective extended haul operational advantages of the system, it's far necessary to find the bran-new plant at a Premiere area.

References

1. Phadke, A.G., Thorp, J.S.: Synchronized Phasor Measurements and Their Applications. Springer, New York (2008). https://doi.org/10.1007/978-0-387-76537-2
2. Das, P.K., Chanda, R.S., Bhattacharjee, P.K.: Combined generation and transmission system expansion planning using implicit enumeration and linear programming technique. J. – Inst. Eng. India Part EL Electr. Eng. Div. **86**, 110–116 (2005)

3. Abido, M.A.: Multiobjective particle swarm optimization for environmental/economic dispatch problem. Electr. Power Syst. Res. **79**, 1105–1113 (2009)
4. Kundur, P.: Power System Stability and Control. McGraw-Hill, New York (1994)
5. Srivastava, L., Singh, S.N., Sharma, J.: Estimation of loadability margin using parallel self-organizing hierarchical neural network. Comput. Electr. Eng. **26**(2), 151–167 (2000)
6. Subramanian, S., Ganesan, S.: A simple approach for emission constrained economic dispatch problems. Int. J. Comput. Appl. **8**(11), 39–45 (2010)
7. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: Proceedings of IEEE Neural Networks, Pisctway, NJ, vol. IV, pp. 1942–1948 (1995)
8. Krusienski, D.J., Jenkins, W.K.: Design and performance of adaptive systems based on structured stochastic optimization strategies. IEEE Circuits Syst. Mag. **5**, 8–20 (2005)
9. Robinson, J., Yahya, R.S.: Particle swarm optimization in electromagnetic. IEEE Trans. Antennas Propag. **52**, 397–407 (2004)
10. Kennedy, J., Eberhart, R.C., Shi, Y.: Swarm Intelligence. Morgan Kaufmann, San Mateo (2001)
11. Hung, H.L., Lee, S.H., Huang, Y.F., Wen, J.H.: Performance analysis of PSO-based parallel interference canceller for MC-CDMA communication systems. European Transactions on Telecommunications (to appear)
12. Wen, J.-H., Lee, S.-H., Huang, Y.-F., Hung, H-L.: A sub-optimal PTS algorithm based on particle swarm optimization technique for PAPR reduction in OFDM systems. EURASIP [6] J. Wirel. Commun. Netw. Article in Press, ID 601346, May 2008
13. Clerc, M., Kennedy, J.: The particle swarm-explosion, stability, and convergence in a multidimensional complex space. IEEE Trans. Evol. Comput. **6**, 58–73 (2002)

Real Time 2D Shape and Object Recognition: A Child Learning Approach

Parisimita Sarma(✉), Subham Kakati, and Shivom Deb

Department of Information Technology, Gauhati University,
Guwahati, Assam, India
Parismita.sarma@gmail.com, subhamkakati@gmail.com,
debshivom6190@gmail.com

Abstract. It is very essential to identify shapes of different objects surrounding us. Specially the early learner children should be able to recognize the geometric shape of any object at a first glance. In this system, we are proposing a method to identify both the geometric shape and object recognition when the child will bring the item nearer to the web camera. Generally it is seen that most of the time shape and object recognition are done in scattered and lengthy way. But in this proposed work we have used two modules for identification and recognition inside one methodo. Canny edge detection algorithm is used to distinguish and check the geometric shape of an object in real time. We have also used SSD and mobilenet to classify some of the common objects.

Keywords: Canny · BGR · Contours · Frozen graph · SSD · Mobilenet

1 Introduction

1.1 Child Learning

Now a days there are numerous ways to teach the children. Among them computer aided technology is supposed to be more fruitful. A three to five years old child should be able to recognize shape of an object. Also he/she should recognize what the object is. Pictorial representation is always a good way of teaching. With the help of this application a child will be able to understand and recognize the geometric shape of a simple object when brought to the front of a digital camera in real time. We have not tried for complex objects or shapes, as initially the children should be clear about only the basic shapes and objects surrounding them. We think it will be best way to make a child understand about appropriate classification of objects according to their shape.

1.2 Shape Recognition

When a child learn to identify objects near by him, at first the shape makes him able to recognize the object. When we teach a child about a shape initially they should be taught about the basic geometric shapes. Circle, triangle, square, rectangle etc. are fundamental geometric shapes which are easy to remember for a child. Basic shape recognition process can be divided into a number of techniques like Detecting and

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 113–122, 2019.
https://doi.org/10.1007/978-981-15-0108-1_12

Indentifying, Miming and Memorizing for future. In this proposed work we are using Open CV to built the system.

1.3 Object Recognition

For a child learner, object recognition means correctly identify the object from its shape, texture and color. Here in this report we are considering commonly seen objects on real time environment. There are a number of algorithms which can be chosen collectively for object processing and identification. We can see huge applications which are based on those group of algorithms. Appropriate model is required for training of the images for accurate object recognition. Recognition process is based on vastness of trained images. Accurate classification of objects from their distinctive features is an important action. Platform to carry out all these tasks are dependent on designer but we have to maintain the correct processing sequences.

2 Literature Review

As our work of shape and object detection is based on OpenCV platform, we have studied a number of recent papers. A short discussion on these papers are as follows.

According to Lew [1] texture is the major feature for shape recognition. It is an optical characteristics and can directly find segments that exists in a certain class. Marr [2] proposed two types of edge detection. One is due to variation in intensity which can range to a considerable scale and the other is surface discontinuities. He explained boundaries can be detected due to intensity changes of the image. Another research paper [3] experimented for identification of 2D geometric shapes of triangle, square and many more with conversion from 3D RGB color model. They took help of red, green and blue color and got result with high accuracy. Patel et al. [4] proposed method for identification of location, which can categorize objects at different level. This paper discusses about some of the methods which are used to detect different geometric shapes from an image. Object detection is done with help of image processing algorithms and compared those identified objects with features of geometric figures. Kumbhar with his assistance [5] used Open CV and transformed objects into some specific coordinate points, finally the objects were detected in a rectangular box. Their method distinguishes an object if it is identified inside a box only. Another paper by Gomez [6], gray scale concept was introduced and according to this method pixel values have to be presented in a definite sequence. Different geometric shapes like rectangle, circle, triangle etc. can be compared with one another to discover the proper shape of the concern object.

3 Methodology

Most of the existing methods of shape and object recognition are done on scattered systems such as camera calibration, motion tracking, feature extraction etc. and they are done individually. It is a time consuming procedure. To remove this cumbersome individual processing we are proposing a method with only two modules. First one deals with recognition of geometric shapes of a real time object and second module recognizes or identifies the object. Object identification is tried with respect to category, type and classification. Open Cv platform used here is compatible with most of the programming policy.

The Fig. 1 shows the common block diagram which shows the flow of different actions. Shape detection in real time depends upon the camera connectivity. It is because capturing of an image with its proper shape is very important. OpenCv is the best tool for real time shape recognition because of its upgrading versions and serial updates. Different libraries of this tool also simply copes up with programming platform like Matlab, Python etc. Moreover Numpy and Python image processing techniques can easily be implemented. At the first stage of the application, the sample image have to read accurately. Predefined inbuilt Python functions perform this work. The attributes of the image are obtainable and properly mentioned such that they can be passed as arguments. For image analysis purpose we are using in built functions such as imread for outline identification. According to Fig. 1, if the user wants to detect the geometric shape of the figure, then "object shape recognition" branch will be executed. The phases in this branch are image preprocessing and enhancement, edges will be detected next, contour generation will identify the shape and label it appropriately. On the other hand if only object should be recognized then dependencies should be loaded and proper set up of the environment should be performed.

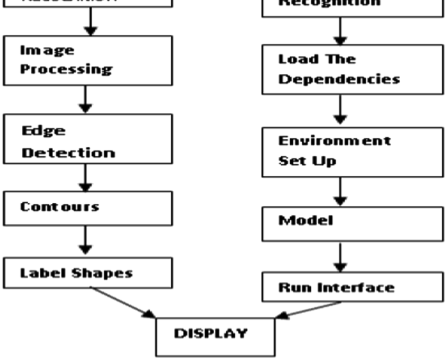


Fig. 1 Single block diagram for shape detection and recognition

3.1 Shape Detection

The Fig. 2 shows the steps to perform shape detection. Basically we are considering this approach to detect five to six geometric shapes which a child should know. Bounding rectangle approach is used for this purpose. It can detect shapes having the sides of triangle, rectangle, square, pentagoan, hexagoan. For sphere and circle detection another approach is used. Further with the help of these basic sides recognition by bounding rectangle technique, recognition of extended figures can also be done. This algorithm basically estimates height and width of shapes. The sizes of the shape assists to identify the correct shape. As pointed out in the block diagram we are using canny edge detection algorithm to recognize edges of the figures. We have evaluated contours, then projected the contour with precision compared to contour parameter, leave out minor or non convex stuff, calculated vertices of the polygons, calculated cos() of all the bends. Then sorting was performed according to the degrees which were calculated, also we computed total number of vertices to find out contour's shape. At last we were able to get the correct geometric shape of the object. We were not trying for very complex geometric figures but were able to recognize some common shape aimed for very young learner. Identification of a spherical shape is done by implementation of contour area and building rectangle process. As shown in Fig. 2, radius of the figure will be found out. Next absolute value range of that radius is assigned with respect to display screen. In our work absolute value of (1 − width/height) is evaluated, if it is less than or equal to two and at the same time absolute value of (1 − (area of circle/math pi * radius * radius)) is less than equal to .2 then the shape is recognized as a circle. We are going to discuss the steps used to identify different geometric shapes.

Defining the Video-Capture Object: We have to generate a video captured object in order to acquire frames. The argument of the video- capture is the index term or device index of the camera. We need to assign the device index to specify the camera in use with the following function. capt = cv2.Videocapture(0).

Defining the Codec: A Four Character Code called Fourcc is used for recognition of color or pixel format, video format etc. It takes 32 bits or 4 bytes as the character used here. It is an efficient one and easy to grip but very limited too. Some of the additional renowned fourcc's include "DIVX", "JPG", "JPEG", "DX50" and so on.

fourcc = cv2.VideoWriterfourcc(*'XVID')

Video Writer Object Creation: The syntax of video writer object function is obj = cv2.VideoWriter ('output file', 'codec', 'fps', (width, height)), as observed from the syntax the first parameter is the output path, second parameter is the fourcc codec, third parameter is the preferred frames of the output file, height and width are the dimension of output file.

Capture Frame by Frame of an Object: ret, frame = cap.read(), we acquire the return value from the camera frame using a Boolean called ret, this value is either true of false. A true value means proceed to succeeding frame and if not we discontinue. With cap as video-captured object and read function we capture frame by frame from our video source.

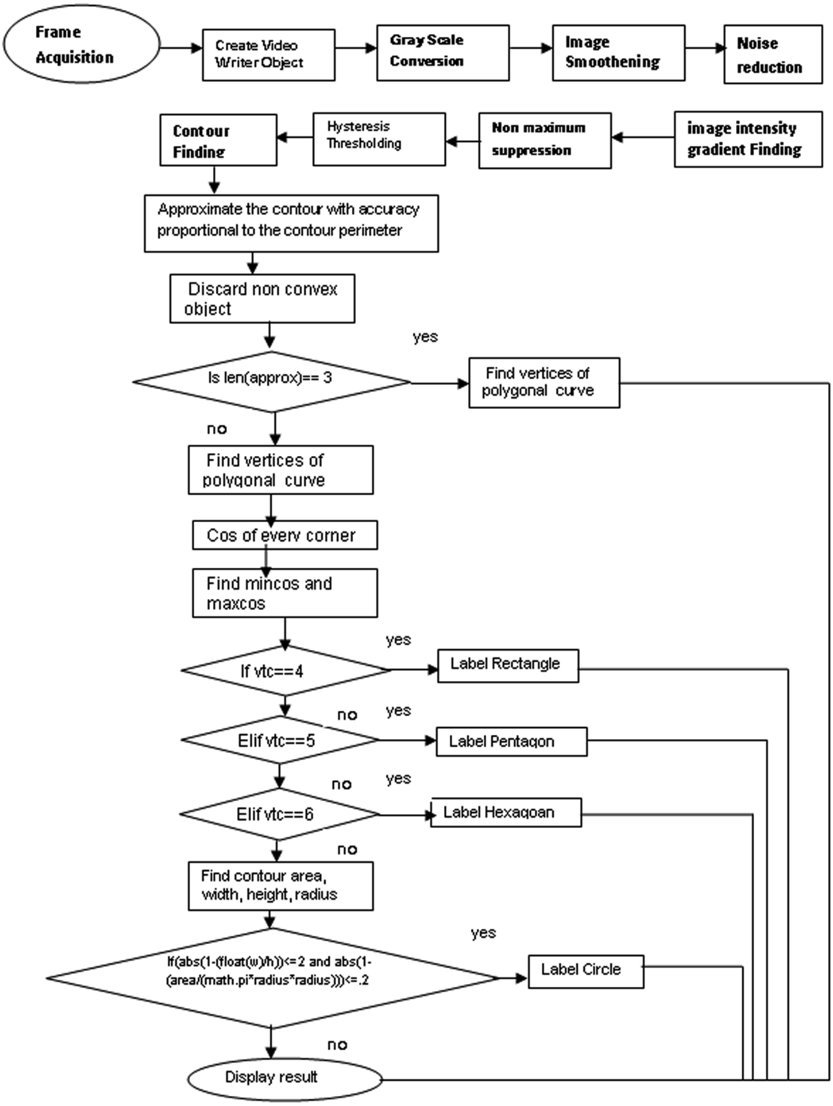

Frame Acquisition
Create Video Writer Object
Gray Scale Conversion
Image Smoothening
Noise reduction
image intensity gradient Finding
Non maximum suppression
Hysteresis Thresholding
Contour Finding
Approximate the contour with accuracy proportional to the contour perimeter
Discard non convex object
Is len(approx)== 3
yes
Find vertices of polygonal curve
no
Find vertices of polygonal curve
Cos of every corner
Find mincos and maxcos
If vtc==4
yes
Label Rectangle
no
Elif vtc==5
yes
Label Pentagon
no
Elif vtc==6
yes
Label Hexagoan
no
Find contour area, width, height, radius
If(abs(1-(float(w)/h))<=2 and abs(1-(area/(math.pi*radius*radius)))<=.2
yes
Label Circle
no
Display result

Fig. 2 Block diagram for shape detection

Canny Edge Detection Algorithm: It is a multi stage algorithm works efficiently to distinguish edges of multiple objects from a digital image. It can work on real time with the help of a camera thus able to recognize the appropriate shape of object. The function signature is: cv2.Canny (frame,82,220,2). The first parameter of the function indicates the frame; second and third parameters means Minvalue and Maxvalue. Last parameter is aperture size which utilizes the image gradients. The fundamental operations of the algorithm is explained below.

a. Noise reduction is the first action done by this algorithm. Noises obstruct in correct detection of edges. A 5 × 5 Gaussian filter is used for smoothening the image. This filter eliminates high frequency noises.
b. Intensity Gradient detection of Image: The smoothen image is now filtered from both the direction with the help of Sobel kernel. This instructs the algorithm to gain derivation in horizontal (Gx) then vertically (Gy). These two images help to find out edge gradient and orientation for every pixel. Its value is finally rounded off to all four directions.
c. Non maximum suppression: After getting the appropriate magnitude and direction the unnecessary pixels have to eliminate from the image. Individual pixel is experimented to know its local utmost in the neighborhood through the gradient directions. Pixel's maximum neighborhood is evaluated in different gradient direction. Next hysteresis thresholding is considered otherwise set to zero (suppressed).
d. Hysteresis Thresholding: With the help of this experiment real edges are determined. Two threshold values minValue and maxValue are considered respectively. If intensity gradient of an edge is more than maxValue then it is considered, if less than minValue then discarded. On the other hand if the value lies in between then further test is performed for connectivity.

Contour Utilize: Another issue for shape detection is to identify the contour. Direct Functions are available that calculates the length of an arc. That function checks whether the shape is closed or only an arc. So it returns a Boolean value true or false. Contour approximation estimates one shape to another with smaller number of vertices. Convexity can be confirmed with another Boolean function.

3.2 Object Detection

We are using a number of imported packages which comprises of all common modules that suits for python 2, python 3 and object detection as well as recognition. Six.moves.urllib is an example of this package. We are discussing here some other modules which are used for object recognition. Figure 3 shows the block diagram for object recognition process. The Tarfile module is used to deal with lots of compressed files in a simple way. Tensorflow is a accuracy based system and used in our application for machine learning aspects of image processing and object detection. Sys module puts admission to the variables and their allied tasks which are specific and maintained by the interpreter. This is a python compatible module. Zipfile is an efficient and useful module that facilitated us in object detection by providing tools to attach, write, create and reading the captured objects. We are using String IO file module that reads and writes string buffer separately. Another efficient module called matplotlib is used in the application for plotting 2D objects. It basically provides many environments necessary for image processing and image types compatible with diverse platforms. This module also makes us possible to generate histogram, graphs. A default dictionary is used to keep track of different objects and later retrieve them. Initially the key value is set to zero. This dictionary will detect more items one after another which are already stored as a set of keys in the dictionary. In python it is passed as a single code. For an application which can detect object, most important is deal of an object detection environment. Reference to the imports in current localities are necessary and it is done by setting up the environment. Next selection of modules are done by carrying out different tests. These activities are performed by few simple code through Model name, Model file and Download Base consequently. Objects can be determined by graphs. Data are necessary to define a graph. A path is precisely allocated to the models for loading. Next number of classes and strings list are entered correctly meant for labeling which must appear after detection of the object. At last the objects should be labeled with their names appropriately. For this a label map is loaded to the environment which helps in proper detection of the object by their category name. Helper code is used to move image closer to the detection area. For this all the parameters like image width, height etc. have to be changed. The algorithmic steps for object detection is as follows:

1. Acquisition of the frame is done.
2. Image dimension expansion is completed.
3. Environment set up.
4. Model specification.
5. Direct the path for frozen detection and label map.
6. Load into memory the frozen tensorflow.
7. Label map loading.
8. For individual image get handles to input and output tensor.
9. Reframe from box co-ordinate to image co-ordinate.
10. Add batch dimension.
11. Run interface.
12. Display the output.

It is experienced that all images are not so easy to detect. Some images need mathematical calculation to detect but some can be detected easily. Some helper codes are necessary for the complicated images to detect. Our object detection job was tried with ssdmobilenet (lightest weight model), SSDs (single shot detectors), fast and efficient MobileNets architecture models.

The required image is searched by defining path or other alternatives. Session objects determines operations of object execution and tensorflow object. Vital tensor names are set as keys, these are consists of detection scores, detection masks, detection boxes, number and class detection. If the displayed object matches with any of the tensor key values then the object is detected. Reframing is performed for translation of mask from box coordinates to image coordinates for the image fitting. As all the detected images are in float Numpy arrays image initially, they should be transformed to proper types such as classes, boxes and scores. Otherwise there will be lack in detection. Initially arrays of pictures are loaded into Numpy array. Next dimension expansion is carried out to fit our object. Then only actual detection process starts by matching it with the dictionary stored database. Output will be finally detected and pictures with associated key values like boxes, scores, class, number etc. will be displayed.

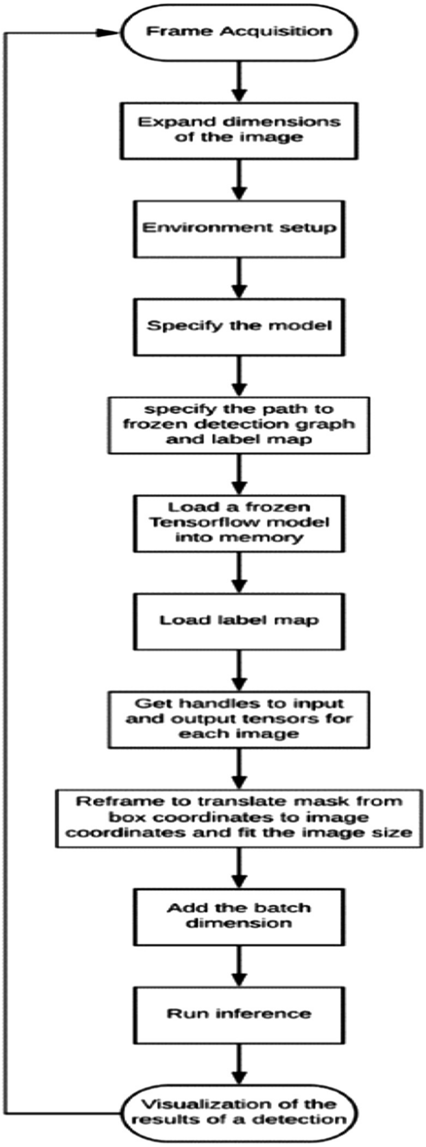

Frame Acquisition
Expand dimensions of the image
Environment setup
Specify the model
specify the path to frozen detection graph and label map
Load a frozen Tensorflow model into memory
Load label map
Get handles to input and output tensors for each image
Reframe to translate mask from box coordinates to image coordinates and fit the image size
Add the batch dimension
Run inference
Visualization of the results of a detection

Fig. 3 Block diagram of object recognition

4 Results and Discussion

Throughout the whole work we have used Open CV which handles different components capable of shape recognition as described in methodology section. We are basically working for detection of simple geometric shapes which is supposed to be very preliminary tool to teach a child about the surrounding objects. Polygon detection is most significant task in shape detection. This method shows the way for an error free geometrical shape recognition. Canny edge detection algorithm is used to find out accurate edges of different geometric shapes. Edges are adjusted using Bounding rectangle formulae. Cos value of the angles and value of contours are evaluated. Thus we get a efficient and correct geometrical shape recognition system captured from real time video frame. As shown in Fig. 4, a circle on a mobile phone is detected when brought nearer to the video camera of a laptop. We have used MS COCO model as training set. All along we have also used "ssd" as localizer with "Mobilenet" as classifier to detect the object. For better classification and recognition graphical proceedings like Frozen inference graph is used.

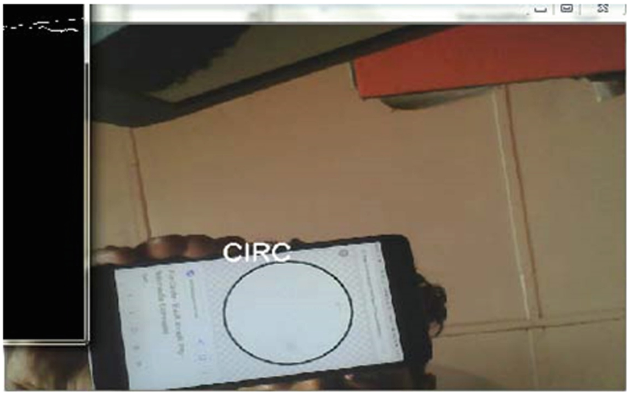


Fig. 4 Snapshot of a circle detection

The near input images are tested very accurately and finally the image inside the frame is recognized with their respective keys such as the box, its class and the score of the detected object as shown in the Fig. 5. This screenshot shows an apple detected from a photograph on a mobile with 65% score. This score means the probability of the detected object to be a apple is 65%. In this way we are able to detect most of the commonly available surrounding things with a good accuracy.

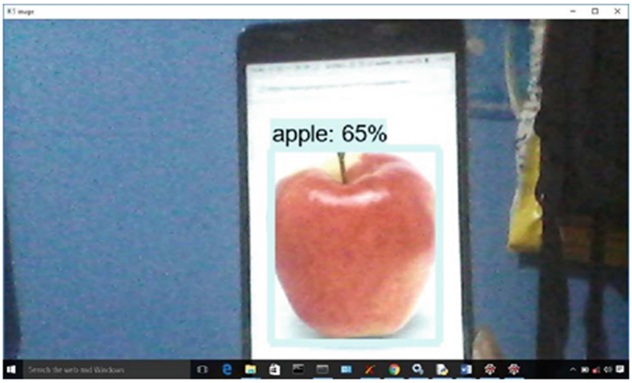


Fig. 5 Snapshot of an apple recognized

5 Conclusion and Future Work

In our system, for shape recognition initially the shapes are examined and determine their geometric outline. We can see that approach for object recognition and geometric shape detection is a fast and precise procedure. We have used Open CV, regular polygon techniques, Canny Edge Detection algorithm, Bounding rectangle formula and Contours for a fast, efficient and correct geometrical recognition of different shapes. On the other hand, in object recognition procedure most important and mandatory is importing different libraries which we have done efficiently. We have used a procedural set up for object detection. MS COCO is used as a training dataset of images. Although this method was able to recognize some shapes and objects, yet we left with some other shapes. Limited number of object recognition is a drawback we have noticed. Moreover sometimes more time is taken by the system to process the acquired frame. There is scope to improve the accuracy of this system in near future by taking a larger database.

References

1. Lew, M.S. (ed.): Principles of Visual Information Retrieval. Springer, Heidelberg (2013)
2. Marr, D., Hildreth, E.: Theory of edge detection. Proc. R. Soc. Lond. Ser. B. Biol. Sci. **207** (1167), 187–217 (1980)
3. Rege, S., et al.: 2D geometric shape and color recognition using digital image processing. Int. J. Adv. Res. Electr. Electron. Instrum. Eng. **2**(6), 2479–2487 (2013)
4. Patel, S., et al.: 2D basic shape detection using region properties. Int. J. Eng. Res. Technol. **2** (5), 1147–1153 (2013)
5. Kumbhar, P.Y., et al.: Object recognition using Open CV. Int. J. Res. Emerg. Sci. Technol. 4 (4), 39–43 (2017)
6. Gomez, S.: Shape recognition using machine learning. In: 2011 6th Colombian Computing Congress (CCC). IEEE (2011)

Adaptive e-Learning System for Slow Learners Based on Felder-Silverman Learning Style Model

Lumy Joseph[1(✉)] and Sajimon Abraham[2]

[1] School of Computer Sciences, Mahatma Gandhi University, Kottayam, India
lumybiju@yahoo.com

[2] School of Management and Business Studies, Mahatma Gandhi University, Kottayam, India
sajimabraham@yahoo.com

Abstract. Adaptive learning plays a significant role in online learning. It enables the students, to decide what to select, how to learn and how to assess themselves. This method provides a personalized learning path and enabling them to involve in, as they advance through the learning resources. To demonstrate, this study has developed an adaptive e-learning system (AeLS), using Lesson activity in Moodle, to teach a course in Computer Graphics, for the undergraduate students of Computer Applications Programme. The Felder-Silverman Learning Style Model has also employed with the intention of integrating diverse learning styles of students. To evaluate the effectiveness of the system, the study has resorted to the use of the statistical method; independent two-sample t-test among two groups of slow learners. Experimental evaluation demonstrates that AeLS is able to achieve comparable performance on a group of slow learners, than who used traditional face-to-face class room teaching method.

Keywords: Adaptive e-learning · Felder-Silverman Learning Style Model · Learning path · Slow learners

1 Introduction

The role of education is highly significant in the development of a nation and many of the countries are considering education as a thrust area for ensuring sustainable development. In the field of teaching-learning, many changes are happening and traditional classroom teaching is moving to e-learning and m-learning. But, there is significant difference in the way teachers give instructions to learners in a traditional class room and in the delivery of instructions to learners in an online and/or blended course. Monitoring learners in a classroom-based learning is relatively easy. But in a self-access learning, it is more difficult because of the individualized nature of the work and the reduced level of teacher contact.

In order to make the teaching–learning process more effective, it is required to have an education model that uses recent technologies. With this intention, this study has developed an efficient learning system which provides suitable materials according to

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 123–135, 2019.
https://doi.org/10.1007/978-981-15-0108-1_13

the interest of learners and helps them to progress their academic performance. An adaptive learning system can address all these issues in an effective way. To achieve this, in this study, an adaptive e-learning system (AeLS) was developed in MOODLE, for slow learners. The learning paths of these learners were also identified using the activity logs obtained from the Moodle weblog data.

The remaining sections of this paper are organized as follows: Sect. 2 briefly reports related works. Section 3 includes a theoretical framework of the study. Section 4 elaborates the research method and processes involved in the development of the proposed system. Section 5 provides the analysis result and its interpretation. Finally, Sect. 6 concludes the study by mentioning the scope for further research.

2 Review of Literature

This section explains about some of the relevant studies happened in this research area. Dealing with slow learners is a difficult task for many of the 21st century teachers [1]. It is a big challenge for the educators to teach slow learners in a suitable way and to provide them quality education [1]. In the current education paradigm slow learners as well as fast learners are forced to learn the same learning resources in a similar manner [2]. At present, teachers adopt their own techniques to handle them. In a traditional classroom environment, most of the time teachers are compelled to follow a prescribed syllabus and they are giving instructions, without considering the learner characteristics like learning style, knowledge level, attitude, etc. But in an online environment, slow learners are compelled to finish the learning process before they get expertise in the respective learning content. Hence, there is a need for further research to provide teachers with certain openings for supporting slow learners.

A study by Ives (2010) describes the significance of changing to a new paradigm in the area of education where educators expected to be worked with students from diverse backgrounds [2] and suggests to apply the instructional theory [2] to facilitate teaching. There exist many studies explaining about the benefits of online learning to disabled people. Since e-learning centers more on individual learners, it provides personalized learning and helps learners to advance their learning from any location [3]. A study on the impact of e-learning with regard to the academic performance of learners and its benefits to higher educational institutions [4], showed that e-learning considered different types of learner perspectives and provides an opportunity to disabled students to progress their learning from any place.

According to the study [5], adaptive e-learning is one of the best method to provide ultimate instruction to learners. An article on the topic ‘Integrating-Machine-Learning-Education-Technology’ written by Harish Agrawal in 2017, explains the advantage of adaptive learning in classrooms. It provides personalized learning and helps learners to improve their academic performance by providing a continuous learner feedback. It also helps them to adopt the mode of presentation of learning material according to the requirements and interests of learners. In teachers’ perspective, they can assess their students individually or as a whole through this method, and can provide extra care individually. Teachers can adjust their speed of teaching as well as they can carry out the delivery of learning material according to the progress of the learners.

The study [6] explained about the need for understanding individual differences and behavioral characteristics of each learner for providing the most suitable learning path. The studies [7, 8] explain the advantage of integrating learning style into a personalized e-learning system, for providing suitable educational resources and to satisfy the request of learners. Hung et al. (2016) mentioned the significance of personalized learning and the role of learning styles during the design of adaptive learning systems [9]. The Felder-Silverman Learning Style Model is a widely used learning style model by many intelligent learning systems to provide support to their tailored learning services [10].

A study on, how to keep learners engaged with adaptive triggering [11], shows high value and importance in the area of adaptive learning concept. An adaptive learning system can be developed in different ways. Creating a course with learning contents of different levels of complexity [12] and designing a learning path based on attributes that describe learning contents and student characteristics [13] are some examples. In order to improve an existing e-learning system used in Moodle, the study [14] developed a method for creating adaptive e-learning systems, by considering learning style as one key factor.

Log data related to student activities is a key resource for acquiring knowledge about students' learning behavior in online environments [15, 16]. A study has made to understand students' online behavior and developed a method using Moodle log data for extracting and visualizing students' learning behavior [16].

There are many studies explaining the application of data mining in an e-learning environment. Poon et al. (2017) used the sequential pattern mining to perform weblog data analysis and collected navigational patterns to get more understanding about the learners [17]. A study on clustering of slow learners' behavior [18] showed the importance of applying clustering algorithms on slow learners' data to find optimal patterns of learning for improving students' learning capability. According to Vieira et al. (2018), students participating in a learning environment follow different learning paths and modeling of students using data mining methods helps to explore their learning paths [19].

All these studies indicate the importance of integrating recent technologies in the teaching and learning scenario to acquire quality education. To achieve this, researcher has developed an adaptive e-learning system (AeLS) to support slow learners, to provide an efficient learning mechanism. It is an instruction design-based course module, developed using Lesson activity in Moodle, to teach the subject Computer Graphics. The system developed was evaluated by comparing variances of learning performance between the group of students who used instructional design-based AeLS and the group of students who used traditional face-to-face class room teaching method. The topic 'scaling' in Computer Graphics was selected to explain the concept, based on the understanding dimension in FSLSM. The learning paths of all learners were identified using activity logs obtained from the Moodle weblog data.

3 Theoretical Background

Realizing learning strategies, using pedagogical approaches is assuming great significance in e-learning and blended learning. Since this work includes of some sort of educational pedagogy, theories related to Felder- Silverman Learning Styles model (FSLSM), slow learners' characteristics, etc. are discussed in this section.

3.1 Felder-Silverman Learning Style Model

Learning style refers to the concept that individuals differ in regard to what mode of instruction or study is the most effective for them [20]. Many studies have established that learning styles not only affect the way a learner learns, but it also exerts immense power on the way knowledge is delivered [21].

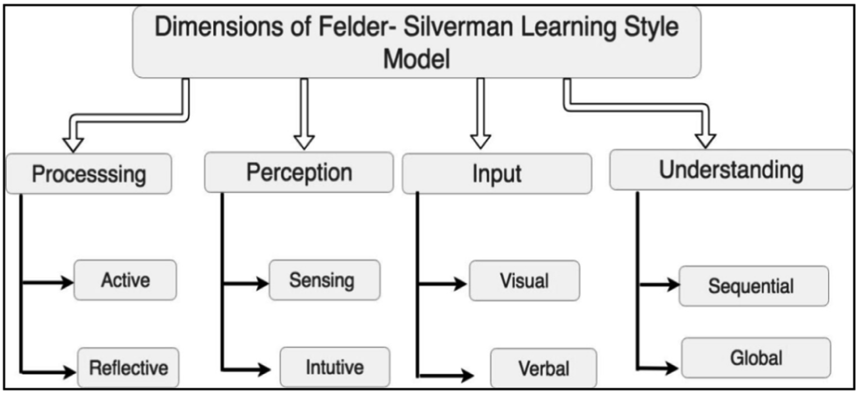


Fig. 1. Dimensions of Felder - Silverman Learning Style Model.

According to Graf and Liu (2009), there exist several learning style models [22]. This study has used the Felder–Silverman Learning Style Model [10] since it has been widely applied in engineering education and research related to learning technologies [23]. Figure 1 describes the four dimensions in FSLM. Perception relates to the type of information a student prefers to identify; processing dimension shows how perceived information is converted into knowledge; understanding describes the manner in which students' progress towards understanding; finally, input considers the way in which students prefer to receive external information [24].

3.2 Characteristics of Slow Learners in an e-Learning Environment

The major difference between slow learners and fast learners is that slow learners always work on all tasks very slowly; they need extra time to read and understand learning materials properly. They prefer a step-by-step learning path and could grasp a concept only when it is linked to the previous topic. There are many hints for identifying slow learners in an online learning environment (see Table 1) and these hints can be used for detecting their learning style [22].

Table 1. Hints used for identifying slow learners.

Behavioral pattern	Meaning
Preference on the usage of learning objects such as outlines, course overview, page/details on a topic	Preference for details
Performance on questions	Performance on questions dealing with facts and concept of a topic
Time spend on learning objects such as outlines, course overview, page/details on a topic	Duration
Students' navigational behavior	Number of skipped objects/Preference to go through the course(step by step or in large leaps)

4 Research Methods

In the process of developing adaptive e-learning system, this study has gone through various stages and considered many factors, like learners' characteristics, educational theory, etc. in the research process (Fig. 2 shows these steps).

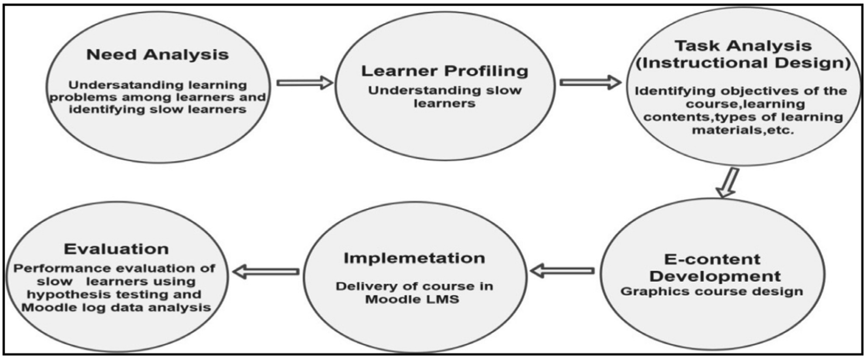

Need Analysis
Undersatanding learning problems among learners and identifying slow learners
Learner Profiling
Understanding slow learners
Task Analysis (Instructional Design)
Identifying objectives of the course,learning contents,types of learning materials,etc.
E-content Development
Graphics course design
Implemetation
Delivery of course in Moodle LMS
Evaluation
Performance evaluation of slow learners using hypothesis testing and Moodle log data analysis

Fig. 2. Various stages involved in the design of the system (AeLS)

4.1 System Description: Adaptive Learning Path Creation Using Lesson Activity in Moodle

This section highlights the significance of proper instructional design for the development of e-content in a course. A task analysis (instructional design) breaks down tasks in the course and shows learners the steps they have to follow to complete the task successfully.

In this stage, proper planning is required for the delivery of the course. Since learners in an online learning environment have different characteristics, proper planning is needed for the selection of appropriate learning contents to deliver them.

In order to consider the diverse learning style of learners, this study has considered the profile of learners, learning material/content types, its difficulty level etc. and also the learning strategies during the design of learning path to this course. This is achieved by considering various types of learning contents like power point presentations, video, images, text files, etc. as well as by bearing in mind the order in which the contents has to be arranged.

The Lesson activity in Moodle has used to provide an adaptive learning path to its learners. In order to offer better learning experience to slow learners, researchers have incorporated the characteristics of visual learning style and sequential learning style dimensions in FSLSM.

In an e-learning environment, a learning path is considered as a sequence of learning activities carried out by the learners while they are moving all the way through learning units. Each learning unit is a theoretical depiction of a course, a lesson, a workshop, or any other formal or informal teaching or learning event. A learning path may be considered adaptive, if proper learning activities can be suggested for each learner [25]. Figure 3 explains the process of adaptive learning through the roles of its components.

According to moodle.org, adaptive ability feature is the most important difference between a lesson and other activity modules in Moodle. The 'Lesson activity' in Moodle helps to create content pages with a variety of alternatives and pathways for learners [26]. This Moodle feature has used in this study to provide an adaptive learning path and to make the course more interactive. The topic 'scaling' in Computer Graphics was used to explain the concept. A Lesson consists of a sequence of HTML web pages with two types-content pages and question pages. Figure 4 is an illustration of the adaptive learning path creation using lesson activity.

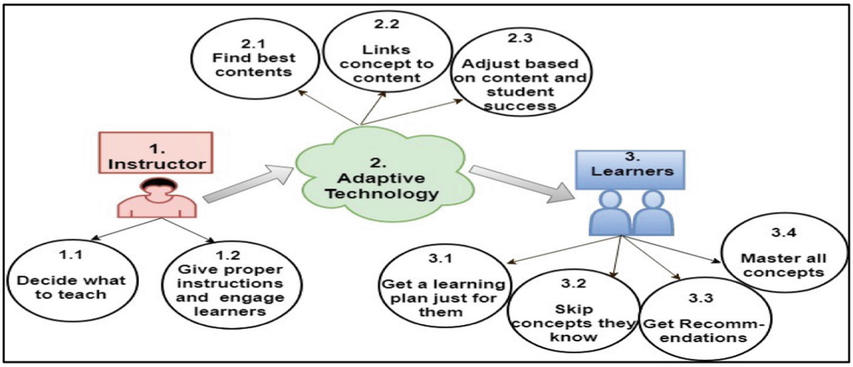

2.1 Find best contents
2.2 Links concept to content
2.3 Adjust based on content and student success
1. Instructor
2. Adaptive Technology
3. Learners
1.1 Decide what to teach
1.2 Give proper instructions and engage learners
3.1 Get a learning plan just for them
3.2 Skip concepts they know
3.3 Get Recomm-endations
3.4 Master all concepts

Fig. 3. Three components in adaptive learning - instructors, adaptive technology and learners

Initially, at Level-0, an introduction about the concept of scaling is explained. In Level-1, further details like definition of scaling, various scaling equations and scaling factors are discussed. In Level-2, evaluation of each learner is done. In this level, students were supposed to answer few questions related to previous topic. If they respond to all questions correctly, they are directed to Level-1 where the subsequent matter is explained, otherwise, they are directed to the Primer page in Level-3 where some remedial content for a detailed study is included. In this page, any number of learning materials can be included. It helps to improve their knowledge in that topic. The possibility of including more pages in lesson activity is an added advantage to all types of learners. In this Lesson activity, different types of resources such as videos, images, power point presentations, text files, a web page in which URLs provide links to learning materials which contains images, power point presentations, etc. are included.

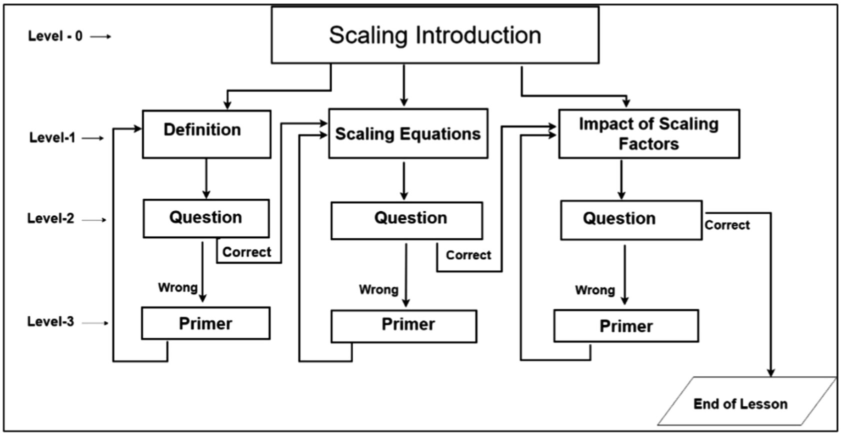


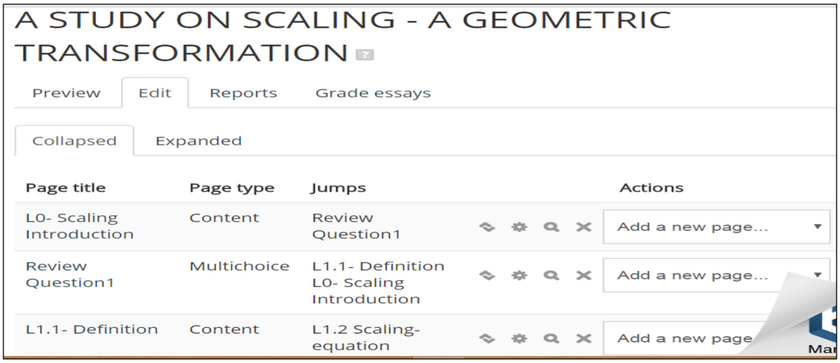
Fig. 4. An adaptive learning path creation using lesson activity for scaling

In this study, the learning path of slow learners is explained with the help of page-ids obtained from the log report of Moodle log data. Since a Lesson is a series of html pages, in order to explain the concept 'scaling', this study has included many web pages. These pages may be either content pages or multiple choice pages. Table 2 shows a detailed description of the pages used to explain the concept of scaling.

Table 2. Pages used in the lesson activity ‘Scaling’ - A summary

Page-id	Page title	Page code	Page type
64	Scaling Introduction	L0	Content
68	Review Question1	RQ1	Multichoice
65	Definition	L1.1	Content
74	Scaling-Equation	L1.2	Content
69	Review Question2	RQ2	Multichoice
70	Detailed Definition	L1.3	Content
72	Review Question3	RQ3	Multichoice
75	Example Scaling	L1.4	Content
78	Review Question4	RQ4	Multichoice
77	Impact of Scaling factor	L1.5	Content
73	Review Question5	RQ5	Multichoice

Any number of pages can be included in the lesson activity and each content page is made up of with text, images, videos, and so on. Students can do the learning process through these pages in a linear pathway or they can select any branching paths. Since a lesson activity has the capability of including questions, students can perform a self-assessment about their learning. In this learning process, students can directly move on to other content pages, feedback pages or further questions, subject to their responses. Figure 5 shows a sample screen shot of the Lesson activity generated from Moodle.



A STUDY ON SCALING - A GEOMETRIC TRANSFORMATION

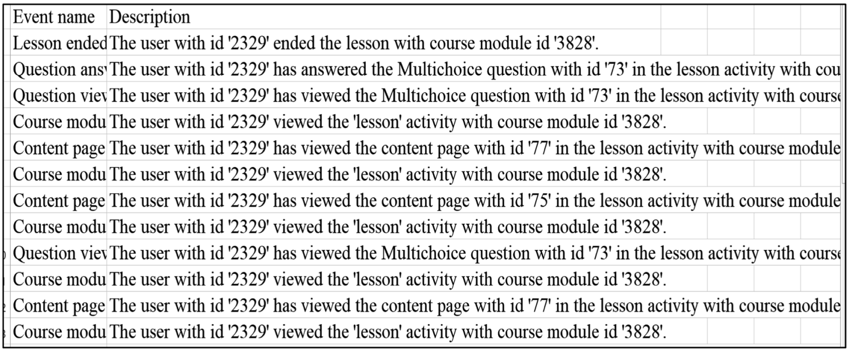
Page title	Page type	Jumps	Actions
L0- Scaling Introduction	Content	Review Question1	Add a new page...
Review Question1	Multichoice	L1.1- Definition L0- Scaling Introduction	Add a new page...
L1.1- Definition	Content	L1.2 Scaling-equation	Add a new page

Fig. 5. A screenshot of the lesson activity created for the small learning object ‘Scaling’.

5 Data Analysis and Interpretation

5.1 Learning Path Analysis of Slow Learners

Log data retrieved from Moodle could be used for understanding the learning path of all learners. It provides information regarding the types of content used, the time spent on each learning material, etc. Figure 6 shows an activity report of a particular learner generated from Moodle.



Event name	Description
Lesson ended	The user with id '2329' ended the lesson with course module id '3828'.
Question ans	The user with id '2329' has answered the Multichoice question with id '73' in the lesson activity with cou
Question viev	The user with id '2329' has viewed the Multichoice question with id '73' in the lesson activity with course
Course modu	The user with id '2329' viewed the 'lesson' activity with course module id '3828'.
Content page	The user with id '2329' has viewed the content page with id '77' in the lesson activity with course module
Course modu	The user with id '2329' viewed the 'lesson' activity with course module id '3828'.
Content page	The user with id '2329' has viewed the content page with id '75' in the lesson activity with course module
Course modu	The user with id '2329' viewed the 'lesson' activity with course module id '3828'.
Question viev	The user with id '2329' has viewed the Multichoice question with id '73' in the lesson activity with course
Course modu	The user with id '2329' viewed the 'lesson' activity with course module id '3828'.
Content page	The user with id '2329' has viewed the content page with id '77' in the lesson activity with course module
Course modu	The user with id '2329' viewed the 'lesson' activity with course module id '3828'.

Fig. 6. Log report of the lesson activity viewed by a student.

A pictorial representation of the learning path generated by the first learner is explained in Fig. 7. Repeated use of some page-ids and varying learning paths shows a slow learner's difficulty in understanding few learning materials. An analysis on the log data showed that they visited some pages many times and spends more time on these pages. This happens, because in the question page if they were not able to answer a question correctly, it provides a chance for revisiting the page containing learning materials.

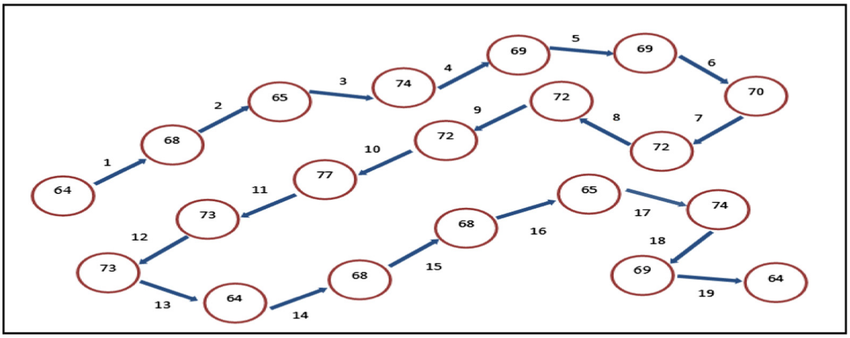


Fig. 7. A learning path showing the movement of a learner based on the page-ids.

5.2 Experimental Method

This section explains about the experimental design done among the two groups; an experimental group and a control group. A comparative study of the performance of both groups has made; one group was taught using the adaptive e-learning system and the other group was taught in the traditional face-to-face classroom teaching method. The two groups were pre-tested and post-tested. The pre-test dealt with the topics Translation and Scaling in Computer Graphics. A post-test on the same topics was also conducted for both the groups after the completion of the course. The results of the pre-test and post-test obtained are explained in Table 3.

Table 3. A descriptive statistics of pre-test and post-test marks obtained.

	Mean		SD	
Results	Pre-test	Post-test	Pre-test	Post-test
Experimental group	10.2	17.85	2.745331	0.961085
Control group	10.425	16	2.763841	2.492093
N	20	20		
t-value	0.23103	2.770504		

5.3 Analysis of the Effectiveness of AeLS for Slow Learners

This section demonstrates the experimental evaluation of the system developed. In order to evaluate the effectiveness of AeLS, the performance variance in the pre-test and post-test marks of both the groups were compared and it is found that there is significant difference in the post-test marks of both groups (see Table 3). Hence to test the effectiveness of the teaching method, a statistical hypothesis testing, independent sample t-test was used.

Table 4. Hypothesis used.

Null Hypothesis H_0:	There is no significant difference between the learning performances of the group of students who used instruction design based adaptive e-learning method and the group of students who used traditional classroom teaching method
Alternative hypothesis $H_{1:}$	There is significant difference between the learning performances of the group of students who used instruction design based adaptive e-learning method and the group of students who used traditional classroom teaching method

Table 4 shows the hypothesis used in the study. To carry out the test, researchers used the statistic t as under:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{\left[(n_1-1)s_1^2 + (n_2-1)s_2^2\right](n_1+n_2)}{(n_1+n_2-2)(n_1 n_2)}}} \tag{1}$$

with d.f. = $n_1 + n_2 - 2$.

where $\overline{x_1}$ gives the mean of the first sample, $\overline{x_2}$ is the mean of the second sample, n_1 is the total number of observations in the first sample, n_2 is the number of observations in the second sample and s_1 and s_2 are the standard deviations of the two samples. Variance is the square of standard deviation. Table 5 shows the evaluation result on the performance of both groups.

Table 5. Evaluation of performance difference of slow learners.

	Experimental group	Control group
Mean	7.725	5.575
SD	2.839918	3.17173
Variance	8.065132	10.05987
N	20	20
t-value	2.201285	
Degrees of freedom (d.f.)	38	
Critical Value for t (two-tailed)	2.0244	

Since the observed t-value (2.201285) is greater than the critical value (2.0244), we couldn't accept the null hypotheses. It shows that there is a significant difference between the two teaching methods. Since, mean value of the learning performance of experimental group is significantly greater than the control group, it can be concluded that the adaptive e-learning is more effective than traditional face-to-face classroom teaching method. Hence this method will be an effective means to improve the learning performance of slow learners.

6 Conclusion

The notion of personalized learning content and the effort to implement the same is becoming a popular trend in e-learning. Hence, it is significant to incorporate the most recent technologies into the process of teaching and learning, to cope up with the new changes happening in a learning environment. In this study, the adaptive e-learning system (AeLS) developed, facilitated the slow learners to improve their academic performance. An analysis of Moodle log data helped to identify the learning path, as well as, to gauge the effectiveness of the use of instructional design-based learning content from the perspective of slow learners. An experimental evaluation done among a group of 40 slow learners showed that the proposed AeLS was highly effective. This study included only a limited number of slow learners, and used only one concept scaling to implement the system. As a future study, this method can be implemented among more number of learners and to a complete course. As a further step, the behavioral features obtained from the log data will be used for detecting the learning style of learners, based on Felder- Silverman Learning Style Model. Integration of suitable machine learning techniques like K-NN, Fuzzy clustering, Bayesian networks, Decision trees, Hidden Markov model, etc. to the Moodle log data, can be used to classify the learners, based on their learning style, and to predict their academic performance.

References

1. Borah, R.R.: Slow learners: role of teachers and guardians in honing their hidden skills. Int. J. Educ. Plan. Adm. **3**(2), 139–143 (2013)
2. Ives, C.: Instructional-Design Theories and Models, Volume III: Building a Common Knowledge Base, pp. 219–221 (2010)
3. Holmes, B., Gardner, J.: E-learning: Concepts and Practice. Sage, Thousand Oaks (2006)
4. Arkorful, V., Abaidoo, N.: The role of e-learning, advantages and disadvantages of its adoption in higher education. Int. J. Instr. Technol. Distance Learn. **12**(1), 29–42 (2015)
5. Roy, A., Basu, K.: A comparative study of statistical learning and adaptive learning. arXiv preprint arXiv:1511.07538 (2015)
6. Klašnja-Milićević, A., Vesin, B., Ivanović, M., Budimac, Z., Jain, L.C.: E-Learning Systems: Intelligent Techniques for Personalization, vol. 112. Springer, Heidelberg (2016). https://doi.org/10.1007/978-3-319-41163-7
7. Yang, J., Huang, Z.X., Gao, Y.X., Liu, H.T.: Dynamic learning style prediction method based on a pattern recognition technique. IEEE Trans. Learn. Technol. **7**(2), 165–177 (2014)
8. Radwan, N.: An adaptive learning management system based on learner's learning style. Int. Arab J. e-Technol. **3**(4), 7 (2014)
9. Hung, Y.H., Chang, R.I., Lin, C.F.: Hybrid learning style identification and developing adaptive problem-solving learning activities. Comput. Hum. Behav. **55**, 552–561 (2016)
10. Felder, R.M., Silverman, L.K.: Learning and teaching styles in engineering education. Eng. Educ. **78**(7), 674–681 (1988)
11. Kljun, M., Krulec, R., Pucihar, K.C., Solina, F.: Persuasive technologies in m-learning for training professionals: how to keep learners engaged with adaptive triggering. IEEE Trans. Learn. Technol., 1–1 (2018). https://doi.org/10.1109/tlt.2018.2840716
12. Premlatha, K.R., Dharani, B., Geetha, T.V.: Dynamic learner profiling and automatic learner classification for adaptive e-learning environment. Interact. Learn. Environ. **24**(6), 1054–1075 (2016)
13. Yang, F.: Learning path construction in e-learning–what to learn and how to learn? Doctoral dissertation, Durham University (2013)
14. Despotović-Zrakić, M., Marković, A., Bogdanović, Z., Barać, D., Krčo, S.: Providing adaptivity in Moodle LMS courses. J. Educ. Technol. Soc. **15**(1), 326–338 (2012)
15. Zaiane, O.R., Luo, J.: Towards evaluating learners' behaviour in a web-based distance learning environment. In: Proceedings IEEE International Conference on Advanced Learning Technologies, pp. 357–360. IEEE (2001)
16. Estacio, R.R., Raga Jr., R.C.: Analyzing students online learning behavior in blended courses using Moodle. Asian Assoc. Open Univ. J. **12**(1), 52–68 (2017)
17. Poon, L.K.M., Kong, S.-C., Wong, M.Y.W., Yau, T.S.H.: Mining sequential patterns of students' access on learning management system. In: Tan, Ying, Takagi, Hideyuki, Shi, Yuhui (eds.) DMBD 2017. LNCS, vol. 10387, pp. 191–198. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61845-6_20
18. Mohammad, T.Z., Mahmoud, A.M.: Clustering of slow learners behavior for discovery of optimal patterns of learning. LITERATURES **5**(11), 102–109 (2014)
19. Vieira, C., Parsons, P., Byrd, V.: Visual learning analytics of educational data: a systematic literature review and research agenda. Comput. Educ. **122**, 119–135 (2018)
20. Pashler, H., McDaniel, M., Rohrer, D., Bjork, R.: Learning styles: concepts and evidence. Psychol. Sci. Public Interest **9**(3), 105–119 (2008)
21. Carpenter, J. P.: Using learning styles to improve student learning. La. Tech Univ. (2004). http://sections.maa.org/lams/proceedings/spring2004/JennaCarpenter.pdf

22. Graf, S., Kinshuk, Liu, T.C.: Supporting teachers in identifying students' learning styles in learning management systems: an automatic student modelling approach. J. Educ. Technol. Soc. **12**(4), 3–14 (2009)
23. Kuljis, J., Liu, F.: A comparison of learning style theories on the suitability for elearning. Web Technol. Appl. Serv. **2005**, 191–197 (2005)
24. Scott, E., Rodríguez, G., Soria, Á., Campo, M.: Are learning styles useful indicators to discover how students use Scrum for the first time? Comput. Hum. Behav. **36**, 56–64 (2014)
25. Yang, F., Dong, Z.: Learning Path Construction in e-Learning. Lecture Notes in Educational Technology. Springer, Heidelberg (2017). https://doi.org/10.1007/978-981-10-1944-9
26. Kc, D.: Evaluation of moodle features at Kajaani University of applied sciences-case study. Proc. Comput. Sci. **116**, 121–128 (2017)

Unsupervised Extraction of Respiration Cycles Through Ballistocardiography

Vibhor Saran[1], Gulshan Kumar[2], and Gaurav Parchani[1](✉)

[1] Turtle Shell Technologies Pvt. Ltd., Bangalore, India
gaurav@dozee.io
[2] Department of Neurophysiology, NIMHANS, Bangalore, India

Abstract. Ballistocardiography (BCG), a non-invasive technique for measuring micro-body vibrations arising from cardiac contractions. It also contains motion arising from breathing, snoring and body movements. Long-term acquisition of respiratory signal finds relevance in various applications such as sleep analysis as well as monitoring of respiratory disorders. Current methods (such as nasal thermistor and Respiratory Inductance Plethysmography) are costly, inconvenient and require technical expertise to setup and analyse. In this paper we assess how BCG based contact-free methods can allow for an accurate, cost-effective and convenient long-term monitoring from the ease of home environment. We propose a novel algorithm to detect breathing cycles from BCG signal, achieving an accuracy of $\sim$95% in determining respiration rate for 30 s epochs with a detection rate of 72.8% compared to current methods. Long-term continuous monitoring of respiratory signals with a high accuracy will allow for detection of abnormalities like respiratory distress and apnea/hypopnea episodes.

Keywords: Ballistocardiography · Respiration rate · Respiratory effort · Contact-free · Clustering

1 Introduction

Continuous monitoring of body vitals is poised to be one of the key elements of future healthcare including monitoring critical illnesses, personalized therapeutics and predictive medicine. Continuous monitoring of patients by a team of healthcare professionals in a hospital setting has proven to be extremely effective over traditional (and sparsely) periodic manual measurements by an attendant [1]. Such manual monitoring, apart from being prone to inefficiency and human error, is also not feasible due to already stressed healthcare systems, especially in developing countries where healthcare needs are already at severe odds with their fast growing ageing populations and an increasing shortage of trained professionals – a projected deficit of 14 million healthcare workers globally by 2030 [2].

We can broadly classify applications for continuous monitoring of body vitals into two classes: (a) short term and (b) long term. The short term class involves uses like generating real-time alerts based on abnormal deviations in vitals to prevent a critical event like a cardiac arrest [3]. On the other hand, long term monitoring collects data

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 136–147, 2019.
https://doi.org/10.1007/978-981-15-0108-1_14

over longer periods of time (days, months) to detect a gradual shift in vitals of an individual from their healthy or expected baseline to monitor recovery (or relapse) or catch a disorder at its onset. While the short term use cases are almost always applicable to point of care establishments like hospitals, long term monitoring enables collecting health data while an individual is at home, work, travelling and so on.

Continuous monitoring explored so far has primarily been for cardiac activity (achieved using a full-sized ECG machine, a holter machine and the likes of these), body temperature and blood pressure. Respiration rate has been a highly neglected body vital [4], and has been shown to indicate health deterioration in critically ill patients, leading to onset of critical events [5]. In a clinical setting, respiration is measured as air flow using nasal thermistor or by respiratory effort using Respiratory Inductance Plethysmography (RIP) belts for abdomen and/or chest. These are usually costly, require a trained professional to set them up and cause discomfort to the individual, making it infeasible for a regular use over long term. Innovative solutions to overcome these challenges in a non-invasive manner can enable long-term continuous monitoring of respiration in hospitals and at home.

Ballistocardiography (BCG) is one such promising technique. While it was originally developed to measure the micro body motions produced during each cardiac contraction [6], the phenomenon is able to detect any physiological parameter that produces a motion – including breathing, snoring and limb movements [7] and [8]. Moreover it can be made to work in a contact-free and non-wearable manner given a medium that can propagate body vibrations, such as when placed on a solid surface under a mattress, while a subject is lying over it. (For the rest of the paper, we assume this setup.)

However, BCG has its own set of challenges. BCG is prone to undesired noise from the setup environment. This could be mechanical vibrations or movement near and around the setup. Even heavy body movements can overpower the cardiac and respiratory signals resulting in a lower detection rate. Apart from these noises, BCG raw signal is a superimposition of respiratory effort, cardiac contractions and vibrations due to snoring; effective signal conditioning is required to segregate these signals for high fidelity analysis. Additionally, the signal can vary for different people and with different mattresses. It can even vary for the same subject in different postures and body positions.

In this paper, we propose a novel unsupervised clustering algorithm that is effectively able to identify each respiratory cycle from an unconstrained BCG signal. In the next section, we discuss the previous related work. Further in Sect. 3, we describe the algorithm with validation methodology and results in Sect. 4 respectively. Future potential work is highlighted in Sect. 5 with conclusion in Sect. 6.

2 Background

A normal respiration signal comprises alternate inhales and exhales, and is typically a sinusoidal waveform. The same is true for respiration effort as captured in a BCG setup. Figure 1(a) shows raw BCG signal in comparison with respiratory effort captured from RIP belts on chest & abdomen along with airflow signal from the nasal thermistor acquired simultaneously.

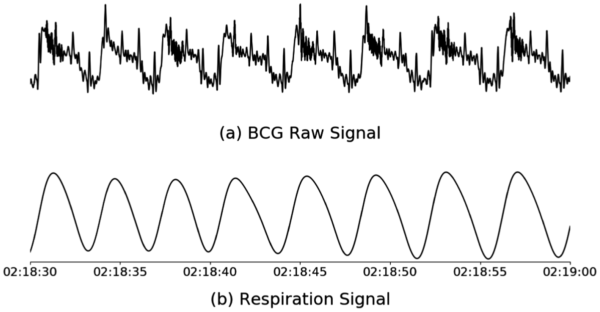


Fig. 1. (a) Raw BCG signal, (b) Respiration signal

Previous studies have attempted to extract respiration rate from a BCG signal. One study [9] describes an algorithm that uses the amplitudes of the filtered BCG signal as the only feature with thresholding to identify inhales and exhales. While this approach works good for normal sinusoidal respiration pattern, it fails to correctly identify respiration cycles in the presence of movements and posture change events that abruptly change the amplitude of the signal. This method is also inconsistent for disorder respiration patterns like hypopnea events, Biot's respiration where the respiration signal has a non uniform amplitude. In another study [10], maximas in the respiration signal, derived after removing the heart signal from the raw BCG signal, were identified as the respiration cycles. This was tested in a controlled environment for only 3 min per subject. The subjects were asked to breathe with an external sound indication for a part of the 3 min. This approach, however, wasn't tested in uncontrolled settings for longer durations such as overnight testing where the subjects are asleep. Another recent study [11] proposes two methods, one to calculate breathing rate using maximas in filtered BCG signal that cross a threshold. However, fixed thresholds don't yield consistent results across various breathing patterns. The other method proposed only calculates an estimate to breathing rate through the most prominent frequency in the fast fourier transform of the filtered BCG signal. The lack of identification of each breath limits this approach for detailed analysis of breathing signal.

3 Proposed Algorithm

In this section we describe our proposed algorithm for identifying respiration cycles in a general BCG signal. The overall flow of the algorithm is shown in Fig. 2.

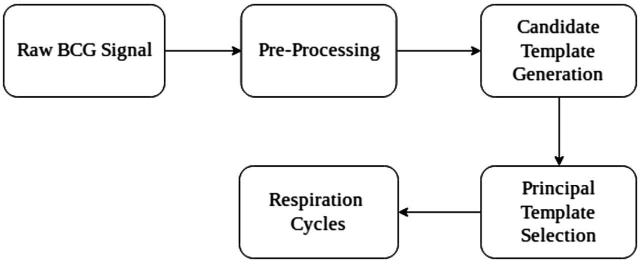


Fig. 2. Proposed algorithm flow

3.1 Pre-processing Raw Signal

The captured BCG signal contains several features in addition to the respiration signal – cardiac contractions, snoring, limb movements, etc. Before identifying respiratory cycles, we pre-process the raw BCG signal to remove such artifacts along with power line noise, if present. The first step of pre-processing involves sifting through the data epoch by epoch, where each epoch is of 30 s duration, for movement artifacts and removing the epochs with movements. Movement artifacts are usually high frequency, high amplitude signals that we identify based on their statistical properties. This results in loss of data for which we do not find respiration cycles, reducing the detection rate; we discuss this further in Sect. 4.4. Non-movement epochs are then passed through a 2nd order 0.1 Hz–0.5 Hz bandpass Bessel filter with a 3 dB cutoff to suppress high frequency components such as cardiac contractions, snoring and 50/60 Hz power line noise. Individual epochs are then further processed separately in the subsequent phases of the algorithm.

3.2 Template Generation

Each pre-processed epoch is used for finding respiration cycles. While using nasal thermistor (for air flow) and RIP belts (for respiratory effort), the respiration cycles are typically identified by finding zero crossings in the acquired signals. However, using the same technique for identifying respiration cycles in the BCG signal does not always yield a high accuracy because of low amplitude secondary mechanical effects and other noises that can still be present after pre-processing the raw signal. This can lead to wrongly identifying additional respiration cycles as shown in Fig. 3(a). This gets even worse if there is disordered breathing, such as apnea events, present in the raw signal – any small non-respiratory signal that crosses zero line also gets counted as a respiration cycle (see Fig. 3(b)).

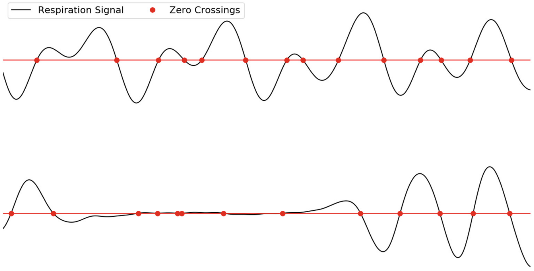
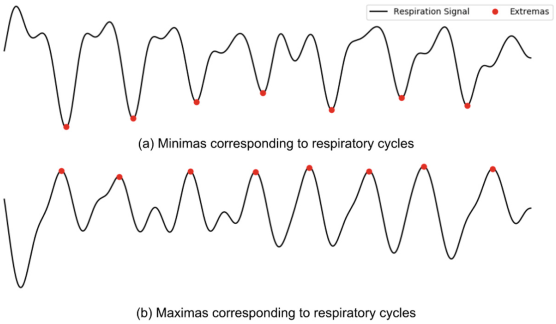


Fig. 3. Pre-processed BCG signal with zero crossings

Instead, our algorithm focuses on local extremas in the signal. To begin with, we identify all the local extremas in the signal. Templates are then generated by clipping the pre-processed signal by 0.5 s on the either side of each local extrema in the signal. If two templates overlap, we remove the (chronologically) later template.



(a) Minimas corresponding to respiratory cycles

(b) Maximas corresponding to respiratory cycles

Fig. 4. Local extremas in pre-processed BCG signal that correspond to actual respiration cycles

There are several extremas in the pre-processed signal and not all of them represent actual respiration cycles. Moreover, the extremas corresponding to respiration cycles can sometimes be easily observed in the maximas of pre-processed signal and in the minimas other times. Figure 4 illustrates an example of each of this case.

3.3 Principal Template Selection

Our goal is to identify extremas that correspond to actual respiration cycles. First, we segregate all the identified templates from the previous step into groups of similar patterns. For this, we use k-means++ [12], a modified form of k-means clustering [13] which augments k-means with a simple, randomized seeding technique. From various iterations, 3 was determined to be the optimum number of clusters for this process. Figure 5 shows the 3 clusters obtained for a pre-processed epoch. If the absolute amplitude of any of the extrema is greater than 15 times the absolute amplitude of either of its neighbors extremas it is dropped before clustering.

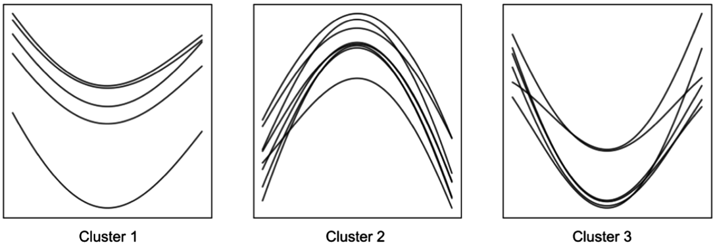


Fig. 5. k-means++ clustering of extremas in a pre-processed epoch

Each cluster now has a set of templates that are similar to each other, but in only one of the clusters each template corresponds to a respiration cycle. To identify this *principal* template, we compute the euclidean distance between all the three centroid templates for each of the clusters. The cluster farthest from the other two clusters is selected as the one to extract respiration cycles. Figure 6 shows the pre-processed signal from Fig. 4(b), along with its extremas clustered into 3 clusters. Based on the euclidean distance, cluster 2 is farthest from the other two and is selected for identifying 8 respiration cycles in the signal.

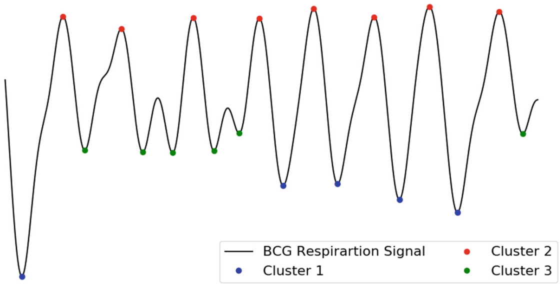


Fig. 6. Clustered extremas

4 Evaluation

4.1 BCG Data Acquisition

We used a commercial non-contact sleep activity and body vitals monitoring device – Dozee, for acquiring BCG data. The system comprises a mesh of Polyvinylidene-fluoride (PVDF) vibroacoustic sensors placed under the mattress to capture micro- and macro-vibrations produced by the body when an individual is lying over it. The sensor array is connected to a data acquisition unit sampling data at a rate of 228 Hz. Figure 7 shows the typical setup of the device in use.

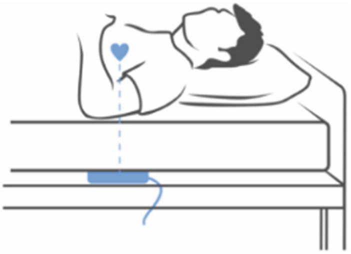


Fig. 7. BCG data acquisition setup

4.2 Validation Data Acquisition

Respiration rate was computed from three different sensors – (a) nasal thermistor for monitoring nasal airflow, (b) RIP belt around the chest and (c) RIP belt around abdomen for respiratory effort. This data was acquired during overnight polysomnography (PSG) recordings conducted at the Human Sleep Research Laboratory, Department of Neurophysiology at National Institute of Mental Health and Neuro Sciences (NIMHANS), Bangalore, India. All the recordings were done using Nihon Kohden [14] Neurofax EEG-1200 machine (24-bit resolution, 1024 Hz sampling rate and 0.1–250 Hz bandpass filter) with 24 electrodes (19-EEG: Electroencephalography electrodes as per Jasper's 10–20 [15] system, 2-EOG: Electrooculography and 3-EMG: Electromyography electrodes as per the guidelines of American Academy of Sleep Medicine [16] and ECG was recorded using bipolar ECG leads.

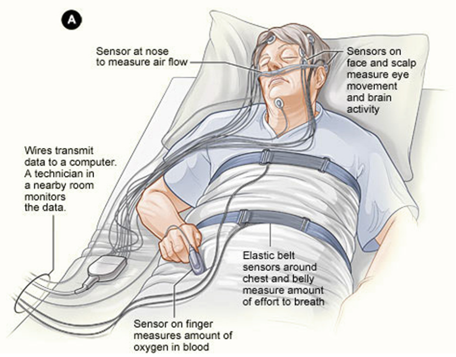

A
Sensor at nose to measure air flow
Sensors on face and scalp measure eye movement and brain activity
Wires transmit data to a computer. A technician in a nearby room monitors the data.
Elastic belt sensors around chest and belly measure amount of effort to breath
Sensor on finger measures amount of oxygen in blood

Fig. 8. Standard Polysomnography Setup [source: https://www.sleep-apnea-guide.com]

The impedance for all the electrodes was kept below 5 KΩ. Finger-pulse oximetry was used to record arterial oxygen saturation. Three way thermocouple transducer (EB Neuro, Italy) was used to record the airflow difference between inhalation and exhalation produced by the subject during sleep. To record subject's thoracic and abdominal movements, one thoracic belt and one abdominal belt (piezoelectric sensors) were placed at the respective positions, i.e. on chest and abdomen. A Snoring sensor recorded the subject's snoring vibrations during sleep with band pass filter of 10–120 Hz. Figure 8 illustrates the typical setup for a PSG recording.

4.3 Validation Methodology

The algorithm was validated for 13 full-night PSG recordings on 10 subjects (for a total of almost 11508 epochs) – details are shown in Table 1.

Table 1. Details of the subjects in the study

Subject	Age	Gender	No. of recordings
Subject 1 (S1)	25	M	2
Subject 2 (S2)	24	M	1
Subject 3 (S3)	58	M	3
Subject 4 (S4)	33	M	1
Subject 5 (S5)	24	F	1
Subject 6 (S6)	25	M	1
Subject 7 (S7)	26	M	1
Subject 8 (S8)	28	F	1
Subject 9 (S9)	27	M	1
Subject 10 (S10)	24	M	1

All the recordings were conducted as described in Sect. 3.2 and scored by an AASM certified expert. During each PSG recording, the BCG data acquisition module was also placed under the mattress. The validation data mentioned in Sect. 4.2 was extracted from the PSG analysis. The raw BCG data was processed through our proposed algorithm to identify each respiration cycle and respiration rate for each 30 s epoch, as described in Sect. 3. Respiration rate for each epoch as computed by the proposed algorithm was then compared against the same extracted from each of the nasal airflow signal, chest respiratory effort and abdomen respiratory effort. The nasal thermistor to monitor airflow is placed inside the nostrils of the subject and is susceptible to getting dislodged while the subject is asleep. The sleep expert conducting the study manually identified such instances and removed the faulty data for them. 732 epochs (6.3% of total data) of nasal thermistor data was dropped. Similarly, incorrect data for the respiratory efforts from both chest RIP belt and abdomen RIP belt was dropped when these belts got loose and didn't have enough tension to monitor the respiratory effort. 866 epochs (7.5% of total data) of respiratory effort data from chest RIP was dropped. Compared to these, in the pre-processing stage of the proposed algorithm, after pre-processing, 3130 epochs of BCG data (27.2% of total data) was not usable due to movements.

The accuracy was measured using the following formula:

$$\epsilon = \sum\nolimits_{i=0}^{n} \left(|\mathrm{BR}_{\mathrm{VAL}i} - \mathrm{BR}_{\mathrm{BCG}i}|\right) \div \mathrm{BR}_{\mathrm{VAL}i}$$

$$Accuracy(\%) = (1 - \epsilon) \times 100$$

where ϵ is error, n is number of epochs, $\mathrm{BR}_{\mathrm{VAL}}$ is breathing rate from validation sensors and $\mathrm{BR}_{\mathrm{BCG}}$ is breathing rate from BCG sensors using the presented algorithm.

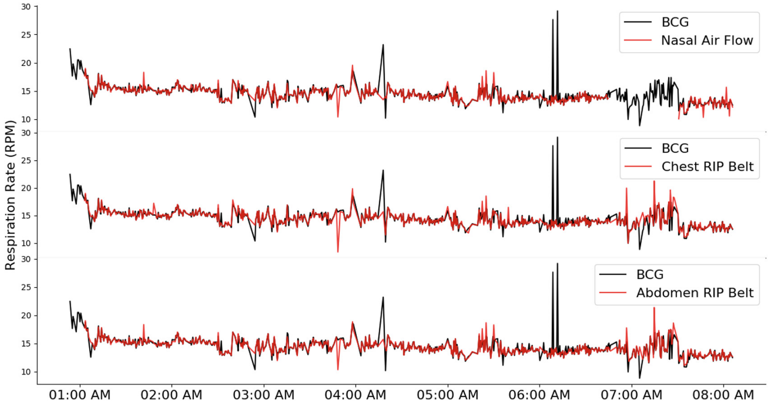


Fig. 9. Whole night respiration rate comparison for Subject 2

4.4 Results

The detection rate for the proposed algorithm varied from 47.4% to 87.7%, with an average of 72.8% for 11508 epochs of BCG data across 13 full night recordings. The detection rate for all the subjects was above the average barring one subject (Subject 3) who was extremely restless throughout the 3 recordings (detection rates of 47.4%, 50% and 53.8%). Without the 3 recordings for Subject 3, the average detection rate increases to 79.3% over roughly 8640 epochs of BCG data across 10 recordings.

Figure 9 shows the respiration rate computed from the proposed algorithm for the whole night overlaid with each of the three validation sensors – nasal thermistor, chest and abdomen RIP belts. For this recording, our algorithm achieved an accuracy of 97.89%, 97.75% and 97.94% from the BCG data when compared to nasal thermistor, chest RIP and abdomen RIP belts respectively with a detection rate of 80.4%.

Table 2 shows the recording duration, detection rate and accuracy of the proposed algorithm in comparison to all the 3 validation sensors. We are able to achieve an average accuracy of 95.48% as compared to the respiration rate from nasal airflow, 94.98% compared to respiratory effort from the chest RIP and 94.88% in comparison to the respiratory effort from the abdomen RIP.

Table 2. Overall results for the proposed algorithm

Subject	Time (h)	Detection rate (%)	Accuracy w.r.t. air flow (%)	Accuracy w.r.t. chest RIP belt (%)	Accuracy w.r.t. Abdomen RIP belt (%)
S1	7.2	74.2	93.89	93.62	93.41
S1	7.1	78.3	96.04	96.12	95.84
S2	7.3	80.4	97.89	97.75	97.94
S3	7.7	47.4	NA	93.43	93.5
S3	9	50.0	NA	92.71	93.28
S3	8	53.8	NA	93.96	94.18
S4	8.3	75.3	96.47	95.86	96.18
S5	8.8	79.4	96.21	95.53	95.4
S6	6.2	80.7	95.92	95.75	95.84
S7	6.8	75.9	90.75	93.39	90.2
S8	7.3	87.7	97.02	96.91	96.75
S9	6.5	82.1	96.34	96.43	96.52
S10	5.7	81.2	94.31	93.34	94.34
Total/Avg	**95.9**	**72.8**	**95.48**	**94.98**	**94.88**

Further, when the respiration rate was computed with zero crossings in the BCG signal over the same dataset, the accuracy was 87.9% when compared to the respiration rate from chest respiratory effort signal, 87.68% when compared to respiration rate obtained from abdomen respiratory effort signal and 88.14% when compared to respiration rate obtained from nasal airflow signal.

5 Future Work

The ease of use of a BCG-based contact-free respiration monitor allows it to be used for long-term data acquisition with critical applications such as enabling continuous monitoring as well as early detection of respiratory disorders and respiratory distress in cardiac patients. The algorithm presented in this paper also needs to be validated on a bigger and diverse group of subjects to study variability with age and gender. Further work and studies can help build systems for detecting apnea and hypopnea events, as well as other breathing abnormalities like Cheyne-Stokes, Biot's, etc.

In fact, Subject 4 was diagnosed with moderate sleep apnea based on the recording conducted in this study with an AHI of 25/h. The proposed algorithm was able to correctly identify a reduced respiration rate during apnea episodes and can be further enhanced to calculate AHI (see Fig. 10).

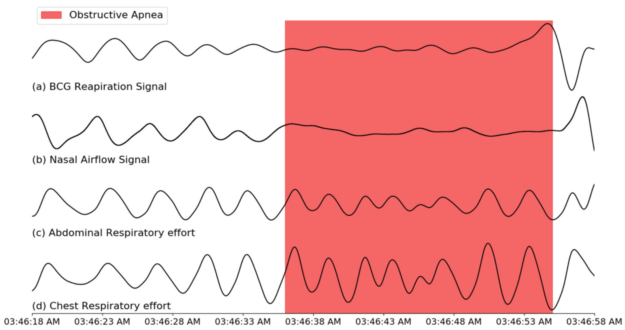


Fig. 10. Apnea detection using BCG and proposed algorithm

6 Conclusions

A novel algorithm to identify respiration cycles and respiration rate in long-term BCG recordings with a high accuracy was presented. To overcome several challenges posed by the variable nature and quality of BCG data, the algorithm used unsupervised clustering methods focussed on the shapes of localized extremas in the respiration signal. Over a cumulative duration of about 96 h, the proposed algorithm was able to achieve an accuracy of 95.48% for respiration rate for 30 s epochs when compared to the respiration obtained from nasal airflow signal, 94.98% when compared to the respiration rate obtained from the chest respiratory effort signal and an accuracy of 94.88% for the respiration rate computed from abdominal respiratory effort signal. This accuracy is significantly higher than that obtained using traditional zero crossing method on BCG data to identify respiration cycles. Further the algorithm showed promising initial results in identifying potential apnea and hypopnea episodes.

Acknowledgement. All the validation PSG studies were conducted at the Human Sleep Research Laboratory of Department of Neurophysiology at NIMHANS. We would like to thank Dr. Bindu M Kutty and her team at the lab for help they provided in conducting these studies and acquiring validation data. This work is partly funded under the Biotechnology Innovation Grant Scheme of the Government of India.

References

1. Zimlichman, E., et al.: Early recognition of acutely deteriorating patients in non-intensive care units: assessment of an innovative monitoring technology. J. Hosp. Med. **7**, 628–633 (2012)
2. World Health Organization: Global strategy on human resources for health: workforce 2030 (2016)
3. Churpek, M.M., Yuen, T.C., Park, S.Y., Meltzer, D.O., Hall, J.B., Edelson, D.P.: Derivation of a cardiac arrest prediction model using ward vital signs. Crit. Care Med. **40**, 2102 (2012)
4. Cretikos, M.A., Bellomo, R., Hillman, K., Chen, J., Finfer, S., Flabouris, A.: Respiratory rate: the neglected vital sign. Med. J. Aust. **188**, 657–659 (2008)

5. McBride, J., Knight, D., Piper, J., Smith, G.B.: Long-term effect of introducing an early warning score on respiratory rate charting on general wards. Resuscitation **65**, 41–44 (2005)
6. Gordon, J.W.: Certain molar movements of the human body produced by the circulation of the blood. J. Anat. Physiol. **11**, 533 (1877)
7. Hwang, S.H., et al.: Polyvinylidene fluoride sensor-based method for unconstrained snoring detection. Physiol. Meas. **36**, 1399 (2015)
8. Alihanka, J., Vaahtoranta, K., Saarikivi, I.: A new method for long-term monitoring of the ballistocardiogram, heart rate, and respiration. Am. J. Physiol.-Regul. Integr. Comp. Physiol. **240**, R384–R392 (1981)
9. Mack, D.C., Patrie, J.T., Suratt, P.M., Felder, R.A., Alwan, M.: Development and preliminary validation of heart rate and breathing rate detection using a passive, ballistocardiography-based sleep monitoring system. IEEE Trans. Inf Technol. Biomed. **13**, 111–120 (2009)
10. Wang, X., Jiang, F., Yang, D., Liao, Y.: Estimation of the respiratory component from ballistocardiography signal using adaptive interference cancellation. In: Chinese Control and Decision Conference (2011)
11. Wusk, G., Gabler, H.: Non-invasive detection of respiration and heart rate with a vehicle seat sensor. Sensors **18**, 1463 (2018)
12. Arthur, D., Vassilvitskii, S.: k-means ++: the advantages of careful seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2007 (2007)
13. Lloyd, S.P.: Least squares quantization in PCM. IEEE Trans. Inf. Theor. **It-28**(2), 129–137 (1982)
14. Kohden, N.: Neurofax EEG-1200. https://us.nihonkohden.com/media/1060/eeg-1200-brochure_nmlb-028-g-co-0163.pdf
15. Jasper, H.H.: Report of the committee on methods of clinical examination in electroencephalography. Electroencephalogr. Clin. Neurophysiol. **10**, 370–375 (1958)
16. AASM: The AASM Manual for the Scoring of Sleep and Associated Events. Rules, Terminology and Technical Specifications. Version 2.0 (2012)

Bank Cheque Validation Using Image Processing

Deepak Chaudhary[1], Prateek Agrawal[1,2(✉)], and Vishu Madaan[1]

[1] Lovely Professional University, Phagwara, Punjab, India
deepak3.14159@gmail.com, prateek061186@gmail.com,
vishumadaan123@gmail.com
[2] University of Klagenfurt, Klagenfurt, Austria

Abstract. Bank cheques, as documents issued by banks can be used as a form of bills capable of monetary exchange, allowing a payee a certain sum of money from the account of drawer. However, due to many fraudulent practices and a need of faster cheque clearance, there had been advances in the process of cheque clearance. Consequently, to aid the process of cheque validation this research work focuses on implementing image processing techniques such as OCR, ANN and Deep Learning to extract key parameters essential for cheque validation. These techniques can be used in sequential manner to automate the task of cheque validation. For extracting machine typographic information Optical Character Recognition is used. Whereas, for the handwritten characters we have used CNN trained using MNIST dataset. The accuracy achieved in handwritten character recognition is 99.14%. For testing purposes IDRBT cheque dataset is used comprising cheque leaflets of different banks.

Keywords: Optical Character Recognition (OCR) · Artificial Neural Network (ANN) · Convolutional Neural Network (CNN) · Reserve Bank of India (RBI) · Morphological operations

1 Introduction

Machine assisted simulation for reading scripts written in different forms and fonts has become a major area of research in the past few years, and it is still undergoing tremendous changes due to the introduction of new image processing techniques. In this research work we would be focusing on bank cheque validations capabilities of an automated document processing system. In present decade with the advent of CTS a major factor in the delay of cheque processing has been countered. However, CTS is incomplete in itself, as even after usage of an image of cheque for monetary transaction, all the remaining work of validation and verification has to be done manually, which need human work force and considerable time and most importantly technical expertise which again require skilled work force. The validation of bank cheque involves ensuring that the leaflet itself is a valid bank cheque leaflet issued to the customer from a recognized bank. This research work would be focused on the validation of bank cheque leaflet using the image processing techniques. In a cheque there are certain key parameters essential for transactions, these parameters are as per the

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 148–159, 2019.
https://doi.org/10.1007/978-981-15-0108-1_15

guidelines of the regulatory body i.e. Reserve Bank of India. Conventionally, the physical transfer of cheque used to cause several problems such as: chances of loss of cheque, damaging of cheque leaflet, unwanted delay in transportation. However, in the modern system after the implementation of CTS (Cheque Truncation system) only an image in the prescribed format and resolution as per the guidelines, needs to be sent and the image itself will be treated as a cheque. Again as the medium of monetary exchange has become digital thus, it demands more fervor in the validation and verification methods which are employed for the funds transfer. The feasibility of such techniques depend on the factor of space time complexity, provided, it satisfies the primary target of matching sample with standard image. As there is less chances of variation in case of machine typographic characters. Thus, Optical Character Recognition technique is employed in order to perform the task of recognition. Whereas, in case of handwritten characters we would be using Deep learning based Convolutional Neural Network due to nuances of variation in script writing which varies from person to person.

2 Literature Review

For efficient image processing operations we need efficient segmentation. In [1, 2] emphasis is laid on surveying and getting high efficiency for segmentation of images. On of such techniques used is Graph distance theory for primary segmentation and SVM for the task of classification. The recognition rate achieved is 99.84%. Whereas, in [3, 4] the segmentation techniques were used with wider implement-ability especially in medical scenario, employing techniques such as grab cut segmentation for desired outcome. Medical imaging is one of the most benefitted domains from the advances in image processing techniques. Thus, in [5] the research have proposed a fully automated tomography image segmentation method which using unsupervised approach is able to evaluate information from visual data which is precise and accurate for the medical prognosis, diagnosis and deduction. In [6] a technique is proposed involving characterization of pixels in images to define the similarity relation between them. The results they achieved validate the method to be efficient for image segmentation using texture analysis. In [7] they have proposed a combination of Niblack and Sauvola techniques, and consequently were able to overcome the limitation of both these techniques by determining the mean of the thresholding values of both methods. The primary function of OCR is to extract character from any document and based on the required preprocessing of segmented document we can achieve higher degree of efficiency, in [8] their research work had been able to obtain an accuracy rate of 93% for English characters. They had employed Otsu's algorithm for binarization and Hough transformation for skew detection. Image segmentation and binarization is the standard when it comes extraction of characters from any form of script however, if we have to identify the different colors present in a segmented image then we would have to perform color image segmentation to get in-depth features from an image. In [9] their research proposed a technique for vessel extraction using image processing techniques. They were able to get the required segmented region of colored retinal blood vessels, which can be used for further medical diagnosis or prognosis. OCR have a widespread application and with a well-defined thorough framework it can be used and complement high-cost radar systems as evident from the research in [10] in which they were able to use OCR to propose a simple yet effective system to identify aircrafts using off-the-shelf cameras making use of the registration number printed on the aircrafts. Similarly in [11] a Markov Random Field framework is proposed for searching document images from books, making it a highly desirable feature. The trained model can also be used to determine required queries from books irrespective of the font size and type. Adding to the applications of OCR in [12] an automated system is proposed for the extraction of Invoice information considering the widespread usage of invoicing causing the additional burden on the human workforce. They have employed image processing, template-matching, OCR and information exporting, they have used normalized coefficient correlation matching based on the experimental evaluation. The accuracy acquired was 95%. In [13] optical flow and eigen values for the precise detection of moving objects is proposed. The results obtained in this research work was satisfactory and as per required standards. Also, whereas, in [14] they have used SIFT and Random Sample Consensus methods for optimizing fused images. The images were fused using discrete wavelet transform. The proposed method was able to generate a robust image which is independent of scale, orientation of camera and light intensity as well.

Image segmentation can be fruitful in the medical diagnosis allowing enhanced efficient disease detection an attempt for early cancer detection. In [15] they performed denoising of the image using wavelet transform and performed the analysis on inverse transformed image they have used Back propagation Neural Network for the training using the images as input. Th results obtained were with an accuracy of 89%. In case of noise present in an image it becomes quite difficult to identify and recognize an image therefore in [16] they have used Harris Corner Detector in order to obtain Interest point to be matched using Fast Retina Key-point technique. The proposed method generated result which was precise and robust.

3 Methodology

In this section, a method based on template matching is put forward in order to identify the key parameters. The flowchart for the entire process is represented in the Fig. 1.

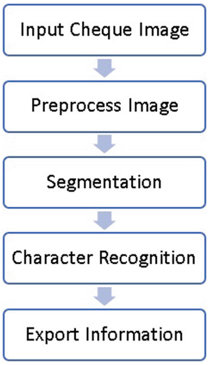


Fig. 1. Flow chart for validation process

3.1 Image Acquisition

The Cheque Image can be acquired using a flat-bed scanner in a standard format.

3.2 Image Pre-processing

In order to ensure that the image is ready for segmentation we need to rotate the image to make it parallel to the set horizontal axis so as to allow correct segmentation of image.

3.2.1 Rotation

If there is any slight variation in the orientation of the image. We need to rotate it. Therefore, in order to rotate two features are necessary: point of rotation and angle of rotation. The point of rotation is set to the mid-point of the image. And the angle of rotation is considered with the location information of the rectangular box of date which is common and standard feature present in all of bank cheques. We need to determine the pixel values of rectangular date box and label each vertices as A, B, C, D. If $A_x > B_x$ we need to rotate clockwise otherwise if, $A_x < B_x$, then we need to rotate counterclockwise as represented in the Fig. 2.

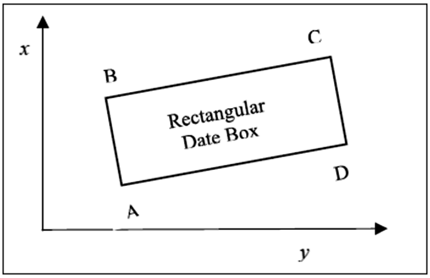
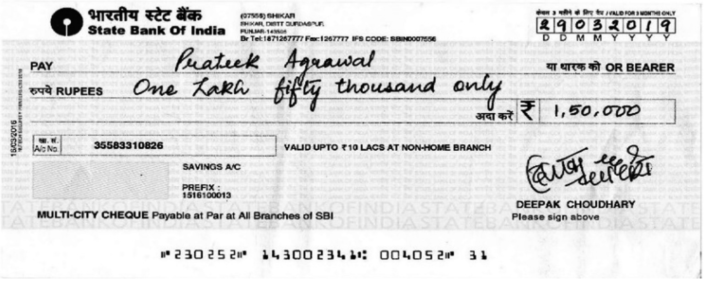


Fig. 2. Rotation of cheque image

3.2.2 Grayscale Operation

The acquired image is with three channels. We will be converting it into one channel. In order to covert a 3-channel image to a 1-channel image, we will use the formula given below as the Eq. (1) and image in Fig. 3.

$$\text{Image} = 0.212671 * \text{R} + 0.71516 * \text{G} + 0.072169 * \text{B} \tag{1}$$



भारतीय स्टेट बैंक
State Bank Of India

(07556) SHIKAR
SHIKAR, DISTT GURDASPUR,
PUNJAB-143505
Br Tel:1871267777 Fax:1267777 IFS CODE: SBIN0007556

केवल 3 महीने के लिए वैध / VALID FOR 3 MONTHS ONLY
2 9 0 3 2 0 1 9
D D M M Y Y Y Y

PAY Prateek Agrawal या धारक को OR BEARER

रुपये RUPEES One Lakh fifty thousand only

अदा करें ₹ 1,50,000

खा. सं. A/c No. 35583310826

VALID UPTO ₹10 LACS AT NON-HOME BRANCH

SAVINGS A/C

PREFIX :
1516100013

16/03/2016

MULTI-CITY CHEQUE Payable at Par at All Branches of SBI

DEEPAK CHOUDHARY
Please sign above

⑈230252⑈ 143002341⑆ 004052⑈ 31

Fig. 3. Grayscale image of bank cheque

3.2.3 Gaussian Filtering

The image obtained after the grayscale operation still contains. We will remove the noise using the gaussian filtering technique. Gaussian filtering technique is a linear smoothening technique. The formula for 2-dimensional gaussian filter is represented in the Eq. (2).

$$\mathrm{G(x,\ y)} = \{\frac{1}{2\pi\sigma^2} * e^{\frac{-x^2+y^2}{2\sigma^2}}\} \tag{2}$$

Here, G(x, y) represents the 2-D normal distribution of the image. Whereas, σ represents the standard deviation of the image distribution.

3.3 Segmentation

When noise is removed the image is segmented into the key components of the image which are required to perform character recognition. The segmented images are shown in Fig. 4 except date all images will be used for OCR.

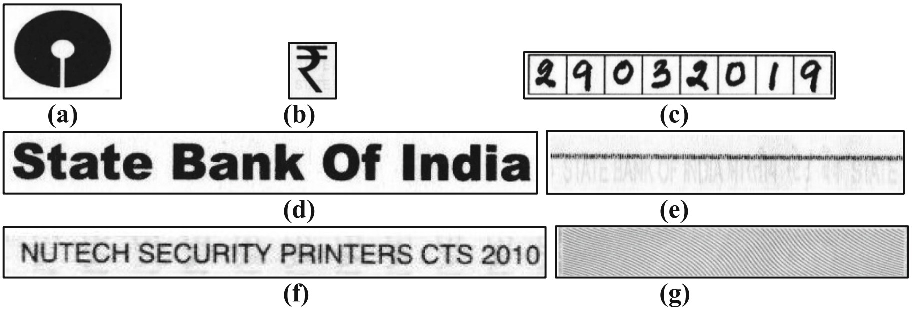


Fig. 4. Segmented Images of a bank cheque. (a) Bank logo. (b) Rupee symbol (c) Date (d) Name of the bank (e) Micro-lettering (f) Printer name. (g) Void Pantograph

3.4 Algorithm for Date Verification

In order to validate whether the cheque is more than three months old we can use the below mentioned algorithm. If it is issued more than 90 days earlier then as per the guidelines of RBI, the cheque will be considered as invalid.

```
      Input: Current Date, Cheque Date
      Output : Date Valid OR Invalid Date
      dCurrent – Current Date
      dCheque – Cheque Date
      function dCorrect(day, month, year)
      Initialise array std[12] = [31, 28, 30…, 31]
      Initialise array Lyear[12]= [31, 29, 30…, 31]
      If year%4 not equal to 0
      { If (day<1 OR day >std[month-1] OR
                  month<1 OR month>12 OR
                  year<0 )
                  return 0
            else
                  return 1 }
            elseif (day<1 OR day >Lyear[month-1] OR
                  month<1 OR month>12 OR
                  year<0)
                  return 0
            else
                  return 1
function totalDays (Date)
      tDays is equal to days passed in Date
      if (dCorrect (date) is equal to 1)
      { If (year%4) is equals to 0
            { for i= 0 to month-1
                  tdays+=(Lyear--)
                  i+=1 }
            else
            {     for i=0 to month-1
                  Tdays+=std[i-1]
                  i+=1 }
            print total days = tdays}
      else
            print Incorrect date
Initialize current days to null
Input dCurrent = days month year
Currentdays =Call totalDays (dCurrent)
Initialize cheque days to null
Input dCheque = ckdays ckmonth ckyear
Chequedays = Call totalDays (dCheque)
      if (ckyear==year)
            if (Currentdays < Chequedays )
                  { if (Chequedays - Currentdays) is less
                  than equals to 90
                        print Date Valid
                  else
                        print Date Invalid}
                        elseif (Currentdays - Chequedays)
                        is less than equals to 90
                              return 1
                              print Date Valid
                        else
                              return 0
                              print Date Invalid}
      else
            print Invalid Date
```

3.5 Algorithm for Cheque Validation

The algorithm mentioned below which considers all the parameters stated in this research to conclude if an image provided is a valid bank cheque or not.

```
Input: Date Validity, Bank Name, Printer Name, Rupee
Symbol, Bank Symbol, Void Pantograph.
Output: Cheque Valid OR Cheque Invalid.
stdBname← Standard Bank Name String
stdPname← Standard CTS 2010 String
stdLogo ← Standard Logo
stdRupee← Standard Rupee Symbol
Date← Validity of Date
Bname← Sample Bank Name String
Pname← CTS 2010 Name String
Logo← Sample Logo image
Rupee← Sample Rupee image
if (Date is valid)
      { if (stdBname==Bname AND
            stdPname==Pname AND )
      else {
            print Invalid cheque }
            if( stdLogo MATCHES Logo AND
                  stdRupee MATCHES Rupee)
                  print Cheque Valid }
print cheque Invalid
```

4 Results and Analysis

Upon performing experimentation using the techniques mentioned above we were able to get desirable results.

4.1 Date

As date is a handwritten parameter for validation of cheque we have used Transfer Learning for digit identification. The segmented digits from date are represented by using the bounding box in Fig. 5.

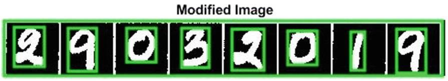
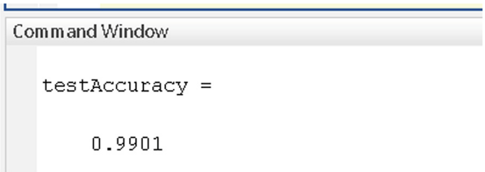


Fig. 5. Segmented digits of date

For testing when we used the MNIST dataset's testing partition, the result obtained is represented in the Fig. 6.



```
testAccuracy =

    0.9901
```

Fig. 6. Testing accuracy of CNN

Upon training the CNN for numeric handwritten digits the accuracy displayed is represented in the Table 1.

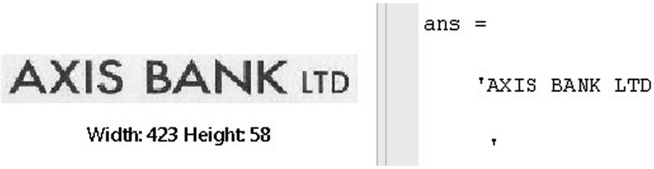
Table 1. Recognition rate for handwritten digits.

Digit types	Samples	Recognition percent
0	50	99.10
1	50	98.98
2	50	99.20
3	50	99.13
4	50	98.96
5	50	99.55
6	50	99.23
7	50	99.11
8	50	99.20
9	50	98.96

The mean recognition rate for digits is 99.14%.

4.2 Bank Name

The technique used for the recognition of the bank name is Optical Character recognition implemented on the MATLAB software, OCR requires that there should be minimal to no noise present on the background. The result of Bank Name recognition is represented in the Fig. 7.

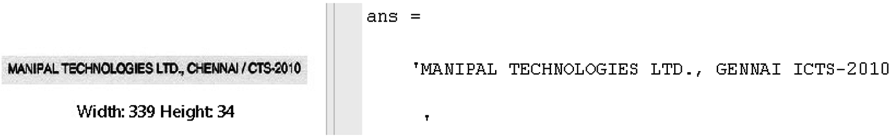


```
ans =

    'AXIS BANK LTD
     '
```

Fig. 7. Bank name recognition using OCR

4.3 Printer Name

This is the name of the printer press where the cheques are printed as per the standard of CTS-2010 and are present on the left side of every standard cheque in vertical format. The results are represented in the Fig. 8.

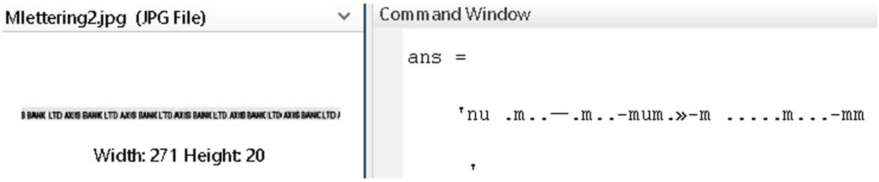


```
ans =

    'MANIPAL TECHNOLOGIES LTD., GENNAI ICTS-2010

     '
```

Fig. 8. Printer name recognition using OCR

4.4 Micro-lettering

As one of the security features to ensure the validity of bank cheques the horizontal lines on the cheque are micro-lettered with the name of the bank. They require either magnification or UV-lamps capable enough to ensure their readability. Using OCR we were not able to get desirable or satisfactory results as it was not able to detect any letters whatsoever present in the micro-lettering image, the result obtained is represented in Fig. 9.


Mlettering2.jpg (JPG File)
S BANK LTD AXIS BANK LTD AXIS BANK LTD AXIS BANK LTD AXIS BANK LTD AXIS BANK LTD A
Width: 271 Height 20
Command Window

```
ans =

    'nu .m..—.m..-mum.»-m .....m...-mm

     '
```

Fig. 9. Micro-lettering recognition using OCR

4.5 Void Pantograph

In the early years of validation measures, void pantograph was a reliable parameter however, as the printers were not fine-tuned and they behaved like Low-pass filters which used to make the Void written in the pantograph to become visible However, with the advent of modern printing like LaserJet printers, it has become possible to obtain exact copy of pantograph without alarming the security measure or making it visible in the photocopy (Fig. 10).

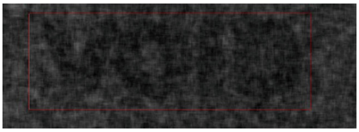


Fig. 10. Void Pantograph on using Low-pass filters.

4.6 Bank Logo and Rupee Symbol

After performing the denoising of the images of Logo and rupee symbol, we use the local feature detection and extraction for feature matching using the SURF and BRISK algorithms. The entire process is represented in Figs. 11 and 12.

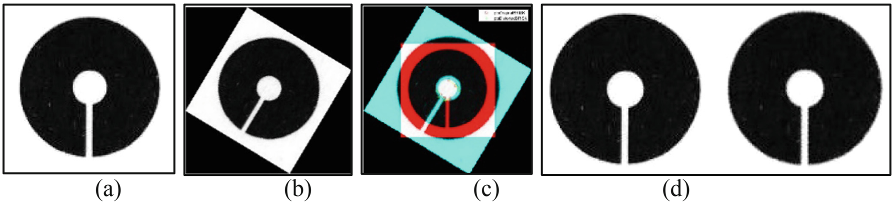


Fig. 11. Bank logo matching using SURF and BRISK algorithm. (a) Original bank logo. (b) Distorted bank logo image. (c) Matching points between original and distorted image. (d) Images after matching between original and distorted image.

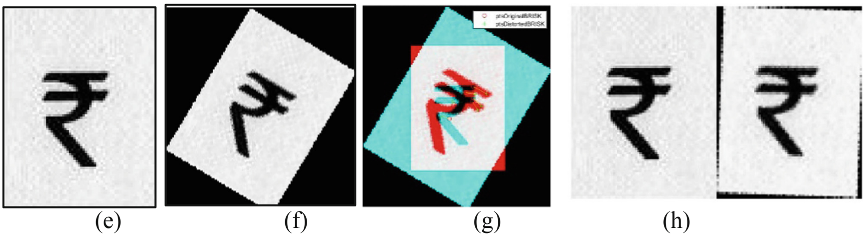
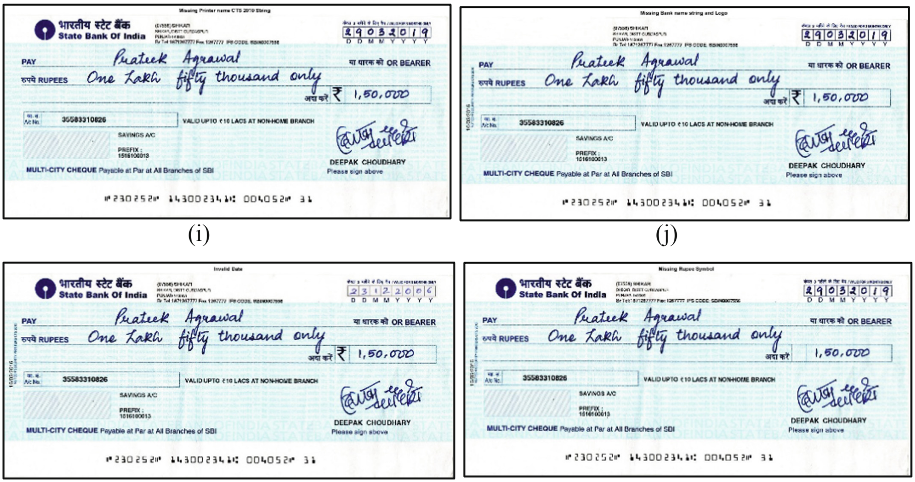


Fig. 12. Rupee symbol images using SURF and BRISK matching algorithms. (e) Original rupee symbol Image. (f) Distorted rupee symbol image. (g) Matching points between original and distorted image. (h) Images after matching between original and distorted images.

As we have performed experiments on all segments of cheque samples images, we can compare a sample cheque to ensure that if it is valid or invalid, the entire process is represented using the cheque images obtained after performing necessary steps. The different types of invalidities possible in a cheque so as to be rendered as an invalid are represented using images in Fig. 13.

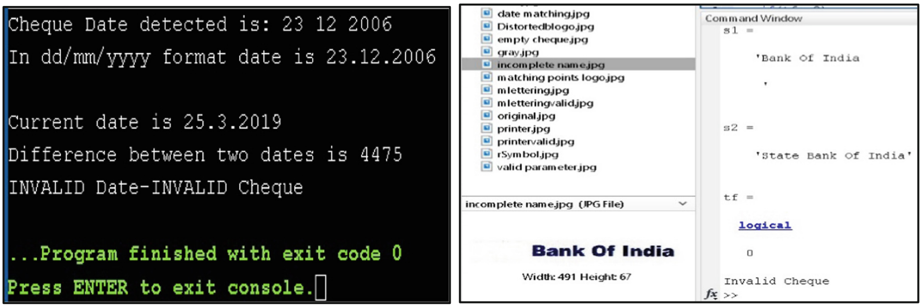
Thus, we can use OCR for the detection of machine typographic characters to a significant degree of accuracy if there is little to no noise present on the image. Also Void Pantographs are not a reliable parameter of validation due to the reason as the general purpose printers nowadays are capable enough to fake it. Therefore it is not used in this research on grounds of reliability (Fig. 14).



(i) (j)

(k) (l)

Fig. 13. Forms of invalid bank cheques for same bank cheque Leaflet. (i) Missing printer name CTS-2010 string. (j) Missing bank name. (k) Invalid date. (l) Rupee symbol missing.


```
Cheque Date detected is: 23 12 2006
In dd/mm/yyyy format date is 23.12.2006

Current date is 25.3.2019
Difference between two dates is 4475
INVALID Date-INVALID Cheque

...Program finished with exit code 0
Press ENTER to exit console.
```

```
s1 =

    'Bank Of India
     '


s2 =

    'State Bank Of India'


tf =

  logical

   0

Invalid Cheque
fx >>
```

(m) (n)

Fig. 14. Cheque validation using invalid cheques. (m) Validation based on backdated cheque. (n) Validation based on fake bank named cheque.

5 Conclusion and Future Work

In this research work we have performed operations on various validation parameters present on a bank cheque leaflet. After the parameters were segmented from the leaflet, pre-processing steps were performed on them based on the techniques which were used for task of recognition and matching. For OCR techniques images were converted into grayscale format and were denoised using the gaussian denoising technique. Whereas, for the task of handwritten digit recognition images were segmented separately. The efficiency obtained was satisfactory. For the task of logo matching, in case of Bank Logo and Rupee Symbol we used SURF and BRISK matching algorithms. The matching was desirable and highly satisfactory as evident from the results obtained. For

Future works we can use UV-lamps for identification of symbols on cheque to achieve higher rate of accuracy as the micro-lettering and Bank Logo present on cheques are UV sensitive.

References

1. Sabu, A.M., Das, A.N.: A survey on various optical character recognition techniques. In: International Conference on Emerging Devices and Smart Systems, pp. 152–155 (2018)
2. Sahare, P., Dhok, S.B.: Multilingual character segmentation and recognition schemes for Indian document images. IEEE Access **7**, 10603–10607 (2018)
3. Li, Y., Zhang, J., Gao, P., Jiang, L., Chen, M.: Grab cut image segmentation on image region. In: 3rd International Conference on Image, Vision and Computing, pp. 311–315 (2018)
4. Lee, S.H., Yang, C.S., Hou, T.W., Yeh, C.H.: An image preprocessing method for fingernail segmentation in microscopy image. In: 2nd International Conference on Signal and Image Processing, pp. 489–493 (2017)
5. Dorgham, O.M.: Automatic body segmentation from computed tomography image. In: 3rd International Conference on Advanced Technologies, pp. 1–5 (2017)
6. Brzoza, A., Muszynski, G.: An approach to image segmentation based on shortest paths. In: International Conference on Systems, Signals and Image Processing, pp. 1–5 (2017)
7. Saddami, K., Afrah, P., Mutiawani, V., Arnia, F.: A new adaptive thresholding technique for binarizing ancient document. In: Indonesian Association for Pattern Recognition International Conference, pp. 57–61 (2017)
8. Agrawal, N., Kaur, A.: An algorithmic approach for text recognition from printed and typed text images. In: 8th International Conference on Cloud Computing, Data Science & Engineering, pp. 876–879 (2018)
9. Latha, M.A., Evangeline, N.C., SankaraNarayanan, S.: Colour image segmentation of fundus blood vessels for the detection of hypertensive retinopathy. In: 4th International Conference on Biosignals, Images and Instrumentation, pp. 206–212 (2018)
10. Vidakis, D.G., Kosmopolous, D.I.: Facilitation of air traffic control via optical character recognition-based aircraft registration number extraction. Inst. Eng. Technol. J. **12**, 965–975 (2018)
11. Yelniz, I.Z., Manmatha, R.: Dependence models for searching text in document images. IEEE Trans. Pattern Anal. Mach. Intell. **41**, 49–63 (2018)
12. Sun, Y., Mao, X., Hong, S., Xu, W., Gui, G.: Template matching-based method for intelligent invoice information identification. IEEE Access **7**, 28392–28401 (2019)
13. Agrawal, P., Kaur, R., Madaan, V., Mukkelli, S.B., Sethi, D.: Moving object detection and recognition using optical flow and eigen face using low resolution video. In: Recent Patents on Computer Science, pp. 1–8. Bentham Science Publisher (2018)
14. Kaur, G., Agrawal, P.: Optimisation of image fusion using feature matching based on SIFT and RANSAC. Indian J. Sci. Technol. **9**, 1–7 (2016)
15. Monica, Singh, S.K., Agrawal, P., Madaan, V.: Breast cancer diagnosis using digital image segmentation techniques. Indian J. Sci. Technol. **9**, 1–5 (2016)
16. Ben-Musa, A.S., Singh, S.K., Agrawal, P.: Object detection and recognition in cluttered scene using Harris corner detection. In: International Conference on Control, Instrumentation, Communication & Computational Technologies (2014)

Diagnosis of Arthritis Using K-Nearest Neighbor Approach

Rupinder Kaur[1], Vishu Madaan[1](✉), and Prateek Agrawal[1,2](✉)

[1] Lovely Professional University, Phagwara, Punjab, India
rjather85@yahoo.com, vishumadaan@gmail.com,
prateek061186@gmail.com
[2] University of Klagenfurt, Klagenfurt, Austria

Abstract. Disease means ‘shortage of comfort’ or ‘lack of ease’. Disease is also known as illness, sickness, rot, affliction or complaint. Disease is caused by tiny living creatures known as microbes. Pathology is the study of diseases and medical experts means doctors study diseases everyday by using wonderful instrument i.e. microscope which is used to make things look bigger. This paper gives a deep cognition about the diseases, their reasons and the Arthritis. It also provides us a sound noesis about the symptoms of the Arthritis. Basal dendrites referring the receptive from where it’s getting the signals. Let’s suppose the response of the inputs is not that which is desired. So, if the actual response is different to desired. Naturally, we must adjust the internal parameters of nerve cells. This paper gives a novel technique for prediction of arthritis on the bases of the arthritis dataset.

Keywords: Disease · Arthritis · Symptoms · Diagnose · Prediction · kNN · Classifier

1 Introduction

Disease means “shortage of comfort” or” lack of ease”. The disease is also known as illness, sickness, complaint, monbus, rot, affiction or complaint. Basically, a distortion in humans, plants or creatures is termed as a disease. Disease is a strange situation or phase that adversely strikes or bears upon a person or a group of people [23, 24]. A disease is one that yields specific symptoms and attacks at a specific location. In other words, it is said that a disease is an uneasiness, trouble or unusual feeling that negatively determine the normal functionality of the body and causes discomfort [1, 17]. Diseases are rapidly growing day by day due to the unhealthy lifestyle of humans or by some other factors. Nobody can run away from disease because the disease can be anywhere like in the area a person breathes, in the ground a person walks on and disease can be in water also. Disease is a worse enemy of everyone [19, 20]. Disease is caused by tiny living creatures known as microbes. Microbes are too small to see means microbes are invisible to the human eye. Pathology is the study of diseases and medical experts means doctors study diseases everyday by using a wonderful instrument i.e. microscope which is used to make things look bigger. Microscope consist of magnifying glasses so that one could clearly see the creatures [2, 16].

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 160–171, 2019.
https://doi.org/10.1007/978-981-15-0108-1_16

1.1 Arthritis

‘Arth’ means joint and ‘itis’ means inflammation and when combine both the terms, it becomes arthritis which means inflammation of joints. Inflammation is a combination of pain, redness, swelling and stiffness. Whenever viruses, bacteria, fungus, etc. enters any part of body or attacks on specific area. The area where they attack the result is swelling, pain, and stiffness. There are more than 100 types of arthritis. The risk factors of arthritis are older age, sex, genetics, obesity, joint injury, certain occupations and bone deformity. Although all the arthritis can present with similar symptoms, the reason can be different [3].

The pain which is concerned with inactivity known as Inflammatory arthritis. Inactivity means in this type of arthritis the pain becomes worse with rest means if one doesn’t perform any physical activity the problem increases, and it improves with activity. Rheumatoid Arthritis and Osteo Arthritis are the best example of inflammatory arthritis. In inflammatory pain the morning stiffness duration is at least 35–40 min [4] (Fig. 1).

Fig. 1. Arthritis effect

Infectious arthritis as per the name implies, it is just like an infection in a joint and this type of arthritis is also known as septic arthritis. It is caused by a virus or bacteria means when an infected virus spread around the joint or infected fluid spread surrounding the joint. Infectious arthritis affects one large joint in the human body, usually the hip or knee. Chills, fatigue, fever, inability to move the limb with the infected joint, severe pain in the affected joint and especially with movement, swelling, warmth are the symptoms of infectious arthritis [5].

1.2 Rheumatoid Arthritis

It is a chronic, auto-immune and inflammatory disease in which joint cartilage is damaged and it equally affects the both small and large joints. Joint cartilage is very soft and smooth white tissues (just like a pillow) surroundings of the joints that covers the terminals of the bones where they come together to form a joint [6].

Cause for Rheumatoid Arthritis: Quite frankly, the reason behind RA is not known. It is not pin-point able yet. There is no certain cause for this disease because no idea who will caught by this disease, but yes, sometimes one blames infections, genetics and hormonal changes. Smoking and bad life style are also linked to the disease.

Symptoms of Rheumatoid: Ra affects the joints of fingers, wrist, knee, ankles, hip, shoulders, etc. This disease often begins slowly and difficult to diagnose in early stages. But early symptoms can include inflammation, minor joint pains, stiffness of joints, joint fatigue (Fig. 2).

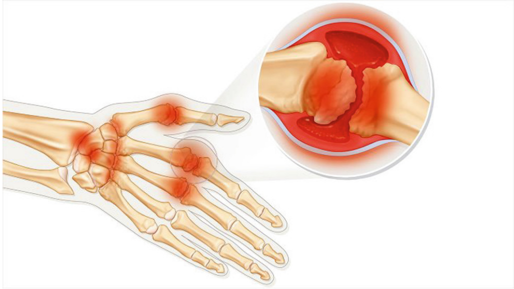


Fig. 2. Rheumatoid Arthritis

Artificial Intelligence: It is the branch of computer science which gives more emphasis to developing automatic as well as automatic machines that think, plan, work and react like humans. In other words, it can be said that artificial intelligence is a broad idea that makes machines same as as human intelligence and it is a field of developing computers or machines which can perform a task requiring human intelligence. [7, 18, 23, 25].

Artificial Neural Networks: These networks derived from human brains or in other words, we can say neural network is inspired by the natural neural network of human nervous system. The inventor of the first Neuro-computer Dr. Robert Hecht Nielson denies that Neural Network is made up of simple but massively interconnected processing parameters. If we talk about the human brain, it is highly complex, nonlinear, massively parallel computer.

Our human brain having billions (approximate 10 billion) of nerve cells with trillion (approximate 60 trillion) of connections. Human brain consists of a cell body, axons, synaptic terminals, basal dendrites, apical dendrites etc. where axons are basically acting as transmission lines that can carry electrical signals. In terms of its electrical signals its having high degree of electrical resistance and large capacitor. Synaptic terminals are basically used for making connections with the other nerve cells. Synaptic inputs receive the signals from other neurons and dictating what is the signal is going to be. The response will be transmitted to other neurons by synaptic terminals. Basal dendrites referring the receptive from where it's getting the signals. Dendrites having more branches than axon, while the link of the axon is larger than dendrites. Now will draw the artificial neural network with the help of biological neurons [8, 9] (Fig. 3).

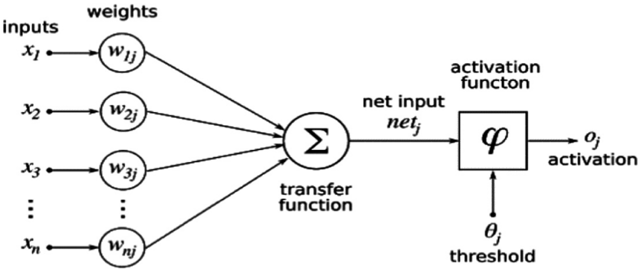


Fig. 3. Model of Artificial Neural Networks

KNN (K Nearest Neighbors): In the past few years of analytics, there can be seen several classifiers and regression models. In this ratio the classification has more models as compared to regression. A widely used classification model is K nearest neighbors (KNN). The reason of being biased towards classification models is These are quite good at making decisions and predicting the outcomes. For diagnosis of the arthritis KNN will play a good role. Which will predict that a person is an arthritis-positive or arthritis-negative. Our focus will be primarily on how the model will work and how the symptoms effect the output/prediction [10, 22].

2 Related Work

Shiezadeh et al. have discussed that arthritis is that chronic disease and its actual causes are not known. For the early diagnosis of arthritis predictive model is proposed. The dataset of arthritis patients are collected from rheumatology clinic which is depend on several features such as elbow and knee joints, gender and ESR test value. After that several classification or decision tree algorithms such that Adaboost and Cuckoo Search applied to these features to check the highest precision. And the new algorithm is also proposed named CS-boost which is the combination of existing algorithms that are adaboost and cuckoo search. Initially adaboost has highest precision but after checking the performance of CS-boost, it is observed that CS-boost enhanced the accuracy of Adaboost in term of diagnosing rheumatoid arthritis [3].

Rathore et al. used the concept of thermography in this paper i.e. one of the technique of analyzing dysfunctions in human body. But this technique had need to improve its performance in case of diagnosing of rheumatoid arthritis and it produce high vulnerability which is responsible for the changes of end results. So, artificial neural network and C-means algorithm is proposed along with thermography for analysis and diagnosis of rheumatoid arthritis [4].

Subramoniam explained that therapeutic approach along with images of affected joints and applied image segmentation algorithm for early diagnosis and to check the severity of the problem. The suspicious joint is observed by radiologist and then apply image segmentation which helps in edge detection and extraction. By using this concept of image processing normal and abnormal joint can be detected [5].

Gobikrishnan et al. uses thermography which one of the best tool in medical field for diagnosis any type of disease. In this paper C-means algorithm is applied and checked the significance of the results based on some parameters. People having features like mean value above 5, standard deviation and kurtosis above 10, skewness below 10 are suffering from rheumatoid arthritis and rest are under control [6].

Majumdar et al. defined that how severity of the rheumatoid arthritis as well as osteoarthritis can determine along with the help of infrared imaging of knee joints. In this paper, after acquisitioning the image of knee joint the feature of the joint get extracted then proposed two algorithms one fuzzy C-means algorithm and second one is region growing segmentation. The result of this proposed method is to detect inflammation of joint [11].

Bhisikar, Kale tried to overcome the headache of radiologist because radiograph analysis of arthritis is very time-consuming process and create accuracy. So, image processing technique is proposed. This paper is the enhanced version of previously discussed paper. The sample of five hand radiograph images are observed for the early examine of the affected joints.

Bhisikar et al. proposed a method which used image processing to detect or measure the space of the hand joints. The proposed algorithm used Gaussian filter for pre-processing of an image. After that otsu's binarization came into existence to separate the background as well as foreground then need to apply morphological filtering to get the skeleton of the binary image (Fig. 4).

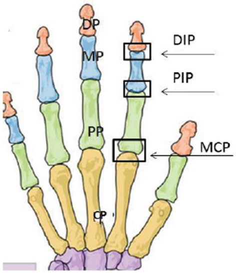


Fig. 4. Joints and bones of a human hand [12]

Gobar filter is applied to extract the minimal space of hand joint automatically [12]. Mean joint accuracy, segmentation accuracy and error are 92%, 80% and 1–3% after performing test on given five images. One of the best part of this proposed model is that the diagnosis of rheumatoid arthritis is totally automatic. It provides good accuracy and consume less time and the severity of the disease can also determine at later stages [21].

Jobin Christ et al. diagnosed two types of arthritis in this paper that is reactive arthritis and septic arthritis using some of the swarm intelligence techniques. For the early diagnosis of these types of arthritis Ant colony optimization, switching optimization and clown fish algorithm is proposed. The main goal behind the early prediction of disease is to reduce the possibility of joint damaging. These meta heuristic algorithms are applied on the hip joint images and compared the result of each proposed algorithm which is used in this paper [13].

Nouri, Amirfattahi et al. proposed nonparametric windows for the early prediction of the arthritis (Fig. 5).

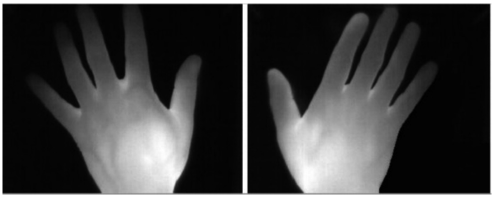


Fig. 5. Bilateral images of human hands [14]

Non-parametric windows processed in such a manner so that it could differentiate a human body which is affected from rheumatoid arthritis hands well as normal human body. In this paper 50 thermography images of hand are collected from rheumatology clinic and estimate the histogram of every image to achieve the mutual distribution of joint [14].

Snekhalatha, Anburajan have used two main concepts in this paper, the first one is watershed algorithm which is used to perform image segmentation on the hand radiograph images of the patients. To classify the images back propagation neural network is used. The proposed method is to extract the feature from grey scale image. This study used the radiograph hand images of ten rheumatoid arthritis patients and five normal persons, then data applied to the neural network which classifies the normal as well as RA patients [15] (Table 1).

Table 1. Review of literature

Year	Author	Title	Techniques
2018	Sumit Sharma et al.	Heart Disease Prediction Using Fuzzy System	Fuzzy system
2017	Entin Martiana et al.	Auto Cropping for Application of Heart Abnormalities Detection Through Iris Based on Mobile Devices	Image Processing
2017	C. Sowmiya et al.	Analytical Study of Heart Disease Diagnosis Using Classification Techniques	SVM, KNN and ANN
2016	kalyani Ohri et al.	Fuzzy Expert System for diagnosis of Breast Cancer	Fuzzy Systems
2016	Hatungimana Gervais et al.	Computer aided screening for Acute Leukemia blood infection using gray level intensity	Image Processing
2016	Ms. Shivani S. Puranik et al.	Morphology Based Approach for Microaneurysm Detection from Retinal Image	Image Processing
2016	Manpreet Singh et al.	Building a Cardiovascular Disease Predictive Model using Structural Equation Model & Fuzzy Cognitive map	Fuzzy Systems
2016	Sakom Mekruksavanich	Medical Expert System Based Ontology for Diabetes Disease Diagnosis	Fuzzy Systems
2016	Dr. R.R. Janghel et al.	Soft Computing Based Expert System for Hepatitis and Liver Disorders	ANN, Soft computing
2016	Gobikrishnan M, et al.	Diagnosis of rheumatoid arthritis in knee using fuzzy c-means segmentation technique	Fuzzy and image processing
2016	Mrs. Swati A. Bhisikar et al.	Automatic joint detection and measurement of joint space width in arthritis	Image processing
2016	Mrs. Swati A. Bhisikar et al.	Automatic analysis of rheumatoid arthritis based on statistical features	Image processing
2016	Ali Nouri	Mutual information-based detection of thermal profile in hand joints of rheumatoid arthritis patients using non-parametric windows	Image processing and non - parametric windows
2016	Meenakshi Sharma et al.	Classification of Uterine Cervical Cancer Histology Image using Active Contour Region based Segmentation	Image processing
2016	Dimple Sethi et al.	X-Tumour: Fuzzy Rule based Medical Expert System to Detect Tumours in Gynecology	Fuzzy system
2016	Ranjit Kaur et al.	Fuzzy Expert System to Calculate the Strength/Immunity of a Human Body	Fuzzy system

(*continued*)

Table 1. (*continued*)

Year	Author	Title	Techniques
2016	Meenakshi Sharma et al.	Classification of Uterine Cervical Cancer Histology Image using Active Contour Region based Segmentation	Image processing
2016	Dimple Sethi et al.	X-Tumour: Fuzzy Rule based Medical Expert System to Detect Tumours in Gynecology	Fuzzy system
2016	Ranjit Kaur et al.	Fuzzy Expert System to Calculate the Strength/Immunity of a Human Body	Fuzzy system
2016	Ranjit Kaur et al.	Fuzzy Expert System for Identifying the Physical Constituents of a Human Body	Fuzzy system
2016	Monica et al.	Breast Cancer Diagnosis using Digital Image Segmentation Techniques	Image processing
2016	Meenakshi Sharma et al.	Classification of Clinical Dataset of Cervical Cancer using KNN	KNN
2016	Dimple Sethi et al.	X-Gyno: Fuzzy Method based Medical Expert System for Gynecology	Fuzzy System
2015	*Subramoniam. M*	A non-invasive method for analysis of arthritis inflammation by using image segmentation algorithm	Image processing
2015	Prateek Agrawal et al.	Fuzzy Rule Based Medical Expert System to Identify the Disorders of Eyes, ENT and Liver	Fuzzy system
2015	Kirandeep Kaur et al.	Classification of Follicular Lymphoma Grades using Support Vector Machine	SVM

3 Methodology

The proposed work in this paper is to collect data from an authorized source and process it to create a dataset for predicting arthritis in humans. The process includes five steps as shown in the Fig. 6.

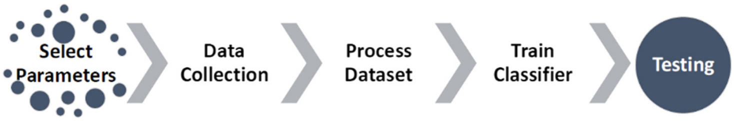


Fig. 6. Block diagram of proposed model

3.1 Selection of Parameters

Every disease has some unique symptoms. On the bases of which the disease can be determined e.g. headache, runny nose and temperature can be symptoms for cold and fever. In this step various doctors are consulted, ansd literatures are reviewed for deciding the symptoms of the Arthritis. After an adequate research and consultation 11 symptoms are determined which will act as parameters in the further processes of this model. These symptoms include gender, Age, Morning Stiffness, Joint Deformity, TLC, ESR etc.

3.2 Data Collection

Once the parametric quantities have been decided then the next footmark is to gather up the information from the real world. This information should be ascertained and examined by an authorized doctor or some Lab specialist. So that, the data should be objurgate. We have collected data on more than 50 affected roles, which include well-nigh every parameter of both male and female patients of all the age groups. This data also admits some Arthritis-positive events and some Arthritis-negative consequences.

3.3 Data Pre-processing

Afterward collecting the data from the empowered laboratory, it should be processed to have the error detached data. While aggregation it's not potential that every individual has gone through all the tests and have data under the heading of every parametric quantity. Which results that the data have some omitting values. To deal with those missing values we have following methods usable and worthy to our trouble: Mean value, Median value and Most commonly used value. We have exploited "Mean value" to occupy the absent values. For that estimate the mean value for Arthritis-cocksure cases and fill it in all the positive cases where measures are missing, same for the Arthritis-veto cases. This operation should be done to every parametric quantity.

3.4 Train Classifier

A dataset creative activity is an anterior foot-mark to the training the classifier. A classifier is used to predict the arthritis in human-being. The aim of this paper is to diagnose Arthritis by using machine learning algorithms. The algorithm put-upon is kNN (k nearest neighbours). KNN is exclusively based on the Euclidean distance. This distance is the length of the straight line between two points which can be calculated by using formula-1.

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2}. \tag{1}$$

where d (p; q) is the Euclidean distance and p; q are two points whose distance is to be calculated. KNN produces two clusters from the training dataset, one for positive results and another for negative results. The training dataset consist of 70% of the global dataset.

3.5 Testing

Once the training process is over, the next step is to test the model with some untrained feature vectors, i.e. those examples of the global dataset which is unknown to the trained model. In this paper 70% of the data is used for training and rest 30% will be used for the testing process. Testing is a very important step because it will tell us that how accurately this model is diagnosing the Arthritis.

4 Results

Total 57 patients' data were collected and processed through this model. For dealing with the missing values mean values were computed separately for the positive cases and negative cases. Then those values were copied to the respective fields. Then a KNN classifier was trained by using MATLAB 2016a.

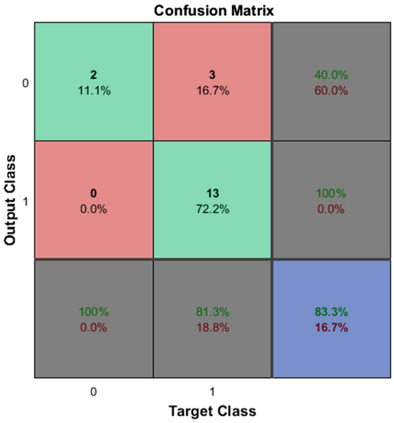


Fig. 7. Confusion Matrix (Color figure online)

In Fig. 7 Green color represents the correct classification, red represents wrong classifications and blue block have the overall accuracy i.e. 83.3%. There are 16.7% cases which are false negative.

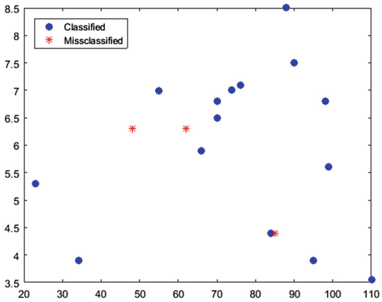


Fig. 8. Clustered feature vectors (Color figure online)

They should be classified to the positive cluster, but they are clustered to the negative cluster. The Fig. 8 shows that how many classes are misclassified and how many are correctly classified. The red color '*' symbol represents misclassification and blue '•' represent classified examples.

5 Conclusion

This study gives deep knowledge about the diseases, their causes and the Arthritis. It also provides us a sound knowledge about the Artificial Intelligence and KNN classifier. Artificial Intelligence comprises of various techniques, i.e. Classification, regression and other Neural networks. This paper gives an overview of these techniques and their working. A complete step by step process is also given in this paper, which will lead us towards a novel technique for the diagnosis of the Arthritis in human beings with accuracy 83.3%. This can be applicable in many areas like in Medical, normal life, etc. where we need to determine that a person is suffering from the problem of Arthritis or not. Arthritis have some specific symptoms like morning stiffness, which means if a person has stiffness in the morning for more than one hour, then this can be counted as a symptom of Arthritis. Likewise, joint pain, Anaemia, joint warmth, high uric acid, etc. are the symptoms of Arthritis. This study involves the methodology by which we can predict that a person is Arthritis-positive or Arthritis-negative.

References

1. Kusumaningtyas, E.M., Barakbah, A.R., Hermawan, A.A., Candra, S.R.: Auto cropping for application of heart abnormalities detection through Iris based on mobile devices. In: 2017 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), pp. 108–113. IEEE (2017)
2. Biyouki, S.A. Turksen, I.B., Zarandi, M.H.F.: Fuzzy rule-based expert system for diagnosis of thyroid disease. In: 2015 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), pp. 1–7. IEEE (2015)
3. Shiezadeh, Z., Sajedi, H., Aflakie, E.: Diagnosis of rheumatoid arthritis using an ensemble learning approach. Comput. Sci. Inf. Technol. (CS & IT) **5**(15), 139–148 (2015)
4. Rathore, S., Bhalerao, S.V.: Implementation of neuro-fuzzy based portable thermographic system for detection of rheumatoid arthritis. In: 2015 Global Conference on Communication Technologies (GCCT), pp. 902–905. IEEE (2015)
5. Subramoniam, M.: A non-invasive method for analysis of arthritis inflammations by using image segmentation algorithm. In: 2015 International Conference on Circuits, Power and Computing Technologies (ICCPCT-2015), pp. 1–4. IEEE (2015)
6. Gobikrishnan, M., Rajalakshmi, T., Snekhalatha, U.: Diagnosis of rheumatoid arthritis in knee using fuzzy C means segmentation technique. In: 2016 International Conference on Communication and Signal Processing (ICCSP), pp. 0430–0433. IEEE (2016)
7. Eiben, A.E., Smith, J.: From evolutionary computation to the evolution of things. Nature **521** (7553), 476 (2015)
8. Engelbrecht, A.P.: Computational Intelligence: An Introduction. Wiley, Hoboken (2007)
9. Kriesel, D.: A brief introduction to neural networks (2007). http://www.dkriesel.com
10. Oliverio, V.: Artificial intelligence applied to aid in the diagnosis of chronic pelvic pain, Doctoral dissertation, University of São Paulo (2018)
11. Majumdar, P., Das, K., Nath, N., Bhowmik, M.K.: Detection of Inflammation from temperature profile using Arthritis knee joint Datasets. In: 2018 IEEE International Conference on Healthcare Informatics (ICHI), pp. 409–411. IEEE (2018)

12. Bhisikar, S.A., Kale, S.N.: Automatic analysis of rheumatoid arthritis based on statistical features. In: 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), pp. 242–245. IEEE (2016)
13. Jobin Christ, M.C., Lakshmi Narayanan, A., Krishnan, R.: Detection of septic arthritis using meta heuristic algorithms, vol. 6, no. 6, pp. 6–7 (2017)
14. Nouri, A., Amirfattahi, R., Moussavi, H.: Mutual information based detection of thermal profile in hand joints of rheumatoid arthritis patients using non-parametric windows. In: 2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1–4. IEEE (2016)
15. Snekhalatha, U., Anburajan, M.: Dual tree wavelet transform based watershed algorithm for image segmentation in hand radiographs of arthritis patients and classification using BPN neural network. In: 2012 World Congress on Information and Communication Technologies, pp. 448–452. IEEE (2012)
16. Sharma, S., Madaan, V., Agrawal, P., Garg, N.K.: Heart disease prediction using fuzzy system. In: Luhach, A.K., Singh, D., Hsiung, P.-A., Hawari, K.B.G., Lingras, P., Singh, P.K. (eds.) ICAICR 2018. CCIS, vol. 955, pp. 424–434. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-3140-4_38
17. Sharma, M., Singh, S.K., Agrawal, P., Madaan, V.: Classification of uterine cervical cancer histology image using active contour region based segmentation. Int. J. Control Theory Appl. **9**(45), 31–40 (2016)
18. Sethi, D., Agrawal, P., Madaan, V.: X-tumour: fuzzy rule based medical expert system to detect tumours in gynecology. Int. J. Control Theory Appl. **9**(11), 5073–5084 (2016)
19. Kaur, R., Madaan, V., Agrawal, P.: Fuzzy expert system to calculate the strength/ immunity of a human body. Indian J. Sci. Technol. **9**(44), 1–8 (2016)
20. Kaur, R., Madaan, V., Agrawal, P., Singh, S.K., Kaur, A.: Fuzzy expert system for identifying the physical constituents of a human body. Indian J. Sci. Technol. **9**(28), 1–8 (2016)
21. Monica, Singh, S.K., Agrawal, P., Madaan, V.: Breast cancer diagnosis using digital image segmentation techniques. Indian J. Sci. Technol. **9**(28), 1–5 (2016)
22. Sharma, M., Singh, S.K., Agrawal, P., Madaan, V.: Classification of clinical dataset of cervical cancer using KNN. Indian J. Sci. Technol. **9**(28), 1–5 (2016)
23. Sethi, D., Agrawal, P., Madaan, V., Singh, S.K.: X-Gyno: fuzzy method based medical expert system for gynecology. Indian J. Sci. Technol. **9**(28), 1–10 (2016)
24. Agrawal, P., Madaan, V., Kumar, V.: Fuzzy rule based medical expert system to identify the disorders of eyes, ENT and liver. Int. J. Adv. Intell. Paradigm (IJAIP) **7**(3–4), 352–367 (2015). InderScience Publications
25. Kaur, K., Singh, S.K., Agrawal, P., Goyal, A.: Classification of follicular lymphoma grades using support vector machine. Int. J. Adv. Eng. Res. (IJAER) **10**(8), 20441–20449 (2015). RIP Publications

A Method for Continuous Clustering and Querying of Moving Objects

A. Nishad[✉] and Sajimon Abraham

Mahatma Gandhi University, Kottayam, India
an.nishad@gmail.com, sajimabraham@rediffmail.com

Abstract. Location sensing moving objects generate a continuous stream of spatio-temporal data. Querying and analysis of this data can give more inference on the mobility patterns of the objects. In this paper, we are proposing a method for evaluating spatio-temporal aggregate queries. For the effective processing of queries continuous clustering is employed on moving objects. The frequency of clustering is determined by the semantic properties of the travel network. Moving objects and queries are combined together and processed incrementally all through the travel network. The cluster membership of an object may change during the course of the journey. By analyzing this, the pattern of the movement of object can be ascertained. Special data structures are maintained to keep track clusters to answer spatio-temporal aggregate queries. We prove that our system can deliver answers to spatio-temporal aggregate queries effectively without processing the entire records of the moving objects.

Keywords: Location based systems · Moving objects · Semantic data processing · Spatio-temporal data mining · Spatio-temporal aggregate queries

1 Introduction

Clustering of moving objects works on two approaches. First, cluster the historical movement data which is generally referred to as trajectory clustering and second approach is based on the concept that, multiple objects move along a constrained path in clusters like locomotion of vehicles in traffic junctions, movement of animals and birds. These moving objects have uniform features such as pattern of movement, change in speed, direction, destination etc. The data of moving objects primarily hold spatial and temporal characteristics. The spatial component represents the latitude and longitude of the object and temporal component represents the time at which the object occupies a space.

The continuous changing nature of the moving objects necessitates special approaches for mining of data on moving objects. In order to deal with the features of moving object clusters such as changing the membership of objects in to

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 172–183, 2019.
https://doi.org/10.1007/978-981-15-0108-1_17

different clusters during its life time, different concepts are discussed in literature. Some of them are cluster block, micro clustering [5], incremental clustering [6] etc. Later new paradigm called data stream processing has evolved as an extension of moving object data management systems [13]. Data Streams are continually flooded data generated from various sources. Due to the limitations of system resources in terms of space and time, it is tedious to process and entire data. In order to extract information from this unbounded flow of data continuous queries are used. Unlike one time queries which evaluates the expression once over a point of time, continuous queries generates a stream of answers that reflect data received over a course of time.

Apart from the explicit spatio-temporal data, moving objects hold lots of contextual information that will give comprehensive outlook on the semantic properties of the objects. These groups of implicit information comprise of the direction of movement, the velocity of the object, stop points etc. and these are also the factors that link the meaning and purpose of the journey. Moving object data are conveniently represented in the form of time stamped locations called trajectories. Enhancing these trajectories with contextual semantic attributes like*name of tourist spot, name and facilities of restaurants and theaters etc.* are examples of semantic enrichment. Recently, semantic enrichment of trajectories are gaining more research attention as it simplifies querying and analysis of spatio-temporal data.

As the number of objects being tracked and frequency of sampling rates increases the number of GPS records also increases. Heavy computational resources will be required to process such data. At this juncture clusters of mobility data can be used for efficient and scalable processing of moving object queries. Moving object queries are SQL like queries that are required to fulfill spatio-temporal attributes. Sequential scanning of entire records to answer such type of queries are not a cost effective. One suitable method is to consider both the queries and moving object clusters as a single unit for processing. In this paper we are following the SCUBA method suggested by Nehme et al. in [7].

The rest of the paper is arranged as follows: in Sect. 2 we narrate some of the studies in dynamic clustering and querying of moving objects. In Sect. 2, important concepts related to our work are explained. Section 3 provides the algorithm. In Sect. 4 we outline experimental evaluation. We conclude the topic by giving some future scopes in Sect. 5.

2 Related Works

2.1 Clustering of Continuous Data

Clustering of moving objects requires exceptional approaches due to its dynamic nature. A challenge in clustering moving objects is the difficulty in managing two dimensional geographical information with the time domain. One of the well-known scheme for moving objects is incremental clustering. According to this location are clustered upon the updation from different moving objects without re-clustering the entire space time components. Incremental clustering avoids the necessity of preserving all location updates, due to this storage requirement can be reduced. Several constraints arises at this stage, especially associated with the cluster centroid, speed of movement and termination condition etc. Li et al. [6] proposes an incremental clustering framework for a continuously updating system called TCMM which is based on the assumption that new locations will not influence the clusters distant from the incoming data. It executes the operation in two steps. Initially the micro clustering phase is used to generate representative trajectory line segments by comparing an incoming trajectory with the exiting clusters. Similar micro clusters are merged together in periodical fashion, which will avoid unnecessary maintenance of micro clusters.

SCUBA [7] is a framework that adopts the idea of grouping moving objects and queries in a single unit according to its spatial temporal features. The method utilizes concept of moving micro clusters for the effective and less expensive evaluation of spatio-temporal queries. The clusters are updated for its membership incrementally in every time units. During the motion cluster member ship of individual objects may vary due to the change in speed. Hence the abstraction of join between and join within are performed for filtration of redundant data. In order to preserve the performance in the execution of multiple objects, authors have proposed semantic load shedding mechanism that rejects insignificant points. This method doesn't consider the splitting of moving objects in different clusters. One of the significant proposal in this approach is the grouping of multiple object movements and queries in single cluster.

Periodical clustering of moving objects with multiple criteria in spatial networks is described in CMON [2]. It proposes a structure called cluster block for maintaining the clusters. The distance between two moving objects are measured as the length of the shortest path connecting them in the spatial network. Unlike SCUBA this method effectively manages split and merge of cluster blocks. In order to manage the termination point it also requires intermediate destinations in the travel path. When an object in cluster block reaches the destination it departures from corresponding block. To reduce the cost the split scheme is managed based on the direction of movement and speed of the objects. However the CMON framework doesn't manage queries on moving objects.

Jensen et al. [4] outlines a new scheme that is capable of incrementally clustering moving objects. This proposal employs a notion of object dissimilarity that considers object movement across a period of time, and it employs clustering features that can be maintained efficiently in incremental fashion. A data structure called clustering feature is set and updates incrementally with the key properties of the moving object cluster. An average radius function is used that automatically detects cluster split events which compared to existing approaches, eliminates the need to maintain bounding boxes of clusters with large amounts of associated violation events. Young et al. [12] presents a fast and stable incremental clustering algorithm that forces minimum memory requirement. The method employs a Winner Take All paradigm, a computational model generally applied in neural networks for competitive learning. In order to update the centroid of moving clusters the starvation trace approach is adopted, it allows idle centroids to garner credit over the time when they are not considered as centroid. Once the centroid is updated it loses this credit. Additional statistical measure is adopted to stabilise the input parameters for clustering process.

2.2 Querying from Moving Object Data

Querying from a dynamic moving object database is rather different from querying from a static database. Dynamic database queries are to strictly adhere with the spatial and temporal factions of querying and queried entities. Based on the parameters used for queries and outputs generated queries can be of different types [9] such as distance queries, K-nearest queries, range queries, trajectory queries, aggregate queries etc. Zhou et al. [17] presents close pair range queries concept of moving objects that is used to identify pairs of objects closer than specified distance during the specified time interval and within user defined spatial range. Special storage and retrieval trajectories schemes based on Multiple TSB-Tree method have adopted in this. Closed pair queries are particularly applicable in air traffic control, battlefield configurations and intelligent transportation systems etc. The query, *Which airplanes were closer to each other than 10 miles during the past month in Massachusetts*, illustrates the purpose of this concept. The SCUBA [7] framework efficiently manages moving queries by combining it with moving objects. A continuously running query in SCUBA is represented with different parameters such as id, location, time, speed, destination location and query attributes. Queries generated from multiple moving objects are also clustered based on the attributes. The advantage is that instead of monitoring individual objects, answers can be obtained from cluster table that is maintained by periodical updation. Adjacently moving queries are merged together by two inter related schemes called join between and join within. Once the cluster that consists of moving object and moving queries reaches a destination, SCUBA assumes that it dissolves and the algorithm doesn't consider possibility of splitting the clusters. It is designed particularly for continuous range queries. Gryllakis et al. [3] outlines a spatio-temporal database engine Hermes@Neo4j for efficient storage and indexing of semantic trajectories. It provides number of utilities that facilitates hybrid indexing for spatio-temporal and textual data. The framework also suggests special types of queries called Spatio temporal-Keyword pattern (STKP) that will result semantic trajectories over a geographical area.

2.3 Moving Object

A moving object at time t can be represented in the form $O(id, x, y, t)$. The coordinates (x, y) indicate latitude and longitude of the moving object and t is the time in which the object resides the location (x, y). The position (x, y) is called spatial component and t_i is its temporal component [1], where $t_{i-1} < t_i < t_{i+1}$. Apart from this implicit details such as speed and direction of movement of the moving object are calculated form moving objects. Traces of movement can be represented in the form of trajectories (Fig. 1).

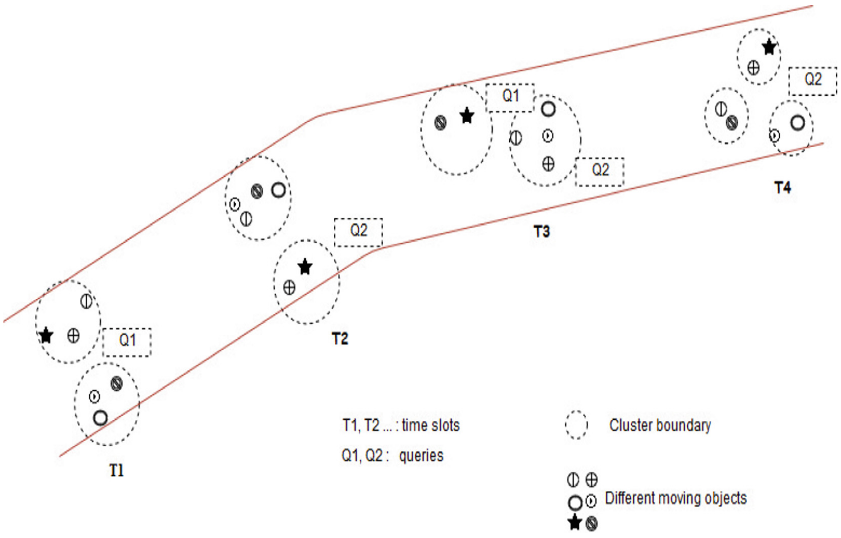


Fig. 1. Representation of periodic clustering of moving objects and queries.

2.4 Semantic Properties

Semantic properties are contextual information of the moving object trajectories. Trajectories augmented with semantic information are called semantic trajectories. The derivation of semantic data depends on the interests of data mining process such as Direction of movement, velocity, speed, maximum and minimum speed, different objects traveled in parallel etc. A study of trajectories annotated with semantic data can be seen in [16].

2.5 Continuous Clustering of Moving Objects

Clustering is the process of dividing given D objects into different classes as $C = \{C_1, C_2, C_3 \cdots C_m\}$ based on definite similarity measures, where $D = \{O_1, O_2, O_3 \cdot \cdot \cdot O_n\}$. C_i. is a cluster in the cluster set C, where $i = 1, 2, 3 \cdots m$ and $C \subseteq D$ [10]. Object moving on the constrained path are clustered periodically according to their geographical distance. According to the mobility behavior, different number of clusters are generated in each clustering time slot. The cluster membership of each objects also changes along the slots according to its semantic attributes such as velocity and direction of movement. Even though objects are considered within a constrained environment, its mobility pattern can change within the limits stipulated by the travel network. The frequency and places of querying is set according to the preloaded locations. These locations, in the case of road network could be major traffic junctions or diversions. New moving objects can be added or existing can departed at this point. Such vital geographical points also could be semantic locations identified according to the model specified in [8].

2.6 Cluster Based Continuous Querying

Queries as to moving object database are classified in to range queries, trajectory queries and aggregate queries etc. Range queries are basic operations on trajectory database [11] that retrieves records between a specific range. Trajectory Queries works on historical spatio-temporal data. Our method employs processing of spatio-temporal aggregate queries, which are aggregate queries specific to moving object data. With respect to the final inference ascertained from both clustering and aggregate queries both the abstractions provides general behavior of a set of data.

Let Q denotes a spatio-temporal aggregate query for a given space time constraints and is represented as $Q(id, x, y, t, Attr, Qtype)$. Given a tuple of attribute values $Attr = (attr_1, attr_2 \cdots attr_d)$ where $attr_i \in dom(Attr)$, the domain value specifies values for *select* and *condition* clause in the query. The coordinates (x, y) indicate latitude and longitude of the moving object and t is the time in which the query resides the location (x, y). Here (x, y) is the spatial component and t is its temporal component, where $t_{i-1} < t_i < t_{i+1}$. Here the triplet (x, y, t) is referred as query feature that describes the query. For the execution of our algorithm we identifies two class of spatio-temporal aggregate queries, that are applicable to the moving objects, as Static Aggregate Queries (SAQ) and Dynamic Aggregate Queries (DAQ). Static aggregate queries are snapshot queries while DAQ requires periodical updation of answers. There are two expiration parameters applicable to DAQ, period(Δt) and distance Δd. Some of the possible queries are listed below.

- Get total number of objects traveled between loc_1 and loc_2 during $time_1$ and $time_2$

```
SELECT COUNT(*)
FROM clus_ds  as c
WHERE c.tim BETWEEN t1 and t1
AND c.location BETWEEN loc1 AND loc2
```

- Average speed of objects traveling between loc_1 and loc_2 in next t seconds.

```
SELECT AVG(duration)
FROM clus_ds  as c
WHERE c.location BETWEEN loc1 AND loc2 UNTIL c.curr_time + t
```

- Number of objects proceeding to loc_1 travelling with a speed of *speed*

```
SELECT COUNT(*)
FROM clus_ds  as c
WHERE c.location BETWEEN loc1 AND loc2
AND c.speed <= speed UNTIL c.curr_time + t
```

The parameter *Attr* specifies values for select and conditional clauses and *Qtype* specifies in which category the query belongs. The selection of query category between SAQ and DAQ are automatically done by the system based

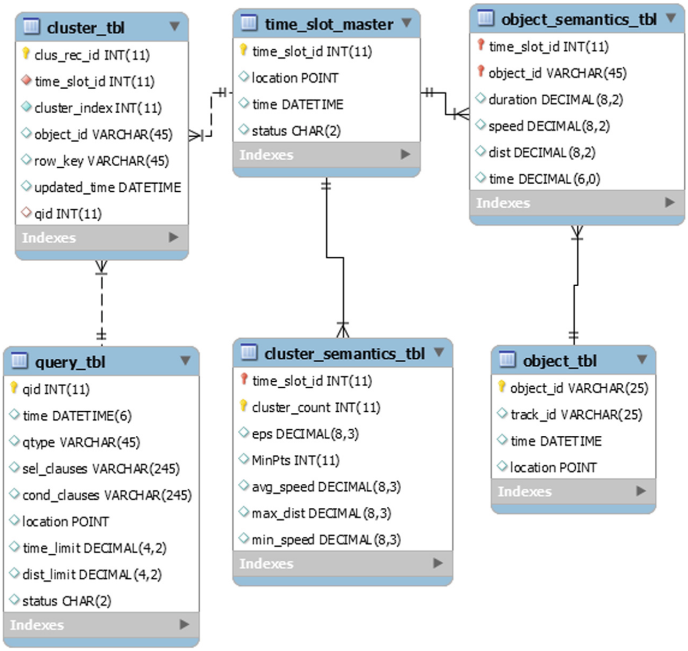


Fig. 2. The data model

on the querying parameters. The cluster results are stored in special structure that keep tracks different parameters in various time slots and queries are applied on this data structure. To reduce the execution overheads with respect to DAQ the queries are grouped according to its spatio-temporal attributes. In order to answer DAQ we first calculate the two expiration values. If queries with attribute values of same domain are already running, such queries are grouped for common segments.

2.7 Calculation of Clustering Parameters

The density clustering method adopted in this paper, DBSCAN, is driven by two vital parameters Eps and MinPt. The Eps indicate minimum distance required to be maintained for including objects in a cluster. MinPts is the minimum number of points required to be there to form a cluster. As the objects we are considering is continuously changing its location, there is no point in using an Eps a constant value. Hence our method dynamically calculates the Eps Value according to the semantic feature of the travel network. As a standard mechanism MinPts is $\ln(n)$, where n is the maximum number of objects. Due to the unavailability of real querying environment we have randomly generated both dynamic and static spatio temporal aggregate queries for the experiment. Our focus is to evaluate the efficiency of querying approach by the use of density based clustering.

2.8 Data Model

The Data model used in the clustering process is given Fig. 2. Spatio-temporal attributes from the moving objects are stored in object_tbl. Moving query also tracked and stored in query_tbl, since query is fired from another moving objects spatio-temporal attributes are valid for moving query as well. The speed of the query is an important parameter that determines the scope and extend of query. Attributes of periodical clustering is stored in cluster_tbl, the cluster_index and time_slot_id are two important attributes of the table that keep tracks cluster membership of various objects in continuous time slots.

3 Algorithm

The pseudo code for the proposed method is given as Algorithm 1. Input for the system comprises of moving object data and query data. Semantic location list (*SemLocList*), number of minimum points to cluster(*MinPts*) and cluster radius (*Eps*) are additional values required for processing. The cluster parameters are dynamically calculated based on the number of objects and semantic properties of the road network. The spatio-temporal values and other parameters of moving object and query are constantly updated in the data structures *obj_tbl* and *qry_tbl* respectively. Accepted values are clustered in each pre-defined semantic locations *SemLocList* and cluster results are maintained in the data structure *clus_ds*, as given in line 4. The queries received from different entities are classified in to two according to the contextual properties. If the query belongs to the category *SAQ* no further processing is required, simply fetch the data from *clus_ds*. If the query belongs to *DAQ* category the scope of queries are to be evaluated (line 10 to 15).

4 Experimental Evaluation

4.1 Environmental Setup

The real world trajectory data set TDrive [14,15] is used for experimentation. Each record in the trajectory consists of latitude, longitude, altitude and temporal information of the moving objects. For the ease of our work we exclude altitude component. Since the traces of constrained paths, especially the road networks is the focus of our study, it is evident that the elevation part is not significant in computing the distances. GPS traces available in various formats are pre-processed and loaded in MySQL database. To enable efficient processing spatial indexing is used to store the coordinate values. The algorithm is implemented in Java 1.8. All experiments are conducted in Intel Core i7 machine with 12 GB RAM.

Algorithm 1. Clustering and Querying of Moving Objects

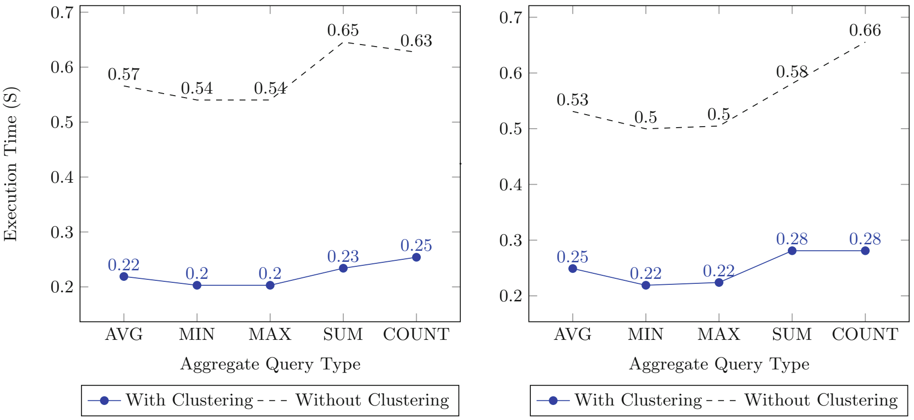
```
Input  : Moving object $O(x, y, t)$, Moving query
         $Q(x, y, t, attr, qtype)$,SemLocList
Output: Query results, Moving object clusters $D$
Accept O and Q over a constrained travel network and update the values in
obj_tbl and qry_tbl
Calculate MinPts, Eps
foreach location in SemLocList do
    $clus\_ds \leftarrow Density\_Cluster(clus\_ds, MinPts, Eps)$
    foreach Q received do
        Identify Q.qtype
        if Q.qtype = SAQ then
            Fetch the result from clus_ds
        end
        else if Q.qtype = DAQ then
            Calculate the query expiry time $\Delta t$
            Calculate the query expiry distance $\Delta d$
            while $\Delta t$ expires OR $\Delta d$ expires do
                Evaluate Q in each SemLoc
            end
        end
    end
end
```

Table 1. Cluster generation statistics for $ObjectCount = 10$.

No. time slots	No. records	Time (Seconds)	No. clusters generated max.
50	3673	0.56	2
100	7273	0.47	3
200	14545	3.5	3
500	35321	17.88	4
1673	49889	33.52	3

Table 2. Comparative study on detection of joining or attrition rate for different number of objects

No.objects	No. of objects detected (attrition/joining)		Rate of variation
	With clustering	Without clustering	
10	5	6	0.166666667
15	11	13	0.153846154
20	15	15	0
25	21	22	0.045454545
30	23	26	0.115384615



(a) $ObjectCount = 10$ (b) $ObjectCount = 15$

Fig. 3. Comparative plot on execution time for different aggregate queries with number of cluster slots = 200.

4.2 Discussion

The Static and Dynamic Spatio-Temporal Aggregate Queries are executed in a simulated environment. The results of DAQ are computed in each semantic locations along the network, which are loaded to the system in advance. The preliminary tests shows that short values of spatial (Δd) and temporal (Δt) may not receive any result. The time taken to cluster moving objects and maximum number of clusters generate are given in Table 1. As the number of semantic locations to cluster increases time for execution also increases. Number of clusters generated depends on the moving pattern of the objects. Less number of objects indicates that objects are traveling close and the speed variation is also less. It is also a fact that number of clusters also depends on the geographical feature of the moving path. During the journey some of the moving objects leave from the path, identifying this attrition/joining rate is a valuable information in mobile data management. This rate is effectively calculated using our approach with minimal variation and in maximum speed. Table 2 reports the statistics of drop off/joining rate of different objects that are detected with and without clustering. Another highlight of our algorithm is that consideration of clusters for answering the aggregate queries. By that heavy computation process can be leveraged. Figure 3 pictures a comparative plot on time consumed for executing different type of aggregate queries for different number of moving objects with and without clustering.

5 Conclusion and Future Work

Clustering of moving object data is gaining more momentum as an area of research due to the generation of huge spatial temporal data from wide range of location sensing devices. In this paper we are proposing an algorithm for the effective processing of spatio-temporal aggregate queries. Our algorithm is devised by ascertaining the fact that inferences obtained from clustering and processing of aggregate queries provides same conclusion. Without processing the entire data and with an incremental density based clustering process our method effectively answer spatio-temporal aggregate queries. This will take care of spatial temporal and semantic aspect of moving objects. In future we are to extend the system by modifying the cluster data structures by accommodating more semantic features of the moving objects.

References

1. Alvares, L.O., Bogorny, V., Kuijpers, B., de Macedo, J.A.F., Moelans, B., Vaisman, A.: A model for enriching trajectories with semantic geographical information. In: Proceedings of the 15th Annual ACM International Symposium on Advances in Geographic Information Systems, p. 22. ACM (2007)
2. Chen, J., Lai, C., Meng, X., Xu, J., Hu, H.: Clustering moving objects in spatial networks. In: Kotagiri, R., Krishna, P.R., Mohania, M., Nantajeewarawat, E. (eds.) DASFAA 2007. LNCS, vol. 4443, pp. 611–623. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-71703-4_52
3. Gryllakis, F., Pelekis, N., Doulkeridis, C., Sideridis, S., Theodoridis, Y.: Searching for spatio-temporal-keyword patterns in semantic trajectories. In: Adams, N., Tucker, A., Weston, D. (eds.) IDA 2017. LNCS, vol. 10584, pp. 112–124. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68765-0_10
4. Jensen, C.S., Lin, D., Ooi, B.C.: Continuous clustering of moving objects. IEEE Trans. Knowl. Data Eng. **19**(9), 1161–1174 (2007)
5. Li, Y., Han, J., Yang, J.: Clustering moving objects. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 617–622. ACM (2004)
6. Li, Z., Lee, J.-G., Li, X., Han, J.: Incremental clustering for trajectories. In: Kitagawa, H., Ishikawa, Y., Li, Q., Watanabe, C. (eds.) DASFAA 2010. LNCS, vol. 5982, pp. 32–46. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12098-5_3
7. Nehme, R.V., Rundensteiner, E.A.: SCUBA: scalable cluster-based algorithm for evaluating continuous spatio-temporal queries on moving objects. In: Ioannidis, Y., et al. (eds.) EDBT 2006. LNCS, vol. 3896, pp. 1001–1019. Springer, Heidelberg (2006). https://doi.org/10.1007/11687238_58
8. Nishad, A., Abraham, S.: Semantic trajectory analysis for identifying locations of interest of moving objects. In: 2017 International Conference on Networks & Advances in Computational Technologies (NetACT), pp. 257–261. IEEE (2017)
9. Niyizamwiyitira, C., Lundberg, L.: Performance evaluation of trajectory queries on multiprocessor and cluster. Comput. Sci. Inf. Technol. Vienna Austria **6**, 145–163 (2016)
10. Portugal, I., Alencar, P., Cowan, D.: Developing a spatial-temporal contextual and semantic trajectory clustering framework. arXiv preprint arXiv:1712.03900 (2017)
11. Xu, J., Lu, H., Güting, R.H.: Range queries on multi-attribute trajectories. IEEE Trans. Knowl. Data Eng. **30**(6), 1206–1211 (2018)

12. Young, S., Arel, I., Karnowski, T.P., Rose, D.: A fast and stable incremental clustering algorithm. In: 2010 Seventh International Conference on Information Technology: New Generations, pp. 204–209. IEEE (2010)
13. Yu, Y., Wang, Q., Wang, X., Wang, H., He, J., et al.: Online clustering for trajectory data stream of moving objects. Comput. Sci. Inf. Syst. **10**(3), 1293–1317 (2013)
14. Yuan, J., Zheng, Y., Xie, X., Sun, G.: Driving with knowledge from the physical world. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 316–324. ACM (2011)
15. Yuan, J., et al.: T-drive: driving directions based on Taxi trajectories. In: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 99–108. ACM (2010)
16. Zheng, B., Yuan, N.J., Zheng, K., Xie, X., Sadiq, S., Zhou, X.: Approximate keyword search in semantic trajectory database. In: 2015 IEEE 31st International Conference on Data Engineering, pp. 975–986. IEEE (2015)
17. Zhou, P., Zhang, D., Salzberg, B., Cooperman, G., Kollios, G.: Close pair queries in moving object databases. In: Proceedings of the 13th Annual ACM International Workshop on Geographic Information Systems, pp. 2–11. ACM (2005)

Analysis of Different Supervised Techniques for Named Entity Recognition

Archana Goyal[1](✉), Vishal Gupta[2], and Manish Kumar[3]

[1] PG Department of Information Technology,
Goswami Ganesh Dutta Sanatan Dharma College, Chandigarh, India
id.archana@yahoo.co.in

[2] Department of Computer Science and Engineering,
University Institute of Engineering and Technology, Chandigarh, India
vishal@pu.ac.in

[3] Department of Computer Science and Applications,
Panjab University Regional Center, Muktsar, Punjab, India
mkjindal@pu.ac.in

Abstract. The enormous growth of information available on the internet poses a great challenge to Information Extraction tasks. Named Entity Recognition is one of such an Information Extraction task which works on locating the presence of entities belonging to some predefined categories. It is a crucial pre-processing tool assisting in diverse Natural Language Processing applications such as Automatic Text Summarization, Machine Translation, etc. In this article, we analyze the performance of different supervised machine learning algorithms worked for Named Entity Recognition task. For this, we use five diverse learning algorithms namely, Conditional Random Field (CRF), Naive Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT) and Multilayer Perceptron (MLP) based neural network to create a variety of models based upon several combinations of different features that are independent of any domain-specific knowledge. These algorithms are evaluated mainly for the Hindi language. The evaluation of experimental results shows how Multilayer Perceptron based neural network approach performs better than other supervised learning methods used in this task.

Keywords: Named Entity Recognition (NER) · Conditional Random Field (CRF) · Naive Bayes (NB) · Support Vector Machine (SVM) · Decision Tree (DT) · Multilayer Perceptron (MLP)

1 Introduction

Named Entity Recognition (NER) works on identification and classification of every word form in a text into some predetermined classes such as Person, Location, Organization, Number, Measurement etc. NER performs a vital act

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 184–195, 2019.
https://doi.org/10.1007/978-981-15-0108-1_18

in various NLP applications such as Question-Answering [1], Machine Translation [2], Automatic Text Summarization [3] etc. The research on Named Entity Recognition was first considered in MUC-6 [4]. Later on, various scientific events such as Automatic Content Extraction (ACE) [5], IREX-2000 [6], CoNLL-2002 [7], CoNLL-2003 [8], IJCNLP-2008 [9] brought revolution in the emergence of NER. Various issues and challenges have been highlighted by [10] in recognizing named entities. These challenges include dealing with nested entities, handling of ambiguous data, annotation of the training dataset, lack of availability of language resources, etc. Different researchers have worked upon different segment representation techniques [11] to deal with the problem of nested entities. Till now, Named Entity Recognition task has been performed in different languages using different techniques. NER techniques can be sorted into three main groups namely, rule-based, machine learning based and hybrid techniques. The accuracy of rule-based techniques [12,13] rely mainly on choosing of handcrafted rules used for extracting named entities. Most of the earlier systems were rule-based systems. Nowadays, machine learning based techniques [14,15] have taken place of rule-based techniques to overcome the difficulties faced by earlier rule-based systems. The success of machine learning based techniques depends upon the effectiveness of features extracted, availability of huge training dataset as well as the choice of the suitable classification algorithm.

The aim of the article is to find the appropriate algorithm suitable for the extraction of named entities. Hybrid techniques [16,17] take the benefit of both rule and machine learning based methods. In this paper, five different machine learning algorithms namely, Naive Bayes [18,19] (Gaussian and Multinomial), Decision tree [20] (Gini index and Information Gain), SVM [21,22], CRF [23] and MLP based neural network [24] has been used for building different models for extracting named entities out of Hindi data. The results obtained using these classifiers are compared. Out of which, MLP based neural network technique is found as the most suitable learning technique for recognizing named entities.

The remaining article is arranged as follows. Related work is discussed in Sect. 2. A brief introduction of various machine learning techniques used in our work is reported in Sect. 3. The approach used for NER task is depicted in Sect. 4. Experimental results of each learning algorithm and discussions are reported in Sect. 5. Finally, we finish with future works in Sect. 6.

2 Related Work

Preliminary work in the field of NER task highlights the progress made in this area. Mahalakshmi et al. [18] presents domain based NER system for Tamil language using Naive Bayes classifier. The authors divide the whole system into two steps namely, analysis step and synthesis step. Analysis step includes preprocessing, morphological analysis and parsing modules whereas the synthesis step focuses on training and testing of Tamil dataset. Gorla et al. [19] presents a generative model for extracting named entities from Telugu news articles using Naive Bayes classifier. Paliouras et al. [20] proposes a Named Entity Recognition system for extracting person, location, organization using dataset obtained from MUC-6 [4] by learning decision trees. The authors make the use of decision trees in constructing NER grammars which further outperform the results of the system. Saha et al. [25] have proposed a hybrid NER system using vote-based classifier ensemble technique. In this article, the authors combine the predictions of seven different classifiers using Multiobjective Optimization (MOO) technique. Michailidis et al. [21] have performed the Named Entity Recognition task for Greek language using Support Vector Machine, Maximum Entropy, and OneTime method. This system exploits the linguistic features to get the desired results. Ekbal et al. [22] have developed a NER System using Support Vector Machine for Hindi and Bengali language by taking advantage of contextual and other linguistic features. Ekbal et al. [23] have presented a CRF based NER System for the Bengali language. Contextual information and other variety of features have been used by the authors to train and test the dataset. Gallo et al. [24] have presented a NER system using a neural sliding window. In this study, an MLP based algorithm is implemented to perform context-based NER approach. The authors have found this approach good for grammatically ill documents. Das et al. [26] have presented a NER system for Bengali language using word embeddings and Wikipedia categories. The authors found that word embedding based NER outperformed CRF based NER. Banik et al. [27] have proposed a NER system for Bangla online newspaper using Gated Recurrent Unit (GRU). The authors have also explored that capability of deep learning algorithms are more profound than traditional supervised learning algorithms in recognition of named entities.

3 Supervised Techniques Used for NER

In this study, five different classification algorithms have been analyzed for recognizing NEs. A brief description of these algorithms is presented below:

3.1 Naive Bayes

Naive Bayes [18] is a popular classification algorithm which is widely used for classifying inputs into binary or multi-classes in the field of Information Retrieval, NLP, Pattern Recognition, etc. It works upon Bayes' Theorem which assumes that the feature vectors used for classification are conditionally independent. Mean and Variance of attributes of each class are the two key parameters of the Naive Bayes algorithm. The performance of the system decreases if the computed variance increases.

Let n is the required target and m is the feature vector. So, our final aim is to compute the possibility that feature vector m belongs to target n.

$$Pi(n|m) = Pi(n) * Pi(m|n)/Pi(m) \tag{1}$$

Here, Pi(n) is the prior probability of target (n). Pi(m) is the feature vector's (m) prior probability. Pi(m|n) is the probability of vector m given target n. Finally, Pi(n|m) is the post probability of target after considering the features. We have experimented with Gaussian NB and Multinomial NB in a Python environment for our work. Out of these two, Gaussian Naive Bayes provides better results than Multinomial Naive Bayes.

3.2 Decision Tree

Decision Tree Classifier [20] is the simple and most commonly used classification algorithm. It has its wide scope in data mining and pattern recognition problems. Decision Tree is a tree-like structure which consists of three types of nodes i.e. root node, internal nodes and leaf nodes where internal nodes point to the feature variables which further expands to its possible values at succeeding level. The leaf node contains the target variable or class corresponding to the feature vector. Let the training data be X = X1, X2, ..., Xn, where each set Xi is a feature vector. For training, another vector is also available T = T1, T2, ..., Tn, where each Ti represent the different category in which each training sample is classified. The algorithm generates sub-trees at every particular node by selecting the features that are most suitable for further expansion. This suitability is determined by calculating normalized information gain. The highest value of normalized information gain becomes the base of the selection of features. This process continues until the final results are found out. We use Classification and Regression Trees (CART) algorithm in Python which is very similar to the decision tree algorithm C4.5. Default parameters of the decision tree are used for the experiments.

3.3 Support Vector Machine (SVM)

Support Vector Machine [28] is one of the supervised machine learning models which produces a linear hyperplane that divides the underlying data either in a positive category or in the negative category as shown in Fig. 1. SVM is a non-probabilistic classifier which is well suited to text categorization problems and can deal with a large number of features with high accuracy. In spite of handling a large number of features, it does not fall into over-fitting. SVM works well with a binary-class problem: (P1, T1, . . ., (PN, TN) where Pi ∈ Fv is the feature vector of the ith example and Ti ∈ {Yes, No} is the class of the ith example in the training data. Basically, the main purpose of SVM is to separate the examples into positive or negative categories with maximum margin.

$$(wt.p) + mx = 0 \qquad\qquad wt \in Fv, mx \in F. \tag{2}$$

Here, two variables define the linear separator i.e weight wt assigned to each feature, and a maximum margin m which determines the distance between hyperplane and the origin.

3.4 Conditional Random Field (CRF)

Conditional Random Field [29] comes into the category of probabilistic graphical models. It is well suited for sequence labeling problems like Named Entity

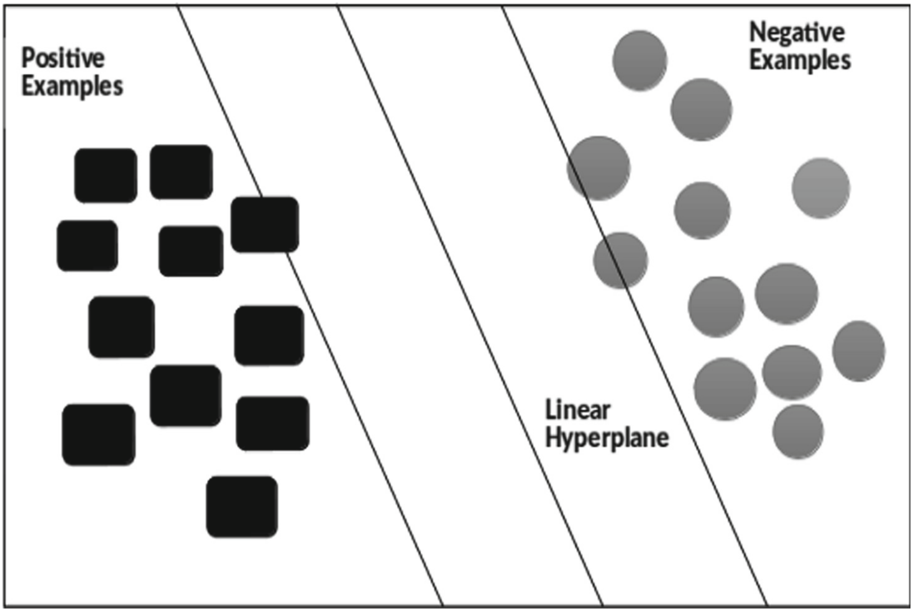


Fig. 1. Support Vector Machine

Recognition, Object Recognition, Part of Speech (POS) Tagging, etc. CRF can work well with a large amount of non-independent features. Conditional Random Field provides conditionally trained model. It has a special feature of considering surrounding examples. CRF calculates the conditional probability of a target sequence t = t₁, t₂, . . ., tₙ given an observation sequence s = s₁, s₂, . . ., sₙ which is calculated as follows:

$$P \wedge (t|s) = \frac{1}{Z_0} exp\left(\sum_{n=1}^{N}\sum_{k=1}^{K} \lambda_k * f_k(n, s, t_{n-1}, t_n)\right) \tag{3}$$

Here, $f_k(n, s, t_{n-1}, t_n)$ is a feature function. This function considers previous and current target label to calculate the probability. λ_k is the learning weight and Z_0 is the normalization factor which is used to sum up conditional probabilities up to 1. We have used sklearn_crfsuite package in Python to build a model which calculates the conditional probabilities to predict the class of unknown data.

3.5 Multilayer Perceptron (MLP)

A Multilayer Perceptron [24] is a class of artificial neural networks. It consists of multiple layers namely input layer, hidden layers, and output layer. Perceptrons are known as neurons based model. Neurons are a basic building block of any artificial neural network. It is a computational unit which applies activation function on weighted input signals and produces output signals. Hidden layer and output layer can contain a number of neurons. In this study, we have used word embeddings to represent our inputs into a real-valued vector and designed a sequential MLP model which is built on the top of embedding layer, dense layer, and dropout layer. The size of feature vector representation is specified as 50 dimensions and a number of neurons used in the hidden layer are aligned to 64. The activation function used in the hidden layer and output layer is Relu and Softmax respectively. Categorical cross_entropy is used as the loss function for output tuning. Adam is used to optimizing the cost of the network. Batch size is fixed to 128 and 50% dropout is used to overcome the problem of over-fitting at passing input to the network and at Softmax layer. The maximum length of the sequence is set to 50 to equalize the length of all input sequences. A total of 20 epochs is used to train the network model.

4 Approach Used for NER

In this section, we present the details about dataset used, pre-processing of the dataset, features extraction and evaluation metrics used to measure the performance of the NER task.

4.1 Dataset and Pre-processing Used for NER

In this work, a named entity tagged corpora of 2,24,687 tokens of Hindi language has been used which is defined for the IJCNLP-08 NER Shared Task [9]. Hindi dataset obtained from IJCNLP-08 Shared Task is available in Shakti Standard Format (SSF) [30] which is converted into a format suitable for training and testing. Hindi corpora are originally annotated with 12 different named entities as shown in Table 1 which is further converted into 4 entity tags namely, Person, Organization, Location and Miscellaneous. Number Expressions (NEN), Time Expressions (NETI) and Measurement Expressions (NEM) are mapped to Miscellaneous Entity. Other entities of the shared task are mapped to 'Other-than-NE' which is denoted by 'O'. In this way, the newly mapped tag set becomes as shown in Table 2. For determining the boundaries of named entities, these four named entities are further tagged into IOB (Inside, outside, begin) format which was followed in CoNLL-2003 Shared Task [8]. For training, 80% of Hindi dataset is used and testing is done on rest 20% of the dataset.

4.2 Feature Extraction

Feature extraction is an important aspect to be considered while training and testing data on supervised learning algorithms. In this work, most of the features are extracted without using any language specific resources and deep domain knowledge.

- Contextual Information: Contextual information is obtained by collecting the surrounding words of the current token. This is based on the hypothesis that context words contain intended information helpful in the identification of named entities.

- Part of Speech Tagging: Part of Speech are the different categories of words which carry similar grammatical properties. It acts as an effective feature for Named Entity Recognition. We have used CRF-based POS tagger [31] for POS tagging.
- Word Prefix and Word Suffix: Word prefixes and word suffixes are the prescribed length sequence of characters stripped from the left side and right side positions of the words. The prefixes and suffixes of length up to 2 is considered for this work. Keeping the observation in view that NEs keep some common prefixes and/or suffixes, this feature has been included.
- Initial Word: This feature contains binary value either '1' or '0' which checks if the current word is the initial word of the sentence. This feature is considered important because, in most of the Indian languages, Subject is likely to be a named entity.
- Last Word: This feature contains binary value either '1' or '0' which checks if the current token is the last word of the sentence. As most of the Indian languages follow Subject-Object-Verb structure. So, verbs are mostly at the last position of the sentence which separates it from named entities.
- Infrequent Word: Most frequent words are rarely a named entity because they are mostly the stop words in any of the language. To separate frequent occurring words from NEs, we maintain a list of words which occurs less than or equal to 20 times in a document. This is the binary feature which is set as '1' if the current token appears in the maintained list otherwise it is set as '0'.
- String Length: This feature contains binary value which checks whether the length of characters in a word form is less than or equal to a predefined threshold value which is set as 3 here. String length feature is included to keep in view that there are fewer chances of very short words to be the named entities.
- Digit Features: Several digit features are included to check the presence of symbols and/or digits in a word. These features include contain_digits (word contains digit only), two_digits (word contains 2 digits), four_digits (word contains four digits), contain_digit_comma (word contains digit and comma), contain_digit_hyphen (word contains digit and hyphen). These features are helpful to identify number, date and measurement entities which are mapped to Miscellaneous category.
- Word Embeddings: Word embeddings [26] are distributed representation of words wherein the words which are in close proximity will have the same feature vector representation. This feature vector represents different aspects of input data and finds the semantic similarity among them. Feature vector representation of those words is the same which are semantically similar. In this study, word embeddings are only used for MLP based neural model and the output dimension of word embedding vector is set to 50.

Table 1. Original named entity tag set used in IJCNLP-08 NER shared task

Named entity tag	Meaning	Example
NEP	Person Name	kavIsha/NEP, kavIsha rAma gupta/NEP
NEL	Location Name	panipAtA/NEL, gt roDa/NEL
NEO	Orgaization Name	pAnjAba bishVbidyAlYa/NEO
NEA	Abbreviation	bi je pi/NEA, em di/NEA
NED	Designation	diRectAra/NED, cheYArmAn/NED
NEN	Number	5/NEN, panCha/NEN
NEB	Brand	fYAntA/NEB
NEM	Measurement	dasa keji/NEM, tina meetrA/NEM
NETI	Time	2008/NETI, 5 ema/NETI
NETE	Term	kemikYAla riYYAkchYAna/NETE
NETP	Title-Person	shrImana/NETP, shrImati/NETP
NETO	Title-Object	AfricAn biUti/NETO

4.3 Evaluation Metrics

For evaluating the results, standard intrinsic metrics namely, Precision, Recall, and F-score have been used.

Precision is the proportion of a number of named entities correctly identified as that of the total number of entities identified by the system.

Recall is the proportion of a number of named entities correctly identified by the system as that of the total number of entities present in the test data.

F-score is the measure of computing harmonic average of Precision and Recall.

Table 2. Tagset mapped in this task

NE tags	Tagset used	Meaning
NEP	Person	Single/Multi token person name
NEL	Location	Single/Multi token location name
NEO	Orgaization	Single/Multi token organization name
NEM, NEN, NETI	Miscellaneous	Single/Multi token miscellaneous name
NEA, NED, NEB, NETE, NETP	Others	Other-than-NEs

5 Experimental Results and Discussions

We build two Naive Bayes models (with Gaussian and Multinomial), two Decision Tree models (with Gini index and information gain), one Support Vector

Machine (SVM) model with standard parameters and one Conditional Random Field (CRF) model using feature set containing part of speech, word prefix, word suffix, initial word, last word, infrequent word, string length, all digit features. Contextual information features are considered only for CRF based model because CRF accommodates contextual information while training and testing the data. One MLP based neural model is also developed with only word embedding features. The results using different classification algorithms are reported in Table 3. The MLP based neural network algorithm has outperformed all other traditional algorithms.

Table 3. Results of different classification algorithms

Algorithm	Precision (in %)	Recall (in %)	F-score (in %)
Naive Bayes (Gaussian)	30	19	18
Naive Bayes (Multinomial)	03	28	05
Decision Tree (Gini index)	24	13	17
Decision Tree (Information Gain)	24	13	17
Support Vector Machine (SVM)	24	25	23
Conditional Random Field (CRF)	74	62.9	67.7
Multilayer Perceptron (MLP)	75	74	75

In this study, difference experiments are conducted on different architectures to evaluate the performance of the NER task. First of all, we have used two models of Naive Bayes namely, Gaussian and Multinomial. Out of these two classifiers, Gaussian Naive Bayes has performed well for NER. The reason behind it that Gaussian Naive Bayes takes continuous data into consideration while Multinomial considers discrete data. Most of the features used in this study are of a continuous nature. However, Naive Bayes is giving somewhat better results than Decision Tree with the same features. Support Vector Machine is performing well than Naive Bayes and Decision Tree but not than Conditional Random Field. CRF considers dependencies among data from state to state and feature to state naturally but SVM does not have such a property. Besides it, CRF takes into account the contextual information while training and testing the data. CRF has been considered as the most suitable traditional classification algorithm for the NER task. But with the emergence of deep learning concepts, traditional machine algorithms are going to be replaced with neural network algorithms. In this study, we find the best results using an MLP based neural network model with F-score value of 75%. Besides it, to achieve this output we rely only on word embedding features and appropriate network tuning parameters.

6 Conclusion and Future Works

In this article, we have presented a NER system for Hindi language using different frameworks such as Naive Bayes, Decision Tree, SVM, CRF, and MLP. The standard dataset of Hindi has been collected from IJCNLP-08 website which is initially in Shakti Standard Format (SSF). First of all, the data set is converted into an appropriate format which is suitable for training and testing. Different language independent features like contextual information, first word, last word, POS, several orthographic features have been used to build different traditional machine learning models which are applied on unknown data to extract the coarse-grained named entities. However, contextual information feature has been used additionally for CRF because it deals with the sequential data implicitly. But for designing the Multilayer Perceptron based neural model, only word embeddings are used as a feature vector. The main motivation of this study is to analyze the performance of different supervised machine learning algorithms for the NER task. Out of all the classification algorithms used in this study, the MLP based neural network has outperformed by giving F-score value of 75%. Multilayer Perceptron networks can work better by overcoming the problem of feature engineering and can be successfully applied to low resource languages.

Till now, several types of research have been conducted on Named Entity Recognition. Numerous rule-based, machine learning based and hybrid approaches have been suggested by different researchers. Named Entity Recognition for different languages/domains and for different entity types like ENAMEX, NUMEX and TIMEX have been explored. But still, there is a scope to reach the new dimensions in the field of Named Entity Recognition. Deep Learning based techniques are rising in the area of Natural Language Processing nowadays. In this study, we have not explored several neural networks like Convolutional Neural Network (CNN) [32], Recurrent Neural Network (RNN) [33] which can improve the performance of NER task so our future work can be to explore these networks for the problem of Named Entity Recognition. Use of different pre-trained word embeddings can be used to improve the performance of deep learning based Named Entity Recognition. Earlier researches have concentrated mainly on the extraction of coarse-grained entities so in future fine-grained entities can be considered for the extraction task. Besides, the development of the NER system for different languages can also be considered in the future.

References

1. Pizzato, L.A., Mollá, D., Paris, C.: Pseudo relevance feedback using named entities for question answering. In: 2006 Proceedings of the Australasian Language Technology Workshop, pp. 83–90 (2006)
2. Babych, B., Hartley, A.: Improving machine translation quality with automatic named entity recognition. In: Proceedings of the 7th International EAMT Workshop on MT and Other Language Technology Tools, Improving MT Through Other Language Technology Tools: Resources and Tools for Building MT, pp. 1–8. Association for Computational Linguistics (2003)
3. Nobata, C., Sekine, S., Isahara, H., Grishman, R.: Summarization system integrated with named entity tagging and IE pattern discovery. In: LREC (2002)
4. Grishman, R., Sundheim, B.: Message understanding conference-6: a brief history. In: COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics, vol. 1 (1996)

5. Doddington, G.R., Mitchell, A., Przybocki, M.A., Ramshaw, L.A., Strassel, S.M., Weischedel, R.M.: The automatic content extraction (ACE) program-tasks, data, and evaluation. In: LREC, vol. 2, p. 1 (2004)
6. Sekine, S., Isahara, H.: IREX: IR & IE evaluation project in Japanese. In: LREC, pp. 1977–1980. Citeseer (2000)
7. Sang, T.K.: Introduction to the CoNLL-2002 shared task: language-independent named entity recognition. In: Proceedings of Conference on Natural Language Learning (2002)
8. Sang, E.F., De Meulder, F.: Introduction to the CoNLL-2003 shared task: language-independent named entity recognition. arXiv preprint cs/0306050 (2003)
9. Singh, A.K.: Named entity recognition for south and south east Asian languages: taking stock. In: Proceedings of the IJCNLP-08 Workshop on Named Entity Recognition for South and South East Asian Languages (2008)
10. Goyal, A., Gupta, V., Kumar, M.: Recent named entity recognition and classification techniques: a systematic review. Comput. Sci. Rev. **29**, 21–43 (2018)
11. Keretna, S., Lim, C.P., Creighton, D., Shaban, K.B.: Enhancing medical named entity recognition with an extended segment representation technique. Comput. Methods Programs Biomed. **119**(2), 88–100 (2015)
12. Aone, C., Halverson, L., Hampton, T., Ramos-Santacruz, M.: SRA: description of the IE2 system used for MUC-7. In: Seventh Message Understanding Conference (MUC-7): Proceedings of a Conference Held in Fairfax, Virginia, 29 April–1 May 1998 (1998)
13. Mikheev, A., Moens, M., Grover, C.: Named entity recognition without gazetteers. In: Proceedings of the Ninth Conference on European Chapter of the Association for Computational Linguistics, pp. 1–8. Association for Computational Linguistics (1999)
14. Saha, S.K., Narayan, S., Sarkar, S., Mitra, P.: A composite kernel for named entity recognition. Pattern Recogn. Lett. **31**(12), 1591–1597 (2010)
15. Wang, Y., et al.: Supervised methods for symptom name recognition in free-text clinical records of traditional chinese medicine: an empirical study. J. Biomed. Inform. **47**, 91–104 (2014)
16. Srihari, R.: A hybrid approach for named entity and sub-type tagging. In: Sixth Applied Natural Language Processing Conference (2000)
17. Yu, X.: Chinese named entity recognition with cascaded hybrid model. In: Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics, Companion Volume, Short Papers, pp. 197–200. Association for Computational Linguistics (2007)
18. Mahalakshmi, G., Antony, J., Roshini, S., et al.: Domain based named entity recognition using Naive Bayes classification. Aust. J. Basic Appl. Sci. **10**(102), 234–239 (2016)
19. Gorla, S., Velivelli, S., Murthy, N.B., Malapati, A.: Named entity recognition for Telugu news articles using Naïve Bayes classifier. In: NewsIR@ ECIR, pp. 33–38 (2018)
20. Paliouras, G., Karkaletsis, V., Petasis, G., Spyropoulos, C.D.: Learning decision trees for named-entity recognition and classification. In: ECAI Workshop on Machine Learning for Information Extraction (2000)
21. Michailidis, I., Diamantaras, K.I., Vasileiadis, S., Frère, Y.: Greek named entity recognition using support vector machines, maximum entropy and onetime. In: LREC, pp. 47–52. Citeseer (2006)

22. Ekbal, A., Bandyopadhyay, S.: Named entity recognition using support vector machine: a language independent approach. Int. J. Electr. Comput. Syst. Eng. **4**(2), 155–170 (2010)
23. Ekbal, A., Haque, R., Bandyopadhyay, S.: Named entity recognition in Bengali: a conditional random field approach. In: Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-II (2008)
24. Gallo, I., Binaghi, E., Carullo, M., Lamberti, N.: Named entity recognition by neural sliding window. In: 2008 The Eighth IAPR International Workshop on Document Analysis Systems, pp. 567–573. IEEE (2008)
25. Saha, S., Ekbal, A.: Combining multiple classifiers using vote based classifier ensemble technique for named entity recognition. Data Knowl. Eng. **85**, 15–39 (2013)
26. Das, A., Ganguly, D., Garain, U.: Named entity recognition with word embeddings and Wikipedia categories for a low-resource language. ACM Trans. Asian Low-Resour. Lang. Inf. Process. (TALLIP) **16**(3), 18 (2017)
27. Banik, N., Rahman, M.H.H.: GRU based named entity recognition system for Bangla online newspapers. In: 2018 International Conference on Innovation in Engineering and Technology (ICIET), pp. 1–6. IEEE (2018)
28. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, New York (2013). https://doi.org/10.1007/978-1-4757-3264-1
29. Lafferty, J., McCallum, A., Pereira, F.C.: Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of the Eighteenth International Conference on Machine Learning, pp. 282–289 (2001)
30. Bharati, A., Sangal, R., Sharma, D.M.: SSF: Shakti standard format guide, pp. 1–25. Language Technologies Research Centre, International Institute of Information Technology, Hyderabad, India (2007)
31. Reddy, S., Sharoff, S.: Cross language POS taggers (and other tools) for Indian languages: an experiment with Kannada using Telugu resources. In: Proceedings of the Fifth International Workshop on Cross Lingual Information Access, pp. 11–19 (2011)
32. Gu, J., et al.: Recent advances in convolutional neural networks. Pattern Recogn. **77**, 354–377 (2018)
33. Li, S., Li, W., Cook, C., Zhu, C., Gao, Y.: Independently recurrent neural network (IndRNN): building a longer and deeper RNN. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5457–5466 (2018)

Parallel Job Scheduling Using Bacterial Foraging Optimization for Heterogeneous Multi-cluster Environment

Navjot Kaur[1], Shikha Jaryal[2], and Sahil Sharma[2](✉)

[1] Department of Computer Science and Engineering, Chandigarh University, Ajitgarh, India
navjot.e7452@cumail.in

[2] Department of Computer Science and Engineering, Lovely Professional University, Phagwara, India
shikhajaryal57@gmail.com,
sahilsharma.191191@gmail.com

Abstract. A variety of clusters are grouped together to form a multi-cluster environment which can tackle the computational needs of a system which cannot be addressed by a single cluster. Studying multi-cluster frameworks is turning challenging day by day as it requires contemporary tools to move alongside with rapidly development and enhanced complexity of one system. Job scheduling in considered as NP hard problem in parallel and distributed computing environments such as cluster, grid and clouds. The way jobs are scheduled by the scheduler is dependent on various factors like number of jobs, processor availability, arrival time etc. Metaheuristics techniques like Genetic Algorithms, Ant Colony Optimization, Artificial Bee Colony, Cuckoo Search, Firefly Algorithm, Bat Algorithm etc. are used by researchers to get near optimal solutions to job scheduling problems. This work addresses a scheduling problem with multiple objectives. The makespan and flowtime are minimized simultaneously solving the issue of optimal job allocation. This work also includes the detailed description of parallel computing and various types scheduling as well as scheduling environments. The performance of the multi-cluster environment is optimized by applying a novel meta-heuristic technique named Bacterial Foraging Optimization Algorithm. This algorithm has better convergence and is not affected by the size of problem. The proposed algorithm was evaluated for different job sets on 3 types of processor configurations. And the final values were compared to those of the existing algorithm. The results show that the proposed algorithm has performed better than the existing one and it can be concluded that the proposed algorithm is feasible and effective for optimal allocation of jobs.

Keywords: Scheduling · Metaheuristics · Multi-cluster · Co-allocation · Genetic algorithm · Bacterial Foraging Optimization Algorithm

1 Introduction

Parallel computing comprises of many different process executions being carried out simultaneously. Larger problems are sub-divided into smaller ones. The smaller problems are then solved individually at the same time. Earlier parallelism was used

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 196–207, 2019.
https://doi.org/10.1007/978-981-15-0108-1_19

exclusively for high-performance computing but now its use is expanding in different fields due to various issues such as power consumption, less memory etc. The results of the smaller problems are then combined upon completion. Each process is broken into independent parts in such a way that each processing element can execute its part simultaneously.

Multi-cluster environments are comprised of a set of connected computers (clusters) that work together to provide high-performance computing for solving large-scale optimization problems. These clusters are connected via a dedicated interconnection network [1].

Over the last few years, clusters and distributed memory multiprocessors have gained a lot of attention. Multi-cluster systems are made up of multiple geographically distributed clusters and hence provide large computation power as compared to single cluster. Due to large groups being able to share the multi-cluster, job turnaround time is reduced and system utilization becomes high. This leads to larger job sizes possible by allowing jobs to use processors in multiple clusters simultaneously i.e. to employ co-allocation [2].

Due to exceptional growth in the number of resources in diverse organizations, efficient algorithms are needed for job scheduling. Scheduling in a heterogeneous multi-cluster environment is considered as an NP-hard problem which is why population-based meta-heuristics are used to obtain near optimal solutions. Scheduling problems can be solved using mainly two types of algorithms: Deterministic Algorithms and Approximate Algorithms. The latter are considered because they tend to give near optimal solutions for large-scale problems in reasonable time. The former however, are not suitable for large-scale problems as they do not guarantee to provide optimal solutions even though they take lesser time.

Genetic Algorithm (GA) and Bacterial Foraging Optimization Algorithm (BFOA) are two such approximate algorithms. In this work, we have compared both the algorithms by using different job and processor sets. The algorithms are evaluated on a real workload trace.

1.1 Contribution

In this work, the BFOA is compared with Genetic Algorithm in terms of makespan and flowtime. Processors of 3 configurations 96, 112 and 128 have been used and 4 sets of parallel jobs have been used i.e. 100, 300, 500 and 700. Since it is a multi-cluster environment, so 5 clusters have been used. The parameters have been measured for different job sets of all the three types of processors. The advantages of BFOA lead to a considerable improvement in all the parameters which have been shown in their graphical representations. The results prove that the proposed technique is better than the existing one in terms of makespan, flowtime.

The rest of the paper is organized as follows. In Sect. 2, related work is discussed. Section 3 discusses the Genetic Algorithm in brief, followed by Sect. 4 that discusses the BFOA in a detailed manner. Experimental setup and results are shown in Sect. 5 that shows the comparison of the existing and proposed technique on various parameters. Finally, the conclusions are presented in Sect. 6.

2 Related Work

Gabaldon et al. [1] have discussed that optimization of performance criteria and scheduling are considered as NP-hard problems. So to solve these problems, a common step taken is using metaheuristic algorithms. The real traces from HP2CN workloads have been used and effectiveness of the work is determined. The experimental analysis shows that by using the GA-MF metaheuristic, makespan can be minimized along with minimizing flowtime. Gabaldon et al. [4] in their work described that since recent years, energy consumption has been a major issue in the field of large-scale computing. Their aim is a multi-objective genetic algorithm (MOGA) based on a weighted blacklist. The MOGA is implemented by making the use of non-dominated sorting genetic algorithm II (NSGA-II) and GridSim simulator is used for carrying out the simulations. MOGA showed a slight reduction in the makespan while keeping the same energy consumption results. Also, the solutions hence found had low dispersion results which conclude the fact that MOGA's robustness is independent of the nature of workloads. Dasgupta et al. [5] propose a novel load balancing strategy which makes the use of Genetic Algorithm (GA). The main motive of this algorithm is to balance the load of the cloud infrastructure and at the same time while try to minimize the makespan. The experimental analysis for a typical sample shows that the proposed algorithm is better than the existing approaches like First Come First Serve (FCFS), Round Robin (RR) and Stochastic Hill Climbing (SHC). Kalra et al. [6] in their paper, review the applications of metaheuristics in the field of scheduling. The environments mainly considered in this work were cloud and grid environments. It is a fact that metaheuristics are generally slower than deterministic algorithms and the solutions that they generate may or may not be optimal. Thus, the authors have done most of their research towards the improvement of convergence speed and solution quality. Enomoto et al. [7] discuss that the GAs having multiple dimensions requires a long time for execution. The authors have in this work proposed a technique to make improvements in the diversity of the individuals in GA by making use of MapReduce. The authors propose a technique to parallelize the RCGA in order to improve the GAs to they can be applied to a problem where solution space has multiple dimensions. The simulation results show an improvement in the solution accuracy. Chana et al. [8] propose a novel bacterial foraging based hyper-heuristic resource scheduling algorithm has been designed to schedule the jobs effectively on resources that are available in a Grid environment. The proposed algorithm's performance has been calculated with already existing common heuristics-based scheduling algorithms using the GridSim toolkit for simulation. The experimental results show that the proposed algorithm outperforms the hybrid heuristics in all the cases. The proposed algorithm minimizes the makespan along with minimizing the cost. Yang et al. [9] developed a new bacterial foraging optimizer by making the use of a newly designed chemotaxis mechanism. In the new BFO-CC, each bacterium swims along one of the many standard basis vector directions. Then the BFO-CC was compared with other state-of-the-art algorithms and so on. The experimental analysis shows that BFO-CC is a very competitive and efficient method for solving single objective optimization problems.

3 Genetic Algorithm

This is an optimization technique based on population and is grounded on Darwin's theory of evolution. In this algorithm, evolution of a population of individuals is done for obtaining better solutions. Each individual possesses certain qualities which can be altered or mutated. Each potential solution in GA, is denoted by a chromosome. A preliminary population is taken arbitrarily and it is used as a beginning point. A fitness function is computed for every chromosome so that it is acknowledged whether the chromosome is appropriate or not. Crossover as well as mutation functions are performed on the chosen chromosomes and offsprings for fresh population are created. This process is repeated until enough offsprings are created [10]. The fundamental steps for genetic algorithm are:

– *Initialization:* An initial population is arbitrarily generated. Although the size of the population is dependent upon the nature of the problem, but usually hundreds or thousands of solutions are generated.
– *Selection:* The fitness function is calculated for each chromosome and the one with best quality are selected.
– *Crossover:* The parent is crossovered to form a new offspring.
– *Mutation:* Mutation is performed to retain a genetic diversity of one generation from the next.

However, GA has a slow convergence speed and there is no guarantee of finding global maxima. Due to this we are motivated to study about a new meta-heuristic technique called Bacterial Foraging Optimization Algorithm (BFOA) and compare both of these techniques.

4 Bacterial Foraging Optimization Algorithm

BFOA, proposed by Kevin Passino in 2002, is a new member in the family of nature inspired optimization algorithms. It has drawn a worldwide attention as it is a high-performance optimizer. This algorithm is based on the foraging behavior of the E.Coli bacteria present in the human intestine. Foraging refers to the act of searching for food. The search of nutrients is done to maximize energy obtained per unit time. Also, the bacteria can communicate with each other by sending signals. The foraging decisions are taken after considering these factors. Animals searching for food and obtaining the nutrients in such a way that their energy intake is maximized per unit time is the basic assumption for the foraging theory. There are four basic steps in BFOA:

– *Chemotaxis*: Chemotaxis is achieved by swimming and tumbling [11]. These actions are performed with the help of tensile flagella present of the surface of the bacterium. When the flagella move in a clockwise direction, then each flagellum pulls on the cell and the bacterium tumbles about. But if the flagella move in a counter-clockwise manner, then they push the bacteria and it swims in one direction. When the bacteria meet a favorable environment, swimming is continued in the same direction. The favorable environment is the one in which nutrients increase as the bacteria swims. If the nutrients are not increasing then the bacteria will tumble. The bacteria keep on switching between these two positions during their entire lifetime.

- *Swarming:* The E.Coli bacteria show an interesting grouping behavior where the cells arrange themselves in the form of a travelling ring by moving up the nutrient gradient. When the cells are stimulated by high level of succinate, they release an attractant aspartate which helps them to aggregate into group and hence move as concentric patterns of swarms. After each chemotaxis step, the bacteria reach new point in space due to their movement. At each present location, fitness of each bacterium is evaluated. Fitness is represented by cost function. Better the fitness, lesser is the cost function.
- *Reproduction:* After calculating the fitness, the half healthy bacteria survive and reproduce. This is an important point which leads to many applications and advantages of BFOA, since only healthy bacteria reproduce and now, we deal with only fit bacteria. They already have more fitness value and hence this leads to fast convergence. The surviving bacterium splits into two and these two are placed at same location keeping the population of bacteria constant.
- *Elimination and Dispersal*: The convergence speed is increased by reproduction and local search is performed by chemotaxis. But this might not be enough to reach the global minimum point. Also, the bacteria might get trapped in local minima. So, to avoid this, elimination and dispersal event is performed. The bacterium having the probability of elimination and dispersal is eliminated and one bacterium is placed (dispersion) at random location. The population still remains constant [12] (Table 1).

Table 1. Nomenclature for the Algorithm

Symbol	Parameter
d	Dimension of search space
Z	Total no. of bacteria in the population
N_{ch}	Count for chemotactic steps
N_{sw}	Swimming length
N_{rp}	Count for reproduction steps
N_{eldp}	Count for elimination-dispersal events
P_{eldp}	Probability for elimination-dispersal
A(e)	Step size taken by tumble in any direction
f	Index of chemotactic step
g	Index of reproduction step
h	Index of elimination-dispersal step
v	Number of variables to be optimized

Firstly, the job matrix is loaded and clusters are initiated. If the number of jobs is less than or equal to the maximum jobs, then it is checked if the required resources are more than available resources. If that is true then fragmentation of jobs into tasks is done. Now BFOA algorithm is initialized. All the parameters are defined and initialized. Now the stopping criterion in this case is returning the best schedule. If stopping criteria is not met, then the chemotactic loop, reproduction loop and elimination-dispersal loop is initialized. Fitness function of the bacterium is calculated at each location. Increase the swim steps and calculate the new objective function. If the bacterium is not moving in the favorable environment, then tumble. Finally, the best schedule is returned and parameters are evaluated. The BFOA works in scheduling in this way:

- Scheduling in multi-cluster environment using BFOA involves allocation of jobs to processors.
- While allocating jobs to processors (chemotactic steps), we search for the schedule which returns the minimum value of parameters and best schedule for job allocation is taken for further evaluation.
- During the swimming, bacteria checks for co-allocation.
- The schedules which are not fit are discarded (elimination-dispersal step) (Table 2).

Algorithm

Let $P(f,g,h) = \{\theta^e\ (f,g,h) |\ e = 1,2,\ldots\ldots Z\}$ denote each member's position in the population of Z bacteria at the f-th chemotactic step, g-th reproduction step and h-th elimination-dispersal step.

(1) Initialize the parameters Z, N_{ch}, N_{sw}, N_{rp}, N_{eldp}, P_{eldp}, A(e), θ^e and v.
(2) Elimination dispersal loop: h= h+1
(3) Reproduction loop: g= g+1
(4) Chemotaxis loop: f= f+1

 (a) For $e = 1,2,\ldots\ldots Z$ a chemotactic step is taken like this for the e bacterium.
 (b) Fitness function $Q(e,f,g,h) = Q(e,f,g,h) + Q_{cc}\left(\theta^e(f,g,h), e(f,g,h)\right)$
 (c) $Q_{last} = Q(e,f,g,h)$ is for saving the value because there is a chance of finding better solutions.
 (d) Tumble: Generate a random vector $\Delta(e) \in R^d$ with each element $\Delta_m(e), m = 1,2\ldots\ldots d$, a random number on [-1, 1].
 (e) Movement: Consider

$$\theta^e = (f+1,g,h) = \theta^e(f,g,h) + A(e)\frac{\Delta(e)}{\sqrt{\Delta^T(e)\,\Delta(e)}}$$

 This means that a step of size A(e) has been taken in the tumble's direction for bacterium e.
 (f) Now evaluate $Q(e,f+1,g,h)$.
 (g) Swim

- Consider m=0 (which is the swim length counter).
- As m < N_{sw}.
  - ➢ Let m= m+1
  - ➢ If $Q(e,f+1,g,h) < Q_{last}$ (i.e if there is any improvement), then suppose $Q_{last} = Q(e,f+1,g,h)$ and let

$$\theta^e = (f+1,g,h) = \theta^e(f,g,h) + A(e)\frac{\Delta(e)}{\sqrt{\Delta^T(e)\,\Delta(e)}}$$

  And to compute the new $Q(e,f+1,g,h)$ use this $\theta^e = (f+1,g,h)$.
  - ➢ Otherwise, let m= N_{sw}. End of while.

 (h) If e ≠ Z, go to next bacterium (e+1).

(5) If f < N_{ch}, this means that chemotaxis step should be repeated. So repeat step 4.

(6) Reproduction:

(a) Consider the health of bacterium e as

$$Q_{health}^{e} = \sum_{f=1}^{N_{ch}+1} Q(e, f, g, h)$$

Now chemotactic parameter A(e) and bacteria are sorted in the ascending order of cost Q_{health}. Low cost means better health and vice versa.

(b) Z_k bacteria die. These are those one which have high values of Q_{last}. The remaining Z_k bacteria split.

(7) If g < N_{rp} then repeat step 3.

(8) Elimination-Dispersal: $e = 1, 2, \ldots \ldots Z$, having probability of P_{eldp}, eliminate and disperse the bacterium in order to keep the population constant. In the optimization domain, eliminate the bacterium and disperse another at any random location. Go to step 2 if h < N_{eldp}. Else end.

Table 2. Analogy for BFOA

Symbol	Full form	Role in proposed algorithm
d	Dimension of search space	Number of clusters
Z	Number of bacteria	Number of random schedules using random distribution
N_{ch}	Chemotactic steps	Assignment of jobs to clusters
N_{sw}	Swimming length	Check for co-allocation
N_{rp}	Reproduction steps	Reproduction from random schedules
N_{eldp}	Elimination-dispersal events	Eliminate those schedules which have lesser fitness value
P_{eldp}	Elimination-dispersal probability	Eliminate those schedules which demand more resources
Q_{last}	Last bacterium	Last job in schedule
Q_{health}	Fitness function	Fitness value

5 Experimental Evaluation

The experimental evaluation has been done in MATLAB environment. The observations for GA have been taken initially, followed by those of BFOA and are then compared with each other. Based on the results, graphs have been designed which are shown as we go further in the chapter. Three configurations of resources have been taken i.e. 96, 112 and 128. They have been divided into a set of 5 clusters. The evaluation is done using real workload traces (Parallel Workload Archives). 100, 300, 500 and 700 are the sets of jobs being used. The value of fitness function is taken as 0.6. All these configurations lead to the calculation of 2 parameters: Makespan and Flowtime (Table 3).

Table 3. Configuration

No. of processors	Cluster configuration
96	[32 16 16 16 16]
112	[32 32 16 16 16]
128	[32 32 32 16 16]

First of all the GA was tuned by changing the values of Mutation Rate (MR) and Crossover Rate (CR). 10 iterations and 30 chromosomes were used. The tuning has been done on 100 jobs and 128 resources. The cluster for 128 resources is [32 32 32 16 16]. Three values of these parameters were used and the one that gave the best results is chosen for further evaluation (Table 4).

Table 4. GA tuning

Values	Makespan	Flowtime
MR-0.01, CR-0.2	205389.4	64176.3
MR-0.03, CR-0.4	197552.3	61194
MR-0.05, CR-0.6	200817.7	63103.8

The above table shows the average values of parameters after taking 10 sets of values. This was done so as to stabilize the results. It could be clearly seen that MR-0.03 and CR- 0.4 gives the least values of Makespan and Flowtime. Hence, these values are used throughout the work (Table 5).

Table 5. GA parameters

Parameters	Values
Initial population	Random
Selection	Linear selection
Crossover	One point crossover
Dimensions	2
Chromosomes	30
Number of generations	10
MR, CR	0.03, 0.4

All further calculations are done using the above parameters. Now, taking the best values of MR and CR and keeping the rest of the values same, the parameters were calculated by varying resources and jobs. Firstly, 100 jobs were executed on 96, 112, and 128 processors. The results show that with the increase in the number of processors, the values of the parameters improve. Then the results for 300, 500 and 700 jobs are taken.

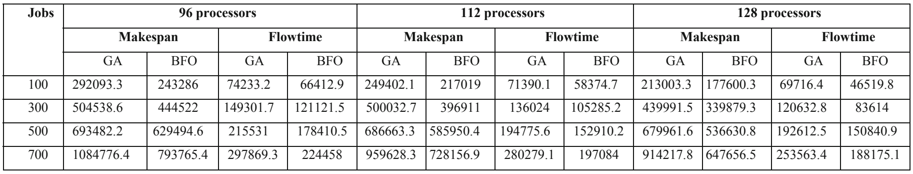
Just as we did in GA, same configurations were followed during studying the behavior of the BFOA. The values taken for BFOA parameters are shown in the following Table 6:

Table 6. BFOA parameters

Parameter	Values
Dimensions	4
No. of bacteria	50
Chemotactic steps	10
Swim length	4
Reproduction steps	4
Elimination-dispersal events	2
Elimination-dispersal probability	0.25
Elimination-dispersal probability	0.25

Initially 100 jobs were run on 96, 112 and 128 processors. Then 300, 500 and 700 jobs were also run on the same processors. Then finally we compare the outputs of both the algorithms. Firstly, the case with 96 processors will be compared taking all 4 types of jobs into consideration. Table 7 shows the readings.

Table 7. Comparison on 96, 112 and 128 processors



Jobs	96 processors				112 processors				128 processors			
	Makespan		Flowtime		Makespan		Flowtime		Makespan		Flowtime	
	GA	BFO	GA	BFO	GA	BFO	GA	BFO	GA	BFO	GA	BFO
100	292093.3	243286	74233.2	66412.9	249402.1	217019	71390.1	58374.7	213003.3	177600.3	69716.4	46519.8
300	504538.6	444522	149301.7	121121.5	500032.7	396911	136024	105285.2	439991.5	339879.3	120632.8	83614
500	693482.2	629494.6	215531	178410.5	686663.3	585950.4	194775.6	152910.2	679961.6	536630.8	192612.5	150840.9
700	1084776.4	793765.4	297869.3	224458	959628.3	728156.9	280279.1	197084	914217.8	647656.5	253563.4	188175.1

Table 7 indicates the comparison of the results of the existing and proposed algorithm on 96 processors and 100 jobs. The results show that there is a very clear improvement in the values of our proposed algorithm. It also shows the comparison for the same number of processors but 300 jobs. An improvement can be seen there as well. In the case of 112 processors, the values are smaller than the previous ones. But also, if GA and BFOA are compared, BFOA performs better than the existing algorithm. The makespan and flowtime are clearly improved in all the cases. The values in the 128-processor set are the best because the number of processors is the highest in this case. Since more number of processors lead to less time spent on execution, so the results are the best in this case (Figs. 1, 2, 3 and 4).

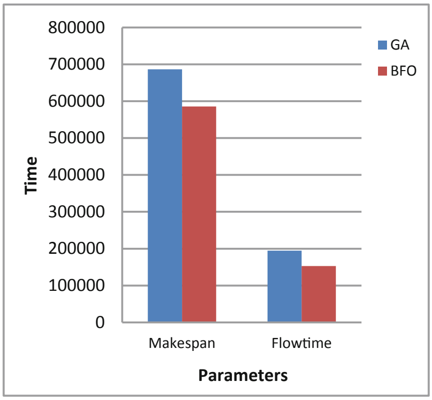


Fig. 1. Comparison on (100, 128) set

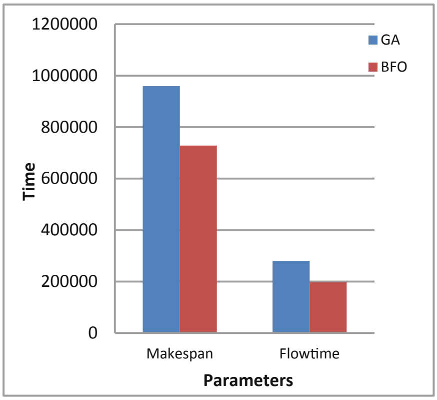


Fig. 2. Comparison on (300, 128) set

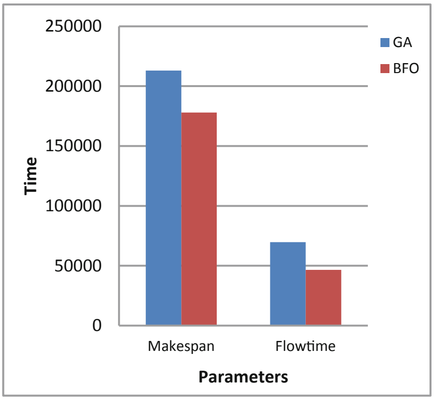


Fig. 3. Comparison on (500, 128) set

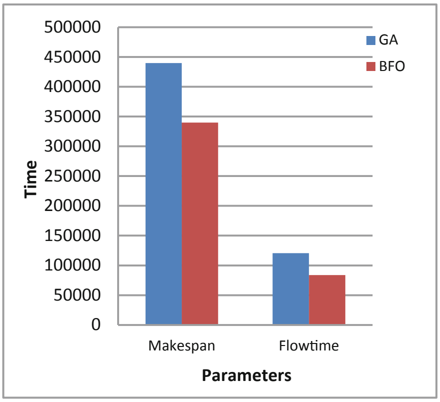


Fig. 4. Comparison on (700, 128) set

After considering the average values it was noted that there is 16.14% and 17.81% improvement in the values of makespan and flowtime respectively for 96 processors. For 112 processors, an improvement of 18.07%, and 22.99% was seen. And lastly for 128 processors, the makespan and flowtime improved by 22.39% and 27.85%. The improvement in 128 processors is the highest which implies that more the number of processors, lesser is the time taken for execution.

In the end it is safe to say that the proposed algorithm has performed better than the existing technique in all the cases. We have successfully shown that the makespan and flowtime are effectively optimized by BFOA better than the Genetic Algorithm for every processor- job configuration (Table 8).

Table 8. Improvement (in %)

Parameter	96 processors	112 processors	128 processors
Makespan	16.14	18.07	22.39
Flowtime	17.81	22.99	27.85

6 Conclusion

Optimizing multiple metrics concurrently is a difficult task especially when they are inter-related. However, multi-criteria optimization is essential to improve the performance of the scheduling techniques. During studying the literature, we noticed that GA suffers from the problem of local optima. Also, it has slow convergence and pre-mature convergence issues while evaluating optimistic values. The performance of the multi-cluster environment is optimized by applying a novel meta-heuristic technique named Bacterial Foraging Optimization Algorithm. Initially we tuned the parameters of GA to find the most optimal values and used those for further calculations. Real workload traces were used for evaluating makespan, flowtime and waiting time using 3 types of processors within 5 clusters. 96, 112 and 128 processors were used evaluation on 100, 300, 500 and 700 jobs. In this we have also considered the bi-objective function which referred as normalization function that works on two parameters i.e. makespan and flowtime.

For 96 processors, we saw 16.14% and 17.81% improvement in the values of makespan and flowtime respectively. For 112 processors, an improvement of 18.07% and 22.99% was seen. And lastly for 128 processors, the makespan and flowtime were improved by 22.39% and 27.85% respectively.

The improvement in 128 processors is the highest which implies that more the number of processors, lesser is the time taken for execution. Hence it can be concluded that BFOA performs better than GA in all the cases be it any processor-job configuration. Bacterial Foraging Optimization however still does have the tendency to get trapped in local minima. So its hybridization with other algorithms will be considered in the future. The study of new optimization metrics could also be considered. The tuning of user-defined parameters will also be done. It can be considered for energy-aware and communication-aware scheduling as well.

References

1. Gabaldon, E., Lerida, J.L., Guirado, F., Planes, J.: Multi-criteria genetic algorithm applied to scheduling in multi-cluster environments. J. Simul. **9**, 1–9 (2015)
2. Roberge, V., Tarbouchi, M., Labonté, G.: Comparison of parallel genetic algorithm and particle swarm optimization for real-time UAV path planning. IEEE Trans. Ind. Inform. **9**(1), 132–141 (2013)
3. Jones, W.M., Ligon III, W.B., Pang, L.W.: Characterisation of band-width aware meta-schedulers for co-allocating jobs across multiple clusters. J. Supercomput. **34**, 135–163 (2005)

4. Gabaldon, E., et al.: Blacklist mufti-objective genetic algorithm for energy saving in heterogeneous environments. J. Supercomput. **73**, 1–16 (2016)
5. Dasgupta, K., et al.: A genetic algorithm (GA) based load balancing strategy for cloud computing. Procedia Technol. **10**, 340–347 (2013)
6. Kalra, M., Singh, S.: A review of metaheuristic scheduling techniques in cloud computing. Egypt. Inform. J. **16**(3), 275–295 (2015)
7. Enomoto, T., Masaomi, K.: Improving population diversity in parallelization of a real-coded genetic algorithm using MapReduce (2014)
8. Chana, I.: Bacterial foraging based hyper-heuristic for resource scheduling in grid computing. Future Gener. Comput. Syst. **29**(3), 751–762 (2013)
9. Yang, C., et al.: Bacterial foraging optimization using novel chemotaxis and conjugation strategies. Inf. Sci. **363**, 72–95 (2016)
10. Garg, R., Mittal, S.: Optimization by genetic algorithm. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **4**(4), 587–589 (2014)
11. Jun, L., Jian-wu, D., Feng, B.: Analysis and improvement of bacterial foraging optimization algorithm. J. Comput. Sci. Eng. **3**(1), 1–7 (2014)
12. Sharma, V., Pattnaik, S.S., Garg, T.: A review of bacterial foraging optimization and its application. In: National Conference on Future Aspects of Artificial intelligence in Industrial Automation, NCFAAIIA 2012 (2012)

Performance Analysis of Different Fractal Image Compression Techniques

Rupali Balpande(✉) and Atish Khobragade

Department of Electronics Engineering,
Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India
rupalibalpande81@gmail.com,
atish_khobragade@rediffmail.com

Abstract. The FIC has the disadvantage of high computational cost. This paper outlines the comparison of different encoding methods to reduce computational complexity while retaining the quality of the image is retrieved. To increase the PSNR of full search method (BFIC), EP-NRS method is introduced in which image is partitioned into range and domain blocks of similar edge property. Then they are mapped to lowest DCT coefficient in a vertical and horizontal direction into 2D coordinate System. In another method new FIC scheme is proposed based on the fact that affine similarity between two blocks is equivalent to the absolute value of Pearson's correlation coefficient (APCC) between them. In comparing to the original technique, the APCC based method gave number of MSE computations less, high PSNR value and high compression ratio in image quality which is acceptable.

Keywords: FIC (Fractal image compression) ·
PIFS (Partition iterated function system) · DCT (Discrete cosine transform) ·
PCC (Pearson correlation coefficient) ·
EPNRS (Edge based-Neighborhood region Search Method)

1 Introduction

The fractal image encoding is the technique which based on the self similarity property within the image to improve the compression. In fractal image encoding no. of computations required are larger due to full search mechanism therefore to reduced the encoding time or the better quality of reconstructed image many schemes are used. The main motivations are behind image compression – encoding speed and data storage. Fractal image compression is a lossy compression technique which was first proposed in 1985 by Barnsley and Demko operating from IFS (Iterated Function System). Fractal image compression was first introduced practically in 1992 by Jacquin [14]. The main disadvantages of all previous method were that they required more encoding time.

Jacquin [14] proposed a full search fractal image encoding method, in which image is partitioned into non-overlapping range blocks and overlapping domain blocks. To search the best Domain block for each range by finding the optimal affine transformation. But this required more time. Furao and Hasegawa [15] develop an encoding method in which for best matching domain block direct assignment of specific domain

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 208–216, 2019.
https://doi.org/10.1007/978-981-15-0108-1_20

pool as optimal matching one. This method speedup the encoding time and gives high compression ratio with poor quality of retrieved image. In Fisher 72 classes scheme which is variance based block sorting scheme in which an image is divided into 72 classes. This method is more efficient in encoding but it is difficult to arrange 72 parts. Wu et al [21] develop a method based on standard deviation called fast fractal image compression. Tong [19] proposed the same method fast fractal image encoding but base on adaptive search. In Wang et al [20] paper used a fitting plane to present a modified gray level transform. As the modified gray level transform reduce the minimum matching error. It gives the high quality of reconstructed image. Another approach like stochastic optimization [1, 2] method such as Genetic algorithm (GA) can be used to speed up the encoder but does not preserve the image quality. Some GA-based [3, 4] algorithms are proposed to improve the efficiency.

Table 1. The 8 Affine transformations

T_0	T_1	T_2	T_3	T_4	T_5	T_6	T_7
$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

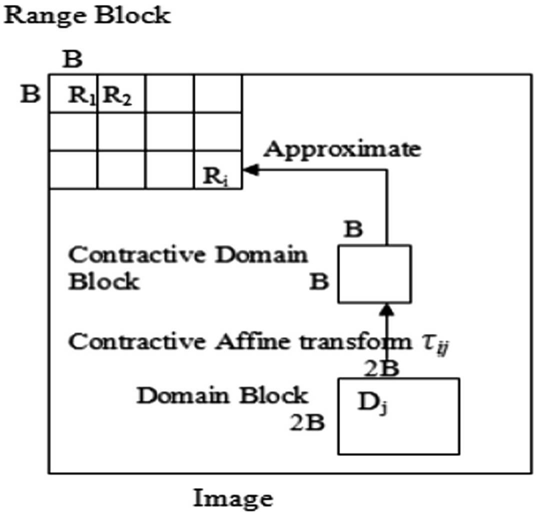
In this paper different methods are used to reduce search space in FIC. In first method edge based Neighborhood region search method using DCT is used to speed up the fractal encoder by gathering image blocks with similar edges in some specific region. It has advantages of relatively high compression ratios and improvement in reconstructed image quality which gives good fidelity [8]. In another method Fisher's classification method is used to classify the partitioned domain pool and range pool blocks into 3 classes. In the meanwhile sorting of domain blocks using preset block by calculating APCC between them [5]. Therefore encoding time is further reduced. For performance analysis of Fractal image encoding algorithms are performed on images using two different methods including the full search method, Edge based-Neighborhood region Search Method and APCC based method.

2 Fractal Image Encoding

Fractal image compression has been proposed various variants and extensions so far. In this paper we consider PIFS which is based on partitioning the original image into smaller and larger pieces. Given an image in the rectangular array of pixels with size (m × m). The original image can be partitioned arbitrarily; however, a natural choice would be to use rectangular blocks. To simplify the partitioning we choose to use square blocks. Thus the underlying image can be partitioned into (n × n) square blocks which are called as range blocks. The lager blocks called as domain blocks. It is seen that the non-overlapping range blocks cover the entire image, whereas the domain blocks may overlap and they are two times greater than range block. Thus in FIC method usually an image is divided in two blocks basically the range and domain block.

The number of range blocks in an image is calculated using equation $(m/n)^2$ from the original gray level image f of size (m × m) which makes up range pool R of all non-overlapping blocks of size (n x n) of the image f. And the no of domain blocks is calculated using $(m - 2n + 1)^2$ which formed the set of all possible blocks of size (2n × 2n) of image f called as domain pool D. Then fractal affine transformation from Table 1 is constructed by searching all of domain block from domain pool D for each range block v from R [5].

The most important step in implementing fractal image compression is searching best similar domain block for every range block over the entire image. A good match for a given range block is obtained if more domain blocks are there. For best range-domain block pair, the next step is to apply an appropriate transformation on the underlying domain block to obtain an approximation for the range block. In FIC the size of domain blocks are reduced to the range block size using contractive transformation as shown below:



The whole D blocks are contracted into blocks of same size as range blocks by averaging four pixel values to one pixel value. The new extended domain pool is obtained by 8 affine transformation operation as shown in Table 1. For each range block in an image, to search best matching domain block from the entire extended domain pool is to minimize following equation of mean square error [5]

$$\mathrm{MSE}(R, sd + o1) = \|R - (sd + o1)\|_2 \tag{1}$$

Where $\| . \|_2$ is the two-norm, s controls the contrast, o controls the brightness and I is constant vector with the values of 1 in all elements. In fractal image encoding above combination of (s, o, and index of d in the domain pool) formed the PIFS subsystem of R [5].

3 Edge Based Neighborhood Region Search Method [8]

Proposed method shows Edge based-Neighbourhood region Search Method to speed up the fractal encoder and which improves the quality of reconstructed image. Instead of block matching, we are only interested in edge [8]. In this method formation of frequency domain two dimensional coordinate system using DCT method & the entire range and domain block into this system is mapped. Then we find the lowest vertical F (1, 0) and lowest horizontal F(0, 1) DCT coefficients [8]. We know the two dimensional DCT of an image f (i, j) of size (N × N) defined by

$$F(m,n) = \tfrac{2}{N} C_m C_n \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i,j) \cos\left(\tfrac{(2i+1)m\pi}{2N}\right) \cos\left(\tfrac{(2j+1)n\pi}{2N}\right),$$

$$\text{where } m,n = 0,1,\ldots,N-1 \text{ and}$$

$$\mathrm{C}_k = \begin{cases} 1/\sqrt{2}, & \text{if } k = 0 \\ 1, & \text{otherwise.} \end{cases}$$

$$\text{Typically, for } N = 8, \text{we have}$$

$$F(1,0) = \tfrac{\sqrt{2}}{8} \sum_{i=0}^{7} \sum_{j=0}^{7} f(i,j) \cos\theta_i \quad \text{and}$$

$$F(1,0) = \tfrac{\sqrt{2}}{8} \sum_{i=0}^{7} \sum_{j=0}^{7} f(i,j) \cos\theta_j$$

Where F(1, 0) and F(0, 1) are lowest vertical and lowest horizontal DCT coefficients [8].

In FIC to find the best domain block, similarity measurement is done between each range block with all 8 affine transformed blocks of domain block. By using this method number of Computations are reduces by embedding the edge property of blocks into neighbourhood region method. Initially we have to divide the whole image into range and domain block by applying PIFS function on the image. Then apply DCT on these blocks to find the coefficients F(1, 0) and F(0, 1). Then we use affine transformations T0 to T7 which are explained in introduction. In this method these transformations are divided into two categories i.e. T0 to T3 of Table 1 in which T0 = T and the relation between transformed block DCT coefficients of other 3 to the original block can be calculated [8].

$$\begin{cases} F_1(1,0) = F(1,0) \\ F_1(0,1) = -F(0,1), \end{cases} \begin{cases} F_2(1,0) = -F(1,0) \\ F_2(0,1) = F(0,1), \end{cases} \begin{cases} F_3(1,0) = -F(1,0) \\ F_3(0,1) = -F(0,1), \end{cases}$$

Above transformed blocks will have the same directional edge since their Fi(1, 0) and Fi(0, 1) are same. After transformation using DCT the transformed blocks coefficients of other 3 to another is T4 to T7 is following [8].

$$\begin{cases} F_4(1,0) = F(0,1) \\ F_4(0,1) = F(1,0), \end{cases} \begin{cases} F_5(1,0) = F(0,1) \\ F_5(0,1) = -F(1,0), \end{cases}$$

$$\begin{cases} F_6(1,0) = -F(0,1) \\ F_6(0,1) = F(1,0), \end{cases} \begin{cases} F_7(1,0) = -F(0,1) \\ F_7(0,1) = -F(1,0) \end{cases}$$

Above 4 transformed blocks will also have the same directional edge since their |Fi (1, 0)| and |Fi(0, 1)| are the same. Thus above Image blocks with same directional edges will be determined in specific two regions Ω1 and Ω2 in coordinate system. If the domain block is from the region Ω1 then for similarity comparison of a given range block vj of the same region only the four affine transformations fk: k = 0, 1, 2, 3 of the domain block is needed. Similarly in the Ω2 region, the vj also need only to perform the similarity comparison with the four affine transformations fk: k = 4, 5, 6, 7 of the domain block. Therefore by searching similar directional edge in specific region, only four MSE computations required using this method instead of eight transformations. Therefore the purpose of speedup can be increase by reducing the number of computations further two times.

4 APCC Based Block Classification Using FISHER'S Scheme [5]

In 1992, Fisher proposed classification scheme of image blocks using 3 classes given in equation below. In this scheme domain pool is reduced with classifying the domain blocks into Fisher's 3 classes to speed up encoding [5]. First step is subdivision of square block image into upper left, upper right, lower left and lower right quadrants, [16] by computing the average pixel intensities and the corresponding variances of each quadrant [17]. Then in second step possible orientation (rotate and flip) of the block in such a way that the average intensities are ordered using canonical orderings in one of the following three ways [5]:

$$\mathrm{a1} \geq \mathrm{a2} \geq \mathrm{a3} \geq \mathrm{a4} \text{ - - - - - - Class 1} \tag{2}$$

$$\mathrm{a1} \geq \mathrm{a2} \geq \mathrm{a4} \geq \mathrm{a3} \text{ - - - - - - Class 2} \tag{3}$$

$$\mathrm{a1} \geq \mathrm{a4} \geq \mathrm{a2} \geq \mathrm{a3} \text{ - - - - - - Class 3} \tag{4}$$

Where a1, a2, a3, a4 are sum of luminance from upper left to lower right of its four sub-blocks. Pearson correlation is calculated between range and domain block of same class. Best matching domain block is searched in the same class in which range blocks classified to reduce the pair wise comparisons further.

In image compression correlation is used for comparison of image to specify the strength of the block's relationship. PCC between two variables is calculated between pixels of image blocks and its absolute value has range of −1 to +1. PCC is defined as covariance of two variables to the product of standard deviation of those two variables. PCC can be calculated as below

$$\rho_{\chi,y} = \frac{cov(x,y)}{\delta_{\chi}\delta_{y}} \tag{5}$$

Where, cov(x, y) is covariance and $\delta_{\chi}\delta_{y}$ is standard deviation of the two variables.

4.1 APCC Based Block Sorting Using Trained Preset Block [5]

In Fractal image encoding to search best domain from domain pool it is necessary to sort each domain class independently after domain block classification as given above section using Fisher's scheme. In APCC based FIC domain blocks are sorted with respect to APCC between range and domain blocks. If Pearson's correlation coefficient absolute value between each range block R and domain block D is approximate to 1 means that domain block D which satisfy this condition $|\rho(\mathbf{R},\ \mathbf{D})\ | \rightarrow 1$ is the approximate domain block for that range block [5]. But it is difficult to search the best approximate domain block **D** therefore it is necessary that using Eq. 5 compute the APCC between each domain block and the preset block **B** and select a proper preset block B. As domain blocks classified into 3 classes, there is need 3 preset blocks. For searching matched Domain block in a set of domain blocks for each range block, there is need of approximation between the APCCs of these domain blocks and preset block and APCC between range block **R** and preset block **B.** Then for a range Block **R** of the each class should be calculated APCC with preset block B. For a preset block **B**, we first calculate the PCC between R and B then between D and B to search the matching domain block for a current range block **R** meeting the Eq. (6) in this domain class [5] (Fig. 1).

$$|\rho(\mathbf{R},\mathbf{B})| \approx |\rho(\mathbf{D},\mathbf{B})| \tag{6}$$

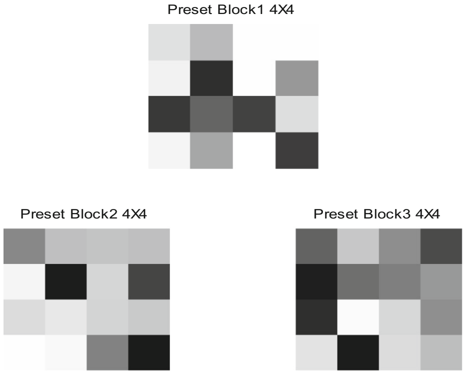


Fig. 1. The 4 × 4 preset blocks trained for cameraman image

5 Experimental Results

In these experimental results the various methods conducted on various images (256 × 256, 512 × 512, 8 bit gray scale) and results are obtained. Figure 2 shows the preset blocks trained for 4 × 4 image blocks of cameraman image corresponding to the classes expressed in Eqs. (2), (3) and (4). Table 2 lists the performance analysis of FIC using the edge property-based neighborhood region search method and absolute value of Pearson's correlation coefficient methods including full search method. The tested images are Lena and cameraman. Here EP-NRS, APCC and FIC are the abbreviations of the edge property-based neighborhood region search method, absolute value of Pearson's correlation coefficient method and baseline fractal image compression method respectively. Experimentation has been performed using MATLAB software on WINDOWS platform. Hardware requirement of proposed approach is general minimum configuration, Pentium III Processor and 256 MB RAM (Figs. 3 and 4).

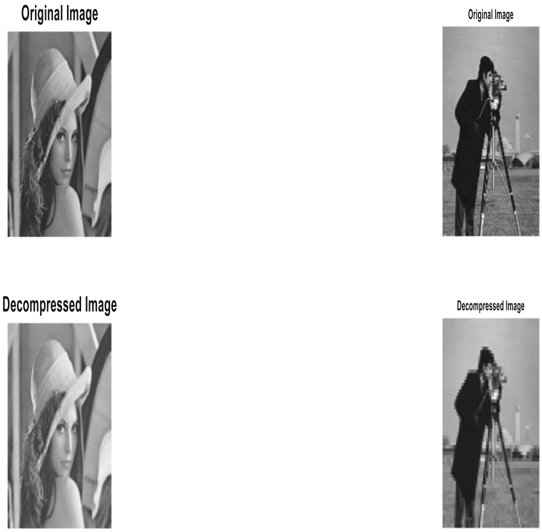


Fig. 2. The original 512 × 512 and 256 × 256 image and decompressed image using EP-NRS method

Table 2. Performance comparison

Image	APCC method			EP-NRS method			BFIC method		
	PSNR	MSE	Number of computations	PSNR	MSE	Number of computations	PSNR	MSE	Number of computations
Lena (512 × 512)	43.54 dB	8.56	53248	38.81 dB	2.88	169998	34.28 dB	24.26	184185
Cameraman (256 × 256)	36.08 dB	16.02	13312	36.03 dB	16.21	41740	34.89 dB	21.08	149152

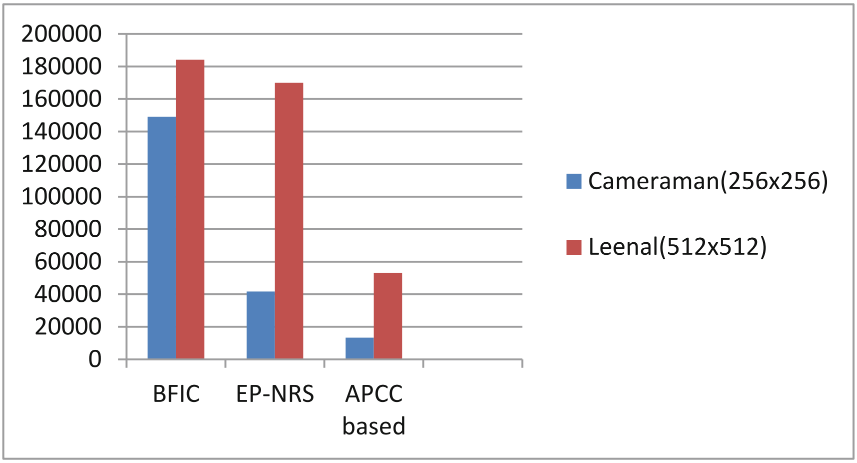


Fig. 3. Number of computations with different methods

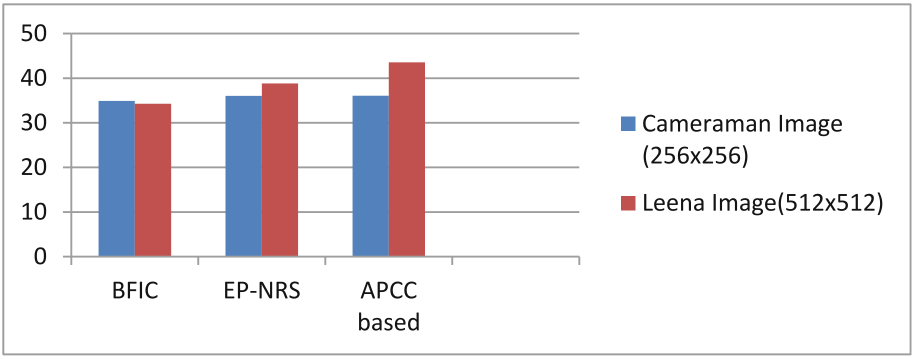


Fig. 4. PSNR values with different methods

6 Conclusion

The FIC has drawback of high computational complexity. To overcome this, I have compared different technique of FIC as shown results. Using Edge based neighborhood region method in which speedup rate can be increases further by reducing the no. of computations by using edge property of block into the search process. Using this method PSNR and compression ratio is also increases as compared with other methods and reconstructed image quality is also improved. The proposed algorithm improved the PSNR as compared to basic method fractal image compression. In this compression scheme main concept is reproduction of image using Iterated Function Systems (IFS). This Iterated Function System can be further optimized using Pearson Correlation coefficient. Numbers of domain blocks are reduce by the APCC method, in which absolute value of Pearson's correlation coefficient of the blocks is calculated. The comparison of these schemes for 512 × 512 images Lena and 256 × 256 Cameraman image with range block size 4 shown in Table 2.

References

1. Distasi, R., Nappi, M., Riccio, D.: A range/domain approximation error-based approach for fractal image compression. IEEE Trans. Image Process. **15**(I), 89–97 (2006)
2. Mitra, K., Murthy, C.A., Kundu, M.K.: Technique for fractal image compression using genetic algorithm. IEEE Trans. Image Process. **7**, 586–593 (1998)
3. Barnsley, M.F., Jacquin, A.E.: Application of recurrent iterated function systems to images. In: Proceedings of SPIE, vol. 1001, pp. 122–131, November 1988
4. Wohlberg, B., de Jager, G.: A review of the fractal image coding literature. IEEE Trans. Image Process. **8**(12), 1716–1729 (1999)
5. Wang, J., Zheng, N.: A novel fractal image compression scheme with block classification and sorting based Pearson's correlation coefficient. IEEE Trans. Image Process. **22**(9), 3690–3702 (2013)
6. He, C., Xu, X., Li, G.: Improvement of fast algorithm based on correlation coefficients for fractal image encoding. Comput. Simul. **12**(4), 60–63 (2005)
7. Wang, J., Liu, Y., Wei, P., Tian, Z., Li, Y., Zheng, N.: Fractal image coding using SSIM. In: Proceedings of 18th ICIP, pp. 245– 248, September 2011
8. Lin, Y.-L., Wu, M.-S.: An edge property-based neighborhood region search strategy for fractal image compression. Department of Information Engineering, I-Shou University, Kaohsiung, Taiwan Elevier (2011)
9. Roy, S.K., Kumar, S., Chanda, B., Chaudhuri, B.B., Banerjee, S.: Fractal image compression using upper bound on scaling parameter. Elsevier Ltd. (2017)
10. Tseng, C.C., Hsieh, J.G.: Fractal image compression using visual-based particle swarm optimization. Image Vis. Comput. **26**, 1154–1162 (2008)
11. Subramanian, P., Indumathi, R.: Fractal image compression technique. Int. J. Comput. Organ. Trends **4** (2014)
12. Truong, T.-K.: A fast encoding algorithm for fractal image compression using DCT inner product. IEEE Trans. Image Process. **9**(4), 529–535 (2000)
13. Barnsley, M.F.: Fractal Everywhere. Academic Press, New York (1993)
14. Jacquin, A.E.: Image coding based on a fractal theory of iterated contractive image transformations. IEEE Trans. Image Process. **1**, 18–30 (1992)
15. Furao, S., Hasegawa, O.: A fast no search fractal image coding method. Signal Process.: Image Commun. **19**(5), 393–404 (2004)
16. Fisher, Y.: Fractal Image Compression: Theory Application. Springer-Verlag, Berlin (1995). https://doi.org/10.1007/978-1-4612-2472-3
17. He, C., Yang, S., Huang, X.: Variance-based accelerating scheme for fractal image encoding. Electron. Lett. **40**(2), 1052–1053 (2004)
18. Wang, X.Y., Wang, Y.X., Yun, J.J.: An improved no-search fractal image coding method based on a fitting plane. Image Vis. Comput. **28**(8), 1303–1308 (2010)
19. Tong, C.S., Pi, M.: Fast fractal image encoding based on adaptive search. IEEE Trans. Image Process. **10**(9), 1269–1277 (2001)
20. Wang, X.Y., Wang, S.G.: An improved no-search fractal image coding method based on a modified gray-level transform. Comput. Graph. **32**(4), 445–450 (2008)
21. Wu, X.W., Jackson, D.J., Chen, H.C.: A fast fractal image encoding method based on intelligent search of standard deviation. Comput. Electr. Eng. **31**(6), 402–421 (2005)

Empirical Aspect to Extract Hidden Features from Unattended Portion of Neuroimaging Modalities for Early Detection of Diseases

Yogesh Kumar Gupta(✉)

Department of Computer Science, Banasthali Vidyapith,
Newai, Rajasthan, India
gyogesh@banasthali.in

Abstract. In this world of digitized era there is immense desideratum of image processing which manipulates the image information for analysis purport. Any image can be characterized by its feature sets, which demonstrate the detailed information of the images and enhance their visual interpretation. The features are extracted corresponding to the whole image comprised of both ROI (region of interest) and Non-ROI regions. Additionally, the region of Non-ROI shows promising results for obnubilated information corresponding to expected disease. So, for that purport we have proposed a method that extract the feature from the ROI and the Non-ROI region corresponding to the mark vertices in an image utilizing mouse click. The proposed work utilizes a model that extracted the categorical region by culling the concrete area and marks the vertices, double click at the last vertices utilizing mouse, the area automatically extracted corresponding to the masking that is discretely exhibited. The proposed work uses two neuroimaging modalities such as CT-Scan and MRI. As per the survey of radiologist, Medico and medicos, most of the time they concentrated only those components of the image i.e. ROI that can be facilely detectable and shows the cause of disease. They have to leave the rest of the component of the image i.e. Non-ROI that can be cause of the further detection of disease. In furthermore we extracted the obnubilated consequential information from the unattended portion of the neuroimaging modalities i.e. Non-ROI.

Keywords: ROI · Non-ROI · Neuroimaging modalities · Hidden features

1 Introduction

Image processing is a method to acquire enhanced images by converting it into digital form, in order to extract some utilizable or obnubilated information from it by performing many computations on an image. The main objective of image processing is to visualize the image objects, image retrieval to find ROI, image renovation and sharpening to engender a better image, apperception or distinguish image objects etc [12]. The final result of the image processing is a numerical data rather than an image. Image processing includes the following stages.

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 217–226, 2019.
https://doi.org/10.1007/978-981-15-0108-1_21

(1) Image acquisition- in this stage enhanced images are acquired and stored in a media for further processing. There are some techniques used to convert analogy data into a digital form [13].

(2) Pre-processing- This is the most essential step of image processing in which enhance the quality of acquired images by noise filtering, enrich edge, incrementing contrast, pseudo-coloring, resizing an image, sharpening the regions and equalize the frequency level utilizing histogram equalization [14].

(3) Segmentation- in this stage digital image divided into multiple segments and each segment contained sundry features such as color, special location, texture, intensity and many other statistical properties. The main purport of segmentation is to ascertain a concrete region of the image that is called region of interest (ROI) [4, 6].

(4) Feature Cull- The main objective of this step is to eliminate the entire extraneous and duplicate feature and cull those features or attribute that are more germane and which reduce the computation time and amend the precision of prognostication that formulate the better or optimized results.

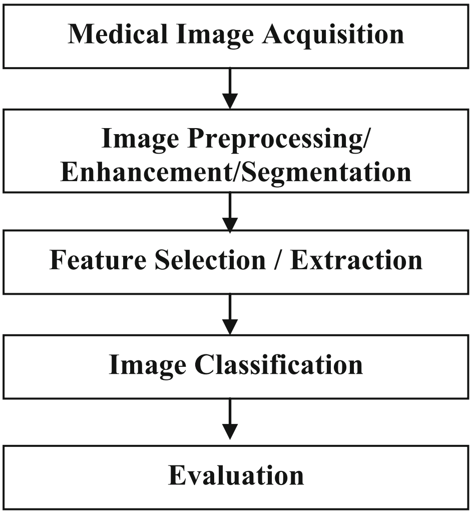

Medical Image Acquisition
Image Preprocessing/ Enhancement/Segmentation
Feature Selection / Extraction
Image Classification
Evaluation

Fig. 1. Model of image processing

(5) Feature Extraction- There are sundry items which provide the information of an image is kenned as features e.g. texture, size, color, shape, composition, location etc [15].

(6) Image Relegation- In this stage extracts the cognizance or information which avails radiologist or medico for decision making. There are sundry kinds of relegation methods such as texture relegation, neural network relegation, data mining relegation etc. which uses machine learning, virtualization, statistical methods, and other manipulation techniques to retrieve the information [1].

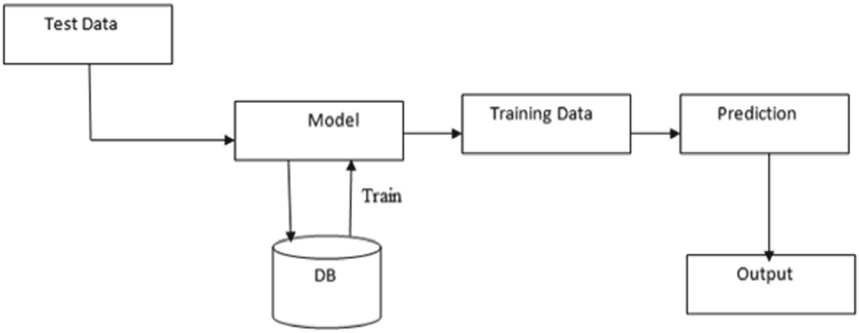


Fig. 2. Classification process model [1]

In this stage there are two types of datasets; training dataset and testing dataset. The training dataset stored in the backend database and input dataset which is called test dataset match with the data stored in database for the relegation and then we got the trained datasets as an outcome that is further analyze and prognosticated for the final results [18].

(7) Evolution- In this stage evaluates the relegated data on the behalf of some statistical parameters that shows the precision factor of the final outcome.

2 Neuroimaging Modalities

There are sundry types of Invasive and Non-Invasive Imaging Modalities such as MRI, CT-Scan, X-Ray, USG etc. but in our research work we have discussed about only CT-Scan and MRI images [5].

2.1 Magnetic Resonance Imaging (MRI)

To record the internal structure and some portion of function in the body, a non-invasive medical imaging modality such as MRI is significantly utilized. MRI is a spectroscopic imaging technique predicated on nuclear magnetic resonance principle, uses non-ionizing electromagnetic radiation. The main advantage of MRI in medical is that it is mainly used to provide in depth analytic imaging of soft tissues in the body such as soft organs and cartilage tissues like encephalon and the heart [16]. MRI is mundanely painless and not uses radiations so MRI scans generally considered safe in gravidity and for children. MRI shows some unique information those other techniques not capable to show and has the great impact for cell tracking and breast tissue regeneration. On the other hand MRI does have many drawbacks such as it has very tight space and an inordinate amount of strepitous. For some people it may be claustrophobic and quite sumptuous. MRI is not for intra-luminal abnormalities and there is desideratum of sedation to little children because they can't remain still [3]. Some sample images of MRI are shown in the Fig. 3.

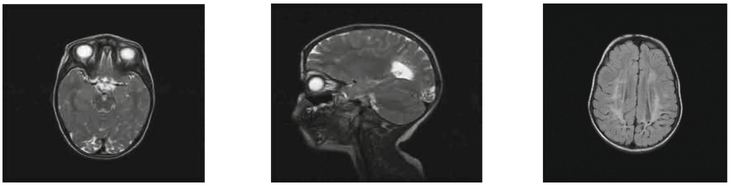


Fig. 3. Brain structure MRI

2.2 Computed Tomography (CT-Scan)

CT-Scan is a diagnostic technology consists of a rotating frame which has two components; one is an X-Ray tube at one side and a detector at antithesis side of the frame. An image is acquired each time when the X-Ray tube and detector consummates one round the patient body and many images will be accumulated from many angles. Every profile of X-Ray beam is reconstructed by the computer to engender 2D image and then scanned. 3D CT-Scan can withal acquire with the avail of spiral CT which is subsidiary in visualization of tumor in three dimensions. Recently 4D CT-Scan introduced to surmount with the quandary of respiratory forms of kineticism. It engenders temporal and spatial information about the organs. It is painless and non invasive method to diagnose the medical quandaries. It captures the picture of bones, soft tissues and blood vessels in our body [7]. CT-Scan additionally provide detailed information about very minute abnormalities even the body don't have the symptoms of it. CT-Scan releases relatively high radiations which has the jeopardy of lung and breast cancer and mainly not recommended for enceinte women and has health issue for unborn babies and fetuses [3]. Some sample images of CT-Scan are shown in the Fig. 4.

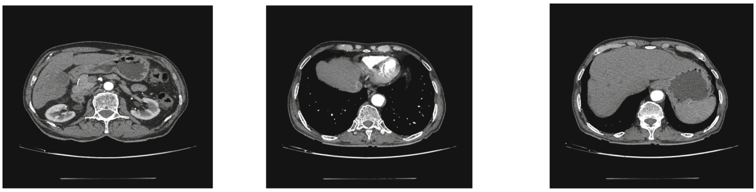


Fig. 4. CT-Scan images showing internal body structures

3 Research Methodology

The first step of our research work is to acquire several medical imaging modalities data by visiting sundry hospitals, clinics and diagnostic centers. For our work, we have accumulated primary data, near about 6000 medical images from two different cities such as Jaipur and Kota. The Neuroimaging Modalities such as MRI and CT-Scan taken from “GETWELL Healthcare Center”, Jaipur and “Bharat Vikas Parishad Hospital and Research Center”, Kota, Rajasthan. The amassed data will be in any hard drive media like pen drive, DVD, CD etc. All the amassed medical image data was stored in own database environment. The size and color of the amassed data is not categorical, depending on the priority they have culled for the imaging. So, for image processing we require to perform some alteration such as resizing and converting the image into grayscale. The size should be 256 × 256. The intensity of grayscale image varies from 0 to 255 where 0 approaches to consummately ebony and 255 approaches to consummately white. We have accumulated images in DICOM format but stored in database in JPEG format.

3.1 Experimental Setup

We have shown the experimental setup for the medical image processing, the research work uses the hardware and software as follows:-

Hardware: Intel® core™ i3-3210 CPU 540 @3.07 GHz, 4.00 GB RAM, 500 GB HDD. Software: MATLAB R13a, SQL database server.

3.2 Proposed Method to Extract Hidden Features from Unattended Portion of Neuroimaging Modalities

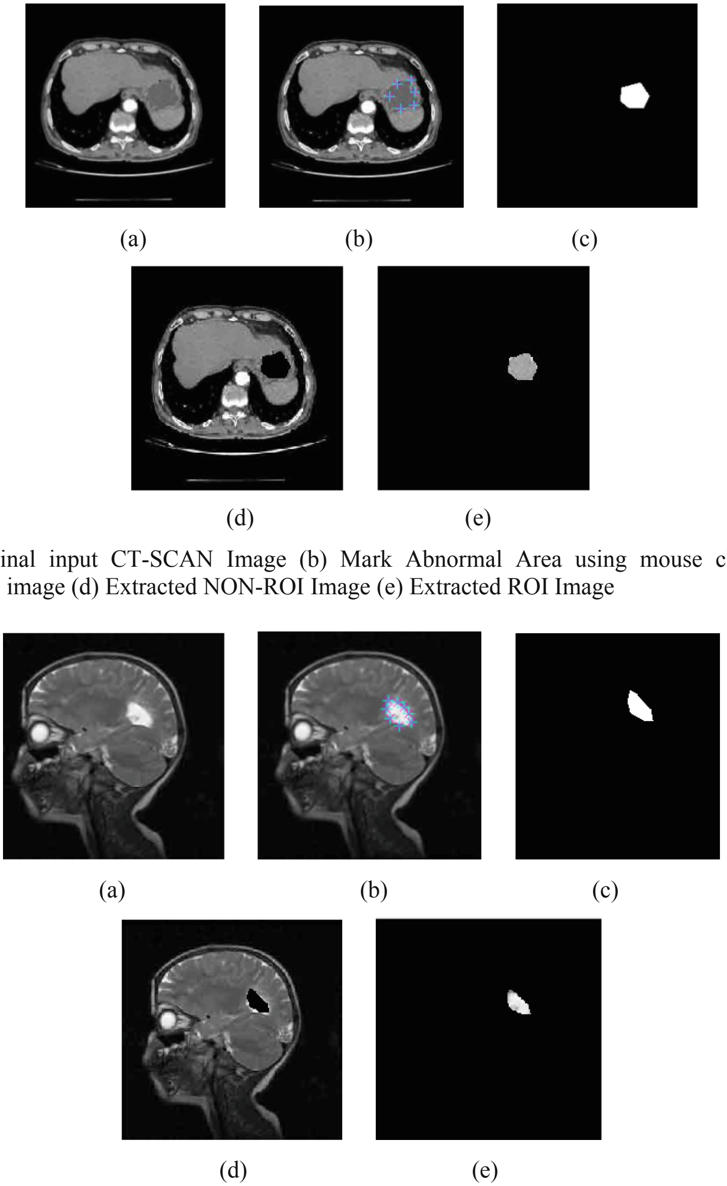
The features are extracted corresponding to the whole image comprised of both ROI and Non-ROI regions. Withal, the region of Non-ROI shows promising results for obnubilated information corresponding to expected disease [9]. So, for that purport we have proposed a model that extracts the features from the ROI and the Non-ROI region corresponding to the mark vertices in an image utilizing mouse click. The proposed work utilizes a model that extracted the concrete region by culling the categorical area and marks the vertices, double click at the last vertices utilizing mouse, the area automatically extracted corresponding to the masking that is discretely exhibited.

Proposed Algorithm

```
Step 1   Designate polygonal region of interest (ROI) for the input image.
Step 2   Create buffer for ROI with 256 * 256.
Step 3   Create buffer for Non-ROI with 256 * 256.
Step 4   Performs the binary masking corresponding to loop, to disunite the ROI and
         Non-ROI region.
Step 5   Display the ROI and Non-ROI images discretely
```

4 Results and Discussion

The results of proposed model as shown in the Fig. 5.



(a) Original input CT-SCAN Image (b) Mark Abnormal Area using mouse click (c) Masking image (d) Extracted NON-ROI Image (e) Extracted ROI Image

(a)Original input MRI Image (b) Mark Abnormal Area using mouse click (c) Masking image (d) Extracted NON-ROI Image (e) Extracted ROI Image

Fig. 5. Extracted ROI and Non-ROI region from Brain CT-Scan and MRI image

The results of both CT and MRI images shows the ROI and Non-ROI region that avails the radiologist to extract the ROI region discretely on mouse click in very efficaciously manner and expeditiously. As per the survey of radiologist, Physicians and doctors, most of the time they concentrate only those parts of the image i.e. ROI that is easily detectable and shows the cause of disease. They have to leave the rest of the part of the image i.e. Non-ROI that can be cause of the further detection of disease. In further more we extracted the hidden meaningful informations from the unattended portion of the medical images i.e. Non-ROI.

The input images are firstly process using the Wavelet Transform with daubechies level four (db4) to obtain the approximated image that holds high intensity values and entire detail band coefficient of pixels of image. There after extract Non-ROI region from the medical images and then apply the model of GLCM (Gray Level Co-occurrence Matrix) and Wavelet Transform to extract the categorical features from the extracted Non-ROI region of the medical images [17, 19].

Now further we are going to elaborate the concepts and formulas of these features:

1. Maximum Probability: Maximum value represents the largest value of PI in matrix.
2. Minimum Probability: Minimum value represents the smallest value of PI in matrix [19].
3. Mean: An average value. $\mu = \frac{1}{N} \sum_{i=1}^{N} A_i$

   Where Ai denotes the ith observed value of the attribute. N is the number of data points, we have in our data set.

4. Median: Middle value from the image matrix set.
5. Skewness:- Third moment of a real-valued random variable of X. The skewness can be positive skewed, negative skewed and unskewed depending on whether Skew(X) is positive, Negative or zero [18].
6. Kurtosis:- Fourth moment of a real-valued random variable of X. It is always Non-negative, in fact strongly positive unless X is a constant.
7. Contrast: It measures variation between pixel and its adjoining pixel in terms of gray scale change. Contrast can be calculated with the help of the following formula-

$$\mathrm{Contra} = \sum \mathrm{a, b}\, |\mathrm{a} - \mathrm{b}| 2\, \mathrm{Pi}\,(\mathrm{a, b})$$

   Where Pi (a, b) represents pixel at position (a, b).

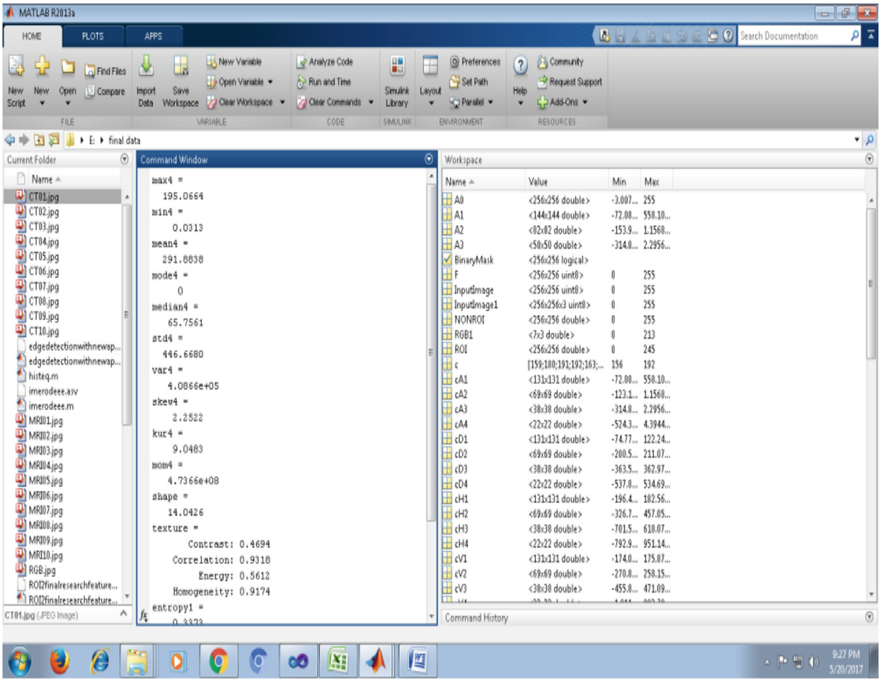
8. Correlation:- It is define the strength of relationship between two variable suppose X and Y [8].
9. Energy: It returns the sum of squared elements in the GLCM. Energy is 1 for a constant image.

$$\mathrm{Energy} = \sum \mathrm{a, b\, Pi}\,(\mathrm{a, b})\, 2$$

10. Homogeneity: It measures changes in gray values. If there are large variation in gray values then homogeneity will also be large and vice-versa [17].

$$\text{Homog} = \sum i,j \frac{\text{Pi}(a,b)}{1 + |a - b|}$$

Now, the output of the M-file program for 4th level decomposition and feature extraction utilizing DWT and GLCM are shown in the Fig. 6.



```
max4 =
  195.0664
min4 =
    0.0313
mean4 =
  291.8838
mode4 =
     0
median4 =
   65.7561
std4 =
  446.6680
var4 =
   4.0866e+05
skew4 =
    2.2522
kur4 =
    9.0483
mom4 =
   4.7366e+08
shape =
   14.0426
texture =
       Contrast: 0.4694
    Correlation: 0.9318
         Energy: 0.5612
    Homogeneity: 0.9174
entropy1 =
```

```
Name ▲         Value                     Min        Max
A0             <256x256 double>          -3.007...  255
A1             <144x144 double>          -72.08...  558.10...
A2             <82x82 double>            -153.9...  1.1568...
A3             <50x50 double>            -314.8...  2.2956...
BinaryMask     <256x256 logical>
F              <256x256 uint8>           0          255
InputImage     <256x256 uint8>           0          255
InputImage1    <256x256x3 uint8>         0          255
NONROI         <256x256 double>          0          255
RGB1           <7x3 double>              0          213
ROI            <256x256 double>          0          245
c              [159;180;191;192;163;...  156        192
cA1            <131x131 double>          -72.08...  558.10...
cA2            <69x69 double>            -123.1...  1.1568...
cA3            <38x38 double>            -314.8...  2.2956...
cA4            <22x22 double>            -524.3...  4.3944...
cD1            <131x131 double>          -74.77...  122.24...
cD2            <69x69 double>            -200.5...  211.07...
cD3            <38x38 double>            -363.5...  362.97...
cD4            <22x22 double>            -537.8...  534.69...
cH1            <131x131 double>          -196.4...  182.56...
cH2            <69x69 double>            -326.7...  457.05...
cH3            <38x38 double>            -701.5...  618.07...
cH4            <22x22 double>            -792.9...  951.14...
cV1            <131x131 double>          -174.0...  175.87...
cV2            <69x69 double>            -270.8...  258.15...
cV3            <38x38 double>            -455.8...  471.09...
```

Fig. 6. Extracted feature using proposed work in MATLAB R13a environment

The extracted features from the Non-ROI part of the medical images that utilizing GLCM and Wavelet transform are shown in the Table 1.

Table 1. Extracted features from the Neuroimaging Modalities

Feature no.	Feature name	Feature value				
		Figure 1	Figure 2	Figure 3	Figure 4	Figure 5
1	Max	191.1	195.06	254.94	254.59	254.59
2	Min	3.0313	0.0313	1.96	0.78	0.78
3	Mean	340.71	291.88	1110.5	524.29	751.04
4	Median	82.1	65.75	939.28	247.86	604.04
5	Skewness	1.02	2.25	0.57	2.7	0.8171
6	Kurtosis	2.94	9.04	2.09	11.09	2.35
7	Contrast	0.3463	0.4694	0.6514	0.8479	0.4737
8	Correlation	0.9384	0.9318	0.9097	0.919	0.913
9	Energy	0.4605	0.5612	0.398	0.5435	0.3854
10	Homogenity	0.9238	0.9174	0.9136	0.9148	0.8985

These features extracted from the Neuroimaging Modalities which are more germane and that avail the radiologist and medico to identify the cause of the disease.

5 Conclusion

During our research work we have proposed a ROI and Non-ROI extraction model that avails the radiologists and medicos to extract the ROI region discretely on mouse click in a very efficaciously manner and expeditiously. As per the survey of medical persons, Medico and medicos, most of the time they concentrated only those components of the image i.e. ROI that can be facilely detectable and shows the cause of disease. They have to leave the rest of the component of the image i.e. Non-ROI that can be cause of the further detection of disease. Furthermore we extracted the obnubilated paramount information from the unattended portion of the medical images i.e. Non-ROI, that avail the radiologist for early detection of the diseases.

References

1. Gupta, Y.K., Jha, C.K.: Study of big data with medical imaging communication. In: International Conferences on Communication and Computing Systems (ICCCS), pp. 993–997. CRC Press/Taylor & Francis Group (2016)
2. Gupta, Y.K., Jha, C.K.: A review on the study of big data with comparison of various storage and computing tools and their relative capabilities. Int. J. Invocation Eng. Technol. (IJIET) **7**(1), 470–477 (2016)
3. Dhabhai, A., Gupta, Y.K.: Empirical study of image classification techniques to classify the ımage using SVM: a review. IJIRCCE **4**(10), 17128–17133 (2016)
4. Dubay, S., Gupta, Y.K., Soni, D.: Comparative study of various segmentation techniques with their effective parameters. IJIRCCE **4**(10), 17223–17228 (2016)
5. Dubay, S., Gupta, Y.K., Soni, D.: Role of big data in healthcare with non-invasive and minimal-invasive medical imaging modality. IJIRCCE **5**(3), 5322–5331 (2017)
6. Singh, P., Chadha, S.R.: A novel approach to image segmentation. IJARCSSE **3** (2013). ISSN 2277-128X
7. Tamilselvan, K.S., et al.: A novel image segmentation algorithm for clinical CT ımages using wavelet transform, curvelet transform and multiple kernel FCM. Appl. Math. Sci. **9** (2015). https://doi.org/10.12988/ams.2015.53216. ISSN 2351-2362
8. Chandrakala, M., Durgadevi, P.: Threshold based segmentation using block processing. IJIRCCE **4** (2016). ISSN 2320-9801
9. Sambasivarao, C., Naganjaneyulu, V.: An efficient boundary detection and image segmentation method based on perceptual organization. IJCTT **7** (2014). ISSN 2231-2803
10. Singh, P., Singh, A.: A study on image segmentation technique. IJRTER **2**(2016). ISSN 2455-1457
11. Yogamangalam, R., Karthikeyan, B.: Segmentation techniques comparison in image processing. IJET **5** (2013). ISSN 0975-4024
12. Saif, J.A.M., Al-Kubati, A.A.M., Hazaa, A.S., Al-Moraish, M.: Image segmentation using edge detection and thresholding. In: ACIT (2012). ISSN 1812-0857
13. Abubakar, F.M.: Study of image segmentation using thresholding technique on a noisy image. IJSR **2** (2013). ISSN 2319-7064

14. Manikannan, A., SenthilMurugan, J.: A comparative study about region based and model based using segmentation techniques. IJIRCCE **3** (2015). ISSN 2320-9801
15. Kumari, R., Sharma, N.: A study on the different image segmentation technique. IJEIT **4**, 284–289 (2013)
16. Eqbal, S., Ansari, A.M.: Medical image feature extraction for computer aided diagnosis of lung cancer. IJARCSSE **5** (2015). ISSN 2277-128X
17. Manojbhai, D.D., Rajamenakshi, R.: Large scale image feature extraction from the medical image analysis. IJAERS **3** (2016). ISSN 2349-6495
18. Rajaei, A., Rangarajan, L.: Wavelet feature extraction for medical image classification. IJES **4** (2011). ISSN 2229-6913
19. Rinky, B.P., Mondal, P.K., Manikantan, K., Ramachandran, S.: DWT based feature extraction using edge tracked scale normalization for enhanced face recognition. ICCCS Procedia Technol. **6**, 344–353 (2012)

Deciphering the Association of Single Amino Acid Variations with Dermatological Diseases Applying Machine Learning Techniques

Jaishree Meena, Aparna Chauhan, and Yasha Hasija(✉)

Department of Biotechnology, Delhi Technological University,
Shahbad Daulatpur, Main Bawana Road, Delhi 110042, India
yashahasija06@gmail.com

Abstract. SAVDerma predicts the association of a Single Amino Acid Variation (SAVs) with dermatological diseases. SAVs are basically non-synonymous SNPs which are sometimes associated with various diseases. A single amino acid mutation can affect various physico-chemical properties of a protein which then effect the normal functionality of a particular protein and leads to disease. Studies have shown that among various mendelian diseases, around 60% of them are due to these types of SAVs. The scope of this study limits with dermatological disorders and the process can be applied to all other types of diseases. We have curated SAV's physico-chemical and sequence-based features which are associated with dermatological disorders and kept them as positive set and vice versa as negative set. This data sets were feed to machine learning classifiers for developing a model which can classify SAV's association with dermatological disorders. Our classifier obtained an accuracy of 87.29%. SAVDerma is a web application where user can find all the curated and machine generated data about SAVs and their predicted association with skin diseases. It has a user-friendly interface where information regarding particular SAV can be retrieved using three different parameters i.e. Gene symbol, Swissprot id and rs id and are provided on the search page of the website (http://savderma.info/). SAVDerma can be used as a single point information retrieval system by clinicians and thus increasing the current knowledge on skin disorders.

.

Keywords: SAVs (Single amino acid variations) · SNPs (Single nucleotide polymorphism) · Machine learning (ML) · SAVDerma

1 Introduction

Skin is always a fascinating organ of human body having distinct roles of protection, sensation, thermoregulation and most interesting metabolism of Vitamin D. It provides body, a complex barrier against external environment which comprises of dermis, subdermal structures, surface keratinocytes and inter-keratinocyte substances. Socio-economic scarcity is one of the major causes of skin diseases which adversely affects

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 227–238, 2019.
https://doi.org/10.1007/978-981-15-0108-1_22

the lifestyle of the patients. Skin disorders are commonly found in different societies influencing for about thirty to seventy percent of population of every age group [1–3]. The International Classification of Disease identifies skin disorders as the most prevalent human disease, recording more than thousand cases, representing a heavy load of skin diseases [4]. However, skin diseases are the most ignored and least discussed in health debate in India. [4]. An incomparable magnitude of data about amino acid substitutions has been shaped using genomic profiling tools like genotyping arrays and next-generation sequencing. Recent data from 1000 Genomes Project has revealed that more than 1.5 crore Single Nucleotide Polymorphisms (SNPs) and approximately 0.3 crore insertions and deletions alongwith twenty thousand structural variants exist in human genome and are still growing in number day by day [5–7]. It is evaluated that there are three to five million single amino acid variations exist in human as specified by the ongoing sequencing statistics of all-inclusive human genome. SAVs are the most enormous type of SNPs that principally generate protein products which contain amino acid substitutions [8]. From different studies on protein's structure and function, it has been observed that some amino acid substitutions may cause pernicious diseases and are responsible for about sixty percent of Mendelian diseases while other are totally harmless having no role in type of disease [9]. Information obtained from studies on SAVs can give insights about the migration pattern of long-gone human and ancestor line of existing humans. Moreover, the findings on SAVs can be utilized to establish the foundation for their use in therapeutic practices by decoding the effect of genomic variations and by setting up the connections with phenotypes [10].

To predict the functional influence of SAVs on protein, many computational approaches like machine learning or statistical methods have been used in past decade which employs protein related features like physico-chemical properties, 3-D structure, amino acid sequence, intricate residue-contact network property, evolutionary evidence and more. The dominant part of these strategies are being utilized by investigators which can be connected as autonomous programming or web applications to outfit unrestricted calculation of useful and functional impact of SAVs for informative and non-business reasons [8]. In the current research paper, information associated with skin diseases have been recovered through different web assets and literature accessible on the web and all the SAVs, neutral along with disease-linked were mined then the curated data were fed to various machine learning classifiers for training and a model was prepared to calculate the linkage of SAVs with skin diseases, the model was applied on the test data. Based on the prediction results, a web application named "SAVDerma: Single Amino Acid Variations related to Dermatological disorders" is developed. SAVDerma consists of data connected to SAVs and tells about the association of an SAV with dermatological disorders. SAVDerma caters the need of research community especially in the area of healthcare. Data of SAVDerma can be used to make easy to use and efficient diagnostic tools for skin disorders. In addition, anticipated unique SAVs whose role in skin disorders have not been perceived, but rather they possibly have role in skin diseases can be additionally researched. This investigation will be useful in unraveling the effect of SAVs in dermatological diseases and in giving machine learning based analytic frameworks for clinical progressions.

2 Methodology

2.1 Extraction of SAVs

Skin disease associated genes were curated from literature and online web resources, and mapped against Uniport ids. Present study focused on two groups of single amino acid variations (SAVs), first group consists of skin disease causing variations in amino acids and second group consists of neutral polymorphisms (having no role in skin diseases) for making positive and negative cases in this study for machine learning purpose [11]. By using HumsaVar.txt [12], Uniport ids were then drawn against 4279 instances of skin disease causing variations and 4895 instances of neutral variations. The protein sequences of all the known human proteins were also extracted from Uniprot and the sequences for specific amino acid substitutions have been extracted by altering the specific amino acid residues at particular positions as demonstrated by variation data from HumsaVar.

2.2 Examination of SAVs

Extracting Physico-Chemical Properties of SAVs

The physico-chemical features were retrieved through PROFEAT which is a web based platform to retrieve features of proteins and peptides on the basis of amino acid sequence. It provides information about amino acid constituents and the physico-chemical properties like hydrophobicity, polarity, solvent accessibility etc. which are exceedingly helpful [13] for the efficient application of statistical learning approaches in predicting essential structural organisation, practically functional communication profiles of proteins irrespective of sequence resemblances.

Extracting Properties Related to Mutational Effect of SAVs

In several human disorders, it is observed that protein aggregation complements or alters the pathology of disease with which it is associated. Mutations has an important role in protein aggregation, so we employed TANGO which performs various calculations on aggregation propensity of unfolded peptide chains which are useful to assess the effect of mutation in protein aggregation. TANGO is hinged on the physico-chemical values associated with generation of beta-sheet, drawn-out by the confidence that the core areas of an aggregated protein are totally buried [14, 15]. LIMBO which is a position specific algorithm was used to distinguish loci of chaperone binding in proteins. LIMBO is based on a PSSM matrix which is made proficient from in vitro peptide binding data and structure modelling [16]. It is highly effective in foreseeing the mutational effect on chaperone binding and their role in disease transformation. To retrieve information of mutational consequences of SAVs on amyloid propensity of protein, a program called WALTZ has been selected. WALTZ is a locus-sensitive prediction algorithm that takes amino acid position, physico-chemical and structural information of protein as input and generates a scoring matrix for protein aggregation propensity. This properties are taken into consideration for assessment of aggregating regions as an effect of SAVs [17]. Lastly, Grantham score is recovered which provides the information of the distance between two amino acid in evolutionary sense and can be utilized to predict harmful and neutral amino acid substitutions. The distance scores for amino acid substitutions go from 5 to 215 in Grantham matrix [18].

2.3 Prediction of SAVs Association with Dermatological Disorders Using Machine Learning

To train and test data related to SAVs, Waikato Environment for Knowledge Analysis (Weka) is used, it is an accessible, user-friendly platform for machine learning tasks. The state-of-the art methods of machine learning like random forest, J48 tree, multilayer perceptron classifier, classification via regression etc. were utilised for classification purpose [19] and evaluated to discover the top performing classifier in terms of accuracy, precision, recall, area under ROC curve and F-measure for prediction by using 10-fold cross validation.

2.4 Construction of SAVDerma

SAVDerma was built utilizing WordPress and Caspio Bridge. SAVDerma incorporates all the data of SAVs retrieved from various databases used in the study. WordPress is an open source, Content Management System (CMS), which is developed to support online blogs and sites. Caspio Bridge is a user-friendly cloud software and dais which helps in building compliant Web databases and structures. Caspio Bridge is useful platform for non-programmers for developing databases recourses both back-end and front-end. It can also integrate with WordPress for wide accessibility through web browsers.

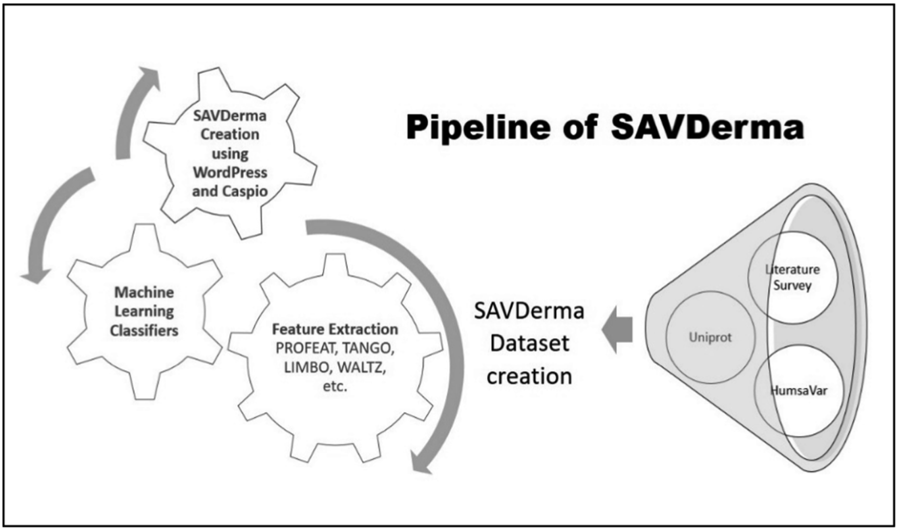

Pipeline of SAVDerma
SAVDerma Creation using WordPress and Caspio
Machine Learning Classifiers
Feature Extraction PROFEAT, TANGO, LIMBO, WALTZ, etc.
SAVDerma Dataset creation
Literature Survey
Uniprot
HumsaVar

Fig. 1. Pipeline of SAVDerma

3 Results and Discussion

Physico-chemical and mutation associated data on SAVs were collected to make training data for machine learning purpose through various tools like PROFEAT, TANGO, LIMBO, WALTZ etc. To make test data, unclassified SAVs, which are not known to be associated with any kind of skin disorder were retrieved from HumsaVar and tested to check their relation with skin diseases.

3.1 Optimum Classifier Selection

Various machine learning classifiers were used to select a best performing optimum classifier on basis of accuracy, MCC value, f-measure, recall, precision and ROC curve. Random forest classifier was selected as optimum machine learning classifier on the basis of obtained results and was used for testing the relationship of unclassified SAVs with skin disorders. It showed the highest accuracy (87.29%) and precision (87.40%) in prediction tasks. Furthermore, it provides outstanding performance based on ROC area (93.40%) amongst other machine learning classifiers, refer Figs. 1, 2 and Table 1.

Table 1. Results achieved for various ML classifiers used in the study.

ML classifiers	Accuracy	Precision	Recall	F measure	MCC	ROC area
Random forest	87.29	87.40	87.30	87.30	74.30	93.40
J48	82.23	82.40	82.20	82.30	64.20	82.80
Function logistic	76.65	76.70	76.70	76.70	52.70	84.50
JRip	80.80	80.80	80.80	80.80	61.00	82.30
Filtered classifier	82.89	83.00	82.90	82.90	65.50	86.70
Classification via regression	84.92	85.00	84.90	85.00	69.60	91.00
Multilayer perceptron	84.86	85.00	84.90	84.90	69.40	90.40

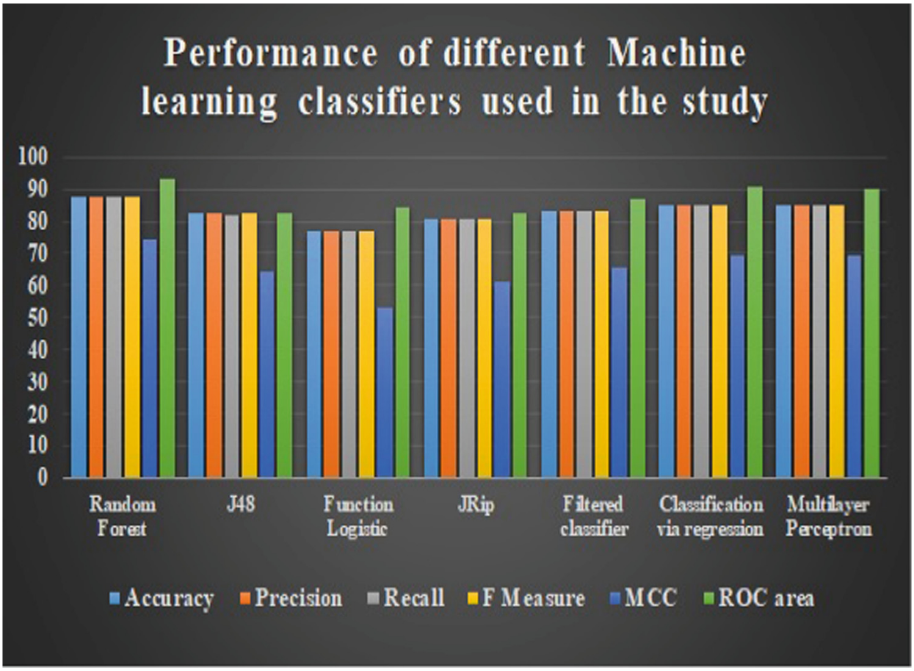


Fig. 2. Bar graph displaying various ML classifiers against their performance.

3.2 Testing of Different Types of Datasets

To test the discrete performance of different tools and SAVDerma, Random forest ML classifier was used and the performance of every tool alongwith SAVDerma was compared as shown in the Table 2. SAVDerma comes out to be the top performer in terms of accuracy, ROC area, precision and other machine learning performance measures. SAVDerma shows the highest accuracy of 87.29% and precision of 87.40%, moreover, it showed ROC area of 93.40% which represents the eminence of binary measure. All these performance results prove that SAVDerma gathers information of high quality. SAVDerma contains 57389 unclassified SAVs which are present in 10204 genes. Out of 57389 SAVs, 5788 SAVs are predicted to be related with dermatological diseases located in 1712 genes in which 77 SAVs are found to have highest prediction margin.

Table 2. Performance of different tools against SAVDerma.

—	Accuracy	Precision	Recall	F measure	MCC	ROC area
PROFEAT	81.59	81.60	81.60	81.60	62.70	81.40
TANGO	57.72	56.70	57.70	53.80	10.50	55.70
WALTZ	57.23	56.40	57.20	50.20	08.40	52.00
LIMBO	56.74	55.30	56.70	51.10	07.40	51.90
Grantham	62.00	61.60	62.00	60.60	21.20	64.70
SAVDerma	87.29	87.40	87.30	87.30	74.30	93.40

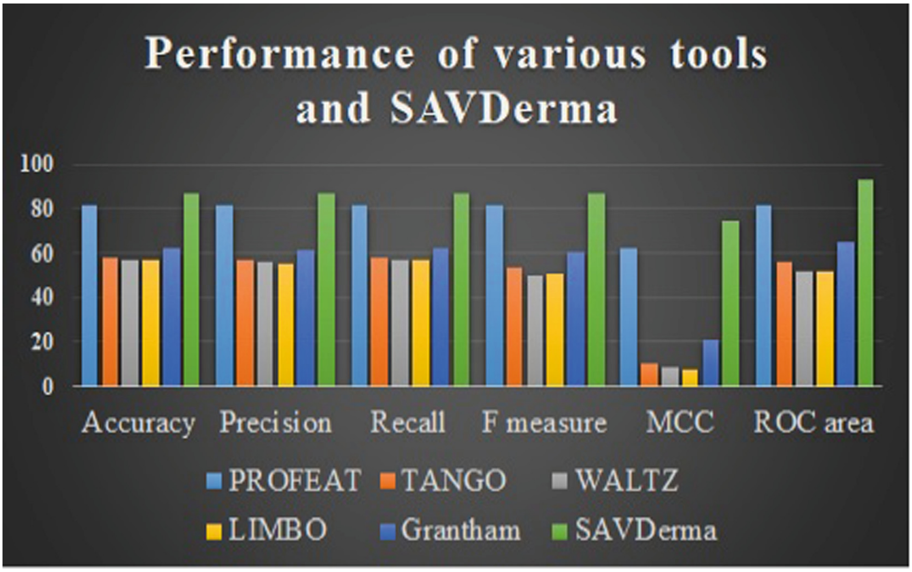
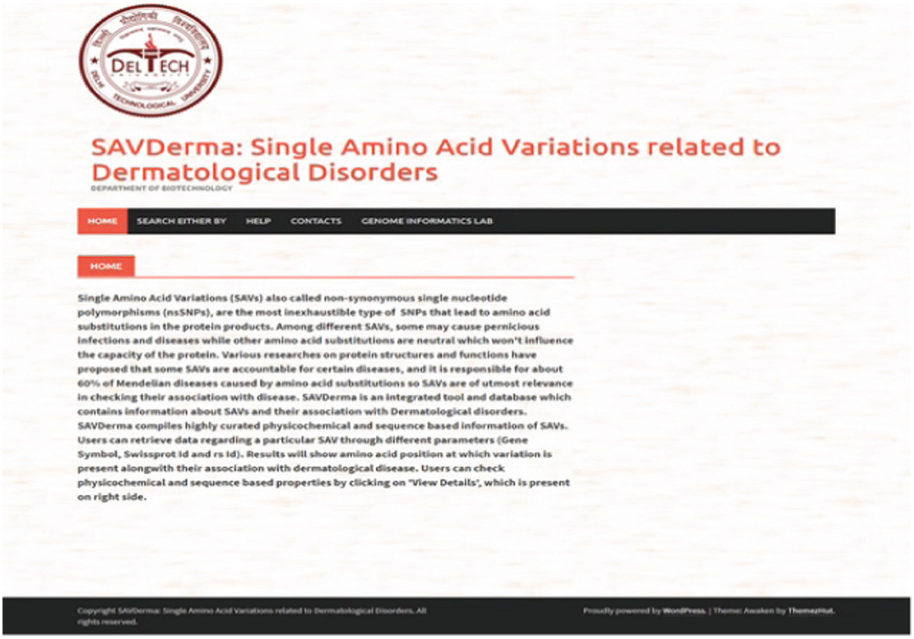


Fig. 3. Bar graph displaying performance of various tools and SAVDerma.

3.3 User Interface Development of SAVDerma

SAVDerma is a web application which holds data about SAVs and their connection with skin conditions. SAVDerma assembles profoundly curated physico-chemical and features of SAVs based on sequence properties. Users can assess data with respect to a specific SAV through parameters like Gene Symbol, Swissprot Id and rs Id. Search outcomes will display amino acid position at which substitution exist along with their relationship with dermatological disease (refer Figs. 3, 4, 5, 6 and 7).



SAVDerma: Single Amino Acid Variations related to Dermatological Disorders

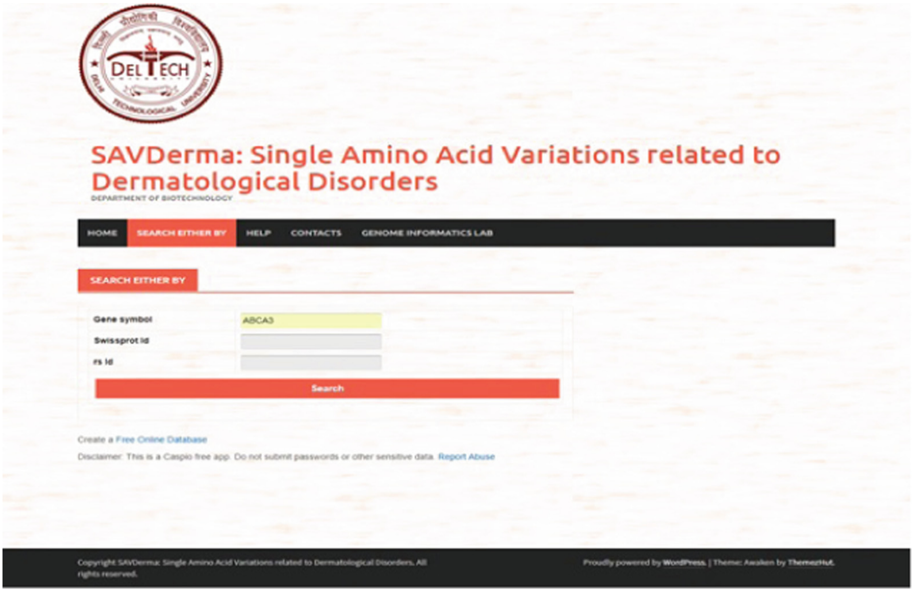
DEPARTMENT OF BIOTECHNOLOGY

Single Amino Acid Variations (SAVs) also called non-synonymous single nucleotide polymorphisms (nsSNPs), are the most inexhaustible type of SNPs that lead to amino acid substitutions in the protein products. Among different SAVs, some may cause pernicious infections and diseases while other amino acid substitutions are neutral which won't influence the capacity of the protein. Various researches on protein structures and functions have proposed that some SAVs are accountable for certain diseases, and it is responsible for about 60% of Mendelian diseases caused by amino acid substitutions so SAVs are of utmost relevance in checking their association with disease. SAVDerma is an integrated tool and database which contains information about SAVs and their association with Dermatological disorders. SAVDerma compiles highly curated physicochemical and sequence based information of SAVs. Users can retrieve data regarding a particular SAV through different parameters (Gene Symbol, Swissprot Id and rs Id). Results will show amino acid position at which variation is present alongwith their association with dermatological disease. Users can check physicochemical and sequence based properties by clicking on "View Details", which is present on right side.

Copyright SAVDerma: Single Amino Acid Variations related to Dermatological Disorders. All rights reserved.

Proudly powered by WordPress. | Theme: Awaken by ThemezHut.

Fig. 4. Homepage of SAVDerma.



SAVDerma: Single Amino Acid Variations related to Dermatological Disorders

DEPARTMENT OF BIOTECHNOLOGY

Gene symbol ABCA3

Swissprot Id

rs Id

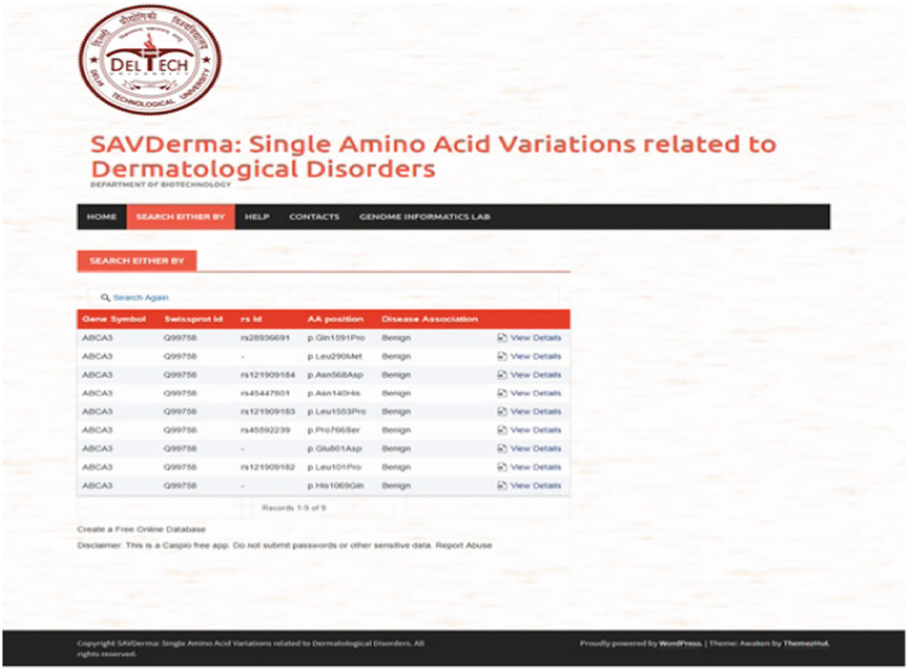
Create a Free Online Database

Disclaimer: This is a Caspio free app. Do not submit passwords or other sensitive data. Report Abuse

Copyright SAVDerma: Single Amino Acid Variations related to Dermatological Disorders. All rights reserved.

Proudly powered by WordPress. | Theme: Awaken by ThemezHut.

Fig. 5. Search page of SAVDerma.



SAVDerma: Single Amino Acid Variations related to Dermatological Disorders

DEPARTMENT OF BIOTECHNOLOGY

SEARCH EITHER BY

Gene Symbol	Swissprot Id	rs Id	AA position	Disease Association	
ABCA3	Q99758	rs28936691	p.Gln1591Pro	Benign	View Details
ABCA3	Q99758	-	p.Leu290Met	Benign	View Details
ABCA3	Q99758	rs121909184	p.Asn568Asp	Benign	View Details
ABCA3	Q99758	rs45447801	p.Asn140His	Benign	View Details
ABCA3	Q99758	rs121909183	p.Leu1553Pro	Benign	View Details
ABCA3	Q99758	rs45592239	p.Pro766Ser	Benign	View Details
ABCA3	Q99758	-	p.Glu801Asp	Benign	View Details
ABCA3	Q99758	rs121909182	p.Leu101Pro	Benign	View Details
ABCA3	Q99758	-	p.His1069Gln	Benign	View Details

Records 1-9 of 9

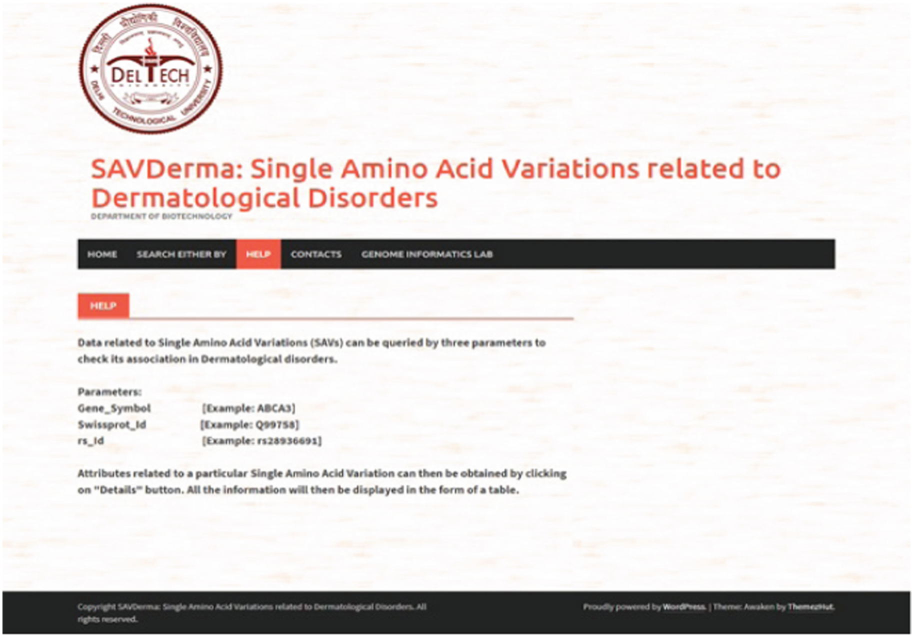
Create a Free Online Database

Disclaimer: This is a Caspio free app. Do not submit passwords or other sensitive data. Report Abuse

Copyright SAVDerma: Single Amino Acid Variations related to Dermatological Disorders. All rights reserved.

Proudly powered by WordPress. | Theme: Awaken by ThemezHut.

Fig. 6. Result page of SAVDerma after searching for a particular query.



SAVDerma: Single Amino Acid Variations related to Dermatological Disorders

DEPARTMENT OF BIOTECHNOLOGY

HELP

Data related to Single Amino Acid Variations (SAVs) can be queried by three parameters to check its association in Dermatological disorders.

Parameters:	
Gene_Symbol	[Example: ABCA3]
Swissprot_Id	[Example: Q99758]
rs_Id	[Example: rs28936691]

Attributes related to a particular Single Amino Acid Variation can then be obtained by clicking on "Details" button. All the information will then be displayed in the form of a table.

Copyright SAVDerma: Single Amino Acid Variations related to Dermatological Disorders. All rights reserved.

Proudly powered by WordPress. | Theme: Awaken by ThemezHut.

Fig. 7. Help page of SAVDerma.

Table 3. Analysis of putative SAVs

Rs ids	Top scoring genes	Proposed role in skin diseases
rs122458138, rs753267653	ACSL4	ACSL4 gives directions to make a protein that forestalls minerals, including calcium, from being kept in body tissues where they don't have a place. Any change in ACSL4 can prompt cole infection
rs72559751, rs35404804	ABCC9	Any alteration in ABCC9 may modify the structure of the potassium channel which may prompts to skin disorder called Cantú syndrome
rs777832794, rs387906592, rs794728021, rs397516685, rs121434528, rs121434527, rs112602953, rs751300489, rs121434526	ACTA2	Any modification to this gene can cause allergic reactions due to impairment of protein found in various types of immune system cells leading to urticaria
rs886043807, rs72549398, rs72549399, rs756220860, rs72549395, rs777469571, rs72549400, rs72551306, rs72551305, rs121908935, rs72549394, rs72551307, rs11568372, rs72549402, rs11568372	ABCB11	Modifications in ABCB11 prompts disabled secretion of bile juice which can cause serious pruritus and yellowing of the skin

4 Conclusion and Discussion

Dermatological diseases are one of the regular conditions that influences human and their way of life. A collection of computational techniques has been created to gauge the functional impact of SAVs on a protein and their association with disease utilizing statistical methods or machine learning approaches. In this study, using the machine learning methods, we tried to make predictions on relationship of SAVs with dermatological diseases and found many SAVs associated with skin disease that are not previously known in any kind of skin disorder (refer Table 3). On-premise of the findings, we have made SAVDerma, which encompasses more than fifty-two thousand SAVs and their linked sequence-based and physicochemical properties. SAVDerma is first of its kind user-friendly web application for the SAVs linked to skin diseases. This will provide biologically substantial information related with SAVs to users, which will undoubtedly working to give understandings about the cause of the skin conditions. We proposed to refresh an ever increasing number of information related with amino acid variations in coming time. The information about novel single amino acid variations that we got from ML methods can be furthermore used to examine their association with skin diseases and can be used to make effectual drugs for management of skin related conditions. Additionally, this investigation will assist researchers and scholars in finding improved machine learning based diagnostic frameworks for fast and unswerving diagnosis of skin diseases.

Acknowledgement. This study has been supported by Department of Biotechnology, Government of India, project file No. BT/PR5402/BID/7/408/2012.

Availability and Implementation. The developed SAVDerma is available as free online resource at [http://savderma.info/].

References

1. Bickers, D.R., et al.: The burden of skin diseases: 2004. A joint project of the American Academy of Dermatology Association and the Society for Investigative Dermatology. J. Am. Acad. Dermatol. **55**, 490–500 (2006)
2. Hay, R.J., Fuller, L.C.: The assessment of dermatological needs in resource-poor regions. Int. J. Dermatol. **50**, 552–557 (2011)
3. Schofield, J., Grindlay, D., Williams, H.: Skin conditions in the UK: a health care needs assessment (2009)
4. Hay, R.J., et al.: The global burden of skin disease in 2010: an analysis of the prevalence and impact of skin conditions. J. Invest. Dermatol. **134**(6), 1527–1534 (2014)
5. Sachidanandam, R., et al.: A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature **409**, 928 (2001)
6. International HapMap Consortium: A second generation human haplotype map of over 3.1 million SNPs. Nature **449**, 851 (2007)
7. Durbin, R.M., et al.: A map of human genome variation from population-scale sequencing. Nature **467**, 1061 (2010)
8. Wang, M., et al.: FunSAV: predicting the functional effect of single amino acid variants using a two-stage random forest model. PLoS ONE **7**(8), e43847 (2012)
9. Botstein, D., Risch, N.: Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nature Genet. **33**, 228 (2003)
10. Fernald, G.H., Capriotti, E., Daneshjou, R., Karczewski, K.J., Altman, R.B.: Bioinformatics challenges for personalized medicine. Bioinformatics **27**, 1741–1748 (2011)
11. Srivastava, I., Gahlot, L.K., Khurana, P., Hasija, Y.: DbAARD & AGP: a computational pipeline for the prediction of genes associated with age related disorders. J. Biomed. Inform. **60**, 153–161 (2016)
12. Famiglietti, M.L., et al.: Genetic variations and diseases in UniProtKB/Swiss-Prot: the ins and outs of expert manual curation. Hum. Mutat. **35**, 927–935 (2014)
13. Rao, H.B., Zhu, F., Yang, G.B., Li, Z.R., Chen, Y.Z.: Update of PROFEAT: a web server for computing structural and physicochemical features of proteins and peptides from amino acid sequence. Nucl. Acids Res. **39**(Suppl. 2), 32–37 (2011)
14. De Baets, G., Van Doorn, L., Rousseau, F., Schymkowitz, J.: Increased aggregation is more frequently associated to human disease-associated mutations than to neutral polymorphisms. PLoS Comput. Biol. **11**, e1004374 (2015)
15. Fernandez-Escamilla, A.M., Rousseau, F., Schymkowitz, J., Serrano, L.: Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat. Biotechnol. **22**, 1302 (2004)

16. Van Durme, J., Maurer-Stroh, S., Gallardo, R., Wilkinson, H., Rousseau, F., Schymkowitz, J.: Accurate prediction of DnaK-peptide binding via homology modelling and experimental data. PLoS Comput Biol. **5**, e1000475 (2009)
17. Maurer-Stroh, S., et al.: Exploring the sequence determinants of amyloid structure using position-specific scoring matrices. Nat. Methods **7**, 237 (2010)
18. Grantham, R.: Amino acid difference formula to help explain protein evolution. Science **185**, 862–864 (1974)
19. Sharma, T.C., Jain, M.: WEKA approach for comparative study of classification algorithm. Int. J. Adv. Res. Comput. Commun. Eng. **2**(4), 1925–1931 (2013)

Classification of Multi Source Ultrasonogram Image of Steatosis

Nivedita Neogi[1], Arunabha Adhikari[2](✉), and Madhusudan Roy[3]

[1] MeghnadSaha Institute of Technology, Kolkata, West Bengal, India
nivedita.neogi@msit.edu.in
[2] West Bengal State University, Barasat, West Bengal, India
arunabha.adhikari@gmail.com
[3] Saha Institute of Nuclear Physics, Kolkata, West Bengal, India
roy.madhusudan57@gmail.com

Abstract. Aim of supervised classification is to obtain an algorithm that works across the data sets collected from different sources. Ideally a classifier, trained with sufficient variations in the input data, should be able to make reasonably good prediction on test data, obtained from arbitrary independent source. Medical images are sensitive to imaging devices and imaging conditions; therefore, classification of medical images is, typically restricted to a single data set where images are collected under similar conditions. In the present communication human normal and fatty liver ultrasonogram images from two different sources are used. When one source is employed for training and the other for testing, the classification accuracy is low. However, addition of a small fraction of data from the testable source to the training set drastically improves the classification accuracy indicating an algorithm to develop a classifier that works on data from different independent sources.

Keywords: Multi-source · Anisotropy feature · PSO · Ultrasound · Steatosis

1 Introduction

Motivation behind the present work is to work out a suitable algorithm leading to a Computer Aided Diagnosis (CAD) tool to look into liver steatosis from the human liver ultrasonogram images (US). Steatosis of human liver, also known as fatty liver, a common disease which often leads to liver cirrhosis and hepatocellular carcinoma [1], if untreated. Ultrasound technology is readily available and less costly. This kind of image identification modality is also useful to find out liver steatosis [1].

Over the years, in the literature, binary classification detecting steatosis was made using different features like Gray Level Co-Occurrence Matrices (GLCM) features [2–11], wavelet transform [12], radon transform and discrete cosine transform [13], GIST descriptor [14] as well as using different classifiers, Multilayer Perceptron (MLP) [2, 5], Support Vector Machine (SVM) [10, 13, 16, 17], Probabilistic Neural Network PNN [13, 15], Decision Tree (DT) [15], fuzzy sugeno [13].

It is reported in literature that the highest accuracy of classification of ultrasound image of normal and fatty human liver is achieved 100% using fatty liver disease index

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 239–248, 2019.
https://doi.org/10.1007/978-981-15-0108-1_23

(FLDI) and FS with 5 features [13].100% accuracy is achieved by Alivar et al. [18] for classification of diffused livers into normal or fatty using 'serial feature fusing mode'. However, these results are produced with test and train fraction drawn from the same single source.

In our previous work we have proposed a set of anisotropy features [19] and it is establish that these features are very efficient in binary classification of ultrasonogram image (US) into fatty (steatosis) and normal classes [20]. Accordingly, we have compared the efficiencies of five classifiers namely, Bayesian, MLP, PNN, SVM, Learning Vector Quantization (LVQ) with the highest accuracy 99% with PNN, and in a subsequent publication, we have found that ANN classifier with feature selection with Particle Swarm Optimization (PSO) improves the classification to perfection [21]. 100% accuracy is attained with only 6 features and the accuracy is maintained by increasing the number of features.

In all the above communications, images are acquired from one source and always results are reported with larger training set compared to the test set. This is the common practice in pattern recognition relating the medical images. In nearly all published reports, the database is built from a single source. As the images critically depend upon the imaging conditions, a multisource database is rare to find. On the other hand, to find whether these classification algorithms have any usability, we have to work on images obtained from any similar sources. Ideally, a classifier, trained with sufficient variations of input data, should be able to make reasonably good prediction on data obtained from a different and independent source.

In a separate communication [communicated], we find that anisotropy- PSO-ANN method can work on images from two different sources with equal elegance. In both the cases, only 6 selected features lead to perfect classification. When images from both the sources are pooled together, the problem becomes harder and 9 features require achieving 100% classification.

In this communication, we have initially noted as to how the single source accuracy depends on the training fraction. Next we have studied how these accuracies are affected by keeping the training fraction same but supplementing with data from another source. Precisely, our motivation is to minimize the training fraction by using of an alternate dataset. We have also explored the effect in accuracy by replacing the classifier ANN by SVM.

2 Data Acquisition

Data are acquired from two data sources which are independent to each other. These data collection centres are

(1) Chittaranjan National Cancer Hospital (CNCI), Kolkata, West Bengal, India. At CNCI ultrasound images of human livers are grabbed by a scanner (Siemens Sonoline Versa Plus) in which abroad bandwidth phased array convex transducer is used. Ultrasound probe frequency is 3.5 MHz and the image field size varies from 6 to 24 cm.

(2) Institute of Post-Graduate Medical Education and Research and Seth Sukhlal Karnani Memorial (IPGMERSSKM) hospital, Kolkata, West Bengal, India. IPGMERSSKM takes these ultrasound images of human livers by Philips HD7 machine using convex transducers with 2 to 5 MHz and ultrasound probe frequency is 3 to 5 MHz, usually they set it 3 MHz.

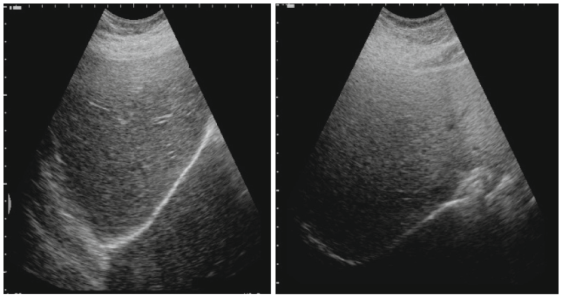
Images from CNCI are categorized as either normal or fatty by the head of the department of radio-diagnosis of CNCI using her visual perception of echo texture and images from IPGMER and SSKM are identified by two radiologists. These images are pathologically correlated and are acquired with proper settings of the echo graphic instrument and stored in the Digital Imaging and Communications in Medicine (DIACOM) format.

2.1 Data Preparation

Square shaped non overlapping sub images are manually cropped from each US image of human liver. This sub images are called Region of Interest (ROI). Different number of ROIs from each image constitutes two datasets which are obtained from two different sources.

(1) Data source-1 (CNCI): 5 ROIs from each US images of 34 patients having normal liver resulting in a total of 170 inputs of normal sample and 10 ROIs from each US of 17 patients are cropped creating total of 170 inputs for fatty liver samples. This is balanced data base where no of data in both classes are same.
(2) Data source-2 (IPGMER and SSKM): 3 ROIs are cropped from 32 normal and 55 fatty US images of IPGMERSSKM hospital yields 96 ROIs of normal human livers and 165 ROIs of fatty human livers.

Figure 1 represents the sample of the normal human liver and fatty human liver of data source-1 and Fig. 2 represents the sample of the normal human liver and fatty human liver of data source-2 respectively.



Normal liver Fatty liver

Fig. 1. Human normal liver and fatty liver from data source-1

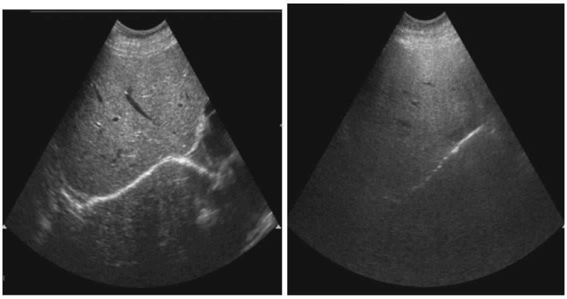

Normal liver
Fatty liver

Fig. 2. Human normal liver and fatty liver from data source-2

3 Feature Extraction

On critical visual observation the texture of a US liver image appears as anisotropic. A set of few novel features based on this anisotropy are introduced by the present authors [19] are also found to be very efficient in classifying the data from data source-1 [20, 21].

These anisotropy features are obtained from different parameters. The parameters are local directionality statistics using Edge Histogram and Line Likeliness [22], Pair Correlation Function (PCF) [23], Grey Level Difference Histogram (GLDH) [24] and randomness of texture using Gray Level Co-occurrence Matrix-chi square (GLCM-χ^2) [25].

From local directionality statistics 4 features are obtained. They are FT [22], Peakiness [19], Skewness [19] and Llike [19].

FT (formulated by Tamura, Mori and Yamawaki [22]):

$$\mathrm{FT} = 1 - \sum\nolimits_{\theta} \left(\theta - \theta_p\right)^2 H(\theta). \tag{1}$$

where $H(\theta)$ is the normalized frequency at the angle θ and θ_p is the angle with maximum frequency. Two other features are proposed for the same purpose by the authors [19].

$$\mathrm{Peakiness} = \max\left(H(\theta) - \overline{H(\theta)}\right). \tag{2}$$

$$\mathrm{Skewness} = \left(\frac{H(\theta) - \overline{H(\theta)}}{\sigma_{H(\theta)}}\right)^3. \tag{3}$$

$$\mathrm{Llike} = \sum\nolimits_{i=1}^{4} \sum\nolimits_{j=1}^{4} P_{ed}(ij) \cos\left|(i-j)\frac{2\pi}{4}\right| / \sum\nolimits_{i}^{4} \sum\nolimits_{j}^{4} P_{ed}(ij). \tag{4}$$

Where an edge code co-occurrence matrix P_{ed} is considered (i, j)$^{\mathrm{th}}$ element of which is the frequency of the pixel pairs with edge code i and j.

20 PCF features are derived according to 20 different neighborhood pixel pair distances (d = 1, 2, 3, 4 and 8) and directions (θ = 0°, 45°, 90°, 135°) [19].

Two features namely, Scale angle and Anisotropy index are calculated from GLDH and GLCM-χ^2.

$$\text{Scale-angle} = f(1,0) - \langle f(d,\theta)\rangle_{d,\theta}. \tag{5}$$

$$\text{Anisotropy index} = \left\langle \langle f(d,0) - f(d,\theta)\rangle_{\theta} \right\rangle_{d}. \tag{6}$$

Where f is parameter of interest f (1, 0) denotes its value at (d, θ) = (1, 0) and the average in the last term is taken over all d and θ's.

Mean and variance of grey level difference vector are determined from GLDH. Therefore, 4 features namely,

Scale-angle index (GLDH mean)
Anisotropy index (GLDH mean)
Scale-angle index (GLDH variance)
Anisotropy indexes (GLDH variance) are predicted.

Similarly Scale angle and Anisotropy index is derived from GLCM-χ^2 and PCF both. Details of these features are given in our previous communication [19, 20]. Total 32 anisotropy features are summarized in Table 1.

Table 1. Name of 32 anisotropy features

Feature number	Name of the anisotropy feature
1.	Peakiness
2.	Skewness
3.	FT
4.	Anisotropy index (GLCM-χ^2)
5.	Scale-angle index (GLCM-χ^2)
6.–25.	PCF (20 different d and θ)
26.	Scale-angle index (PCF)
27.	Anisotropy index (PCF)
28.	Scale-angle index (GLDH mean)
29.	Anisotropy index (GLDH mean)
30.	Scale-angle index (GLDH variance)
31.	Anisotropy index (GLDH variance)
32.	Llike

4 Feature Selection

With the help of the appropriate features set, classifier performs their best. Therefore, feature selection is the important phase of the CAD system. In this study, a meta heuristic optimization method called Particle Swarm Optimization (PSO) [26] is used to select feature. Elaborate description of this algorithm is described at our previous communication [21].

5 Classification Procedure

Two supervised classification algorithm namely Artificial Neural Network of Levenberg-Marquardt back propagation algorithm (ANN) [27] with 10 neurons in one hidden layer and Support Vector Machine (SVM) [28] are used to train the system. For SVM, linear kernel is chosen.

6 Classification Measurement

The performance of the classification is measured by the following parameter

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} * 100 \tag{7}$$

where TP is true positive (true fatty), TN true negative (true normal), FP false positive (false fatty) and FN is false negative (false normal). High accuracy is always desirable for a successful classification system.

7 Result

In this study, the proposed method is developed by MATLAB 2016 environment in a standalone personal computer using Intel i5 8^{th} generation 8250U processor@1.80 GHz with 8 GB RAM and 64 bit Windows 10 operating system. The accuracies reported here are averages accuracies obtained in five independent runs.

Experimental data are collected from two different sources Data set 1: 170 normal and 170 fatty data sample collected from Data Source-1 and Data set 2: 96 normal and 165 fatty data collected from Data Source-2.

In an initial experiment a traditional PSO selected ANN classifier is run on Dataset 1. In our previous communication we found that only 6 features (out of 32 total) selected by PSO can achieve 100% classification accuracy for this data set [21] when 70% of the data are used in training under 5 fold cross validation. Here we changed the training and test fraction of the data and found that accuracy decreases slowly as the training fraction is reduced. Table 2 summarises the results. However, with only 10% data in the training set, 95% accuracy is achievable. This is an important observation since reports of training with such a small fraction of data does not exist so far at least in the field of liver ultrasonogram. This efficiency is a combination of the properly designed feature set, the chosen classifier and the feature selection method.

In the next experiment we ask the question whether the deterioration of the classification accuracy with small training fraction be resisted by supplementing data from another source instead of increasing the training fraction of the dataset-1. Different fraction of dataset-1 is used for training and the other fraction for testing but the entire dataset-2 is used as a supplementary data included in the training set. Table 3 describes the results which are obtained by using different data ratio from dataset-1 with fixed size of dataset-2. Clearly the inclusion of the supplementary data increases the accuracy. Comparing Tables 2 and 3, it is found that for 30% data in the training set the accuracy goes up to 99% from 96% and 50% data in the training set the accuracy goes up to 100% from 99% with the support of the supplementary data. Number of features could also be reduced from 6 to 5. However, if the training set does not contain any data from dataset-1 the accuracy drastically drops. Training with dataset-2 and testing with dataset-1 can

Table 2. Accuracy for only Dataset 1

Train: tests (percentage of data used)	Accuracy
70:30	100%
50:50	99%
30:70	96%
10:90	95%

Table 3. Accuracy with two Datasets number of features = 5

Training set Dataset 1% + Dataset 2%	Test set Dataset 1	Accuracy
70% + 100%	30%	100%
50% + 100%	50%	100%
30% + 100%	70%	99%
0% + 100%	100%	81%

Table 4. Accuracy comparison of two classifiers

Number of features	Accuracy	
	ANN	SVM
5	81%	74%
12	70%	69%
15	68%	69%
20	68%	69%

maximum achieve 81% accuracy in this model. Further improvement needs a modification in this model – either in the features or the classifier or the feature selection method.

We explore the role of classifier in this model and change it to SVM from ANN. The training is with dataset-2 and testing with dataset-1. Results in Table 4 shows that SVM turns the maximum accuracy significantly lower, by increasing the number of features both ANN and SVM perform similar and the accuracy decreases.

8 Conclusion

In this study, a CAD model is explored to identify human normal liver against fatty liver from ultrasonogram images by developing a classifier that uses a novel set of 32 anisotropy features. A good CAD system should work on any ultrasonogram images of human liver independent of its source and settings during imaging. In reality most of the published report in the field of CAD of medical images, the training and testing are restricted within a single source. In the present work we address the question how well a classifier, trained with images from one source works on the images from a different source. Data samples are collected from more than one independent data sources. We use anisotropy features as they are of higher order statistics and they are found to be efficient for classification of data from a single source [21].

In our previous paper [communicated], we show that our proposed CAD model consisting of anisotropy features, PSO dependent feature selection and ANN classifier achieves 100% accuracy on single source data for both dataset-1 and dataset-2 with only 6 anisotropy features. Increasing number of features retains the accuracy level. When dataset-1 and dataset-2 are pooled and fed to this model, 100% accuracy is achieved and this time minimum number of features is 9.

In the present article, we use two sets of data collected from two independent sources and our primary question is as to how the training with the data from one source helps learn the classification characteristic of the other source. We find that if the training is done by entire dataset-2 and dataset-1 is tested for classification the accuracy achieved is 81%. This accuracy level is increased by including a small fraction of the data from the dataset-1 set during the training. Only 30% of total data of dataset-1, if included in the training set the test achieves 99% accuracy. However, if the training is with 30% data of dataset-1 alone then the accuracy level is 96%. This increase in the accuracy level comes from seeing the data of the other set during training. In other words this is a kind of transfer of experience across the data sets.

In another experiment we replace ANN by SVM classifier but the accuracy of the test drastically decreases. This indicates that the necessity of selecting features and classifiers for better performance of this model. Precisely it is our seminal idea behind the work and that is partially able to work across image sources. When a new source of data arrives, inclusion of only a few samples from the new set would suffice a perfect prediction. It is quite possible that, if only a handful of data from few other sources are added to this training set, the model may become truly source independent.

As per our knowledge, multisource medical image data classification, with such simple algorithm, so far is not reported in the earlier literature.

Acknowledgement. We would like to thank Dr. Suparna Majumdar, Department of Radio diagnosis, Chittaranjan National Cancer Institute, Kolkata-700026, India and Dr. Archana Singh, Department of Radio diagnosis, IPGMERSSKM, Kolkata-700020, India for providing the USG of livers for this study. We would like to thank Dr. Tarun Kumar Roy, Senior Consultant Radiologist, Surakha Diagnostic Center, Phoolbagan, Kolkata-700054, India.

References

1. Ascha, M.S., Hanounch, I.A., Lopez, R., Tamini, T.A.R., Feldstein, A.F., Zein, N.N.: The incidence and risk factors of hepatocellular carcinoma in patients with non-alcoholic steatohepatitis. Hepatology **51**, 1972–1978 (2010)
2. Ogawa, K., Fukushima, M., Kubuta, K., Hisa, N.: Computer aided diagnosis system for diffuse liver diseases with ultrasonography by neural networks. IEEE Trans. Nucl. Sci. **45**(6), 3069–3074 (1998)
3. Poonguzhali, S., Ravindran, G.: Automatic classification of focal lesions in ultrasound liver images using combined texture features. Inf. Technol. **7**(1), 205–209 (2008)
4. Li, G., Luo, Y., Deng, W., Xu, X., Liu, A., Song E.: Computer aided diagnosis of fatty liver ultrasonic images based on support vector machine. In: EMBS 2008, 30th Annual International Conference of the IEEE, pp. 4768–4771(2008)
5. Mittal, D., Kumar, V., Saxena, S.C., Khandelwal, N., Kalra, N.: Neural network based focal liver lesion diagnosis using ultrasound images. Comput. Med. imaging Graph. **35**(4), 315–323 (2011)
6. Kumar, S.S., Moni, R.S., Rajeesh, J.: An automatic computer-aided diagnosis system for liver tumours on computed tomography images. Comput. Elect. Eng. **39**, 1516–1526 (2013)
7. Mitrea, D., et al.: Abdominal tumor characterization and recognition using superior order cooccurrence matrices, based on ultrasound images. Computat. Math. Methods Med. **2012**, 1–17 (2012)
8. Horng, M.H., Sun, Y.N., Lin, X.Z.: A diagnostic image system for assessing the severity of chronic liver disease. In: 20th Annual International Conference on Proceedings of Engineering in Medicine and Biology Society. IEEE (1998)
9. Xian, G.: An identification method of malignant and benign liver tumors from ultrasonography based on GLCM texture features and fuzzy SVM. Expert Sys. Appl. **37**, 6737–6741 (2010)
10. Andrade, A., Silva, J.S., Santos, J., Belo-Soares, P.: Classifier approaches for liver steatosis using ultrasound images. Procedia Technol. **5**, 763–770 (2012)
11. Virmani, J., Kumar, V., Niranjan, N.K.: A comparative study of computer-aided classification systems for focal hepatic lesions from B-mode ultrasound. J. Med. Eng. Tech. **37**(4), 292–306 (2013)
12. Lee, W.L., Chen, Y.C., Hsieh, K.S.: Ultrasonic liver tissues classification by fractal feature vector based on M-band wavelet transform. IEEE Trans. Med. Imaging **22**(3), 382–392 (2003)
13. Acharya, U.R., et al.: An integrated index for identification of fatty liver disease using radon transform and discrete cosine transform features in ultrasound images. Inf. Fusion **31**, 43–53 (2016)
14. Acharya, U.R., et al.: Decision support system for fatty liver disease using GIST descriptors extracted from ultrasound images. Inf. Fusion **29**, 32–39 (2016)

15. Acharya, U.R., et al.: Automated characterization of fatty liver disease and cirrhosis using curvelet transform and entropy features extracted from ultrasound images. Comput. Biol. Med. **79**, 250–258 (2016)
16. Yeh, W.C., Huang, S.W., Li, P.C.: Liver fibrosis grade classification with B-mode ultrasound. Ultrasound Med. Biol. **29**, 1229–1235 (2003)
17. Jiang, Z., Yamauchi, K., Yoshioka, K., Aoki, K., Kuroyanagi, S., Iwata, A.: Support vector machine-based feature selection for classification of liver fibrosis grade in chronic hepatitis C. Med. Syst. **30**, 389–394 (2006)
18. Alivar, A., Danyali, H., Helfroush, M.S.: Hierarchical classification of normal fatty and heterogeneous liver diseases from ultrasound images using serial and parallel feature fusion. Biocybern. Biomed. Eng. **36**(4), 697–707 (2016)
19. Neogi, N., Adhikari, A., Roy, M.: Anisotropy of the texture in the ultra-sonogram of human livers. In: 2016 International Conference on Information Science (ICIS), pp. 114–119. IEEE (2016)
20. Neogi, N., Adhikari, A., Roy, M.: Use of a novel set of features based on texture anisotropy for identification of liver steatosis from ultrasound images: a simple method. Multimedia Tools Appl. **67**(3), 1–23 (2018)
21. Neogi, N., Adhikari, A., Roy, M.: Fatty liver identification with novel anisotropy features selected by PSO. J. Imaging Graph. **6**(2), 160–166 (2018). ICMVA (2018)
22. Tamura, H., Mori, S., Yamawaki, T.: Textural features corresponding to visual perception. IEEE Trans. Syst. Man Cybern. **8**(6), 460–473 (1978)
23. Lehoucq, R., et al.: Analysis of image vs. position, scale and direction reveals pattern texture anisotropy. Front. Phys. **2**, 84 (2015). FPHY (2014)
24. Virmani, J., Kumar, V., Niranjan, N.K.: A comparative study of computer-aided classification systems for focal hepatic lesions from B-mode ultrasound. J. Med. Eng. Technol. **37**(4), 292–306 (2013)
25. Zucker, S.W., Terzopoulos, D.: Finding structure in co-occurrence matrices for texture analysis. Comput. Graph. Image Process. **12**(3), 286–308 (1980)
26. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: IEEE Conference on Neural Networks, pp. 1942–1948 (1995)
27. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Cogn. Model. **5**(3), 1 (1988)
28. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)

Role of Perseverance and Persistence for Retaining and Stimulating MOOC Learners

Pankaj Deep Kaur(✉), Jyoteesh Malhotra(✉), and Megha Arora

Department of Computer Science and Engineering,
Guru Nanak Dev University Regional Campus, Jalandhar, India
pankajdeepkaur@gmail.com, jyoteesh.ecejal@gndu.ac.in,
arorامegha587@gmail.com

Abstract. Massive Open Online Courses are a progressing advancement in web based training went for boundless intrigue and open access with the assistance of web. They are a possibly creating innovation, changing how guidance is passed on and financed far and wide. In current and contemporary world a critical increment is seen in the quantity of members learning through MOOCs, supplemental online instruction and e-learning courses. Be that as it may, the effective course fulfillment rate of the students and their execution is far underneath than the conventional instruction framework. Hereafter, this paper endeavors to unfurl the learning examples of MOOC members by distinguishing numerous parameters, for example, Engagement, Persistence, Completion, Attention, Relevance, Confidence and Satisfaction and group them as: Active learners, Passive learners and Bystander learners This paper incorporates an exceptional data file of 30,000 understudies relating to learning courses. The data file has been used to accomplish various objectives: a. to process the connection between the members and courses b. to analyze the learning example of the MOOC with various recognized parameters c. to anticipate the student classification for MOOC courses utilizing Data mining techniques.

Keywords: MOOC · Engagement · Persistence · Learners · Retention

1 Introduction

Massive Open Online Courses (MOOCs) are the rapidly developing strategy of instructive arrangement, holding the possibility to open up access to world class instructing with geological and social limits. These courses are intended for humongous accumulation of students that can be recovered by anybody, anyplace as long as they have a virtual network are available to all and one without section capabilities, and prescribe a full course understanding for nothing. Most importantly benefits are especially high for those understudies for whom the consumption of educational cost to go to eye to eye training at best colleges would be restrictive. The rising notoriety of MOOCs has lead to declare as a trend setting innovation and a genuine danger to organizations of higher instructions. In addition, regardless of the capability of MOOCs, degrees of consistency in general are ordinarily very low (figure of 10% are

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 249–259, 2019.
https://doi.org/10.1007/978-981-15-0108-1_24

seen). This Study hence embarked to investigate the parameter which interlinked with the pupils retention. An examination was led with students who finished MOOCs as full just as the individuals who dropped out, in this way enabling correlation with be made between these subgroups. The goal of this examination is to assess the connection among members and courses. In the light of previously mentioned, this paper tends to the investigation of Active, Passive and Bystander Learners by examining the learning designs with the use of various parameters.

2 Literature Survey

The development of web education and innovative advances connote a significant opportunity to improve education's access. In this regard, the United Nations Educational Consultants, Scientific, and Cultural Organization have built up that free access to educational resources is a technique to upgrade the quality of education, to interchange knowledge and to develop skills. This writing survey elaborates alternate points of view. Advanced education suppliers are adjusting to later and fast worldwide changes, markets and innovations, trying to keep up and improve quality, just as create partnership with their students. MOOCs are resources that could potentially support numerous positive changes. MOOCs might offer motivation for advance education suppliers to revisit both learners and staff engagement, and what quality assurance and enhancement entail in the MOOC context. This mirrors a discussion about whether MOOCs ought to be intended for prominent engagement. From that point forward many number of MOOC courses and MOOC giving stages were developing at a high rate. MOOCs are unique because of the gigantic number of members and open to any client who is engrossed for learning. MOOC courses offer for nothing out of pocket, some MOOCs are issuing certifications or check the genuineness at a significant lower cost.

MOOCs naturally have few regular trademark; short recordings, tests, peer base or/and self-assignments and online discussions yet there are pedagogical differences in courses even in the same platform. Offering or taking part in a MOOC has advantages to each party. However, it is occurring because there are higher dropouts in MOOC, which implies merely 7–13% of pass rate or sometimes less than that finish the courses. MOOC has plenty of success stories since its birth. However, the MOOC style learning has raised many issues along with its rapid development. The Main Problem Present in MOOC is Student retention. MOOC dropout rates are known as 90% or even higher. This research find out the persistence pattern of the student in the courses with using regression analysis [1]. Second Research indicates corresponding variations in the learning outcomes experienced by MOOC participants [2]. They used cluster analysis to determine four prototypical engagement patterns for learners in MOOCs: auditors, samplers, completers and disengages. They analyzed the low dimension of engagement and did not find the user's behavior with data mining techniques. Additionally, dropout in an effort either to anticipate dropout or to depict the sorts of students liable to dropout [4] yet these investigations did not analyze examination of data set. They found that separate motivation is the best factor that enhance and enrich the participation in the MOOC. Most broadly learners motivation can be classified as either intrinsic or extrinsic. [5] To date studies have documented many specific intrinsic and extrinsic motivations among MOOC participants. These include intrinsic motivational factor such as interest in the subject matter connecting with other learners and personal development [6].

This paper focuses on types of learner behavioral factors and how these factors analysis the learner pattern. One Prominent body of such research has centered on differences in MOOC outcomes by Learning Characteristics. In holistic view to literature found Interactivity, colloborativeness, pedagogy and technology has a significant role in making a MOOC effective to a learner. It is stated by in their research that peer interactions and bonds impact on the dropout rates [10]. They discovered that fostering a supportive and positive peer influence will reduce the drop outs. But, this paper provide a dropout pattern of the MOOC and further how to the identify learners behavior so this paper find all the learners characteristics.

3 Supplementary Classification of Learning

3.1 Online Social Learning Platforms

Web based learning is an elective type of open training that gives a chance to students to acquire information and abilities in their very own learning way. Open instruction was first presented around 50 years back, be that as it may, it is currently not the same as the training in the 21st century since we are currently in an advanced age where students can associate with the world by means of innovation and social platform.

3.2 Social Platform

Social stages are considered as a pivotal device for MOOC students to interface and speak with friends and course facilitators explicitly outside MOOC stages.

Facebook and twitter are among the most widely recognized social stages utilized in MOOCs for systems administration and information sharing among students and companions or students and courses facilitators. Social stages additionally help advance online social connections among students, course facilitators, and course assets. There are three types of cooperation of “student communication, student course facilitator collaboration, and student course asset connection”, and are considered as online social association since they happen in online social learning stages.

4 MOOC and Participants Characteristics

This examination further inspected learning results in MOOC by various Types of students with the goal so that retention issue could be resolved. Three kinds of MOOC Learners were named Active Learners Passive Learners and Bystander Learners. Activity Learners who submitted their work on time and as often as possible watched video address recordings demonstrated a high fulfillment rate and a superior evaluation in the courses. Activity students are on the track (“finished undertaking on time”).

MOOC Learners who partook in address recordings however constrained cooperation on courses and less endeavor assignments are Passive Learners. Spectator Learners are who neither watch video nor finishing the assignments. The ward's various leveled or k-implies non progressive bunching strategies were utilized to arrange sorts of learner's behavior while they are occupied with the learning exercise on the MOOC stage.

Crucial attributes of a MOOC are being open, participatory and distributed:

4.1 Open

Investment in a MOOC is free and open to any individual who approaches the web. One may take more than one course and all the substance is available to course takers. At long last there is receptiveness as far as the student's job. The courses are effectively accessible on the web and join into exchange and contribute in the development of learning inside a specific field.

4.2 Participatory

The learning in a MOOC is improved by cooperation both in the creation and sharing of individual commitments and in the associations with the commitment of others however the investment is deliberate.

4.3 Distributed

MOOC depends on the methodology wherein information ought to be disseminated over a system of member's. Most of the courses movement happens in learning conditions, where members interface with the material.

5 Perseverance Analysis in MOOC

This Analysis assesses the commitment and steadiness utilizing courses, addresses and Learners qualities across over one of the informational collection utilized in the MOOC. Diverse examples of understudy conduct in MOOC require distinctive qualities and Framework. An increasingly efficient and expansive scale examination of student engagement persistence in MOOC is imperative to recognize the behavior of students.

There are different terms utilized in the MOOC. Commitment alludes to any example when an understudy connects with the any courses that imply Downloading or observing any video address in the MOOC enlistment. The online recordings, evaluations and online forms assume a critical job in MOOC. The understudies which are enthusiasm for MOOC observe expansive number of address recordings and adhere to the guidelines of MOOC educator. Presently the persistence is the prolonged Engagement. To what extent the students include in the courses and watching a quantities of video over a week. These data discover the characteristic inspiration of students. In the wake of watching video address step by step the Completion is characterized as engagement with the Course until the finish of Course. It implies that watching address video through the most recent week Or Earning a Certificate. The Credit Consideration by state Agencies and placement Opportunities can give to the Active Learners that implies those students very include in the course. Presently the MOOC student selectivity focused on data that was firmly identified with their objectives for courses and overlooked all other data.

Relevance of MOOC was self chosen to a substantial broaden. Students should be all around educated why they need the content and how content identified with student need. Relevance is consider as a Self Awareness and principle factor consider a the enthusiasm for the subject matter, connecting with different students and the interest. In Confidence the MOOC students were certain that they accomplish the objectives. The educator's Sympathy for the students in video were effective in expanding the certainty. So the cooperation between the teacher and student ought to be happen in MOOC or the extraordinary enthusiasm for the course fabricate the self-confidence in the student. After certainty the following stage go under as a fulfillment. In Satisfaction implies that the information picked up from the courses Goal of inspiration more often than not coordinate the investment and premium and the objective. Besides these components supports the conduct of students. Active Learners are more include in this investigation and Bystander students are not intrigued by the Course and this examination evaluates the conduct of various students and furthermore with the utilization of Data Mining techniques that evaluates-patterns.

6 Research Methodology

This examination utilized a blended strategies look into configuration in structure an analytical example in which distinctive information mining procedures are utilized. Colleges can catch and record the manner in which understudies associated with the courses (Fig. 1).

Numerous Student end up falling courses or pulling back and this contextual analysis have combined tables utilizing one of a kind identifiers. In this examination, we connected the informative successive structure, which begins with the gathering and investigation of quantitative information pursued by the accumulation of subjective investigation. The quantitative information was gathered from a vast informational index and in this investigation an id_student distinguish the understudy. Such information can give valuable and noteworthy bits of knowledge into understudies learning conduct which colleges can use to improve understudy execution by furnishing them with extra help at whatever point necessary. The database plot portrays data gathered for understudies in various gatherings as Student Demographic, Student Activities and Module Presentation. Final datasets is consisting by joining the different tables. The student Info table contain demographic details of students, Student VLE and VLE tables contain virtual learning environment information, Student Registration contain information on when the students registered and unregistered for the course student Assessment tables contain information on assessments. Course table is also consisting of code module and code presentations. Information from VLE tables were summarized to get the total sum clicks for various types of activities the student undertake for a course module. Each student undergo several assessment over the duration of course.

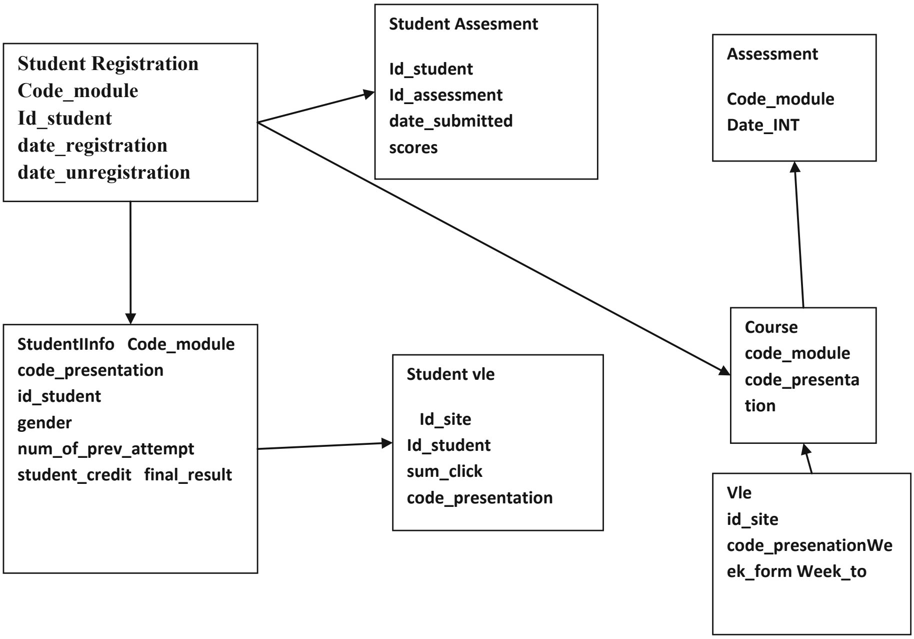


Fig. 1. Database schema

Assessment were weighted and students may opt to drop out of courses by withdrawing from courses at any point the university deems fit. Data is in the form of rows and columns for further justifications (Fig. 2).

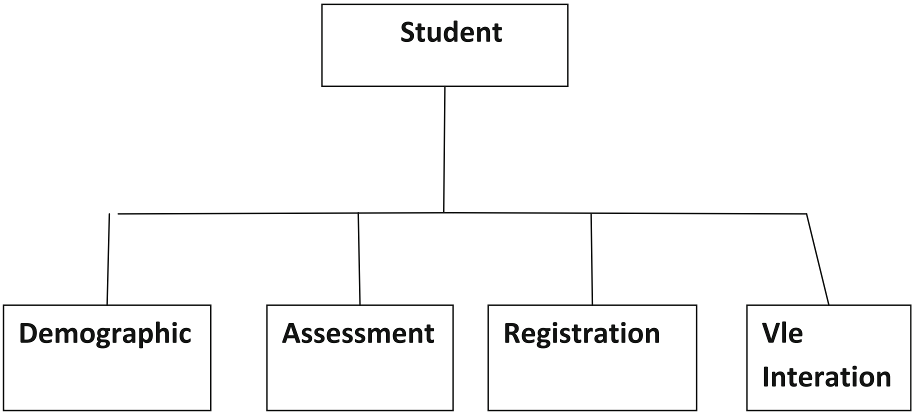
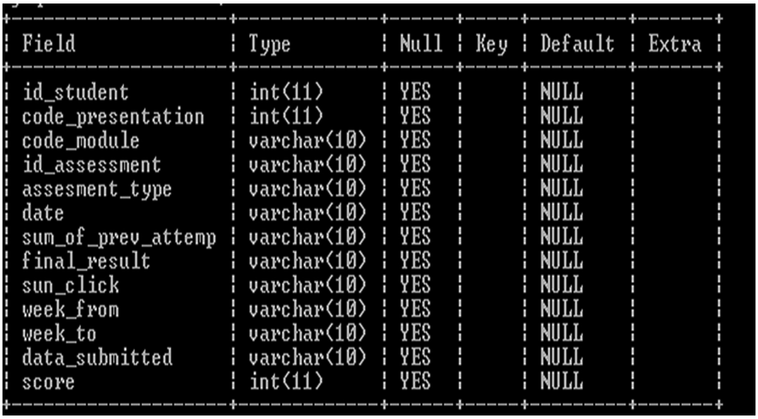


Fig. 2. Data preparation methodology

6.1 Research Settings

The study was conducted in the setting of MOOC, we would be able to explain the important variables of an analysis by observing the top segment of the most decision tree and analyzing the variable important matrix (Fig. 3).



```
+--------------------+-------------+------+-----+---------+-------+
| Field              | Type        | Null | Key | Default | Extra |
+--------------------+-------------+------+-----+---------+-------+
| id_student         | int(11)     | YES  |     | NULL    |       |
| code_presentation  | int(11)     | YES  |     | NULL    |       |
| code_module        | varchar(10) | YES  |     | NULL    |       |
| id_assessment      | varchar(10) | YES  |     | NULL    |       |
| assesment_type     | varchar(10) | YES  |     | NULL    |       |
| date               | varchar(10) | YES  |     | NULL    |       |
| sum_of_prev_attemp | varchar(10) | YES  |     | NULL    |       |
| final_result       | varchar(10) | YES  |     | NULL    |       |
| sun_click          | varchar(10) | YES  |     | NULL    |       |
| week_from          | varchar(10) | YES  |     | NULL    |       |
| week_to            | varchar(10) | YES  |     | NULL    |       |
| data_submitted     | varchar(10) | YES  |     | NULL    |       |
| score              | int(11)     | YES  |     | NULL    |       |
+--------------------+-------------+------+-----+---------+-------+
```

Fig. 3. Different fields in table final

Data_submitted, score, code_modules and sum_clicks are important variables and later we will use these to build the decision tree. The classification and clustering (k-means) method used to identify the behavior of the learner. (Active, Passive and Bystander learners). Here the data can be analyzed by different attributes. We analyzed engagement pattern at the lecture level by usage the methods and find the different type of learners. We group the student assessment_type into one form and find the patterns from the data mining methods.

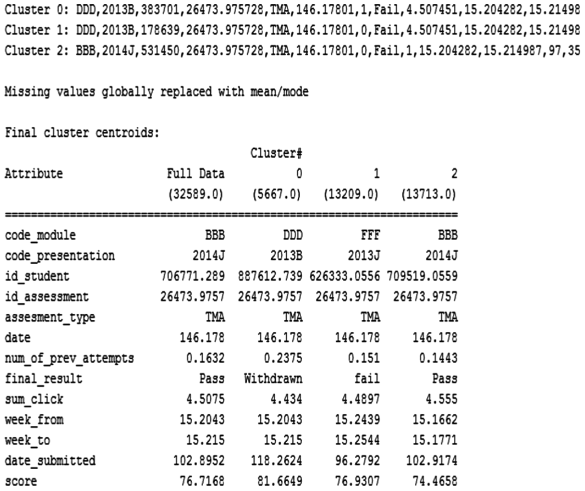
Findings

K-means is the algorithm which is come under the unsupervised learning technique and this technique is used when user have the labeled dataset. The main aim of this method is to find out the groups in the dataset and the number of groups represented by the variable k. The algorithm works as assigning the each data point to one of k groups based on the features that are provided. Data points are clustered based on the feature similarity. The conclusion of the k-means clustering algorithm is as follow: The centroids of k-means, which is used to label new data and add labels on the training dataset. We can also use the objective function for the cluster analysis. This is one of the versatile algorithms for the grouping of the data.

The objective function

$$J = \sum_{j-1}^{k} \sum_{j-1}^{H} \left\| x_j^{(j)} - c_j \right\|^2$$

where $\left\|x_i^{(j)} - c_j\right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j, is an indicator of the distance of the n data points from their respective cluster centres (Fig. 4).



```
Cluster 0: DDD,2013B,383701,26473.975728,TMA,146.17801,1,Fail,4.507451,15.204282,15.21498
Cluster 1: DDD,2013B,178639,26473.975728,TMA,146.17801,0,Fail,4.507451,15.204282,15.21498
Cluster 2: BBB,2014J,531450,26473.975728,TMA,146.17801,0,Fail,1,15.204282,15.214987,97,35

Missing values globally replaced with mean/mode

Final cluster centroids:
                                      Cluster#
Attribute                Full Data           0           1           2
                         (32589.0)    (5667.0)   (13209.0)   (13713.0)
======================================================================
code_module                    BBB         DDD         FFF         BBB
code_presentation            2014J       2013B       2013J       2014J
id_student              706771.289  887612.739 626333.0556 709519.0559
id_assessment           26473.9757  26473.9757  26473.9757  26473.9757
assesment_type                 TMA         TMA         TMA         TMA
date                       146.178     146.178     146.178     146.178
num_of_prev_attempts        0.1632      0.2375       0.151      0.1443
final_result                  Pass   Withdrawn        fail        Pass
sum_click                   4.5075       4.434      4.4897       4.555
week_from                  15.2043     15.2043     15.2439     15.1662
week_to                     15.215      15.215     15.2544     15.1771
date_submitted            102.8952    118.2624     96.2792    102.9174
score                      76.7168     81.6649     76.9307     74.4658
```

Fig. 4. Clusters of active, passive and bystander learners

The four cluster are finding in the clustering analysis. Cluster 0 is for Withdrawn (Passive Learners) Cluster 1 is called as (Passive learners) and cluster 2 is known as (Passive learners). In the cluster analysis the full data conclude that the few participants are passed in the examination and complete the MOOC courses with full active interest and enthusiasm.

These groups pursue the various parameters like dynamic students pursue the high steadiness, culmination, satisfaction and significance however the inverse is on account of uninvolved and onlooker students. Detached students has high rate yet low as contrast with the dynamic students. K-implies bunching is a sort of unsupervised realizing, which is utilized when you have unlabeled information. The objective of this calculation is to discover bunches in the information, with the quantity of gatherings spoken to by the variable K.

The calculation works iteratively to allocate every datum point to one of K bunches dependent on the highlights that are given. Information focuses are grouped dependent on highlight comparability. The aftereffects of the K-implies bunching calculation are as the centroids of the K groups, which can be utilized to mark new information and names for the preparation data (each information point is allocated to single group.

7 Conclusion

Massive open online course (MOOC) is a free Web-based separation learning program that is intended for the cooperation of expansive quantities of topographically scattered understudies. MOOC enrollment is huge, free and not confined to understudies by age or geographic area. They need to pursue the arrangement of a course i.e., incorporate a prospectus and calendar and offer the direction of one or a few educators. Our Study Represent a vital commitment to the examination on MOOC exceptionally identified with the maintenance and investigation conduct of the Learners. Our examination arrange the diverse students as Active Learners, Passive Learners and Bystander students. This paper incorporates extraordinary data values of 30,000 understudies relating to learning courses. The data values has been used to accomplish various objectives: a. to register the connection between the members and courses b. to inspect the learning example of the MOOC with various recognized parameters c. to anticipate the student classification for MOOC courses utilizing information mining systems d. to recommend components for holding and invigorating MOOC students according to as distinguished class. The commitment, Persistence and consummation are the most grounded indicator to disclose the MOOC aim to utilize. This paper will discover the technique for the Retention increases that is the MOOC has increasingly number of Active Users as contrast with Bystander.

References

1. Evans, B.J., Baker, R.B., Dee, T.S.: Persistence patterns in massive open online courses (2016)
2. Reeves, T.D., Tawfik, A.A., Msliu, F., Simsek, I.: What's in it for me? Incentives, learning and completion in massive open online courses (2007)
3. Alarimi, K.M., Zo, H., Cignek, A.P.: Understanding the MOOC continuance: the role of openness and reputation (2015)
4. Doijode, V., Singh, N.: Predicting student Success based on interaction with virtual learning environment
5. Tseg, S.-F., Tsao, Y.-W., Lai, K.R.: Who will pass? Analyzing learners behavior in MOOC (2016)
6. Suguna, K., Nandhini, K.: Literature review on data mining techniques (2015)
7. Barak, M., Watted, A., Haick, H.: Motivation to learn in massive open online courses: examining aspect of language and social engagement (2015)
8. Song, J., Zhang, Y., Duan, K., Hossain, M.S.: TOLA: topic-oriented learning assistance based on cyber-physical system and big data (2016)
9. Gere, M., Goel, S.: Data mining techniques, methods and algorithms: a review on tools and their validity (2015)
10. Dutt, A., Aghabozrgi, S., Ismail, M.A.B., Mahroeian, H.: Clustering algorithms applied in educational data mining. Int. J. Inf. Electron. Eng. **5**(2), 112 (2015)
11. Anderson, A., Huttenlocher, D., Kleinberg, J., Leskovec, J.: Engaging with massive online courses. In: Proceeding of 23rd International Conference (2014)
12. Merceron, A., Yacef, K.: Educational data mining: a case study. In: Proceedings of the 12th International Conference on Artificial Intelligence in Education, AIED 2005 (2005)

13. Kumar, J.: A comprehensive study of educational data mining. Int. J. Electr. Electron. Comput. Sci. Eng. Spec. Issue - TeLMISR (2015). ISSN 2348–2273
14. Jindal, R., Borah, M.D.: A survey on educational data mining and research trends. Int. J. Database Manag. Syst. (IJDMS) **5**(3), 53 (2013)
15. Oskouei, R.J., Askari, M.: Predicting academic performance with applying data mining techniques **3**, 79–88 (2014)
16. Osmanbegović, E., Suljić, M.: Data mining approach for predicting student performance. Econ. Rev. J. Econ. Bus. **X**(1), 3–12 (2012)
17. Ng, K.K., Chen, Z.: Retention and intention in massive open online courses: in depth www.educause.edu/ero/article/retention-and-intention-massive-open-online-courses-depth-0
18. Blake, D.: What type of MOOC student are you? (2014). http://moocs.com/index.php/what-type-of-mooc-student-are-you. Accessed 10 Aug 2015
19. Sharif, A., Magrill, B.: Discussion forums in MOOCs. Int. J. Learn. Teach. Educ. Res. (2015)
20. Margaryan, A., John, A.: Instructional quality of massive open online courses (MOOCs). Comput. Educ. **80**, 77–83 (2015)
21. Thai-Nghe, N., Drumond, L., Krohn-Grimberghe, A., Schmidt-Thieme, L.: Recommender system for predicting student performance. Procedia Comput. Sci. **1**(2), 2811–2819 (2010)
22. Gardner, J., Brooks, C.: Dropout model evaluation in MOOCs. IEEE Access **2018**, 7906–7912 (2018)
23. Umer, R., Susnjak, T., Mathrani, A., Suriadi, S.: Prediction of students' dropout in MOOC environment. IEEE Access **3**(2), 43–47 (2017)
24. Wen, M., Yang, D., Rosé, C.P.: Sentiment analysis in MOOC discussion forums: what does it tell us? IEEE (2011)
25. Jothi, N., Rashid, N.A., Husain, W.: Data mining in healthcare - a review. Procedia Comput. Sci. **72**, 306–313 (2015)
26. Ţăranu, I.: Data mining in healthcare: decision making and precision. Database Syst. J. **5**(4), 33–40 (2015)
27. Naraei, P., Street, V., Street, V., Street, V.: Application of multilayer perceptron neural networks and support vector machines in classification of healthcare data. IEEE Access 848–852 (2016)
28. Huang, F., Wang, S., Chan, C.: Predicting patterns by using data mining based on healthcare information system. In: 2012 IEEE International Conference on Granular Computing are Predicted, pp. 12–15 (2012)
29. Yan, K., You, X., Ji, X., Yin, G., Yang, F.: A hybrid outlier detection method for health care big data. In: 2016 IEEE International Conferences on Big Data Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications, pp. 157–162 (2016)
30. Farooq, M.A., Azhar, M.A.M., Raza, R.H.: Automatic lesion detection system (ALDS) for skin cancer classification using SVM and neural classifiers. In: 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering, pp. 301–308 (2016)
31. Sivagowry, S., Durairaj, M., Persia, A.: An empirical study on applying data mining techniques for the analysis and prediction of heart patterns. In: 2013 International Conference on Information Communication and Embedded Systems, pp. 265–270 (2013)
32. Anusha, C., Vinay, S.K., Raj, H.J.P., Ranganatha, S.: Medical data mining and analysis for heart patterns dataset using classification techniques. In: National Conference on Challenges in Research & Technology in the Coming Decades (CRT 2013), p. 1.09 (2013)
33. Palaniappan, S., Awang, R.: Intelligent heart patterns prediction system using data mining techniques. In: 2008 IEEE/ACS International Conference on Computer Systems and Applications, pp. 108–115 (2008)

34. Armstrong, L.J.: Rice crop yield forecasting of tropical wet and dry climatic zone of india using data mining techniques, pp. 357–363 (2016)
35. Delibašić, B., Marković, P., Delias, P., Obradović, Z.: Mining skier transportation patterns from ski resort lift usage data. IEEE Access 1–6 (2016)
36. Krupa, P., Myers, J.: Data mining root cause analysis findings, pp. 1–4 (2017)
37. Chuan, W.: Data mining and data mining technology under, pp. 2–6
38. Mukhlas, A., Ahmad, A.: Data mining technique: towards supporting local co-operative society in customer profiling, market analysis and prototype construction, pp. 109–114, May 2016
39. Liu, F., Wang, L., Qian, Y., Wu, Y.: Analysis of MOOCs courses dropout rate based on students' studying behaviors. IEEE Access **83**(Hss), 139–144 (2017)
40. De La Garza, A.L., Gomez-Zermeno, M.G.: Research analysis on MOOC course dropout and retention rates. IEEE Access 3–14 (2016)
41. Wang, W.: Deep model for dropout prediction in MOOCs. IEEE Access (2015)
42. Chen, J., Feng, J., Sun, X., Wu, N., Yang, Z., Chen, S.: MOOC dropout prediction using a hybrid algorithm based on decision tree and extreme learning machine. IEEE Access **2019** (2019)
43. Onah, D., Sinclair, J.: Dropout rates of massive open online courses: behavioural patterns dropout rates of massive open online courses: behavioural patterns. IEEE Access (2014)

MRI Image Compression Using Asymmetric Wavelet Analysis

Ekta Soni[1] and Rashima Mahajan[2](✉)

[1] GD Goenka University, Gurgaon, India
ekta.soni@gdgoenka.ac.in
[2] Manav Rachna International Institute of Research and Studies,
Faridabad, India
rashima.fet@mriu.edu.in

Abstract. An exponential rise in amount of medical image data generation has been recorded since two decades. In telemedicine based applications, this enormous amount of data is initially stored and later, transmitted to medical experts for required diagnosis. This whole storage and transmission process may require extremely high memory space and large transmission bandwidth. An attempt has been made to compress MRI (magnetic resonance imaging) data by exploiting neighbouring pixel redundancy found in images along with the relevant diagnostic information. A detailed analysis of MRI image compression using asymmetric wavelet based frequency transformation techniques has been carried out in MATLAB. It implements discrete wavelet transform using asymmetric mother wavelet daubechies (db) in conjunction with a hierarchal embedded bit coding scheme, set partitioning in hierarchal trees. Distinct daubechies (db6, db12, db16 and db20) mother wavelets have been applied to a set of ten MRI image (512 × 512) dataset. The performance of applied compression scheme has been evaluated by analyzing Compression-Ratio achieved, Peak-Signal-to-Noise-Ratio and Bits-Per-Pixel. A high compression ratio (CR - 3.84) with considerably high peak signal to noise ratio (PSNR - 43.35) is achieved with db20 wavelet. This high value of PSNR may help in preserving the diagnostic details in input MRI image data to a great extent and high CR would lead to less storage space followed by small transmission bandwidth requirement as compared to earlier compression techniques. It could definitely lead to development of more efficient remote healthcare systems.

Keywords: Asymmetric wavelets · Compression · Compression-Ratio · Image · MRI · Peak-Signal-to-Noise-Ratio · Wavelets

1 Introduction

Medical imaging is a technique of acquiring and analyzing diagnostic information through the visual representation. Magnetic resonance imaging (MRI) and Computed Tomography (CT) provides high spatial resolution of a brain among distinct medical imaging techniques including X-Ray, Positron Emission Tomography (PET) and Single-Photon Emission Computed Tomography (SPECT) [1]. Magnetic resonance imaging is a non-invasive technique and is dominantly adopted for brain tissue classification. These

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 260–270, 2019.
https://doi.org/10.1007/978-981-15-0108-1_25

techniques have become important in certain body disorder diagnostics; however, massive data generated may result in inefficient storage space utilization [2]. This huge amount of data is required to be compressed to optimally utilize the available storage space and transmission bandwidth. The high degree of correlation between neighboring pixels is exploited to obtain higher compression values [2, 3].

The compression techniques are mainly categorized as irreversible (lossy) and reversible (lossless). The irreversible methods save storage space, require low bandwidth and provide high compression ratio (CR). However, reversible (lossless) methods are just the reverse of above and generally used for compression of medical images [2, 4, 5]. The maximum CR limit in lossless compression is around 4:1 while in lossy methods this limit can be as high as 20:1 [2]. The compression can either be done in the spatial (time domain) or transform domain (frequency transform domain) of the image. The Fourier transform is the basic frequency transformation technique. It transforms the input MRI data but with loss in phase related details of an input image data and only provides information about frequency components. It can be replaced by using short time Fourier transform (STFT) which is a windowing technique. It uses a certain size of the window to get details in both domains. However due to fixed size window, it cannot provide precise information every time [6]. In addition to these drawbacks the frequency transformation techniques may also have blocking effects. Discrete cosine transform (DCT) is a technique which reduces the blocking effects because it possesses the property of even symmetric extension [7] and gives only real terms after transformation. However, it is a lossy compression technique and generally band limited. To completely eliminate blocking effects wavelet transform (WT) can be utilized. It can transform and compress the whole image as a singular data object [8]. Unlike Fourier transform, it preserves information both in time and frequency domain. It also eliminates issues of constant windowing technique which was the case of STFT, as it utilizes a Variable-Sized windowing method [6]. It works on the principle of recording the neighboring signals at different scales and then calculates the difference between them. To implement WT by only using the dyadic scales and positions, discrete wavelet transform (DWT) can be used [9].

DWT has a multi-resolution feature close to characteristics of the human visual system [3]. At the highest level of compressions, it is capable of reducing more artifacts than DCT. It may reduce memory requirements since fixed point arithmetic is involved instead of floating point and thereby reducing the number of arithmetic computations [10]. In DWT the image is divided into four sub bands *viz.* LL, LH, HH and HL, by using horizontal and vertical filters [11]. These sub-bands are of different spatial domain and of independent frequency. The LL shows the low frequency coefficient containing approximation information while the remaining three are high frequency components containing detailed information of the image [8]. The LL coordinate can be divided again and again into LL, LH, HH and HL as per the required depth of the wavelet transform. Once a desired level of depth is achieved, the further decomposition of the last LL approximation coordinate will be terminated [6]. This paper includes the frequency transformation of input MRI images using asymmetric wavelets db6, db12 and db20. These are also called mother wavelet or basis functions of daubechies wavelet.

The compressed image obtained through frequency transformation technique should go through the encoding process for further compression. The encoding of an image can be done by considering storage space and neighboring pixel redundancy [12]. Encoding methods related to DWT are classified as Vector & Scalar methods of encoding. Embedded zero tree wavelet (EZW), and Set Partitioning In Hierarchal Trees (SPIHT) [13, 14] are examples of vector encoding schemes while Joint-Photographic-Experts-Group-2000 (JPEG 2K) and Embedded-Block-Code-for-Optimized-Truncation (EBCOT) are scalar methods of encoding [3]. Encoding can be done with Embedded-Zero-Tree structures obtained by analyzing similarity in distinct sub bands [10, 15].

SPIHT algorithm is also a kind of embedded coding and has a zero tree structure [16]. The zero-trees are spatial orientation tree pattern consisting of coefficients of different sub band level. It is also a progressive transmission method like EZW which can be stopped at any point. However, it provides higher compression ratio and better performance. The partitioning of quad trees in SPIHT can be done on threshold values by processing the image towards lowering threshold [13, 14, 16, 17]. Initially a threshold value is defined and then the transformed coefficient tree is created by comparing its value with the threshold value. The threshold value can be given as [14]

$$T_n = (T_{n-1}/2) \tag{1}$$

It decreases after each of the iterations.

If the coefficient is at the highest level of the tree and is insignificant in comparison to choose a threshold value it would be ignored. Its descendants are also proved to be insignificant which are at lower levels hence ignored. In this way by only checking one coefficient value many coefficients can be eliminated automatically. This will save the time of encoding and may contribute towards achieving higher CR [18]. The pictorial representation of DWT segmented image and its SPIHT encoded form is shown in the Fig. 1(a and b).

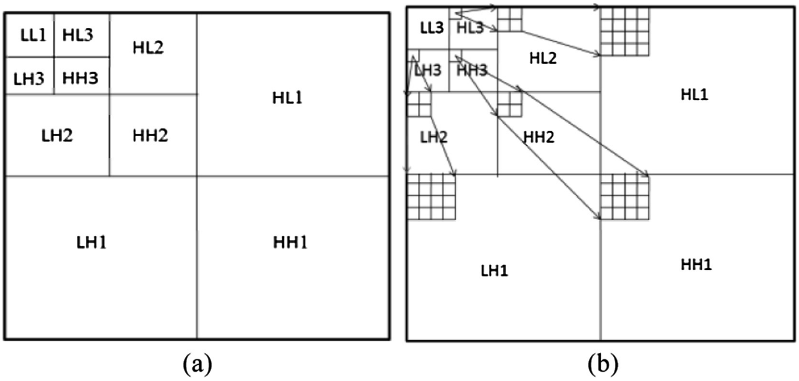


Fig. 1. (a) DWT segmented image and (b) hierarchal representation by SPIHT [16].

In SPIHT, large subsets are used to keep the insignificant information coefficients together; hence partition of the coefficients is being done [19]. A vector quantization based DVQ-DCCR (Dynamic vector quantization with distortion-constrained codebook replenishment) method for compression has also been suggested in literature [12]. However, the compression ration achieved was very low. Table 1 presents the summary of distinct medical image compression techniques.

Table 1. Comparison of different medical image compression methods

<table>
<tr><th>S. No.</th><th>MRI image compression technique</th><th colspan="12">Parameters</th><th>Image Specification</th><th>Remarks</th></tr>
<tr><td rowspan="3">1.</td><td rowspan="3">DCT + vector zig-zag order + DPCM + MSVQ + entropy encoding [7]</td><td colspan="6"><b>PSNR</b></td><td colspan="6"><b>BPP</b></td><td rowspan="3">512x512</td><td rowspan="3">Better scheme than JPEG and vector quantization &hybrid DCT-vector quantization</td></tr>
<tr><td colspan="6">34.26</td><td colspan="6">0.49</td></tr>
<tr><td colspan="6">33.58</td><td colspan="6">0.40</td></tr>
<tr><td rowspan="4">2.</td><td rowspan="4">Set Partitioning in Hierarchal Trees [14]</td><td colspan="6"><b>PSNR</b></td><td colspan="6"><b>BPP</b></td><td rowspan="4">512x512</td><td rowspan="4">SPIHT is better than EZW in terms of PSNR</td></tr>
<tr><td colspan="6">37.2</td><td colspan="6">0.5</td></tr>
<tr><td colspan="6">34.1</td><td colspan="6">0.25</td></tr>
<tr><td colspan="6">31.9</td><td colspan="6">0.15</td></tr>
<tr><td rowspan="3">3.</td><td rowspan="2">a. JPEG [17]</td><td colspan="4"><b>CR</b></td><td colspan="4"><b>MSE</b></td><td colspan="4"><b>PSNR</b></td><td rowspan="3">512x512</td><td rowspan="3">JPEG2K provides better reconstruction of compressed images</td></tr>
<tr><td colspan="4">8.14</td><td colspan="4">16.11</td><td colspan="4">35.85</td></tr>
<tr><td>b. JPEG2K [17]</td><td colspan="4">9.12</td><td colspan="4">11.27</td><td colspan="4">46.60</td></tr>
<tr><td rowspan="3">4.</td><td rowspan="3">DCT+CSPIHT [11]</td><td colspan="3"><b>BPP</b></td><td colspan="3"><b>SPIHT (PSNR)</b></td><td colspan="3"><b>JPEG2K (PSNR)</b></td><td colspan="3"><b>DCT+ CSPIHT (PSNR)</b></td><td rowspan="3">512x512</td><td rowspan="3">Computational complexity reduced but with increase in transmission bandwidth</td></tr>
<tr><td colspan="3">0.025</td><td colspan="3">33.4</td><td colspan="3">36.1</td><td colspan="3">39.5</td></tr>
<tr><td colspan="3">0.50</td><td colspan="3">49.8</td><td colspan="3">50.5</td><td colspan="3">52.0</td></tr>
<tr><td rowspan="5">5.</td><td>For bit rate 3 [4]</td><td colspan="4"><b>3D-SPIHT</b></td><td colspan="4"><b>BISK</b></td><td colspan="4"><b>SPECK</b></td><td rowspan="5">3D medical images</td><td rowspan="5">Cdf9/7 with 3D-SPIHT outperforms db6, db12 and db20 in terms of PSNR and also crosses the limit of PSNR which is generally set at 50.</td></tr>
<tr><td>Daubechies 4,</td><td colspan="4">56.89</td><td colspan="4">53.9</td><td colspan="4">53.98</td></tr>
<tr><td>Daubechies 6,</td><td colspan="4">56.26</td><td colspan="4">54.36</td><td colspan="4">53.67</td></tr>
<tr><td>cdf9/7</td><td colspan="4">56.23</td><td colspan="4">55.23</td><td colspan="4">55.3</td></tr>
<tr><td>cdf5/3</td><td colspan="4">56.18</td><td colspan="4">55.18</td><td colspan="4">54.78</td></tr>
<tr><td rowspan="6">6.</td><td>Chen Y. T. et al [5]</td><td colspan="6"><b>MRI encoding time (s)</b></td><td colspan="6"><b>MRI decoding time (s)</b></td><td rowspan="6">512x512</td><td rowspan="6">A method of adaptive prediction implemented for lossless compression.</td></tr>
<tr><td>SPIHT</td><td colspan="6">1.8</td><td colspan="6">1.8</td></tr>
<tr><td>JPEG2000</td><td colspan="6">0.8</td><td colspan="6">0.8</td></tr>
<tr><td>CALIC</td><td colspan="6">3.2</td><td colspan="6">2.4</td></tr>
<tr><td>SSM</td><td colspan="6">4.8</td><td colspan="6">3.3</td></tr>
<tr><td>WCAP</td><td colspan="6">3.4</td><td colspan="6">2.1</td></tr>
<tr><td rowspan="2">7.</td><td rowspan="2">SPIHT encoded image transmission [18]</td><td colspan="4"><b>BER, BPP</b></td><td colspan="4"><b>PSNR</b></td><td colspan="4"><b>Rate</b></td><td rowspan="2">512×512</td><td rowspan="2">Image transmission scheme for compressed SPIHT over BSC channel</td></tr>
<tr><td colspan="4">0.01, 0.252</td><td colspan="4">32.41</td><td colspan="4">0.72</td></tr>
<tr><td>8.</td><td>DWT(hardware implementation) [19]</td><td colspan="12">---</td><td>512×512</td><td>Resulted in small memory requirements and fixed point arithmetic implementation</td></tr>
</table>

Different MRI compression techniques have been explored in detail and it has been observed that compression using daubechies wavelet and its basis functions db6, db12, db16 and db20 have not been explored extensively. An attempt has been made in this research to present a detailed analysis of MRI image compression using asymmetric wavelet based frequency transformation technique. Frequency transform technique is implemented to reduce the number of coefficients representing an image, hence to reduce the size of an input image. It uses discrete wavelet transform using asymmetric mother wavelet *i.e.,* daubechies (db) in conjunction with a hierarchal embedded bit coding scheme *i.e.,* set partitioning in hierarchal trees. Real time brain images (512 × 512) of human subjects were collected from a neural department of a hospital. It compares results obtained by applying daubechies db6, db12, db16 and db20 mother wavelets to compress the collected MRI image data. The performance of applied compression scheme is evaluated by analyzing Compression-Ratio achieved, Peak-Signal-to-Noise-Ratio and Bits-Per-Pixel.

2 Materials and Methods

The functional block diagram of MRI image acquisition and its compression is shown in the Fig. 2. This block diagram describes the proposed compression method by fusion of DWT transform, SPIHT encoding, inverse transform and finally reconstruction of the image.

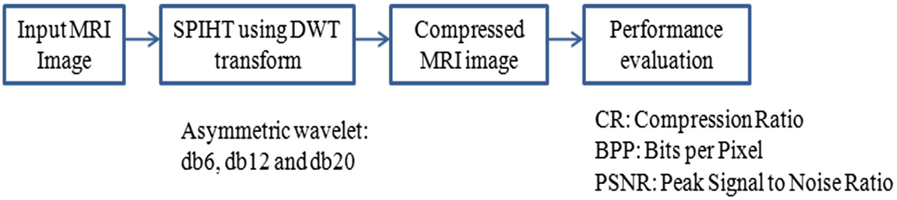


Fig. 2. Block diagram representation of medical image compression using asymmetric wavelets

2.1 Input MRI Image

The real time MRI images were acquired from the neural department of a local hospital. The acquired brain images (MRI) were in DICOM (Digital-Imaging-and-Communications-in-Medicine) format and are not a readable format for MATLAB. Input images were first converted into .jpg (Joint Photographic Experts Group) format using DICOM converter software and then undergone through a compression process which is defined in the next section. Input images in .jpg format have been exported to MATLAB workspace for subsequent analysis.

2.2 MRI Image Transformation

In the MATLAB environment the images were first of all frequency transformed using DWT's basis functions db6, db12, db16 and db20. These are asymmetric wavelet functions. Individual input images were first of all transformed through each of the above mentioned wavelet and after getting transformed the images were individually encoded through a common encoding method i.e. SPIHT. The given process resulted in the compressed images which could either be stored or transmitted for further use. At the receiver end decompression of the transmitted image is done. Inverse of DWT and inverse of SPIHT has been used to convert the image into original form. After getting the original image, the performance of the applied compression scheme has been evaluated.

2.3 Performance Metrics

Once the decompressed images are recovered, the performance of applied compression methods has been checked using different performance evaluation metrics like bits per pixel, compression ratio, mean squared error and peak signal to noise ratio.

BPP (Bits-Per-Pixel): It is used to detail the number of colors in an image. It is calculated using the relation 2^{bpp}.

CR (Compression-Ratio): The CR of an image can be given in terms of the ratio number of bits in uncompressed image to the compressed image [20]. It represents the amount of compression of an image and is calculated as per Eq. (2):

$$\mathrm{CR} = \frac{\text{Number of bits in uncompressed image}}{\text{Number of bits in compressed image}} \tag{2}$$

With the increase in compression ratio, image resolution would be decreased. Hence the compression ratio should be chosen carefully in order to protect the diagnostic information contents to be lost [21]. It should be at a moderate level.

MSE (Mean Squared Error): It is the measure of distortion in decompressed image and defined in the terms of 'average of the square error' [21]. Its value should be low in order to decrease error or noise of the signal [14]. The MSE is computed using following Eq. (3) as:

$$\mathrm{MSE} = \frac{\sum_{\mathrm{M,N}}[\mathrm{I}_1(\mathrm{m,n}) - \mathrm{I}_2(\mathrm{m,n})]^2}{\mathrm{M} \times \mathrm{N}} \tag{3}$$

It represents two m × n size grayscale images I_1 and I_2, and one image is the compressed version of the other. Where 'M' represents total rows and 'N' represents total columns of an input medical image. $\mathrm{I}_1(\mathrm{m,n})$ is the input image and $\mathrm{I}_2(\mathrm{m,n})$ is the decompressed output image.

PSNR (Peak Signal to noise ratio): This is also a quality measure like CR and used to measure the resolution of reconstructed or decompressed image. It is defined as the ratio between the maximum possible power of signal and the power of corrupting noise and expressed in terms of the logarithmic decibel scale [14, 21]. It is computed as given by following Eq. (4) [20]:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{(\text{Dynamics of image})^2}{\mathrm{MSE}}\right)\ (\mathrm{dB}) \tag{4}$$

where 'Dynamics of image' is defined as the maximum fluctuation in the input image data type. For 8-bit unsigned integer data type images, its value is predefined and set to 255. PSNR varies in inverse proportion to CR and MSE and directly proportional to the image quality, as if the value of the PSNR increases so is the resolution. Its range varies in between 30 and 50 dB values [21].

3 Results and Discussions

This section presents and discusses the results obtained by applying four DWT asymmetric basis wavelets: db6, db12 and db20 in fusion with SPIHT encoding on 512 × 512 MRI image data set of ten images. MATLAB has been chosen to develop an algorithm for compression and its implementation. The acquired images were in 'DICOM' format which is not acceptable for image processing in MATLAB. For this reason DICOM converter software has been used before loading the images into MATLAB workspace. The ten input images are imported into MATLAB and are shown in the Fig. 3. An algorithm has been presented to compress input MRI images using four wavelets db6 db12, db16 and db20. MATLAB was used for compression and decompression by using function 'wcompress' and by inserting parameters as 'waveletname' and 'maxloop'. The 'maxloop' is a function which defines the number of loops applied for transformation and for the proposed work set to 12. It should not be below than '3' [14]. The resultant images have been encoded using SPIHT technique. Three compression parameters CR, PSNR and BPP are obtained for all the ten acquired MRI images and are tabulated in Table 2 with respect to each of the wavelet used. Figure 4a–e represent compression of Fig. 5 with the above mentioned methodology using db6, db12, db16 and db20 mother wavelets.

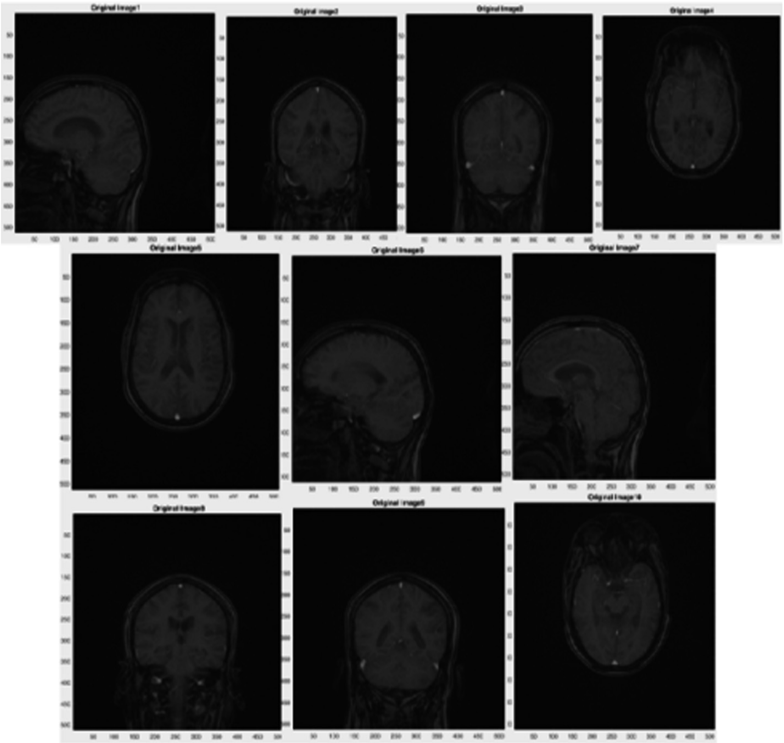


Fig. 3. Original real time MRI image dataset (1 to 10)

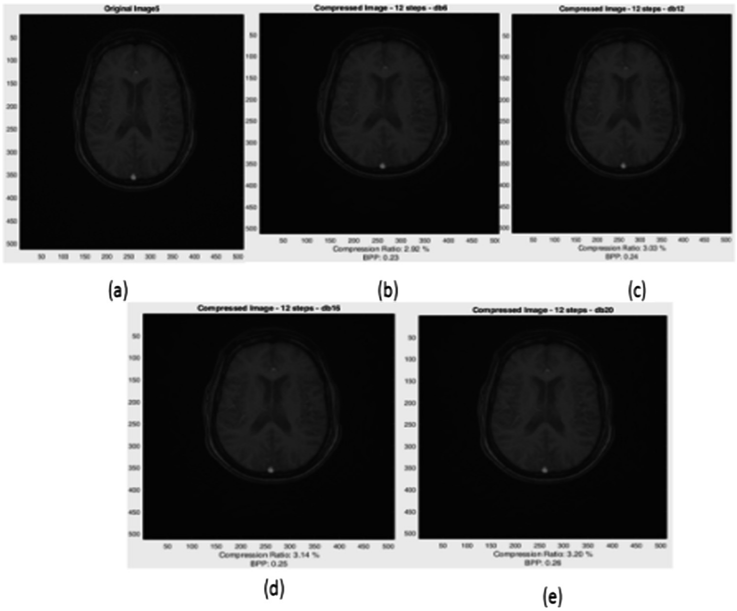

Original Image5
Compressed Image - 12 steps - db6
Compression Ratio: 2.92 %
BPP: 0.23
Compressed Image - 12 steps - db12
Compression Ratio: 3.03 %
BPP: 0.24
(a)
(b)
(c)
Compressed Image - 12 steps - db16
Compression Ratio: 3.14 %
BPP: 0.25
Compressed Image - 12 steps - db20
Compression Ratio: 3.20 %
BPP: 0.26
(d)
(e)

Fig. 4. (a) Original image 5, (b) image compressed through db6, (c) db12, (d) db16 and (e) db20.

During MRI image compression followed by decompression using wavelets db6, db12, db16 and db20, distinct performance parameters have been calculated and tabulated in Table 2 for subsequent evaluations.

Table 2. MRI image compression parameters using DWT db6, db12, db16 and db20 mother wavelet in fusion with SPIHT

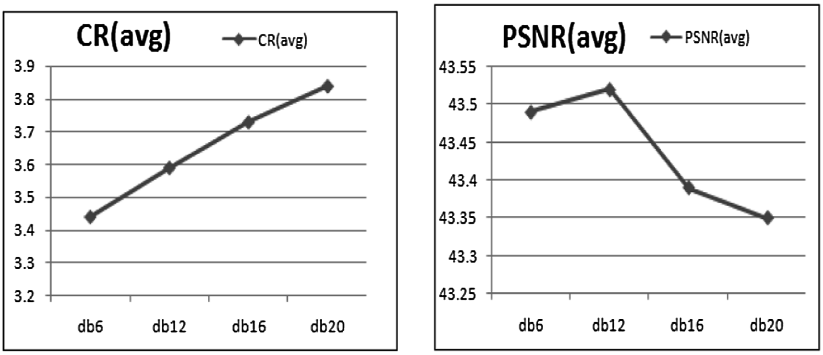
Wavelet used	Parameters	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6	Image 7	Image 8	Image 9	Image 10	Average
db6	CR	3.65	3.77	2.88	3.16	2.42	3.55	4.36	3.86	3.46	3.30	3.441
	MSE	3.00	3.06	2.48	3.02	2.55	2.87	3.52	3.00	2.84	2.92	2.926
	PSNR	43.36	43.27	44.18	43.33	44.06	43.56	42.66	43.35	43.60	43.48	43.485
	BPP	0.29	0.30	0.23	0.25	0.23	0.28	0.35	0.31	0.28	0.26	0.278
db12	CR	3.73	3.87	3.06	3.24	3.03	3.59	4.41	3.97	3.57	3.44	3.59
	MSE	3.02	2.95	2.61	2.69	2.54	2.93	3.53	3.20	2.85	2.89	2.92
	PSNR	43.33	43.44	43.96	43.83	44.09	43.47	42.65	43.22	43.58	43.60	43.51
	BPP	0.3	0.31	0.24	0.26	0.24	0.29	0.35	0.32	0.29	0.27	0.29
db16	CR	3.89	3.98	3.16	3.40	3.14	3.71	4.63	4.05	3.74	3.55	3.73
	MSE	3.08	3.19	2.59	2.85	2.57	2.99	3.44	3.16	3.11	2.94	2.99
	PSNR	43.25	43.09	44.00	43.57	44.04	43.38	42.76	43.13	43.21	43.45	43.39
	BPP	0.31	0.32	0.25	0.27	0.25	0.30	0.37	0.32	0.30	0.28	0.30
db20	CR	4.00	4.13	3.28	3.47	3.20	3.84	4.74	4.23	3.88	3.63	3.84
	MSE	3.28	3.20	2.68	2.79	2.66	3.08	3.59	2.85	2.92	2.85	2.99
	PSNR	42.97	43.08	43.84	43.68	43.88	43.24	42.58	43.17	43.47	43.58	43.35
	BPP	0.32	0.33	0.26	0.28	0.26	0.31	0.38	0.34	0.31	0.29	0.31

The Table 2 includes the compression parameters of ten images by applying db6, db12, db16 and db20 mother wavelet functions, respectively. This table also provides average value of each of the parameters. The average values of CR, PSNR and BPP are found to be 3.40, 43.49 dB and 0.28, respectively with wavelet function db6. Similarly, the average values of CR, PSNR and BPP are obtained as 3.56, 43.54 and 0.28 (same as db6), respectively with wavelet function db12. Although the BPP values are same, but CR and PSNR values obtained using wavelet function db 12 are bit higher. Hence it is found to be better than db6 in terms of performance. The average compression parameters for db16 mother wavelet are found to be 3.73, 2.99, 43.39, and 0.30 for CR, MSE, PSNR and BPP respectively. Its CR is higher and PSNR is lower than the above two methods. For db20 mother wavelet, the average CR value is 3.81 which is higher than db6 and db12 both. But the average PSNR is 43.32 which is a little bit lower than db12 but higher than db6. Average value of BPP is 0.31 which is higher than the above two methods.

Table 3. Average value of performance parameters for all the mother wavelets

Evaluation parameters/mother wavelet	db6	db12	db16	db20
CR (avg)	3.44	3.59	3.73	3.84
MSE (avg)	2.93	2.92	2.99	2.99
PSNR (avg)	43.49	43.52	43.39	43.35
BPP (avg)	0.28	0.29	0.30	0.31

Table 3 summarizes the average values of all four parameters obtained using distinct mother wavelet functions. The averaged compression ratio and peak signal to noise ratio values with respect to distinct mother wavelet function db6, db12, db16 and db20 have been plotted in Fig. 5a and b, respectively for subsequent comparative analysis.



a. b.

Fig. 5. a. Comparison of average CR and b. Comparison of average PSNR with all four wavelet functions

It has been observed that the MRI image compression process using daubechies wavelet's basis functions (db6, db12, db16 and db20) embedded with SPIHT encoding technique is capable of providing improved results in terms of CR and PSNR. The compression ratio values calculated using DVQ-DCCR method were 2.7578, 2.7279 and 2.7314 [12] which are lower than the values achieved in the proposed method. Further, DCT in fusion with SPIHT encoding method was used for compressing images and PSNR values of 32.41 for 0.252 bits per pixel (BPP) were obtained [13]. Researchers have also implemented Cohen-Daubechies-Feauveau (cdf) wavelet for 512 × 512 image compression and obtained low average PSNR values of 35.80 with 0.75 BPP [14].

4 Conclusion

A detailed compression analysis of distinct MRI images has been presented using asymmetric wavelet functions. It utilized the properties of asymmetric wavelet functions viz. db6, db12, db16 and db20 along SPIHT. The improved values of compression ratio and peak signal to noise ratio has been obtained with all of the four wavelets used. The db20 wavelet gives a little bit higher CR (3.81) than the remaining two, but a lesser value of PSNR (43.54) than db12 wavelet. The mother wavelet function db16 provided good results in terms of CR but PSNR values obtained are little low. It reveals the importance of higher order basis function. In future cdf5/3 and cddf9/7 wavelets will be explored for lossless and lossy image compression respectively on a similar set of MRI images. Efforts shall be attempted to attain large dataset of medical MRI images for more reliable compression application and analysis.

References

1. He, C., Liu, Q., Li, H., Wang, H.: Multimodal medical image fusion based on IHS and PCA. Proceedia Eng. **7**, 280–285 (2010)
2. Strintzis, M.G.: A review of compression methods for medical images in PACS. Int. J. Med. Inform. **52**(1), 159–165 (1998)
3. Sriraam, N., Shyamsunder, R.: 3-D medical image compression using 3-D wavelet coders. Digit. Sig. Process. **2**(1), 100–109 (2011)
4. Bairagi, V.K., Sapkal, A.M.: ROI-based DICOM image compression for telemedicine. Sadhana Indian Acad. Sci. **38**(1), 123–131 (2013)
5. Chen, Y.T., Tseng, D.C.: Wavelet-based medical image compression with adaptive prediction. Comput. Med. Imaging Graph. **31**(1), 1–8 (2007)
6. Zhang, Y., Dong, Z., Wu, L., Wang, S.: A hybrid method for MRI brain image classification. Expert Syst. Appl. **38**(8), 10049–10053 (2011)
7. Zhou, X., Bai, Y., Wang, C.: Image compression based on discrete cosine transform and multistage vector quantization. Int. J. Multimedia Ubiquitous Eng. **10**(6), 347–356 (2010)
8. Jiansheng, M., Sukang, L., Xiaomei, T.: A digital watermarking algorithm based on DCT and DWT. In: Proceedings of the 2009 International Symposium on Web Information Systems and Applications (WISA 2009), Nanchang, PR China, pp. 104–107 (2009)

9. Korde, N.S., Gurjar, A.D.: Wavelet based medical image compression for telemedicine application. Am. J. Eng. Res. (AJER) **3**(1), 106–111 (2014). e-ISSN 2320–0847
10. Savakis, A.E., Carbone, R.: Discrete wavelet transform core for image processing applications. In: Electronic Imaging, International Society for Optics and Photonics, vol. 5671, pp. 142–151 (2005)
11. Rawat, P., Rawat, A., Chamoli, S.: Analysis and comparison of EZW, SPIHT and EBCOT coding schemes with reduced execution time. Int. J. Comput. Appl. **130**(2), 24–29 (2015)
12. Miaou, S.G., Chen, S.T., Chao, S.N.: Wavelet-based lossy-to-lossless medical image compression using dynamic VQ and SPIHT coding. Biomed. Eng. Appl. Basis Commun. **15** (6), 235–242 (2003)
13. Chen, Y.Y.: Medical image compression using DCT-based subband decomposition and modified SPIHT. Int. J. Med. Inform. **76**(10), 717–725 (2007)
14. Beladgham, M., Bessaid, A., Abdelmounaim, M.L., Abdelmalik, T.A.: Improving quality of medical image compression using biorthogonal CDF wavelet based on lifting scheme and SPIHT coding. Serb. J. Electr. Eng. **8**(2), 163–179 (2011)
15. Maly, J., Rajmic, P.: DWT-SPIHT image codec implementation. Department of Telecommunications, Brno University of Technology, Brno, Czech Republic (2003)
16. Said, A., Pearlman, W.A.: A new, fast and efficient image codec based on set-partitioning in hierarchical trees. IEEE Trans. Circ. Syst. Video Technol. **6**(3), 243–250 (1996)
17. Lakhdar, A.M., Rachida, M., Malika, K.: Robust image transmission performed by SPIHT and turbo-codes. Serb. J. Electr. Eng. **5**(2), 353–360 (2008)
18. Chen, Y.Y., Tai, S.C.: Embedded medical image compression using DCT based subband decomposition and modified SPIHT data organization. In: IEEE Symposium on Bioinformatics and Bioengineering, Taichung, Taiwan, pp. 167–175 (2004)
19. Sanchez, V., Abugharbieh, R., Nasiopoulos, P.: Symmetry-based scalable lossless compression of 3D medical image data. IEEE Trans. Med. Imaging **28**(7), 1062–1071 (2009)
20. Raid, A.M., Khedr, W.M., El-Dosuky, M.A., Ahmed, W.: JPEG image compression using discrete cosine transform - a survey. Int. J. Comput. Sci. Eng. Surv. (IJCSES) **5**(2), 39–47 (2014)
21. Roy, A., Saikia, L.P.: A comparative study on lossy image compression techniques. Int. J. Current Trends Eng. Res. (IJCTER) **2**(6), 16–25 (2016)

Can Tweets Predict Election Results? Insights from Twitter Analytics

Prabhsimran Singh[1](✉), Kuldeep Kumar[1], Karanjeet Singh Kahlon[2], and Ravinder Singh Sawhney[3]

[1] Department of Computer Engineering and Technology,
Guru Nanak Dev University, Amritsar, India
prabh_singh32@yahoo.com, kuldeep8437@gmail.com
[2] Department of Computer Science, Guru Nanak Dev University,
Amritsar, India
karankahlon@gndu.ac.in
[3] Department of Electronics Technology, Guru Nanak Dev University,
Amritsar, India
sawhney.ece@gndu.ac.in

Abstract. Social media has emerged as a powerful tool where people can share their common point of interest. Political events such as elections are one such event that attracts many social media users, to share their common view point and support towards a political party. This paper tries to capitalize this power of social media in order to predict the outcome of 2018 Pakistan General elections using Twitter as a tool. We fetched 33,468 tweets related to Pakistan Election over a span of 10 days. To get better insights of the results various social media analytic techniques were employed. Our results depict that Imran Khan led PTI was clear favorite among masses, which actually coincided with actual election results, where PTI emerged as a single largest party, making Imran Khan its Prime Minister.

Keywords: Election · Sentiment analysis · Social media analytics · Twitter · Tweets

1 Introduction

Social media has seen a tremendous growth in last one decade [1–4], it has provided a perfect environment for people to share their common interest (Like Business, Economic, Political & Social Issues) on these virtual platforms [5, 6]. One Such common interest which attracts a lot of people is elections, and social media's especially Twitter is extensively used for discussion for predicting the outcome of future elections [7, 8]. This paper explores one such election, which attracted a lot of people around the world, the 2018 Pakistan General Election. The 2018 Pakistan General Election took place on July 25, 2018. We collected data 10 days prior to the elections, as it is the perfect time in many ways like election campaign is in full swing and people show their full support to their desired candidates [9]. The aim of this paper was is to explore the use of Twitter by people for showing their support and checking whether the Twitter data can be used

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 271–281, 2019.
https://doi.org/10.1007/978-981-15-0108-1_26

as an effective tool for election prediction. The main three parties for 2018 Elections were Pakistan Muslim League-(N) (PML), Pakistan Peoples Party (PPP), and Pakistan Tehreek-e-Insaf (PTI). PML was led by Shahbaz Sharif [10], PPP was led by Bilawal Bhutto Zardari [11], while PTI was led by cricketer turned politician Imran Khan [12].

2 Review of Literature

From last one decade Twitter has been used as an extensive tool for predicting election outcome. With time this relationship has grown stronger, with approximately 326 million active Twitter users worldwide [13]. This strong online community constantly post intentionally or unintentionally regarding political activities [14]. In fact, almost 33% of the total discussion being carried out on Twitter is related to politics [15]. Tumasjan et al. [16], were the first to make use of Twitter to predict the outcome of 2009 German Federal Elections. Though, their technique was quite simple yet, it drew huge criticism specially, from Jungherr et al. [17] and Gayo-Avello [18, 19]. Gayo-Avello, further insisted to make use of sentiment analysis, as simply counting the number of tweets was not the solution of the problem. Various subsequent studies by DiGrazia et al. [20], Franch [21], Ceron et al. [22], Caldarelli et al. [23], Burnap et al. [24], Grover et al. [8] and Singh et al. [4]. Hence these studies provide us ample amount of motivation to carry out our research on 2018 Pakistan General Election. Through this research we try to address the following research questions:

(a) Can Twitter predict correct election outcomes?
(b) Do people prefer different parties in different provinces?
(c) What factors influence the difference in number of Twitter users in different provinces?

3 Research Methodology

This paper aims to predict the outcome of 2018 Pakistan General Elections, using Twitter as a tool. For this we use a very simple methodology, data for prediction is being fetched Twitter using Twitter API [25] in an authenticated manner on daily basis starting from July 15, 2018 to July 24, 2018 (10 days). Tweets were fetched based upon specific (#) hashtags associated with parties (#PML, #PPP, #PTI). A total of 33,468 tweets were collected in these 10 days. Since, these tweets contain a lot of unwanted stuff therefore, cleaning (preprocessing) of these tweets was necessary [26–28]. This preprocessing process included removing punctuations, removing web links, removing stop words, removing extra white spaces and converting the tweet into lower case. Once this preprocessing process was done, these tweets were ready for applying various social media analytics techniques [29], to give better insights about the election outcome. The details of these techniques are explained in upcoming section (See Sect. 4). R-language has been used to perform the entire experimentation.

4 Social Media Analytics

Social media analytics is a branch of data analytics which deals with extracting important information from data gathered from social media websites and using this important information for decision making [29]. This section is divided into various sub sections, each sub section depicting a social media technique. All these social media analytic techniques helps us to get better insights of 2018 Pakistan General elections.

4.1 Tweet Statistics

Tweet statistics deal with various statistics associated with tweets [30]. A total of 33,468 were fetched from 21,382 different users. This indicates that 5,871 users tweeted more than one tweet. Table 1 shows the results of tweet statistics.

Table 1. Tweet statistics

	PML	PPP	PTI
Total tweets	10351	8335	14782
Average tweets per day	103.51	83.35	147.82
Total unique senders	6967	5386	9029
Average tweets per sender	1.48	1.54	1.63

Imran Khan led PTI received maximum number of tweets (14,782), followed by PML (10351), while PPP were distant third with 8335 tweets. Similarly, maximum unique senders tweeted in favor of PTI, followed by PML and PPP.

4.2 Sentiment Analysis

Since, simply relying on tweet volume is not an efficient way for predicting results [18, 19]. Hence, sentiment analysis comes into play. Sentiment analysis is branch of text mining which deals with mining of people's opinion towards an entity [31, 32]. It is further categorized into (a) Polarity analysis [33, 34] (b) E-Motion Analysis [35, 36].

(a) **Polarity Analysis:** It deals with identifying the polarity (Positive or Negative) associated with given piece of text. The results of Polarity analysis is shown in Fig. 1. PTI not only got maximum tweets but also got the maximum number of positive tweets. Figure 2 shows top positive and negative words in form of a wordcloud.

(b) **E-Motion Analysis:** In E-motion analysis words are classified based upon eight emotions (Anger, Anticipation, Disgust, Fear, Joy, Sadness, Trust, and Surprise). Figure 3 shows the result of E-motion analysis. The results clearly show, PTI had greater appealing among masses in the country as compared to other two parties. The wordcloud of top words is shown in Fig. 4.

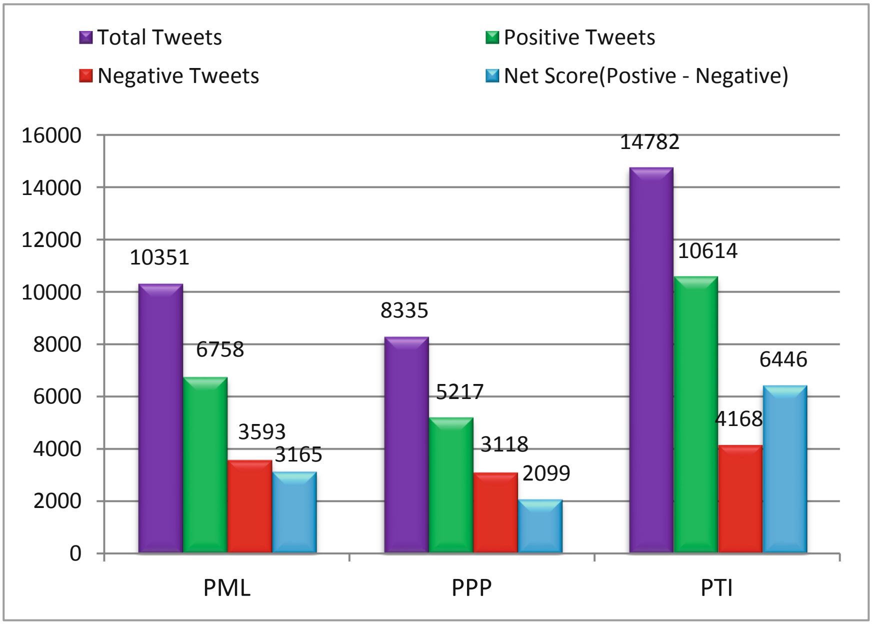


Fig. 1. Result of Polarity analysis

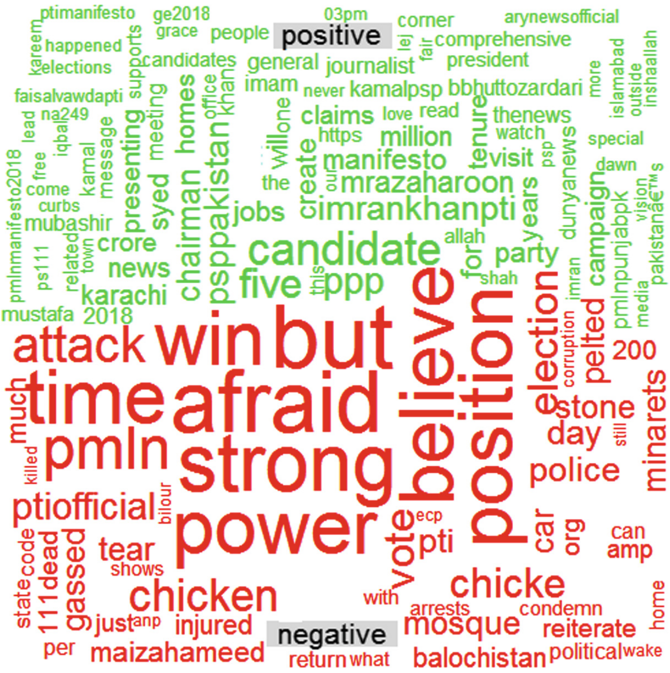


Fig. 2. Wordcloud of positive and negative words

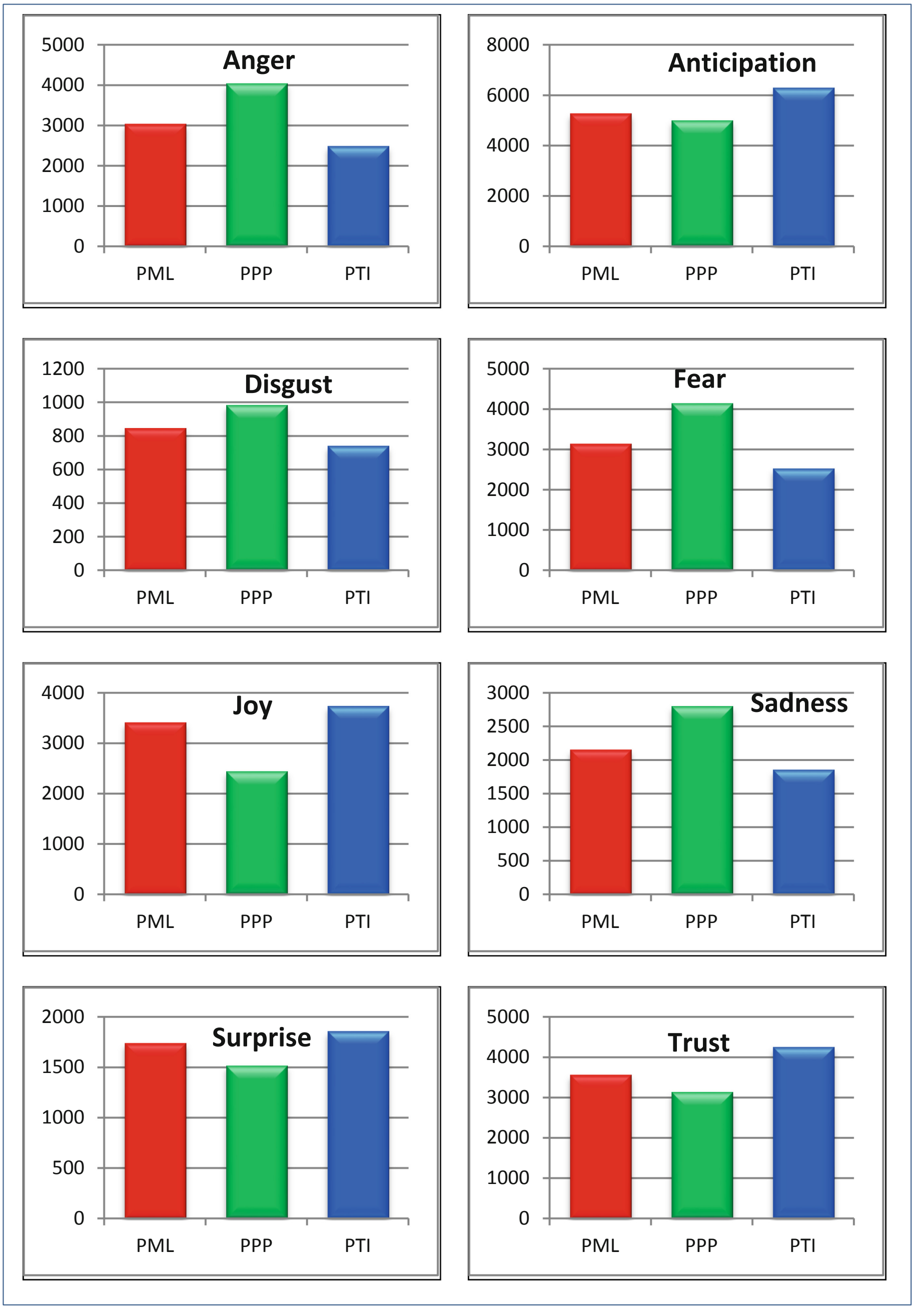


Fig. 3. Results of E-Motion analysis

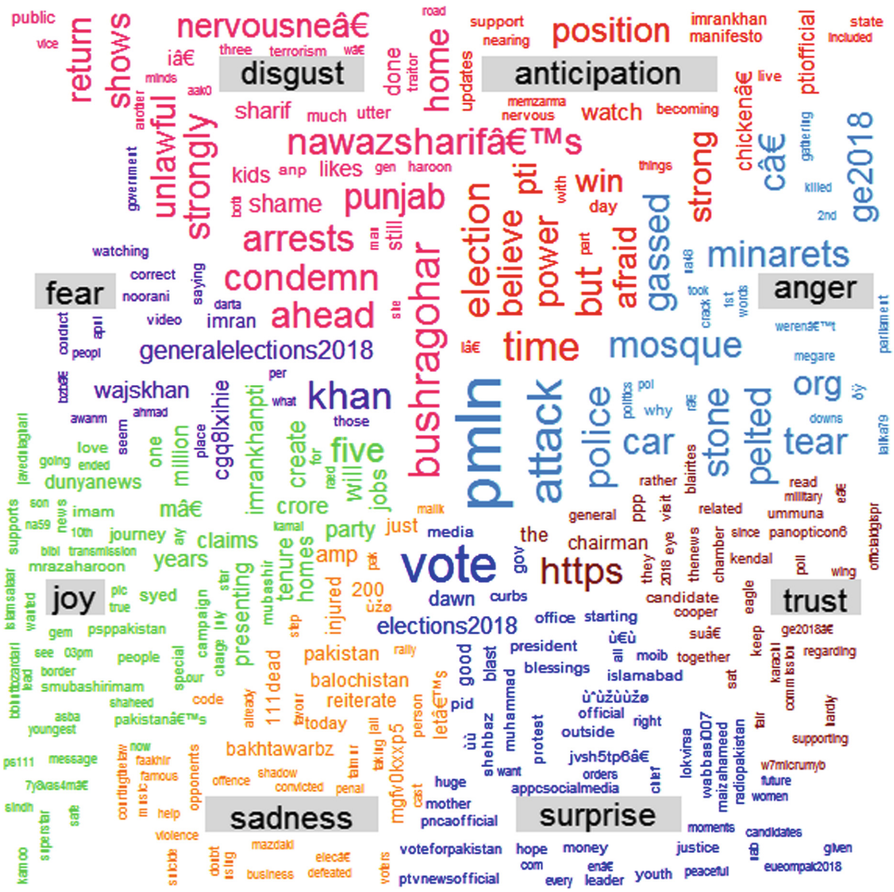


Fig. 4. Wordcloud of top words of E-Motion analysis

4.3 (#) Hashtag Analysis

(#) Hashtags are the indicator of checking what is trending on Twitter [37–39]. In total there were 6,749 unique hashtags, with #PML having maximum occurrences followed by #PTI. Though, hashtags such as #VoteForPTI and #ImranKhan were also prominent, showing trends in favor of PTI. In hashtag analysis our core focus was on the hashtag of three party names, in which PML got maximum occurrences. The results of hashtag analysis are shown in Fig. 5.

4.4 Geo-Location Analysis

Location based analysis is an important tool for gathering information, while mapping public opinion towards an event [40–42]. Given the diversity of country like Pakistan this type of analysis becomes utmost important. Since all the tweets fetched by us were not geo-tagged because of the Twitter privacy policy. Hence this result consisted of only those tweets which were geo-tagged. A total of 13,724 tweets were geo-tagged. Punjab and Sindh provinces contributed maximum, as these two have greater access to resources, better infrastructure and have higher literacy rate [43]. Province wise contribution of each province is shown in Fig. 6. In Punjab province PTI emerged as single largest party followed PML. In Sindh province PPP was the most tweeted party followed by PTI, while PML was distant third. In Khyber Pakhtunkhwa province PTI again got the maximum share of tweets, followed by PML. In Balochistan province PTI again emerged as a winner in terms of tweet share, followed by PML [44]. This showed that each party had different support base in different provinces. PTI was not only favorite party among masses at national level but also attracted the people in all provinces. Due to this PTI was one of top parties in each province, which made it the single largest party at the national level.

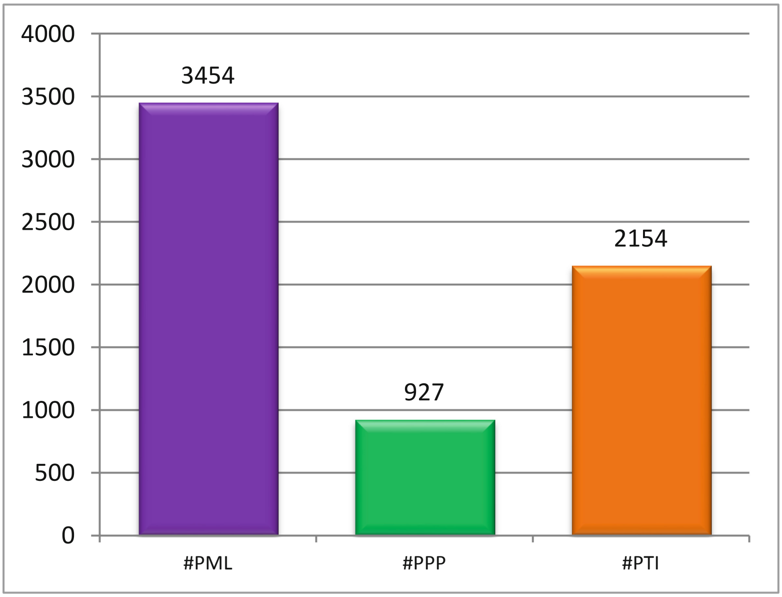


Fig. 5. Results of (#) Hashtag analysis

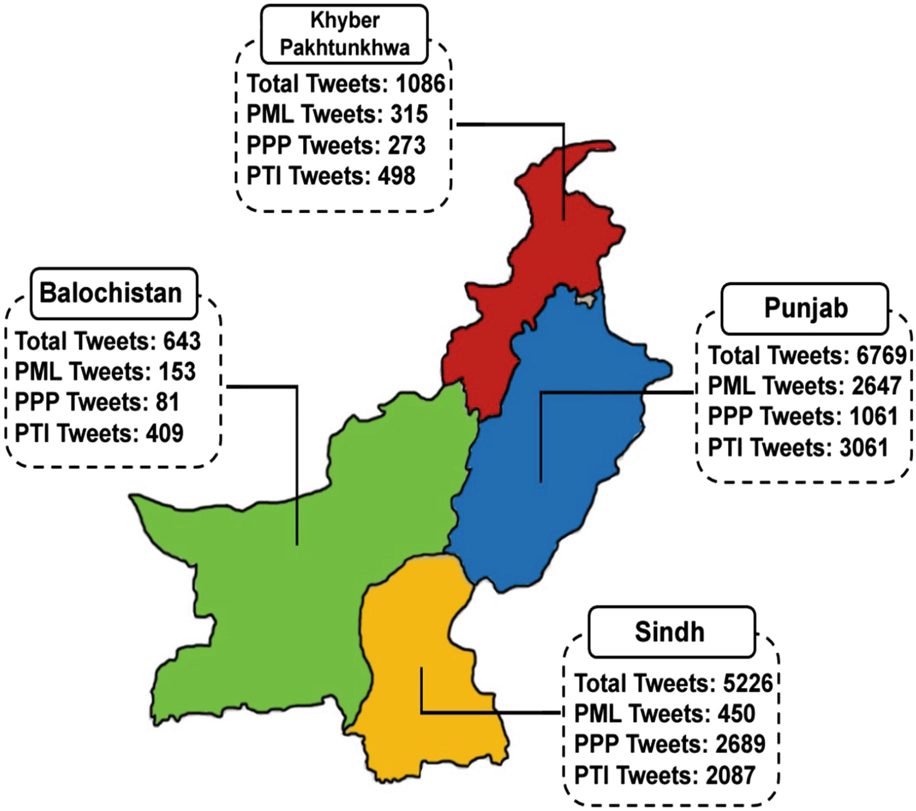


Fig. 6. Results of Geo-Location analysis

5 Conclusions

The aim of this paper was to predict the outcome of 2018 Pakistan General Elections, using Twitter as a tool. Our results clearly showed that Imran Khan Led PTI enjoyed a greater support as compared to the other two main opposition parties i.e. PML and PPP. This is evident from the fact that they not only got more number of tweet share, but also enjoyed better positive support as compared to their counter parts. These results do coincided with the actual election results, where PTI emerged as a single largest party [44]. Though results of Geo-Location analysis indicated that each province preferred one particular party, however PTI was among top parties in every province. So the first research questions that we tried to answer through this research i.e. ***"Can Twitter predict correct election outcomes?"***. The answer is yes, as depicted by our results; Twitter is indeed a handy tool to predict the election outcomes. Next research question ***"Do people prefer different parties in different provinces?"***. The answer is yes, people do prefer different parties in different provinces as depicted by the results of Geo-Location analysis. This also means each party has different support base in different parts of the country. Finally, the research question ***"What factors influence***

the difference in number of Twitter users in different provinces?”. The factors like infrastructure, growth rate and access to better facilities are the main contributors which affect the number of Twitter users in different provinces. Another important factor that was identified in our research was literacy rate of the province. Since Punjab and Sindh province had better literacy rate, their contribution was maximum in number of tweets.

Though we made a concrete effort to carry out our research in a fruitful manner, yet it still suffers from some limitations. One of the major limitations was the tweet collection period of 10 days. Though this 10 days data gave us good depiction of the results, however the data collected over a longer period would have been more useful in providing better results. Another important limitation of our work was that the data of the only top three political parties was considered for analysis. Further, no mechanism was adopted for bot detection [39]. So, these issues need to be addressed in future work while working on election predictions.

References

1. Beier, M., Wagner, K.: Social media adoption: barriers to the strategic use of social media in SMEs. In: ECIS, p. ResearchPaper100, June 2016
2. Shen, Y., Chan, H.C., Heng, C.S.: The medium matters: effects on what consumers talk about regarding movie trailers (2016)
3. Stieglitz, S., Bunker, D., Mirbabaie, M., Ehnis, C.: Sense-making in social media during extreme events. J. Contingencies Crisis Manag. **26**(1), 4–15 (2018)
4. Singh, P., Sawhney, R.S., Kahlon, K.S.: Forecasting the 2016 US presidential elections using sentiment analysis. In: Kar, A.K., et al. (eds.) I3E 2017. LNCS, vol. 10595, pp. 412–423. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68557-1_36
5. Kapoor, K.K., Tamilmani, K., Rana, N.P., Patil, P., Dwivedi, Y.K., Nerur, S.: Advances in social media research: past, present and future. Inf. Syst. Front. **20**(3), 531–558 (2018)
6. Singh, P., Sawhney, R.S.: Influence of Twitter on prediction of election results. In: Saeed, K., Chaki, N., Pati, B., Bakshi, S., Mohapatra, D.P. (eds.) Progress in Advanced Computing and Intelligent Engineering. AISC, vol. 564, pp. 665–673. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-6875-1_65
7. Bruns, A., Stieglitz, S.: Towards more systematic Twitter analysis: metrics for tweeting activities. Int. J. Soc. Res. Methodol. **16**(2), 91–108 (2013)
8. Grover, P., Kar, A.K., Dwivedi, Y.K., Janssen, M.: Polarization and acculturation in US Election 2016 outcomes–can Twitter analytics predict changes in voting preferences. Technol. Forecast. Soc. Change **145**, 438–460 (2018)
9. Singh, P., Sawhney, R.S., Kahlon, K.S.: Predicting the outcome of Spanish general elections 2016 using Twitter as a tool. In: Singh, D., Raman, B., Luhach, A., Lingras, P. (eds.) ICAICR 2017. CCIS, vol. 712, pp. 73–83. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-5780-9_7
10. Dawn. https://www.dawn.com/news/1416141
11. Tribune (a). https://tribune.com.pk/story/1746890/1-bilawal-kicks-off-election-campaign-karachi/
12. Tribune (b). https://tribune.com.pk/story/1740378/1-pti-kick-start-election-campaign-mianwali/
13. Statista. https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/
14. Hossain, M.A., Dwivedi, Y.K., Chan, C., Standing, C., Olanrewaju, A.S.: Sharing political content in online social media: a planned and unplanned behaviour approach. Inf. Syst. Front. **20**(3), 485–501 (2018)

15. Singh, P., Dwivedi, Y.K., Kahlon, K.S., Sawhney, R.S.: Intelligent monitoring and controlling of public policies using social media and cloud computing. In: Elbanna, A., Dwivedi, Y., Bunker, D., Wastell, D. (eds.) TDIT 2018. IFIPAICT, vol. 533, pp. 143–154. Springer, Cham (2018)
16. Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welpe, I.M.: Predicting elections with Twitter: what 140 characters reveal about political sentiment. In: Fourth International AAAI Conference on Weblogs and Social Media, May 2010
17. Jungherr, A.: Tweets and votes, a special relationship: the 2009 federal election in Germany. In: Proceedings of the 2nd Workshop on Politics, Elections and Data, pp. 5–14. ACM, October 2013
18. Gayo-Avello, D., Metaxas, P.T., Mustafaraj, E.: Limits of electoral predictions using Twitter. In: Fifth International AAAI Conference on Weblogs and Social Media, July 2011
19. Gayo-Avello, D.: "I wanted to predict elections with Twitter and all I got was this Lousy paper"–a balanced survey on election prediction using Twitter data. arXiv preprint arXiv: 1204.6441 (2012)
20. DiGrazia, J., McKelvey, K., Bollen, J., Rojas, F.: More tweets, more votes: social media as a quantitative indicator of political behavior. PLoS ONE **8**(11), e79449 (2013)
21. Franch, F.: (Wisdom of the crowds) 2: 2010 UK election prediction with social media. J. Inf. Technol. Politics **10**(1), 57–71 (2013)
22. Ceron, A., Curini, L., Iacus, S.M., Porro, G.: Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. New Media Soc. **16**(2), 340–358 (2014)
23. Caldarelli, G., et al.: A multi-level geographical study of Italian political elections from Twitter data. PLoS ONE **9**(5), e95809 (2014)
24. Burnap, P., Gibson, R., Sloan, L., Southern, R., Williams, M.: 140 characters to victory?: using Twitter to predict the UK 2015 general election. Electoral Stud. **41**, 230–233 (2016)
25. Twitter API. https://www.nuget.org/packages/TweetinviAPI/
26. Liu, Y., Chen, Y., Wu, S., Peng, G., Lv, B.: Composite leading search index: a preprocessing method of internet search data for stock trends prediction. Ann. Oper. Res. **234**(1), 77–94 (2015)
27. García, S., Luengo, J., Herrera, F.: Data Preprocessing in Data Mining. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-10247-4
28. Haddi, E., Liu, X., Shi, Y.: The role of text pre-processing in sentiment analysis. Procedia Comput. Sci. **17**, 26–32 (2013). https://doi.org/10.1016/j.procs.2013.05.005
29. Stieglitz, S., Dang-Xuan, L.: Social media and political communication: a social media analytics framework. Soc. Netw. Anal. Min. **3**(4), 1277–1291 (2013)
30. Purohit, H., Hampton, A., Shalin, V.L., Sheth, A.P., Flach, J., Bhatt, S.: What kind of# conversation is Twitter? Mining# psycholinguistic cues for emergency coordination. Comput. Hum. Behav. **29**(6), 2438–2447 (2013). https://doi.org/10.1016/j.chb.2013.05.007
31. Mishra, N., Singh, A.: Use of Twitter data for waste minimisation in beef supply chain. Ann. Oper. Res. **270**, 337–359 (2018)
32. Liu, B.: Sentiment analysis and opinion mining. Synth. Lect. Hum. Lang. Technol. **5**(1), 1–167 (2012). https://doi.org/10.2200/S00416ED1V01Y201204HLT016
33. Mohammad, S.M., Turney, P.D.: Emotions evoked by common words and phrases: using mechanical turk to create an emotion lexicon. In: Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, pp. 26–34. Association for Computational Linguistics, June 2010
34. Ou, G., et al.: Exploiting community emotion for microblog event detection. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1159–1168 (2014)

35. Saif, H., Fernandez, M., He, Y., Alani, H.: Evaluation datasets for Twitter sentiment analysis: a survey and a new dataset, the STS-Gold (2013)
36. Yuan, H., Xu, W., Li, Q., Lau, R.: Topic sentiment mining for sales performance prediction in e-commerce. Ann. Oper. Res. **270**, 553–576 (2018)
37. Chae, B.K.: Insights from hashtag# supplychain and Twitter analytics: considering Twitter and Twitter data for supply chain practice and research. Int. J. Prod. Econ. **165**, 247–259 (2015). https://doi.org/10.1016/j.ijpe.2014.12.037
38. Singh, P., Sawhney, R.S., Kahlon, K.S.: Sentiment analysis of demonetization of 500 & 1000 rupee banknotes by Indian government. ICT Express **4**(3), 124–129 (2018)
39. Singh, P., Kahlon, K.S., Sawhney, R.S., Vohra, R., Kaur, S.: Social media buzz created by# nanotechnology: insights from Twitter analytics. Nanotechnol. Rev. **7**(6), 521–528 (2018)
40. Singh, P., Dwivedi, Y.K., Kahlon, K.S., Sawhney, R.S.: Intelligent monitoring and controlling of public policies using social media and cloud computing. In: Elbanna, A., Dwivedi, Y.K., Bunker, D., Wastell, D. (eds.) TDIT 2018. IAICT, vol. 533, pp. 143–154. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-04315-5_11
41. Singh, P., Sawhney, R.S., Kahlon, K.S.: Twitter based sentiment analysis of GST implementation by Indian government. In: Patnaik, S., Yang, X.-S., Tavana, M., Popentiu-Vlădicescu, F., Qiao, F. (eds.) Digital Business. LNDECT, vol. 21, pp. 409–427. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-93940-7_17
42. Amirkhanyan, A., Meinel, C.: Density and intensity-based spatiotemporal clustering with fixed distance and time radius. In: Kar, A.K., et al. (eds.) I3E 2017. LNCS, vol. 10595, pp. 313–324. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68557-1_28
43. Pakistan Literacy Rate. https://pakobserver.net/literacy-in-pakistan-2/
44. Pakistan Election Results. https://www.ecp.gov.pk/default.aspx

Deep Learning in Image Processing: Revolutionizing Diagnosis in the Field of Dermatology

Richa Nayak and Yasha Hasija(✉)

Delhi Technological University, Bawana Road, Shahbad, Delhi, India
ricnl996@gmail.com, yashahasija06@gmail.com

Abstract. The modern-day society is increasingly dependent on computer-aided tools and techniques. Digital imaging techniques have a tremendous impact on our day-to-day lives. Image processing is a vital component in the field of biological sciences and has the potential to drastically change the computer-human interface. Image processing refers to the conversion of an image into a digital form followed by enhancement of the image in order to extract useful information from it that are indiscernible by human ocular perceivers. Rapid advances in image processing, computerized reconstruction of an image and allied advancements in image analysis algorithms and the application of artificial intelligence has spurred a revolution in the field of medical and diagnostic imaging. Deep learning, a type of Artificial Neural Network (Machine Learning), is resurfacing as a powerful tool for its utilization in big healthcare data. The integration of deep learning techniques to image processing has the potential to add momentum to the dermatological imaging and promote early and accurate diagnosis of skin lesions. This review attempts to discuss the fundamentals of image processing, its importance, various clinical imaging modalities in use in the field of dermatology and application of deep learning algorithms in dermatological imaging, accentuating the inadequacies and future research prospects.

Keywords: Image processing · Image segmentation · Dermatological imaging · Deep learning

1 Introduction

The advancements in computer availed diagnostic systems began in the 1980s. Owing to the advent of computer vision techniques and rapid advancements in the field, analyzing medical imaging data gained momentum. Today, significant research in digital image processing has provided us with better diagnostic techniques and early prediction of diseases with an accuracy that was not possible before. The modern-day diagnostic system is colossally dependent on computer-aided techniques. Medical image processing is the core field that combines bioinformatics, medical informatics and neuroinformatic. Biomedical image processing indulges an interdisciplinary approach including a myriad of fields like applied mathematics, physics, statistics, computer science, medicine etc. Since technological advancement is directly proportional to

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 282–293, 2019.
https://doi.org/10.1007/978-981-15-0108-1_27

advancement in the field of medical sciences, computer aided image processing got quickly adopted as an important part of the clinical routine. Imaging modalities and image processing makes use of non-invasive methods for diagnostic and research purposes and thus is gaining popularity in the field of biomedical sciences. As it involves computer-aided processing, the chances of human error are minimized. It plays an important role in dermatological imaging. The skin is the most exposed and easily diagnosable organ, it would appear so, but often severity of skin lesions is underestimated or misdiagnosed. Also, the currently available techniques for diagnosis of dermatological disorders are not only time consuming but also invasive in nature. The gold standard for detection of skin lesions has been dermoscopy which has limited diagnostic accuracy due to the complexity of visual inputs and the its dependency on physician's skills. Thus, a diagnostic upscaling was due. Digital image processing is an important non-invasive method for diagnosis of various dermatological disorders including melanomas, which is accessible and affordable. The application of deep learning techniques to image processing owing to the availability of big data further propagates more accurate diagnosis.

This review aims to present a cumulated overview of image processing and the application of deep learning algorithms, correlating it with dermatological image processing.

1.1 Digital Image Processing

Image Processing is a process of converting an image i.e. a defined array or matric of square pixels arranged in rows and columns, into a digital form to which certain algorithms can be applied so as to get an improved image that will provide us with desirable information. With image processing we aim to go beyond the two-dimensional and understand the underlying intricacies of an image. Mathematically, processing of a 2D image by a computer as a function of two variables t(x, y) with an amplitude such as brightness of an image at coordinate point (a, b). Image processing entails processing of an input image, image analysis and image extraction that gives us an output. A complete digital image processing system comprises of both software and hardware components. The following steps are involved in standard image processing:

(a) Image acquisition: Image is acquired using appropriate sensors to detect field or radiation and capture all possible features of the image. In case of an analog image (continuous image), it is first digitized using an analog-to-digital converter.
(b) Storage of image: Image is stored for further processing, either in an image format, using RAM (read/write memory devices) or, in CD or flash drives.
(c) Manipulation or processing of the image: This is the most vital step and involves image reconstruction, image enhancement, image analysis, compression and synthesis.
(d) Image display: Image is then displayed on a monitor. It comprises lines of continuously varying (analog), intensity. This is achieved using a digital-to-analog converter.

The input image is sampled on a separate lattice and each sample or pixel is quantized by a fixed number of bits [1]. For a digitized image to be shown, it first needs to be converted into a signal (analog) i.e. scanned onto an output.

In the case of medical imaging systems, the input signals are obtained from different body parts of patient, using ultrasound reflection or simply x-ray attenuation. The images then obtained can either be analog (continuous), or digital (discrete) and upon digitization, the prior could easily be transformed into the latter. It must be kept in mind that the objective is to acquire an image in the output that is a precise depiction of the signal in input, and later to analyse and extract information that can help in diagnosis from the image (Table 1).

Table 1. Examples of computerized image processing operations [2]

Classes	Operation
Enhancement of image	Enhancement of brightness, adjustment of contrast, convolution, averaging of image, etc.
Restoration of image	Inverse filtering and photometric correction
Analysis of image	Feature extraction and segmentation
Compression of image	Lossy compression and lossless compression
Synthesis of image	Reconstruction, tomographic imaging

In this review, we discuss several basic deep learning-based techniques of image recognition and image processing that prove useful for analyzing biological images.

1.2 Deep Learning in Image Processing

Deep learning or deep-structured learning is a Machine Learning method that is an extension of Artificial Neural Network that bears a resemblance to the connected, multitiered and extensive human cognition system. In the last decade, there has been a huge increase in accessibility of Big Data, enhanced computational strength with GPU (graphics processing units) and advanced algorithms used in training DNNs (Deep Neural Networks). DNN draws inspiration of their structure by mimicking human neural network and have exhibited impressive performances in a plethora of fields, which include medical imaging. A major task achieved by this boost was classification of disease categories based on detection of structural abnormalities. Ever since the introduction of ML, numerous new algorithms with different logical theories, mathematical equations, and implementations have been executed to perform classification and regression tasks [3]. Until last decade numerous computer-aided detection (CAD) systems were developed and incorporated in the working clinicals. But these were not fool proof and were often prone to more false positive errors than a diagnosis by an expert, which led to a greater assessment time. Therefore, the advantages of using CAD for diagnostics purpose was debatable. Current deep learning technology are expected to overcome these limitations and attain pronounced precision in detection and also aid experts in diagnosis by doing this repetitive run-of-the-mill job for them, so that they may be more productive.

This new Artificial Intelligence-based technology is well suited to handle medical big data, and can put into perspective the large amount of data that has accumulated due to high throughput technology. Deep learning has the potential to compose preliminary reports by performing automatic lesion detection and differential diagnoses [3].

1.2.1 Artificial Neural Networks

Machine Learning is an application of AI that empowers the systems with the ability to perform tasks by automatically learning and improving with experience, without being explicitly programmed. For this purpose, the system needs to be familiarized with a dataset, called training data. The techniques of machine learning are classified into two broad categories- supervised and unsupervised learning. In supervised learning a function is generated that yields an output based on inference drawn from training data based on a certain set of instructions that are provided at the time of training. The training process is called *regression,* when the output data yields a continual value [3] and *classification,* if it has a categorical value. Whereas unsupervised learning involves generating a function by assessing hidden features in unlabeled input data. The training phase involves pre-processing of training data set and meaningful feature extraction. The pre-processing in case of image analysis comprises of operations such as image reconstruction, noise reduction, image enhancement, feature extraction etc. Feature extraction is a challenging task and thus, it is pre-requisite to design such unique features for every novel application, especially when it is of medical importance. This methodology is commonly known as feature "*hand-crafting*" in the literature of deep learning [4]. Based on the feature vector $x \in R^n$, the classifier has to predict the correct class y, which is typically estimated by a function $\hat{y} = f(x)$ that directly results in the classification result $\hat{y}$. The classifier's parameter vector θ is determined during the training phase and later evaluated on an independent test data set.

Artificial Neural Network is a popular regression and classification algorithm in Machine learning, that models' multiple layers of computing units by mimicking the architecture and signal transduction mechanism of neurons in human brain. It is a rendition of the human neural network and consists of a network of interconnected artificial neurons. Each artificial neuron gears a relatively basic classifier model that yields a decision signal as output on the basis of the given "weighted sum of evidences". To establish a full-fledged ANN, some hundreds of such models or computing units are put together. Like human brain relies on external stimuli to perform tasks, similarly, weights in networks are trained based on a learning algorithm like back propagation. Features in machine learning can be defined as the numerical and nominal values used in the input data. Defining meaningful and powerful features is cardinal to machine learning studies. ANN has shown remarkable outputs in diverse fields, but has also had several shortcomings such as a overfitting and decreased local minima all through optimization technique [4]. In recent years, prognostic methods based on ANN have shown exceptional ability in solving non-linear modelling problems, but predominantly shallow architectures were used as training deep networks was a challenging task. Deep architecture has drawn a lot of attention recently owing to fast learning algorithms that have been proposed and deep ANN have proved to outperform the conventional methods in classification, pattern recognition and other domains of machine learning. DNN comprises of a series of stacked layers. Prediction is made based on the first layer (input) and output (last layer) predicting a class or a value. The underlying layers between output values and input values are called '*hidden layers*', since their condition does not accord to observed data. The multi-layered structure of the ANNs enables them to foster decisions that are often complex and undecipherable. To cite an example, a complicated task like classification of tumors from pixel to curve to shape and to feature is possible using a neural network model that has deeper hidden layers [3]. For specific training samples, each edge necessitates optimized weights- These weights use a sum of a wide range of parameters that are initialized randomly and configured by an optimization algorithm, like 'gradient descent', in order to find a local minima of a function by steps which are relative to the negative gradient of the function at the point [3]. Once the training samples have been applied to the network, a regression value or loss function is evaluated between the input class and the prediction class. The ultimate goal is minimization of this loss function and the parameters are updated accordingly. Deep learning is further modified and applied using various approaches. DNN has proved to perform better in pattern recognition and prediction studies of the complex type and thus its adoption in the field of medical imaging is well justified.

2 Image Processing Progression

This section outlines various steps involved in image processing.

2.1 Enhancement of Image

Using software tools to manipulate a stored image digitally is the method underlying image enhancement. It encompasses adjustment of brightness, contrast or manipulation of the gray scale and RGB color patterns of an image [5]. It also entails smoothing of an image in case it contains a lot of noise or speckle. More sophisticated types of image enhancement tools are capable of applying changes to certain parts of an image such as de-blurring, filtering etc. Diving into the technicalities of image enhancement, we encounter:

Gray-Level Transformation

Gray-level transformation is an image processing technique that is used in the conversion of a gray-level value to another value. A mapping function is used to keep track of the gray-scale transformations, called *tone curve* [6]. This curve has the ability to convert dark gray values to darker ines, and light ones to lighter, thereupon, making the difference between brighter and darker regions, more apparent. In order to visualize different aspects, one can play with the shape of the tone curve.

Tone Curve Alignment

Tone curve alignment follows suit. It is a technique used to remove the brightness difference between two images captured at different intervals. This difference occurs due to automatic aperture and shutter control during image acquisition or due to either of them [6]. It is advisable to nullify this difference before continuing with additional processes. In a case where such a difference is not nullified, a value of a parameter suiting image X may not suit image Y. It could so happen, that they are recognized as very different images altogether since the difference in their brightness levels is non-negligible.

Binarization

Binarization is another important component of image enhancement. It is an operation that is used to convert an original grayscale image into a black-and-white image [6]. It has numerous applications such as, distinction between a bright target object from a dark background and vice versa. The resulting image contains regions of black and white pixels. Each of these regions is called a *connected component*. It is possible to analyse the shape and size of each target object by observing its corresponding connected component.

Smoothing

Smoothing is a feature used to lower the noise and speckles in an image since noise causes a difference in the gray-level of neighbouring pixels. Smoothing aims to minimize this difference in gray-level. It is pivotal to image analysis.

Edge Detection

Edge detection is another important filter used in image processing. The definition of an ‘edge’ is a number of pixels with a significant difference in the pixel value. If on a black background an image has a white filled centre, then the boundary between the white centre and black background is the edge. By using edge detection filter, pixels on the edge are highlighted. The essential thought behind filtering of edge detection is computing the first or the second-order derivatives of pixel values.

2.2 Restoration of Image

Image restoration is the process of reversing the degradation caused by the imaging system such as unequal illumination, distortion caused by poorly focused lenses, noise, improper acquisition due to non-linear detectors etc. Image restoration would require modelling of the degradation caused to the image by any of the aforementioned inadequacies during image acquisition and then applying an inverse operation to reverse the damage. Image reconstruction techniques play a major role in medical imaging. They can make use of sets of 1D projections to make 2D and 3D images. Image reconstruction in Computer Tomography (CT) scans, is a mathematical algorithm that takes x-ray projection data acquired from different angles of a patient and generates a tomographic image for the same. With advancements in image reconstruction techniques there are more efficient ways to compute the attenuation coefficients of different x-ray absorption paths that are obtained as a set of data. There are various algorithms for image reconstruction. The ones that are of importance to medical imaging like CT are iterative algorithm that doesn't take into account statistical modelling and analytical algorithm with statistical modelling.

2.3 Analysis of Image

Image analysis is the most important component of image processing. It is the step where feature extraction is carried out. To put it in simpler words, image analysis involves identifying objects within an image and taking its measurements. The first step is to identify and isolate the objects of interest from the lot. This is called segmentation of the image. The second step is classification based on features such as shape, size and texture. Classification permits the categorization of an object, whether or not it belongs to a particular group based on a set range of tolerance for each group. This process is called *pattern recognition*. Pattern recognition can be used in the classification of images of biological samples. Like in case of lesions or suspicious clusters of micro-calcifications, pattern recognition helps classify them into benign or malignant.

Segmentation

Image Segmentation is crucial to image processing of biological samples. It involves partitioning of an input image into various regions. It has multiple purposes; for example, counting objects, measuring the 2-D (or 3-D) distribution of different objects, measuring their appearance or shape, recognizing them individually, localizing of objects for their tracking, removal of unnecessary regions, etc. [6]. This makes image segmentation the most challenging part of image processing. Even though human eyes have the ability of performing image segmentation with much ease, the error value could be too much for it to be of any significance. On the other hand, computers encounter difficulty in performing this task but with advanced and improving techniques, computers are much more reliable to give more accurate predictions than humans. For a very long time we have not perfected an algorithm for segmentation of an image, even for simple images like facial recognition. Biological images are much more complex. In case of biological images, the target objects have ambiguous boundaries and thus we are faced with a challenge while separating the object from other objects and the background. We discuss various segmentation techniques below that are used based on the requirements.

i. ***Background Subtraction*** - It is a method of separation of target object from an input image, using a background image. Background image is an image that has no target object. For example, in the biological images, let us say we have to examine a tumor cells then the background image would be an image of a normal tissue taken from the same angle as that of the tissue to be examined otherwise we need to take into account the distortion before making any analysis.

ii. ***Watershed method*** - It is a method of segmenting images for biological images, such as - nuclei and cell segmentation [7]. "Watershed" refers to the ridge lines of a 3D surface like a ground-surface [6]. In the watershed method, a region surrounding a closed ridge line is considered as a partitioned region. As a result, its property depends on the method of derivation of an input image. By accounting a gray-scale value at a pixel, as height at the pixel location, any gray-scale image can be represented as a 3D surface; this representation, however, is not appropriate for the watershed method. This is because its local peaks often do not correspond to a segmentation boundary. For example, consider an image where a white-filled centre lies on a black background. The boundary pixel of the central object will not correspond to a local peak of the gray-scale surface.

iii. ***Region growing*** - In this process each pixel is considered as a segment. Neighbouring segments that have any similar characteristics are merged into a new segment. This merging procedure is repeated on loop until convergence and segmentation result is obtained.

iv. ***Clustering*** - It is a method of partitioning a set into its subsets based on some specified criteria [8]. Elements having similarities are clustered in a subset. Clustering can help discover latent groups when a large number of samples are available. For example, when dealing with biological images, we usually encounter a large number of datasets that often have latent groups that need to be identified for proper diagnosis.

v. ***Active contour model*** - It is the method to find a closed contour of a target object [9]. It aims to find out the most plausible contour from all possible contours based on certain specified criterion.

vi. ***Recognition method based on template matching*** - Image segmentation is basically an image pattern recognition problem. For example, in cell image segmentation, various organelles can be assigned a class such as golgi class, nucleus class, lysosome class, mitochondria class etc. Template matching is a simple realization of image segmentation that makes use of pattern recognition techniques. It can be used in various ways. It is possible to use a small region i.e. a defined block as the unit of recognition, instead of a pixel. One can perform segmentation taking into account specific textures like spatial patterns and other elaborate similarity evaluation indices. Class consistency of neighbouring pixels can also be considered.

vii. ***Markov random field*** - It is a more integrated method as compared to other methods, to optimize the segmentation result by taking into account the similarity between neighbouring pixels. It is similar to binarization. As in case of binarization, each pixel was assigned to either of the two classes i.e. black or white. Segmentation results using MRF allows the assignment of each pixels to arbitrary classes, exceeding two classes.

After segmentation, the image is ready to be saved often in a compressed format. Image compression aims at reduction of data needed to describe an image without losing its important properties. Compression is achieved as images often contain redundant or repetitive information. Medical images are preserved in compressed file format, the kind of compression that is lossless.

2.4 Image Synthesis

Image synthesis involves the creation of a novel images from various different images or from non-image data. A major example is of CT, it involves rebuilding of 1D projections obtained by x-ray, to give rise to tomographic image.

Processing of image is not a one step process. It involves a number of steps that must be performed in a sequential pattern in order to extract relevant information from an image. The hierarchy in the processing steps is unique to each requirement.

3 From Algorithm to Application

Advancement in image processing algorithms and deep learning techniques has facilitated early diagnosis of various diseases. Some of the major applications are listed in Table 2.

Table 2. Various applications of digital imaging in the field of medical sciences.

Field	Application
Medical diagnostic imaging	This involves x-ray computed tomography (CT) using transmission of x-rays and radiography projection, digital subtraction angiography (DSA) which produces improved images of the blood vessels, mammography which produces images of the soft tissue in the breast Nuclear medicine also makes use of image processing using gamma ray emission from radiotracers that are injected into the body including emission computed tomography and planar scintigraphy (SPECT and PET) Ultrasound imaging is another important application that uses the technique of reflecting ultrasonic waves inside the body MRI or magnetic resonance imaging uses the precession of spin systems in a large magnetic field which includes fMRI or functional MRI
Biological imaging	Analysis, classification and matching of 3D topology of the genome Automated classification of cell type based on morphology Visualization of proteins in vivo

As we all know, early diagnosis is the key to effective treatment of any disease. Computer aided image processing has played a vital role in diagnosis of dermatological disorders, replacing the outdated method of dermoscopy and analysis by the medical practitioner, which is prone to human error. Computerized pre-screening of suspicious skin lesions and moles for malignancy accelerates the diagnosis process and helps in the determination of the treatment regime.

The color of the lesion on patient, in dermatology, determines strongly - the diagnosis step. Example – pigmentation of the color black or brown with uneven distribution of color is an indicator of melanoma. An accurate skin lesion diagnosis requires a preliminary examination of subtle differences in the color of lesions, manual analysis of the lesions is not very efficient and might not give accurate predictions [10]. Use of deep learning algorithms has made the process more effortless. Automatic segmentation of the surrounding skin from melanoma region is a necessary step in computational analysis of dermoscopic pictures [11]. Nevertheless, this chore is not unimportant since melanoma has a varied forms of appearance and comes with different, shape, size and color along with variations with skin types and texture [12] (Fig. 1).

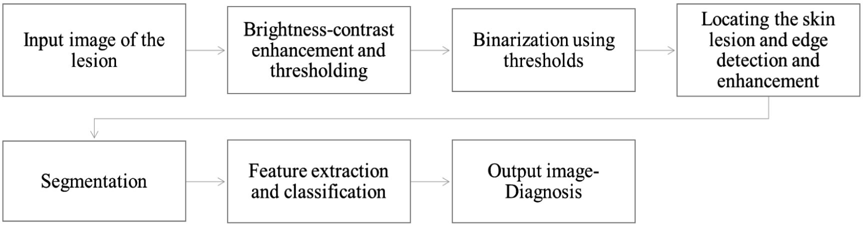

Input image of the lesion
Brightness-contrast enhancement and thresholding
Binarization using thresholds
Locating the skin lesion and edge detection and enhancement
Segmentation
Feature extraction and classification
Output image-Diagnosis

Fig. 1. Flow diagram depicting the basic steps of image processing in dermatology.

According to the WHO statistics of skin cancer, the occurrence of melanoma and non-melanoma skin cancers, both, has been on the rise in the past several decades. Each year, about 132,000 melanoma skin cancers and 2–3 million non-melanoma skin cancers occur around the world. About 1 in every 3 cancers diagnosed is a skin cancer, as per estimates. Due to global warming and depletion of ozone layer at a concerning rate, the atmosphere fails to filter the harmful solar UV radiations. It is projected that a decrease in 10% of ozone levels will result in an increased incidence of melanomas. With this alarming incidence rate of skin cancers, a fast and effective technique for diagnosis is important.

In 2017, a group of researchers [12] devised methods for automatic melanoma recognition using deep fully convolutional networks with Jaccard distance. It is a completely self-sufficient atomized method for segmentation of skin lesion which makes use of an extensive 19-layer deep CNN (convolutional neural network) that does not rely on prior knowledge of the data and has been trained end-to-end. They proposed several strategies to make sure of efficient and operative learning with limited data for training. Moreover, a novel loss function was designed by them based on the 'Jaccard distance' to remove the need to re-weight the sample, which is a typical method while using cross entropy as the loss function for segmenting the image because of a strong inequality between the number of pixels in the background and the foreground. Furthermore, their method was found more superior when its efficiency, effectiveness and the generalization ability of the framework proposed was evaluated on two publicly available databases. In 2017, computer scientists at Stanford also set out to create a machine learning based algorithm for detection of skin cancer and with inspiring accuracy. A new method has been proposed based on deep neural networks is proposed for accurate extraction of region of a lesion [14]. In 2018, another group of researchers [13] employed deep learning framework consisting of two fully convolutional residual networks to simultaneously produce the segmentation result and the coarse classification result for dermoscopic images of skin lesions.

4 Discussion and Future Directions

The use of deep learning in dermatological image analysis is not exhaustive, rather it opens up a wide range of possibilities, bringing diagnosis to the home desk. There are still various challenges that need to be addressed. Improved accuracy and efficiency of computer aided diagnosis by use of deep learning technologies would greatly assist dermatologists for improved diagnosis of challenging lesions and design a better treatment regime for patients. But before deep learning can transfigure dermatological image processing, in its full capacity, certain lacunas need to be addressed. Most deep learning algorithms focus on supervised learning but a vast majority of medical big data is unannotated thus advanced data augmentation techniques need to be developed. Outsourcing annotations, transfer learning etc. can be used to address data labelling dilemma. The "black box" issue i.e. the CNN models do not reveal the features they rely upon to arrive at decisions and thus the proverbial black box issue needs to be dealt with. Rigorous prospective validation and more efficient algorithms need to be developed, before such modalities can be fully implemented in clinical practice.

5 Conclusion

The field of digital image processing and its application in the field of biomedical sciences is unlimited in dimensions. The nature of the application-oriented and specific domain-dependent image processing has suggested the use of deep learning, computer vision and artificial intelligence. It is evident that these approaches will add momentum to diagnostic methods and will not only generate faster, more precise biomedical images but will also facilitate incorporation of contextual information for better diagnostic capabilities utilizing one or more knowledge databases. In this paper, we have only covered few prominent deep learning applications that have been employed with commonly known image processing and analysis techniques for dermatological imaging, which is far from exhaustive. Despite rapid advancement in the field, we are still a long way from accurate and fully automated diagnosis of diseases.

References

1. Basavaprasad, B., Ravi, M.: A study on the importance of image processing and its applications, 155–160 (2014)
2. Okada, D.R., Blankstein, R.: Digital Image Processing for Medical Applications (2009)
3. Lee, J.G., et al.: Deep learning in medical imaging: general overview. Korean J. Radiol. **18**, 570–584 (2017)
4. Maier, A., Syben, C., Lasser, T., Riess, C.: A gentle introduction to deep learning in medical image processing. Zeitschrift fur Medizinische Physik **29**, 86–101 (2019)
5. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: Digital Image Processing Using Matlab (2004)
6. Uchida, S.: Image processing and recognition for biological images, 523–549 (2012)
7. Coelho, L.P., Shariff, A., Murphy, R.F.: Nuclear segmentation in microscope cell images: a hand-segmented dataset and comparison of algorithms. In: Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009 (2009)

8. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM Comput. Surv. **31**(3), 264–323 (1999)
9. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. Int. J. Comput. Vis. **1**, 321–331 (1988)
10. Herbin, M., Venot, A., Devaux, J.Y., Piette, C.: Color quantitation through image processing in dermatology. IEEE Trans. Med. Imaging **9**, 262–269 (1990)
11. Ganster, H., Pinz, A., Röhrer, R., Wildling, E., Binder, M., Kittler, H.: Automated melanoma recognition. IEEE Trans. Med. Imaging **20**, 233–239 (2001)
12. Yuan, Y., Chao, M., Lo, Y.C.: Automatic skin lesion segmentation using deep fully convolutional networks with Jaccard distance. IEEE Trans. Med. Imaging **36**, 1876–1886 (2017)
13. Li, Y., Shen, L.: Skin lesion analysis towards melanoma detection using deep learning network. Sensors (Switzerland) **18**, 556 (2018)
14. Yuan, Y., et al.: Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma. Comput. Methods Programs Biomed. **0062**(1), 1–11 (2017)

Human Age Classification System Using K-NN Classifier

Vaishnawi Priyadarshni[1](✉), Anand Nayyar[2], Arun Solanki[1], and Archana Anuragi[3]

[1] Department of Computer Science and Engineering,
Gautam Buddha University, Greater Noida, India
Vaishnawipriyadarshni02@gmail.com,
Ymca.arun@gmail.com

[2] Graduate School, Duy Tan University, Da Nang, Viet Nam
anandnayyar@duytan.edu.vn

[3] Department of Computer Science, Government Girls Polytechnic,
Charkhari, Mahoba, India
archana_cs@rediffmail.com

Abstract. The age classification system is used to categorize age into various groups determined by facial features. Many researchers have used different techniques for the age classification system, namely Geometric Ratio, Wrinkle Ratio, Local Gabor Pattern, Histograms of Oriented Gradients, Adaptive Neuro-Fuzzy Inference system, Texture feature, Support Vector Machine and Artificial neural network. The current system does not give great precision in light of the fact that these strategies are wasteful for classification. The proposed work represents the classification of images depending on the groups. This work classifies the age into old, adult and child groups. This work completes in three steps, i.e., preprocessing, facial aging feature extraction and classification. In preprocessing, the RGB image is converted into a grayscale level image; the facial aging feature extraction is performed by skin texture feature extraction and wrinkle analysis. LGBPH and the density of wrinkles accomplished the wrinkle analysis. LGBPH is a histogram of LGBP which is calculated by LBP. The LGBP feature is not dynamic for features like rotation, alignment, and the illumination of the image. The proposed work selected three different skins for skin texture feature and five different skin textures for wrinkle investigation. This work uses K-NN classifier due to its fast processing; classifies the testing dataset from training dataset. The LGBPH provides better results and estimates the accuracy as compared to the previous technique used by the researcher. This classification is a triumphant work as it gives an efficient classification using facial aging feature extraction.

Keywords: Age classification system · Facial aging feature extraction · K- Nearest Neighbor (K-NN) · Local Gabor Binary Pattern Histogram (LGBPH) · Wrinkle analysis

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 294–311, 2019.
https://doi.org/10.1007/978-981-15-0108-1_28

1 Introduction

Age is important as well as basic information about humans, which determines the age of a person. Human age is affected by internal and external factors. A person may live in different places having different environments. Therefore, environmental effects on human skin are noticeable. The skin texture of human varies from person to person. Numerous variables are influencing the age, for example, eating schedules, dietary patterns, a way of life, use of cosmetics and medical problems. This study proposed a structured age estimation framework. Presently, many researchers used distinctive databases to different age groups. The age determination is not easy by the human as well as by the machines. Age information can be beneficial for applications such as in the medical diagnosis, computer security; biometrics and video surveillance which are time-consuming and find difficulty in gaining better results in limited time. Many researchers give their contribution to the human age classification system. In computer vision, the classification of age system is a growing area in the research of the past decade [2, 11, 12, 14].

2 Related Work

In the past, researchers have used different techniques and classifiers for feature extraction and classification of images. Kwon and Lobo [15] were the first investigators to work on the age classification system. Researchers have dealt with the problem of the small database which is insufficient for classification. Horng et al. [8] dealt with the problem of a relatively small database. However, the researcher faced the following three problems. First, estimation of age is not accurate for the system. Second, feature extractions depend on the size of the image which varies and change in the size of the image reduces the performance; and last, the unmanaged age groups. Kalamani et al. [13] used fuzzy lattice neural model for the age-related classification system. The researcher used various features that include wrinkle density, average skin variance, and wrinkle depth. The work of Kwon and Lobo [15] was the first in the area of the age-related classification problem. The researcher classified images into various age groups, namely adults, elderly and babies. This approach assessed on small sample size. It computes the mouth, virtual top of the head, eyes, chin, and nose as primary features. These feature ratios differentiate babies from the remaining age groups. The secondary feature was a wrinkle. These secondary feature ratios discriminate adults from the other categories. These classified categories had feature ratio and wrinkle analysis, and this was the first achievement to classify human age in different age groups. Horng et al. [8] are amongst the inventor to identify the age classification system. The inventor classified input images into four grayscale facial images age groups such as seniors, adults, young adults, babies, and middle-aged adults. Three steps follow the age-related classification procedure; the first step is localization, the second step is feature extraction, and the last step is age classification. The classification system uses neural networks. All categories are different from each other. The researcher used a private database. Kanno et al. [14] worked on age-related classification system. The authors used a private database to categorize 440 young male face images. The method used mosaic features and neural network. The age groups were showing 80% accuracy. Hayashi et al. [7] categorized human ages in a span of ten-year. The author used the histogram for equalization of the wrinkle and extraction of face wrinkles. Three hundred images database shows the accuracy of 27%. The age groups below 15 which have negligence changes, but have some recognizable changes in face done by analysis. Lanitis et al. [16] designed an age estimation algorithm. 0 to 35 years age estimation was taken. Above 35 years, the face did not provide better results. Iga et al. [9] used human and object interaction processing database which categorized 101 images into five different age groups. The information of skin, color, support vector machine and Gabor wavelet used with an accuracy of 58.4%. Takimoto et al. [20] used Human and Object Interaction Processing as the database. The author took images of different gender such as as 139 females and 113 males. The images categorized into six different age groups with the help of Principal Component Analysis neural network and Gabor wavelet which gained 54.7% and 57.3% accuracy for female and male facial images respectively. Ueki et al. [23] used WIT database for age classification. The method used a two-phase technique named 2DLDA and LDA. For feature extraction and description of a projection to maximize the proportion within the class and LDA did the segregation of the class. The accuracy rate o achieved for different range such as five years ranges, age group of 46.3% accuracy, ten years range age group of 76.8% accuracy and 15-year range age group accuracy of 78.1%. Yang and Ai [24] used different types of databases in an age classification system uses face recognition technology. The study had three thousand five hundred forty images. Six hundred ninety-six images are taken for study from Pose, Illumination, and Expression databases. These were three age groups in the database. This method used by Local Binary Pattern histogram technique. The accuracy was 92.12% for face recognition technology database and 87.5% Pose, Illumination and Expression database. Günay and Nabiyev [5] took a private database of 350 facial images and face recognition technology database. The method used Local Binary Pattern and Nearest Neighbor (KNN) classifier for classifying images into six age groups with 80% accuracy. Gao and Ai [4] used a fuzzy classifier, and Gabor features for grouping 6386 images in 4 different groups. The authors obtained 91% accuracy. Dehshibi and Bastanfard [2] proposed an age-related classification for four various age groups. The researcher used Iranian face Database (IFDB). The significant enhancement in the age-related classification achieved 86.64% accuracy. The researcher used 498 images for the age classification system. Tonchev et al. [21] developed an age group estimation system based on the subspace projection algorithm and vector classifier. Hajizadeh and Ebrahimnezhad [6] used Histograms on Probabilistic Neural Network and Oriented Gradients for analyzing 377 facial from the Iranian face database. Authors achieved 87.02% accuracy in categorizing images into age groups. Liu and Liu [18] designed an approach for the age-related classification system. The authors categorized images into five different age-groups. Support Vector Machine classifier used during the last stage [15]. Nithyashri et al. [19] used the Adaptive Resonance Theory Network (ART) method for classification. The author used the FG-NET database to categorize images into many different age groups such as senior adult, adult, young and child. Thukral et al. [22] used FG-NET database to classify images into several age groups. Fard et al. [3] classified 575 facial images into a different age group from the Productive Aging Lab face database.

Authors used Histogram of Local Binary Pattern, Oriented Gradients, and Adaptive Neuro-Fuzzy Inference methods and achieved 88.01% accuracy. Izadpanahi et al. [10] proposed a system of age classification for Iranian Face Database, Face and Gesture Recognition Research Network database and classified the face image into seven age groups. The feature used geometric ratios and wrinkle analysis. The classifier used Support Vector Classifier. The authors obtained accuracy at 92.62%. Lee et al. [17] classify images into different age groups based on local age group modeling. Kalansuriya et al. [12] classify facial images corresponding to gender and age. Images used by the author in the range of 8–63, 14–25, 26–45, 46–60 provide the gender classification of 100% and accuracy of age classification 90%, 50%, 40%, 90%. Jagtap and Kokare [11] classify input images into grayscale facial images age groups. The researchers used PAL face databases. The proposed system provided better accuracy of 93.75% for the age classification system.

3 Proposed Approach

Following points comprise the proposed age classification system.

3.1 Preprocessing

It is the first process of image processing considered as a general process. The objective of image preprocessing is to enhance and remove unwanted distortion, edges, some unnecessary information and remove changeable color quality. It provides the effect on the contrast of an image.

3.2 Facial Based Aging Feature Extraction

The facial aging feature extraction performed using skin texture feature extraction and wrinkle analysis. Skin texture features accomplished by local Gabor Binary Pattern Histogram (LGBPH) technique, and Wrinkle performs analysis of wrinkle features Extraction using LGBPH and wrinkle density.

3.2.1 Skin Textural Feature Extraction

The skin textural feature is extracted using LGBPH. LGBPH techniques are as follows.

3.2.1.1 Local Gabor Binary Pattern Histogram

LGBPH is a technique used for recognizing the face, which is first invented by Zhang et al. [25]. LGBPH is attained through applying Gabor filters of various orientations and frequencies, then using the LBP are applied over processing image and histograms are calculated. LGBP is a combination of Gabor feature and LBP feature.

- *Gabor Feature*

A collection of Gabor filter applied in image different frequencies and different orientation which extract feature, these features are called Gabor feature. A 2-D kernel is represented as follows:

Complex

$$g(x,y;\lambda,\theta,\varphi,\sigma,\gamma)=exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)exp\left(i\left(2\pi\frac{x'}{\lambda}+\varphi\right)\right) \tag{1}$$

Real

$$g(x,y;\lambda,\theta,\varphi,\sigma,\gamma)=exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)cos\left(2\pi\frac{x'}{\lambda}+\varphi\right) \tag{2}$$

Imaginary

$$g(x,y;\lambda,\theta,\varphi,\sigma,\gamma)=exp\left(-\frac{x'^2+\gamma^2 y'^2}{2\sigma^2}\right)sin\left(2\pi\frac{x'}{\lambda}+\varphi\right) \tag{3}$$

Where,

$$x' = xcos\theta + ysin\theta \tag{4}$$

$$\sin\theta(1+x)^n = 1+\frac{nx}{1!}+\frac{n(n-1)x^2}{2!}+\ldots \tag{5}$$

$$y' = -xsin\theta + ycos\theta \tag{6}$$

In this equation λ represent the wavelength of the wave vector, θ describes the orientation of Gabor kernel φ the phase offset, σ is standard deviation of the Gaussian envelope and γ describes the spatial aspect ratio and specifies the ellipticity of the support of the Gabor function. Gabor filter bank of 40 filter made of five different scales, 0, 1, ……, 4, and eight orientation 0, 1…….. 8.

Gabor filter with convolution in image performed to extract the feature as the following equation

$$G(z) = \varphi_{\mu,v}(z) * J \tag{7}$$

Where,

G(z) = result of convolution
J(z) = input image
z = (x, y).

- *Local Binary Pattern*

LBP is a very easy and efficient method that related to texture. Because of computational simplicity and differential power, LBP technique is a very famous technique in many application areas. LBP has the ability to solve a computational problem that makes easy to examine images in real-world challenges. The LBP calls to non- uniform patterns that might be used to implement rotation-invariant descriptor and overcome the length of the feature vector. A pattern is said to be a uniform local binary pattern if the pattern takes no of transitions from 0 to 1 in bitwise or vice versa. If total no of the pattern is 256, then 58 patterns are used uniform pattern and 1 non-uniform pattern. The LBP feature vector created in the following way.

a. Partition the examined window into cells that each cell is having 16 × 16 pixels.
b. Consider the pixels along a circle, in which each pixel within a cell compares with its all pixel to 8 neighbors.
c. Value of the center pixel is more than the value of the neighbor pixel value. It means center pixel "1" neighbor's value is 0 otherwise, "1".

LBP are calculated by the following equation:

$$\mathrm{LBP}_{\mathrm{P,R}} = \sum\nolimits_{j=0}^{P-1} s\left(g_p - g_c\right)2^p \tag{8}$$

where,

$$s(k) = \begin{Bmatrix} 1 & k \geq 0 \\ 0 & k < 0 \end{Bmatrix}$$

where,

g_p = pth pixel value
g_c = center pixel value
P = no of neighboring pixel
R = distance of neighboring pixels from center pixel.

3.2.2 Wrinkle Analysis

The wrinkle is a primary feature of the face used in the wrinkle analysis. As the age grows, the wrinkles are determined to face easily. A wrinkle shows that growing age progression. Wrinkle grows, according to the age progression. Wrinkles are inescapable information for a human being. The wrinkle analysis is calculated using LGBPH as a discussion in Sect. 3.2.1 and wrinkle density. Wrinkle helps in the age classification system. Wrinkle density is determined by

$$V_d = \frac{\boldsymbol{Wn}}{\boldsymbol{Tn}} \tag{9}$$

where,

V_d = wrinkle density
W_n = count the wrinkle pixel
T_n = total no of pixel in image.

3.3 Classification

K-Nearest Neighbor is very easy and straightforward to understand. It is a non - parametric method that applies to both classification and regression. In classification, an object is categorized from a majority vote of about their neighbors and using the object is allocated to class between k nearest neighbors (k is a small positive integer). In regression, the output is used for an object that is the average of their k-NN value [24, 26]. The feature vector space is part of the training and training sample of class labels. K-NN is to determine the numbers of k samples. The training has identified K samples. Categorization of samples that reused as test samples into a class. K-NN is used to categorize different age groups. It can be more efficient for large training data sets. The Euclidean distance D determined by

$$D = \sqrt{\sum\nolimits_{i=1}^{N} (x_i - y_i)^2} \tag{10}$$

where, x_i = element of X, y_i = element of Y.

4 Process Flow and Working of System

The process flow of age classification system s divided into three parts, i.e. preprocessing, skin aging feature extraction and classification. The Fig. 1 represent the process flow diagram of the age group classification system. The first process is preprocessing in which RGB image converted into the Grayscale image. The second process is facial aging feature extraction in which skin texture feature extracted using LGBPH and wrinkles are analyzed with the help of LGBPH and the density of wrinkle. The third process is a classification of images in the relevant age groups with the help of K-NN classifier.

4.1 Image Preprocessing

Image preprocessing is a critical and essential process in the image processing. It is an initial and primary step in any process of image processing. The objective of image preprocessing is to enhance and remove unwanted distortion, edges, some unnecessary information, and changeable color quality. Image preprocessing provides the effect on the contrast of an image. It gives a significant impact on image analysis and results. It converted RGB image into a grayscale level, crop facial area on gray scale level image and normalized as shown in Figs. 1, 2 and 3.

4.2 Facial Aging Feature Extraction

The second process is the facial aging feature extraction process for skin texture extraction. After that analysis of wrinkle is accomplished. A face has eight regions; used for facial aging feature extraction. There are three face regions; used for skin textural feature extraction. There are five face regions; used for wrinkle analysis. These are the particular areas of the face, which show identifiable age changes in Fig. 4.

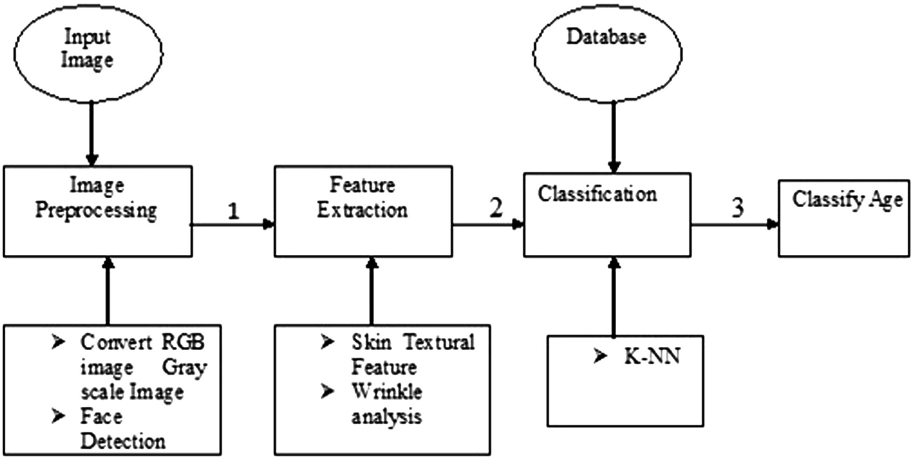


Fig. 1. Process flow of age group classification system

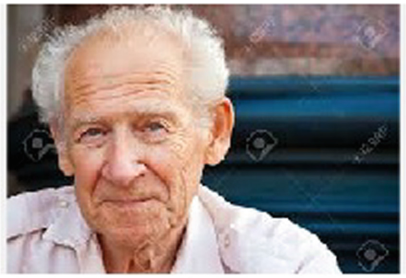


Fig. 2. RGB image

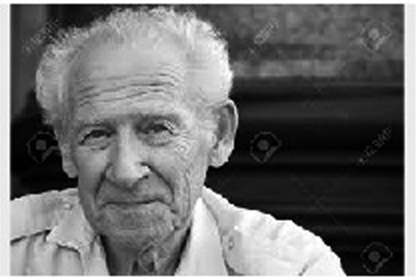


Fig. 3. Gray scale level image

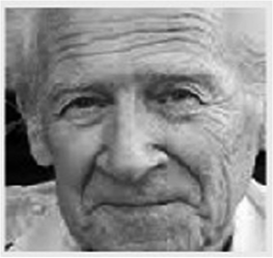


Fig. 4. Cropped face area

4.2.1 Skin Textural Feature Extraction

For Skin textural features are extracted by LGBPH using the following steps.

Step (a): In the first step of skin texture feature extraction, proposed work builds a bank of 40 filters which used Gabor filter as shown in Fig. 5.

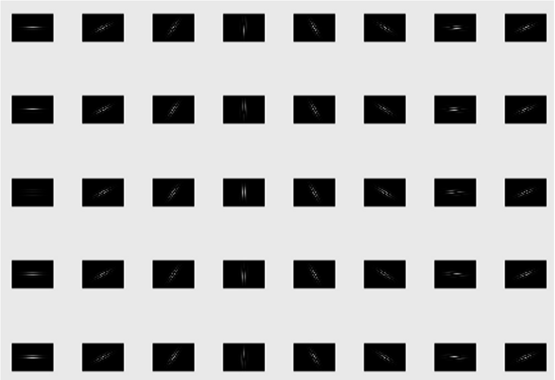


Fig. 5. Gabor filter bank used 40 filters with 5 scales and 8 orientations

Step (b): The used region is cropped from the grey level facial image.
Step (c): Calculate the convolution of the cropped region (Eq. 7) as shown in Fig. 6.

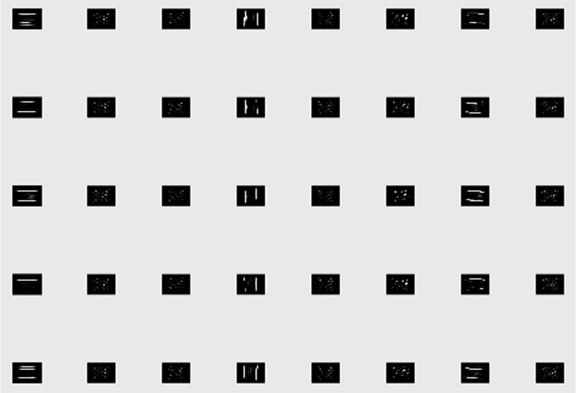


Fig. 6. Convolution of crop region with 40 Gabor filter bank

Step (d): This cropped region proposed system used Gabor filter bank with convolution to determine the Gabor feature. Local binary pattern technique is used along Gabor filter to calculate local gabor binary pattern as shown in Fig. 7.

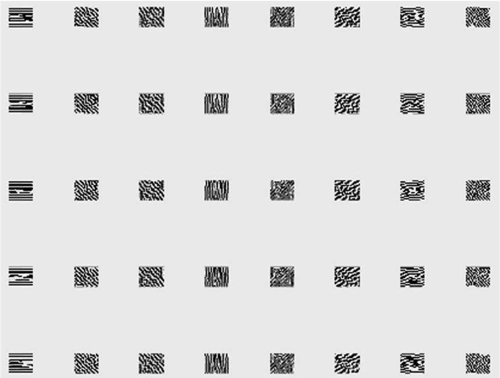


Fig. 7. LBP over Gabor features to create LGBP.

Step (e): Evaluate the histogram of LGBP which is referred LGBPH as shown in Fig. 8.

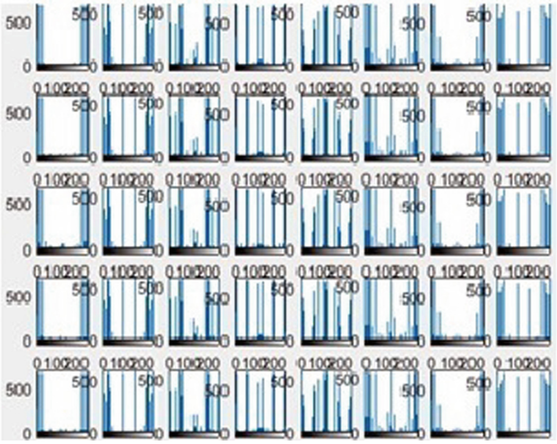


Fig. 8. Histogram of LGBP

4.2.2 Wrinkle Analysis

The wrinkle analysis evaluates wrinkle density and LGBPH.

(A) The following steps calculate the LGBPH for wrinkle analysis as shown in Fig. 12.

Step (a): Build a bank of 40 filters which use Gabor filter.
Step (b): The cropped region is used from grey level face images.
Step (c): Calculate convolution of crop image through 40 filter bank for extracting Gabor feature.

Step (d): Determine the magnitude of every Gabor features.
Step (e): This step Take the first 12 features of Gabor and applying LBP for evaluation of LGBP.
Step (f): Evaluate the histogram of LGBP which referred LGBPH.

(B) The following steps is performed for the detection of wrinkle.

Step (a): Firstly cropped the region of interest used for wrinkle analysis and used a Gaussian filter to remove noise from an image as shown in Fig. 9.
Step (b): This step performed the canny operator to edge detection of an image and performed the morphological operation for reducing unnecessary edges as shown in Figs. 10 and 11.

LGPBH is used differently for wrinkle analysis and skin textural feature extraction. This study applies the LBP on 40 Gabor features for skin texture feature extraction Whereas 12 Gabor features applied to wrinkle analysis with a large amount of amplification and highest correlation using input image.

Fig. 9. Forehead region

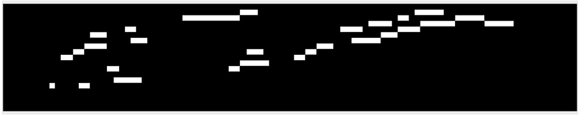


Fig. 10. Edge detection by canny operator

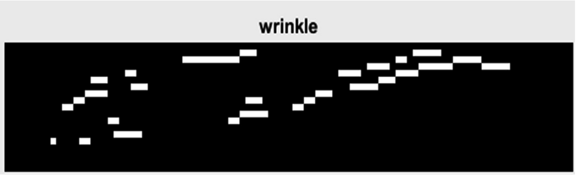


Fig. 11. Wrinkle detection

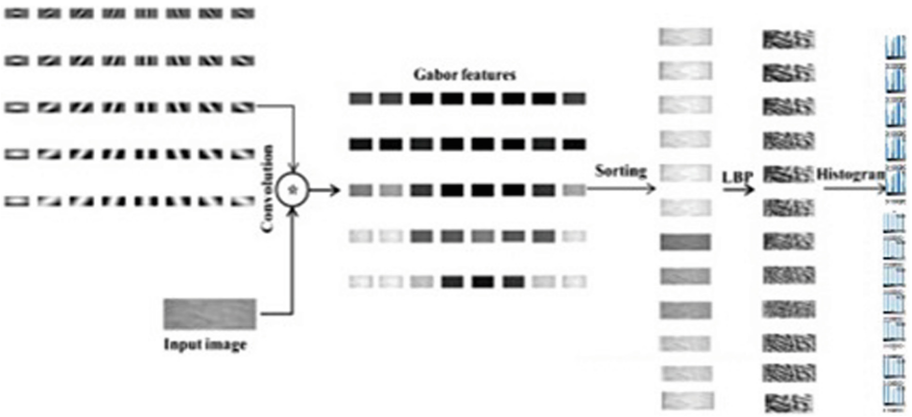


Fig. 12. LGBPH evaluation analysis of wrinkle

Gabor filter's response shows the highest amount of correlation using the input image to match the orientation of the face wrinkle to the orientation of its answers. There are following steps which used to extract facial aging feature extraction.

i. Crop the three regions of skin texture namely nose, left cheek below eyes, right cheek below eyes, according to dimension mention as [5, 5, 0.33, 0.4], [5, 5, 0.33, 0.3], [5, 5, 0.33, 0.3] pixels as shown in Fig. 13.
ii. Calculate the LGBPH of three skin textural region of the face image.
iii. Concatenate LGBPH of three skin textural regions of face area which stored in the following vector

$$V = [v_1, v_2, v_3] \tag{11}$$

Where
V = skin textural feature matrix

iv. Crop the five regions of skin texture namely left eye's corner, right eye's corner, forehead, left below eyes, right below eyes, according to dimension mention as [5, 4, 0.11, 0.25], [10, 4, 0.11, 0.25], [3, 3, 1, 0.25], [10, 7, 0.33, 0.166], [10, 7, 0.33, 0.166] pixels as shown in Fig. 14.
v. Determine the wrinkle analysis by calculating LGBPH and wrinkle density for each particular five regions.
vi. Concatenate LGBPH of five skin textural regions of face area which stored in the following vector

$$Wa = [w_1, w_2, w_3, w_4, w_5] \tag{12}$$

Where
Wa = LGBPH of wrinkle analysis feature matrix

vii. Concatenate same for wrinkle density and stored in the following vector

$$V_d = [wd_1, wd_2, wd_3, wd_4, wd_5] \tag{13}$$

Where
V_d = Wrinkle density feature vector

viii. Combine the LGBPH of these three regions of the face area and LGBPH of the wrinkle analysis feature vector

$$Vc = [V, Wa] \tag{14}$$

ix. After the feature extraction, reduce dimension due to large matrix size. For dimension reduction, Principle Component Analysis (PCA) used as named Vp.

x. Before combining wrinkle density and LGBPH feature concatenation, normalization performed by the following equation

$$F(i) = \frac{F_{ki} - \min(F_k)}{\max(F_k) - \min(F_k)} \tag{15}$$

After normalization, both features are fusioned.

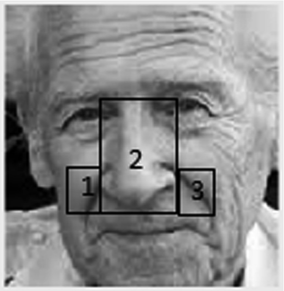


Fig. 13. Regions used for skin textural feature extraction

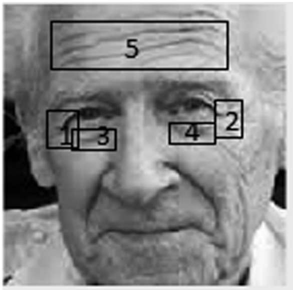


Fig. 14. Regions used for wrinkle analysis

4.3 Classification

The last process is classifications which are very important to categorize input face images into age groups. A training dataset comprised the feature extracted from the face and labeled to the data set pair; forms a model. The testing dataset obtained these feature in the same fashion. The extracted feature with its age group applied to classify which determine the age group. K-NN is an easy and fast way to implement, and its nearest k neighbors classify the subject. After that last feature classifies the input image into the relating age group. The model has been used these features to categorize the age groups. K-NN classifier categorizes the testing dataset images in which it belongs to one of the age groups such as young, adult and old [1].

5 Results

The proposed system used a private database which has 270 images having testing and training data set. Three regions of an image, namely nose, left cheek below eyes, right cheek below the eyes for skin texture feature extraction uses LGBPH. The skin texture LGBPH features stored in matrix named as V. For wrinkle analysis firstly LGBPH calculates the five regions namely left eye's corner, right eye's corner, forehead, left below eyes, right below the eyes is used. Then calculates wrinkle density. The LGBPH feature of wrinkle analysis stored in matrix named as Wa and the length of wrinkle density feature stored in a vector named as Vd. Skin texture LGBPH features stored in a matrix called as V and LGBPH feature of wrinkle analysis stored in matrix named as Wa. Both matrices are combined and stored in a matrix named as Vc. The size of both feature matrixes is large; therefore PCA applied for reducing a dimension of a matrix to store in a vector named as Vp. The reduced vector is normalized as the feature lies in range 0 to 1, used in the classification. The whole mechanism used in both training and testing stages. The proposed system takes 180 images to use for training datasets, and 90 images choose to apply for testing datasets. The last feature classifies the input image to the relevant age group. The model used these features to categorize into different age groups. K-NN classifier categorized the testing dataset images into groups young, adult and old.

- *Recognition Results*

As shown in Table 1, there are three groups, namely child, adult and old. The total numbers of the testing dataset have 90 images having 30 images of each group, 28 images from child group matched accurately. The recognition rate of the child group obtained is 93.3% Twenty-nine images are matched accurately from the adult group. The recognition rate of the adult group obtained is 96.6%. Thirty images are matched accurately from the adult group. The recognition rate of the old group obtained is 100%. The average recognition rate of proposed system which is obtained by K-NN classifier = (93.3 + 96.6 + 100)/3 = 96.6%. Figure 15 shows the accuracy of child, adult and old groups. The x-axis denotes the groups, and Y-axis denotes accuracy of different groups. The accuracy of child, adult and old represented in the graph is 93.3%, 96.6%, and 100% respectively.

Table 1. Recognition rate of input image

Groups	Recognition results			
	Child	Adult	Old	Recognition rate
Child	28	0	2	93.3%
Adult	0	29	1	96.6%
Old	0	0	30	100%

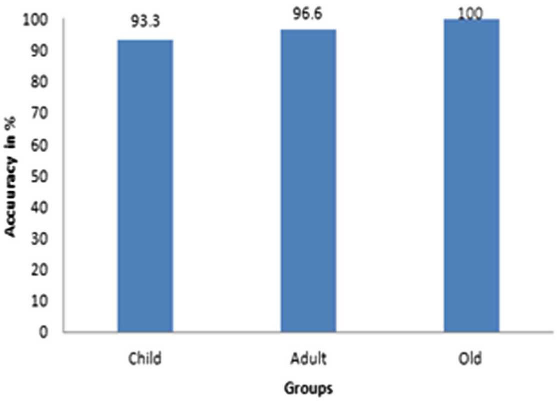


Fig. 15. Accuracy of groups

- *Comparison with Existing Techniques and Classifiers*

Table 2. Comparison of existing techniques and classifier

Author's name	Features	Tools of classification	Accuracy in %
Kwon and Lobo [15]	Geometric ratio and wrinkle analysis	Similarity measures	100
Kanno et al. [14]	Mosaic feature	ANN classifier	80
Horng et al. [8]	Geometric ratio and wrinkle analysis	ANN classifier	81
Iga et al. [9]	Gabor feature	SVM classifier	58.4
Gao and Ai [4]	LBP feature	FUZZY LDA classifier	91
Dehshibi and Bastanfard [2]	Geometric ratio and wrinkle analysis	ANN classifier	86.64
Hajizadeh and Ebrahimnezhad [6]	Histogram of Oriented Gradient	PNN classifier	86.64
Fard et al. [3]	LBP and HoG features	ANFIS classifier	88.01
Izapandhi and Toygar [10]	Geometric ratio and wrinkle analysis	SVC classifier	92.62
Kalansuriya and Dharmaratne [12]	Geometric ratio and parameter calculation	ANN classifier	90
Solanki and Priyadarshni (Proposed method)	Facial aging feature	K-NN classifier	96.6

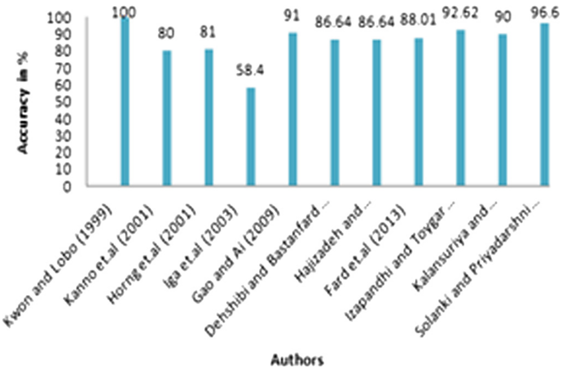


Fig. 16. Comparison of accuracy achieved by authors

Many researchers worked on the age classification system. The researchers used many techniques of classification. The work of Kwon and Lobo [15] was first in the area of age-related classification problem which used geometric ratio, wrinkle analysis feature and similarity measures as a tool of classification. The researcher classified images into various age groups namely adult, seniors and babies with 100% accuracy. Kannoet al. [14] used mosaic features and neural network for the age classification system. The accuracy obtained with four age groups with 80% accuracy. Horng et al. [8] classify input images into four grayscale facial images with ANN classifier having 81% accuracy. Iga et al. [9] used a Gabor filter for extracting the feature and SVM as a tool of classification. Images divide into five different age groups. The analysis was done [4] with an accuracy of 58.4%. Gao and Ai used a fuzzy classifier, and Gabor features for grouping images in 4 different groups. The accuracy obtained 91%. Hajizadeh and Ebrahimnezhad [6] used Histograms of Probabilistic Neural Network and Oriented Gradients for analysis with 86.64% accuracy [14]. Fard et al. classified facial images in a different age group. Izadpanahet al. [10] classify face image into seven age groups using Support Vector Classifier. Analysis obtained accuracy 92.62%. Kalansuriya and Dharmaratne classify facial images using geometric ratio, wrinkle analysis and ANN classifiers to corresponding to gender and age. Images used by the author in the range of 8–63, 14–25, 26–45, 46–60 which provide the gender classification of 100% and accuracy of age classification 90%, 50%, 40%, 90%. Figure 16 represents the comparison of accuracy by various authors. Proposed work provides better results and estimates the accuracy as compared to the previous technique used by the different researcher in Table 2.

6 Conclusion

Human face provides adequate information that is easily recognizable by a human. The human face also provides expression, moods, and state of action. As a human grows, noticeable change perceives in age progression. The age classification is also challenging for the machine to categorize face images in age groups. In this work, the age-group classification system has been classified age into three different groups which include a child, adult and old. A method is used to solve the age classification related problem. The proposed work is completed in three stages, namely preprocessing, facial aging feature extraction and classification. Preprocessing is an important and essential stage that RGB image converted into the grayscale image because it removes unwanted distortion, edges, some unnecessary information, and changeable color quality. Facial aging feature extracted by skin textural feature extraction with the help of LGBPH and Wrinkle analysis method. For skin, texture feature used three regions (nose, left cheek below eyes, right cheek below eyes) and for wrinkle analysis used five regions (left eye's corner, right eye's corner, forehead, left below eyes, right below eyes). LGBPH accomplishes skin texture feature extraction, and Wrinkle analysis is performed using LGBPH and wrinkle detection. LGPBH is handled differently for wrinkle analysis and skin textural feature extraction. LBP applied on 40 Gabor features for skin texture feature extraction whereas 12 Gabor features applied to wrinkle analysis with a large amount of amplification and highest correlation using the input image. Gabor filter's response shows the highest amount of correlation using the input image to match the orientation of the face wrinkle to the orientation of its answers. LGBPH is unvarying for rotation illumination and translation. The facial aging features which used the region of interest from face images to categorize into different age groups. K-NN classifier is a fast learning classifier, used in the proposed work. The proposed system classified into three groups namely child, adult and old which obtained accuracy 93.3%, 96.6% and 100% respectively. The overall accuracy of the proposed system is 96.6%. The proposed methods provide a better result and improve existing age classification system. LGBPH give a better result and estimate the accuracy as compared to the previous technique used by the researcher. This classification is successful work which provides an efficient classification using facial aging feature extraction. The Proposed system provides a better result and improves existing age classification system. In the future, for an age classification system used the most extensive database for improving the real-time system.

References

1. Alkhateeb, J.H., Khelifil, F., Jiani, J., Ipsonl, S.S.: A new approach for off handwritten Arabic word recognition using K-NN classifier. In: IEEE International Conference on Signal and Image Processing Application (2009)
2. Dehshibi, M.M., Bastanfard, A.: A new algorithm for age recognition from facial images. Sig. Process. **90**, 2431–2444 (2010)
3. Fard, H.M., Khanmohammadi, S., Ghaemi, S., Samadi, F.: Human age-group estimation based on ANFIS using the HOG and LBP features. Electr. Electron. Eng. **2**, 21–29 (2013)
4. Gao, F., Ai, H.: Face age classification on consumer images with gabor feature and fuzzy LDA method. In: Tistarelli, M., Nixon, M.S. (eds.) ICB 2009. LNCS, vol. 5558, pp. 132–141. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-01793-3_14
5. Günay, A., Nabiyev, V.: Automatic age classification with LBP. In: 23rd International Symposium on IEEE Computer and Information Science (2008)
6. Hajizadeh, A., Ebrahimnezhad, H.: Classification of age groups from facial image using histograms of oriented gradients. In: 2010 Proceedings of the 7th Iranian Conference on Machine Vision and Image Processing, pp. 1–5 (2010)

7. Hayashi, J., Yasumoto, M., Ito, H., Niwa, Y., Koshimizu, H.: Age and gender estimation from facial image processing. In: Proceedings of the 41st SICE Annual Conference, pp. 13–18 (2002)
8. Horng, W.B., Lee, C.P., Chen, C.W.: Classification of age groups based on facial features. Tam kang J. Sci. Eng. **4**, 183–192 (2001)
9. Iga, R., Izumi, K., Hayashi, H., Fukano, G., Ohtani, T.: A gender and age estimation system from face images. In: Proceedings of the SICE Annual Conference, pp. 756–761 (2003)
10. Izadpanahi, S., Toygar, O.: Human age classification with optimal geometric ratios and wrinkle analysis. Int. J. Pattern Artif. Intell. (IJPRAI) **28**, 1–17 (2014)
11. Jagtap, J., Kokare, M.: Human age classification using facial aging features and artificial neural network. Cogn. Syst. Res. **40**, 116–128 (2016)
12. Kalansuriya, T.R., Dharmaratne, A.T.: Neural network based age and gender classification for facial images Int. J. Adv. ICT Emerg. Reg. **7** (2014)
13. Kalamani, D., Balasubtamani, P.: Age classification using fuzzy lattice neural network. In: Proceedings of Sixth International Conference on Intelligent Systems Design and Application (ISDA 2006), vol. 3, pp. 225–230 (2006)
14. Kanno, T., Akiba, M., Teramachi, Y., Nagahashi, H., Agui, T.: Classification of age group based on facial images of young males by using neural networks. IEICE Trans. Inf. Syst. **84**, 1090–1104 (2001)
15. Kwon, Y.W., Lobo, N.V.: Age classification from facial images. Comput. Vis. Image Underst. J. **74**, 1–21 (1999)
16. Lanitis, A.: On the significance of different facial parts for automatic age estimation. In: 14th International Conference on Digital Signal Processing, vol. 2, pp. 1027–1030 (2002)
17. Lee, S.H., Ro, Y.M.: Local age group modeling in unconstrained face images for facial age classification. In: IEEE International Conference on Image Processing (2014)
18. Liu, L., Liu, J., Cheng, J.: Age-group classification of facial images. In: 11th International Conference on IEEE Machine Learning and Applications (ICMLA) (2012)
19. Nithyashri, J., Kulanthaivel, G.: Classification of human age based on neural network using FG-NET aging database and wavelets. In: Fourth International Conference on IEEE Advanced Computing (ICoAC) (2012)
20. Takimoto, H., Mitsukura, Y., Fukumi, M., Akamatsu, N.: A design of gender and age estimation system based on facial knowledge. In: Proceedings of the SICE-ICASE International Joint Conference, pp. 3883–3886 (2006)
21. Tonchev, K., Paliy, I., Boumbarov, O.: Human age-group classification of facial images with subspace projection and support vector machines. In: IEEE 6th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), vol. 1 (2011)
22. Thukral, P., Mitra, K., Chellappa, R.: A hierarchical approach for human age estimation, acoustics. In: IEEE International Conference on Speech and Signal Processing (ICASSP) (2012)
23. Ueki, K., Hayashida, T., Kobayashi, T.: Subspace-based age-group classification using facial images under various lighting conditions. In: 7th International Conference on IEEE Automatic Face and Gesture Recognition (2006)
24. Yang, Z., Ai, H.: Demographic classification with local binary patterns. In: Lee, S.-W., Li, S. Z. (eds.) ICB 2007. LNCS, vol. 4642, pp. 464–473. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-74549-5_49
25. Zhang, W., Shan, S., Gao, W., Chen, X.: A novel non statistical model for face representation and recognition. In: IEEE International Conference on Computer Vision, pp. 786–791 (2005)
26. Alzubi, J., Nayyar, A., Kumar, A.: Machine learning from theory to algorithms: an overview. J. Phys: Conf. Ser. **1142**(1), 012012 (2018)

Bilingual Machine Translation System Between Hindi and Sanskrit Languages

Neha Bhadwal[1], Prateek Agrawal[1,2](✉), and Vishu Madaan[1]

[1] Lovely Professional University, Phagwara, Punjab, India
bhadwalneha21@gmail.com, prateek061186@gmail.com, vishumadaan123@gmail.com
[2] University of Klagenfurt, Klagenfurt, Austria

Abstract. Machine translation (MT) is increasingly becoming popular today as it reduces a great deal of efforts when it comes to the task of translation from one language to other. Not only is it time saving, but secure and cost-effective in the long run. A lot of work has been done in the field of MT lately. Various machine translation systems (MTS) have been developed for numerous languages. Sanskrit is one of the oldest languages in the world and is still in existence. There is loads of literary, scientific and mathematical work written in Sanskrit during ancient times. Despite of its important part in Indian culture and history, not much work has been done for translation to or from Sanskrit language. This proposed work is an effort to bridge the language gap between Sanskrit and Hindi, one of the most widely spoken languages of the world. The idea is to develop an MTS which will serve as a tool to perform two-way translation between Hindi and Sanskrit language. The system once developed can be used for learning and understanding the grammar and structure of both the involved languages.

Keywords: Machine translation system · Bilingual translation · Bilateral MT · Natural language processing · Hindi translation · Sanskrit translation

1 Introduction

MT comes under the area of natural language processing (NLP). NLP is the field which studies the human-machine interaction. MT is the use of computer systems to perform the translation task, instead of deploying human experts A machine translation system (MTS) is basically an application software that deals with the translation of languages, whether given in oral or written form. The MTS can work for two languages or it can also involve multiple languages depending upon the requirements. Designing of an MTS depends upon the syntax and semantic knowledge of the languages involved. This requires a deep and thorough understanding of the divergence and grammatical similarities of the concerned languages. In this era of globalization, when the world is coming closer and there is burst of information, MT can prove to be a powerful tool

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 312–321, 2019.
https://doi.org/10.1007/978-981-15-0108-1_29

to aid this phenomenon by providing people the information in the language that they understand. It involves being able to translate web pages and other available online information to the languages of interest. For an MTS, the efforts in developing the system and accuracy of the system are major concerns. These are the basis to differentiate among approaches for machine translation. The current work makes use of rule-based approach to perform two-way machine translation using a common bilingual corpus. Rule-based approach makes use of the knowledge of grammatical features of the two languages as it is the basis of designing rules for translation. Table 1 presents a comparison of four major MTS approaches, namely, rule-based, statistical, example-based and neural machine translation. Section 1.1 provides a brief review of related works in the MTS field. Section 2 gives a detailed view of the proposed methodology. Section 3 contains the conclusions and future scope of the proposed work.

Table 1. Comparison of various MT technique

	RBMT	SMT	EBMT	NMT
Basis	morphological, semantic and syntactic analysis	statistical models	translation by analogy	machine learning paradigm
Requirements	manually designed set of rules	correlations between source text and translations	existing translation pairs of source and target	neural networks consisting of nodes
Pros	no need of large bilingual corpus	no need of manually creating linguistic rules	high translation quality when examples are similar	dynamic and complex nature allows context-based translation
Cons	time-consuming, labor-intensive unable to cover each conflict or linguistic phenomenon	need of large and organized bilingual corpus	translation quality can be very low if no similar example is found	continuous learning requires high processing capability
Training	rules define the mapping between source and target languages	engine needs to be trained with the text similar to the source text	limitation of translated examples which are retrieved after matching the source	a web of complex relationships between nodes that may represent a word, phrase or sentence

1.1 Related Work

Singh et al. [1] propose an MTS based on genetic algorithm (GA) from Sanskrit to Hindi language. The system works in two phases where the initial phase deals with tokenization and analysis of the input sentence. The next phase is the generator phase where genetic algorithm is applied to the input. The categorization of inputs is done based on their size and complexity. An analysis of different MTS is presented by Raulji and Saini [2]. They try to compare various MTS which involve Sanskrit in any way, whether a source, target or a support language. They further discuss the morphological, semantic and syntactic features of Sanskrit language and how they can be beneficial to carry out translation.

Bahadur et al. [3] discuss the linguistic structure of Sanskrit followed by a comparison between Sanskrit grammar and context-free grammar (CFG). Subsequently, a system for English to Sanskrit language is proposed after a brief discussion of the grammatical divergence between the two languages. Saini and Sahula [4] present a review of various MTS developed in India and different approaches used for machine translation (MT). A comparative analysis is performed on various MTS meant for translation among Indian languages. Rathod [5] proposes a model for an MTS which translates from English to Sanskrit language using rule-based as well as example based approach. The input sentences are categorized on the basis of their length as small, large and very large. The input is tokenized before translation.

Mane et al. [6] develop a system for English to Sanskrit translation using rule based approach. The translation process is aided by a bilingual dictionary which contains morphological information of both the involved languages. The input after being tokenized and tagged is converted to an ordered structure, such as, a parse or syntax tree. Mishra and Mishra [7] present a study of example-based machine translation (EBMT). Different approaches to EBMT are discussed and compared. A detailed comparison of English and Sanskrit language is performed. The challenges in English to Sanskrit translation are discussed followed by translation divergence between the two languages (Table 2).

Shahnawaz [8] present a comparison between Hindi and Urdu languages. He discusses the problems arising while performing the transliteration from Hindi to Urdu. Also, it is concluded that direct MT approach is the most suitable for given language pair.

Mishra and Mishra [9] present a rule-based approach for MT from English to Sanskrit language. They make use of the markings formed by using morphological properties of Sanskrit language for parts of speech (POS) tagging. Linguistic features of Sanskrit are discussed followed by a brief comparison of English and Sanskrit language. The system is evaluated for various classes of inputs using standard evaluation schemes, such as, BLEU, Meteor, F-measure, Unigram precision. The input is classified based on the complexity of the sentences. Gupta et al. [10] present an algorithm for MT from Sanskrit to English language using rule-based approach. The knowledge representation of the translation process is also presented by firstly generating lexemes from the input and then comparing and replacing them with the corresponding translation in the output.

Table 2. Comparison of Hindi and Sanskrit grammars

Basis of difference	Hindi	Sanskrit
Alphabets	11 vowels, 35 consonants	13 vowels, 33 consonants
Numbers	Singular, Plural	Singular, Dual, Plural
Genders	Feminine, Masculine	Feminine, Masculine, Neuter
Persons	First, Second, Third	First, Second, Third
Dependence of verb	Person, Number, Gender	Person, Number
Vowel Types	Short, Long	Short, Long, Elongated
Inflecting	No	Yes
Word order	Important (SOV)	Not important
Tenses	Past, Present, Future	Past Aorist, Perfect, Past Imperfect, Present, First future, Second future
Moods	Indicative, Presumptive, Subjunctive	Indicative, Potential and Imperative

Mane and Hirve [11] perform a survey of various MTS developed in India. Different approaches to MT are discussed followed by a tabular comparison between four MTS based on features, approaches, advantages and limitations. Shukla et al. [12] propose that there are seven types of translation divergence between Hindi and Sanskrit. Five of them arise due to Sanskrit grammar and rest two because of Hindi language. Goyal and Sinha [13] study translation divergence between English-Sanskrit and Hindi-Sanskrit language pairs. Dorr's classification for divergence has been taken as the basis for this study. Gehlot et al. [14] discuss a transfer-based approach for machine translation of documents from Hindi to English language. The advantages of using transfer rules for MT are discussed over corpus-based MTS. Mall and Jaiswal [15] present a rule-based approach for MTS from Hindi to English language. Various existing MTS are studied along with their features and limitations. Rule-based approach eliminates the requirement of using any intermediate representations.

Agrawal [16] proposes and designs a rule-based MTS from Sanskrit to Hindi in his Ph.D. thesis. Jain and Agrawal [17] propose a method to perform transliteration from English to Sanskrit language text. The idea is to do direct mapping of phonetically similar set of characters from two languages. They further proposed another method to identify and extract root words from source Hindi sentences [18]. The process of identification is independent of the context. Sethi et al. [19] design an algorithm to paraphrase input Hindi sentences without losing their actual context. This holds great importance in the field of NLP. Sugandhi et al. [20] present a parser model which suggests possible translation for an English sentence based on the syntactical information gathered such as, case, gender and number. Ambiguity may arise due to difference in grammatical structures of languages.

Tapaswi and Jain [21] perform morphological and lexical analysis of Sanskrit sentences. It involves extracting root words after separating the markers and deriving its meaning. This model is applicable to morphologically rich languages similar to Sanskrit. A comparative study of various MTS and their techniques is performed by Ashraf and Ahmad [22]. They are compared based on their ease of use, complexity and efficiency. A comparative study has been done by Chand [23] of paragraph translations performed by various online tools. They are categorized as rule-based and statistical systems. Bharati et al. [24] give a detailed introduction to natural language processing in this book. This is done based on the Ashtadhyayi grammar of Sanskrit language. This book has been a source of origin for much of the work done in the field of modern linguistics.

1.2 Comparison of Hindi and Sanskrit

Sanskrit is one among the 22 official languages of India. The word Sanskrit itself means redefined or created with perfection. It belongs to Indo-European family of languages. It is believed to be highly systematic and technical language. A large number of works of drama, poetry and literature have been written in Sanskrit. Much of the work in philosophy, science, mathematics, astronomy and logic can be found in Sanskrit. Briggs [25] emphasized the importance of Sanskrit language. He explained how its structure was similar to semantic nets and hence suitable for use in the field of artificial intelligence. On the other hand, Hindi is the fourth most widely spoken language in the world along being the national language of India. Table 1 represents a brief comparison of grammatical features of Hindi and Sanskrit languages (Fig. 1).

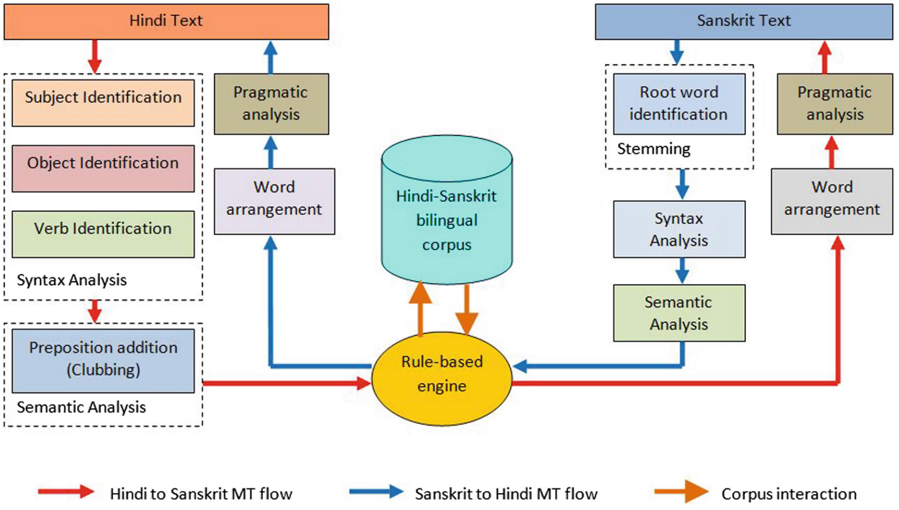


Fig. 1. Architecture of the proposed Hindi-Sanskrit bilateral MTS

2 Proposed Methodology

The proposed system is a bilateral MTS for Hindi and Sanskrit based on a rule-based approach. A common bilingual (Hindi-Sanskrit) dictionary is designed which contains grammatical information about words and phrases. The system can operate in either direction based on the language provided in the input. The key feature is the use of a common database which consists of rules for governing the translation process. This process consists of mainly four important steps. They are discussed as follows.

1. Taking the input sentence and tokenization (Parts-of-speech tagging)
2. Performing syntax and semantic analysis.
3. Performing mapping of the tokens using the rule-based engine.
4. Arranging the words in order and performing pragmatic analysis to generate final output.

2.1 Hindi to Sanskrit MT Flow

Initially, the input sentence is parsed. The nouns (Subject and object) and verb are identified under syntax analysis. The prepositions are added after nouns according to the cases under semantic analysis. Mapping from Hindi to Sanskrit is performed using the bilingual corpora according to the rule-based engine. After arranging the words, pragmatic analysis is performed on the sentences to yield the final output. The flow of Hindi to Sanskrit translation process is depicted in Fig. 2(a).

2.2 Sanskrit to Hindi MT Flow

Firstly, the input sentence is parsed and stemming is performed. Stemming involves segregating root word from inflections (mostly suffixes). The nouns (subject and object) and verb are identified. Then, after performing the semantic analysis, mapping is performed from Sanskrit to Hindi using the bilingual corpus. At last, words are arranged followed by pragmatic analysis of sentences to form the final Hindi output. This flow from Sanskrit to Hindi language is shown in Fig. 2(b).

2.3 Rule-Based Engine

The rule-based engine contains the set of rules which govern the Hindi-Sanskrit translation in both ways. It directly deals with a bilingual corpus which contains Hindi-Sanskrit mapping plus morphological information of possible words. The source language words are given as input, mapping is performed from source to target language and finally, corresponding target language words are obtained as input.

2.4 GUI Design

A graphical user interface (GUI) for the proposed model has been developed in Java which can provide an easy-to-use environment plus detailed explanation of the MT process in both the directions. Figures 3, 4, 5, 6 and 7 depict the working

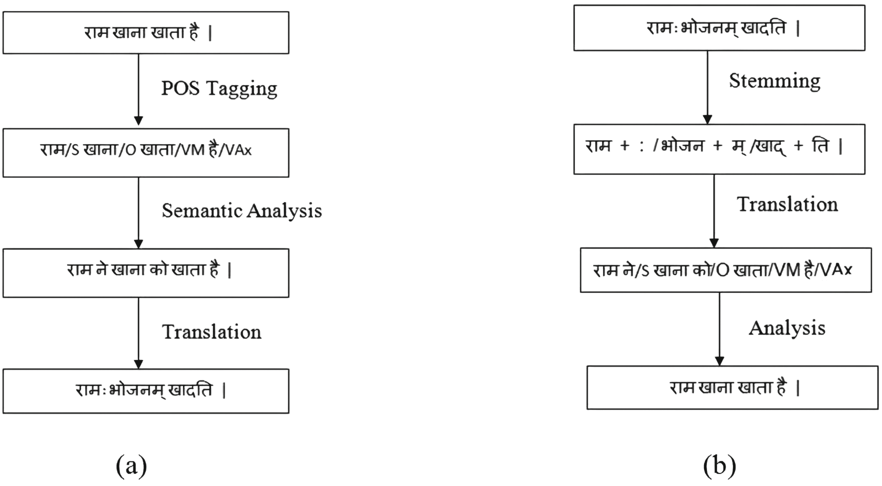


Fig. 2. Depiction of bilateral MT flow between Hindi and Sanskrit, (a) Hindi to Sanskrit and (b) Sanskrit to Hindi

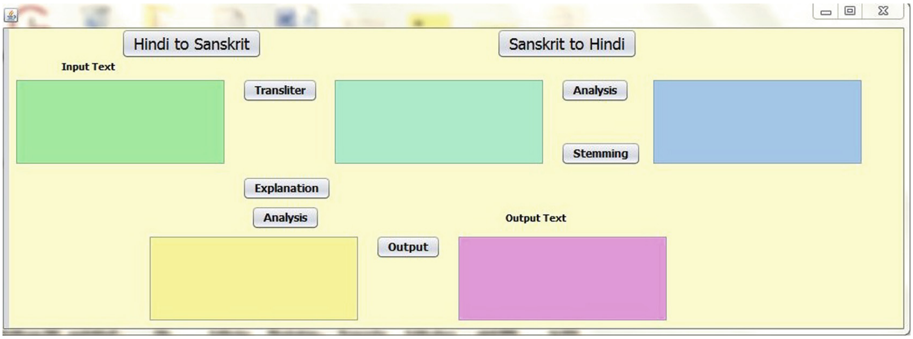


Fig. 3. GUI of the proposed MTS in Java

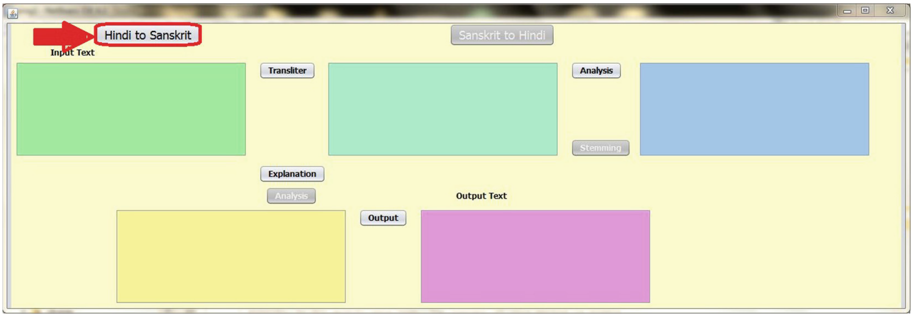
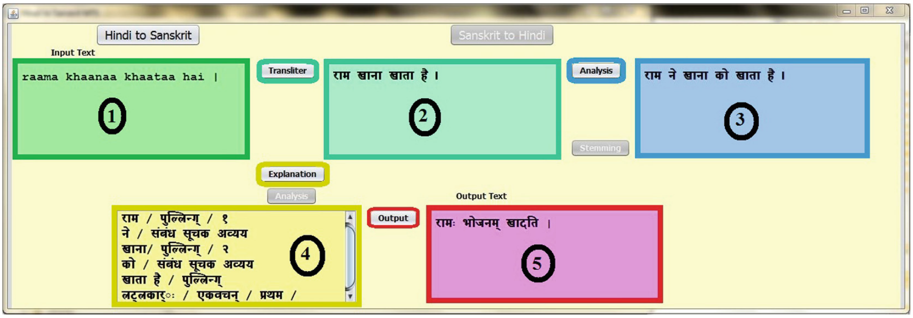


Fig. 4. Enabling Hindi to Sanskrit MT flow



raama khaanaa khaataa hai |

राम खाना खाता है ।

राम ने खाना को खाता है ।

राम / पुल्लिन्ग् / १
ने / संबंध सूचक अव्यय
खाना/ पुल्लिन्ग् / २
को / संबंध सूचक अव्यय
खाता है / पुल्लिन्ग्
लट्लकार्ः / एकवचन् / प्रथम /

रामः भोजनम् खादति ।

Fig. 5. (Hindi to Sanskrit MT flow) output of the system

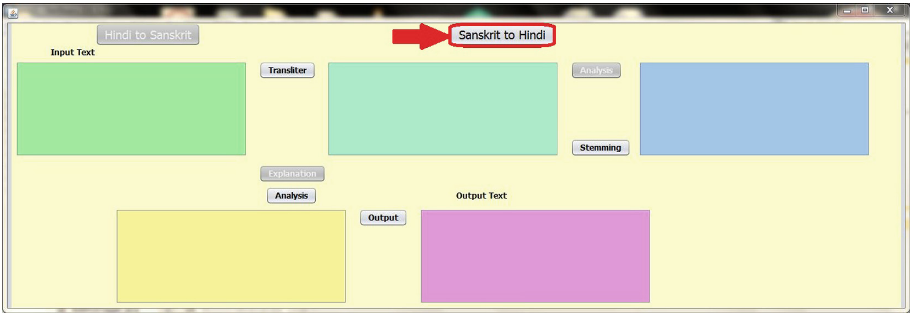
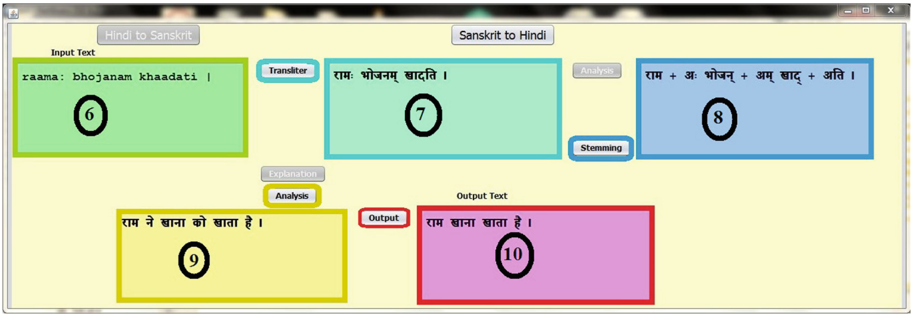


Fig. 6. Enabling Sanskrit to Hindi MT flow



raama: bhojanam khaadati |

रामः भोजनम् खादति ।

राम + अः भोजन् + अम् खाद् + अति ।

राम ने खाना को खाता है ।

राम खाना खाता है ।

Fig. 7. (Sanskrit to Hindi MT flow) Output of the system

of bilateral MTS. Firstly, the user needs to select the MT flow. According to his choice certain buttons are disabled. Figures 4 and 5 depict the Hindi to Sanskrit MT flow. Similarly, Figs. 6 and 7 depict the Sanskrit to Hindi MT flow. Figure 4 depicts what happens when Hindi to Sanskrit MT flow is selected. There are 5 fields and a button is associated to four of them, as shown in Fig. 5. Field 1 represents the input Hindi text to be entered. Field 2 and the corresponding button represent the transliteration phase. Field 3 and the corresponding button represent analysis phase. Field 4 and the button provide the explanation of the analysis phase. The final translation output is provided by field 5 and its corresponding button. Similarly, Fig. 6 shows what happens when Sanskrit to Hindi MT flow is selected. Field 6 represents the input Sanskrit text to be entered. Field 7 and the associated button represent the transliteration phase. Field 8 represent the stemming phase. Field 9 represent the analysis phase. Filed 10 provide the final translation output in Hindi.

3 Conclusion and Future Work

The proposed system performs two-way translation between Hindi and Sanskrit language. The translation from either sides require parsing of source language text, syntax analysis, semantic analysis, mapping (source to corresponding target), arrangement, pragmatic analysis and then the final output in the form of target language text. The system is easy to design once the two involved languages are thoroughly studied and understood. Their grammatical properties are pivotal here. Since the system provides a detailed step by step explanation of each phase involved, it can be used as a learning and teaching pedagogy tool. The beginners can get a deep understanding of the translation process and the grammatical structure of the involved languages. The system when designed and implemented will prove to be an excellent tool for teaching and learning. Furthermore, the accuracy of such a system can be improved by constantly updating the corpus for every exception or a new case that may arise.

References

1. Singh, M., Kumar, R., Chana, I.: GA-based machine translation system for Sanskrit to Hindi language. In: Khare, A., Tiwary, U.S., Sethi, I.K., Singh, N. (eds.) Recent Trends in Communication, Computing, and Electronics. LNEE, vol. 524, pp. 419–427. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-2685-1_40
2. Jaideepsinh, K., Saini, J.: Sanskrit machine translation systems: a comparative analysis. Int. J. Comput. Appl. **136**, 1–4 (2016)
3. Bahadur, P., Jain, A., Chauhan, D.S.: Architecture of English to Sanskrit machine translation. In: 2015 SAI Intelligent Systems Conference (IntelliSys), pp. 616–624. London (2015)
4. Saini, S., Sahula, V.: A survey of machine translation techniques and systems for Indian languages. In: 2015 IEEE International Conference on Computational Intelligence & Communication Technology, pp. 676–681, Ghaziabad (2015)
5. Rathod, S.G.: Machine translation of natural language using different approaches: ETSTS (English to Sanskrit translator and synthesizer). Int. J. Comput. Appl. **102**(15), 26–31 (2014)
6. Mane, D.T., Devale, P.R., Suryawanshi, S.D.: A design towards English to Sanskrit machine translation and sythesizer system using rule base approach. Int. J. Multidiscip. Res. Adv. Eng. (IJMRAE) **2**(2), 405–414 (2010)
7. Mishra, V., Mishra, R.B.: Study of example based English to Sanskrit machine translation. Polibits **37**, 43–54 (2008)

8. Shahnawaz: Conversion between Hindi and Urdu. In: International Conference on Computing, Communication & Automation, pp. 309–313 (2015)
9. Mishra, V., Mishra, R.: English to Sanskrit machine translation system: a rule-based approach. Int. J. Adv. Intell. Paradig. **4**(2), 168–184 (2012)
10. Gupta, V.K., Tapaswi, N., Jain, S.: Knowledge representation of grammatical constructs of Sanskrit Language using rule based Sanskrit Language to English Language machine translation. In: 2013 International Conference on Advances in Technology and Engineering (ICATE), pp. 1–5, Mumbai (2013)
11. Mane, D., Hirve, A.: Study of various approaches in machine translation for Sanskrit Language. Int. J. Adv. Res. Technol. **2**(4), 2278–7763 (2013)
12. Shukla, P., Shukl, D., Kulkarni, A.: Vibhakti divergence between Sanskrit and Hindi. In: Jha, G.N. (ed.) ISCLS 2010. LNCS, vol. 6465, pp. 198–208. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-17528-2_15
13. Goyal, P., Sinha, R.M.K.: Translation divergence in English-Sanskrit-Hindi language pairs. In: Kulkarni, A., Huet, G. (eds.) ISCLS 2009. LNCS (LNAI), vol. 5406, pp. 134–143. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-93885-9_11
14. Gehlot, A., Sharma, V., Singh, S., Kumar, A.: Hindi to English transfer based machine translation system. Int. J. Adv. Comput. Res. **5**(19), 198–204 (2015)
15. Mall, S., Jaiswal, U.C.: Developing a system for machine translation from Hindi language to English language. In: 4th International Conference on Computer and Communication Technology (ICCCT), pp. 79–87, Allahabad (2013)
16. Agrawal, P.: A machine translation system for Sanskrit to Hindi language. Ph.D thesis report. Discipline of Computer Science Engineering, IKG-Punjab Technical University (2018)
17. Jain, L., Agrawal, P.: English to Sanskrit transliteration: an effective approach to design natural language translation tool. Int. J. Adv. Res. Comput. Sci. (IJARCS) **8**(1), 1–10 (2017)
18. Jain, L., Agrawal, P.: Text independent root word identification in Hindi language using natural language processing. Int. J. Adv. Intell. Paradig. (IJAIP) **7**(3–4), 240–249 (2015)
19. Sethi, N., Agrawal, P., Madaan, V., Singh, S.K.: A novel approach to paraphrase Hindi sentences using natural language processing. Indian J. Sci. Technol. **9**(28), 1–6 (2016)
20. Sugandhi, R.S., Shekhar, R., Agarwal, T., Bedi, R.K., Wadhai, V.M.: Issues in parsing for machine aided translation from English to Hindi. In: World Congress on Information and Communication Technologies (WICT), pp. 754–759, Mumbai (2011)
21. Tapaswi, N., Jain, S.: Morphological and Lexical analysis of the Sanskrit sentences. MIT Int. J. Comput. Sci. Inf. Technol. **1**(1), 28–31 (2011)
22. Ashraf, N., Ahmad, M.: Machine translation techniques and their comparative study. Int. J. Comput. Appl. **125**(7), 25–31 (2015)
23. Chand, S.: Empirical survey of machine translation tools. In: International conference on Research in Computational Intelligence and Communication Networks (ICRCICN), pp. 181–185, Kolkata (2016)
24. Bharati, A., Chaitanya, V., Sangal, R.: Natural Language Processing: A Paninian Perspective, 7 Printing edn. Prentice Hall of India, Delhi (2010)
25. Briggs, R.: Knowledge representation in Sanskrit and artificial intelligence. AI Mag. Spring **6**(1), 32–39 (1985)

Throughput Enhancement and Loss Classification in Wireless Networks Using Machine Learning

Mohammad Ummer Chopan(✉), Pooja, and Amit Upadhyay

Department of Computer Science, Sharda University, Greater Noida, India
2017011245.mohammad@pg.sharda.ac.in,
{pooja.1,amit.upadhyay}@Sharda.ac.in

Abstract. TCP is problematic in wireless systems since it reacts to packet loss brought about by congestion in the same fashion as it does when brought about by a link error. It cuts the rate at which the information is sent which reduces the throughput in wireless networks. Since TCP was developed long before wireless networks were even conceived, it was not designed to consider the physical characteristics of wireless networks. A strategy for upgrading TCP execution over wireless networks is to classify loss causes, and consequently enabling TCP to react just to those losses perceived as being brought about by congestion. Our point, in this paper is to propose a solution which aims at improving the throughput in wireless networks without altering the existing TCP congestion control mechanism, by utilizing machine learning techniques including hybrid methods to classify loss causes and thereby investing the existing TCP with a mechanism that allows it to react only to those loss whose loss cause is classified as congestion and disregard those losses whose loss cause is classified as link error.

Keywords: TCP · Wireless networks · Machine learning · Loss classification

1 Introduction

TCP is one of the extensively used protocols in today's world. TCP was at first structured and advanced for wired systems. The underlying target of TCP was to productively utilize the available bandwidth in the network and to abstain from overburdening the system (and the resultant packet loss) by properly throttling the senders' transmission rates [1]. TCP congestion control mechanism in vogue is relies on the notion of congestion window and the size of this congestion window which is adjusted as and when required according to a specific procedure. This mechanism works by first sending the packets slowly, then increasing the rate steadily (additively) and in the event of packet loss, reduce it more abruptly (multiplicatively). Network congestion is deemed to be the underlying reason for packet losses. Thus, TCP execution which comes up short on the capacity to isolate congestion losses from link error losses on the wireless link is regularly inadmissible when utilized in wireless systems and therefore requires different enhancement strategies [2]. In recent years, as wireless connections have ended up being progressively essential in the world of networks, the subject of

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 322–333, 2019.
https://doi.org/10.1007/978-981-15-0108-1_30

TCP conduct over wireless links has advanced toward getting to be significant. In such a scenario, unmodified TCP seems to malfunction in wireless systems because the existing TCP congestion control mechanism follows exactly the same procedure of congestion control in wireless systems as it does in wired systems. Whenever TCP perceives a congestion event it minimizes its congestion window and as a consequence the transmission rate or packet flow gets reduced [3, 4]. Such a reaction to a congestion event is admirable but only in wired systems. However, link errors which occur more frequently in wireless networks, and the TCP which has no mechanism of recognizing a link error as distinct from a congestion event, treats it likewise and in response decreases the congestion window which consequently reduces the packet flow and hence the throughput, without any justification.

1.1 TCP Congestion Control

TCP congestion control mechanism involves different aspects of add-on increase and multiplicative decrease schemes. TCP utilizes the notion of congestion window (cwnd) maintained by the sender. Cwnd is maintained for each TCP session [5] and represents the maximal volume of data that can be sent across the network. After each acknowledgment is received the sender increases the cwnd size by one MSS (Maximum segment size). MSS is determined during connection establishment by using an option of the same name. Sender keeps sending packets as it receives acknowledgments for each packet it has sent, implying that the window starts slowly but grows exponentially. This slow start phase doesn’t continue indefinitely. Rather, it stops as soon as the cwnd size reaches ssthresh (slow start threshold).

A packet loss is detected when a sender receives three duplicate acknowledgments. It then responds immediately by dividing the congestion window (cwnd) which therefore reduces multiplicatively the traffic flowing through the network. This mechanism does well in wired networks but it leads to drastic throughput problem in wireless networks where packet losses occur, more often due to link error.

The rest of the paper is organized as follows: Sect. 2 presents related work. Section 3 discusses methodology followed by machine learning techniques in Sect. 4 and the last, Sect. 5 provides the results of our work.

2 Related Work

Several researchers like Barmann, Matta, Biaz and Vaidya [6–9] have already worked with the loss classification related issues in wireless networks. Fuzzy based solution have likewise been proposed in the recent past by ElAlarag [10], it also aims at improving TCP by blessing it with a classifier but more importantly with a fuzzy based Classifier [11]. The Idea is to classify the loss causes and henceforth help TCP better comprehend the genuine congestion loss from the seemingly one. El Khayat [12] in her research on enhancement of TCP in wireless/wired networks has focused on using the information attainable at the transport layer. She has used supervised learning algorithms to classify loss causes and make predictions. Among the solutions considered, we have adopted to use the one that uses Information available at the end-systems because such an information can easily be harnessed and put to use using machine learning techniques. we have also used a hybrid approach of machine learning techniques, appending one algorithm over the other. The three parameters that are employed to anticipate packet loss causes in networks are the round trip times, the inter arrival times and the one-way delays [13]. Round trip time which is the total time which involves a packet to be sent and to receive back an acknowledgement is used in Hidden Markov Models. Inter-arrival time which is passage or duration between the arrival of packets is utilized by Biaz and one-way delay represents the total duration a packet takes to reach from source to destination, has been used by Alarag. The investigation hitherto demonstrated that the inter-arrival times and the one-way delays are correlative with regard to anticipate loss causes. Moreover, any alteration in the return path influences the round trip time without evenly influencing the packet loss cause in a network. In every one of these works, rules have been derived on ad hoc basis and tuned physically. Moreover, according to Westwood, should the current round-trip-time fall below 1.4 RTTmin the loss is classified as a result of link error. RTTmin represents the minimal round trip- time evaluated since the start [14]. As opposed to this Veno [15] gauges the excess by utilizing just the inter-arrival times and terms the loss as a result of congestion if the backlog is higher than three.

Our methodology, on the other hand broadens the contemporary work in a few ways. Rather than utilizing just a single indicator, we will join a few of them in our rules of classification. Moreover, the rules formulated as such shall be tuned automatically from a vast database of losses while availing the assistance from our learning algorithms.

2.1 Machine Learning Concept

Machine learning is a category of algorithms that enables programming applications to become more accurate in foreseeing results. Machine learning is by far the most exciting technology that ever existed. It provides the computers/machines the ability to learn just like humans. The idea is that systems can learn from the underlying pattern of the data and consequently arrive at solutions without human mediation. Machine learning algorithms are mainly categorized as supervised and unsupervised.

Supervised Machine learning: Supervised learning is a labeled learning in which algorithms anticipate future events by utilizing the knowledge it has gained from the past data. We can also say that the result is already known.

Unsupervised Machine learning: It is a kind of unlabeled learning, output is not presented. In this method the algorithm makes decision by learning from the hidden structure of the data.

3 Methodology

In this portion we are going to discuss which category of machine learning along with the machine learning techniques have been used for actually classifying these two distinct losses. We have focused on using supervised machine learning techniques for our problem of loss classification.

3.1 Supervised Learning

Supervised learning is that sort of automatic learning in which both input and desired output are present. The objective of this learning process is to demonstrate a mapping from some input contribution to the output or yield. Usually, the learning sample (LS) which is available for observation consists of input/output pairs LS = {x_1/y_1, x_n/y_n}. Xi represents the input vector relating to the ith observation, which is to be fed into the machine. Likewise y_i represents the output vector value. The output or yield produced may either be discrete or continuous. A problem is termed either classification or regression based on the output it takes. When the output is in a discrete form it is termed as classification and when it is continuous it is referred to as regression [15].

This problem of classification can be solved by using supervised learning algorithms. There are various such supervised machine learning algorithms. The algorithm generates a hypothesis while learning from the input data. Supervised learning algorithms tend to learn through iterative optimization of objective function to anticipate the output related to new unknown data. Using cross entropy function to compute loss, learning algorithms are inclined to minimize the loss function while using gradient descent by adjusting weights/parameters that result in the optimal solution. After adequate training, the model is able to correctly label new data. The learning algorithm can likewise collate its generated output with expected one and find possible flaw so as to alter the model in the like manner. The fundamental standard employed to evaluate learning algorithms is their ability to accurately make predictions i.e. how effectively they generalize on the test data. Another essential factor is interpretability followed by computational efficiency of the learning algorithm. However, there is a trade-off between these three criteria. The algorithm returns a function f (say) from some input data on which we will breeze through test informational index to discover the prediction accuracy of the model.

3.2 Dataset Description

To check the feasibility of involving machine learning for classifying losses and enhancing throughput in wireless networks, we tried various popular machine learning algorithms utilizing python as the machine learning framework and the simulated data generated by the "Montefiore Institute Belgium".

In our examination, we have used a dataset of losses generated by the "Montefiore Institute Belgium". Our dataset is a combination of both congestion and link error loss events, amounting to 35441 such losses corresponding to an ample number of different random topologies. Such large diversity of such topologies are meant to ensure the generality of the rules obtained. Figure 1 gives a simple representation of losses. The data was generated through simulation using ns-2 simulator [16] by selecting arbitrary number of nodes (somewhere in the range of 10 and 600) involving haphazard associations between these nodes. Parameters such as buffer size, propagation delay and the bandwidth all of them are chosen haphazardly between these connections. Bandwidth is allowed between 56 kbps to 10 Mbps and propagation delay varies between 0.1 ms to 500 ms.

Since our dataset is a combination of losses. Figure 1 gives a histogram representation of our target values. The two different types of losses correspond to a vast majority of events.

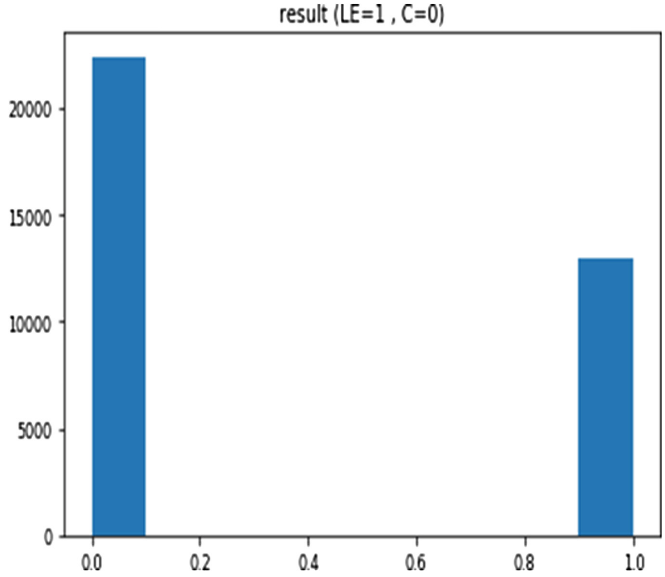


Fig. 1. Represents a diagrammatic view of the loss events

3.3 Choice of Inputs

The decision about the input variables is coordinated by a few limitations. The information obtained as such should succinctly foretell congestion. For the classification models to be essentially of any use, these factors ought to likewise be quantifiable (at the end systems). At either on the sender or the receiver side, the data required to foretell a congestion event are the one-way delays and the inter arrival times.

To figure our input variables, we utilize the data obtained from triple packets following the loss and the packet that precedes it. In addition to this, we calculate maximum, minimum, standard deviation and average of the inter-arrival time and the one- way delays for the packets sent during one round-trip-time and we store these values for the last two consecutive round trip times before the present time. Our input therefore consists of the packet preceding the loss, three packets following the loss and the packets for the two preceding round trip times. We likewise incorporate in our inputs the ratios or the proportion between our calculated minima, maxima and average of our indicators (one way delays and inter arrival time) in all pairs of such categories. The parameters so obtained can be used for training the model for acquiring higher accuracy because machine learning techniques are able to perform better when there is a multitude of data. We will use these parameters to train our model and to predict the loss cause

3.4 Data Pre-processing

To begin with, this process involves the collection of data. After collecting the data the said data needed to be cleaned for null values and possible outliers. We performed the data reduction and several other processes to make the available data apt for use in machine learning algorithms. This process essentially involved label encoding and several other techniques to convert qualitative data into quantitative data, for machine learning works only with numbers.

4 Machine Learning Techniques

After performing data pre-processing the classification techniques can then be applied. For this reason we divided our dataset into training and testing sets, allocating 80% of our data for training purposes and the rest 20% for testing. This implies maximum portion of data is used to manufacture the model and the remaining portion of this data is utilized to test the learner performance. Also the K-fold cross validation (for K = 5) has been used, dividing the training set into five folds and randomly selecting instances from each category. Several different learning algorithms have been utilized for the purpose of classification in our project and these are as follows:

KNN or K-Nearest Neighbor: It is a machine learning algorithm which is used for the classification of data into target values (in our case error due to congestion or due to link error). It stocks all available cases and classifies [17] new unknown cases based on similarity measure using distance function. It usually takes into consideration the Euclidean distance between data points. In KNN a case is classified as belonging to a particular class based on the majority vote using distance vector as a metric. Distance vectors are calculated by certain distance functions such as:

$$\sqrt{\sum_{i=1}^{k} (x_i - y_i)^2} \tag{1}$$

This is Euclidean distance formula and is valid for continuous variables. KNN can also be used for regression purposes but is mainly used for classification purposes. An imperative downside of this technique is that it resource demanding and results may vary while altering the values of K.

Logistic Regression: It is a classification technique that is generally used for exploring the data [18] in which there are one or more independent variables that dictate the outcome. The outcome contains 0 or 1 in a binary classification problem. In logistic regression the input data is passed through a sigmoid function which limits the output data between 0 to 1 ranges. In our case Logistic regression can predict the probability of an outcome that can only have two values as is indicated above (C or LE) for 1 and 0 respectively. Logistic regression involves the use of maximum likelihood estimation to obtain the model coefficients that relate predictors to their target.

Random Forests: Random forest is a flexible, easy to use learning algorithm that provides even without hyper parameter tuning outstanding results all the time. Random forests improve both accuracy and computational efficiency. Random forest classifier handles missing values and prevents over fitting. It is a combination of different tree predictors. From these forests of trees generated using separate feature set, proximities are computed. These proximities are use in detecting outliers replacing missing values and providing low dimensional view of the data. One of the significant attribute of random forests is that there is no need for performing cross validation to get an estimate of an error. Such an error is estimated internally in random forests.

Decision Trees: This is a standout amongst the most prominent learning algorithms. Decision tree is an easy to understand algorithm with each internal node depicting a test on an attribute and each branch depicting an output of the test [19]. Parameters like information gain and classification error in ID3 and Gini index in cart is used as a means for splitting nodes. Splitting of nodes cease when a particular criterion in met such as each terminal node yielding single value (such as C or LE). The terminal node contains the class label.

The problem with the decision trees is that they over fit and for that reason decision tree variant like bagging, boosting and random forests are preferred for classification problems. Boosting is by far the best method that produces results with much higher accuracy with least over fitting.

5 Results and Analysis

The results are obtained using a confusion matrix in terms of accuracy, precision, recall, F1 score and cross validation mean of the algorithms. More importantly accuracy measure is used in identifying the classifier performance. Figure 2 shows the relationship between the different feature columns of a dataset and the respective their impact on our target value. Features that generate distinct regions of these two error types are preferred over those that create overlapping regions. Parameters which are overwhelmingly overlapped are less preferred. Larger values of these parameters show a correlation with the error type and are therefore retained to be used in algorithms. The advantage of this procedure is that it empowers us to choose just those parameters that impact the outcomes and thusly drop the less compelling ones from a dataset with huge sections.

Although, we have included around forty parameters in our input but for the sake of illustration we have only shown here in Fig. 2 some features columns of our dataset and their corresponding impact on the target value (which is C and LE).

Based on the observation from the histogram plots important features are extracted and utilized accordingly. A classification model is built with each learning method and its error due to congestion (denoted by C) and due to link error (denoted by LE) is calculated and without any bias by setting P_{th} exactly at 0.5.

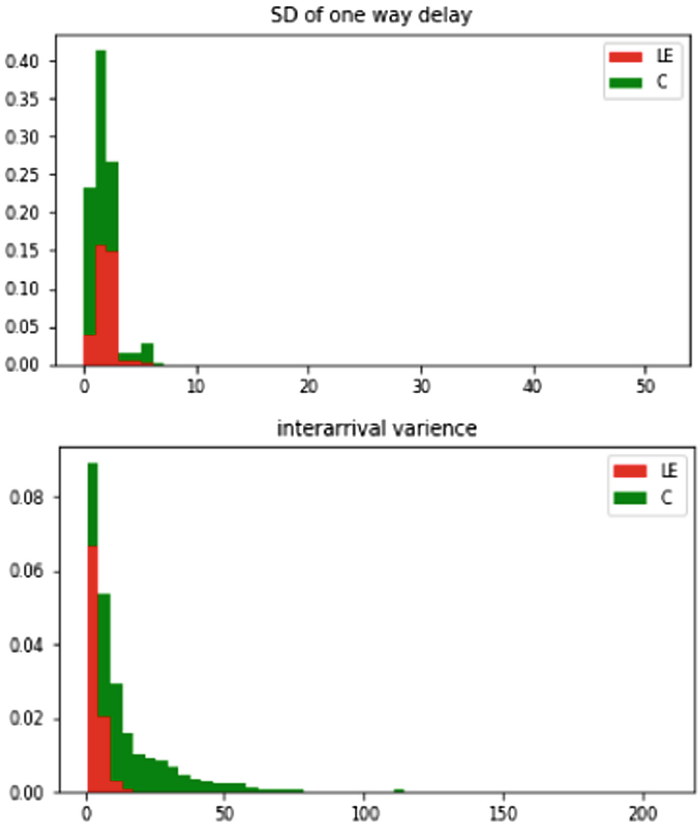


Fig. 2. Histogram representation of feature importance

The accuracy of a model is defined as how accurately a classifier has distinguished or classified these different losses. We have in our case used four supervised learning algorithms like Logistic regression, K-nearest neighbor, decision trees and random forests with certain characterized parameters following accuracies were recorded. Figure 3 gives the accuracy and cross validation mean of the different learning algorithms used. It shows that Random forests produced the highest 90% accuracy and the cross validation mean of 89%.

Furthermore decision trees over fit and produce cent percent accuracy. This over fitting is mitigated through pruning.

Moreover, one of the standouts of the random forest is that it generates the matrix of feature importance along with their corresponding ranking.

While accuracy is used as a general tool for determining the classifier performance precision, recall and F1 score measure shown in Fig. 4 are also used to give the overview of the classifier performance in terms of error rate. Precision is a good measure to determine performance when the cost of false positive is high and f1 score is needed when we need to strike a balance between precision and recall.

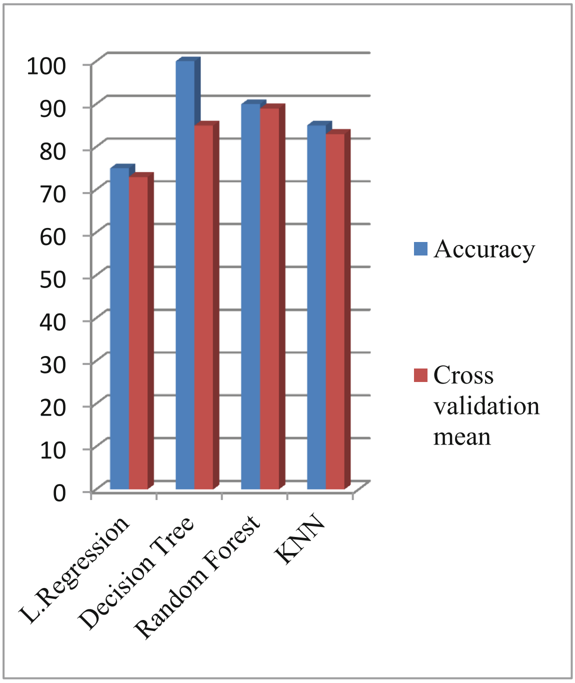


Fig. 3. Algorithm performance comparison

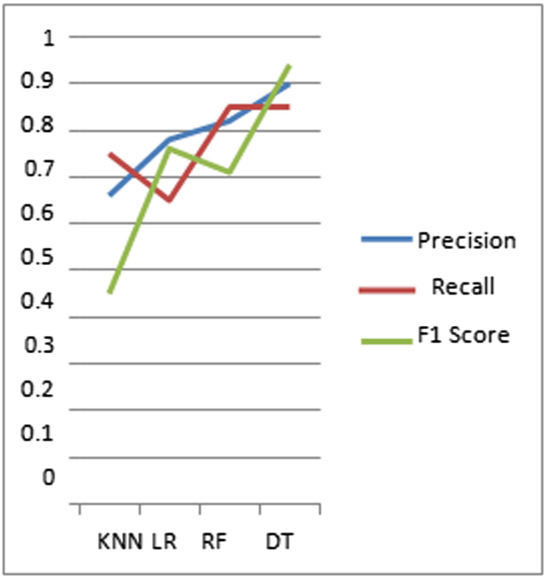


Fig. 4. Precision, Recall and F1 score of the algorithms

5.1 Hybrid Approach

It is much reliable to use various different models instead of only one. Voting ensemble is one of the simplest methods of combining the prediction of multiple machine learning algorithms. Voting consists in training our model using various different algorithms and then ensemble them to predict the final output. A voting classifier is used to wrap up different sub models and average their predictions when approached to make predictions for new data. It is a meta classifier for joining similar or conceptually distinct machine learning classifiers for classification by way of majority or plurality voting. This method employs hard and soft voting. It predicts the final class label in the hard voting, as the class label that has been predicted most frequently by the classification model and in soft voting it predicts the final class labels by averaging the class probabilities.

Up until this point, we have used one algorithm at a time and accordingly evaluated our classifier accuracy. As can be seen from the Fig. 3 in our experiment that algorithms like logistic regression in our classification problem, doesn't seem to generalize well with the unseen data and therefore provide minimum accuracy when compared with the other three algorithms used in our approach.

Using ensemble voting approach, for instance logistic regression when hybridized with other algorithms greatly improves its accuracy of classifying losses. Figure 5 shows the hybridized version of different machine learning algorithms and their corresponding accuracies. A Voting classifier generally has a higher accuracy than the individual classifiers. Such a technique of optimizing the output using different algorithms simultaneously can have a significant impact in improving the performance of certain classifiers and their derivatives. However we need to make sure that the models which a susceptible to similar types of errors don't aggregate the errors.

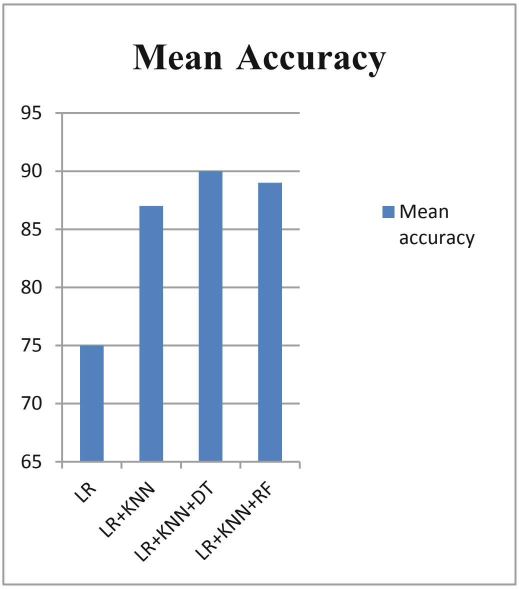


Fig. 5. Classifier performance with hybrid approach

6 Conclusion

In our work, we have used machine learning along with its ensemble schemes as a basis for detecting TCP packet losses over congestion prone wireless networks. We have in the main, centered around using supervised learning techniques to infer a mechanism (or a classifier) that can automatically recognize losses brought about by link error from those occurred due to congestion and consequently guide TCP when it ought to reduce the traffic flow across the network. We used algorithms like K-nearest neighbor, Logistic regression, random forest and decision trees on the learning sample and each one of them furnished an alternate model with very great error rates. Of all the algorithms used random forests produces the best outcomes individually. Moreover, combining or aggregating these classifiers significantly improves the accuracy of certain classifiers to predict the packet loss. We learned from our work that there is still very much required to be done in the future in order to make our system more compatible with still higher accuracy of making prediction. For much better results complex hybrid machine learning algorithms can be invoked in future works. Our work serves an impetus to the use of hybrid algorithms for classifying losses in wireless networks.

References

1. Balakrishnan, H., Padmanabhan, V., Seshan, S., Katz, H.: A comparison of mechanisms for improving TCP performance over wireless links. In: Conference Proceedings on Applications, Technologies and Protocols for Computer Communications, pp. 256–269. ACM Press (1996)
2. Xylomenos, G., Mahonen, P., Saaranen, M.: TCP performance issues over wireless links. IEEE Commun. Mag. **39**, 52–58 (2001)
3. Wu, H., Long, K., Cheng, S., Ma, J.: Performance of reliable transport protocol over IEEE 802.11 wireless LAN. In: Proceedings of INFOCOM 2002, Joint Conference of the IEEE Computer and Communications Societies, vol. 2, pp. 599–607, March 2002
4. ElAarag, H., Wozniak, M.: Improving TCP performance over mobile networks. J. ACM Comput. Surv. **34**(3), 357–374 (2002)
5. Biaz, S., Vaidya, N.H.: Distinguishing congestion losses from wireless transmission losses. In: Proceedings of IC3N, New Orleans (1998)
6. Barman, D., Matta, I.: Effectiveness of loss labeling in improving TCP performance in wired/wireless networks. In: Proceedings of the 10th IEEE International Conference on Network Protocols, pp. 2–11. IEEE Computer Society (2002)
7. Liu, J., Matta, I., Crovella, M.: End-to-end inference of loss nature in hybrid wired/wireless environment. In: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (2003)
8. Parsa, C., Garcia-Luna-Aceves, J.: Differentiating congestion vs. random loss: a method for improving TCP performance over wireless links. In: Proceedings of IEEE WCNC, pp. 90–93, April (2000)
9. ElAarag, H., Bassiouni, M.: Performance evaluation of TCP connections in ideal and non-ideal network environments. Comput. Commun. J. **24**(18), 1769–1779 (2001)
10. ElAlarag, H., Wozniak, M.: Using fuzzy inference to improve congestion control in wireless networks (2012)

11. Shi, K., Yantai, S.: Receiver centric fuzzy logic congestion control for TCP throughput improvement over wireless networks. In: Proceedings of International Conference on Fuzzy Systems and Knowledge Discovery, Tianjin, April 2011
12. El Khayat, I., Geurts, P., Leduc, G.: Enhancement Of TCP over wired/wireless networks with packet loss classifiers inferred by supervised learning. Wirel. Netw. **16**(2), 273–290 (2010)
13. Altman, E., Barakat, C., Ramos, V.: Analysis of AIMD protocols over paths with variable delay. In: Proceedings of IEEE INFOCOM, March 2004
14. Wang, R., Valla, M., Sanadidi, M., Gerla, M.: Efficiency/friendliness tradeoffs in TCP westwood. In: 7th IEEE Symposium on Computers and Communications (2002)
15. Fu, C.P., Liew, S.: TCP enhancement for transmission over wireless access networks. IEEE J. Sel. Areas Commun. **21**, 216–228 (2003)
16. El Khayat, I., Geurts, P., Leduc: http://www.montefiore.ulg.ac.be/Boosting-DT
17. Hall, P., Park, B., Samworth, R.: Choice of neighbor order in nearest neighbor classification. Ann. Stat. **36**, 2135–2152 (2008)
18. Hao, J., Priestley, J.L.: A comparison of machine learning techniques and logistic regression method for the prediction of past-due amount, September 2016
19. Sharma, H., Kumar, S.: A survey on decision tree algorithms of classification in data mining (2013)

Adaptive eLearning System: Conceptual Framework for Personalized Study Environment

Bhawna Dhupia(✉) and Abdalla Alameen

Department of Computer Science, College of Arts and Science,
Prince Sattam bin Abdulaziz University,
Wadi Addwasir, Kingdom of Saudi Arabia
{b.dhupia, a.alameen}@psau.edu.sa

Abstract. e-Learning system provides web-based learners the facility to access the courses online anytime anywhere through internet. Traditional e-Learning system does not provide individualistic model for learner and is not suitable for all types of learner opted for e-Learning course. Therefore, the satisfaction level of the learner is not very high in maximum cases. This problem of individual learning can be taken care by implementing Adaptive e-Learning. Adaptive learning is more interactive and popular these days. This learning is the combination of intelligent tutoring and machine learning. The delivery of the learning material in Adaptive e-Learning is based on the knowledge, goal, experience, interest and background of the learner. This paper gives a detailed introduction to the concept adaptive e-learning environment and its components. A student can choose the way of learning and according to that a suitable path is assigned to them. This way of learning increase the interest of the learner and positively influence the learning outcome of the course. Keeping this new technique of e-learning, a framework for the implementation of adaptive system is also proposed based on the basic principles of adaptive e-learning.

Keywords: e-Learning · Adaptive learning system · Conceptual framework

1 Introduction

The advancement of technology has affected the field of education extensively. Internet as one of the advancements in technology has raised the standard of teaching through numerous applications available online. Online learning is popularly known as e-learning. E-learning is a web-based instruction through which learning material is delivered to the learners. Many applications are available on the internet, but they only provide HTML pages and are not concerned about the need or learning pattern of the user. These learning materials are used by many users and can have different aims and goals of learning. The material provided in eLearning was not able to solve the problem of all the users due to the common syllabus and common learning styles. Therefore a new e-learning system came into existence called Adaptive E-learning System, which took care of various shortcomings in a traditional e-learning environment.

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 334–342, 2019.
https://doi.org/10.1007/978-981-15-0108-1_31

Education is the field which is more affected by the advancement of technology. Adaptive e-learning system (AES) has come to the fore in the last ten years. This e-learning system is more effective and efficient web-based intelligent way of delivering the education online. The AES is based on the user model, which gathers information about the users, their goals of learning, learning pattern, knowledge level and their aptitude. After gathering this information, it offers the best suitable style of learning to the learners [19]. The AES is the combination of an intelligent system to adapt the learning style of users and Web-Based Instruction (WBI), which is the combination of audio, video, graphics, plain text and hyperlinks known as hypermedia to provide relevant learning environment on the courses offered [1]. Adaptive learning is based on the theory that identification of learning style is a very crucial tool to improve the individual learning of a learner especially in online learning. In this system learner expects to have a unique style of teaching that helps achieve the learning goal. Most vital feature of adaptive learning is its ability to change the learning style and pedagogical environment for the learner depending on the continuous action of the user during complete learning process. The change on the learning strategy is directly related to the knowledge, goal, experience, interest and background of each and every person [31].

Education enables a country to develop and achieve its goals of development and prosperity. These days almost all the universities in the world are offering Hi-tech tools such as Blackboard to impart the e-learning mode of education to educate the youth but this need more to offer the personalized learning. Implementation of adaptive e-learning system will prove to be more effective in achieving the set goals. A number of applications are prevalent to support the adaptive e-learning pattern. To mention a few; AES-CS [15], INSPIRE [14] and ILASH [16], are developed to implement the adaptive e-learning system [5]. But the shortcoming of these is that most of the work is done on the user's level of knowledge and lacks many other features like learning styles which have to be accompanied to implement the real adaptive e-learning system [6].

This paper reviews the earlier research done in the field of AES. The various models offered to implement the adaptive learning environment; the framework of the conceptual adaptive learning model is also discussed. It also explains components of the adaptive learning environment.

2 Related Work

According to the Blooms Learning Taxonomy learning is improved when it is delivered in the learning style of individual [7]. The basis of adaptive learning lies in this statement. The most crucial point is to implement the adaptive leaning system is to first understand the learning style of each individual and then to formulate the teaching strategy for each one of them separately. Cronbach mentioned that the results of the learning are based on the characteristics of a person and their adaptability to learn the things [8]. Bloom's later findings state that one-to-one delivery of instruction is more effective than a conventional group- teaching. Bloom believed that all students can achieve high learning outcome provided the learning style is suitable to learner [9].

Education has undergone a phenomenal change in the last decade. E-learning was considered to be the best way to impart the training online and this system was adapted by many educational institutions. But, as discussed in the introduction e-learning cannot provide the customize education to the leaners. So, to overcome the shortcomings in e-learning, adaptive learning is being accepted by the educational organizations. Many current researchers have conducted research on the benefits of delivering education using adaptive e-learning system. According to Duong et al., implementation of personalized learning system for Adaptive E-learning is formalized by knowing the ontology of an individual. Ontology of a person helps to find out the character traits and helps to design the personalized training pattern for them [10]. It also includes the user personal profile and track of the user's general searching and browsing behaviors. This helps to decrease the burden of extra information which is not relevant to the learner profile. By providing the learning pattern and material linked with the interest of the learner provides greater confidence and satisfaction which results in achieving higher learning outcomes [11].

Pashler et al. [17] have comments on learner styles that adapting the learner styles gives better learning outcome. Learning styles are commonly defined as a group of emotional, psychological and cognitive factors which collectively give an idea about the learner characteristics, interaction and response to the learning environment offered to them. Learning style also comprehends the information type, presentation style such as visual, auditory or kinesthetic and learning actions. Sonwalkar recent study [12] states that adaptive leaning model is highly contributed by artificial intelligence, intelligent tutor, and adaptive control. The first step in adaptive learning implementation is to find out the behavior of a learner towards learning, then to conduct the analysis of the learning behavior and finally to offer a learning model which suits to the learning pattern of the learner. In order to understand the learning pedagogy of the users, he suggested a learning cube will be discussed subsequently. Snow and Farr [27] learning methodology is incomplete if the preference of the learner is not considered while imparting the education. If the interest of the learner is not included during the course delivery, the learning process is incomplete and unsuccessful. Russel [28] also suggested that educator should identify the varied learning styles of the learner and implement the respective style for imparting the education. According to a researcher [32] the change of pattern during the learning process maintains enthusiasm to learn among the learner. This method greatly affects the learning outcome of the learner. Educator should also use latest technologies to make the learning process interesting and interactive.

3 The Framework

3.1 Pedagogical Framework of Adaptive e-Learning System

The pedagogy of learning deals with the learning pattern of the learner based on their knowledge, experience, situations, and environment. This also covers the learning goals of every individual. As mentioned before, every learner has an own way of learning. Following the pattern adapted provides the desired learning outcome. The pedagogical framework is responsible for the representation of knowledge through metadata [12]. The data is captured through a questionnaire which helps to rate the knowledge and behavior of the learner. The data is analyzed by an intelligent system and based on the formative evaluation a unique learning strategy is offered.

3.2 Conceptual Framework

Adaptation in learning model can be implemented in a better way by formulizing the models required for adaptive e-learning system in an efficient manner. To define the conceptual framework for the system, three models required need to be defined namely; Domain Model, Learner Model and Adaptive Model. The domain model includes the knowledge and information about the courses offered and learner model includes all the information related to the learner and adaptive model deals with the techniques to offer the learning patterns and many more material for the related course. These models will be studied in the later course of the paper.

Domain Model: Domain model is a combination of both behavior and data. Domain specific information shows the skills and status of the learner in specific subjects [31]. It can be called as a main conceptual model of any project. In the concept of an adaptive environment, it contains information about the knowledge part of the course contents to cooperate the application for the adaptive course delivery. It has all the details of the course including topics, contents, modules, links, images, books which design the structure of the information incorporated [20]. The main purpose of the Domain Model is to classify the data based on some criteria fixed by the programmer. The domain model perhaps contains information regarding student learning activities. This model is capable of observing the activity of the learner and connects them to the type of learning styles suitable for that particular learner. Adaptive learning material includes text, graphics, audio, video, animation and simulations which are used to design the suitable learning model for the learner [25]. There are two main components of a Domain Model: course material and learning methods. Depending on the learning style adapted by the learner, course material should support the learner. It should be in accordance with the plan proposed by the system for the learner. The course material and the learning method should be closely linked with each other to get the desired result. It also focuses on the design of the audio-visual content and suitable links to browse the course contents. The design should be user friendly for all the users and browsing of information should also be easily accessed.

Learner Model: It contains all the information regarding the learner such as their domain, level of knowledge, learning pattern and other personal information. It not only saves the information but also tracks the learning records of the users [21, 30]. Information on learner model is divided into two major categories; domain-specific information and domain-independent information [6]. Domain-specific information tells everything related to the knowledge level, skills, understanding level, track records learning behavior, evaluation records, etc. [29]. Domain-independent information may include learning goals, which help to measure the performance of the learner to achieve their goals. It may include cognitive capabilities such as thinking and reasoning skills and team learning skill, motivational states that drive the learners, background, and experience, and preferences. According to Schiaffino et al., There are two methods for managing the behavior of the user. First method is to track the repetitive behavior of the learner during learning process. This tracking of behavior helps in designing a preference model for that individual learner [34]. The second approach is based on cognitive behavior of the learner. This helps to understand the psychology of the learner towards learning. This information helps to define the learning styles of the learner, later can be assigned according to the analysis of the behavior [35].

Adaptive Model: This model is based on teaching-learning activities of the learner [23]. The learning activities of the users are analyzed and according to their behavior, they are categorized into various groups. According to one of the author, Sonwalker [12, 13], the learners can be divided into five categories depending on their analysis, namely; apprentices, incidental, inductive, deductive and discovery. Based on the learning styles any one pattern can be offered to the learners. This model is responsible for analyzing the behavior of the learner and to offer the learning techniques based on the behavior. This model works by selecting the learning style according to the need of the students and assigning the appropriate category for the learning process. These categories are divided into different types of knowledge; such as basic knowledge, procedural knowledge- to understand the concept in a systematic way, conceptual knowledge- to link the topics by implementing it into a real example [4]. Utilizing this information a complete learning module is assigned to the students until they reach the end. They can have the capability or permission to reiterate the nodes for more clarification and also for revision. The adaptive model also contains the link for the various references which gives clarity about the concept that the learner is dealing with. Adaptive model keep on reviewing the learning pattern of the student till the end. In case, any change in the learning pattern is found, new learning style is offered to the learner. According to the style, learning material, learning path, evaluation system is also get changed.

4 Proposed Framework for Adaptive eLearning System

This section discusses the proposed framework to implement the adaptive e-learning system. Adaptive e-learning system (AES) is an established area of research which is an integrated technology with the concepts of CAI, ICT, and hypermedia [24]. Computer Aided Information, Information and Communication Technology and Hypermedia are combined to give an environment for adaptive learning. The adaptive learning system has become very popular due to its ability to cater the needs of the heterogeneous type of users. It is known that the learners don't have the same level of IQ levels. Hence it is useless to offer them the services which are perhaps above or below their level. Adaptive learning offers the service according to the individual's profile. It helps learner to understand the concept according to their IQ level which helps the student to understand the concept clearly and make the learning process interesting for all categories of learners. As discussed above, there are three basic models for adaptive learning environment namely; domain model, learner model, and group model. Therefore, the framework of the adaptive learning revolves around this basic structure.

The basic model of the AES framework is domain-model. This is primarily a source for all types of data required for the system to carry forward. It includes all the information relevant to the course offered in the system with all the details such as syllabus, course content, quizzes, audio/video content, assignment, tests, etc. The most important part of the model is Adaptive Engine. It is responsible for discovering individual learning behavior of the learner. It will also track the preferences, achievement and activities of the learning during whole learning process. According to VanLehn and du Boulay [33], it works as a combination of loops, outer loop decides which task should be offered to the learner and inner loop organize the steps to complete the task assigned to the learner by outer loop. This process in the adaptive engine is implemented with the help of learning algorithm. In our model also, it will be working as a loop till the completion of the learning process. Behavior of the learner will be tracked and changes if any found will be implemented in the preferred learning style of the user.

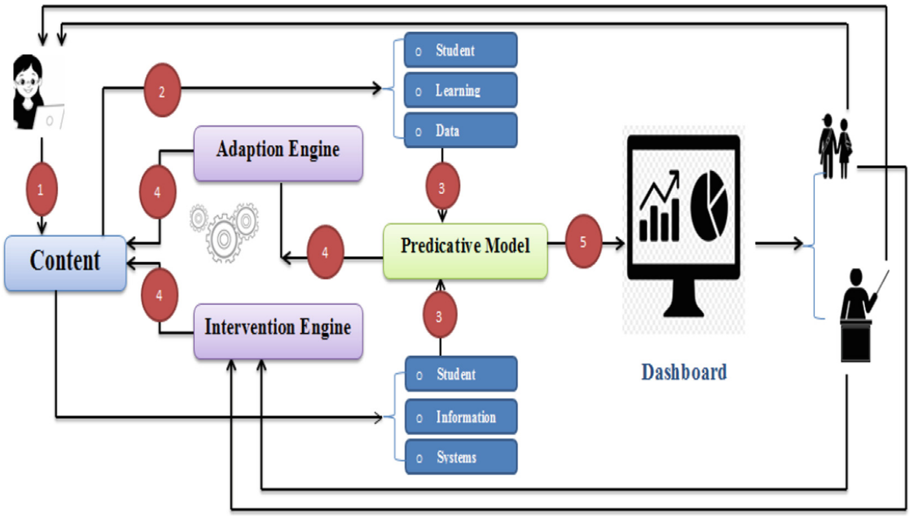


Fig. 1. Adaptive e-Learning conceptual model

Figure 1 is a detailed conceptual framework for the project. The working of the model is elaborated in five steps. At first step is for the user to enter the information regarding the user profile which includes the personal information and the course which a learner wants to adopt. After creating the profile, the user is prompted to fill a questionnaire. This questionnaire consists of questions that enable to learn the learner's behavior. In the second step the data fed by the user will be segregated into student information, learning, and course data. This information gathered in the third step will be transferred to the predictive model of the system. Predictive model analyses the data gathered about the user preference for learning style. The questionnaire to learning preference which is offered in the start of the registration the learner gets analyzed in predictive phase. Based on the result, a learning pattern is offered to the learner. Later this data gets transferred to the adaptive engine, where the course material and delivery system will be assigned. During the learning process, the learning pattern gets analyzed and if there is any change found in the learning style, the new pattern is assigned to the same user. This change is implemented by intervention engine. There will be dashboard available to each users and faculty member to track the progress of the course conducted. As suggested by Sonwalker [11, 12], the learning pattern of the learner keeps changing and needs to be tracked on regular basis. These learning patterns are categorized by him into five categories, namely; Apprenticeship, Incidental, Inductive, Deductive and Discovery. The change of learning style after analyzing the behavior of learner during the course, helps the learner to increase the interest of the learner and result in the effective outcome of the teaching [24]. Other researchers commenting on the learning style mentioned that student learns in a better way if the teaching styles of the students are adopted [26]. To know the preferred learning style of the student there are few techniques to be followed such as a questionnaire or psychometric test. These tests consist of questions related to the learner personality, behavior and attitude. By analyzing these tests, domain system can be chosen as to which learning styles can be adopted for a specific learner Depending on the result, preferred style can be visual, auditory and kinesthetic [21]. The learner model dynamically records the learning aspects of the user such as preferences, interest and knowledge level. The more accurate the analysis of the data related to the leaner, the better will be the solution or the method of teaching provided. Further, they can be grouped into various categories of learners depending upon the criteria like knowledge level, cognitive skills, learning styles, motivation, age, gender, goal and learning objective.

5 Conclusion and Future Work

Adaptive Learning is an enhancement on the traditional e-learning system to make e-learning more interactive and user centric. The main purpose of this paper is to study about the adaptive e-learning system, benefits and proposal of a conceptual model based on basic principle of adaptive learning. This paper covers the detailed introduction of the adaptive e-learning system along with its components. Depending upon the principles a conceptual model is also proposed. Universities are investing lots of capital in implementing e-learning system. So, this has become very crucial to evaluate the utilization of the LMS being implemented in the system. No doubt, BB is a well-known and efficient LMS but whether it has been exploited in its fullest or not which matters the most. Next step for the research is to evaluate the usage of the existing system. This will be done by taking survey from the users of the e-learning system. There are many open LMS are available in the market. In future work, a review of open LMS will also be done and most suitable will be chosen for the implementation based on proposed conceptual model.

Acknowledgement. This project was supported by the deanship of scientific research at Prince Sattam bin Abdulaziz, University Kingdom of Saudi Arabia under the research project # 2017/01/7978.

References

1. Khan, B.H.: Web-based instruction (WBI): what is it and why is it? In: Khan, B.H. (ed.) Web-Based Instruction, pp. 5–18. Educational Technology Publications, Englewood Cliffs (1997)
2. Brusilovsky, P.: Developing adaptive educational hypermedia systems: from design models to authoring tools. In: Murray, T., Blessing, S., Ainsworth, S. (eds.) Authoring Tools for Advanced Technology Learning Environments: Toward Cost-Effective Adaptive, Interactive, and Intelligent Educational Software. Ablex, Norwood (2002)
3. Lee, J., Park, O.: Adaptive instructional systems. In: Spector, J.M., Merrill, M.D., van Merrienboer, J., Driscoll, M.P. (eds.), Handbook of Research on Educational Communications and Technology, 3rd edn., pp. 469–484. Taylor Francis, New York (2008)
4. Fouad, K.M., Nagdy, N.M., Harb, H.M.: Adaptive E-learning system based on semantic search and recommendation in the Arab World. In: Information Systems Applications in the Arab Education Sector, pp. 254–283. IGI Global (2013)
5. Drissi, S., Amirat, A.: An adaptive E-learning system based on student's learning styles: an empirical study. Int. J. Distance Educ. Technol. (IJDET) **14**(3), 34–51 (2016)
6. Stash, N., De Bra, P.: Incorporating cognitive styles in AHA! (The Adaptive Hypermedia Architecture). In: Proceedings of the International 28 International Journal of Library and Information Science Conference Web-Based Education (IASTED), Innsbruck, Austria, pp. 378–383 (2004)
7. Bloom, B.S.: Mastery Learning: Theory and Practice, pp. 47–63. Holt, New York (1971)
8. Cronbach, L.J.: The two disciplines of scientific psychology. Am. Psychol. **12**(11), 671 (1957)
9. Bloom, B.S.: The 2 sigma problem: the search for methods of group instruction as effective as one-to-onetutoring. Educ. Res. **13**(6), 4–16 (1984)
10. Duong, T.H., Uddin, M.N., Li, D., Jo, G.S.: A collaborative ontology-based user profiles system. In: Nguyen, N.T., Kowalczyk, R., Chen, S.-M. (eds.) ICCCI 2009. LNCS (LNAI), vol. 5796, pp. 540–552. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04441-0_47
11. Akbulut, Y., Cardak, C.S.: Adaptive educational hypermedia accommodating learning styles: a content analysis of publications from 2000 to 2011. Comput. Educ. **58**(2), 835–842 (2012)
12. Sonwalkar, N.: The first adaptive MOOC: a case study on pedagogy framework and scalable cloud architecture—Part I. In: MOOCs Forum, vol. 1, no. P, pp. 22–29. Mary Ann Liebert, Inc., USA (2013)
13. Sonwalkar, N.: Adaptive individualization: the next generation of online education. Horizon **16**(1), 44–47 (2008)
14. Papanikolaou, K.A., Grigoriadou, M., Kornilakis, H., Magoulas, G.D.: Personalising the interaction in a web-based educational hypermedia system: the case of INSPIRE. User Model. User-Adap. Inter. **13**(3), 213–267 (2003). https://doi.org/10.1023/A:1024746731130
15. Trantafillou, E., Pomportsis, A., Georgiadou, E.: AES-CS: adaptive educational system based on cognitive styles. In: Proceedings of the Workshop on Adaptive Systems for Web-based Education, Held in Conjunction with AH 2002, Malaga, Spain (2002)

16. Bajraktarevic, N., Hall, W., Fullick, P.: ILASH: incorporating learning strategies in hypermedia. In: Proceedings of the Fourteenth Conference on Hypertext and Hypermedia, Nottingham, UK, 26–30 August 2003
17. Pashler, H., McDaniel, M., Rohrer, D., Bjork, R.: Learning styles: concepts and evidence. Psychol. Sci. Public Interest **9**, 105–119 (2008)
18. Brown, J.D.: Testing in language programs: a comprehensive guide to English language assessment (New edition) (2005)
19. Esichaikul, V., Lamnoi, S., Bechter, C.: Student modelling in adaptive e-learning systems
20. Paramythis, A., Loidl-Reisinger, S.: Adaptive learning environments and eLearning standards (2004)
21. Nguyen, L., Do, P.: Learner model in adaptive learning (2008)
22. Kim, K., Choi, Y.-J., Kim, M., Lee, J.-W., Park, D.-S., Moon, N.: Teaching-learning activity modeling based on data analysis. Symmetry **7**(1), 206–219 (2015)
23. Saberi, N., Montazer, G.A.: A new approach for learners' modeling in e-learning environment using LMS logs analysis. In: 2012 Third International Conference on E-Learning and E-Teaching (ICELET), pp. 25–33. IEEE (2012)
24. da Silva, D.P., Durm, R.V., Duval, E., Olivié, H.: Concepts and documents for adaptive educational hypermedia: a model and a prototype. In: Presented at the 2nd Workshop on Adaptive Hypertext and Hypermedia HYPERTEXT 1998, Pittsburgh, 20–24 June 1998
25. Cannataro, M., Cuzzocrea, A., Mastroianni, C., Ortale, R., Pugliese, A.: Modeling adaptive hypermedia with an object-oriented approach and XML. In: Presented at the 2nd International Workshop on Web Dynamics (WebDyn 2002) in Conjunction with the 11th International World Wide Web Conference (WWW 2002), Honolulu, Hawaii (2002)
26. Rasmussen, K.L., Davidson-Shivers, G.V.: Hypermedia and learning styles: can performance be influenced? J. Educ. Multimedia Hypermedia **7**(4), 291–308 (1998)
27. Snow, R., Farr, M.: Cognitive-conative-affective processes in aptitude, learning, and instruction: an introduction. Conative Affect. Process Anal. **3**, 1–10 (1987)
28. Tseng, J.C.R., Chu, H.-C., Hwang, G.-J., Tsai, C.-C.: Development of an adaptive learning system with two sources of personalization information. Comput. Educ. **51**(2), 776–786 (2008)
29. Honey, P., Mumford, A.: The Manual of Learning Styles. McGraw-Hill, Maidenhead (1982)
30. Dung, P.Q., Florea, A.M.: An approach for detecting learning styles in learning management systems based on learners' behaviours. In: International Conference on Education and Management Innovation, vol. 30, pp. 171–177 (2012)
31. Froschl, C.: User modeling and user profiling in adaptive e-learning systems. Master thesis, Graz, Austria (2005)
32. Becker, S.A., Cummins, M., Davis, A., Freeman, A., Hall, C.G., Ananthanarayanan, V.: NMC horizon report: 2017 higher education edition (2017)
33. Vanlehn, K.: The behavior of tutoring systems. Int. J. Artif. Intell. Educ. **16**(3), 227–265 (2006)
34. Schiaffino, S., Amandi, A.: Intelligent user profiling. In: Bramer, M. (ed.) Artificial Intelligence An International Perspective. LNCS (LNAI), vol. 5640, pp. 193–216. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-03226-4_11
35. Gena, C., Weibelzahl, S.: Usability engineering for the adaptive web. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The Adaptive Web. LNCS, vol. 4321, pp. 720–762. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72079-9_24

Evaluation and Improvement of Load Balancing Using Proposed Cuckoo Search in CloudSim

Deepak Garg(✉) and Pardeep Kumar

Department of Computer Science and Applications, Kurukshetra University, Kurukshetra, Haryana, India
erdeepakgarg21@gmail.com, pmittal@kuk.ac.in

Abstract. Cloud computing offers many services by permitting end users the usage of infrastructure (like networks, servers, storage), platform (like operating systems & middleware services), & softwares (like online gaming) delivered by cloud providers (like Amazon, Google) at short cost. Furthermore, Cloud Computing permits its users to operate resources in on-demand manner. It is based on the pay as you go pricing model that causes cloud users to pay corresponding to user's requirement only. But, demand of Cloud computing is increasing. So, Cloud Service Providers want to provide high Quality of Services (QoS) to cloud users but at same time they want to minimize cost. There are many challenges to achieve this aim. Optimum load balancing or load scheduling is one of technique which helps to achieve the Quality of Service requirements of cloud users which majorly includes Cloudlet Response Time, Cloudlet Makespan improvement. This paper consists of evaluation and improvement corresponding to these two parameters by utilizing and comparing Randomized scheduling, Round Robin scheduling, Shortest Job First Scheduling, Genetic Algorithm and last but most important Cuckoo Search With Levy Flight Algorithm scheduling results. Out of these five algorithms, genetic algorithm and cuckoo search are meta-heuristic where cuckoo search is considered to give global search due to levy flight and considered as most applicable in all application areas because of single parameter Pa. On the basis of comparisons a new Cuckoo Search has been proposed, implemented and showed better results in comparison to all above approaches. These experimental results have been attained with the help of CloudSim 3.0.3 toolkit.

Keywords: Cloud Computing · Cloudlet · CloudSim 3.0.3 toolkit · Cuckoo Search · Data center broker · Load balancing · Quality of Services

1 Introduction

Cloud computing is providing hosted services over internet as utilities on demand. Here pay as you go model is followed. It enables companies to use the resources such as virtual machines as a utility rather than needing to develop & maintain computing infrastructure in the house. According to definition by NIST [1], "*CC is a model for providing convenient, ubiquitous, on-demand admittance of the network to the shared pond of configurable computing assets (e.g., network, server, application, storage &*

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 343–358, 2019.
https://doi.org/10.1007/978-981-15-0108-1_32

other services) that can be quickly delivered and free with the minimal management exertion or with the minimal interaction of the service provider". The main role behind cloud computing is Virtualization that allows parallel execution of many tasks on the shared hardware platform [2].

In Cloud the load is in form of incoming tasks in the form of CloudLets from cloud users and resources are in form of VM (virtual machine). Virtual Machine is like abstract machine which executes on the PM (physical machine) [3]. Co-Located tasks do not interfere with the each other tasks while running on their virtual machines thus VM gives feeling of working in isolated environment to the user. This isolation among virtual machines is managed by Virtual Machine Manager or Hypervisor. It manages multiple different OS (operating systems) or several instances of the same operating system by assigning required no. of resources & gives an similar abstraction view of the basic hardware to VM in full type virtualization [1]. Virtual Machine runs in parallel to reduce response time of cloud user.

1.1 Our Contribution

In this paper, to improve the QoS requirements of cloud users the focus has been taken on Cloudlet Response Time, Cloudlet Makepsan Time. Corresponding to these time parameters, 5 various load balancing schemes have been evaluated and compared (on small scale and large scale) consisting of randomized scheduling, Round Robin, Shortest Job First, Genetic Algorithm, Cuckoo Search using Levy Flight scheduling. For implementing these algorithms, CloudSim 3.0.3 has been used and changes have been done in Data Centre Broker Class. Furthermore, a new Cuckoo Search has been proposed, implemented and it showed better results.

1.2 Organization

Organization of the remaining paper is given as—Sect. 2 talks about load balancing in the Cloud computing and Sect. 3 deals with CloudSim toolkit. Section 4 describes load balancing algorithms used in this paper with their theoretical pros and cons. Sections 5 and 6 represents experimental settings and experimental results comparison. Section 7 proposed modified Cuckoo Search. Finally Sect. 8 concluded & discussed the work that can be done in future.

2 Load Balancing or Scheduling

Load balancing or load scheduling is one of the major aims to achieve QOS in cloud computing. Load balancing is covers all approaches, letting an even dispersal of workloads among the presented Virtual Machines in the cloud. Or we can say that this comprises [3, 4] all the cloudlet scheduling or allocation approaches in cloud [4].

2.1 Load Balancing Need

L.B. is required to improve the overall time of that system in terms of Cloudlet Response time, Cloudlet Makespan. The improvement of Cloudlet response time will beneficiate cloud user as they will start receiving early results of sent cloudlet. While improvement of cloudlet makespan is to evenly use of VMs w.r.t cloudlet knowledge. Thus cloudlet will complete early on free or less busy virtual machine. This will be helpful for Cloud service Provider in terms of operational cost, CO2 emission [5–7].

2.2 Cloud and Its Components in Load Balancing

A Cloud consist of no. of Cloud datacenters, which are further partitioned into no. of nodes & every node have a no. of VMs [8–10]. Here, the incoming requests are allotted on the Virtual Machines. Figure 1 shows overview of the general architecture of cloud. VM is like a software program or a operating system that not only demonstrates functions of a separate computer system, but it is also capable to perform functions like running programs & applications as like individual computer. A VM, usually called as a guest which is made into an another computer environment called as “host.” Multiple VMs can be present in a single host.

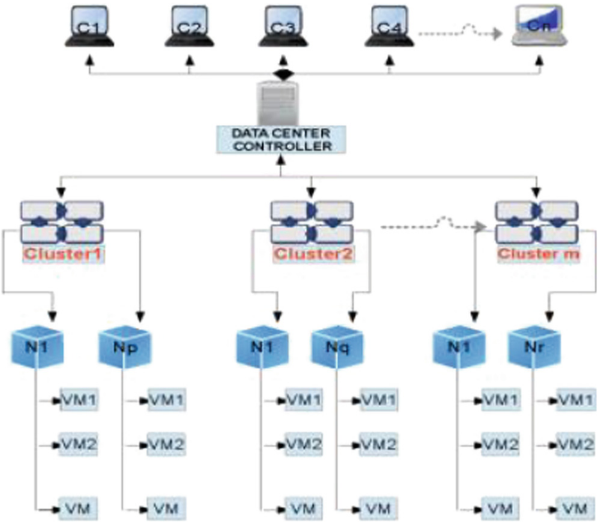


Fig. 1. Generalized architecture of cloud

2.3 Time Parameters Under Load Balancing

i. Cloudlet Response Time: It is required to be minimum. It is cloudlet starting execution time under the load balancing schedule. When there are more than one cloudlet then average cloudlet response time is evaluated as done in this paper.
Response Time of CloudLet i = RTCLi = STi (Starting Time of CLi)
Response Time of Schedule of Cloudlets = RTSCLs = Avg. (STi) for i = 1 to n CLs in that Schedule

ii. Cloudlet Makespan: Cloudlet Makespan is sum of time taken to process a cloudlet for its complete execution. When there are more than one cloudlet or talking about

a set of Schedule of Cloudlets then average of all cloudlets Makespan is evaluated (in that schedule) as done in this paper.
Makespan of Cloudlet i = MSCLi = CTi (Completion Time)-STi (Starting Time)
Makespan of Schedule of Cloudlets = MSSCLs = Avg. (CTi-STi) for i = 1 to n CLs in that Load.

3 CloudSim

There are many grid simulators [8–12] like GridSim, Sim-Grid & the GangSim, which are proficient for modelling and simulating the grid applications in the distributed situation but these fails to upkeep the infrastructure level & application level requirements required by the cloud computing environment [13]. Because Grid simulators and others are not able to segregate the multilayer services as required by cloud computing or grid simulators don't support for modelling of virtualization enabled resources and application management environment [13].

Cloudsim [8–11] is a modelling and simulation toolkit for cloud infrastructure. It is established on Grid-Sim toolkit [11] and provides provision of monetary determined resource administration, scheduling of cloudlets, supervision of bandwidth, management of cost and so on. The main feature that differentiates the infrastructure of Cloud Computing infrastructure as compared to the infrastructure of grid computing is huge organization of the virtualized infrastructure or the virtualization layer that acts as an execution, hosting & management atmosphere for application services which is infeasible for grids. Cloudsim is customizable thus helps researchers to develop researcher's new or modified procedures related to functioning of cloud by amending cloudsim base classes.

Current study targets at encompassing CloudSim by using the broker class because it acts as decision making procedure with the help of which a specific can search and choose the optimum link amid VM & incoming cloudlet.

Few Classes or entities of Cloudsim (which are relevant in this paper) are given below:

CloudUser: End user of cloudenvironment (who uses the cloud services).
DataCenter: It comprises collection of physical nodes or hosts. It acts like provider of IaaS.
Cloudlet: It is like container of tasks sent from clouduser to cloud. It have several attributes like clodlet ID, output file size, cloudlet length, required PE.
Virtual Machine: It is a multiuser shared resources that delivers varied facility to all.
HostMachine/Physical Machine: It is a physical node, existed in all data center. It comprises processing elements cores.
CloudletScheduler: It is a package in CloudSim toolkit. It tells about scheduling scheme between cloudlets and virtual machine. CloudletSchedulerTimeShared and CloudletSchedulerSpaceShare are types under it.
VMScheduler: It is a package in CloudSim toolkit. It tells about scheduling scheme between virtual machine and Host Machine. VMSchedulerTimeShared and VMSchedulerSpaceShared are types under it.

3.1 CloudletSchedulerTimeShared

It is a package in Cloudsim Toolkit. It is like parallel processing of more than one cloudlets in a single VM. Here VM is considered as resource and each unit time of resource is shared among allocated or binded Cloudlets. All allocated Cloudlets are started parallely under that VM. It causes improvement in Cloudlet response time.

3.2 CloudletSchedulerSpaceShared

It is also package in Cloudsim Toolkit. It is like Sequential processing of cloudlet in a single VM. Here VM is considered as resource and each unit time of resource is not shared among more than one allocated or binded Cloudlets. Next Allocated Cloudlet can only start after complete execution of current cloudlet under vm. It causes improvement in Cloudlet finishing time.

4 Load Balancing Algorithms

Cloud within cloud computing is network of the geographically scattered datacenters & datacenter consists of collection of huge no. of servers. When a cloud user gives a task (known as the cloudlet) then it is managed by the DCC (termed as data center controller) which utilizes a VM Load Balancer policy to search about which Virtual Machine should be designated to assign the incoming task request for processing. Below are few of existed load balancing algorithms [14–17].

4.1 Randomized

It involves random allocation of available VMs to incoming Cloudlets. Execution of Cloudlets are done after binding Cloudlets to the allocated VMs. So, it is type of Static Scheduling.

Pros: The best advantage of this algorithm lies within its simplicity and less time taken for selection of VM. Furthermore, it is also used to prove any dynamic approach as better one if that approach gives better performance as compared to it. *Cons:* It does not take benefit of knowledge of Cloudlet and virtual machine parameters like Cloudlets size, VM MIPS (Processing Power), PEs required. That's why, it does not evenly distributes load [18].

4.2 Round Robin

It maintains a list of Virtual Machines corresponding to VMs_IDs and maintains Cloudlets corresponding to CLs_ID. and then Firstly, it randomly picked a VM from the list of VMs, allocates first Cloudlet request to it. After completion of that cloudlet, it puts that VM to last within list of VMs and then next request are allocated in rotation basis [19].

Pros: Benefit of this approach is its simplicity and give even chance to each VM so even distribution corresponding to randomized schedule.

Cons: Its disadvantage is that it don't utilizes the processing speed of virtual machine and Cloudlets length which are different. So not good distribution of load [4].

4.3 Shortest Job First

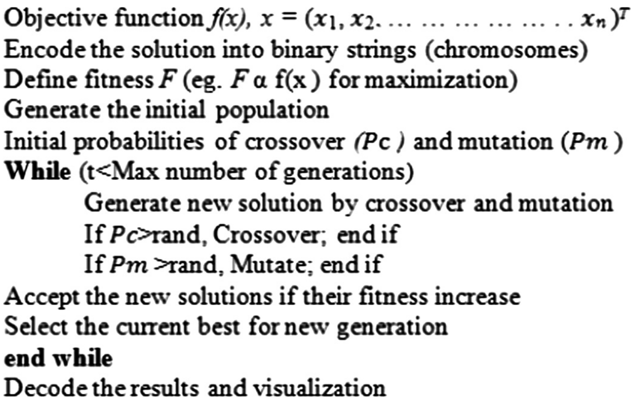
It maintains a list of Virtual Machines corresponding to VMs_IDs and maintains Cloudlets corresponding to Cloudlet Length Million of Instructions (MI) in increasing order. Then it allocates the Cloudlets in this order to the list of Virtual Machines on the basis of rotation. If Cloudlet length of all Cloudlets are same then maintain cloudlets list corresponding to CLs_ID [8].

Pros: By making use of cloudlet length knowledge there is improvement in makespan.
Cons: It don't consider Virtual machine processing power (MIPS). Furthermore, it suffers from starvation. Long cloudlet is waiting for a long time behind short cloudlet.

4.4 Genetic Algorithms

It is a stochastic global searching method & optimization based method. It is originated by Darwin's theory of evolution [17]. Steps are shown in Fig. 2. For load balancing task scheduling Wang et al. has done work on it as shown in below sub-sections [18]. Genetic Algorithm delivers multiple solutions rather than a single one.

Pros: Due to combination of exploration and exploitation features, it gives good results. Furthermore, it return multiple solution rather than a single one, so good for multimodal problem.
Cons: Its results are dependent upon no. of parameters like Mutation probability, crossover probability (Pm, Pc), N (no. of iterations), Population Size (s).



```
Objective function $f(x)$, $x = (x_1, x_2 \ldots \ldots \ldots \ldots \ldots x_n)^T$
Encode the solution into binary strings (chromosomes)
Define fitness $F$ (eg. $F \alpha$ f(x) for maximization)
Generate the initial population
Initial probabilities of crossover ($Pc$) and mutation ($Pm$)
While (t<Max number of generations)
        Generate new solution by crossover and mutation
        If $Pc$>rand, Crossover; end if
        If $Pm$ >rand, Mutate; end if
Accept the new solutions if their fitness increase
Select the current best for new generation
end while
Decode the results and visualization
```

Fig. 2. Working of Genetic Algorithm

4.5 Cuckoo Search with Levy Flight

CS algorithm is motivated from reproduction strategy by cuckoo birds. For mimicking the Cuckoo Search breeding behavior, Yang and Deb [6, 19, 20] use 3 rules given below.

i. Every cuckoo birds lays one egg at that time & put it in host bird nest chosen randomly. Cuckoo bird lays egg which aims to be as similar and good as of chosen host bird in terms of color, texture, weight;
ii. On the basis of survival of fittest, from each nest a good fit egg will be chosen either it be cuckoo egg or host bird egg;
iii. When Host bird came back he can find an alien egg in his nest with a discovery probability pa $\in$ [0, 1]. If discovered then host bird leave nest and build new one using local random walk else nest will carry over to the generation.

Steps corresponding to this algorithm is shown in Fig. 3. The algorithm comprises two salient equations, one is for the renewal of cuckoo egg (Eq. 1) and other is for reestablishment of host nest (Eq. 2). Equation 1 mimics global random walk while Eq. 2 mimics local random walk. These are given below:

$$x_i^{t+1} = x_i^t + \alpha \oplus \mathrm{Levy}(\beta); \qquad \alpha = \alpha 0(x_i^t - x_{best}^t) \tag{1}$$

$$x_i^{t+1} = x_i^t + \varepsilon \cdot (x_j^t - x_k^t) \tag{2}$$

Value of is from 1, 2, …, N. where N depicts total no. of nests; while xi, xj & xk represents some host nest. xbest best solution in that population at iteration t; α represents step size while value of α0 is 0.01; *Levy* (β) denotes random walk by Lévy distribution, ε denotes zoom factor whose value lies from 0 to 1.

Here the factor α0 is 0.01 which is assumed from the property that it should be hundredth of the length scale of the problem; but if it is not chosen like this then levy flight can become either less or high convergent.

Lévy Random or Levy (β)

There have been many for random walk but for true random walk Yang and Deb [19] generated walk from Mantegna's algorithm, whose important equations are given below:

$$s = u/|v|^{1/\beta} \tag{3}$$

$$u \sim N(0, \sigma_u), v \sim N(0, \sigma_v) \tag{4}$$

$$\sigma_u(\beta) = [(\Gamma(1 + \beta)\mathrm{Sin}(\pi \beta/2))/\beta\Gamma((1 + \beta)/2).2^{((\beta-1)/3)}]^{(1/\beta)}, \sigma_v = 1 \tag{5}$$

In these equations u and v are normal random vector with variance as given in Eq. 5. β or beta termed as distribution parameter value between 0.3 and 1.99 but in standard cuckoo search it has been taken as 3/2. Γ also termed as gamma function which is factorial of one value less than the input argument.

Cuckoo Search Algorithm (Pa)

```
begin
Input: N, pa
   Objective function Minimize f(X), x= (x1, x2,….xp)^T
   Xi (i=1, 2,….N)is the initial population of N nests.
   And evaluate its fitness fi
   Gbest=min(f1,f2….fN)
   While (t <MaxIterations of stop criterion)
      Get a host nest randomly say i with fitness fi
      Generate a cuckoo egg corresponding to choosen host nest using eq. 1 (global
      random walk)
      Evaluate fitness of cuckoo egg say Fi
      If (Fi<fi)
               Replace xi by generated cuckoo egg.
      Endif
      If (rand<pa=0.25)// discovering probability is greater
               That nest is abandoned
               Built new nest using eq. 2 (local random walk) accept it & Evaluate its
               fitness Fk.
      Endif
      G*=min(f1,f2….fN)
      If (G* < Gbest)
      Update Gbest, Xbest
      Endif
   Endwhile
   Output:Gbest, Xbest
   end
```

Pros: After deep literature survey it has been find out that CS is superior to other metaheuristic Furthermore, many other researchers also concluded the same [21–27] with following conclusions:

- Levy Flights is supposed to give infinite mean and variance. Thus, CS obeys global random walk traversing search space in better way as compared to other walk.
- CS has only one parameter Pa. so, very less parameters as compared to other Algorithms means it is generally applicable.
- It return multiple solution rather than a single one, so can be used in multimodal search problems.
- If no. of local solutions are less than No. of nest in CS than CS can uncover all optimal solutions.

Cons: Step Length is substantial Tailed in CS & thus any Biggest Step is Probable which is not good for convergence.

5 Proposed Cuckoo Search

After implementation and comparisons of above existed approaches in Sect. 7. It has been found out that with just having only one parameter i.e. pa, cuckoo search has capability to give good results furthermore it is most applicable in all optimization problems as compared to every other approach. But as step length is substantial tailed and any largest step is thinkable so slow convergence is big issue in this, which has been improved in below proposed approach named as hybrid cuckoo search. Within this three modification have been done. Rather than picking host nest randomly (for local random walk) picked host nest with worst fitness. Pa has been set self adaptive i.e. high in start and decreases as iteration increases mimicking theory of evolution. Rather than accepting any cuckoo egg nest from Eq. 2 (global random walk) directly. Evaluated worst cuckoo egg nest again and replaced or abandon corresponding to it. Steps of proposed algorithm is given below.

Proposed Cuckoo Search Algorithm (Self adaptive Pa)

```
begin
Input: N, pa
    Objective function Minimize f(X), x= (x1, x2,….xp)$^T$
    Xi (i=1, 2,….N)is the initial population of N nests.
    And evaluate its fitness fi
    Gbest=min(f1,f2….fN)
    While (t <MaxIterations of stop criterion)
        Get a host nest with Worst Fitness say i with fitness fi= Gworst=max(f1,f2….fN)
        Generate a cuckoo egg corresponding to choosen host nest using eq. 1 (global
        random walk)
        Evaluate fitness of cuckoo egg say Fi
        If (Fi<fi)
                Replace xi by generated cuckoo egg.
        Endif
        If (rand<$\frac{tmax-t}{tmax}$ ×pa ) //Self Adaptive Discovery Probability
                That nest is abandoned
                Built new nest using eq. 2 (local random walk) accept it & Evaluate its
                fitness Fk.
                Find worst nest again say w with fitness
                fw = Gworst=max(f1,f2….fN)
                If (Fk<fw)
                        abandon Xw & replace it by Xk
                else
                        reject Xk and accept Xw.
                Endif
        Endif
        G*=min(f1,f2….fN)
        If (G* < Gbest)
        Update Gbest, Xbest
        Endif
    Endwhile
    Output:Gbest, Xbest
    end
```

6 Experimental Settings

In this paper, Algorithms Evaluation & comparisons has been done with the support of Cloudsim 3.0.3 toolkit. Corresponding to space constraint 6 Cloudlets & 2 Virtual Machines has been taken. Two processing elements has been taken At infrastructure level while for every Virtual Machines and CL (Cloudlets) processing element is taken as one. MIPS of first PE is 1500 MIPS while 2500 MIPS is of second one. No. of schedules or Population size is taken as four.

Table 1 tabulates 6 cloudlets, their size in MIPS and ids.

Table 2 depicts 2 VMs and its processing capabilities in MIPS.

Table 1. Properties of cloudlets

Cloudlet id CL-id	Length (Millions of Instruction)
0	2127
1	1793
2	2825
3	1681
4	2114
5	2486

Table 2. Processing capability of VMs

VM-id	Millions instruction per second (MIPS)
0	1500
1	2500

Table 3 depicts a mockup schedule from random scheduling algorithm as an example & below this table code for random scheduling algorithm has been given.

Table 3. Load after randomized policy

POP [schedule, CL]	CL0	CL1	CL2	CL3	CL4	CL5
0	1	0	1	0	0	1
1	0	1	1	0	0	1
2	1	1	1	1	0	1
3	0	0	1	1	1	1

Randomize Scheduling Code:

```
c-No. of Cloudlets=6
v-No. of VMs=2
s-No. of Schedules=4
int pop[ ][ ]=new int[s][c];
Random robj1=new Random();
for(int k=0;k<s;k++)
      {
         for(int l=0;l<c;l++)
                {
                    pop[k][l]=robj1.nextInt(v+1);
                 }
       }
```

In above, robj1.nextInt(v+1) function outputs a uniform distributed pseudo random value 0 (inclusive) to v+1 (exclusive).

To implement scheduling for Shortest Job Burst changes has been done in submitCloudlets() function, existing in the DatacentreBroker. Java file. Below is the sample code for binding.

Sample Code for Binding:

```
DatacenterBroker broker = createBroker();
int z=0;//z represents cloudlet and kk represents schedule
broker.bindCloudletToVm(cloudletList.get(0).getCloudletId()
, pop[kk][z]);z++;
broker.bindCloudletToVm(cloudletList.get(1).getCloudletId()
, pop[kk][z]);z++;
broker.bindCloudletToVm(cloudletList.get(2).getCloudletId()
, pop[kk][z]);z++;
broker.bindCloudletToVm(cloudletList.get(3).getCloudletId()
, pop[kk][z]);z++;
broker.bindCloudletToVm(cloudletList.get(4).getCloudletId()
, pop[kk][z]);z++;
broker.bindCloudletToVm(cloudletList.get(5).getCloudletId()
, pop[kk][z]);
```

7 Experimental Results

Experimental results have been obtained for 2 criterias. First is Average Makespan of Load of Cloudlets and second is Average Response time of Cloudlets. These criterias are witnessed on 5 existing algorithms (Randomized Approach, Round Robin approach, Shortest Job First, Genetic Algorithm, CS with Levy Flight) and 1 proposed algorithm (Accelerated Cuckoo Search) in 2 categories to validate results. First category is under CloudletSchedulerSpaceShared and second category is CloudletSchedulerTimeShared.

To prove results in population based algorithms, average has been taken for the time parameters in y axis and x axis depicts increase in iterations for load scheduling.

Thus X axis and Y axis is as below:

> X axis: Increase in No. of Iterations; Y axis: Population average Response or Maksespan Time; MSSCL-Denotes Makespan time of schedules of Cloudlets; RTSCL-Denotes Response time of schedules of cloudlets.

7.1 Cloudlet_Scheduler_Space_Shared_Mode

i. For Randomized vs Round Robin vs Shortest Job First vs Genetic Algorithm vs Cuckoo Search vs Proposed Cuckoo Search corresponding to MSSCL

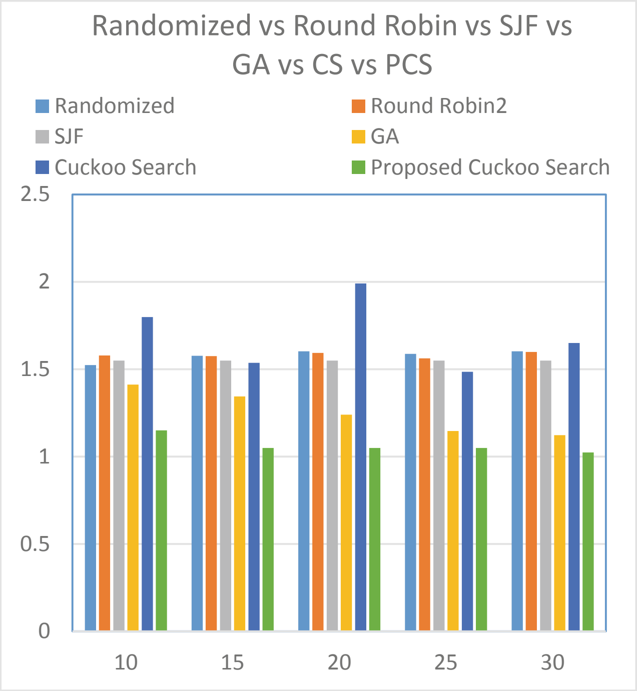


Fig. 3. Comparison of all six algorithms (five existed and one proposed) corresponding to Avg. MSSCL (X axis - Iterations of Load Scheduling, Y axis-Avg. of MakespanTime of four schedules of Cloudlets as population or nests after n iterations)

ii. For Randomized vs Round Robin vs Shortest Job First vs Genetic Algorithm vs Cuckoo Search vs Proposed Cuckoo Search corresponding to RTSCL (Fig. 4).

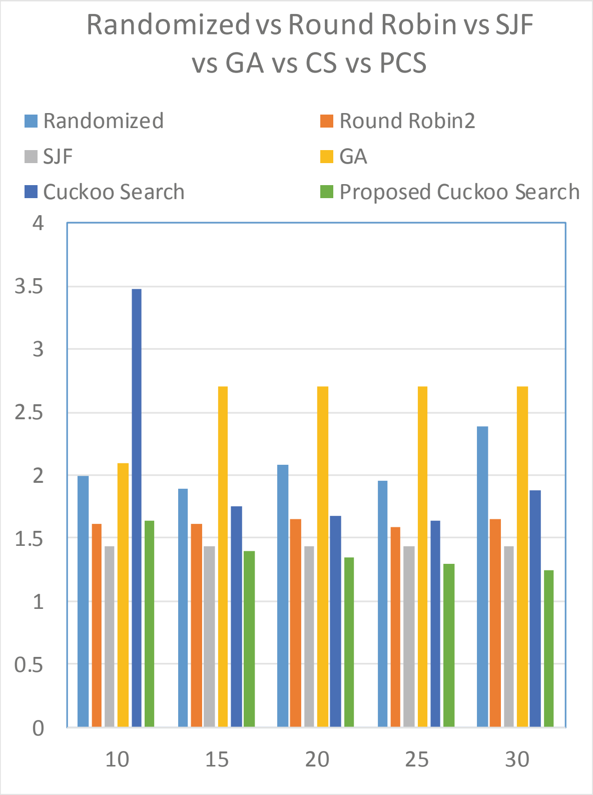


Fig. 4. Comparison of all six algorithms (five existed and one proposed) corresponding to Avg. RTSSCL (X axis: Iterations of Load Scheduling, Y axis-Denotes Average Response Time of the four schedules of Cloudlets)

7.2 CloudletSchedulerTimeShared Mode

In this category as Cloudlet Scheduling policy is time shared, so VM is shared parallaly between all allocated cloudlets. Thus all cloudlets will start executing at starting point of time (by default 0.1). That's why in all below scheduling policies Average response time of schedule of cloudlets is 0.09999. Which is very small thus good for cloud users (Fig. 5).

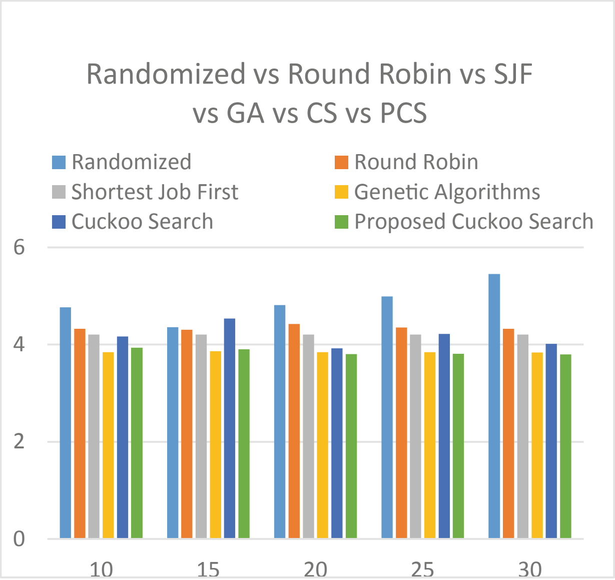


Fig. 5. Comparison of all six algorithms (five existed and one proposed) corresponding to Avg. MSSCL (X axis-Iterations of the Load Balancing, Y axis-Average. of MakespanTime of four schedules of Cloudlets as population or nests after n iterations)

8 Conclusion

In the paper, load balancing under cloud computing is studied including need of load balancing, components under load balancing, time parameters as objective, Cloudsim as tool and five load balancing approaches have been studied. To improve Quality of Service requirements of cloud users which majorly includes Cloudlet Response Time, Cloudlet Makespan improvement evaluation and improvement corresponding to these two parameters has been done by utilizing and comparing Randomized scheduling, Round Robin scheduling, Shortest Job First Scheduling, Genetic Algorithm and last but most important Cuckoo Search With Levy Flight Algorithm scheduling results. Out of these five algorithms, genetic algorithm & cuckoo search are meta-heuristic where cuckoo search is considered to give global search due to levy flight and considered as most applicable in all application areas because of single parameter Pa. On the basis of comparisons proposed a new Cuckoo Search by introducing three changes. First change is rather than picking host nest randomly (for local random walk) picked host nest with worst fitness, second is that Pa has been set to self adaptive corresponding to increase in iterations (i.e. it is high in starting supporting exploration and low in ending supporting exploitation supporting theory of evolution) & third says that on discovery rather than accepting any host bird egg nest directly compare it with worst cuckoo egg nest again and thus replace or abandon accordingly). Implementation showed better results corresponding to proposed cuckoo search as comparison to all other approaches in both space shared and time shared cloudlet scheduler policy w.r.t make span time and response time. All of these experimental have been done by utilizing CloudSim 3.0.3 toolkit after enhancing few base classes.

References

1. Shaw, S., Singh, A.: A survey on scheduling and load balancing techniques in cloud computing environment. In: International Conference on Computer & Communication Technology (ICCCT), pp. 87–95. IEEE (2014)
2. Gkatzikis, L., Koutsopoulos, I.: Migrate or not? Exploiting dynamic task migration in mobile cloud computing systems. IEEE Wirel. Commun. **20**, 24–32 (2013)
3. Achar, R.: Load balancing in cloud based on live migration of virtual machines. In: Annual IEEE India Conference, pp. 85–92 (2013)
4. Ren, X., Lin, R., Zua, H.: A dynamic load balancing strategy for cloud computing platform based on exponential smoothing forecast. In: Proceedings of IEEE CCIS, pp. 220–224 (2011)
5. A particle of dynamic network load balancing cluster [EB/OL]. http://www.linuxaid.com.en/articles/1/4/14251644.shtml
6. Banerjee, S., Adhikari, M., Kar, S., Biswas, U.: Development and analysis of a new cloudlet allocation strategy for QoS improvement in cloud. Arab J. Sci. Eng. **40**(1), 1409–1425 (2015)
7. Amalarethinam, D.I.G., MalaiSelvi, F.K.: A minimum makespan grid workflow scheduling algorithm. In: International conference on Computer Communications and informatics, pp. 1–6 (2012)
8. Abudhagir, S., Shanmugavel, S.: A novel dynamic reliability optimized resource scheduling algorithm for grid computing system. Arab J. Sci. Eng. **39**(1), 7087–7096 (2014)
9. Armbrust, M., et al.: A berkeley view of cloud computing. Technical report No. UCB/EECS-2009-28, pp. 1–23. University of California at Berkley, USA (2009)
10. Aymerich, M., Enul, G., Surcis, S.: An approach to a cloud computing network. In: IEEE, vol. 113, no. 1, pp. 113–118 (2008)
11. Buvya, R., Ranjan, R., Calheiros, R.N.: Modeling and simulation of scalable cloud computing environment and the CloudSim toolkit-challenges and opportunities. In: Proceedings of the 7th High Performance Computing and Simulation Conference, Germany, pp. 1–11 (2009)
12. Bhatia, W., Buvya, R., Ranjan, R.: CloudAnalyst-a Cloudsim based visual modeller for analysing cloud computing environments and applications. In: 24th IEEE International Conference on Advanced Information Networking and Applications, pp. 446–452 (2010)
13. Mohammad, M.K., Analoui, M.: Resource scheduling in desk top grid by grid JQA. In: 3rd International conference on Grid and Pervasive Computing, pp. 63–68. IEEE (2008)
14. Wickremasinghe, B., Calheiros, R., Buyya, R.: CloudAnalyst-a CloudSim-based visual modeller. In: International Conference on Analysing Cloud Computing Environments and Applications, pp. 446–452 (2010)
15. Mohialdeen, I.A.: Comparative study of scheduling algorithms in cloud computing environment. In: International conference on Challenges in Cloud Computing, pp. 252–263 (2013)
16. James, J., Verma, B.: Efficient VM load balancing algorithm for cloud computing environment. Int. J. Comput. Sci. Eng. **4**(9), 1658–1663 (2012)

17. Deepan Babu, P., Amudha, T.: A novel genetic algorithm for effective job scheduling in grid environment. In: Krishnan, G.S.S., Anitha, R., Lekshmi, R.S., Kumar, M.S., Bonato, A., Graña, M. (eds.) Computational Intelligence, Cyber Security and Computational Models. AISC, vol. 246, pp. 385–393. Springer, New Delhi (2014). https://doi.org/10.1007/978-81-322-1680-3_42
18. Garg, D., Garg, P.: Basis path testing using SGA & HGA with ExLB fitness function. Procedia Comput. Sci. **70**, 593–602 (2015)
19. Yang, X.S., Deb, S.: Cuckoo search via Lévy flights. In: IEEE Conference Publication World Congress on Nature & Biologically Inspired Computing (NaBIC), pp. 210–214 (2009)
20. Guo, Q., Gao, Y., Cui, L., Zhang, J.: Cuckoo search algorithm based on three random walks. In: 3rd IEEE International Conference on Computer and Communications, pp. 2180–2186 (2017)
21. Yang, X.S., Karamanoglu, M.: Multi-objective flower algorithm for optimization. In: International Conference on Computational Science, pp. 861–868. Elsevier Science (2013)
22. Yang, X.S., Deb, S.: Cuckoo search-recent advances and applications. Neural Comput. Appl. **24**(1), 169–174 (2014)
23. Gandomi, A.H., Yang, X.S., Alavi, A.H.: Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems. Eng. Comput. **29**(1), 17–35 (2013)
24. Shakya, A.ku., Garg, D., Nayak, P.Ch.: Hybrid live VM migration: an efficient live VM migration approach in cloud computing. In: Luhach, A.K., Singh, D., Hsiung, P.-A., Hawari, K.B.G., Lingras, P., Singh, P.K. (eds.) ICAICR 2018. CCIS, vol. 955, pp. 600–611. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-3140-4_54
25. Garg, D., Kumar, P.: A survey on metaheuristic approaches and its evaluation for load balancing in cloud computing. In: Luhach, A.K., Singh, D., Hsiung, P.-A., Hawari, K.B.G., Lingras, P., Singh, P.K. (eds.) ICAICR 2018. CCIS, vol. 955, pp. 585–599. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-3140-4_53
26. Nayak, P.Ch., Garg, D., Shakya, A.Ku., Saini, P.: A research paper of existing live VM migration and a hybrid VM migration approach in cloud computing. In: IEEE 2nd International Conference on Trends in Electronics and Informatics (ICOEI 2018), pp. 721–726 (2018)
27. Harkawat, A., Kumari, S., Pharkya, P., Garg, D.: Load balancing task scheduling based on variants of genetic algorithms: review paper. In: Kaushik, S., Gupta, D., Kharb, L., Chahal, D. (eds.) ICICCT 2017. CCIS, vol. 750, pp. 318–325. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-6544-6_29

Intelligent System to Classify Peanuts Varieties Using K-Nearest Neighbors (K-NN) and Support Vector Machine (SVM)

V. G. Narendra and K. Govardhan Hegde(✉)

Computer Science and Engineering, Manipal Institute of Technology,
Manipal Academy of Higher Education, Manipal 576104, India
{narendra.vg, govardhan.hedge}@manipal.edu

Abstract. In the world, India is the second biggest producer of peanuts or groundnuts, and it is also our country's major oilseed crop. In India, the existing peanuts crop varieties are GAUG-1, Kuber, Amber, PG-1, BG-1, T-64, GAUG-10, BG-2, Chandra, Kadri-2, Chitra, Kadri-3, Prakash, T-28, Kaushal, etc. Presently, peanuts are having only 75–80% of India's average market value. Because, the peanuts kernel quality assessment, as well as identifying varieties, are done manually by skilled labors, which leads costly. In this research, an affordable method is proposed to assess the peanuts kernel quality and identifying the different varieties quickly with undamaged, repeatability with low cost, and accurately with high distinguishing rate. Also, to meet the quality of peanuts kernel as per the international market standards and to increase the income of the former. The proposed system relies on computer vision and machine learning. The obtained overall accuracies were K-nearest neighbors (93.33%) and Support vector machine (93.82%). These percentages are discriminating peanuts variety as the best predictive model.

Keywords: Peanuts · K-Nearest Neighbors · Support vector machine · Computer vision · Image processing

1 Introduction

In current years, agricultural production played an in crucial role and became a significant area to meet the endless demands (food, raw industrial materials, and energy) by humans. To increase the volume in the area of agricultural production, the investors have enforced to gather and incorporate information and knowledge from various sources. The experts in the field of the agricultural output take a significant role in providing information and decisions making. But, when the discussion is needed, an existing agrarian expert is not available. Hence, a robust control expert intelligent system has to promise to solve the difficulties faced in agricultural production [5].

Today, peanuts is crucial oilseed and the leading food crop, and also called the poor man's nut. South America is the native of the peanut plant. The Arachis hypogaea Linn. is the botanical name for groundnut and is derived from two Greek words: Arachis (a legume) and hypogaea (below ground mentioning to the pods' information in the soil). Peanut is a flat annual plant and is distributive in the sub-tropical, tropical,

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 359–368, 2019.
https://doi.org/10.1007/978-981-15-0108-1_33

and warm temperature regions. Peanuts are grown between the latitudes 40 °N and 40 °S [15].

Presently, the groundnut oil is using traditionally as a dietary component as well as a cooking medium in a few countries of the world. The oil processing industries from groundnuts occurs in many countries. These countries are Nigeria, Gambia, India, Senegal, and Sudan; processed into different products such as Vanaspati and peanut cake.

Peanut is the major crop for India farmers to increase their agricultural income. It is mainly due to the sparse peanuts kernel's quality. Because the automation level of testing peanuts kernel's quality done by workers manually. A load of workers is so high. It requires them to have abundant testing experience [7]. The testing of such quality based on image processing is very innovative along with computer vision. It is a low cost, the high distinguishing rate, which used in a batch test. Studies shown on by several researchers on the wheat [4], maize [22], paddy rice [19], testing of peanut kernels' quality [7] etc. There is still exists scope to undergo similar experiments along with other factor like variety.

A new one testing method is based on computer vision and image processing, are used. It is speedy and undamaged, repeatability with fatigue. It has high distinguishing rate, with low cost. It can be used in the batch test. It helps in evaluating of peanut kernels' quality. Many related studies are in similar field like wheat [4], paddy rice [19], and maize [22] etc. shown good results.

Grain seed analysis is very important. The five corn varieties classification is proposed by a method based on CV [2]. A flat scanner is used for image acquisition for non-touching corn kernels. From each corn images, the different types of features were extracted. The stepwise discriminant analysis (DA) used to obtain an optimized subset of the features. A grouping of BPNN and discriminant analysis is used for this purpose. This method is applied for mainly variety detection. The five rice variety was identified by using color and morphological features [16]. The multilayer perception and neuro-fuzzy artificial neural networks method reported an average accuracy score is about 99% for rice-kernel classification. Zapotoczny [21] investigated discriminating eleven varieties of wheat grain. The texture features are extracted using mass surface color images, and feature reduction methods are implemented to establish a set of features which are optimized. The different classifiers are also evaluated and noted their performances. The classification of rapeseed varieties being carried out. It was using a (NIRS) is reported in Zou et al. [23]. They performed a PCA for feature extraction. BPNN and DA are used to discriminate 5 variabilities of rapeseeds. They attained a classification accuracy of 100%. Similarly, the classification of three rapeseed varieties by using (FTIR-PAS). Based on FTIR-PAS measurements, support vector machines, and partial least squares discriminant analysis are used to develop classification models [10] and [11]. The classification accuracies were in between 90% to 100% reported by many researchers. In spite of spectrophotometric techniques robustness, these methods involve costly tools and skilled people to operate them. It is needed to develop cost-effective as well as a convenient system for peanuts classification using image processing as well as computer vision. In our study, a regular color-imaging approach is used to classify peanut verities. The classification process is based on machine learning. It is being carried out in 2 phases training and testing using the train and test datasets. The predictive model is used to prevent overfitting risk it uses K-fold cross-validation for classification.

The main objective of this study, we have used image processing, computer vision, and machine learning. We have used color and texture features to classify peanut varieties.

2 Materials and Methods

2.1 Samples of Peanuts

In this study, the Seven different varieties (Chitra, Chandra, Kaushal, GAUG-1, Amber, Kuber, and Prakash) of peanuts were used. During the 2017–18 growing season, all the aforesaid peanut varieties were acquired from the University of Agricultural Sciences. The place is in Dharwad, Karnataka, India. The samples were cleaned in an air-screen. From these cleaned samples, all external matter like stones, chaff, dirt, dust, broken and immature peanuts are removed. Then, peanuts were kept in plastic bags under room temperature between 25 °C to 29 °C.

2.2 Acquisition of a Peanuts Kernel's Image

To conduct identification and classification of peanuts varieties, the mass surface imaging method is selected in this research. It requires less knowledge and finally, the proposed model can be used by any person. Hence, under laboratory conditions, the reasonable image acquisition system was consists of a flatbed scanner (charge-coupled-device Scanjet 3770 of 24-bit color with 1200 × 1200 dots-per-inch) and a Hewlett-Packard Computer as a personal computer. A 24 × 24 mm square frame was used to acquire a uniform peanuts kernels' image scene. The peanuts kernels' were spread through the mass surface in a square frame (10 mm to 15 mm), to avoid scanned light passing over the mass-layer. The scanner glass is thoroughly cleaned before scanning of each sample. Thus for each peanut sample was obtained with an image resolution of 1600 × 1600 pixels and saved in.jpg file format. A total of 210 peanuts kernels' image were obtained by scanning (shown in Fig. 1) individual variety of 30 each (shown in Fig. 2). The MATLAB software (R2016a) is used for the image processing program, which involves the image pre-processing such as noise removal using median filtering, morphology operating, segmentation, and the conversion of the color space of samples image [6].

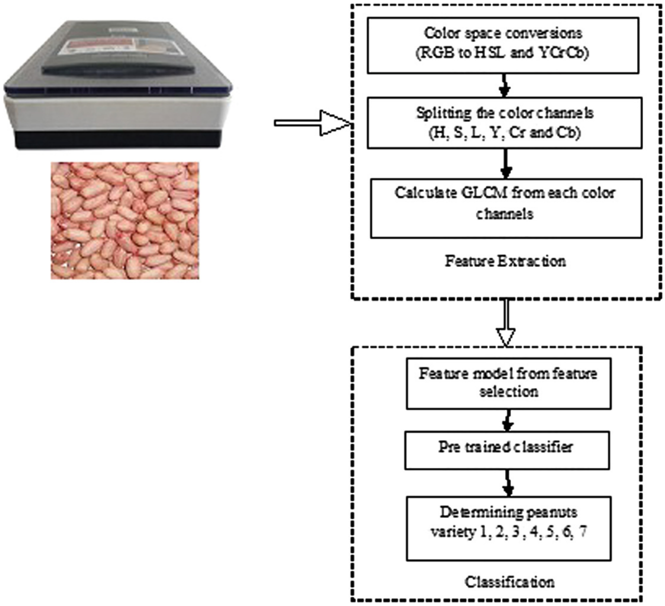

Color space conversions (RGB to HSL and YCrCb)
Splitting the color channels (H, S, L, Y, Cr and Cb)
Calculate GLCM from each color channels
Feature Extraction
Feature model from feature selection
Pre trained classifier
Determining peanuts variety 1, 2, 3, 4, 5, 6, 7
Classification

Fig. 1. The proposed system diagram for image collecting of different peanuts variety and classification.

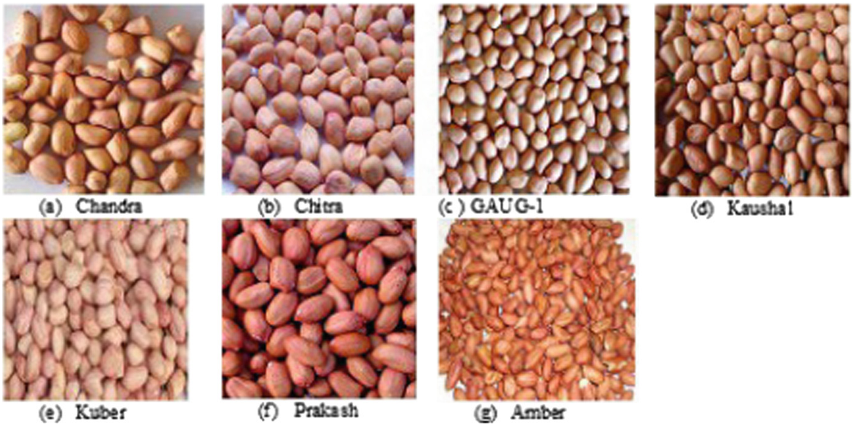


Fig. 2. The different varieties of peanut image.

2.3 Feature Extraction

From various studies [1, 18] and [20], have used computer vision, and image processing for the extraction of color texture features has great advantage as it calculates at same time. In this study, from each image, the extraction of features is conducted by computing the Gray-level co-occurrence matrix-based texture features. It uses different color models such as HSL and YCrCb. To achieve this, initially from each acquired image are converted from RGB to two color models: HSL and YCrCb, and decomposed into individual color components. The separated six different color channels (Y, Cb, Cr, H, S, and L) are obtained from each image [5]. A total of 72 features (6 color channels multiplied with 12 GLCM features) is obtained from each color component. The particulars of these features are described in the next section. Before applying to a machine-learning algorithm, every feature was normalized between 0 and 1.

2.3.1 GLCM Method

It is a pure statistical approach. It can help in identifying and analyzing image texture [1, 18] and [20]. We get different co-occurrence probabilities. The co-occurrence probabilities are calculated by:

$$\Pr(x) = \{P(i,j)|(d,\vartheta)\} \tag{1}$$

$$P(i,j) = \frac{P_{ij}}{\sum_{i,j=1}^{G} P_{ij}} \tag{2}$$

where the Prob(x) gives the measure of probability [3]. There are 12 texture-features [8] are calculated as follows

$$\mathrm{Contrast} = \sum\nolimits_{k,l=0}^{M_g-1} P_{k,l} \quad |k-l|^2 \tag{3}$$

$$\mathrm{Correlation} = \sum\nolimits_{k,l=0}^{M_g-1} \left[\frac{(k-\mu_k) \quad (l-\mu_l)}{\sqrt{(\sigma_k^2) \quad (\sigma_l^2)}} \right] \tag{4}$$

$$\mathrm{Angular\ Second\ Moment} = \sum\nolimits_k \sum\nolimits_l p(k,l)^2 \tag{5}$$

$$\mathrm{Energy} = \sqrt{\mathrm{Angular\ Second\ Moment}} \tag{6}$$

$$\mathrm{Homogeneity} = \sum\nolimits_{k,l=0}^{M_g-1} \frac{P_{k,l}}{1+(k-l)^2} \tag{7}$$

$$\mathrm{Dissimilarity} = \sum\nolimits_{k,l=0}^{M_g-1} P_{k,l}|k-l| \tag{8}$$

$$\mathrm{Entropy} = \sum\nolimits_{k,l=0}^{M_g-1} P_{k,l}(-\ln P_{k,l}) \tag{9}$$

$$\mathrm{Cluster\ \ Shade} = \sum\nolimits_{k,l=0}^{M_g-1} ((k-\mu_k)+(l-\mu_l))^3 P_{k,l} \tag{10}$$

$$\mathrm{Cluster\ \ Performance} = \sum\nolimits_{k,l=0}^{M_g-1} ((k-\mu_k)+(l-\mu_l))^4 P_{k,l} \tag{11}$$

$$\text{Smoothness} = 1 - \frac{1}{(1 + \sigma^2)} \tag{12}$$

$$\text{Third Movement} = \sum\nolimits_{k,l}^{M_g - 1} \left(P_{k,l}\right)(k - \mu_l)^3 \tag{13}$$

$$\text{Maximum Probability} = \max_{k,l}\left(P_{k,l}\right) \tag{14}$$

2.4 Recursive Feature Extraction (RFE) and Feature Model

In this research, the RFE is used as a supervised feature elimination method [24] constructed on ranking feture. In the RFE, the removal of feature begins with a whole feature set, and recursively features are selected, while considering reduced sets of features by using a SVM as a central classier [17, 25]. In this study, the cross-validation approach generated as third subclass which helps in classifying peanut varities.

2.5 Classification

An optimum classifier is directly proportional to the type of the dataset. Its capability to understand interactions between the features in pattern classification tasks. In general, K-NN and SVM classifiers increases the overall performances to classify peanut varieties. Additionally 10-fold cross-validation is implemented in this study.

2.5.1 K-Nearest Neighbors (K-NN)

Here, K is a extremely data-dependent. It is also a tuning parameter, which represents the number of neighbors. In this work, uniform and distance weight function were used [17]. The K values were used as 3, 5 and 7. Table 1 depicts the same.

Table 1. K-NN classifier used in our study.

Parameter	Possible entries
K	7, 5, 3
Weight function	Distance, Uniform
Algorithm	Brute force

2.5.2 Support Vector Machine (SVM)

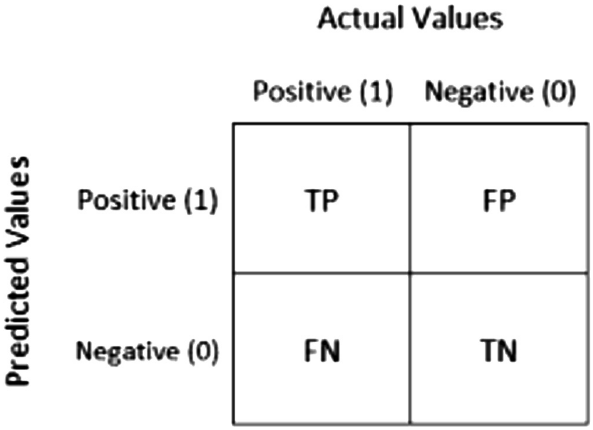
It's tries to find suitable boundaries between distinct classes [9]. In this work, the SVM is trained with seven training sets, which represents a seven variety of peanuts. The "one-against-one" approach is used to discriminate the seven different varieties of peanuts and it is used to translate binary classification to multiclass-learning. The constraints are revealed in Table 2 to discover an actual classifier for peanut classification.

Table 2. Tuned parameters of the SVM used in this study.

Parameter	Possible entries
Regularization parameter C	1000, 100, 10, 1
Type of Kernel function	Radial basis, Polynomial, Linear
Kernel coefficient c (for radial basis and polynomial)	0.0001, 0.001
Polynomial kernel function d	1, 2, 3

3 Results and Discussion

Confusion matrices are used in this study to evaluate performances of classifiers for evaluating multiclass predictive models. Typical confusion matrix used for our study shown in Fig. 3.



		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Fig. 3. Confusion matrix.

Performance formulation parameters of confusion matrix are shown in Eqs. (15)–(18).

$$\text{Accuracy (for a certain class)} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{15}$$

$$\text{Overall Accuracy} = \frac{\sum \text{TP} + \sum \text{TN}}{\sum \text{TP} + \sum \text{TN} + \sum \text{FP} + \sum \text{FN}} \tag{16}$$

$$\text{Omission} = \frac{\text{FP}}{\text{TP} + \text{FP}} \tag{17}$$

$$\text{Comission} = \frac{\text{FN}}{\text{TP} + \text{FP}} \tag{18}$$

Table 3 gives Confusion matrix which illustrates experimental results carried out by K-NN classifier. We got only seven misclassifications in discriminating peanuts varieties. First, among 30 samples of Chandra variety, corrected samples were only 27, but Chitra and Kaushal are misclassified respectively 2 and 1. Finally, the accuracy of Chandra variety was 90%. The second wrongly classified Chitra variety is GAUG-1 with accuracy of 96.67%. Third, among 30 samples of GAUG-1 variety, 28 samples were properly classified, one sample of each was misclassified as Chitra, and Kaushal respectively. As there exists commission error, the accuracy score for GAUG-1 variety being 93.33%. The fourth misclassification is accounted for Kaushal variety. Here, it misclassified one sample as GAUG-1. This variety's accuracy score was 96.67%. Fifth, KNN classifier's accuracy score for discriminating Kuber variety was 93.67%. Sixth, among 30 samples of Prakash variety, only 28 samples are classified correctly, whereas two samples were misclassified as Amber. The obtained accuracy score for discriminating Prakash variety was 93.33%. Seventh, among 30 samples of Amber variety, 27 samples were classified correctly where as one sample misclassified as GAUG-1, one sample was incorrectly classified as Kaushal, and one sample was incorrectly classified as Prakash. SVM classifier's accuracy score for discriminating Amber variety was 90.00%. The overall classification accuracy rate of KNN classifier is found to be 93.33%. Similarly, overall performance of SVM classifier for peanut variety identification on dedicated validation set is shown in Table 4. The overall SVM classifier's accuracy score for discriminating Peanut varieties was found to be 93.82%.

Table 3. Confusion matrix of K-NN classifier.

Ground truth	Predicted							Omission	Commission	Accuracy
	Chandra	Chitra	GAUG-1	Kaushal	Kuber	Prakash	Amber			
Chandra	27	2	0	1	0	0	0	10.00%	10.00%	90.00%
Chitra	0	29	1	0	0	0	0	3.33%	3.33%	96.67%
GAUG-1	0	1	28	1	0	0	0	6.66%	3.33%	93.33%
Kaushal	0	0	1	29	0	0	0	3.33%	0.00%	96.67%
Kuber	0	0	0	1	28	1	0	6.66%	3.33%	93.67%
Prakash	0	0	0	0	0	28	2	6.66%	6.66%	93.33%
Amber	0	0	1	1	0	1	27	10.33%	3.33%	90.00%
Overall accuracy										**93.33%**

Table 4. Confusion matrix of SVM classifier.

Ground truth	Predicted							Omission	Commission	Accuracy
	Chandra	Chitra	GAUG-1	Kaushal	Kuber	Prakash	Amber			
Chandra	29	0	0	0	0	1	0	03.33%	3.33%	96.67%
Chitra	0	27	0	03	0	0	0	10.00%	10.00%	90.00%
GAUG-1	0	1	28	1	0	0	0	6.66%	3.33%	93.33%
Kaushal	0	0	0	30	0	0	0	0.00%	0.00%	100.00%
Kuber	0	0	2	0	26	2	0	13.33%	6.66%	86.67%
Prakash	0	0	0	0	0	28	2	6.66%	6.66%	93.33%
Amber	0	0	0	0	0	1	29	3.33%	3.33%	96.67%
Overall accuracy										**93.82%**

In our research, using RFE, the color texture features were reduced from 72 to 14 features in order to avoid overtraining the classifier models.

Finally, we are able to design a low cost, efficient, robustness method for classifying peanut varieties by using office scanner. The mass surface imaging technique was used for feature extraction. Great classification accuracies can be achieved by using a limited or reduced number of feature sets. Furthermore, the approach we followed can be used in the peanut industry which can determine its variety in an efficient manner.

4 Conclusions

The intelligent system experimented based on image processing and computer-vision, and machine learning (KNN and SVM) classifiers to achieve better results to classify 7 peanut varieties. The demonstrated intelligent system has outperformed greater than 90% (93.82% for the best predictive model) in discriminating peanut variety. Finally, only 14 features were only minimal features required to classify.

Acknowledgments. The authors are greatly indebted to the Department of CSE MIT, MAHE, Manipal, INDIA-576104, for providing dedicated lab facilities to achieve this great work.

References

1. Burks, T.F., Shearer, S.A., Payne, F.A.: Classification of weed species using color texture features and discriminant analysis. Trans. Am. Soc. Agric. Eng. **43**(2), 441–448 (2000)
2. Chen, X., Xun, Y., Li, W., Zhang, J.: Combining discriminant analysis and neural networks for corn variety identification. Comput. Electron. Agric. **71**, 48–53 (2010)
3. Clausi, D.A.: An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote Sens. **28**, 45–62 (2002)
4. Dubey, B.P., Bhagwat, S.G., Shouche, S.P.: Potential of artificial neural networks in varietal identification using morphometry of wheat grains. Bio-Syst. Eng. **95**, 61–67 (2006)
5. Kurtulmus, F., Unal, H.: Discriminating rapeseed varieties using computer vision and machine learning. Expert Syst. Appl. **42**, 1880–1891 (2015)
6. Han, Z., Zhao, Y.: A cultivar identification and quality detection method of peanut based on appearance characteristics. J. Chin. Cereals Oils Assoc. **24**, 123–126 (2009)
7. Han, Z., Li, Y., Liu, J., Zhao, Y.: Quality grade-testing of peanut based on image processing. In: 2010 Third International Conference on Information and Computing, pp. 333–336 (2010)
8. Haralick, R.M.: Statistical and structural approaches to texture. Proc. IEEE **67**(5), 786–804 (1979)
9. Keuchel, J., Naumann, S., Heiler, M., Siegmund, A.: Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data. Remote Sens. Environ. **86** (4), 530–541 (2003)
10. Lu, Y., Boukharouba, K., Boonært, J., Fleury, A., Lecoeuche, S.: Application of an incremental SVM algorithm for on-line human recognition from video surveillance using texture and color features. Neurocomputing **126**, 132–140 (2014)
11. Lu, Y., Du, C., Yu, C., Zhou, J.: Classifying rapeseed varieties using Fourier transform infrared photoacoustic spectroscopy (FTIR-PAS). Comput. Electron. Agric. **107**, 58–63 (2014)

12. Mollazade, K., Omid, M., Tab, F.A., Kalaj, Y.R., Mohtasebi, S.S., Zude, M.: Analysis of texture-based features for predicting mechanical properties of horticultural products by laser light backscattering imaging. Comput. Electron. Agric. **98**, 34–45 (2013)
13. Nnorom, I.C., Jarzyńska, G., Drewnowska, M., Dryżałowska, A., Kojta, A., Pankavec, S.: Major and trace elements in sclerotium of Pleurotus tuberregium (Óstu) mushroom—dietary intake and risk in southeastern Nigeria. J. Food Compos. Anal. **29**(1), 73–81 (2013)
14. Omid, M., Mahmoudi, A., Omid, M.H.: An intelligent system for sorting pistachio nut varieties. Expert Syst. Appl. **36**(9), 11528–11535 (2009)
15. Pattee, H.E., Young, C.Y.: Peanut Science and Technology. American Peanut Research and Education Society Inc., Yoakum (1982)
16. Pazoki, A.R., Farokhi, F., Pazoki, Z.: Classification of rice grain varieties using two artificial neural networks (MLP and neuro-fuzzy). J. Anim. Plant Sci. **24**(1), 336–343 (2014)
17. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O.: Scikit-learn: machine learning in python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
18. Pydipati, R., Burks, T.F., Lee, W.S.: Identification of citrus disease using color texture features and discriminant analysis. Comput. Electron. Agric. **52**(1–2), 49–59 (2006)
19. Sakai, N., Yonekawa, S., Matsuzaki, A.: Two-dimensional image analysis of the shape of rice and its application to separating varieties. J. Food Eng. **27**, 397–407 (1996)
20. Shearer, S.A., Holmes, R.G.: Plant identification using color co-occurrence matrices. Trans. Am. Soc. Agric. Eng. **33**(6), 2037–2044 (1990)
21. Zapotoczny, P.: Discrimination of wheat grain varieties using image analysis and multidimensional analysis texture of grain mass. Int. J. Food Prop. **17**, 139–151 (2014)
22. Zhao, C., Han, Z., Yang, J.: Study on application of image process in ear traits for DUS testing in maize. Acta Agronomica Sinica **42**, 4100–4105 (2009)
23. Zou, Q., Fang, H., Liu, F., Kong, W., He, Y.: Comparative study of distance discriminant analysis and Bp neural network for identification of rapeseed cultivars using visible/near infrared spectra. In: Li, D., Liu, Y., Chen, Y. (eds.) CCTA 2010. IAICT, vol. 347, pp. 124–133. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-18369-0_15
24. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. J. Mach. Learn. Res. **3**(2003), 1157–1182 (2003)
25. Atas, M., Yardimci, Y., Temizel, A.: A new approach to aflatoxin detection in chilli pepper by machine vision. Comput. Electron. Agric. **87**, 129–141 (2012)

Intelligent System to Evaluate the Quality of Orange, Lemon, Sweet Lime and Tomato Using Back-Propagation Neural-Network (BPNN) and Probabilistic Neural Network (PNN)

V. G. Narendra and K. Govardhan Hegde(✉)

Computer Science and Engineering, Manipal Institute of Technology,
Manipal Academy of Higher Education, Manipal 576104, India
{narendra.vg, govardhan.hegde}@manipal.edu

Abstract. The quality assessment and sorting millions of fruits as well as vegetables by manual is usually slower. But also costly and cannot give an accurate result. In this research, to increase the quality of food above products were developed by using a vision-based quality inspection and sorting system. The quality assessment and sorting process analyzes taken image for its quality (good). It discards the defected one (bad). The image can be of vegetables or fruits. Four different systems for different food products (Orange, Lemon, Sweet Lime, and Tomato) have been developed. We have used a dataset of one thousand two hundred images which can be used to train as well as test the image systems. All images of 300 in the count. The obtained overall accuracy ranges between 85.0% to 95.00% for Orange, Lemon, Sweet Lime, and Tomato by using soft-computing techniques such as Backpropagation neural network and Probabilistic neural network.

Keywords: Quality inspection of fruits and vegetables · Backpropagation neural network · Probabilistic neural-network

1 Introduction

Nowadays, good food products quality is needed for human health. At present, to arrive desired quality assessment as good or bad of food products is a challenge because of an increasing in demand due to a large population. Presently, quality assessment and sorting tons of fruits/vegetables are done manually, and it leads to slow as well as cost also. Finally, it ends with an inaccurate process. Hence food quality assessment shows a very major role in providing complete defect-free food products to the consumers—quality, which usually defines the external as well as internal characteristics of materials of interest. In food quality, the Color and Texture are considered as external features. In industries where food is being processed, the food products are continuously tested such that hundreds of food products are scanned in nano seconds. The movement of the food products are monitored by CCD cameras, and finally, the defected materials are thrown away from the sieves.

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 369–382, 2019.
https://doi.org/10.1007/978-981-15-0108-1_34

For several years, to increase product quality control and also an attempt to reduce operation costs, the food industries have adopted automated vision-based scrutiny and sorting systems [6]. Nondestructive detections, like Near Infrared Spectroscopy, photoelectric detection, X-ray analysis, the electromagnetic characteristics analysis, computer vision etc. have been used progressively more in the food as well as agricultural industries for inspection and their evaluation purposes by means of which, it is able to provide suitably economic, economical, rapid as well as objective assessment [13]. The industries who rely on computer vision to monitor the food products have been identified, and later, the food industries are ranked among the top ten industries based on these technology [2]. An inspection system called Vision-based is used to reduce human interaction with these examined goods and then classify as fast as a human can interpret, thereby providing high accuracy in classifying the food products [6]. There exist several vision systems that have been developed for inspecting different varieties of food products, such as Apples, Tomatoes, Potatoes, Eggs, Corn, Rice, and many other food products [1, 2] and [13]. An apple grading system was being developed [19] by using a tool called vision-box-hardware, which provides high precision as well as high automatization [20]. Kohonen's self-organizing map was used for identifying baking curves of baked goods [21]. The primary sources of information being used in various literature are based on morphological, texture, and color based features for agricultural commodity [9], which focuses on sorting and grading and classification. A Computer vision based systems have been very successfully used to recognize or to classify food quality and identify various foods' parameters like its color, size, and shape, which are used in various agricultural industries. Food commodities include coffee [8], dry beans [14], the seed of soya beans [16], peanuts [12], and brazil-nuts [14].

In this research, soft computing techniques were used to develop an intelligent system to inspect the quality of the food products based on Color and Texture features. The proposed system is applied to four different types of food products, namely Tomatoes, Lemons Sweet Limes, and Oranges. As there exist several similarities between these four products mentioned above, different design and training are required for each food product.

The organization of the work is as follows: Sect. 2.1 defines the mechanism for quality inspection and their sorting system and usage of different components to sort food products by using computer vision technique. Section 2.2 depicts the source of data being used to classify these food products. Section 2.3 summarises image processing techniques like feature extraction and feature classification process which required to determine the type of the food and also give the conclusion or decision to accept or reject. Section 3 is about reports the results of the experiments. Finally, Sect. 4 concludes the overall research.

2 Materials and Methods

2.1 The Different Components Used in Quality Inspection and Sorting System

The quality assessment based on computer vision consists of diverse subsystems. The various components of the system are shown in Fig. 1. High processing cameras are being used to capture the high contrast of food image. A mirror can be attached to a single camera to inspect different sides. The usage of multiple cameras is used, which are fixed in different directions to get more clarity in the image. To overcome the variation in lighting problem, an isolated-box with light is being used so that we get better clarity in image. Once the image is captured, these high-quality images are sent to the computer for further processing and monitored in real time. The interfacing circuit which receives electronic signal saying the decision “pass” or “fail” The circuit has an electronic valve which drives the path of the food products using open and close where the closed path indicate bad product allowing only high-quality food products to pass the store. There could be more than two classes of foods which represent its quality in other degrees.

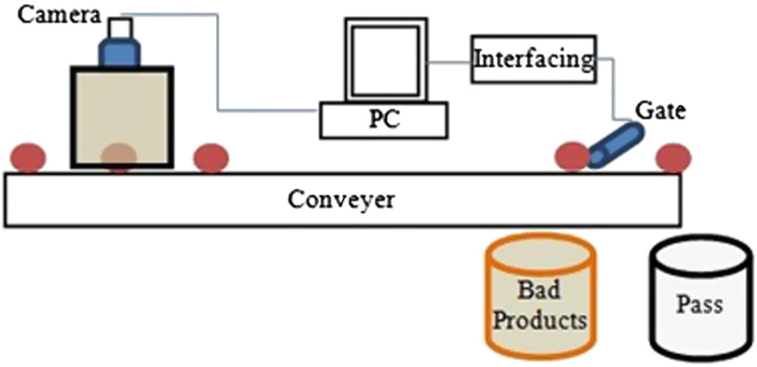


Fig. 1. The different components used in the Quality Inspection and Sorting System.

There are several modules in a computer vision based system which processes the image in real time. The various components used are shown in Fig. 2 for food quality inspection and their sorting. There will be an image acquisition module. It captures an image and then stored in computer memory. The quality of image and its file size is directly proportional to the speed of the sorting system. High-quality resolution images possess many details but require substantial time for its processing as well as classifying. Usage of low-quality resolution images require smaller processing time, but we fail to get high accuracy of the system. We use the suitable resolution image to give an acceptable speed with the best efficiency for inspecting and sorting.

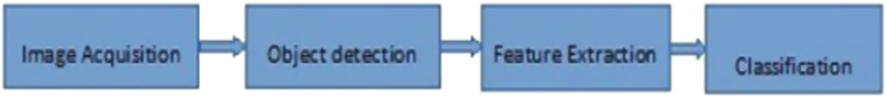


Fig. 2. The modules used in Computer-Vision System.

The first step, we need to identify the locations and the borders of the food product, to process and sort the captured image. Using image segmentation, it is divided into two classes, namely object, and background. Once the object is detected, we analyze again for its area to detect if there are any damages in the food product. As this process is nature dependent, it requires proper classification. We also identify cracks and holes in the food product; it's color using an image processing technique to extract various features. For clear decision, we have used trained classifier in the last step. The further sections respectively present the dataset being used, feature extraction, followed by the classification method.

2.2 Data Set

The FoodCast Research Image Database (FRID) is an attempt at standardizing food-related objects. It can bakery products, fruits, edible-nuts, any leafy or non-leafy vegetables as dataset. In the dataset, all images are of size (530 by 530 pixels). These are standardized and stored as the (.jpg) file format. In our study, we have considered 1200 food-related images and categorized into a lemon (300 images), Orange (300 images), Sweet Lime (300 images), and Tomato (300 images). The sample dataset is shown in Fig. 3.

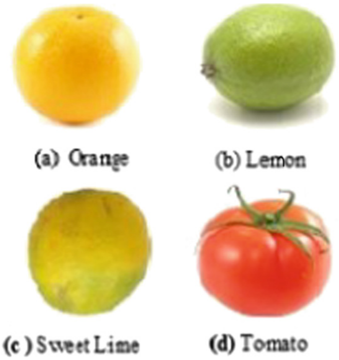


Fig. 3. Sample dataset.

2.3 Feature's Extraction

The feature's extraction is a significant phase in this research. We have used the segmented images of a different category from the FRID dataset. Then we have developed a feature extraction method to extract the features as Morphological, Color, and Texture. The HLS color space is used to extract the color characteristics of a categorized food product to measure luminance and chrominance. The measured color features are as follows [18]:

i. The Luminance (L): It describes the "achromatic" component, representing the brightness of an image.
ii. The Chrominance (C): Gives various color information of an image, and it is usually represented as two color-difference components.
iii. Hue (H): It represents the dominant color.
iv. Color Distance Metric (ΔE): It is a metric of difference between colors.

The brightness and contrast of each color-component are determined statistically as follows.

v. Mean (μ): The overall average brightness of each color component of a captured image.
vi. The Standard Deviation (): It gives the average distance from the mean of the total perceived brightness and contrast of each individual color component in a given image.
vii. Range (r): It gives the range of max and min perceived brightness of each color component in a given image.

2.3.1 HSL Color Space

There are two main aspects to realizes the importance of HSL color space: firstly, the chrominance components are hue and saturation, which are separated from luminosity, and secondly, how much color spectrum human perceived are given by these chrominance components [22]. The high color values for colors assigned in this space, which are approaching the white color with a bounded saturation. In this color space, color purity is measured by hue (H), the degree of white color embedded in a particular color is measured by saturation (S), and the brightness of color is measured by lightness (L).

The following steps are illustrating the conversion of RGB to HSL color space

1. Chroma calculation: The R, G, and B component values are divided by 255 to change the range from 0, …., 255 to range between 0, …., 1:

$$R' = \sum_{k=1}^{r} \sum_{l=1}^{c} R(k,l)/255 \tag{1}$$

$$G' = \sum_{k=1}^{r} \sum_{l=1}^{c} G(k,l)/255 \tag{2}$$

$$B' \sum_{k=1}^{r} \sum_{l=1}^{c} B(k,l)/255 \tag{3}$$

The Da-Wen-Sun, [18] is given chroma definition as “colorfulness relative to the brightness of a similarly illuminated white.”

$$C_{HSL} = Cmax - Cmin \tag{4}$$

where $Cmax = max(R', G', B')\ and\ Cmin = min(R', G', B')$

2. Hue calculation: [18] is given hue definition as “attribute of a visual sensation according to which an area appears to be similar to one of the perceived colors: red, yellow, green, and blue, or to a combination of two of them.”

$$H^{\circ}_{HSL} = \begin{cases} 60^{\circ} * \left(\frac{G'-B'}{C_{HSL}} mod 6\right), C_{max} = R' \\ 60^{\circ} * \left(\frac{B'-R'}{C_{HSL}} + 2\right), C_{max} = G' \\ 60^{\circ} * \left(\frac{R'-G'}{C_{HSL}} + 4\right), C_{max} = B' \end{cases} \tag{5}$$

3. Lightness or Luminosity calculation:

$$L_{HSL} = (Cmax + Cmin)/2 \tag{6}$$

4. Saturation calculation: The ratio of colorfulness to brightness or of chroma to lightness is saturation.

$$S_{HSL} = \begin{cases} 0, & C_{HSL} = 0 \\ \frac{C_{HSL}}{1-|2L_{HSL}-1|}, & C_{HSL} \neq 0 \end{cases} \tag{7}$$

5. The mean, standard deviation, and range of each color components of Hue, Luminosity, and Saturation are determined.

i. The mean, standard deviation, and range of hue component are determined using the following Eqs. 8, 9, and 10.

$$\mu_{HSL_H} = \frac{1}{r * c} \sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} H^{\circ}_{HSL}(k, l) \tag{8}$$

$$\sigma_{HSL_H} = \sqrt{\frac{1}{r * c} \sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} \left(H^{\circ}_{HSL}(k, l) - \mu_{HSL_H}\right)^2} \tag{9}$$

$$r_{HSL_H} = \max\left(\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} H^{\circ}_{HSL}(k, l)\right) - \min\left(\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} H^{\circ}_{HSL}(k, l)\right) \tag{10}$$

ii. The mean, standard deviation, and range of Luminosity component are determined using the following Eqs. 11, 12, and 13.

$$\mu_{HSL_L} = \frac{1}{r * c} \sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} L(k,l) \tag{11}$$

$$\sigma_{HSL_L} = \sqrt{\frac{1}{r * c} \sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} \left(L(k,l) - \mu_{HSL_L}\right)^2} \tag{12}$$

$$r_{HSL_L} = \max\left(\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} L(k,l)\right) - \min\left(\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} L(k,l)\right) \tag{13}$$

iii. The mean, standard deviation, and range of saturation component are determined using the following Eqs. 14, 15, and 16.

$$\mu_{HSL_S} = \frac{1}{r * c} \sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} S_{HSL}(k,l) \tag{14}$$

$$\sigma_{HSL_S} = \sqrt{\frac{1}{r * c} \sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} \left(S_{HSL}(k,l) - \mu_{HSL_S}\right)^2} \tag{15}$$

$$r_{HSL_S} = \max\left(\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} S_{HSL}(k,l)\right) - \min\left(\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} S_{HSL}(k,l)\right) \tag{16}$$

6. The color distance metric is calculated separately of Hue, Saturation, and Luminosity by using Eqs. 17, 18, and 19. And all three combined color distance metric is calculated using Eq. 20.

$$\Delta E_{HSL_H} = \sqrt{\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} \left(\mu_{HSL_H} - H^{\circ}_{HSL}(k,l)\right)^2} \tag{17}$$

$$\Delta E_{HSL_S} = \sqrt{\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} \left(\mu_{HSL_S} - S_{HSL}(k,l)\right)^2} \tag{18}$$

$$\Delta E_{HSL_L} = \sqrt{\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} \left(\mu_{HSL_L} - L(k,l)\right)^2} \tag{19}$$

$$\Delta E_{HSL} = \sqrt{\sum\nolimits_{k=1}^{r} \sum\nolimits_{l=1}^{c} \left(((\mu_{HSL_H} - H^{\circ}_{HSL}(k,l))^2 + ((\mu_{HSL_L} - L(k,l))^2 + ((\mu_{HSL_S} - S_{HSL}(k,l))^2\right)} \tag{20}$$

We have the extracted 14 features from each sample, which are listed in Table 1.

Table 1. The total extracted color measurement of HSL color space from each sample.

Number	Measurement	Code	Number	Measurement	Code
1	Mean of hue component	μ_{HSL_H}	10	Chroma of HSL	C_{HSL}
2	Mean of saturation component	μ_{HSL_S}	11	Color distance metric of hue component	ΔE_{HSL_H}
3	Mean of the luminance component	μ_{HSL_L}	12	Color distance metric of saturation component	ΔE_{HSL_S}
4	The standard deviation of the hue component	σ_{HSL_H}	13	Color distance metric of luminance component	ΔE_{HSL_L}
5	Standard deviation of saturation component	σ_{HSL_S}	14	Color distance metric of HSL	$\Delta \mathrm{E}_{HSL}$
6	Standard deviation of luminance component	σ_{HSL_L}			
7	Range of hue component	r_{HSL_H}			
8	Range of saturation component	$r_{HSL_S},$			
9	Range of luminance component	r_{HSL_L}			

2.3.2 Gray-Level-Cooccurrence Matrix (GLCM) Approach

It is a statistical approach. It can help in identifying and analyzing image texture [3, 15], and [17]. We get different co-occurrence probabilities.

$$\Pr(x) = \{P(i,j)|(d,\vartheta)\} \tag{21}$$

$$P(i,j) = \frac{P_{ij}}{\sum_{i,j=1}^{G} P_{ij}} \tag{22}$$

where the Pr(x) gives the measure of the probability. $P(i,j)$ gives the cooccurrence probability between grey-levels of i and j. Pij gives the number of the occurrence of the grey-levels [7]. There are 12 texture features [11] are calculated as follows.

$$\mathrm{Contrast} = \sum_{k,l=0}^{M_g-1} P_{k,l}|k-l|^2 \tag{23}$$

$$\mathrm{Correlation} = \sum_{k,l=0}^{M_g-1} \left[\frac{(k-\mu_k) \quad (l-\mu_l)}{\sqrt{(\sigma_k^2) \quad (\sigma_l^2)}} \right] \tag{24}$$

$$\mathrm{Ang\,Sec} - \mathrm{Moment} = \sum_k \sum_l p(k,l)^2. \tag{25}$$

$$\text{Energy equals}\sqrt{\text{Angular Second Moment}}. \tag{26}$$

$$\text{Homogeneity} = \sum\nolimits_{k,l=0}^{M_g-1} \frac{P_{k,l}}{1+(k-l)^2} \tag{27}$$

$$\text{Dissimilarity} = \sum\nolimits_{k,l=0}^{M_g-1} P_{k,l}\,|k-l| \tag{28}$$

$$\text{Entropy} = \sum\nolimits_{k,l=0}^{M_g-1} P_{k,l}\left(-\ln P_{k,l}\right) \tag{29}$$

$$\text{Cluster Shade} = \sum\nolimits_{k,l=0}^{M_g-1} \left((k-\mu_k)+(l-\mu_l)\right)^3 P_{k,l} \tag{30}$$

$$\text{Cluster Performance} = \sum\nolimits_{k,l=0}^{M_g-1} \left((k-\mu_k)+(l-\mu_l)\right)^4 P_{k,l} \tag{31}$$

$$\text{Smoothness} = 1-\frac{1}{(1+\sigma^2)} \tag{32}$$

$$\text{Third Movement} = \sum\nolimits_{k,l}^{M_g-1} \left(P_{k,l}\right)(k-\mu_l)^3 \tag{33}$$

$$\text{Maximum Probability} = \max_{k,l}(P_{k,l}) \tag{35}$$

In this study, the texture features based on GLCM are calculated for each given sample. Color texture extraction features are led by computing the textural features. It is used for HSL color-components. Feature-vector uses twelve GLCM features. Each image can be represented with only 14 + (12 × 3) = 50 features. Normalization of each feature is taken care before applying any machine learning approach which will remove the weakness of significant dissimilarities between feature sizes. The normalization is accomplished between the range 0 and 1.

2.4 Feature Selection Model

We use optimum feature selection model. It represents the best set of data. It is a subset of relevant features. In our study, we have used a statistical procedure called principal component analysis PCA as a feature model which can implement orthogonal transformation thereby converting samples that are correlated features into another set of values which are of linearly uncorrelated features known as Principle Components (PCs). It is also known as eigenvectors. With this, we can represent the data very efficiently. The ordinary way of obtaining significant features is using only the first three PCs [10] which is again used as the second feature set for further classifying food category as Lemon, Orange, Sweet Lime, and Tomato.

2.5 Soft Computing Techniques

An optimum classifier being used in pattern classification. It depends on the type of data set, how it able to interpret interactions among various features. In Overall approach, we test performances of the various classifiers. We then fine-tune their specific parameters. In this study, different Soft computing techniques (Backpropagation neural network and Probabilistic Neural Network) and are also the optimal classification model being built.

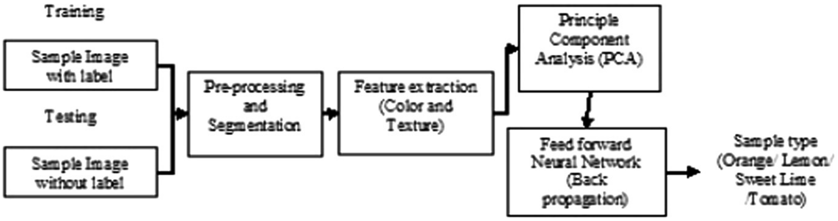

Training
Sample Image with label
Testing
Sample Image without label
Pre-processing and Segmentation
Feature extraction (Color and Texture)
Principle Component Analysis (PCA)
Feed forward Neural Network (Back propagation)
Sample type (Orange/ Lemon/ Sweet Lime /Tomato)

Fig. 4. Schematic representation of fruits and vegetable classification using BPNN.

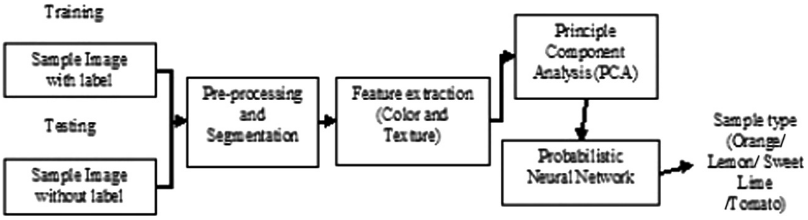

Training
Sample Image with label
Testing
Sample Image without label
Pre-processing and Segmentation
Feature extraction (Color and Texture)
Principle Component Analysis (PCA)
Probabilistic Neural Network
Sample type (Orange/ Lemon/ Sweet Lime /Tomato)

Fig. 5. Schematic representation of fruits and vegetable classification using PNN.

Multilayer feedforward neural networks using backpropagation learning (as shown in Fig. 4) and Probabilistic Neural Network (as shown in Fig. 5) are implemented for classification of four varieties; Lemon, Orange, Sweet Lime, and Tomato. An approach which is based on the combination of ten-fold cross-validation method being adopted for training, testing, and also tuning the specific parameters of the various classifiers.

After suitable training steps, the weights of the network are altered. After that employed for cross-validation. It helps to decide the net performance of the model. To reduce BPNN training time, a single hidden layer is considered in a neural network. Exhaustive search ranging from 1 to 50 is being carried out to determine the number of neurons that must be present in the hidden layer. The 28 nodes Neural Network in the hidden layer had high stability and the least standard deviation error. To develop BPPN models, linear function and the non-linear hyperbolic tangent function at both hidden as well as output layer are being used as a transfer function. Learning rate considered to be

0.9. It is throughout the momentum learning the rule. Data set is divided into two random sets 50% of which for training and 50% for testing, which acts as an additional guard alongside over-fitting. The neural network tool box present in MATLAB 2016a software is being used for designing and testing of the overall BPNN model.

The Probabilistic Neural Network (PNN) is completely based on Bayes decision rule, which uses Gaussian Parzen windows for estimation of probability density functions (pdf) as required in Bayes rule. It needs a single spread value for the pdf estimation. It is proportional to the Gaussian window width. Spread parameter or smoothing factor is directly proportional to the standard deviation (sd) of the Gaussian Parzen window. Proper spread parameters are selected for good performance of the PNN. In this study, to develop PNN, we have considered the whole features set to classify the four varieties such as Orange, Lemon, Sweet Lime, and Tomato. We have used the empirical spread parameter as a constant value (0.89) for training as well as test the features of the sample, which belongs to any one of the types, for classification. Data set is divided into two random sets 50% of which for training and 50% for testing, which acts as an additional guard alongside over-fitting. The neural network tool box of MATLAB 2016a software is used for designing and testing of the PNN model.

3 Results and Discussions

The classification tests are conducted on the color as well as texture features set. Among 1200 total samples of which 300 images each of Lemon, Orange, Sweet Lime, Tomato. (from each category: 150 samples as training and 150 samples as a test set), are chosen randomly. Ten-Fold cross-validation is being used for training as well as testing. For each fold, the proportion between the data used for training and data used for testing is 90:10%. The investigation is to the identification of food products into a category, namely Lemon, Orange, Sweet Lime, and Tomato. The obtained results of samples category are presented in Table 2 and also shown pictorially in Fig. 6.

Table 2. Classification results for sample category.

Category	BPNN		PNN	
	Training	Test	Training	Test
	Accuracy in %	Accuracy in %	Accuracy in %	Accuracy in %
Lemon	93.89	91.58	89.07	90.58
Orange	92.09	90.90	90.90	92.90
Sweet Lime	92.57	92.00	88.27	89.23
Tomoto	94.03	90.00	92.43	93.80

The obtained prediction accuracy for the training set by using BPNN is found to be as follows: Lemon (93.89%), Orange (92.09%), Sweet Lime (92.27%) and Tomato (94.03%). The obtained prediction accuracy for the test set by using BPNN is found to be as follows Lemon (91.58%), Orange (90.90%), Sweet Lime (92.00%) and Tomato (90.00%).

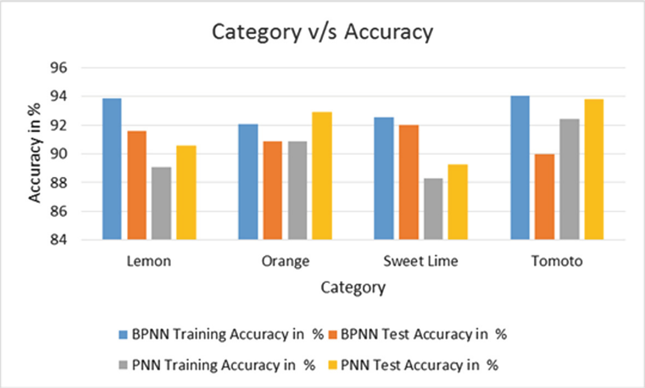


Fig. 6. Shows the category versus accuracy pictorially.

The obtained prediction accuracy for the training set by using PNN is found to be as follows Lemon (89.07%), Orange (90.90%), Sweet Lime (88.27%) and Tomato (92.43%). The obtained prediction accuracy for the test set by using PNN is found to be as follows Lemon (90.58%), Orange (92.90%), Sweet Lime (89.23%) and Tomato (93.80%).

The comparative analysis is made with earlier research work [5], tabulated in Table 3. The test set accuracies are considered for comparative analysis. It has been observed the proposed methods are outperformed compared to the report in the literature.

Table 3. Comparative analysis with earlier research work.

Category	Feature set	BPNN		PNN	
		Reported accuracy %	Proposed accuracy %	Reported accuracy %	Proposed accuracy %
Lemon	Color, texture	90.45	91.58	–	90.58
Orange	Color	88.89	90.90	90.00	93.90
Sweet Lime	Color, texture	89.30	92.00	–	89.00
Tomoto	Color, texture	86.76	90.00	–	93.80

4 Conclusions

An overall color, as well as texture features, were extracted from each sample image, proved to be the very precise method in recognizing characterized one. The study was limited to Lemon, Orange, Sweet Lime, and Tomato; therefore, further studies on more individual food products like fruits and vegetables are needed. A high accuracy and prediction performance of the results helped us to develop a quality evaluation of four food products and their classification.

Acknowledgments. The authors are much obliged to the Department of Computer Science and Engineering, Manipal Institute of Technology - Manipal Academy of Higher Education, Manipal, INDIA-576104, for providing outstanding lab facilities to succeed in this work.

References

1. Arivu, C.V.G., Prakash, G., Sarma, A.S.S.: Online image capturing and processing using vision box hardware: apple grading. Int. J. Mod. Eng. Res. **2**(3), 639–643 (2012)
2. Brosnan, T., Sun, D.W.: Improving quality inspection of food products by computer vision —a review. J. Food Eng. **61**(1), 3–16 (2004)
3. Burks, T.F., Shearer, S.A., Payne, F.A.: Classification of weed species using color texture features and discriminant analysis. Trans. Am. Soc. Agric. Eng. **43**(2), 441–448 (2000)
4. Castelo-Quispe, S., Banda-Tapia, J.D., López-Paredes, M.N., Barrios-Aranibar, D., Patino-Escarcina, R.: Optimization of Brazil-nuts classification process through automation using color spaces in computer vision. Int. J. Comput. Inf. Syst. Ind. Manag. Appl. **5**, 623–630 (2013)
5. Du, C.-J., Sun, D.-W.: Recent developments in the applications of image processing techniques for food quality evaluation. Trends Food Sci. Technol. **15**(5), 230–249 (2004)
6. Chetima, M.M., Payeur, P.: Automated tuning of a vision-based inspection system for industrial food manufacturing. In: Instrumentation and Measurement Technology Conference (I2MTC), pp. 210–215. IEEE, May 2012
7. Clausi, D.A.: An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote Sens. **28**, 45–62 (2002)
8. Soedibyo, D.W., Ahmad, U., Seminar, K.B., Subrata I.D.M.: The development of automatic coffee sorting system based on image processing and artificial neural network. In: The International Conference on the Quality Information for Competitive Agricultural Based Production System and Commerce, pp. 272–275 (2010)
9. Du, C.J., Sun, D.W.: Recent developments in the applications of image processing techniques for food quality evaluation. Trends Food Sci. Technol. **15**(5), 230–249 (2004)
10. Kurtulmus, F., Unal, H.: Discriminating rapeseed varieties using computer vision and machine learning. Expert Syst. Appl. **42**, 1880–1891 (2015)
11. Haralick, R.M.: Statistical and structural approaches to texture. Proc. IEEE **67**(5), 786–804 (1979)
12. Chen, H., Wang, J., Yuan, Q., Wan, P.: Quality classification of peanuts based on image processing. J. Food Agric. Environ. **9**(3&4), 205–209 (2011)
13. Jin, J., Li, J., Liao, G., Yu, X., Viray, L.C.C.: Methodology for potatoes defects detection with computer vision. In: International Symposium on Information Processing, pp. 346–351, August 2009
14. Kumar, M., Bora, G., Lin, D.: Image processing technique to estimate geometric parameters and volume of selected dry beans. J. Food Measur. Charact. **7**(2), 81–89 (2013)
15. Pydipati, R., Burks, T.F., Lee, W.S.: Identification of citrus disease using color texture features and discriminant analysis. Comput. Electron. Agric. **52**(1–2), 49–59 (2006)
16. Namias, R., Gallo, C., Craviotto, R.M., Arango, M.R., Granitto, P.M.: Automatic grading of green intensity in soybean seeds. In: 13th Argentine Symposium on Artificial Intelligence, ASAI, pp. 96–104 (2012)
17. Shearer, S.A., Holmes, R.G.: Plant identification using color co-occurrence matrices. Trans. Am. Soc. Agric. Eng. **33**(6), 2037–2044 (1990)

18. Sun, D.-W.: Computer Vision Technology for Food Quality Evaluation. Food Science and Technology, International Series. Elsevier Inc., Amsterdam (2008)
19. Narendra, V.G., Hareesh, K.S.: Quality inspection and grading of agricultural and food products by computer vision-a review. Int. J. Comput. Appl. **2**(1), 43–65 (2010)
20. White, D.J., Svellingen, C., Strachan, N.J.C.: Automated measurement of species and length of fish by computer vision. Fish. Res. **80**(2–3), 203–210 (2006)
21. Yeh, J.C., Hamey, L.G., Westcott, T., Sung, S.K.: Color bake inspection system using hybrid artificial neural networks. In: Proceedings of IEEE International Conference on Neural Networks, vol. 1, pp. 37–42, November 1995
22. Plataniotis, K., Venetsanopoulos, A.N.: Color Image Processing and Applications. Springer, Heidelberg (2000). https://doi.org/10.1007/978-3-662-04186-4

Comparative Study of RBF and Naïve Bayes Classifier for Road Detection Using High Resolution Satellite Images

Anand Upadhyay, Santosh Singh, Ajay Kumar Pandey,
and Nirbhay Singh(✉)

Thakur College of Science and Commerce,
Kandivali (E), Mumbai 400101, India
anandhari6@gmail.com, sksingh14@gmail.com,
ajaypanday678@gmail.com, nirbhaysingh69682@gmail.com

Abstract. The detection of the road is one of an area of satellite image classification. The satellite image classification plays a vital role in various area of monitoring different resources available on the earth surface. Here, the high-resolution satellite data from Google earth is acquired from a different region of Mumbai, Maharashtra, India region for detection of road. This research paper used two different algorithms i.e. radial basis function neural network and Naive Bayes classifiers for the detection of reading features from the high-resolution satellite image. Both algorithms are implemented using the Matlab simulation toolbox. Radial Basis Function and Naïve Bayes is a supervised classification technique applied on High-Resolution Satellite Image. Extraction of Road from the satellite image is a very difficult task because in the rural areas there are many unstructured roads which may consist of mud and concrete. After applying the algorithms on the image high-resolution satellite, the accuracy of classifiers is calculated using confusion matrix and Kappa coefficient. The accuracy of Naive Bayes found to be 91% with Kappa Value 0.698 and the accuracy of radial basis function found to be 99% with a Kappa value of 0.9831. The accuracy calculation using confusion matrix and Kappa value shows that the radial basis function neural network classifier is better than Naive Bayes classifiers for the detection of the road using high-resolution satellite image.

Keywords: Naïve Bayes · Radial Basis Function · Google Earth · Satellite Image · Matlab · Remote Sensing

1 Introduction

The remote sensing is one of the major inventions of science and technology which is used to solve many problems on a daily basis or major problems related to climate changes, national securities and different issues related to the monitoring of different resources available on the earth surface. The remote sensing is widely used technology. In remote sensing, there is a satellite which is far away from the earth surface and which consist of a high-resolution powerful camera it captures the image of the various parts of the earth surface. So, these particular images are used for different purposes i.e.

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 383–392, 2019.
https://doi.org/10.1007/978-981-15-0108-1_35

climate change and monitoring, land use land cover classification, change detection and crop monitoring by a different organization, agencies, and other research organization for their study and planning of different resources. Road plays an important role in transportation in India because in India there are 85% of passenger traffic is carried by the Roads which is greater than the other transportation. The roads have a different type which we have to identify for good transport. Roads are very useful in the self-driving car which is future of the new transportation because if we cannot detect the road with 0% error then we cannot allow the self-driving car on the road if we allow it may cause the accident on the road. Satellite image consists of much information like the forest, water, building, ocean, etc. A problem in road extraction since unstructured road boundaries may be incorrectly identified or simply hinder the road detection process increase to a higher false rate detection. While referring the research paper and the title upon the road importance and the problem in road detection we decided to work on the road detection technique that can help the people to find the road very simple. There are many research papers based on the road detection in which many Arthur used many techniques to detect the road but to minimize that problem, we introduced the new way of road detection by using the naïve Bayes algorithm. Radial Basis Function (RBF) is part of an artificial neural network it comes under supervised machine learning, this algorithm is used to classify the image. Naïve Bayes algorithm is also used to classify the image and both the algorithm compare with each other. Paper presents a road detection method based on Naïve Bayes classification and radial basis function neural network classifier. The proposed method consists of two phases, in the first phase, the preprocessing methods are applied and classification feature data sets are collected to make training file. In the second phase, Naïve Bayes classifiers and Radial Basis Function are applied on an image by utilizing color features of every pixel in a satellite image from a different area of a satellite image. The implementation of both the algorithm compares the training value to the actual image and accordingly based on the training value the thematic map of detected roads are generated. The above-implemented work helps to various people from a different area to make their project related to road detection and monitoring purpose. This work also helps normal people also to get information about the roads in rural and urban areas (Fig. 1).

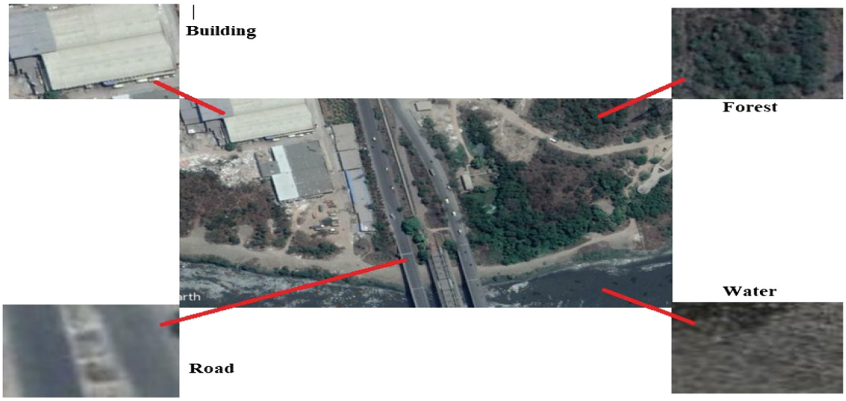


Fig. 1. Road and other information

2 Literature Review

Finding the road from the satellite image is very difficult as well as very time-consuming task. To extract the road from satellite image many techniques are developed which reduces the time, some approaches are mention here.

An automatic procedure is developed for the detection of changes between an existing road and a newly registered satellite image. In this paper there is existing database is used to detect the road and delineation of the corresponding road with the road available in the image. The each road taken separately for comparison with the actual road in image and database. Here the position errors are considered which is randomly distributed. This proposed technique optimizes the delineation process [1]. These days a vital issue in consolidating geospatial information from various sources is that they infrequently adjust. In this paper, we present a methodology for vector-picture conflation and build up a calculation which identifies street convergences from both datasets as control focuses by utilizing picture surface characterization. With this strategy, we first train the framework on a little territory of the orthoimagery to gain proficiency with the street surface conveyance, at that point we can get its division as indicated by its surface, and lastly, the framework finds street crossing point focuses. The last advance is to adjust vector information and pictures by utilizing diverse systems [2]. The advancement and development in the area of remote sensor and satellite technologies produce high-resolution satellite images like IKONOS satellite images. The authors have used IKONOS high-resolution satellite images and proposed a new method for extraction of road network from high-resolution satellite images. The proposed method is based on the binary and grayscale mathematical morphological and line segments matching based techniques. At the first stage, the outline of the road network is detected based on the gray morphological characteristics and later road network is detected based on the line segment match method [3]. Here, the semi-automatic methods are used for road detection. This semi-automatic method uses black and white aerial photographs or panchromatic high-resolution satellite image band for detection. The geometrical characteristics are used for detection of road along with various road detection algorithms [4]. Road extraction from satellite image is a very vital task in terms of research and practices both. Here the improvised model is developed based on the human vision system and principles of perceptual organization [10]. Here the automatic mapping of urban and suburban roads from high-resolution satellite images is implemented. Here the road extraction techniques are improvised by using the fusion of data from multiple data sources. The fusion of satellite data from multiple sources improvised the accuracy of road detection. Here proposed method models the urban and suburban differently [11].

3 Data Characteristics and Field Study

Data plays an important role in research work. If you have accurate and up-to-date data then research will be very effective and impactful. Remote Sensing is a technique which provides the data of any place from the remote place without being making physical contact with that place. In proposed research the remotely sensed S data is collected from Google Earth. Google Earth Provides accurate and up-to-date date of any place. Google Earth provides the data of any place which is recently captured one or two days before. The data which is collected from Google Earth is high resolution satellite image of the road. The image taken from the Google Earth is an RGB image. In India, approximately 1.35 million people die each year because of accident. The total length of National highways are 66,754 km^2, State highways length are 128,000 km^2, District roads are 470,000 km^2 & Rural road 2,650,000 km^2. 1 way lane, intermediate lane are length of 18,350 km^2 i.e. 27% of total road, 2 way lanes is 39,079 km^2 i.e. 59% of total road, 4–8 ways lanes is of 9,325 km^2 i.e. 14% of total road [14]. As only 59% of total road is 2 way lanes so there is need of increasing 2 way lanes & 4–8 way lane in India to reduce the number of accident. There are different reasons of accident like drink and drive, high speed, bad road conditions, no rules and regulations of driving vehicle. As in India human population is increasing, the use of vehicle is also increasing. Many roads in India contain potholes so in night time & in rainy season it is not visible properly by human naked eye, which leads to more probability of accident. This type of roads should be identified and reconstructed. In the rural areas there are many unstructured roads which may consist of mud and concrete. So due to these reasons in rural area many problems are occurring in transportation. Remote sensing is technique which provides the data of these rural areas to help in mapping the roads from remote places without being physically present at that place.

4 Methodology

The algorithm is developed in MATLAB R2010. It uses a Radial Basis Function and Naive Bayes algorithm for the classification of roads from the high resolution satellite image. High resolution satellite image are taken from Google Earth. Google Earth provides image which consist the features in the form of RGB (Red, Green, Blue) value. Below diagram gives the working of system (Fig. 2).

A. Feature Extraction

The RGB image is obtained from Google earths which consist of many features which are stored in the form of RGB value. RGB stands for R (Red), G (Green), and B (Blue). So the first step is to extract the RGB value of Road and others i.e. building, vehicle, garden etc. After extracting the RGB value of road training file is prepared which consist of pixel values of road area and others then give label to them. After applying algorithm system will train the model.

B. Radial Basis Function

The Matlab is very powerful mathematical computational model which provides the tool for RBF artificial neural network based classification. The RBF artificial neural network is designed by using the RBF neural network toolbox. Radial Basis Function algorithm is a type of neural network. It performs classification by measuring the inputs Similarity to examples from the training set. Every radial basis function neuron stores a prototype, which is just one of the examples from the training data. When we want to classify a new input, each neuron computes the Euclidean distance between the input and its prototype. In our study, we made a training file that consists of the RGB pixel

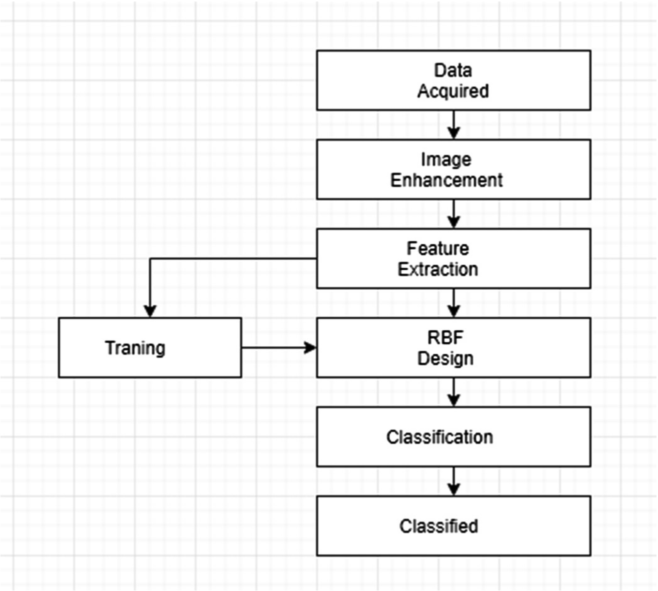


Fig. 2. Flow of the classification

value of the road and other and we give then different tag to each other, we passed the training and Target file to the algorithm where the machine train according to the training file and then classified the road on the given target file.

Algorithm

$$f(x) = \sum_{j=1}^{m} w_j h_j(X)$$

$$h_j(x) = \frac{ex(-(x - c_j)2)p}{r_j 2}$$

Where-

- $C_{j\text{-}}$ is center of region
- $r_{j\text{-}}$ is width of the receptive field

C. Naïve Bayes Algorithm

Naive Bayes classifiers are basically used in the classification of the object in the image and it is a collection of classification algorithms based on Bayes Theorem and it works on conditional probability. Using the conditional probability, we can find the probability of an event using its prior knowledge, in our research we have taken the pixel value of the road and non-road, we pass our training file and the target file as input to the algorithm, algorithm take pixel value of both files and compare with each other, whenever it find the pixel value that matched with the pixel value of the road it give then a different tag.

Below is the formula for finding the conditional probability.

Algorithm

$$P\left(\frac{c}{x}\right) = \frac{P\left(\frac{x}{c}\right)P(c)}{P(x)}$$

D. Classification

Based on the training model the system will perform the classification of the image in two groups Road and others. The system will show the Road portion in the form of Red color and remaining portion will be same. Based on this classification confusion matrix, accuracy and Kappa coefficient is obtained.

5 Experimental Results

The RBF trained with the sample data. After training, testing is performed using testing dataset. After training accuracy and Kappa value is obtained by confusion matrix and classification report is generated (Table 1).

Table 1. Confusion matrix of trained RBF

Classes	Road	Non-road	Total	User accuracy
Road	382	2	384	99.73%
Other	3	237	240	99.37%
Total	385	239	630	
Produce accuracy	99.22%	99.16%		

$$\begin{aligned}\text{Accuracy} &= (619/624) * 100\\ &= 99.19\%\end{aligned}$$

$$\mathrm{K} = \frac{\mathrm{N}\sum_{\mathrm{i}=1}^{\mathrm{r}} \mathrm{xii} - \sum_{\mathrm{i}=1}^{\mathrm{r}}(\mathrm{Xi} + \ \to \mathrm{X} + \mathrm{i})}{\mathrm{N2} - \sum_{\mathrm{i}=1}^{\mathrm{r}}(\mathrm{Xi} + \ \to \mathrm{X} + \mathrm{i})}$$

$$\mathrm{K} = 0.9831\,(\text{very good})$$

Above is the Confusion matrix and Kappa coefficient (K) for the dataset which is tested by RBF. It shows very good accuracy of classification (Figs. 3, 4 and Table 2).

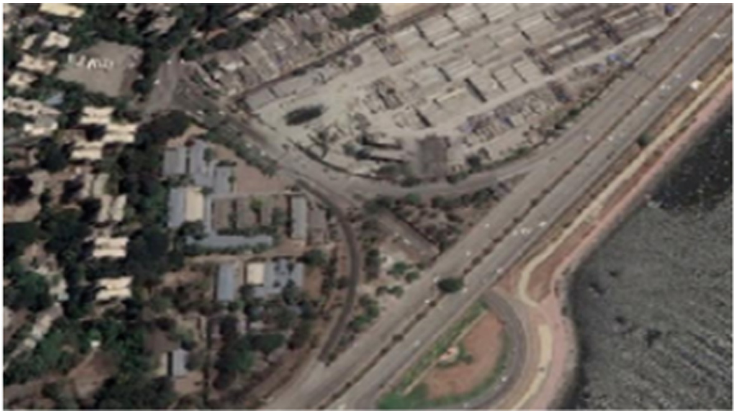


Fig. 3. Image before classification

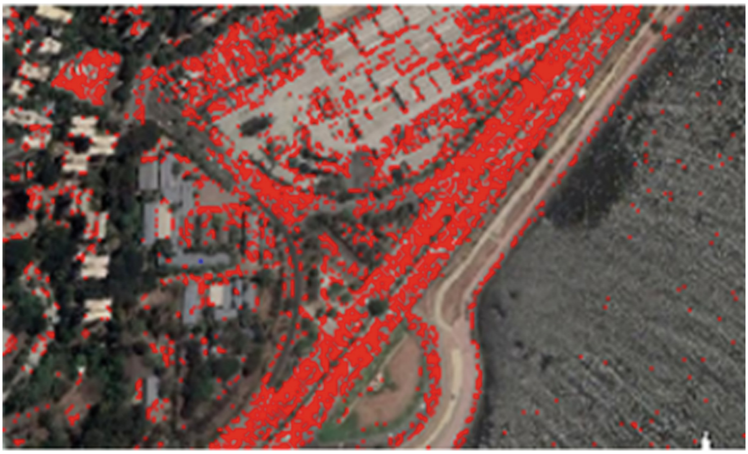


Fig. 4. Image after classification

Table 2. Confusion matrix of trained Naïve Bayes

Classes	Road	Non-road	Total	User accuracy
Road	274	23	297	92.25%
Other	5	41	46	89.13%
Total	279	64	343	
Produce accuracy	98.20%	64.06%		

$$\begin{aligned}\text{Accuracy} &= (315/343) * 100 \\ &= 91.20\%\end{aligned}$$

$$\mathrm{K} = \frac{\mathrm{N}\sum_{\mathrm{i}=1}^{\mathrm{r}} \mathrm{xii} - \sum_{\mathrm{i}=1}^{\mathrm{r}} (\mathrm{Xi}+\ \rightarrow \mathrm{X}+\mathrm{i})}{\mathrm{N2} - \sum_{\mathrm{i}=1}^{\mathrm{r}} (\mathrm{Xi}+\ \rightarrow \mathrm{X}+\mathrm{i})}$$

$$\mathrm{K} = 0.698\ (\mathrm{Good})$$

Above is the Confusion matrix and Kappa coefficient (K) for the dataset which is tested by Naïve Bayes algorithm. It shows good accuracy of classification (Figs. 5 and 6).

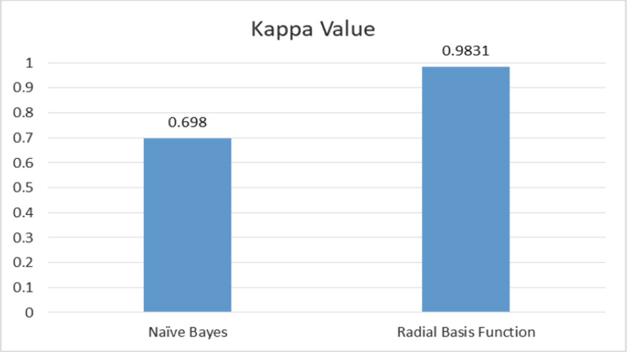


Fig. 5. Camparative study using kappa value

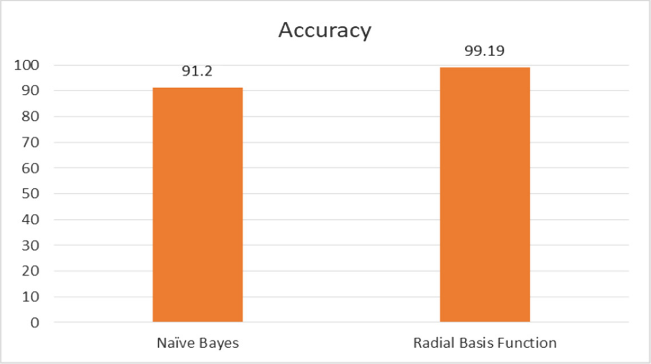


Fig. 6. Comparative study using accuracy

6 Conclusion

In this paper we successfully achieved the road detection with the help of the Naïve Bayes classifier and Radial Basis Function algorithm which extracted the road from the High-Resolution satellite image and gives the different color to road and it is visible to the end user. After applying both the algorithm classification is performed this shows the different accuracy on both the algorithm. Radial Basis Function gives the good accuracy than the Naive Bayes algorithm. From this study it is concluded that Radial Basis Function algorithm is giving higher accuracy than the Naive Bayes algorithm. So Radial Basis algorithm is good classification algorithm.

7 Future Enhancement

The above study is done using Radial Basis Function Neural Network and Naive Bayes for the Road Detection of high resolution satellite image. For the future study, other classification algorithms like Support Vector Machine (SVM), Artificial Neural network (ANN) etc. can be used for road detection using high resolution satellite image and for their comparative study.

References

1. Klang, D.: Automatic detection of changes in road data bases using satellite imagery. Int. Arch. Photogram. Remote Sens. **32**, 293–298 (1998)
2. Ruiz, J.J., Rubio, T.J., Urena, M.A.: Automatic extraction of road intersections from images based on texture characterization. Surv. Rev. **43**(321), 212–225 (2011)
3. Zhu, C., et al.: The recognition of road network from high-resolution satellite remotely sensed data using image morphological characteristics. Int. J. Remote Sens. **26**(24), 5493–5508 (2005)
4. Gruen, A., Li, H.: Semi-automatic linear feature extraction by dynamic programming and LSB-snakes. Photogram. Eng. Remote Sens. **63**(8), 985–995 (1997)
5. Mena, J.B.: State of the art on automatic road extraction for GIS update: a novel classification. Pattern Recogn. Lett. **24**(16), 3037–3058 (2003)
6. Quackenbush, L.J.: A review of techniques for extracting linear features from imagery. Photogram. Eng. Remote Sens. **70**(12), 1383–1392 (2004)
7. Mayer, H., Hinz, S., Batcher, U., Batavia's, E.: A test of automatic road extraction approaches. Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci. **36**(3), 209–214 (2006)
8. Poullis, C., You, S.: Delineation and geometric modeling of road networks. ISPRS J. Photogram. Remote Sens. **65**(2), 165–181 (2010)
9. Hu, J., Razdan, A., Femiani, J.C., Cui, M., Wonka, P.: Road network extraction and intersection detection from aerial images by tracking road footprints. IEEE Trans. Geosci. Remote Sens. **45**(12), 4144–4157 (2007)
10. Yang, J., Wang, R.S.: Classified road detection from satellite images based on perceptual organization. Int. J. Remote Sens. **28**(20), 4653–4669 (2007)
11. Jin, X., Davis, C.H.: An integrated system for automatic road mapping from high-resolution multi-spectral satellite imagery by information fusion. Inf. Fusion **6**(4), 257–273 (2005)

12. Guo, Y., et al.: Genetic algorithm and region growing based road detection in SAR images. In: Third International Conference on Natural Computation (ICNC 2007), vol. 4. IEEE (2007)
13. Hu, X., Tao, C.V., Hu, Y.: Automatic road extraction from dense urban area by integrated processing of high resolution imagery and lidar data. Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci. **35**(B3), 288–292 (2004)
14. Mohan, D., Tsimhoni, O., Sivak, M., Flannagan, M.J.: Road safety in India: challenges and opportunities (2009)

Singer Identification Based on Artificial Neural Network

Sharmila Biswas(✉) and Sandeep Singh Solanki

ECE Department, BIT-Mesra, Ranchi, Jharkhand, India
biswas.sharmila80@yahoo.com, sssolannki@bitmesra.ac.in

Abstract. Music is a vocal or instrumental sounds (or combined) presented in such a way as to create, harmony, and expression of emotion. Present days accurately singer identification is necessary for music indexing and retrieval purpose. This paper proposes a unique features extraction algorithm for singer identification. In this paper, seven singers with five vocal songs are considered for singer identification. The most potential Mel-Frequency-Cepstral Coefficient (MFCC) based feature extraction algorithm, and an artificial neural network (ANN) classifier has been applied for the singer identification purpose. Four multilayer neural network training algorithm such as Levenberg-Marquardt, Bayesian regularization Backpropagation, Scaled Conjugate Gradient, and One-step secant Backpropagation algorithm has been used to classify the seven different singers voices. Three different feature extraction technique, such as MFCC, MFCC with five different musical features and MFCC with ten different musical features has been considered for the feature extraction. The highest training and testing accuracy have been achieved through this algorithm is 98.3% and 88.6%. Classification accuracy varies with musical features and classification algorithms.

Keywords: Artificial Neural Networks (ANN) · Mel-Frequency Cepstral Coefficients (MFCC) · Music Information Retrieval (MIR) · Music signal

1 Introduction

Singer identification algorithms [1–36] are useful for music information retrieval (MIR) purpose [37]. The feature extraction part of the proposed algorithm contains Mel-Frequency Cepstral Coefficients (MFCC) and Musical features [38]. The extracted features are subjected to the neural network classifier [39, 40] for the classification purpose. The paper presents a comparative analysis of classification accuracy based on four different Artificial Neural Networks (ANN) learning algorithms such as LM: stands for Levenberg-Marquard algorithm, BR: stands for Bayesian Regularization algorithm, SCG: stands for Scaled Conjugate Gradient algorithm and OSS: stands for one-step secant backpropagation. Due to the adaptive and non-linear data-driven nature the ANN classifier is applied for classification. Out of the various research areas of music signal processing, a significant amount of research has been done on speaker identification due to the wide applicability. The Gaussian Mixture Model (GMM) [35]

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 393–401, 2019.
https://doi.org/10.1007/978-981-15-0108-1_36

achieves 92.3% with background removal and 78.3% without background removal respectively for each singer. The reference [41] using perceptual features achieves 81% accuracy on various Indian male and female PS's songs.

The paper is organized as follows: Sect. 2 provides the experimental set-up of this article; Sect. 3 describes the proposed methodology; Sect. 4 reports the result and discussions, and finally, Sect. 5 explains the conclusion and future work.

2 Experimental Set-Up

Five popular Hindi songs have been validated the proposed algorithm. In this research work, five Hindi songs of seven different singers have been recorded by using Behringer mic, Behringer mixer, and Creative 5.1 sound card, with sonic foundry sound forge 7.0 software. The same type of five Hindi songs of seven singers is recorded in the WAV format. All songs are segregated into ten different frames of 5 s duration. Every single frame contains 350 samples. The input audio files have various attributes such as file type (WAV), sampling rate (44.1 k), audio type (mono). The framing process has been performed by using MATLAB software.

3 Proposed Methodology

The block diagrammatic representation of the proposed singer identification process is presented in Fig. 1. The recorded music's have been separated into the two-part, such as training and testing signal or data. The training data has been used for the training of the classifier.

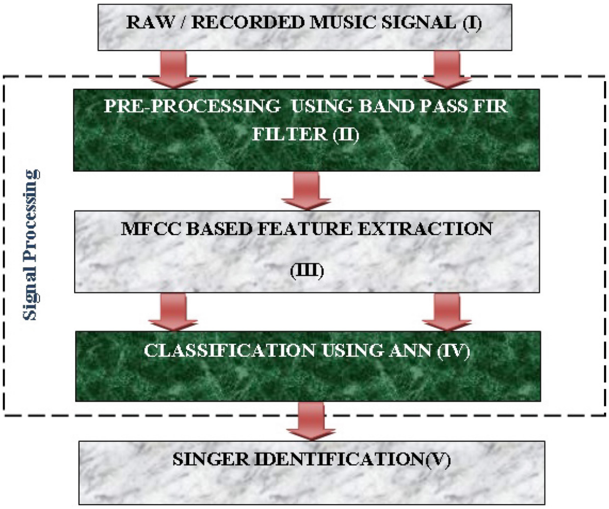

RAW / RECORDED MUSIC SIGNAL (I)
PRE-PROCESSING USING BAND PASS FIR FILTER (II)
MFCC BASED FEATURE EXTRACTION (III)
CLASSIFICATION USING ANN (IV)
SINGER IDENTIFICATION(V)
Signal Processing

Fig. 1. Proposed algorithm for singer recognition

The trained classifiers are applied to analyzed the testing data. The proposed algorithm has been validated by segregating the experiment into two parts. The training of the classifier has been performed in the first part, and the testing of the proposed system has been carried out in the second part. The recorded music signals (displays in the block I) are contaminated with various kinds of noises. The FIR filter has been applied to remove the eco noises (displays in block II) present in the signal.

The filtered music signal is subjected to the MFCC based feature estimation process (displays in step number III of the block diagram). The feature matrixes are further processed to the ANN classifier (shown in step number IV). Classification decision is further fed to the singer identification system (displays in step number IV of the block diagram).

Preprocessing

Recording of music signals has been carried out in the laboratory atmosphere. However, the audio wave of the music performed in a hall room was distributed in all directions and finally recorded it with many times reflection in all directions in a different time. As a result, the echo generated and mixed with the signal. The recorded music is contaminated with the echo noise. In order to retrieve the original music from the recorded musical signal, the signal passes through an FIR filter to remove the echo.

Feature Extraction

The filtered music signal passes through the MFCC based feature extraction process. Three different technique has been applied to extract the features from the recorded music signal. Ten different musical feature like RMS Energy, Event Density, Tempo, Pulse Clarity, Zero Crossing rate, Spectral Irregularity, Spectral Centroid, Skewness, Kurtosis, Shanon Entropy, are considered in feature extraction.

Mel Frequency Cepstral Coefficients (MFCC) describe the spectral shape of an audio input. It is a multiprocessing system. First, the frequency bands are logarithmically positioned. This is called as Mel scale. A method that has the energy compaction capability called as Discrete Cosine Transform (DCT). DCT that considers only the real numbers. By default first 13 components are taken here. The linear frequency scale to the Mel scale frequency mf is given below

$$\mathrm{m_f} = 1127.01048\log_{\mathrm{e}}\left((1+\mathrm{f}/700)\frac{\mathrm{f}}{700}\right), \tag{1}$$

where f is the frequency in hertz for a linear scale.

Root Mean Square (RMS) energy denotes the square root of the mean square of the amplitude values of the audio signal

$$x_{rms} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} \tag{2}$$

where $\mathrm{x_i}$ denotes the magnitude of the $\mathrm{i^{th}}$ sample and the total number of samples is n.

Event Density gives the average frequency of events. It is defined by the onsets per second. Tempo is estimated by identifying periodicities from the onset detection curve. Event Density gives the average frequency of events, i.e., the number of note the tempo is evaluated by detecting periodicities from the onset detection curve [35].

Pulse Clarity is defined as the amount of rhythmic clarity or the strength of the musical beats [35].

Spectral Irregularity is estimated by the degree of variation between successive peaks.

$$\mathrm{I} = \frac{\sum_{\mathrm{k}-1}^{\mathrm{N}} (\mathrm{a_k} - \mathrm{a_{k-1}})^2}{\sum_{\mathrm{k}-1}^{\mathrm{N}} \mathrm{a_k^2}}, \tag{3}$$

where, a_k denotes k$^{\mathrm{th}}$ spectral coefficients magnitude and total no of spectral coefficients is N.

Zero Cross Rate is the number of times the signal crosses X-axis or changes its sign i.e. positive to zero and zero to negative or negative to zero and zero to positive. It is a very vital feature to identify percussive sounds

$$\mathrm{Z} = \frac{1}{\mathrm{T}} \sum_{\mathrm{t=m-T+1}}^{\mathrm{m}} \frac{|\mathrm{sgn(s_t) - sgn(s_t - 1)}|}{2} \mathrm{w(m - t)}, \tag{4}$$

where T is the length of the time window, $\mathrm{s_t}$ is the magnitude of the t$^{\mathrm{th}}$ time domain sample, and w is a rectangular window.

Spectral Centroid gives the weighted mean of the frequencies component present in a signal.

$$Centroid = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}, \tag{5}$$

where, k$^{\mathrm{th}}$ spectral coefficients magnitude is $\mathrm{a_k}$ and total no of spectral coefficients is N.

Skewness is the measure of the symmetry or asymmetry nature of a signal.

$$\mathrm{Skewness} = \frac{\sum_{\mathrm{i=1}}^{\mathrm{N}} (\mathrm{Y_i} - \bar{\mathrm{Y}})^3}{(\mathrm{N} - 1)\mathrm{s}^3}, \tag{6}$$

where $\bar{\mathrm{Y}}$ is the mean, s is the standard deviation and N is the number of data points.

Kurtosis determines of the noisiness nature of the signal. It shows whether the data is peaked or flat relative to a normal distribution.

$$\text{Kurtosis} = \frac{\sum_{i=1}^{N} (Y_i - \bar{Y})^4}{(N-1)s^4}, \tag{7}$$

where $\bar{Y}$ is the mean, s is the standard deviation and N is the number of data points.

Shannon Entropy gives the following equation for information theory.

$$H(x) = -\sum_{(i=1)}^{N} p(x_i)\log_2 p(x_i), \tag{8}$$

Spectral Entropy gives information about the signal and indicates whether it contains predominant peaks or not.

Neural Network Based Supervised Classification
The authors have been chosen ANN classifier due to its non-linearity and adaptive nature. It classifies inputs into a defined set of target group. Feed-forward network of two-layers with hidden sigmoid and softmax output neurons and hidden layers can classify vectors arbitrarily. Pattern recognition networks classify inputs according to the target classes. Four multilayer training functions such as Levenberg-Marquard (LM), Bayesian Regularization (BR), Scaled Conjugate Gradient (SCG) and one-step secant back-propagation (OSS) are used to classify the music features.

4 Result and Discussions

Based on the feature extraction algorithm, the experimental process has been segregated into three parts. In the 1st part, the classifiers have been trained by using the 13 MFCC feature. The maximum accuracy achieved through 13 MFCC is 97.1% in training using Bayesian Regularization and 80% in testing using Levenberg-Marquard and One-step Secant Back-propagation. In the 2nd part, the classifier has been trained by using 13 MFCC with 5 Musical features. The maximum accuracy achieved through this feature is 98% in training and 85.7% in testing using Bayesian Regularization. Performance evaluation of the proposed algorithm over the training dataset using hidden layer size 30 and 40, in terms of classification accuracy were estimated for singer identifications shown in Table 1. Table 2 depicts the performance evaluation of the proposed algorithm over the testing dataset using hidden layer size 30. The results are shown in terms of the classification accuracy. Three different feature data size has been considered for the experiment 35, 42, 49.

Table 1. Performance evaluation of the proposed algorithm over the training dataset using hidden layer size 30 and 40, in terms of classification accuracy

Algorithms	13 MFCC (%)		13 MFCC + 5 Musical features (%)		13 MFCC + 10 Musical features (%)	
	30	40	30	40	30	40
Levenberg - Marquard (LM)	92.0	93.4	96.3	96.3	95.1	96.3
Bayesian Regularization (BR)	97.1	97.1	98.0	98.0	98.3	97.7
Scaled Conjugate Gradient (SCG)	84.6	88.6	97.7	94.0	95.4	94.9
One-step Secant Back-propagation (OSS)	87.1	84.3	97.7	93.4	94.0	93.7

Table 2. Performance evaluation of the proposed algorithm over the testing (with data size 35, 42, 49) dataset using hidden layer size 30, in terms of accuracy

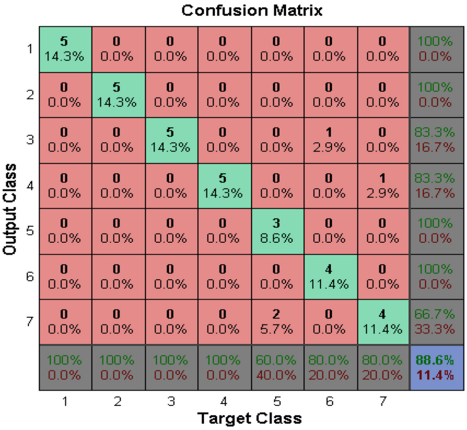
Algorithms	13 MFCC (%)			13 MFCC + 5 Musical features (%)			13 MFCC + 10 Musical features (%)		
	35	42	49	35	42	49	35	42	49
Levenberg - Marquard (LM)	80.0	85.7	71.4	80.0	83.3	83.7	88.6	85.7	75.5
Bayesian Regularization (BR)	74.3	78.6	79.6	85.7	81.0	81.6	85.7	83.3	83.7
Scaled Conjugate Gradient (SCG)	77.1	83.3	77.6	82.9	83.3	85.7	88.6	85.7	89.8
One-step Secant Back-propagation (OSS)	80.0	73.8	77.6	80.0	83.3	85.7	85.7	88.1	87.8

Table 3 shows the performance evaluation of the proposed algorithm over the testing dataset using hidden layer size 40. The results are shown in terms of the classification accuracy. Three different feature data size has been considered for the experiment 35, 42, 49.

Table 3. Performance evaluation of the proposed algorithm over the testing (with data size 35, 42, 49) dataset using hidden layer size 30, in terms of accuracy

Algorithms	13 MFCC (%)			13 MFCC + 5 Musical features (%)			13 MFCC + 10 Musical features (%)		
	35	42	49	35	42	49	35	42	49
Levenberg - Marquard (LM)	85.7	73.8	73.5	80.0	85.7	77.6	88.6	83.3	91.8
Bayesian Regularization (BR)	78.6	81.0	77.6	80.0	78.6	79.6	85.7	81.0	85.7
Scaled Conjugate Gradient (SCG)	83.3	73.8	77.6	82.9	83.3	75.5,	88.6	90.5	83.7
One-step Secant Back-propagation (OSS)	73.8	76.2	79.6	77.1	81.0	77.6	85.7	92.9	87.8

In the 3rd part, the classifier has been trained by using 13 MFCC with ten musical features. The maximum accuracy achieved through this feature is 98.3% in training using Bayesian Regularization and 88.6% in testing using Levenberg - Marquard and One-step Secant Back-propagation. Classification accuracy depends on the neural network function used to classify the signal. Figure 2 validates the results of the 13 MFCC with ten musical features.



Confusion Matrix

Output Class								
1	5 14.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
2	0 0.0%	5 14.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
3	0 0.0%	0 0.0%	5 14.3%	0 0.0%	0 0.0%	1 2.9%	0 0.0%	83.3% 16.7%
4	0 0.0%	0 0.0%	0 0.0%	5 14.3%	0 0.0%	0 0.0%	1 2.9%	83.3% 16.7%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 8.6%	0 0.0%	0 0.0%	100% 0.0%
6	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	4 11.4%	0 0.0%	100% 0.0%
7	0 0.0%	0 0.0%	0 0.0%	0 0.0%	2 5.7%	0 0.0%	4 11.4%	66.7% 33.3%
	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	60.0% 40.0%	80.0% 20.0%	80.0% 20.0%	88.6% 11.4%
	1	2	3	4	5	6	7	

Target Class

Fig. 2. Confusion matrix of testing results for the 13 MFCC with 10 musical features

5 Conclusion

The proposed algorithm outperforms the results of the different well-established algorithms of related works. In this paper, the authors classified seven different singers using ANN classifier. The experimental results support that the volume of the feature matrix is directly proportional with the classification accuracy, i.e., 13 MFCC with ten musical features shows higher accuracy than the only 13 MFCC features. Out of the four neural network functions, Bayesian Regularization function shows better performance in the singer identification process.

Acknowledgments. The authors would like to acknowledge the University Grant Commission (Ministry of Human Resource Development) for financial support in carrying out the research work.

References

1. Sleit, A., Serhan, S.: A histogram based speaker identification technique, Computer Science Department - King Abdulla II School for Information Technology, University of Jordan. IEEE (2008)
2. Harris, F.J.: On then use of windows for harmonic analysis with the discrete Fourier transform. Proc. IEEE **66**, 51–83 (1978)

3. Benesty, J., Chen, J., Huang, Y., Doclo, S.: Study of the wiener filter for noise reduction. In: Benesty, J., Makino, S., Chen, J. (eds.) Speech Enhancement. Signals and Communication Technology, pp. 9–41. Springer, Heidelberg (2005). https://doi.org/10.1007/3-540-27489-8_2
4. Farrell, K.R., et. al.: Speaker identification using neural tree networks. CAIP Center, Rutgers University Piscataway. IEEE (1994)
5. Wang, L., Minami, K.: Speaker identification by combining MFCC and phase information in noisy environments, Toyohashi University of Technology, Japan (2010). 978-1-4244-4296-6/10
6. Martinez, J., et al.: Speaker recognition using Mel frequency Cepstral Coefficients (MFCC) and Vector quantization (VQ) techniques. In: Electrical Communications and Computers (CONIELECOMP). IEEE (2012)
7. Durey, A.S., Clements, M.A.: Features for melody spotting using hidden Markov models. In: 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing, 13 May 2002, vol. 2, p. II-1765 (2002)
8. Akeroyd, M.A., Moore, B.C., Moore, G.A.: Melody recognition using three types of dichotic-pitch stimulus. J. Acoust. Soc. Am. **110**(3), 1498–1504 (2001)
9. Herrera, P., Amatriain, X., et al.: Towards instrument segmentation for music content description: a critical review of instrument classification techniques. In: International Symposium on Music Information Retrieval, vol. 9, October 2000
10. Herrera Boyer, P., et al.: Towards instrument segmentation for music content description a critical review of instrument classification techniques (2000)
11. Eronen, A.: Musical instrument recognition using ICA-based transform of features and discriminatively trained HMMs. In: 2003 Proceedings of Seventh International Symposium on Signal Processing and Its Applications, vol. 2, pp. 133–136, July 2003
12. Tzanetakis, G., Cook, P.: Musical genre classification of audio signals. IEEE Trans. Speech Audio Process. **10**(5), 293–302 (2002)
13. Genussov, M., Cohen, I.: Musical genre classification of audio signals using geometric methods. In: 2010 18th European Signal Processing Conference, pp. 497–501, August 2010
14. Xu, C., Maddage, N.C., et al.: Musical genre classification using support vector machines. In: 2003 Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP 2003), vol. 5, p. V-429, April 2003
15. Byrd, D., Crawford, T.: Problems of music information retrieval in the real world. Inf. Process. Manag. **38**(2), 249–272 (2002)
16. Hsu, J.L., Liu, C.C., Chen, A.L.: Discovering nontrivial repeating patterns in music data. IEEE Trans. Multimedia **3**(3), 311–325 (2001)
17. Kim, Y.E., Whitman, B.: Singer identification in popular music recordings using voice coding features. In: Proceedings of the 3rd International Conference on Music Information Retrieval, vol. 13, p. 17, October 2002
18. Liu, C.C., Huang, C.S.: A singer identification technique for content-based classification of MP3 music objects. In: Proceedings of the Eleventh International Conference on Information and Knowledge Management, pp. 438–445, November 2002
19. Zhang, T.: Automatic singer identification. In: 2003 Proceedings of International Conference on Multimedia and Expo, ICME, vol. 1, p. I-33 (2003). (Cat. No. 03TH8698)
20. Cai, W., Li, Q., Guan, X.: Automatic singer identification based on auditory features. In: 2011 Seventh International Conference on Natural Computation, vol. 3, pp. 1624–1628, July 2011
21. Kroher, N., Gómez, E.: Automatic singer identification for improvisational styles based on vibrato, timbre and statistical performance descriptors. In: ICMC (2014)
22. Bartsch, M.A., Wakefield, G.H.: Singing voice identification using spectral envelope estimation. IEEE Trans. Speech Audio Process. **12**(2), 100–109 (2004)

23. Li, T., Ogihara, M.: Toward intelligent music information retrieval. J. IEEE Trans. Multimedia **8**(3), 564–574 (2006)
24. Yuan, T.T., Cao, M.M.: Voice-recognition-based music retrieval system. J. Bull. Sci. Technol. **31**(7), 156–159 (2015)
25. Liu, F.Y., Wang, S.H., Zhang, Y.D.: Survey on deep belief network model and its applications. J. Comput. Eng. Appl. **54**(1), 11–18 (2017)
26. Gong, A., Jing, M.B., Dou, F.: Music mood classification method based on deep belief network and multi feature fusion. J. Comput. Syst. Appl. **26**(9), 158–164 (2017)
27. Lv, L.L.: Singing voice detection in songs based on clustering of MFCC. J. Comput. Knowl. Technol. **12**(31), 170–171 (2016)
28. Xu, X.X.: Study on gesture recognition based on PCA and DBN. J. Artif. Intell. **36**(13), 55–58 (2017)
29. Salakhutdinov, R., Hinton, G.: An efficient learning procedure for deep Boltzmann machines. J. Neural Comput. **24**(8), 1967–2006 (2012)
30. Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. J. Sci. **313**(5786), 504–507 (2006)
31. Yi, L., Ya, E.: Based on Gabor feature and deep belief network face recognition methods. J. Comput. Simul. **34**(11), 417–421 (2017)
32. Brosch, T., Tam, R.: Efficient training of convolutional deep belief networks in the frequency domain for application to high resolution 3D images. J. Neural Comput. **27**(1), 211–227 (2015)
33. Chen, W.H.: Emotion classification of music based on support vector machine. J. Softw. Eng. **19**(12), 20–23 (2016)
34. Tsai, W.H., Chieh, H.: Singer identification based on spoken data in voice characterization. IEEE Trans. Audio Speech Lang. Process. **20**(8), 2291–2300 (2012)
35. Tsai, W.H., Lin, H.P.: Background music removal based on cepstrum transformation for popular singer identification. IEEE Trans. Audio Speech Lang. Process. **19**(5), 1196–1205 (2010)
36. Mesaros, A., Virtanen, T., Klapuri, A.: Singer identification in polyphonic music using vocal separation and pattern recognition methods. In: Proceedings of International Conference on Music Information Retrieval, ISMIR, pp. 375–378 (2007)
37. Datta, A.K., Solanki, S.S., Sengupta, R., Chakraborty, S., Mahto, K., Patranabis, A.: Music information retrieval. In: Datta, A.K., Solanki, S.S., Sengupta, R., Chakraborty, S., Mahto, K., Patranabis, A. (eds.) Signal Analysis of Hindustani Classical Music. SCT, pp. 17–33. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-3959-1_2
38. Mahato, K., Hota, A., et al.: A study on artist similarity using projection pursuit and MFCC: identification of Gharana from Raga performance. Presented at Proceeding of the IEEE International Conference on Computing for Sustainable Global Development, New Delhi, India, pp. 647–653 (2014)
39. Kumar, P., Sharma, P.: Artificial neural network-a study. Int. J. Emerg. Eng. Res. Technol. **2** (2), 143–148 (2014)
40. Maind, S.B., Wank, P.: Research paper on basic of artificial neural network. Int. J. Recent and Innov. Trends Comput. Commun. **2**(1), 096–100 (2014)
41. Ratanpara, T., Patel, N.: Singer identification using perceptual features and cepstral coefficient form of an audio signal from Indian video songs. EURASIP J. Audio Speech Music Process. **2015**(1), 1–12 (2015)

An Optimized Selective Scale Space Based Fuzzy C-Means Model for Image Segmentation

Geetika Sharma(✉), Nandini Sethi, and Pooja Rana

Lovely Professional University, Phagwara, Punjab, India
sharmageetika64@gmail.com,
nandinisethi2104@gmail.com, poojarana.info@gmail.com

Abstract. Clustering is a widely used technique for segmentation of images. In this paper, a method is proposed for image segmentation, which used an Optimized Selective Scale Based Fuzzy c-Means approach. This approach is used to improve the quality of image segmentation. An algorithm named Fuzzy c-means (FCM) is used for data clustering in which an element can belong to multiple clusters. This algorithm results in the transformation of data elements in such a way that closer elements will come more closer and remaining elements will scatter farther. Genetic Algorithm is used as an optimization technique in this model. Genetic Algorithm is one of the commonly used methods to decide the optimal value of a criterion. The optimal value is determined by simulating the evolution of population until the best fitted individuals among the population is not encountered. It is obtained by mutation selection and crossover of individuals from the existing population.

Keywords: Clustering · Fuzzy C-Means · Image segmentation · Space scaling

1 Introduction

The process of Image Segmentation is defined as partitioning or dividing the image into multiple parts (also known as segments) in such a way that pixels in one segment are similar to each other with respect to some characteristic and dissimilar from the ones in other segments. Image segmentation has several applications like Machine vision, Medical imaging, Object detection and many more.

There are several algorithms and techniques available for image segmentation [1, 2]. The simplest and oldest method of image segmentation is Thresholding method. Other famous approaches are based on Detection of Edges, Histogram-based methods, Compression-based methods, Clustering methods and Region-growing methods.

Clustering is a widely used technique for segmentation of images [3]. Clustering further is classified into different categories. An algorithm named Fuzzy c-means (FCM) is used for data clustering in which an element can belong to multiple clusters [4]. Every element in the population has a set of membership values. These membership values determine the intensity of association among the element and the depicted cluster. These membership values are stored in a partition matrix.

For data refinement, a technique named Scale Space Filter is used in which the data elements are transformed to higher dimensions [5]. The objective of using this

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 402–410, 2019.
https://doi.org/10.1007/978-981-15-0108-1_37

techniques is to separate the data elements on the basis of proximity and other separability measures. The technique results in the transformation of data elements in such a way that closer elements will come more closer and remaining elements will scatter farther.

In the following approach, a method is designed for image segmentation in which Optimized Selective Scale Based Fuzzy c-Means approach is used. This approach is design to improve the quality of segmentation of images.

2 Methodology

2.1 Selective Scale Spaced FCM Method Based on Standard Deviation

Roy [6], given an algorithm in which they proved that for scale space transformation all parameters are not suitable. So in that algorithm, firstly the standard deviation of each parameter is computed and the value is compared with a threshold value. Only for lower standard deviation value scale space transformation is applied but for the standard deviation value greater than the threshold value scale space filter technique will not be applied on those parameters [23].

Multiple functions are present which can be used for scale space filters. Among them, Polynomial and Gaussian functions are most popular ones. In this methodology, Gaussian transformation function is applied. This function is defined as follows:

$$f(x,\sigma) = \frac{1}{\left(\sigma\sqrt{2\pi}\right)^2} e^{-\frac{E^2(x-x_i)}{2\sigma^2}} \tag{1}$$

Suppose n are the number of objects which are needed to be clustered and the feature vector is represented as S = {p_1, p_2,….., p_n}. Let m are the number of attributes that each point has and C are the number of cluster, C = {k_1, k_2,…., k_c}. Cluster Validity is measured by using Xie-Beni index [7]. We can define this index by using Eq. 2.

$$XB(P,C) = \frac{v_x(P,C)}{n\,d(C)} = \frac{\sum_{i=1}^{C}\sum_{j=1}^{n}\mu_{ij}^2 E^2(c_i,x_j)}{n\,min_{i\neq j}E^2(c_i,c_j)} \tag{2}$$

The value of member function is given by Eq. 3.

$$\mu_{i,k} = \frac{1}{\sum_{j=1}^{C}\left(\frac{E(c_i-x_k)}{E\left(c_j-x_k\right)}\right)^{\frac{2}{m-1}}}, 1 \le i \le C,\ 1 \le n \tag{3}$$

The Center is updated using Eq. 4.

$$c_i = \frac{\sum_{k=1}^{n} (\mu_{ik})^m x_k}{\sum_{k=1}^{n} (u_{ik})^m} \tag{4}$$

The generic fuzzy c-means algorithm can be described as follows:

```
Declare the Matrix as $F_{n \times C} S$.
Declare ObjValue as Real
Initialize the Cluster Center, C
while $ObjValue \leq Benchmark$ do
Populate Fuzzy Partition Matrix using (3)
Update Cluster Center using (4)
Update the ObjValue using ObjectiveFunction(C, $F_{n \times C}$)
end while
```

The fuzzy membership has to be:

$$\sum_{k=1}^{n} \sum_{i=1}^{C} \mu_{ik} = n$$

where $0 \leq \mu_{ik} \leq 1, i = 1, 2, \ldots\ldots, C\, and\, j = 1, 2, \ldots\ldots, n.$

The Selective Scale Spaced FCM algorithm based on Standard deviation can be described as follows:

```
Choose the image and convert it into gray scale image.
Get the Standard Deviation value for all parameters and save it in sd.
if $sd < sdThreshold$ then
    Apply Gaussian Scale Space using (1)
end if
Initialize the Cluster Center, C
while $ObjValue \leq Benchmark$ do
    Populate Fuzzy Partition Matrix using (3)
    Update Cluster Center using (4)
    Add value of Xie Beni Function to ObjValue using (2)
end while
```

2.2 Proposed Method

In proposed method, an optimized approach of selective scale based fuzzy c-means algorithm is presented. There are several optimization techniques. In this model, Genetic Algorithm is applied.

Genetic Algorithm is one of the commonly method used to decide the optimal value of a criterion [8]. The optimal value is determined by simulating the evolution of population until we will not encounter the best fitted individuals among the population. It is obtained by mutation selection and crossover of individuals from the existing population.

Genetic Algorithm consists of following essential data:

(a) Genotype: It is resulted segmented image which is considered as an individual.
(b) Initial Population: It is a set of individuals characterized on the basis of genotype.
(c) Fitness Function: A function which is used to quantify the fitness of individual.
(d) Operators on genotype: There are three different operators. These are mutation, selection and cross-over.
(e) Stopping Criterion: It allows to stop this evolution process.

The proposed algorithm is given as follows:

```
1:  Select the image and convert it into gray scale image.
2:  Get the Standard Deviation value for all parameters and save it in $sd$.
3:  if $sd < sdThreshold$ then
4.  Apply Gaussian Scale Space using (1)
5:  end if
6:  Define the initial population and compute the fitness value
7:  Apply mutation and cross-over operators
8:  Apply Selection operator
9:  Evaluate and Go to Step 7 only if stopping criterion is not satisfied
10: Initialize the Cluster Center, C
11: while $ObjValue \leq Benchmark$ do
12:   Populate Fuzzy Partition Matrix using (3)
13:   Update Cluster Center using (4)
14:   Add value of Xie Beni Function to $ObjValue$ using (2)
15: end while
```

3 Results

The designed algorithm is applied on the MRI brain images. The results of designed algorithm are compared with existing approach using AE (Average Error), NED (Normalized Euclidean Distance) and SNR (Signal to Noise Ratio) [9]. I(i, j) is considered here as input image where i = 1, 2, 3, 4 …, M and j = 1, 2, 3, 4 …, N and O(i, j) is considered as output image, then the criterion is calculated as follows:

(i) Average Error (AE):

$$AE = \frac{1}{M \cdot N}\sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{(I(i,j) - O(i,j))^2} \qquad (5)$$

(ii) Normalized Euclidean Distance (NED):

$$NED = \frac{1}{M \cdot N}\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i,j) - O(i,j))^2} \tag{6}$$

(iii) Signal to Noise Ratio (SNR):

$$SNR = 10\log_{10}\left[\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} I^2(i,j)}{\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i\cdot j) - O(i,j))^2}\right] \tag{7}$$

The comparison between proposed and existing algorithm is shown in Table 1.

Table 1. Results of proposed method in comparison to Selective Scale Spaced FCM.

Image	Method	AE	NED	SNR
Image 1	Optimized Scale Space FCM	3.0568	0.0181	3.2904
	Selective Scale Space FCM	3.1363	0.0260	3.9102
Image 2	Optimized Scale Space FCM	4.3390	0.0226	0.4754
	Selective Scale Space FCM	4.3599	0.0395	0.8709
Image 3	Optimized Scale Space FCM	4.7892	0.0250	2.0683
	Selective Scale Space FCM	4.7903	0.0352	1.0926
Image 4	Optimized Scale Space FCM	3.3162	0.0197	0.1721
	Selective Scale Space FCM	3.3435	0.0282	0.6321
Image 5	Optimized Scale Space FCM	3.2115	0.0185	3.0884
	Selective Scale Space FCM	3.3878	0.0261	2.9163
Image 6	Optimized Scale Space FCM	5.5819	0.0273	2.6966
	Selective Scale Space FCM	5.6282	0.0393	1.7749
Image 7	Optimized Scale Space FCM	3.5480	0.0222	2.2132
	Selective Scale Space FCM	3.6833	0.0391	1.0872
Image 8	Optimized Scale Space FCM	2.1031	0.0149	2.4182
	Selective Scale Space FCM	2.1524	0.0216	1.1560
Image 9	Optimized Scale Space FCM	3.4872	0.0191	3.3245
	Selective Scale Space FCM	3.5340	0.0283	4.2031
Image 10	Optimized Scale Space FCM	5.0069	0.0265	3.8600
	Selective Scale Space FCM	5.0349	0.0381	4.3960

The values shown in Table 1 are also plotted in graphs to compare the values of AE, NED and SNR for both the techniques. In graph, the Optimized Scale Spaced FCM is represented with black line and the Selective Scale Spaced FCM is represented with blue line. The graph for comparison according to AE, NED and SNR values are shown in Figs. 1, 2 and 3 respectively.

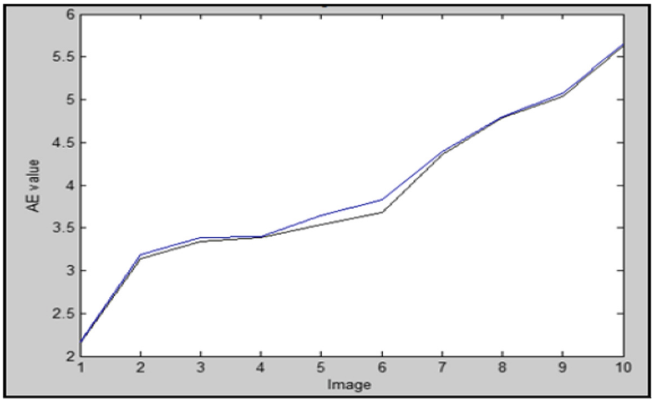


Fig. 1. Comparison based on AE values (Color figure online)

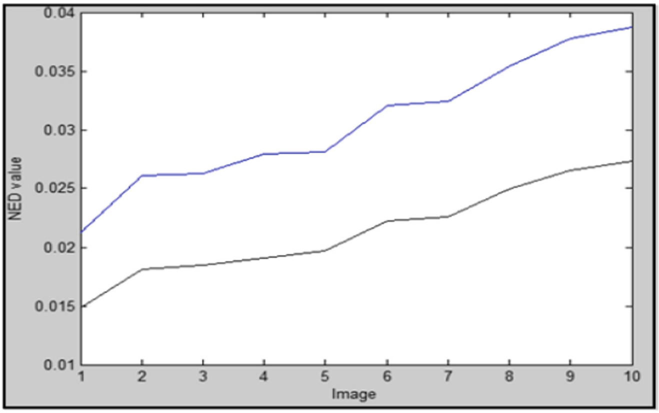


Fig. 2. Comparison based on NED values (Color figure online)

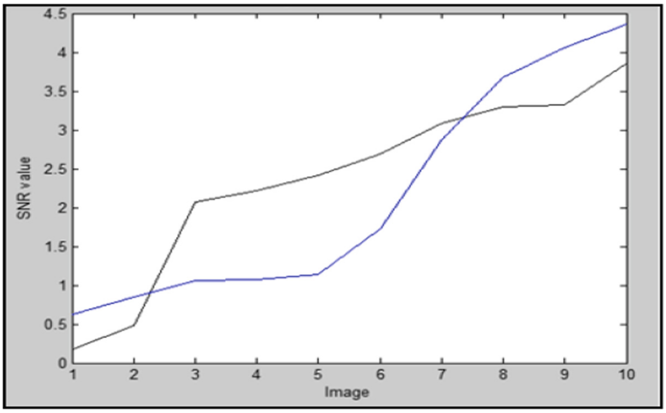


Fig. 3. Comparison based on SNR values (Color figure online)

The MR images after segmentation using proposed method are shown in Fig. 4.

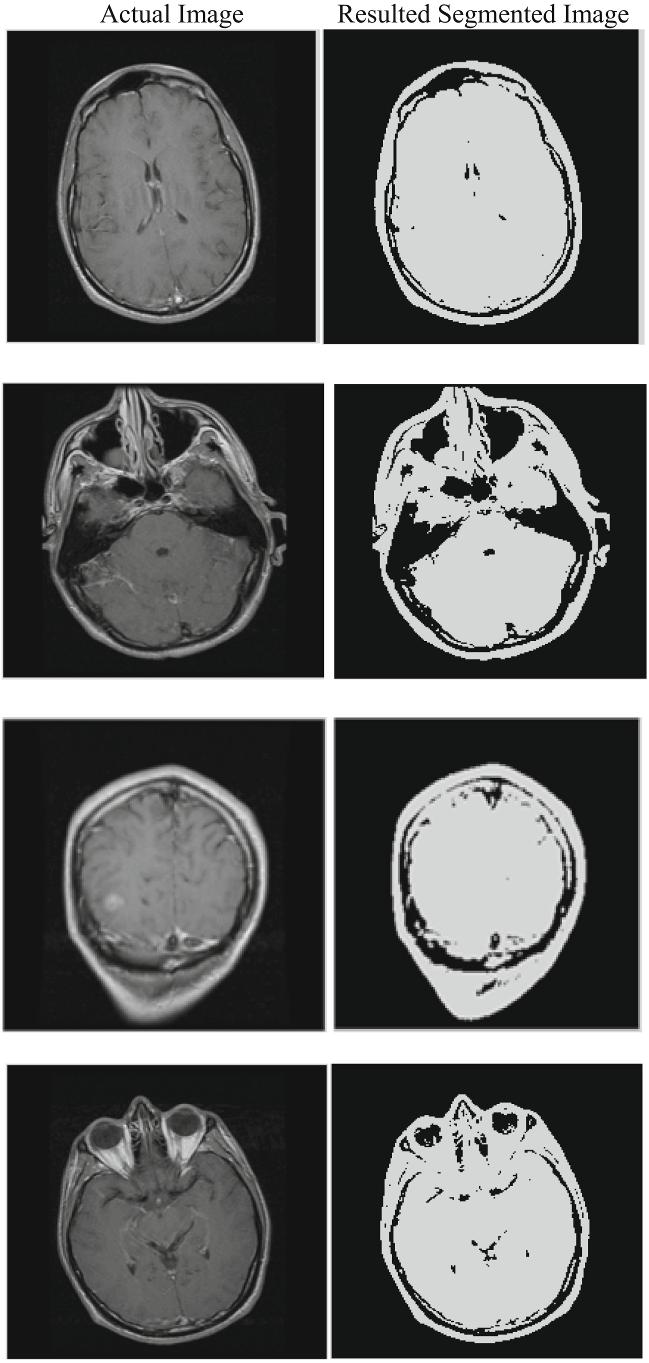


Fig. 4. Results of segmentation with optimized scale space FCM method.

4 Conclusion

In this paper, an algorithm named Fuzzy c-Means is surveyed along with its variation i.e. selective scale space FCM based on standard deviation. An Optimized scale spaced FCM method is proposed in this paper. This proposed algorithm achieves better segmentation results as compared to the existing method.

There is scope of betterment of this algorithm. Firstly, the computational time of this method is higher than the original Fuzzy C-Means algorithm. So, one can try to reduce this computational time in future work. Secondly, in this method Genetic Algorithm is used for optimization purpose. Other optimization techniques can also be used and the results can be compared.

References

1. Haraliek, R.M., Shapiro, L.G.: Image segmentation techniques. CVGIP **29**, 100–132 (1985)
2. Pal, N.R., Pal, S.K.: A review on image segmentation techniques. Pattern Recogn. **26**, 1277–1294 (1993)
3. Wu, Z., Leahy, R.: An optimal graph theoretic approach to data clustering: theory and it's applications to image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **15**(11), 1101–1113 (1993)
4. Iain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM Comput. Surv. **31**(3), 264–323 (1999)
5. Leung, Y., Zhang, J.S., Xu, Z.B.: Clustering by scale-space filtering. IEEE Trans. Pattern Anal. Mach. Intell. **22**(12), 1396–1410 (2000)
6. Roy, P., Mandal, J.K.: A novel selective scale space based fuzzy C-means model for spatial clustering. Procedia Tecnol. **10**, 596–603 (2013)
7. Xie, X.L., Beni, G.: A validity measure for fuzzy clustering. IEEE Trans. Pattern Anal. Mach. Intell. **13**(8), 841–847 (1991)
8. Chabrier, S., Rosenberger, C., Emile, B., Laurent, H.: Optimization based image segmentation by genetic algorithms (2008)
9. Upadhyay, P., Chhabra, J.K.: Modified self organizing feature map neural network (MSOFM NN) based gray image segmentation. Procedia Comput. Sci. **54**, 671–675 (2015)
10. Ahmed, M.N., Yamany, S.M., Mohamed, N., Farag, A.A., Moriarty, T.: A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data. IEEE Trans. Med. Imaging **21**(3), 193–199 (2002)
11. Gulhane, A., Paikrao, P.L., Chaudhari, D.S.: A review of image data clustering techniques. Int. J. Soft Comput. Eng. **2**(1), 212–215 (2012)
12. Hall, L.O., Ozyurt, I.B., Bezdek, J.C.: Clustering with a genetically optimized approach. IEEE Trans. Evol. Comput. **3**(2), 103–112 (1999)
13. Hasanzadeh, M., Kasaei, S.: Multispectral brain MRI segmentation using genetic fuzzy systems, pp. 2–5 (2007)
14. Sharma, M., Mukherjee, S.: Fuzzy C-means, ANFIS and genetic algorithm for segmenting astrocytoma – a type of brain tumor **3**(1) (2014)
15. Zhu, L., Chung, F.L., Wang, S.: Generalized fuzzy c-means clustering algorithm with improved fuzzy partitions. IEEE Trans. Syst. Man Cybern. Part B (Cybern.) **39**, 578–591 (2009)

16. Shasidhar, M., Raja, V.S., Kumar, B.V.: MRI brain image segmentation using modified fuzzy C-means clustering algorithm, pp. 473–478 (2011)
17. Bezdek, J.C., Hathaway, R.J.: Optimization of fuzzy clustering criteria using genetic algorithms, pp. 589–594 (1993)
18. Pal, N.R., Bezdek, J.C.: On cluster validity for fuzzy c-means model. IEEE Trans. Fuzzy Syst **3**(3), 370–379 (1995)
19. Clark, M.C., Hall, L., Goldgof, D.B., Clarke, L.P., Velthuizen, R.P., Silbiger, M.S.: MRI segmentation using fuzzy clustering techniques. IEEE Eng. Med. Biol. **13**, 730–742 (1994)
20. Kaus, M.R., Warfield, S.K.: Automated segmentation of MR images of brain tumors, pp. 586–591 (2001)
21. Maulik, U., Bandyopadhyay, S.: Fuzzy partitioning using a real-coded variable-length genetic algorithm for pixel classification. IEEE Trans. Geosci. Remote Sens. **41**(5), 1075–1081 (2003)
22. Chen, D., Li, X., Cui, D.-W.: An adaptive cluster validity index for the fuzzy c-means. Int. J. Comput. Sci. Netw. Secur. **7**(2), 146–156 (2007)
23. Monica, Singh, S.K., Agrawal, P., Madaan, V.: Breast cancer diagnosis using digital image segmentation techniques. Indian J. Sci. Technol. **9**(28), 1–5 (2016)

Diagnosis of Psychopathic Personality Disorder with Speech Patterns

Deepti Jain[1](✉), Sandhya Arora[2], and C. K. Jha[1]

[1] Department of Computer Science, AIM and ACT, Banasthali Vidyapith, Banasthali, India
deepti.kheterpal@gmail.com, ckjhal@gmail.com

[2] Department of Computer Engineering, Cummins College of Engineering, Pune, India
sandhya.arora@cumminscollege.in

Abstract. This paper proposes a novel approach to recognize psychopaths. Among various personality disorders psychopathic personality disorder is very dangerous as psychopaths behave like normal human beings but they lack conscience and guilt. Not performing at emotional level they commit heinous and brutal crimes with ease and without any remorse. Thus identifying these people is crucial for society. In this context this paper solves the problem in two ways: First, by studying the underlying psychology of psychopathic behavior and second proposing a novel approach to effectively recognize psychopaths by analyzing their speech patterns using Neural network with promising results.

Keywords: Psychopath · Personality disorders · Speech patterns

1 Introduction

Over the last few decades, crime rates have been increased. More and more criminal cases are being registered and come into picture with the advancement of media. Society was never crime free completely but what is more alarming is increasing number of brutal and in-human crimes. Latest Crime in limelight in India, Guru Gram is Ryan International School crime, in which student of class eleven brutally murdered class two boy just to postpone the exam [1]. According to his relatives his behavior was normal after the event. Another most highlighted crime in India was 16 Dec 2014 gang rape case. Four accused involved in the gang rape case show no remorse and reflected signs of psychopathic tendencies [2]. The Noida serial murders (also known as the Nithari Kand) have been described as among the most notorious crimes ever perpetrated in India. During the years 2005 and 2006, more than 16 children and young women from labor class went missing and found killed in Nithari village outside Delhi. Their skeletal were found in the house of Moninder Singh Pandher. He and his servant Surindra Koli were the main accused [3]. Many cases can be taken from history as well. People such as Joanna Dennehy, a 31-year-old British woman who killed three men in 2013 and who the year before had been diagnosed with a psychopathic personality disorder, or Ted Bundy, the American serial killer who is believed to have murdered at least 30 people [4].

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 411–421, 2019.
https://doi.org/10.1007/978-981-15-0108-1_38

All the above cases show the common pattern. All the crimes are heinous, brutal and purposeful in nature. The Criminal is very charismatic, magnetic and attractive and have no emotions. The brutality of the crime reveals the remorseless nature of the criminal which is in-human and very rare among humans. These criminals fall into the category of personality disordered people known as psychopaths as they lack social emotions like empathy, remorse and guilt and involve in aggressive behaviors [5].

These unique but horrible cases have given rise to multiple questions. What is the reason behind these brutal crimes? Who are psychopaths? Why these crimes are so in-human? Why psychopath's brains are not able to feel empathy towards other human beings? There is ample research in psychiatry and psychology but none of the questions and techniques is addressed by computer science community to recognize psychopaths based on their speech patterns. To answer these questions and get deeper insight of personality disorder, we analyzed the psychopath demographics. Based on obtained insights, we propose features that can differentiate psychopath criminals from non-psychopaths ones. Furthermore, we believe our contributions could lead to generation of tools or treatments that can have a significant impact on reducing the crimes committed by people having psychopathic disorder.

The organization of the paper is organized as follows; Sect. 2 presents related literature. We then discuss personality disorders especially psychopathic disorder in detail in Sect. 3. In Sect. 4 we discuss the diverse set of features to study psychopaths. Section 5 covers the proposed methodology and flowchart in detail. We discuss applicability, limitations and other observations in Sects. 6 and 7 concludes the paper with future research.

2 Related Work

Investigating psychopathy has been popular in US since 1980 when great Psychologist Robert-Hare invented Psychopathy Checklist named as PCL-R (Psychopathy Checklist-Revised) [6]. This checklist is now a standard tool used by researchers, forensic psychologists and judicial system to identify psychopathic behaviors [7]. Among Indian researchers this area is still unaddressed [8]. The Brutality and in-human nature of crime yet attractive, charming personality of psychopaths have drawn some of the researchers to study psychopaths from different perspectives. Table 1 presents related work in this arena.

Researchers in the area of Computer Science have tried to identify psychopathic behavior based on Socio-linguistic features, twitter data and MRI data. Same set of features have been studied in Psychiatry and Psychology. In this paper psychopathic tendencies have been analyzed from a different perspective. Our work is the first one to characterize psychopaths based on their speech patterns.

Table 1. Related work

Paper title	Features identified	Technique applied	Results
Wald [9]	Socio-linguistic features like Agreeableness', Conscientiousness, Extroversion, Neuroticism and openness.	Data mining, ensemble learning, SelectRusBoost, four classification learners, four feature selection techniques.	AUC (Area Under Curve) value of 0.736 is achieved
Pearce [10]	Brain MRI	SVM, decision trees, k-nearest neighbor, ensemble methods(to improve the accuracy)	SVM with 73% accuracy in identifying psychopathic patients
Sato [11]	Brain MRI	SVM, MDLA	Achieved 80% accuracy with SVM and MLDA in identification of psychopaths
Steele [12]	Data from PCL-R (measure of psychopathy), TBI questionnaire, Intelligent quotient, MRI scans	SVM	Classification of high and low psychopathic traits achieved 69.23% overall accuracy
Sumner [13]	Linguistic features like frequency of words in pre-defined categories, data from Ten Item Personality Inventory	SVM, Random Forest, J48, Naive Bayes Classifier.	Machine learning is an effective tool to predict Narcissism, Machiavellian, and psychopathic traits.
Eisenbarth [14]	PPI-R items (Psychopathic Personality Inventory- Revised)	Genetic Algorithm	Genetic Algorithm proved to be a new instrument to measure psychopathic traits efficiently.
Hastings [15]	Emotion, Faces, expression	PCL-SV, Facial affect recognition emotion task, Scoring of the facial affect recognition task	Results showed that psychopaths have difficulty in identifying happy and sad emotions.

3 Understanding Personality Disorders and Psychopathy

Personality, defined psychologically, is enduring patterns of thoughts, feelings, behavioural and mental traits that distinguish humans. Hence, personality disorders are defined by long lasting and rigid patterns of experiences and behaviours. Due to the inflexibility and rigidity of these acquired patterns, people having these disorders cause serious problems to the society. Personality disorders are a class of mental disorders. These patterns are inflexible, develop early and are associated with relationship problems. Those diagnosed with a personality disorder are not able to cope with everyday stress and experience difficulties in interpersonal functioning, cognition &

emotiveness. They have poor impulse control. Some of personality disorders are listed as Narcissist, Borderline personality Disorder, Anti Social personality, Psychopaths and Serial killers. Brief description of these disorders is as follows.

Narcissists: Narcissistic Personality Disorder INVOLVES around ego-centric and arrogant behaviour. It is a mental condition in which people seek self admiration and attention, have fragile ego, a lack of empathy for other people. People having narcissistic personality disorder are not able to make meaningful relationships at work and in personal life. Related Personality Disorders: Antisocial, Borderline, and Histrionic. Narcissism is a less extreme version of Narcissistic Personality Disorder. Related personality traits include: Psychopath, Machiavellianism.

Borderline Personality Disorder (BPD): In Borderline Personality Disorder (or emotionally unstable), there is a pattern of intense but unstable relationships, emotional instability, outbursts of anger and violence (especially in response to criticism). People with BPD are involved in self-harm and suicide.

Antisocial Personality Disorder: This is a mental condition in which the person disregards society rules and obligations and shows no regard for right and wrong. The person having this disorder uses wit and charm to manipulate others for personal gain, lacks guilt and concern for others, and fails to learn from experience.

Psychopaths: Psychopathic behaviour and antisocial personality disorders both are closely related to crime [16]. People having psychopathic disorder have very little or no'conscience', lack remorse or guilt, and are profoundly selfish. They are known for impulsivity, sensation seeking, predatory behaviour and need for control thus share common traits with serial killers. Studies suggest that [17] psychopaths have defect in the amygdale part of the brain, which is responsible for emotions and impulse control.

3.1 Challenges in Recognizing Psychopaths

Out of all these disorders, psychopathic disorder is most dangerous [7]. The main challenge in recognizing the psychopaths and other personality disordered people is that they are very attractive, charming and good communicators, but at the same time they can be highly manipulative because they manipulate their victim at psychological level [18]. Research by [5, 19 and 20] concludes that psychopaths lie on the continuum of wide range of other mental disorders like antisocial, depression, sociopath disorder etc. which makes it difficult to identify them. Study conducted by [20] reveals the fact that psychopathic personality disorder covers a wide range of symptoms and there is no exclusively defined diagnostic criterion for them [21].

Thus combination of personality and mental disorders make psychopaths diagnosis a difficult task. Moreover they are very intelligent and psychopaths with high IQ can even manipulate personality tests [22]. They look normal as other people; behave normally in a crowd so they gel very well in the society [23]. They may be around us in any relation, like they may be our co-workers, guards, taxi drivers and any close friend [24, 25]. Due to these complex personality traits, recognition of psychopaths is highly challenging. They can commit crime any time, anybody can be their victim as they get impulsive and aggressive [5] without giving any cue or any reason.

4 Feature Aggregation

Speech: Speech is communication of our thoughts to other person. Choice of words provides significant insight about one's psychology (Gottschalk, Pennebaker) and various mental disorders (Graybeal, Pennebaker). Research [26] suggests that language is related to the frontal lobe part of the brain. Studies [17] suggest that psychopathic offenders have dysfunctional frontal lobe. As per conclusion of [27, 28] psychopaths are more like to use function words like “I”, “me”, “my” etc. and these function words are produced unconsciously. They are more likely to use past tense in describing their crime scene [27]. A study conducted by [29] concludes that psychopath who is skilled at faking emotions without actually feeling them. So, this feature plays very important role to recognize the most charming and attractive predators like psychopaths because speech is produced by unconscious part of the brain [26]. In our present study we have analysed psychopath's language closely by watching his speech patterns.

Callousness and Lack of Empathy: Research by Decety [30, 31] suggests that psychopaths have callous and unemotional traits. They can't feel emotions when they see other people in pain. They only imitate emotions.

Egocentric and Poor Behaviour Controls: They have very fragile ego, which can provoke aggressive behaviour in them [31].

Lack of Remorse/Guilt: Studies concluded by [27] reveals that while describing crime they frequently use words like “because”, “since”, “so that”, “umh” and “hmm” words and they give justification of their crime which indicates that they do not have guilt or feel remorse of their actions.

Amygdale Dysfunction: Research by [17] states that brain MRI of psychopaths are different from other people. The part of brain which is responsible for human emotions regulations is Amygdale. Psychopaths have less activity in their amygdale thus have impaired human emotion recognition.

5 Proposed Methodology

The basic idea of proposing this model is to facilitate the task of identifying and recognizing psychopaths. Forensic psychologists have to go through tedious process of analysing psychopaths interviews. There may be chances of mistake due to human intervention. The proposed algorithm may assist forensic psychologists by automatically converting the video to text and then analyse the text by searching function words, pronouns and past tense used by the psychopaths while describing their crime.

This part involves study of speech patterns. Due to sensitivity of the subject and scarcity of data in this area, we have collected some publicly available interviews. All collected videos are first pre-processed and converted into text. Then text is analysed for occurrence of function words like “I”, “We” and emotion words like “calm”, “tensed”, and “happy”, words having past tense and conjunction words. These parameters are summarized in Table 2.

Table 2. Parameters for speech analysis

S. No.	Words	S. No	Words
1	I	16	satisfied
2	we	17	nervous
3	me	18	relaxed
4	my	19	content
5	because	20	worried
6	since	21	confused
7	had	22	frightened
8	Went	23	secure
9	was	24	indecisive
10	sad	25	steady
11	calm	26	pleasant
12	tensed	27	failure
13	happy	28	cool
14	comfortable	29	inadequate
15	upset	30	strained

A category of the words [27, 28] is summarized in Table 3. Words related to anxiety have been taken form STAI (State-Trait-Anxiety-Inventory) form [32]. Flow-chart of the proposed approach is explained in Fig. 1.

Table 3. Words category

S. No.	Language feature	Words	Meaning
1	1st person singular Subjective personal pronoun	I, me, my	Self -centered attitude
2	1st person singular objective personal pro noun	we	Relates to family, friends and society
3	Subordinate Conjunction	Because, Since	Reflects justification of the crime
4	Past tense words	Had, went, was	Psychological detachment from the crime incident
5	Words that states anxiety	Sad, calm, tensed, worried, nervous, frightened	The person is in state of anxiety.
6	Positive words	Happy, comfortable	Positive words

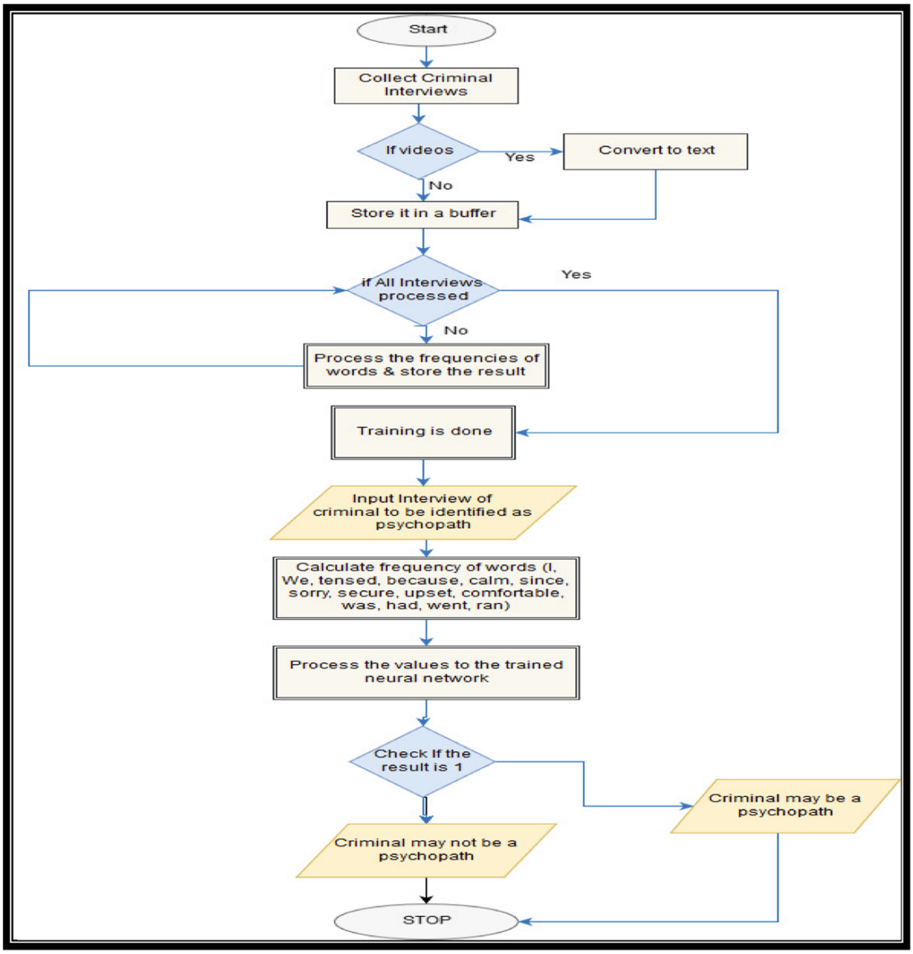

Start
Collect Criminal Interviews
If videos
Yes
Convert to text
No
Store it in a buffer
if All Interviews processed
Yes
No
Process the frequencies of words & store the result
Training is done
Input Interview of criminal to be identified as psychopath
Calculate frequency of words (I, We, tensed, because, calm, since, sorry, secure, upset, comfortable, was, had, went, ran)
Process the values to the trained neural network
Check If the result is 1
Criminal may be a psychopath
Criminal may not be a psychopath
STOP

Fig. 1. Flowchart of proposed algorithm for text analysis

6 Simulation Results

Simulation of Speech Patterns

Simulation of the algorithm for analysing the speech patterns of psychopaths consist of two steps. First the video is converted into text and text file is stored in notepad. From the text file we have calculated total number of words spoken by psychopath to be identified. To identify the patterns from words we have taken parameters from Table 2. The Text file is then passed to the program which has been written in Python Language. We calculated total words spoken by psychopaths and non- psychopaths. We then calculated different category words and observed that psychopaths used more Personal

Pronouns than non-psychopaths. The corresponding meaning [12, 40] of using different category of words is explained in Table 3. The percentage for singular subjective personal pronoun comes out to be 4.97% while for non psychopaths it was 0.70%. The results can be seen in Table 4. As we can see psychopaths were also using more past tense words than non psychopaths while describing the crime, it shows the psychological detachment from the incident. This category showed higher percentage 1.97%. Psychopaths tried to give justification of their crime by Subordinate Conjunction like because and since. In our approach we have used some of the annotated dataset to train the classifier, while the remaining dataset being used for result validation.

Table 4. Results obtained by calculating the frequency of different categories of words spoken by psychopath's criminals and non psychopaths.

	Psychopaths			Non psychopaths		
Language feature	Total Words	Obs.	%age	Total Words	Obs.	% age
1st person singular subjective personal pronoun **I**	18965	943	4.97	19276	136	0.70
1st person singular objective personal pronoun **We**	18965	40	0.21	19276	36	0.18
Subordinate conjunction (because, since)	18965	46	0.24	19276	40	0.21
Past tense (had, went, was)	18965	375	1.97	19276	129	0.66
Emotion words (comfortable, tensed, sorry, secure, upset, calm)	18965	1	0.005	19276	15	0.07

After getting results, we compared previous results produced by LIWC(Linguistic Enquiry Word Count). This software is used by psychologists to find different word categories like positive and negative words, pronouns, verbs, conjunctions etc. However it ignores sarcasms, idioms, irony and language contexts. We tried to train our model to find contexts in the speech. We were able to find certain patterns in speech with the help of machine learning and thus overcome the drawback of LIWC which is based on word's dictionary. Overall percentage of different category of words can be seen in Fig. 2.

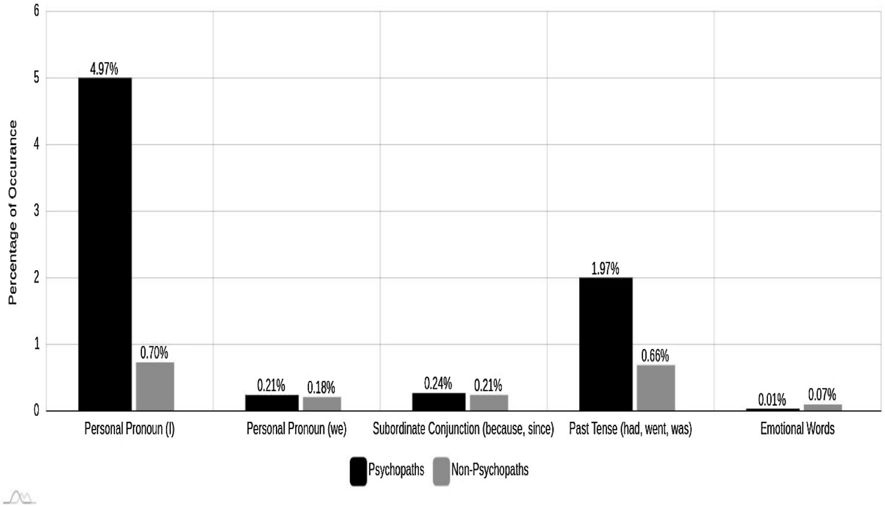


Fig. 2. Percentage of different words category

7 Discussion, Conclusion and Future Scope

In this Paper we have tried to study and analyze psychopathic behavior from different perspective. From the results we conclude that we can identify psychopaths by analyzing their speech patterns with the help of neural network, which proves that machine learning, can become useful tool in psychopath identification.

This study has practical significance as well because it suggests that characteristics unique to the psychopath can be found through speech and textual analysis. The PCL-R (Psychopathy Checklist-revised) is a valid and reliable measure of psychopathy, but it has one significant drawback—it requires lengthy interviews and can only be used on convicts (Hare and Neumann, 2006). Therefore, linguistic analysis could be useful in law enforcement because the speech sample could be obtained from suspects who are not prisoners, which could help identify those with psychopathic tendencies. Linguistic analysis could be potentially helpful in identifying psychopaths in the workplace.

For future work, psychopath personality disorder can be studied thoroughly and the model may be trained to identify subtle patterns so that we can recognize psychopaths before they commit crime. That would help society to be proactive and reduce the crime rates.

Acknowledgement. I would like to thank Prashant for helping me in executing code in python.

References

1. India Today News. http://indiatoday.intoday.in/story/ryan-murder-case-16-year-old-cbi-suspect-investigation/1/1087211.html
2. Wikipedia, the Free Encyclopedia. https://en.wikipedia.org/wiki/2012_Delhi_gang_rape
3. Outlook, the Fully Loaded Magazine. https://www.outlookindia.com/magazine/story/devil-in-the-flesh/233671
4. The Independent News. https://www.independent.co.uk/news/long_reads/serial-killers-shipman-brady-hindley-berkowitz-sutcliffe-dennehy-ted-bundy-a8021181.html
5. Glenn, L.A., Raine, A.: Psychopathy and instrumental aggression: evolutionary, neurobiological, and legal perspectives. Int. J. Law Psychiatry **32**, 253–258 (2009)
6. Wikipedia, the Free Encyclopedia, Psychopathy Checklist
7. Kiehl, A.K., Hoffman, B.M.: The criminal psychopath: history, neuroscience, treatment and economics. J. Jurimetr. **51**, 355–397 (2014)
8. Sepaha, P.M.: Psychopaths: an unrevealed area in Indian judicial system. Nirma Univ. Law J. 4(1) (2014)
9. Wald, R., Khoshgoftaar, T., Sumner, C.: Using twitter content to predict psychopathy. In: 11th International conference on Machine learning and applications, Florida, USA. IEEE (2012)
10. Pearce, M., Sung, I.: Classifying psychopathy patients using machine learning methods on magnetic resonance imaging (MRI) Data (2015). Manuscript
11. Sato, R.J., de Oliveira, R.: Identification of psychopathic individuals using pattern classification of MRI images. J. Soc. Neurosci. **6**(5–6), 627–639 (2011)
12. Steele, V.R., Rao, V., Calhoun, V.D., Kiehl, K.A.: Machine learning of structural magnetic resonance imaging predicts psychopathic traits in adolescent offenders. J. Neuroimage **145** (Pt B), 265–273 (2015)
13. Sumner, C., Byers, A., Boochever, R., Park, G.J.: Predicting dark triad personality traits from twitter usage and a linguistic analysis of tweets. In: 11th International Conference on Machine learning and applications. IEEE (2012)
14. Eisenbarth, H., Lilienfeld, S.O., Yarkoni, T.: Using a genetic algorithm to abbreviate the Psychopathic Personality Inventory-Revised (PPI-R). J. Psychol. Assess. **27**(1), 194–202 (2015)
15. Hastings, M.E., Tangney, J.P., Stuewig, J.: Psychopathy and identification of facial expressions of emotion. Pers. Individ. Differ. **44**(7), 1474–1483 (2008)
16. Everett, C.D.: Antisocial personality disorder vs. psychopathy. A thesis, Auburn, Alabama (2006)
17. Shamay-Tsoory, S.G., Harar, H., Aharon-Peretzc, J., Levkovitzb, Y.: The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. Res. Rep. J. Cortex J. Devoted Study Nerv. Syst. Behav. **46**(5), 668–677 (2010)
18. Wallisch, P.: Psychopaths in our midst what you should know, Elsevier connect (2004)
19. Oguz, M., Tutuncu, R., Ates, A.: The relationship between co morbid psychiatric illnesses and psychopathy levels on male individuals with antisocial personality disorder in the turkish community. J. Eur. Psychiatry **30**, 406 (2014)
20. Hart, S.D., Cook, A.N.: Current issues in the assessment and diagnosis of psychopathy. J. Neuropsychiatry **2**(6), 497 (2012)
21. Johnson, S.C., Elbogen, E.B.: Personality disorders at the interface of psychiatry and the law: legal use and clinical classification. J. Clin. Neurosci. **15**(2), 203–211 (2013)

22. Bate, C., Boduszek, D.: Psychopathy, intelligence and emotional responding in a non-forensic sample: an experimental investigation. Res. Art. (2014)
23. ChangingMinds. http://changingminds.org/explanations/personality/disorders/psychopath.htm
24. Babiak, P., Hare, R.D.: Snakes in Suits When Psychopaths Go To Work. HarperCollins, New York (2006)
25. Babiak, P.: When psychopaths go to work: a case study of an industrial psychopath. J. Appl. Psychol. **44**(2), 171–188 (1995)
26. Chung, C., Pennebaker, J.: The psychological function of function words. J. Soc. Commun. **1**, 343–359 (2011)
27. Hancock, J.T., Woodworth, M.T., Porter, S.: Hungry like the wolf: a word-pattern analysis of the language of psychopaths. J. Leg. Criminol. Psychol. **18**(1), 102–114 (2013)
28. Morrow, R.: Psychopathic storytelling: the effect of valence on self and time in psychopathic language use. Honors Thesis Presented to the College of Agriculture and Life Sciences, Social Science Program, Research Advisor: Jeffrey Hancock (2008)
29. Porter, S., Brinke, T.L., Baker, A., Wallace, B.: Would I lie to you? "leakage" in deceptive facial expressions relates to psychopathy and emotional intelligence. J. Pers. Individ. Differ. **51**, 133–137 (2011)
30. Decety, J., Chen, C.: An fMRI study of affective perspective taking in individuals with psychopathy: imagining other in pain does not evoke empathy. J. Front. Hum. Neurosci. **7**, 489 (2013)
31. Hare, R.D.: Without conscience: The Disturbing Words of Psychopaths Among us. Guilford Press, New York (1999)
32. Spielberger, C.D.: State-Trait Anxiety Inventory for Adults. Mind Garden Inc., Menlo Park (1983)

Pragmatic Analysis of Machine Learning Techniques in Network Based IDS

Divya Nehra, Krishan Kumar, and Veenu Mangat

University Institute of Engineering and Technology,
Panjab University, Chandigarh, India
divyanehra@gmail.com, k.salujauiet@gmail.com,
veenumangat@yahoo.com

Abstract. In providing defense to computer networks the network intrusion detection system (NIDS) plays a very essential role. To cope up with the demands of contemporary networks various concerns like performance evaluation and others related to the networks should be taken under consideration. Proposed work presents a pragmatic analysis of machine learning techniques for network based IDS. The performance analysis over two benchmark datasets i.e. KDD-Cup'99 and NSL-KDD by using five supervised machine learning techniques (RFC, Naïve bayes, J48, Bayes Net and SVM) has been prepared. To assess the performance network based intrusion detection system various metrics such as accuracy, recall, F1-score and precision has been computed and analyzed. Therefore, the summary of the work suggests that no single technique is smart enough to identify all attack classes to conventional levels. Most of the techniques provided poor results for minority attack class(es). To estimate and assess the supervised classifier a blind set of investigation with 10-fold cross validation has been performed. The results achieved are promising and provides a new direction to researchers of the intrusion detection domain.

Keywords: Network security · Intrusion detection system · Security performance · Network intrusion detection · Security and protection

1 Introduction

Network based IDS are the software employed within the networks at some deliberated point to analyze network circulation on the whole subnet. The traffic log is matched through the database of recognized attacks and if a spasm is spotted or security policy violation is detected, a signal is passed to the network supervisor. NIDS are classified as On-line NIDS and Off-line NIDS. On-line NIDS are those which are able to work with the real-time networks whereas the Off-line ones are those who works over the repository of data and analyze the data in such a way to identify the attacks and normal instance.

In recent trend, the main attention of researchers has been inclined towards the machine learning techniques and neural network techniques like Random Forest, Support Vector Machines, Naïve Bayes and Decision Trees [1]. These techniques have been achieving better and improved performance in detection accuracy for network1

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 422–430, 2019.
https://doi.org/10.1007/978-981-15-0108-1_39

intrusion detection system. Machine learning taxonomy has given a whole new meaning to the field of intrusion detection when used up to its potential [2, 3].

To address the improvement required in the field of intrusion detection this new strategy is proposed.

2 Background

This section provides the related material which is necessary to realize the stimuluses and the idea behind the anticipated work in this work.

2.1 Network Intrusion Detection System

Now a day dependence over the organizations that relies on gradually demanding application of information technology is increasing rapidly. Thus service provider software is more prone to vulnerabilities and the errors involved are economically high in cost to be solved. This scenario leads to the need and innovation of a strong network monitoring system which can deal with the following pertinent concerns:

- **Dimensionality of data:** The dimensionality of stored as well as passing by data over network is increasing massively and will be continue to increase. According to the forecast made in [4] the amount of data will reach up to 44ZB by 2020. Deploying NIDS to deal with such big amount of data is a major challenge.
- **Reliability:** To achieve desired levels of reliability in terms of accuracy, the existing techniques are somewhere lacking. Hence more granular datasets, more visualization of data is required to achieve more promising results.
- **Mélange:** The present scenario is focusing on developing ensemble and customized protocols using various algorithms and network attributes. Consequently, identification of nefarious and normal behavior is becoming a cumbersome task.
- **Imbalanced datasets:** This problem arises when datasets consist of such classes which has fewer or smaller number of instances. Due to it, NIDS becomes unable to precisely predict such classes and becomes more prone to errors.

2.2 Machine Learning

According to Wikipedia, machine learning is subclass of artificial intelligence in the domain of computer science that empowers the computers with the ability to “learn” the data by using the statistical techniques, without being explicitly programmed [5]. Therefore, machine learning is programming the computers to enhance a benchmark efficiency via past practice or stored data. Machine learning make uses of the philosophy of statistics to build up mathematical prototypes to make out a corollary from an illustration. Various example of machine learning applications is basket analysis using learning associations which says 70% of customers who buy bread also buy butter, classification problem in which two or more classes are present and by making use of machine learning algorithms the appropriate class of the instance is predicted, pattern recognition which consists of face recognition, medical diagnosis and speech recognition etc. [6]. Machine learning algorithms are divided as following types:

- Supervised learning: the aim of this learning is to memorize the patterns or mapping of input to output whose labels or results are provided by the supervisor himself [7].
- Unsupervised learning: in this type of learning no supervisor is present and only input is provided. Here, the goal is to discover the symmetries in the input. The concept of clustering is used here to make clusters of similar patterns [8].
- Reinforcement learning: this learning selects an action out of sequence of actions and learns the policy which was being used by the sequence of actions to reach the goal. Here the aim is to learn the goodness of policies and generate a policy [6, 9].

3 Existing Work

In this section, the most recent prominent works has been discussed.

The goal NIDS using machine learning is to breed a minimal rule set to detect malicious actions deviating from past behaviors. There are quite a few existing workings in the field of Network IDS. The work by [8] propose a new method to Network intrusion detection and achieved a FNR = 1.15%, FPR = 0.09% and detection accuracy of 98.76% in comparison to another SVM based scheme they've achieved FPR = 4.2%, FNR = 7.77% and detection accuracy of 88.03%. [10] propose a machine of generating learning model for NIDS by comparing five machine learning based models and achieved detection accuracy of 99.4%. They've compared the results with reduced feature set and without reduced feature set. Moreover, one more comparison is made between 10 fold cross-validation results and percentage split results.

[11] propose a machine learning based approach using SVM with augmented features. They have implemented the marginal density ratios transformation method to obtain improved detection rate for SVM. The dataset used is NSL-KDD and the results shows the robust performance results. [12] proposes an IDS on the basis of performance comparison between SVM, RFC and ELM to resolve concerns of performance. The use of these techniques shows limitations of large datasets, huge traffic data and gives an efficient classification technique. [13] analyzed methods for management of datasets related to imbalacing and they concludes that minority classes are not capable for learning as compare to majority classes. [14] has discussed problems regarding learning with skewed class scatterings and effect of it over performance of classifiers. The analysis was conducted for artificial intelligence and computational intelligence and confirms the requirement of building efficient intrusion detection systems. In [23], analysis of artificial NN, decision tree, support vector machine, Bayesian networks and a self-organizing map has been done. Even though high and desirable results have been achieved using machine learning but still machine learning consists of some vulnerabilities, such as misclassification of network data due to poison learning. Such vulnerabilities in the system affect performance. So such problems of machine learning need to be addressed.

4 Classifier Used

In our proposed work, following five algorithms have been used on two different datasets i.e. KDD Cup'99 and NSL-KDD. 10-fold CV approach has been applied with the help of Scikit Learn.

- Random Forest Classifier: these classifiers are from the family of ensemble or forest of decision trees. This family generally have low bias and high variance and are perfect contenders for ensemble method. The bootstrap aggregating or bagging technique is generally used in this classifier to achieve increased variance without altering the bias [15].
- J48: it is a predictive learning technique which make predictions for the new instance on the basis of prior available information. It creates a decision tree using the values of available data [16].
- Naïve Bayes: these classifiers belongs to the family of probabilistic classifier. It uses bayes rule of conditional probability. Naïve bayes observes each feature individually as well independently of other features contained by model [17].
- SVM: these classifiers are best suited for multiclass classification problems for big datasets and one of the superfast machine learning classifier with low computational resources [18]. This family supports classification as well as regression.
- BayesNet: These are the sub set of Bayesian networks with nominal attributes and no missing values [19].

5 Calculations

Related to most of the existing research, our proposed work was implemented using Python. All evaluation was performed using 64-bit Windows 10 Pro with an Intel® Core™ i5-8250 CPU @ 1.60 GHz 1.80 GHz with 8.00 GB RAM and an NVIDIA GeForce MX150 GPU. Two of the benchmark datasets of the domain of intrusion detection i.e. KDD Cup'99 as well as NSL-KDD datasets are used for performance evaluation.

The used metrics are as follows:

- True Positive(TP) – those occurrences which are correctly categorized as an intrusion.
- False Positive(FP) – those occurrences which are incorrectly categorized as an intrusion.
- True Negative(TN) – those occurrences which are correctly categorized as normal.
- False Negative(FN) – those occurrences which are incorrectly categorized as normal.

Performance of the proposed work is calculated by using the following measures:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

The measure of accuracy is appropriately identified instances to the total number of records.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

The precision is the measure of correctly identified records to the incorrectly identified records.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

The recall is the measure of correctly identified records to the number of missed records.

$$\text{F1} - \text{Score} = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4}$$

The F1-Score is the measure of harmonic mean of recall and precision.

5.1 Datasets

Two datasets have been used i.e. NSL-KDD and KDD-Cup'99. They are publically available benchmark datasets and have been massively used by the researchers of intrusion detection domain.

KDD Cup'99: In 1998 MIT Lincoln Labs prepared the intrusion detection assessment program named as DARPA IDS evaluation program. The network log consisting of intrusions imitated in military network environment for survey purpose was conducted [20]. Later on, the KDD Cup'99 dataset utilized it. This dataset contains 4900000 number of records with 41 type of features (e.g. duration, flag, land) and these features are broadly classified into three main classes. As it is a labelled dataset so each record is labelled as normal or attack (attack type). Most of the researchers make use 10% subset of original dataset as working with it requires less computation. The dataset needs to be pre-processed before usage. The pre-processing consists of transformation of string or symbolic values to numeric values to make learning easier.

NSL-KDD: The NSL-KDD is the improvement over KDD Cup'99 with reduced number of redundancy. The number of features is same as of KDD Cup'99 [20], [21]. Though this dataset has also faced criticism but still it is being used extensively worldwide. Whole of the dataset has been used for 5-class classification. Following are the various reason to use NSL-KDD (Table 1):

1. Redundant records are not present in train dataset so classifier is free from producing biased results.
2. Test dataset is free from duplicate records which helps in better reduction rates.

6 Results and Discussions

Results obtained are indicating that out of all the classifier used, RFC is performing the best in terms of Accuracy, Precision, Recall and F1-Score. Moreover, one more analysis is made regarding the number of records available for the R2l and U2r class are less as compare to other classes so is the accuracy and other metrics is also low.

Table 1. NSL-KDD 5-Class Performance

Classifier	Metrics	DoS	NORMAL	PROBE	R2L	U2R
RFC	Accuracy (%)	91.59	99.29	98.45	93.56	15.5
	F1-score (%)	92.00	99.04	99.78	97.00	19.57
	Precision (%)	99.54	99.79	99.89	97.89	66.99
	Recall (%)	91.59	99.29	98.45	94.56	15.5
J48	Accuracy (%)	96.32	93.54	76.44	6.9	15.86
	F1-score (%)	97.00	94.54	77.10	10.54	19.42
	Precision (%)	99.00	99.00	99.00	97.00	66.99
	Recall (%)	96.32	93.54	76.44	6.9	15.86
BayesNet	Accuracy (%)	94.5	97.5	83.45	55.75	14.78
	F1-score (%)	95.00	98.07	84.9	55.95	19.33
	Precision (%)	99.00	99.00	97.12	97.45	65.50
	Recall (%)	95.99	97.99	84.97	55.97	15.25
Naïve Bayes	Accuracy (%)	84.3	96.50	78.56	57.44	19.44
	F1-score (%)	85.4	96.99	78.99	57.95	21.50
	Precision (%)	99.41	99.74	99.00	93.00	41.41
	Recall (%)	84.3	96.50	78.56	57.44	19.44
SVM	Accuracy (%)	90.49	94.71	96.39	83.71	13.59
	F1-score (%)	91.48	94.94	97.83	94.17	14.00
	Precision (%)	99.00	94.14	97.78	93.11	15.05
	Recall (%)	91.05	95.02	97.99	94.41	14.00
Total Instances		45927	67343	11656	995	52

6.1 KDD Cup'99 Evaluation

This section provides the evaluations made on KDD Cup'99 dataset.

5-Class Classification: 5-Class classification consists of the standard 5 classes i.e. Normal, DoS, U2r, Probe, R2l. 10% subset of KDD Cup'99, which is a common practice, has been used. The results indicate that 2 out of 5 classes shows poor performance i.e. R2l and U2r. The rest of the classes offer significant level of accuracy, precision, recall and f1-score. Moreover, it can also be observed from the results that the overall performance of Random Forest Classifier is the best and SVM also outperforms whereas naïve bayes is the worst performer in terms of accuracy (Table 2).

Table 2. KDD-Cup'99 5-Class Performance

Classifier	Metrics	DoS	NORMAL	PROBE	R2L	U2R
RFC	Accuracy (%)	95.34	98.97	96.96	81.95	12.9
	F1-score (%)	96.11	97.99	96.99	81.95	14.75
	Precision (%)	98.98	97.98	96.96	82.94	12.66
	Recall (%)	95.34	98.97	96.99	81.96	12.9
J48	Accuracy (%)	61.96	96.97	95.96	60.85	14.73
	F1-score (%)	62.96	97.96	96.98	61.94	15.75
	Precision (%)	66.96	95.95	96.96	64.95	17.86
	Recall (%)	61.95	96.96	95.99	60.96	14.67
BayesNet	Accuracy (%)	87.97	76.97	68.87	55.19	13.23
	F1-score (%)	88.98	77.97	69.98	56.96	14.25
	Precision (%)	90.98	80.97	70.96	60.93	10.86
	Recall (%)	87.97	76.96	68.99	55.96	13.67
Naïve Bayes	Accuracy (%)	56.60	90.95	70.76	50.53	15.54
	F1-score (%)	57.95	91.97	70.97	50.95	10.66
	Precision (%)	59.96	90.96	70.95	50.94	10.78
	Recall (%)	56.96	90.97	70.98	50.96	15.77
SVM	Accuracy (%)	94.93	92.95	80.94	70.26	14.17
	F1-score (%)	94.92	93.95	80.97	70.95	13.86
	Precision (%)	95.97	90.96	81.96	70.97	13.85
	Recall (%)	94.97	92.96	80.94	70.54	14.00
Total Instances		391458	97278	4107	1126	52

7 Conclusion and Future Work

This work has used the benchmark datasets KDD Cup'99 and NSL-KDD to make performance evaluations. The comparisons have made between 5-class classification of both the datasets. On comparison, we found that the RFC is performing the best in both scenarios. Moreover, it may also be noted that the classes like U2r and R2l are not giving very promising results because of the number of instances available for training. It suggests that efforts for refining the performance of present techniques for rare attack classes needs instant addressing by scholars. Moreover, the results obtained also suggests that for a particular attack class, some classifiers perform better than the others. The significant reason for that is different algorithms are designed differently to work with their particular characteristics.

In future work, the improvement will be made in the direction of dealing with class imbalancing problem. We will work upon improvement of existing evaluations by utilizing more efficient methods like shallow learning and deep learning. Hence we can extend the proposed work to achieve more and more merits out of it.

References

1. Dong, B., Wang, X.: Comparison deep learning method to traditional methods using for network intrusion detection. In: 8th IEEE International Conference Communication Software Networks, pp. 581–585 (2016)
2. Axelsson, S.: Intrusion detection systems: a survey and taxonomy. Tech. Rep. **99**, 1–15 (2000)
3. Hodo, E., Bellekens, X., Hamilton, A., Tachtatzis, C., Atkinson, R.: Shallow and deep networks intrusion detection system: a taxonomy and survey. CoRR, abs/1701.0, pp. 1–43 (2017)
4. Executive summary: Data growth, business opportunities, and the IT imperatives—The digital universe of opportunities: Rich data and the increasing value of the Internet of Things. https://www.emc.com/%0Aleadership/digital-universe/2014iview/executive-summary.htm
5. Machine learning. https://en.wikipedia.org/wiki/Machine_learning
6. Alpaydin, E.: Introduction to Machine Learning. MIT Press, Cambridge (2010)
7. Sommer, R., Paxson, V.: Outside the closed world: on using machine learning for network intrusion detection. In: Proceedings of the IEEE Symposium Security Private, pp. 305–316 (2010)
8. Chowdhury, M.N., Ferens, K., Ferens, M.: Network intrusion detection using machine learning, pp. 30–35 (2010)
9. Alpaydın, E.: Introduction to machine learning. Methods Mol. Biol. **1107**, 105–128 (2014)
10. Kumar, S., Viinikainen, A., Hamalainen, T.: Machine learning classification model for network based intrusion detection system. In: 2016 11th International Conference for Internet Technology and Secured Transactions (ICITST), pp. 242–249 (2016)
11. Wang, H., Gu, J., Wang, S.: An effective intrusion detection framework based on SVM with feature augmentation. Knowl.-Based Syst. **136**, 130–139 (2017)
12. Ahmad, I., Basheri, M., Iqbal, M.J., Rahim, A.: Performance comparison of support vector machine random forest and extreme learning machine for intrusion detection. IEEE Access **6**, 33789–33795 (2018)
13. Kotsiantis, S., Kanellopoulos, D., Pintelas, P.: Handling imbalanced datasets: a review. In: GESTS International Conference on Computer Science and Engineering, vol. 30, pp. 25–36 (2006)
14. Monard, M.C., Batista, G.E.A.P.A.: Learning with skewed class distribution. In: Advances in Logic, Artificial Intelligence and Robotics, Sao Paulo, SP, pp. 173–180. IOS Press (2002)
15. Random Forest Classifier. https://www.stat.berkeley.edu/~breiman/randomforest2001.pdf. Accessed 19 April 2018
16. Sahu, S.: Network intrusion detection system using J48 decision tree. In: 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 2023–2026 (2015)
17. Belavagi, M.C., Muniyal, B.: Performance evaluation of supervised machine learning algorithms for intrusion detection. Procedia Comput. Sci. **89**, 117–123 (2016)

18. Li, Y., Xia, J., Zhang, S., Yan, J., Ai, X., Dai, K.: An efficient intrusion detection system based on support vector machines and gradually feature removal method. Expert Syst. Appl. **39**(1), 424–430 (2012)
19. Kumar, G.: AI based supervised classifiers : an analysis for intrusion detection. In: Proceedings of the International Conference on Advances in Computing and Artificial Intelligence, pp. 170–174 (2011)
20. Tavallaee, M., Bagheri, E., Lu, W., Ghorbani, A.A.: A detailed analysis of the KDD CUP 99 data set. IEEE Symp. Comput. Intell. Secur. Def. Appl. CISDA, 1–6 (2009)
21. Dhanabal, L., Shantharajah, S.P.: A study on NSL-KDD dataset for intrusion detection system based on classification algorithms. Int. J. Adv. Res. Comput. Commun. Eng. **4**(6), 446–452 (2015)
22. Zamani, M., Movahedi, M.: Machine learning techniques for intrusion detection. Comput. Sci. **2**, 1–11 (2015)
23. Sharma, R.K., Kalita, H.K., Borah, P.: Analysis of machine learning techniques based intrusion detection systems á supervised learning. In: Proceedings of 3rd International Conference on Advanced Computing, Networking and Informatics, vol. 44, pp. 485–493 (2016)
24. Chowdhury, M.N., Ferens, K., Ferens, M.: Network intrusion detection using machine learning. Int. Conf. Secur. Manag. **4**, 30–35 (2010)
25. Hamid, Y.: Machine learning techniques for intrusion detection : a comparative analysis. In: ICIA, vol. 7, pp. 0–5. ACM (2016)
26. Ambusaidi, M., He, X., Nanda, P., Tan, Z.: Building an intrusion detection system using a filter-based feature selection algorithm. IEEE Trans. Comput. **65**(10), 2986–2998 (2016)
27. Singh, R., Kumar, H., Singla, R.K.: An intrusion detection system using network traffic profiling and online sequential extreme learning machine. Expert Syst. Appl. **42**(22), 8609–8624 (2015)
28. Proti, D.D.: Review of KDD cup, NSL-KDD and kyoto, datasets. Mil. Tech. Cour. **66**, 580–596 (2006)
29. Angelo, P., Resende, A., Drummond, A.C.: A survey of random forest based methods for intrusion detection systems. ACM Comput. Surv. **51**(3), 52:1–52:27 (2018)
30. Devaraju, S., Ramakrishnan, S.: Performance analysis of intrusion detection system using various neural network classifiers. Int. Conf. Recent Trends Inf. Technol. ICRTIT **2011**, 1033–1038 (2011)

A Recent Survey of DCT Based Digital Image Watermarking Theories and Techniques: A Review

Ranjeet Kumar Singh[1(✉)] and Anil Kumar Singh[2]

[1] National Institute of Technology, Jamshedpur 831014, Jharkhand, India
2014rsca002@gmail.com
[2] School of Management Sciences, Varanasi 221011, Uttar Pradesh, India
singhanilkumar@zoho.com

Abstract. Digital image watermarking is one of the most discussed research areas. It plays an essential role in the field of digital information authentication and security. Based on the study of watermarking systems, image watermarking is divided into two modules: One is watermark embedding, and the other is watermark extraction. This paper reviews the theoretical and experimental analysis and performance measurement of a representative digital image watermarking system in spatial and transforms domains. The key characteristics of a digital image watermarking scheme are robustness, capacity, imperceptibility, the security of the hiding place, and false positive of the watermarking algorithm. Comparison of the different watermarking techniques employing the discrete cosine transform (DCT) is given. This paper presents various details of the algorithm of digital image watermark embedding and extraction process in a different domain and also explains their advantages and disadvantages.

Keywords: Discrete cosine transformation · Digital encryption · Image processing · Watermarking

1 Introduction

In the contemporary era of digital information sharing, it is only desirable to have a robust and safe ecosystem for it to thrive. The main advantage of exchanging data on digital media is the accuracy with which information can be passed on [1]. Digital image watermarking provides security against unauthorized access and confirms digital data ownership [2]. Cryptography is one of the approaches available for information protection and system security. In a conventional cryptographic system, the encrypted information can be decrypted only by an authentic user who would be aware of the decryption key. But in a situation when this information is deciphered by a hacker or any other unintended person, we are left with minimal choices to shield the information and track its illegal distribution. It is a significant shortcoming of conventional cryptography. Watermarking is one efficient way to protect intellectual property and overcome the issue above. Watermarking is an approach which is used to embed information into a cover image to provide authentication of the digital data. It facilitates ownership of digital information.

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 431–440, 2019.
https://doi.org/10.1007/978-981-15-0108-1_40

Some of the watermarking applications are given below [3, 4]:

- **Copyright Protection:** To secure intellectual property, the owner of data can hide watermark information used for authentication of data. The hidden watermark is treated as a piece of evidence, e.g., in case of willful infringements of copyright.
- **Fingerprinting:** To search for the source of protected copies, the authentic user can use a mechanism which is known as fingerprinting. By using this technique, the genuine user can hide different watermark data as replicas of the data that are delivered to different users. Fingerprints can be matched by inserting a serial number, and this serial number can be used to validate the authentic user for the data concerned.
- **Copy protection:** The data hid in the watermark can manipulate the recording devices for copy protection of digital data. In such a situation, the watermark denotes a copy-prohibited bit, and watermark sensors in the recorder determine if the data input to the recorder is saved or not.
- **Data Hiding:** Watermark approach is useful in transmitting secret information. Since many data admins do not use the encryption facilities, users can hide information in other data.
- **Medical safety:** Hiding the patient's identity and their medical conditions that are usually stored in images could be a useful safety measure.
- **Non-perceptibility:** Non-perceptibility refers to the watermark embedded into information that cannot be seen by users. It will be identifiable only through dedicated circuits.
- **Robustness:** Watermark, which is presented in the original data, can continue its existence even after a different image processing attack. The watermark is robust and can defend itself against image processing attacks. Image processing attacks can be classified mainly into rotation, scaling, and noise.

1.1 Classification of Digital Watermarking

Digital watermarking is a mechanism used to embed data, known as a watermark, into multimedia or a digital object. The insertion process is such that the hidden information can be detected later and recovered to assert the object. The digital items in which the watermark information is inserted are generally known as the host signals, the original, or simply, the work.

Figure 1 depicts the various forms of watermarks. Watermarks are classified into four broad categories, viz., video, audio, image, and text. Based on human perception, watermarking can be classified into a visible, invisible-robust, invisible-fragile, and dual watermark. Visible watermarking is robust, but the area of application is small and limited. In invisible watermarking, the watermark cannot be detected by the human eye. Only the authentic user knows the watermark. Another user cannot identify the watermark, and they have no means to change the watermark. A robust watermarking technique is generally used to sign copyright of digital content. The hidden and inserted watermark data can resist several forms of images processing.

Any signal processing attack does not damage the watermark data, and it can be detected for official certifications. Integrity protection fragile watermarking is typically used because it is very complex to alter the signal. In the case of the invisible-fragile watermarking scheme, the watermark information is inserted in a way that any change or manipulation of the picture results in ruining the watermark. In case of the dual watermarking scheme, it is a combined approach of the visible watermarking and the invisible watermarking schemes. Here, the watermarks are the blend of visible watermarks and the invisible watermarks. The invisible watermark is used in this case to back up the visible watermark.

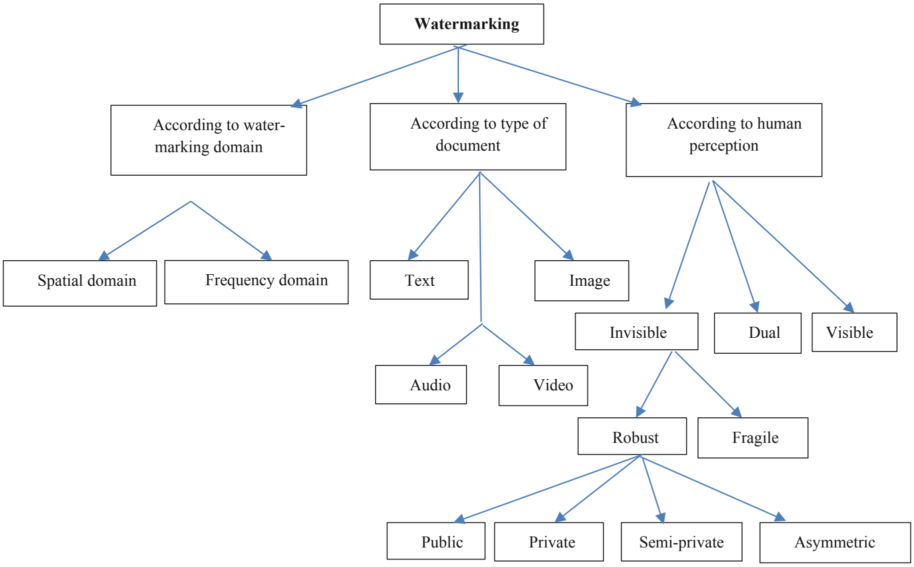


Fig. 1. Different watermarking schemes

Robust watermarking is categorized as follows:

- **Private watermarking scheme:** In the case of private watermarking, we need the host image for detection of watermark. Private watermarking is classified into two types: Type I system and Type II system.
- **Type-I systems** recover the watermark information from the tested, and probably, the tattered image. The host image is required to find the locality of the watermark data in the disrupted image.
- **Type-II systems** need a duplicate of the inserted watermark data for watermark identification, and they are capable of expressing whether a piece of specified watermark information in the verified image exists or not.

Both the schemes are used to create knowledge about the embedding key (Private Key). The embedding key is private data that inserts the watermark into the cover object.

- **Semi-private watermarking:** In this scheme of watermarking, the original image is not required for the detection of the watermark. It gives information about the watermark, which is present or not in the watermarked image.
- **Public watermarking:** Also known as the blind watermarking, there is no need for the cover image and inserted watermark for the watermark extraction procedure. This watermarking scheme requires a more complex watermark technology, and its field of application is wide.
- **Asymmetric watermarking:** This watermarking is known as public-key watermarking. In contrast to the public watermarking mechanism, in this scheme, everyone knows the detection scheme and the detection key. The knowledge about the public key either prevents finding the private key, or it restricts the removal of the watermark.

1.2 Copyright Protection Watermarking and Tampering Tip Watermarking

Watermarking is basically a process of imbibing certain media inside the original information [5]. In copyright protected watermarking, an authentic user would want other users to see the watermark data; then the watermark data can be seen by other users after embedding watermark data to the original information. According to the watermarking domain, it is mainly divided into spatial and frequency domain. In case of a spatial domain watermarking, watermark information is directly put together with the cover image. One of the spatial domain watermarking schemes is the LSB (Least significant bits) mechanism. In this mechanism, the watermark information is inserted into the least significant bits of the cover image. But the spatial domain watermarking suffers a serious issue of limited sturdiness [6]. In the frequency domain scheme, the watermark information is added to the original image. Image transform is then used to modify the image coefficients. Normally, masking based transformed domain techniques are more robust than LSB mechanism from the image processing attacks such as compression and cropping viewpoint.

2 Discrete Cosine Transform (DCT)

Another transform domain watermarking approach is a discrete cosine transform (DCT). Compared to the discrete Fourier transform (DFT), DCT is a better digital watermarking technique because it involves only the orthogonal transformation of real numbers, unlike the DFT where a digital image is computed as a part of a complex number. DCT has the advantage of high-compression ratio and low error-rates [7]. Based on frequency components, this approach allows dividing an image into three parts, viz., low, high, and middle-frequency bands. We can insert watermark in any frequency band. The literature survey discloses that usually the middle frequency components are used to add watermark because, in the middle-frequency component, the information stored in watermark often cannot be scattered. If the watermark is rooted in a low-frequency component, the mechanism tends to be resistant against malicious image-processing attacks, but it is a daunting task to hide the watermark. But in case of the watermark embedded in a high-frequency band, i.e., a perceptually insignificant component, watermark hiding scheme is more straightforward. However, the system is less resistant to image processing attacks [8–11]. The equations given below describe the 2D-DCT and 2D-IDCT. 2D-DCT (two-dimensional discrete cosine transform) *F [u, v]* of a digital image matrix *f [m, n]* is:

$$F[U,V] = \sum_{n=1}^{N}\sum_{m=1}^{M}[u,v]\,f[m,n]\cos\frac{\pi(2m-1)(u-1)}{2M}\cos\frac{\pi(2n-1)(v-1)}{2N}$$

Where,

$$[u,v] = \begin{cases} \frac{1}{\sqrt{MN}} & When\ u=1\ and\ v=1 \\ \sqrt{\frac{2}{MN}} & When\ u=1\ and\ v\geq 2\ or\ v=1\ and\ u\geq 2 \\ \sqrt{\frac{4}{MN}} & else \end{cases}$$

Where, *w [u, v]* is known as weight factor of the transform, *n* and *v* vary from 1 to N, and *m* and *u* vary from 1 to M.

Liu et al. [7] developed a novel watermarking algorithm by a serial amalgamation of fractal encoding and discrete cosine transformation (DCT) techniques. Through this dual encryption method, the authors proposed an improved DCT encryption technique. A cover image is first encoded using the fractal encoding, followed by the second encryption of the encoded parameters using DCT. With the help of only two dimensions of scaling and offset, and applying three types of attacks, the authors tested their technique to conclude that this new embedded technique is more robust and effective.

Roy and Pal [12] came up with a DCT technique, embedded with a repetition code approach of color watermarking for copyright ownership and validation. The authors eliminated the ‘blocking artifact,’ a significant drawback of the block-based DCT, by making use of zigzag scanning of each RGB component’s DCT block. The purpose of this work was to use the error correcting code (ECC), called the repetition code for preserving one watermark bit in every decomposed non-overlapping RGB component’s block. This multiple image watermarking technique demonstrated an imperceptibility property and yielded higher PSNR value and better robustness. But these benefits were achieved only at the price of higher computational complexities.

Singh et al. [13] made use of a hybrid scheme comprising of SVD, DCT, and nonsubsampled contourlet transform (NSCT) to derive a robust, high capacity, and imperceptible watermarking of confidential medical images. In this algorithm, the electronic patient record (the secret message) is rooted into the sub-band of the cover image (the medical image) with a chosen gain factor, resulting in an improved capacity, imperceptibility, and robustness. They determined that clubbing NSCT and SVD with DCT yields a more secure and high-quality image watermarking.

Zear et al. [14] made use of multiple watermarking schemes comprising of DCT, DWT, and SVD for application in the healthcare industry. They went on to use the Back Propagation Neural Network (BPNN) on the recovered image watermark to suppress noise residuals on the recommended grayscale image for watermarking.

2.1 Inverse Discrete Cosine Transform (DCT)

The two-dimensional (2D) inverse discrete cosine transform (IDCT) is defined in below equation:

$$f[m,n] = \sum_{v=1}^{N}\sum_{u=1}^{M} w[u,v]F[u,v]\cos\frac{\pi(2m-1)(u-1)}{2M}\cos\frac{\pi(2n-1)(v-1)}{2N}$$

W [u, v] = Fore mentioned weight factor

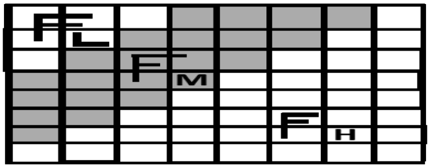


Fig. 2. Various frequency regions in the DCT domain

In Fig. 2, the DCT blocks having low-frequency components are depicted by FL, elements having high-frequency bands are represented by FH, and the middle-frequency bands are represented by FM [15].

In recent years, a lot of digital image watermarking schemes have been proposed to provide security and authentication to multimedia data. Various watermarking approaches have so far been suggested for images. Some DCT based image watermarking algorithms are jotted in the following literature survey.

Al-Baloshi et al. [11], addressed a DCT based image watermarking approach for image security and authentication. In their approach, the watermark was used as a form of visually meaningful binary pattern. Here authors divided the original image into N × N block size, then applied DCT on each block and inserted each DCT block to 1-bit of the watermark.

Al-Afandy et al. [16] proposed a DSWT and DCT based watermarking technique. Initially, the original image is converted its red, green, and blue (RGB) component, and DCT is applied to each color component. DCT output is divided into four sub-bands by using Discrete Stationary Wavelet Transform (DSWT). These frequency bands are represented by A, H, V, and D matrices with the same image size. Here, the watermark is inserted in matrix A.

A new blind multiple watermark scheme was presented by Ahmed N. Al-Gindy et al. [17], where two watermarks are used. Two vectors are created by two watermarks, and then they are merged. Low-frequency bands of the DCT Domain are chosen for insertion of watermark data. The cover image is divided into 8 × 8 block size, and sixteen coefficients of the host image of 8 × 8 sub-blocks are used for the addition of sixteen bits of the merged watermarks.

Nowadays, a binary watermark scheme has found utility in image security. Gindya et al. [18] explained the binary watermarking in a color image. In this technique, the host image is converted into its RGB components, and the green component is chosen for watermarking. The green part is distributed into 8 × 8 sub-blocks, and DCT is applied on each block. The low-frequency component is selected for watermarking. Firstly, the watermark image is scrambled with the help of a private key and then changed to a vector. Before the commencement of the inserting scheme, the binary watermark vector is shifted with a specific shift for each time. This algorithm provides dual level security. First, the watermark is scrambled, and then another watermark is inserted in a particular color component.

For providing more robustness and protection to the digital content, several YC_bC_r color space-based watermarking is available. A. Al-Gindy et al. [18] had focused on the binary watermarking scheme in color image. The original colored image is converted into its RGB components to YC_bC_r color space. Y–channel is used for watermarking. First, the Y-channel ID has divided into N × N block sizes, and DCT is used on each block. The authors in this paper elected low-frequency component for embedding watermarks.

Arnold Cat Map is a widely used robust and efficient technique in image watermarking [19]. A combined approach of YC_oC_g-R color space and Arnold transform for data security was proposed by Moosazadeh and Ekbatanifard [20]. The authors used Arnold transform for multi-level security. The encryption of the watermark image was done by using Arnold transformations five times. The cover picture was transformed into its RGB color components. Then the RGB was converted to YC_oC_g-R and then separated the Y, C_o and C_g components. The Y component of the image was also scrambled by using Arnold transformations and turning its 8 × 8 block size. DCT is applied on each block to find a low-frequency band for insertion of a watermark.

Badran et al. [21] had explained the image watermarking algorithm based on the Expectation Maximization (EM) algorithm, which was used for image segmentation and DCT methods. All image segments used were divided into 8 × 8 block sizes and randomly reordered. The DCT was applied on each block, and a pseudorandom sequence of real numbers was inserted in each image segment of the DCT domain.

Gupta et al. [22] explained a new watermarking scheme built on DCT and LSFR. The host image was divided into N × N blocks size, and DCT was applied on each block. The watermark was converted in a bit sequence and stored in a one-dimensional array. Next, a pseudo-random number was generated by using a Linear Feedback Shift Register (LFSR). Here, the watermark bit was added to the low-frequency component.

Shuifa et al. [23] through their paper presented a binary image watermarking technique. First, the cover image was filtered by a Gaussian filter and then divided it into 8 × 8 blocks size. By applying DCT on each block and calculating the average DCT, the DC component of the whole image for block selecting and watermark data are inserted in these blocks.

The basic idea of DCT based image watermarking approach:

Embedding Process

Step 1: Select the watermark image and host image.
Step 2: Divide the original image into N × N blocks size.
Step 3: convert the original message and watermark image into double.
Step 4: Determine the size of Host image and watermark image.

Step 5: Find the numbers of the block in the host image, i.e., Max message size.

Step 6: Find the length of the watermark message.

Step 7: If the length of the host image message > length of watermark message gives an error

Step 8: Pad the watermark message.

Step 9: Apply DCT in each block

Step 10: Chose the middle-frequency band for embedding the watermark.

Step 11: Watermarked image.

Recovery process

Step 1: Chose a watermarked image and divided N × N blocks size.

Step 2: convert the watermarked image and watermark image into double.

Step 3: Determine the size of the watermarked image and watermark image in the form of several rows and columns.

Step 4: Find the maximum length of the watermarked image.

Step 5: Apply the DCT of each block and applying the inverse embedding procedure to get recover watermark (Fig. 3).

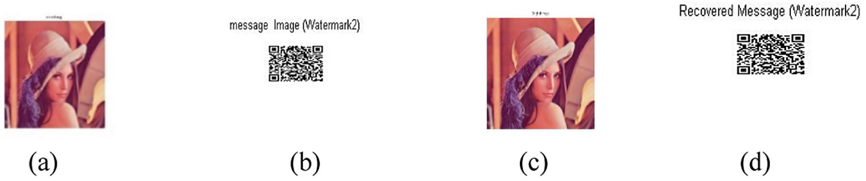


Fig. 3. (a) Original image (b) Watermark image (c) Watermarked image (d) Recovered watermark image

The main advantage of the DCT based image watermarking is that the watermark is embedded into the color channels of the original image. There are two critical issues in DCT based image watermarking. First, by selecting the high-frequency component, the filtering operation can remove the watermark information from the image. The next question is based on how much data modifications were made on DCT coefficients to insert the watermark data. These variations made on factors influence the hiddenness of the watermark information's and destroy the image to a huge extent.

3 Conclusion

In this paper authors explain the detailed working of discrete cosine transformation-based image watermarking for the purpose of providing the security and authentication of digital data. The detailed updated survey of DCT based watermarking is shown in this paper. Basically, DCT decomposes a matrix or a image matrix into low, high, and middle frequency components. Most of the research work carried out so far used middle frequency component for watermark embedding because human eye is less receptive to identifying changes in this component.

References

1. Singh, R.K., Kumar, B., Shaw, D.K., Khan, D.A.: Level by level image compression-encryption algorithm based on quantum chaos map. J. King Saud Univ.-Comput. Inf. Sci. (2018)
2. Singh, R.K., Shaw, D.K., Sahoo, J.: A secure and robust block based DWT-SVD image watermarking approach. J. Inf. Optim. Sci. **38**, 911–925 (2017)
3. Cox, I.J., Miller, M.L., Bloom, J.A., Honsinger, C.: Digital Watermarking. Morgan Kaufmann, San Francisco (2002)
4. Xuehua, J.: Digital watermarking and its application in image copyright protection. In: International Conference on Intelligent Computation Technology and Automation, pp. 114–117. IEEE (2010)
5. Singh, R.K., Shaw, D.K., Jha, S.K., Kumar, M.: A DWT-SVD based multiple watermarking scheme for image based data security. J. Inf. Optim. Sci. **39**, 67–81 (2018)
6. Singh, R.K., Shaw, D.K., Alam, M.J.: Experimental studies of LSB watermarking with different noise. Procedia Comput. Sci. **54**, 612–620 (2015)
7. Liu, S., Pan, Z., Song, H.: Digital image watermarking method based on DCT and fractal encoding. IET Image Process. **11**, 815–821 (2017)
8. Tewari, T.K., Saxena, V.: An improved and robust DCT based digital image watermarking scheme. Int. J. Comput. Appl. **3**, 28–32 (2010)
9. Bedi, S., Kumar, A., Kapoor, P.: Robust secure SVD based DCT-DWT oriented watermarking technique for image authentication. Int. J. Comput. **17**, 46.1–46.7 (2009)
10. Hernandez, J.R., Amado, M., Perez-Gonzalez, F.: DCT-domain watermarking techniques for still images: detector performance analysis and a new structure. IEEE Trans. Image Process. **9**, 55–68 (2000)
11. Al Baloshi, M., Al-Mualla, M.E.: A DCT-based watermarking technique for image authentication. In: International Conference on Computer Systems and Applications, pp. 754–760. IEEE (2007)
12. Roy, S., Pal, A.K.: A blind DCT based color watermarking algorithm for embedding multiple watermarks. AEU-Int. J. Electron. Commun. **72**, 149–161 (2016)
13. Singh, S., Singh, R., Singh, A.K., Siddiqui, T.J.: SVD-DCT based medical image watermarking in NSCT domain. In: Hassanien, A.E., Elhoseny, M., Kacprzyk, J. (eds.) Quantum Computing: An Environment for Intelligent Large Scale Real Application. SBD, vol. 33, pp. 467–488. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-63639-9_20
14. Zear, A., Singh, A.K., Kumar, P.: A proposed secure multiple watermarking technique based on DWT, DCT and SVD for application in medicine. Multimed. Tools Appl. **77**, 4863–4882 (2018)
15. Li, Y.-L., Li, J.-P., Ren, Q.-B.: Based on chaotic encryption and SVD digital image watermarking. In: International Conference on Apperceiving Computing and Intelligence Analysis Proceeding, pp. 285–289. IEEE (2010)
16. Al-Afandy, K.A., Faragallah, O.S., El-Rabaie, E.-S.M., El-Samie, F.E.A., Elmhalawy, A.: A hybrid scheme for robust color image watermarking using DSWT in DCT domain. In: IEEE International Colloquium on Information Science and Technology, pp. 444–449. IEEE (2016)
17. Al-Gindy, A., Al-Ahmad, H., Qahwaji, R., Tawfik, A.: A novel blind Image watermarking technique for colour RGB images in the DCT domain using green channel. In: Mosharaka International Conference on Communications, Computers and Applications, pp. 26–31. IEEE (2008)

18. Al-Gindy, A., Al-Ahmad, H., Qahwaji, R., Tawfik, A.: Watermarking of colour images in the DCT domain using Y channel. In: International Conference on Computer Systems and Applications, pp. 1025–1028. IEEE (2009)
19. Kumar, B., Singh, R.K., Singh, A.K.: A noble watermarking scheme based on spatial domain approach with pixel exchange and compressive sensing. SSRN 3350904 (2019)
20. Moosazadeh, M., Ekbatanifard, G.: Robust image watermarking algorithm using DCT coefficients relation in YCoCg-R color space. In: IEEE Conference on Information and Knowledge Technology, pp. 263–267. IEEE (2016)
21. Badran, E.F., Ghobashy, A., El-Shennawy, K.: DCT-based digital image watermarking VIA image segmentation Techniques. In: 4th International Conference on Information and Communications Technology. IEEE (2006)
22. Gupta, G., Joshi, A.M., Sharma, K.: An efficient DCT based image watermarking scheme for protecting distribution rights. In: Eighth International Conference on Contemporary Computing, pp. 70–75. IEEE (2015)
23. Sun, S., Ling, J., Dong, F., Wan, J.: A new general binary image watermarking in DCT domain. In: International Seminar on Future BioMedical Information Engineering, pp. 34–36. IEEE (2008)

Context Aware Self Learning Voice Assistant for Smart Navigation with Contextual LSTM

J. Silviya Nancy[1](✉), S. Udhayakumar[2], J. Pavithra[3], R. Preethi[4], and G. Revathy[5]

[1] PES University, Bangalore, India
jsilviyanancy@gmail.com
[2] Rajalakshmi Engineering College, Chennai, India
mailtoudhay@gmail.com
[3] Lucid Technologies and Solutions, Chennai, India
pavithrajothi1996@gmail.com
[4] Tata Consultancy Solutions, Chennai, India
preethirv97@gmail.com
[5] Hexaware Technologies, Chennai, India
revathyrevs3197@gmail.com

Abstract. The gift of vision for humans is a valuable blessing but regrettably there are around 37 million people who are visually impaired. Among them, 15 million people are from India. They undergo numerous challenges in their daily lives. They are always dependent on others for traveling to different places. It is noted that context-awareness has a key responsibility in the lives of the visually impaired. There are many mobile applications contributing to ease them, but due to dependence on many additional resources, it has become a nightmare. To sophisticate the above challenge, the proposed mobile-cloud context-aware application will act as a voice chat-bot that provides context-aware travel assistance to the visually challenged people which is implemented in specific public environments. It is an interactive application and provides them with a help desk where they can query their necessary information through speech interface. This application relies on the Location based services including providers and Geo-coordinates for manipulating the latitude and longitude of places. The present location of the user is tracked by using location services. The distance from the user's exact location to the destination is pre-determined and this application will assist them with the route to travel through audible directions. This would completely assist them with the travel by replying to the queries asked by them and it helps them to travel independently. The application flow initially takes the voice instruction and converts that into the text instructions. The contextual LSTM (Long-short Term Memory) in the application takes care of the conversational strategy, analyzes it and advocates all the users with answers for whatever questions are been posted. It also drives the visually handicapped to destination by identifying the obstacle and detection of the object in the way. The application uses the computational resources from cloud servers such as location specific resources and in turn pushes all the data in cloud server for reference and future usage.

Keywords: Voice assistant · Contextual LSTM · Location based service · Visually challenged · Smart navigation · Voice bot

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 441–450, 2019.
https://doi.org/10.1007/978-981-15-0108-1_41

1 Introduction

Loss of Vision is the major hurdle which prevents the people from performing their own personal tasks independently. Travel can be daunting for people with visual impairment. They tend to rely on others for support and backing. This may give them an unbalanced lifestyle and they would deficit in mastering self-empowerment and confidence. But technology has made numerous advancements over the recent years. Today, in India there is world's largest number of blind people which accounts to 1.5% of the Indian population (15 million people). According to surveys of World Health Organization (WHO), 37 million are blind around the world and about 246 million are affected with half-blindness and 285 million are visually impaired. In general, vision loss is a major obstacle in day to day life. They have to undergo a various challenges such as reading and writing [4], traveling around in public places [2, 16] and interacting with people and surroundings. They will always require a companion to travel. It's counted that, advocating the mobile technology in the lives of visually impaired, helps to achieve a better standard in life. It bridges the gap between the necessity and the available technology, because mobility has standardized the quality of life with the increased use of smartphones. Presently the domain of Cloud Computing has taken a huge leap in recent years contributing to all the areas of computer science. Alongside, the mobile technology also took its legacy by servicing the users in all the aspects of life in the form of applications. These together gave rise to a new prototype called Mobile Cloud Computing [15].

Neural Networks at present masters the arena by providing the best solution for many problems by predicting the future itself with the collaboration of supervised and un-supervised learning. It has developed a great impact in the field of pattern recognition, image processing, speech processing and recognition, etc. But the traditional neural networks are quiet unclear in understanding the sequence of events and it may not be able to handle the previous events properly. In the course of time, Recurrent Neural Networks address this issue by persistence of data in the hidden layers. The idea is to connect this to LSTM which manages the Long-term dependencies with the chain of repeating layers.

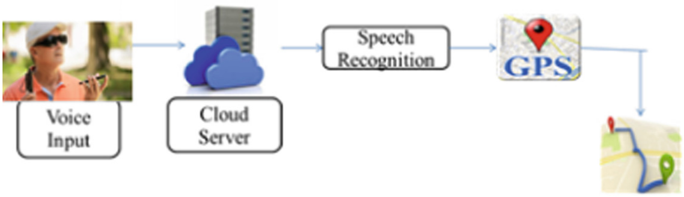


Fig. 1. Background of the proposed model

In this research work, we have proposed a novel voice bot application with collaboration of mobile cloud [13] and LSTM to provide travel assistance to the people of visual deficit as in Fig. 1. Our approach is based on speech recognition technology to provide direct communication facility that would give confidence to the visually challenged [14] community and to enable them to communicate their requirements and travel independently. Speech processing and recognition platform uses the voice message [8] and converts it to textual phrase that is understandable. This application can be deployed in public places which can advocate them to travel around on their own.

2 Literature Review

This section is focused on the literature survey which insights the need for the development of the application assistance that can be used for navigation of visually challenged automated by Mobile Cloud Computing platform. The key aspect of this research is to emphasize the need for the smart application to assist the blind people. There were many researchers conducted to develop an interface application that helps the visually challenged to travel independently. Bansode et al. [1] and the co presenters has proposed and developed a smart routing system for the people who lack vision by using voice as input. They have designed the model with a micro-controller and used location based services which helps them to travel just about easily.

Guentert [3] built a routing supporter that focused on helping the vision loss community, where the current location of the user is identified, which is implemented with map annotation to recognize using computer vision techniques. Gawari and Bakuli [5] have proposed a prototype using location based navigator and voice recognition for evading hindrances and complication and to clearly guide the blind. Sánchez and Oyarzún [6] developed an application of cognitive intelligence with an user-friendly interface which renders a platform that supports the blind community with needed information of the places the person is travelling through.

Chumkamon et al. [7] and the researches of this work have developed an assistant for blind using RFID sensors in limited surroundings. Here the exact current location of the person is captured with the location tag, and then the destination is manipulated with the routing server and that calculates the minimal distance to reach faster. The author, Gulati [9] proposed a vigilant voice based structure with an inbuilt micro-controller. This would help them by indicating about the location and the nearby spots. The paper [10] discusses about the importance of context awareness for the visually impaired by implementing traffic-light detectors. This review of literature gives an ample motivation and knowledge on the requisite for developing a technology that helps the visually handicapped population. Azzouni and Pujolle [11] propose LSTM RNN architecture for predicting traffic matrix in huge networks. The authors [12] Sak et al., stated an efficient architecture for speech processing and recognition by justifying LSTM RNN is better than DNN in acoustic modeling. The discussed review gives various insights on building a voice assistant for visually challenged people for directing them and moderating the app with RNN LSTM for prediction based on the context.

3 Proposed Context Aware Voice Bot System

The context-aware navigation components that should be taken care while designing are to avoid obstacles, walking in right direction, knowing the place where they have reached and the destination, Google maps for outdoor navigation and pedestrian mode and speech recognition interface. The architecture in Fig. 2, explains how well the application can advocate the person with lack of vision, as follows…

1. The person with lack of vision can use this mobile application as a personal assistant.
2. Input Voice (Destination) is recognized and the voice is converted to text.
3. The current and exact location of the user is captured via GPS.
4. Distance and direction between current location and the stored location (destination) are calculated using Google Maps.
5. The text received from the voice instruction is used for question framing and analyzing conversational strategies.
6. This is done by RNN – LSTM with which the user will be guided through answers for all the queries that is posted.
7. This also enables detection of obstacles and answering the queries and helps the user to reach from source to destination.
8. The direction is given as voice output to the visually challenged user.

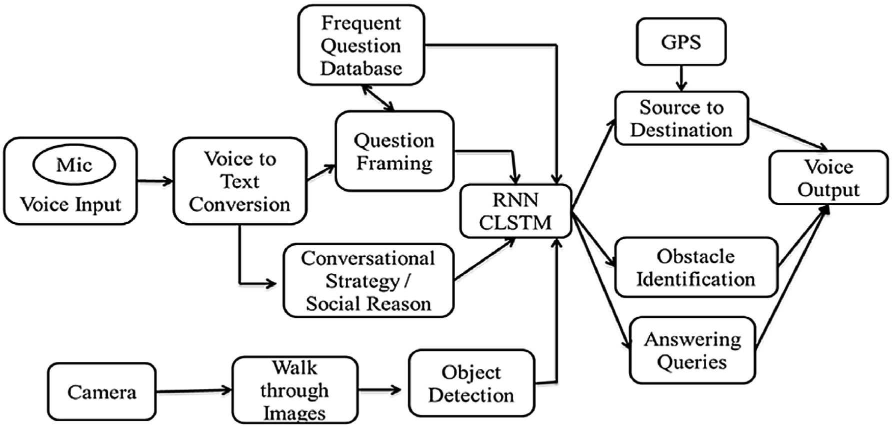


Fig. 2. Architecture of smart navigation for the visually challenged

3.1 Preprocessing Technique

This system eliminates the barrier faced by the visually challenged people while interacting with the mobile application. An audio pre-processor named as Noise Suppression (NS) is used to eliminate the background noise from the voice signal that is facilitated from the application through the input. The main functionality of NS is to facilitate voice commute applications. The unwanted noise may be from the vehicles that are in motion or the conversional chats of other people nearby. To attach the Noise Suppressor to a particular Audio Record (voice input – destination) recorded by the visually challenged people, a specific id is generated at a particular time when the Noise Suppressor is been used. Then that particular word spoken by the user is recognized.

3.2 Speech Processing

Speech is the easiest form of communication. The speech recognition system helps to analyze and understand the spoken information through various stages such as analysis, feature extraction, modeling and decoding. Extraction of feature is an integral component of the speech recognition system. The main task of the extractor is to extort some relevant features that are received from the voice input. It compresses the magnitude of the input signal without causing any harm to the original signal. The speech recognition system has a key building block called Acoustic model. An acoustic model obtains the statistical depiction of different sounds may constitute a word. This is called as phoneme. The HMM (Hidden Markov Model) is distinct for each and every phoneme. The speech decoder then understands the diverse sounds and phrases received from the user by finding a matching HMM as depicted in Fig. 3.

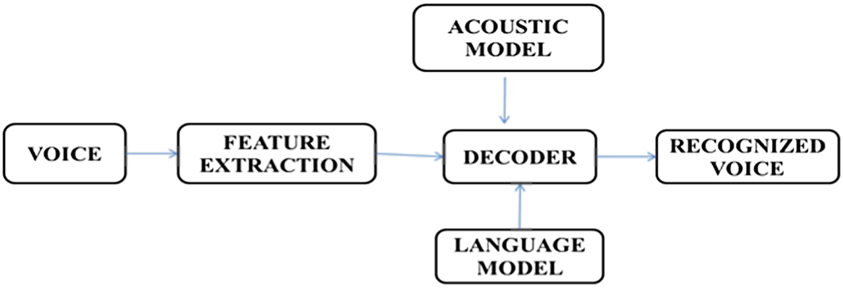


Fig. 3. Sequence of steps in speech recognition

3.3 Steps in Conversion of Speech to Text

- The user would speak into the mobile phone.
- Then the input word (destination) is converted to text and stored.
- The screen shows the word spoken by the user.

3.4 Steps in Navigation Using GPS

- Store the word spoken by the user in Database.
- Obtain the users' present coordinates by means of GPS.
- Retrieve the word from the database and match it with the current location to get the direction between the source and destination.

- The directions are given as voice output to the user which guides the user to reach the destination.

3.5 Recurrent Neural Networks – Long-Short Term Memory

The Long short-term Memory Networks keeps in track of the entire walk through by managing the information/data for longer period of time. This is done with the repeated neural network structure as in RNN. Here in this implementation of voice-bot for the visually challenged, the mobile application assists the person with the voice navigation based on the inputs of the user. Here the application predicts the next possible word that would be put forth by the user based on the previous ones. Moreover all the conversion of the applications is stored in the cloud server, through which an application can predict the words pretty good. The application in-turn, queries with the user and navigate to the location accordingly where the user can be more comfortable with application and avoid confusion.

4 Implementation of Voice Bot

The Fig. 4, that is depicted below clearly explains the development flow of an application in Android. The main activity takes the voice instruction as the input with the interfaces from Location-based services and speech synthesizing API. The storage of voice instructions, queries and answers, are stored in cloud through cloud server API.

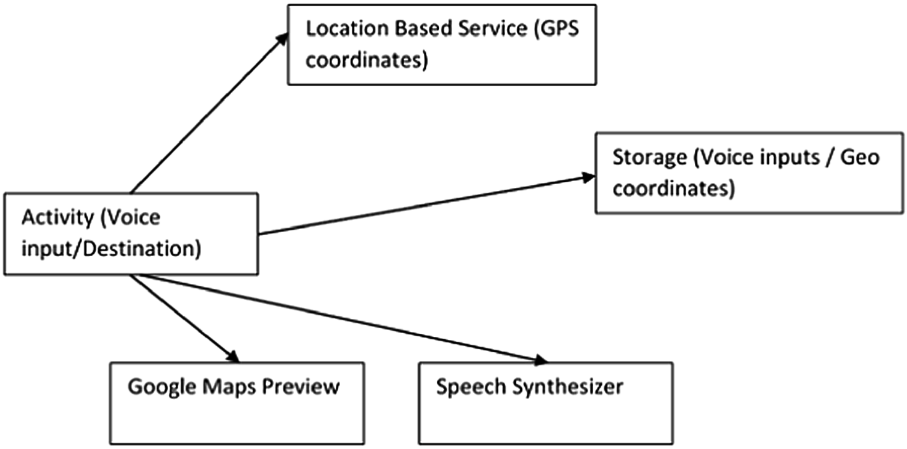

Location Based Service (GPS coordinates)
Storage (Voice inputs / Geo coordinates)
Activity (Voice input/Destination)
Google Maps Preview
Speech Synthesizer

Fig. 4. Application flow in Android

4.1 Workflow of the Application

The application will automatically take the current location (both latitude and longitude) of the user and then the user has to give the destination as voice input. It will automatically fetch the coordinates of the destination. Finally, a route is drawn between the current location and destination. The user receives the directions through voice navigation to reach his destination independently. The overall workflow is shown in the Fig. 5.

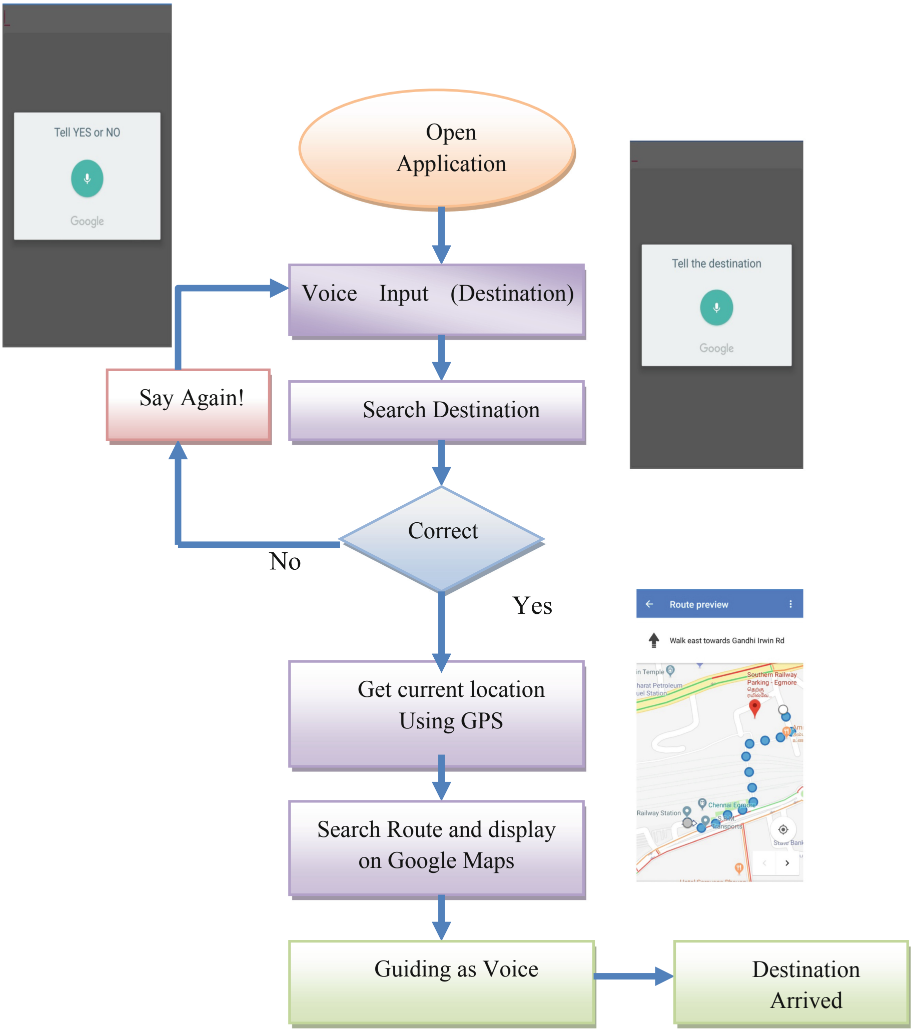


Fig. 5. Workflow of the application

4.2 Pseudo Code for Context-Aware Voice-Bot

This section elaborates the pseudo code for the working of voice-bot.

```
begin{
do{
function voice_input(){
    if ( voice_input == app_name){
        access_to_app = true;
    }if else{
        repeat voice_input = true;
    }else{
       access_to_app = false;
    }
}

fun database(){
    string questionframe;
    if (voicetotext == true){
         questionframe = 1;
      } else{
        return -1;
      }
}

fun CLSTM(string destination, string obstacle, string queries){
         string destination;
         string queries;
    if(voice_recognized == true && GPS == "ON"){
         getlatitude();
         getlongitude();
         display destimation;
    }else{
         return -1;
           }

   if(obstacle == 0){
         queries = "accepted";
         return destination;
         else
          return -1;
          }
}
```

4.3 Context Aware and Self Learning Feature

The voice recognizer is not immediately trained for perfection, hence the CLSTM captures the local features of the phrases as well many other global and temporal sentence semantics. The text are stored along with the location of the GPS. This is used in future whenever, the user visits the same location again. Then based on the location, the existing model searches for the keywords used earlier and gives a suggesting of the expected outcome.

5 Conclusion

The mobile application that is developed helps a visually impaired person to navigate independently and act as a travel assistant. As future work, voice navigations could be done as per the walking speed of the visually challenged person. Inertial sensors (accelerometers and/or gyroscopes) can be used to estimate the walking speed. And voice navigation could be said as per the speed calculated for the user. This would make the application context-aware. It is also extended to find the obstacle while walking which will help the user to walk faster without any hindrance. The queries by the user is captured in server.

References

1. Bansode, M., Jadhav, S., Kashyap, A.: Voice recognition and voice navigation for blind using GPS. Int. J. Innov. Res. Electr. Electron. Instrum. Control Eng. **3**(4), 91–94 (2015)
2. Sanchez, J., Espinoza, M., de Borba Campos, M., Merabet, L.B.: Accessibility for people who are blind in public transportation systems. In: Human Interfaces for Civic and Urban Engagement, pp. 8–12 (2013)
3. Guentert, M.: Improving public transit accessibility for blind riders: a train station navigation assistant. In: ASSETS 2011 (2011)
4. Rituerto, A., Fusco, G., Coughlan, J.M.: Sign based indoor navigation system for people with visual impairments. In: ASSETS 2016 (2016)
5. Gawari, H.: Voice and GPS based navigation system for visually impaired. Int. J. Eng. Res. Appl. **4**, 48–51 (2014)
6. Sánchez, J., Oyarzún, C.: Mobile audio assistance in bus transportation for the blind. Int. J. Disabil. Human Dev. **10**, 365–371 (2011)
7. Chumkamon, S., Tuvaphanthaphiphat, P., Keeratiwintakorn, P.: A blind navigation system using RFID for indoor environments. In: Proceedings of ECTI-CON, pp. 765–768 (2008)
8. Cha, J.S., Lim, D.K., Shin, Y.-N.: Design and implementation of a voice based navigation for visually impaired persons. Int. J. Bio-Sci. Bio-Technol. **5**(3), 61–68 (2013)
9. Gulati, R.: GPS based voice alert system for the blind. Int. J. Sci. Eng. Res. **2**(1), 1–5 (2011)
10. Angin, P., Bhargava, B.: Real-time mobile cloud computing for context-aware blind navigation. Int. J. Next-Gener. Comput. **2**(2), 405–414 (2011)
11. Azzouni, A., Pujolle, G.: NeuTM: a neural network-based framework for traffic matrix prediction in SDN. In: NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium (2018)

12. Sak, H., Senior, A., Beaufays, F.: Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In: INTERSPEECH (2014)
13. Uma Nandhini, D., Tamilselvan, L., UdhayaKumar, S., Silviya Nancy, J.: Client aware opportunistic framework of rich mobile applications in mobile cloud environment. Int. J. u-e- Serv. Sci. Technol. **10**(1), 281–288 (2017)
14. Poovam Raj, T.T., UdhayaKumar, S., Silviya Nancy, J.: Smart city based mobile application for seamless communication of differently-abled. Int. J. Multimedia Ubiquit. Eng. **12**(1), 177–190 (2017)
15. Nandhini, U., TamilSelvan, L.: Computational Analytics of Client Awareness for Mobile Application Offloading with Cloud Migration. KSII Trans. Internet Inf. Syst. **8**(11), 3916–3936 (2014). https://doi.org/10.3837/tiis.2014.11.014
16. Avanthika, U., Sundar, S., Nancy, S.: An interactive mobile application for the visually imparied to have access to listening audio books with handy books portal. Int. J. Interact. Mobile Technol. **9**(1), 64–66

Analysis of Graph Cut Technique for Medical Image Segmentation

Jyotsna Dogra(✉), Shruti Jain, and Meenakshi Sood

Department of Electronics and Communication,
Jaypee University of Information Technology, Solan, India
Jyotsnadogra1989@gmail.com, jain.shruti15@gmail.com,
meenusood9@gmail.com

Abstract. Segmentation plays an important role in image analysis as it is used to identify and differentiate foreground and background regions. Image segmentation in brain MRI analysis performs several roles like extraction of abnormal region for better diagnosis of the disease aiding in the therapy planning. Various brain tumors comprise diverse properties like their shapes, intensity distribution and location, hence reducing the possibility of developing a single general algorithm. In this paper authors have illustrated two methods for performing extraction which includes histogram thresholding and centroid based graph cut segmentation. On the basis of their potential, advantages and limitation comparison is made, that emphasize better performance of centroid based graph cut segmentation method. To measure the performance some quality parameters are evaluated. This paper also solves the problem of initial seed selection by using graph cut segmentation technique.

Keywords: Segmentation · Threshold · Fuzzy C-Mean clustering · K-mean clustering · Split and merge · Graph-Cut

1 Introduction

In the last few decades one of the most growing technique for providing the non-invasive image analysis for the diagnosis is Magnetic Resonance Imaging (MRI) technology. The huge amount of data analysis with good quality aids the expert radiologist in the study of brain anatomy. This huge brain data analysis maybe tedious with various difficulty due to the complex brain stricture but MRI is the most convenient and useful application in the imaging technology.

The formation of brain tumors is due to the irregular growth of the tissue mass that destroys the brain cells. The early diagnosis of tumor helps in increasing the survival rate [1]. Different growth rate of the tumor distributes the tumor in various grades belonging to the benign and malignant classes [2]. All these classes comprise different intensity distribution, locations and vary in size [3]. Brain tumor are one the most fatal disease and a high death rates have been observed in the developed countries [4]. Various Computer aided diagnosis methods provided for benefitting the clinicians with better diagnosis and testing are developed. Among them image segmentation is popularly used as it segments the image into different fragments generating a better

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 451–463, 2019.
https://doi.org/10.1007/978-981-15-0108-1_42

diagnosis of targeted region. Some of the conventional methods employed redundantly are intensity based, clustering based and some hybrid techniques.

A binary conversion of the MR image is done in threshold based segmentation [5–7] method. The iterative methods such as split and merge, region growing [8] are intensity dependent and only considers uniformity of the homogenous regions. Region growing method initializes with seed points and grouping is performed through predefined criteria such as similar intensity, color and texture. The selection of seed points is a hurdle and many algorithms are illustrated by researchers [9, 10]. The second iterative method is the split and merge technique [11] divides the complete image in smaller region and merges all the homogenous region to form one single targeted region. Some of the hybrid approaches studied by the authors [12, 13] include region growing and edge detection. The popular and most generally used clustering methods are: k-mean Clustering and Fuzzy clustering. A membership function is given to all the pixels that provide amount of degree present in an element in the fuzzy clustering also known as soft clustering [14]. On the other hand in k-mean Clustering [15] popularly also known as hard clustering generates different grouping of the image in such a way that each group shall depict highest similarity or minimum Euclidean distance from the mean value of every group. These techniques are completely dependent on the initial points that is the biggest struggle [16–18] and are iterative in nature. Due to the iterative nature, these methods are time consuming and face the problem for the initial seed point selection.

One of the prominent technique used is the graph based method proposed by Boykov *et al.* [19]. In graph cut technique minimum value of the objective function is obtained on the formation of the segmented image. Authors in [20] developed a graphical representation of the image that used cost function enabling the division of the images. The objective function introduced by authors [21–24] provided identical segmentation even though different combinatorial algorithm are used for evaluating minimum cut. Authors analyzed the efficiency in [23] for illustrating all the problems involved when both the 2D and 3D images are used for segmentation. Graph cut extends up to ND segmentation of images due to its ability of integrating the regional and boundary term. Other techniques like live wire and intelligent scissor [25, 26] evaluate the lowest weighted edge for the partition.

In this research paper we have compared two segmentation technique and evaluated the performance metric. The first technique is one of the simplest technique that is thresholding. The second technique is graph cut segmentation that has limitation of seed selection for the initialization of the algorithm. This limitation is avoided in the proposed framework by developing an automatic graph cut segmentation.

The organization of the paper is given as: detail explanation along of the two techniques along with their algorithms and mathematical explanation are described in Sect. 2. The simulation result, image database is described in Sect. 3. Quality measure of the segmented image is verified by four parameters which are briefly described and the values calculated are framed. Conclusion for obtained results from the proposed framework is explained in the Sect. 4.

2 Methodology

In this section flow diagram, algorithm and a detail explanation for the threshold segmentation and proposed framework for graph cut segmentation is provided. In the histogram thresholding method threshold value is evaluated from the histogram of difference intensity values. These values project the pixel location where the tumor may have occurred then, partition is done based on the threshold value and binary values are assigned to different regions. Many region based methods face problem for the initialization of seed points. To overcome this problem, graph cut segmentation explains image in a graphical form and automatic centroid/seed values are evaluated by exploiting the symmetrical nature of the brain. Different labelling is done for the object and the background region. Partition is performed by breaking the edges with lower thickness and s-t graph is formulated. Both methods are compared using mean square error (MSE), peak-signal-to-noise ratio (PSNR), similarity index (SSIM) and dissimilarity index (DSSIM). Experimental results show that our approach of centroid based graph cut segmentation outperforms the histogram thresholding technique.

2.1 Histogram Thresholding

Thresholding technique is one of the popular techniques due to its simplicity. It converts MR image into a binary image, distinctly displaying the segmented region. Complete process of thresholding relies on the selection of the threshold value. This value is evaluated by means of number of pixels corresponding to a particular pixel value which are obtained by histogram. There are three types of thresholding technique: global threshold, variable threshold and regional threshold. In this paper global thresholding is employed and a single threshold value is evaluated by Otsu's method for the entire image.

Flow diagram of the histogram thresholding algorithm is given in Fig. 1. The MR image which is in RGB form is converted into gray scale image. Single threshold value is calculated for the entire image in two steps: the image is divided vertically, then, difference in the pixel intensity between the vertical sections is calculated and their histogram is computed. Using Otsu's method, the difference value is used to calculate the threshold value. All the pixels are compared about the threshold value. Pixels are labelled as '1' (representing white intensity value) if they belong to object region and as '0' (representing black intensity value) if they belong to background region. Hence, this algorithm provides a binary partition of the MR images. This partition is only possible if the accurate knowledge of the threshold value is obtained. If a single threshold value is selected, then the image is partitioned in binary form but on selecting higher value of the threshold the number of regions formed also increase. Therefore, the accurate knowledge of the threshold value is of critical importance. After the pixels are labelled the targeted tumor region is extracted from the MRI image and finally the segmentation is done.

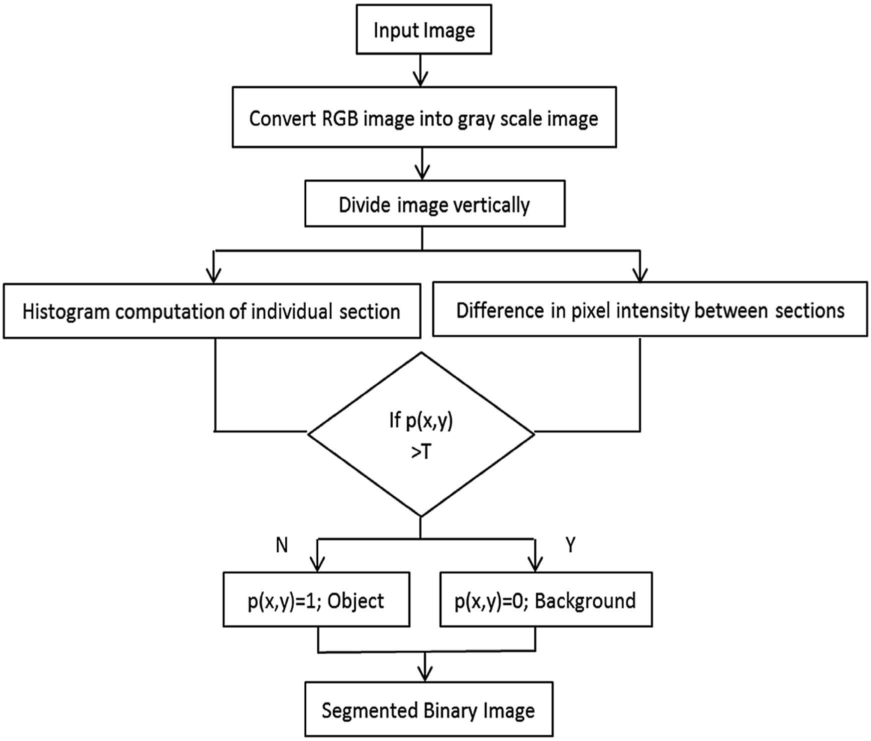

Input Image
Convert RGB image into gray scale image
Divide image vertically
Histogram computation of individual section
Difference in pixel intensity between sections
If p(x,y) >T
N
Y
p(x,y)=1; Object
p(x,y)=0; Background
Segmented Binary Image

Fig. 1. Flow diagram of overall algorithm

2.2 Graph Cut

This segmentation technique divides the complete set of the image in two subgroups in such a way that the targeted region is presented as the foreground and the remaining as the background region. Graph cut represents the image in a graphical form such that each pixels are represented as nodes, illustrated in Figs. 2 and 3.

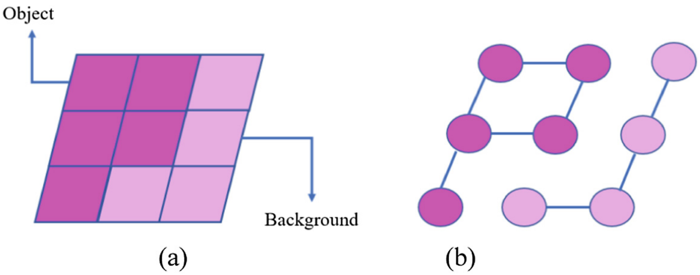


Fig. 2. (a) Graphical form of image (b) segmented image.

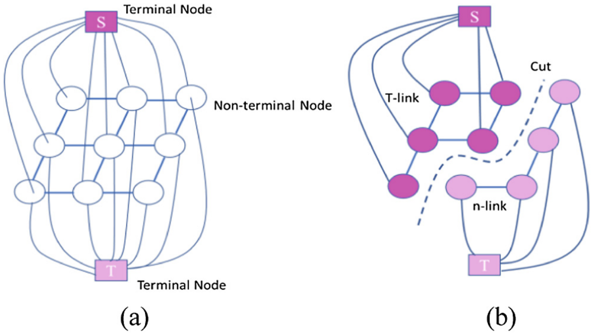


Fig. 3. (a) Source terminal and sink terminal (b) cut to segment object and background.

The basic representation of the image in the graphical form is given in Fig. 3(a) was proposed by Boykov *et al.* in [27] as:

$$G = <V, E> \tag{1}$$

$$\mathcal{V} = \{s, t\} \cup \mathcal{P} \tag{2}$$

where, graph G contains nodes/pixels $\mathcal{V}$ and edges/neighboring distance $\mathcal{E}$. The pixels $\mathcal{P}$ are nodes and the seed values are the non-terminal nodes (s, t) (s: source/object terminal; t: sink/background terminal) that are connected to each other via link. There are two types of links present: $t - link$ and. The $t - link$ connects the terminal nodes to the non-terminal nodes and $n - link$ connects all the neighboring nodes to each other. The partition takes place when two homogeneous regions are formed that show the highest similarity with the terminal nodes. This cut is shown in Fig. 3(b) that provides two exclusive regions. The flow diagram of this methodology is depicted in Fig. 4. The RGB image is converted into gray scale image and the image is divided in vertical sections as previously done in histogram thresholding. Both the symmetrical halves are compared and the pixel values with highest difference are obtained. These are the centroid points that provide the seed point for initializing the graph cut segmentation. Once the seed points are known the segmentation is performed by assigning the weights and labels are assigned corresponding to them.

3 Simulation Results and Discussion

3.1 Study Area and Dataset

The experiments were performed on MATLAB 2013a on the standard database [28, 29]. The clinical data was provided from radiology department of PGI Chandigarh as shown in Fig. 5(c). The sizes of three images are 180 × 218, 800 × 450 and 960 × 1280 respectively as given in Fig. 5(a), (b) and (c).

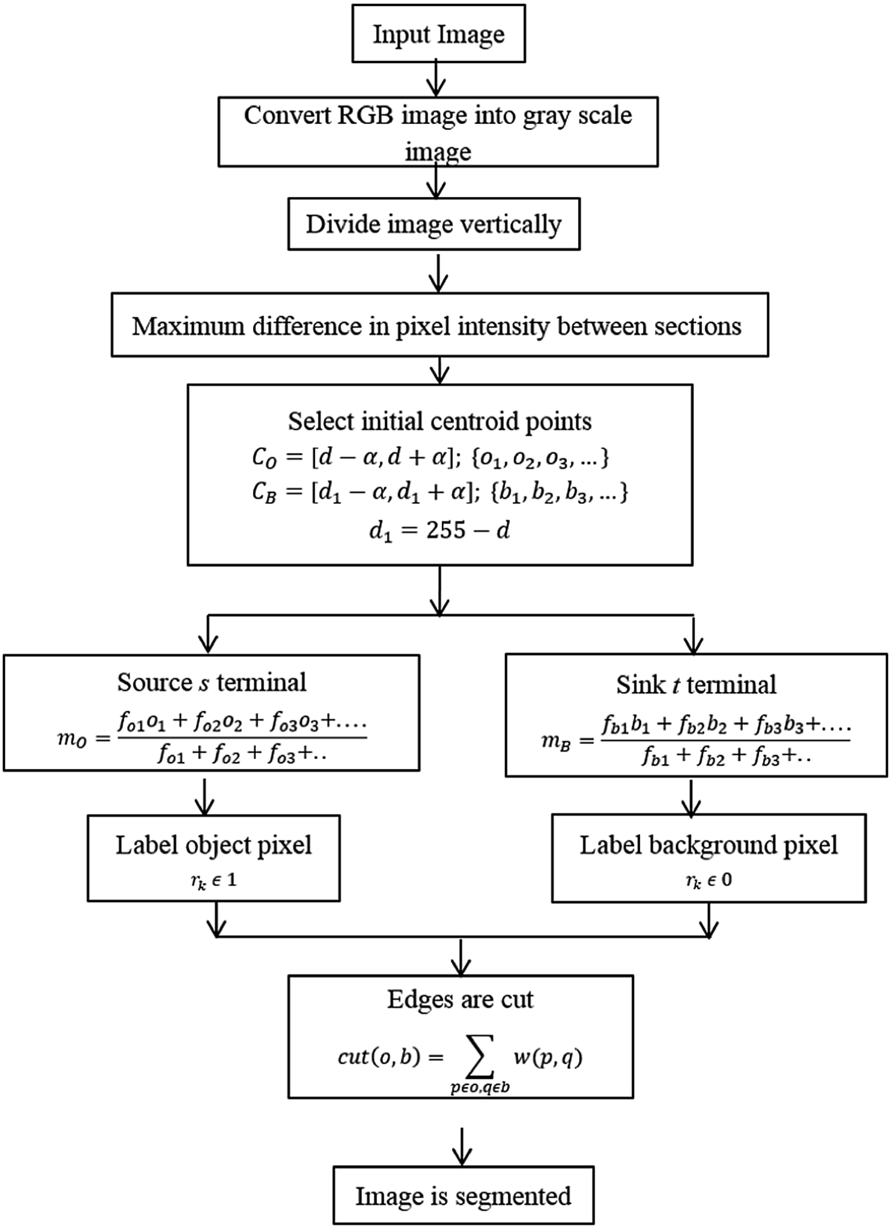

Input Image
Convert RGB image into gray scale image
Divide image vertically
Maximum difference in pixel intensity between sections
Select initial centroid points
$C_O = [d - \alpha, d + \alpha]; \{o_1, o_2, o_3, \ldots\}$
$C_B = [d_1 - \alpha, d_1 + \alpha]; \{b_1, b_2, b_3, \ldots\}$
$d_1 = 255 - d$
Source *s* terminal
$m_O = \frac{f_{o1}o_1 + f_{o2}o_2 + f_{o3}o_3 + \ldots.}{f_{o1} + f_{o2} + f_{o3} + ..}$
Sink *t* terminal
$m_B = \frac{f_{b1}b_1 + f_{b2}b_2 + f_{b3}b_3 + \ldots.}{f_{b1} + f_{b2} + f_{b3} + ..}$
Label object pixel
$r_k \epsilon 1$
Label background pixel
$r_k \epsilon 0$
Edges are cut
$cut(o, b) = \sum_{p\epsilon o, q\epsilon b} w(p, q)$
Image is segmented

Fig. 4. Flow chart of automatic graph cut.

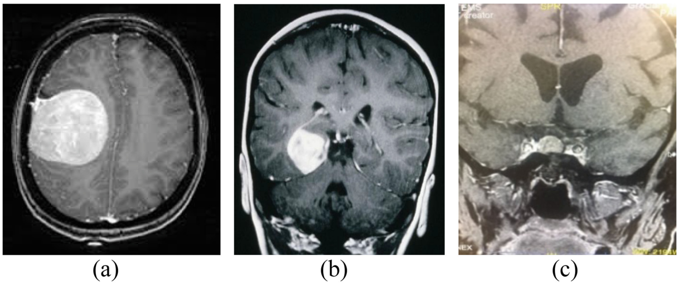


Fig. 5. Different set of original images (a), (b) and (c)

3.2 Quality Measure

For the performance analysis and quality measurement of the image, following parameters obtained from literature are evaluated: Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM) and Dissimilarity Index (DSSIM).

3.2.1 Mean Square Error

MSE measures the average difference of the pixels throughout the image. Higher value implies greater difference of the segmented image from the original image. It is calculated as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(I_{xi} - I_{yi}\right)^2 \tag{3}$$

where, n, is the total no of pixel, I_{xi} is the intensity value of the original image and I_{yi} is the intensity value of the segmented image.

3.2.2 Peak Signal to Noise Ratio

PSNR computes how much noise is present in the resultant image due to segmentation and it is measured in decibel. It is computed as follows:

$$PSNR = 10\log\frac{255 * 255}{MSE} \tag{4}$$

Ideally the PSNR value is infinite but practically it is 20 to 30 dB.

3.2.3 Structural Similarity Index (SSIM)

Authors in [30] presented SSIM as a quality measure parameter that calculates any disturbance in the segmented region of interest if created by the simulation. It correlates the local patterns of pixel intensities that are normalized.

$$SSIM = \frac{\left(2m_x m_y + K_1\right)\left(2\sigma_{xy} + K_2\right)}{\left(m_x^2 + m_y^2 + K_1\right)\left(\sigma_x^2 + \sigma_y^2 + K_2\right)} \tag{5}$$

where, $\boldsymbol{m_x}, \boldsymbol{m_y}$ are the mean value of the original region of interest (infected region in the original image) and the segmented region of interest (infected region in the segmented image) respectively, σ_x, σ_y are the variance respectively. K_1, K_2 are the substitution values, in this paper we have taken these values 0.01 and 0.03 which are calculated by Wang *et al.* in [30]. The values should range between 0 and 1, exhibiting higher similarity if the values are nearby 1 and lower similarity for values nearby 0.

3.2.4 Dissimilarity Index

This parameter evaluates the pixels that have dissimilar intensity value from the original image. It computes adverse effect from the SSIM. Similar to SSIM its value also range between **0** and **1** but differs in the implications. For values near to **0** represent lower dissimilarity and value near to **1** show higher dissimilarity. This value is calculated as follows:

$$\boldsymbol{DSSIM} = \mathbf{1} - \boldsymbol{SSIM} \tag{6}$$

3.3 Results and Discussion

In this research work, tumor affected MRI image is segmented by thresholding and graph cut method as shown in Fig. 6. In threshold segmentation resulting image is binary image that is segmented at threshold value of T = 0.489. It is observed that region of interest is extracted but erroneous pixels also get included and it is difficult to differentiate between tumor region and other part of the brain. Whereas, the brain segmented by graph cut technique has a clear visibility of the tumor region. The segmented image is not binary, and extracted region's intensity values are similar to the tumor region in the original image.

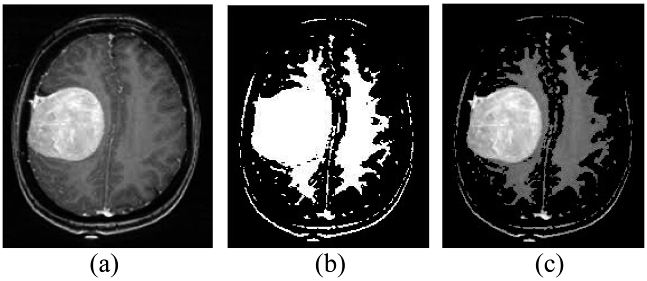


Fig. 6. Segmentation results of Image 1 (a) original image, (b) thresholding, (c) graph cut

Segmented results of the Image-2 are shown in Fig. 7. The original MRI image is of a young child who is diagnosed by astrocytoma which occurs due to the astrocyte cells. Image segmented by threshold technique is a binary image. Tumor region is not extracted properly due to inclusion of pixels whose intensity value is close to the pixels in the tumor region. On the contrary, image obtained by graph cut segmentation has a clear picture of tumor region and no erroneous pixels are included in the ROI.

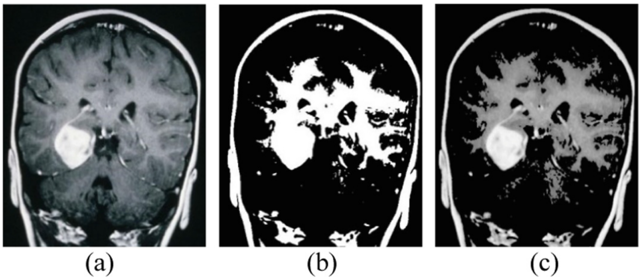


Fig. 7. Segmentation results of Image 2 (a) original image, (b) thresholding, (c) graph cut

In Fig. 8 segmented results of the third image are shown using both the techniques. Image obtained by thresholding technique is binary but it is very difficult to locate the tumor region. Most of the region are holding the value 1 hence majority region in the image is white.

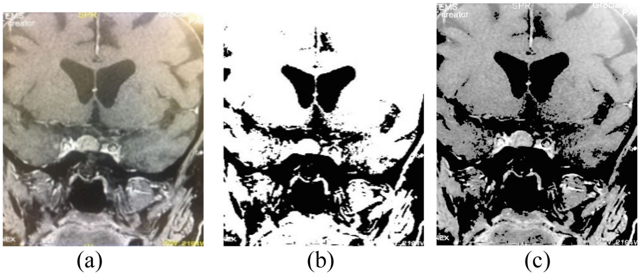


Fig. 8. Segmentation results of Image 3 (a) original image, (b) thresholding, (c) graph cut

This implies that majority region in the original image are of high contrast. Tumor is successfully segmented by the proposed Graph cut segmentation technique. From the results it is observed that resultant image obtained by thresholding process extracts the tumor portion completely comprising of binary intensity values i.e. 0 and 1.

Threshold value of $T = 0.4980$ is obtained by calculating difference between two horizontal half of the image. Although this technique is one of the fast, simplest and oldest but the infected tumor region loses its original intensity value and returns value **0** or black color in the infected region. In the second method object and background centroid are automatically initialized, where 20 seed points for object/tumor region and 40 seed points for the background region are selected in this method. Variation in number of seed points results in similar mean/centroid value. It is observed from the images that the ROI/tumor region hold their original pixel value which is an advantage for the clinical application where the pixel intensity of the region of interest is measured to calculate the various parameter for the diagnosis of tumor. To analyze performance of the two methods MSE and PSNR quality measure are evaluated and are tabulated in Table 1. It is observed that MSE values for threshold segmentation method are very high for all three images. High MSE signifies that difference level between resultant images and original images is very high. The results also show very low PSNR value for threshold segmented images which ideally should be between 20 to 30 dB. However, results of graph cut segmented images show better result as the MSE is low and PSNR value is between 20 to 30 dB. These quality measures illustrate that graph cut method gives better result than threshold segmentation method.

Table 1. MSE and PSNR values of segmented images (by Threshold and Graph Cut method).

	Image-1		Image-2		Image-3	
Parameter	Threshold	Graph cut	Threshold	Graph cut	Threshold	Graph cut
MSE	36252.75	29.284	36018.45	37.506	26307.72	48.78
PSNR	2.503	27.202	2.218	26.197	3.895	25.125

In Table 2 similarity and dissimilarity index of the three images are evaluated. It is observed that SSIM values for threshold segmented images have very low similarity index and a very high dissimilarity index. This implies that the extracted image is quite different from the original image. Whereas, graph cut method shows high similarity index and low dissimilarity index as shown in Table 2.

Table 2. SSIM and DSSIM values of segmented images (by Threshold and Graph Cut method).

	Image-1		Image-2		Image-3	
Parameter	Threshold	Graph cut	Threshold	Graph cut	Threshold	Graph cut
SSIM	0.603	0.007	0.709	0.008	0.058	0.009
DSSIM	0.397	0.993	0.291	0.992	0.942	0.991

For ease of better clarification and understanding of performance of both techniques on the basis of quality measure are represented in graphical representation. The comparison of the two techniques with respect to the parameters in Tables 1 and 2 are shown in Figs. 9 and 10 respectively.

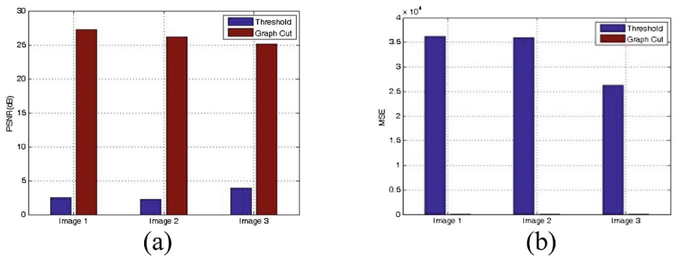


Fig. 9. Comparison of threshold and graph cut segmentation on the basis of parameter (a) MSE and (b) PSNR

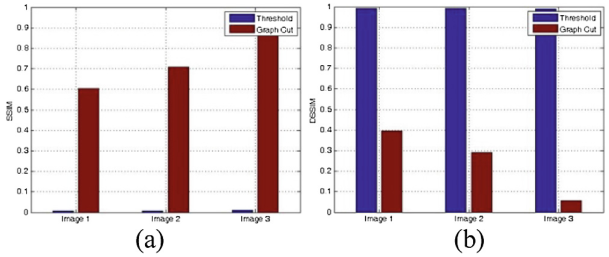


Fig. 10. Comparison of threshold and graph cut segmentation on the basis of parameter (a) SSIM and (b) DSSIM.

4 Conclusion

In this paper we have proposed two techniques for image segmentation and evaluated the effectiveness on the basis of quality measure parameters. In the first technique single threshold value is used for segmenting the three images and corresponding binary images are obtained it is observed that tumor region is not clearly visible for all three images. This is because erroneous pixels are included in the infected region which may result in false diagnosis. This is validated by the evaluated MSE and PSNR. Better segmentation is observed from the second technique of automated graph cut method that generates the seed points resulting in the efficient and most accurate segmentation. The quality measure evaluated for the validation of automated graph cut method for MSE, PSNR parameter are 29.3, 27.2 for image 1; 37.5, 26.2 for image 2 and 48.8, 25.1 for image 3 respectively. The SSIM and DSSIM values obtained are 0.007, 0.993 for image 1; 0.008, 0.992 for image 2 and 0.009, 0.991 for image 3 respectively. The evaluated results imply effectiveness of the proposed graph cut segmentation irrespective of any irregular tumor shape and a clear visibility of extracted tumor.

References

1. Kotsas, P.: Non-rigid registration of medical images using an automated method, pp. 199–201. IEC, Prague (2005)
2. Nagalkar, V., Asole, S.: Brain tumor detection using digital image processing based on soft computing. J. Signal Image Process. **3**, 102–105 (2012)
3. Mirajkar, G., Barbadekar, B.: Automatic segmentation of brain tumors from MR images using undecimated wavelet transform and gabor wavelets. In: 17th IEEE International Conference on Electronics, Circuits, and Systems (ICECS), pp. 702–705 (2010)
4. Lin, C.-T., Yeh, C.-M., Liang, S.-F., Chung, J.-F., Kumar, N.: Support-vector-based fuzzy neural network for pattern classification. IEEE Trans. Fuzzy Syst. **14**, 31–41 (2006)
5. Cheriet, M., Said, J.N., Suen, C.Y.: A recursive thresholding technique for image segmentation. IEEE Trans. Image Process. **7**, 918–921 (1998)
6. Sezgin, M., Sankur, B.: Selection of thresholding methods for nondestructive testing applications. In: Proceedings of International Conference on Image Processing, pp. 764–767 (2001)
7. Li, Z., Liu, G., Zhang, D., Xu, Y.: Robust single-object image segmentation based on salient transition region. Pattern Recogn. **52**, 317–331 (2016)
8. Manousakas, I., Undrill, P., Cameron, G., Redpath, T.: Split-and-merge segmentation of magnetic resonance medical images: performance evaluation and extension to three dimensions. Comput. Biomed. Res. **31**, 393–412 (1998)
9. Adams, R., Bischof, L.: Seeded region growing. IEEE Trans. Pattern Anal. Mach. Intell. **16**, 641–647 (1994)
10. Fan, J., Yau, D.K., Elmagarmid, A.K., Aref, W.G.: Automatic image segmentation by integrating color-edge extraction and seeded region growing. IEEE Trans. Image Process. **10**, 1454–1466 (2001)
11. Hancer, E., Karaboga, D.: A comprehensive survey of traditional, merge-split and evolutionary approaches proposed for determination of cluster number. Swarm Evol. Comput. **32**, 49–67 (2017)
12. Gambotto, J.-P.: A new approach to combining region growing and edge detection. Pattern Recogn. Lett. **14**, 869–875 (1993)
13. Pavlidis, T., Liow, Y.-T.: Integrating region growing and edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **12**, 225–233 (1990)
14. Pal, N.R., Pal, S.K.: A review on image segmentation techniques. Pattern Recogn. **26**, 1277–1294 (1993)
15. Jain, A.K.: Data clustering: 50 years beyond k-means. Pattern Recogn. Lett. **31**, 651–666 (2010)
16. Dhanachandra, N., Chanu, Y.J.: Image segmentation method using k-means clustering algorithm for color image. Adv. Res. Electr. Electron. Eng. **2**(11), 68–72 (2015)
17. Despotovic, I., Vansteenkiste, E., Philips, W.: Spatially coherent fuzzy clustering for accurate and noise-robust image segmentation. IEEE Signal Process. Lett. **20**, 295–298 (2013)
18. Chuang, K.-S., Tzeng, H.-L., Chen, S., Wu, J., Chen, T.-J.: Fuzzy c-means clustering with spatial information for image segmentation. Comput. Med. Imaging Graph. **30**, 9–15 (2006)
19. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient ND image segmentation. Int. J. Comput. Vis. **70**, 109–131 (2006)
20. Wu, Z., Leahy, R.: An optimal graph theoretic approach to data clustering: theory and its application to image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **15**, 1101–1113 (1993)

21. Ford Jr., L.R., Fulkerson, D.R.: Flows in Networks. Princeton University Press, Princeton (2015)
22. Goldberg, A.V., Tarjan, R.E.: A new approach to the maximum-flow problem. J. ACM (JACM) **35**, 921–940 (1988)
23. Boykov, Y., Kolmogorov, V.: Computing geodesics and minimal surfaces via graph cuts. In: null, p. 26 (2003)
24. Heimowitz, A., Keller, Y.: Image segmentation via probabilistic graph matching. IEEE Trans. Image Process. **25**, 4743–4752 (2016)
25. Mortensen, E.N., Barrett, W.A.: Interactive segmentation with intelligent scissors. Graph. Models Image Process. **60**, 349–384 (1998)
26. Falcão, A.X., Udupa, J.K., Miyazawa, F.K.: An ultra-fast user-steered image segmentation paradigm: live wire on the fly. IEEE Trans. Med. Imaging **19**, 55–62 (2000)
27. Boykov, Y.Y., Jolly, M.-P.: Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. In: Proceedings of Eighth IEEE International Conference on Computer Vision, pp. 105–112 (2001)
28. Stubberfield, L.: Big Picture. https://bigpictureeducation.com
29. MathWorks. https://in.mathworks.com/matlabcentral.com
30. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**, 600–612 (2004)

A Buyer and Seller's Protocol via Utilization of Smart Contracts Using Blockchain Technology

Priyanka Kumar(✉), G. A. Dhanush, D. Srivatsa, A. Nithin, and S. Sahisnu

Department of Computer Science and Engineering,
Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India
k_priyanka@cb.amrita.edu, {cb.en.u4cse17415,
cb.en.u4cse17459, cb.en.u4cse17443,
cb.en.u4cse17450}@cb.students.amrita.edu

Abstract. In this paper we have proposed and implemented a buyer and seller's protocol on ethereum platform via utilization of smart contracts using blockchain technology. We have discussed the shift towards the digitization of registering lands and the necessity to use Blockchain instead of traditional storage technologies and provide a brief description and working implementation of the proposed model.

Keywords: Smart contract · Land registry · Blockchain technology · Ethereum · Ether

1 Introduction

Blockchain technology is consensus-based computing which register all transactions without involvement of any Third Party [1]. With the advent of Blockchain technology, we see a shift towards the use of it in systems where transactions frequently occur. It is decentralized, distributed and immutable in nature. Because of these properties blockchain technology has become so popular and it has changed a lot of things [2]. In recent years digitalization is on extreme demand where vital documents such as passports, birth records, medical records, land records and even transactions are being digitized. Digitalization has not just improved security but also time and effort to maintain all records has reduced [4]. It has a great impact on every aspects of life [3]. For example, in banking and finance department, governmental parties, Chambers of Commerce and Land Registry and in healthcare department [5] (Fig. 1).

At present there are many malpractices happening in the field of Land registry. We have many users claiming ownership over the same piece of land. We also have another problem as the history of the piece of land at present is difficult to be traced back and takes longer time even if it is possible. This paper deals with the present scenario of transaction of lands and also presents a model for buying and selling land

A. K. Luhach et al. (Eds.): ICAICR 2019, CCIS 1075, pp. 464–474, 2019.
https://doi.org/10.1007/978-981-15-0108-1_43

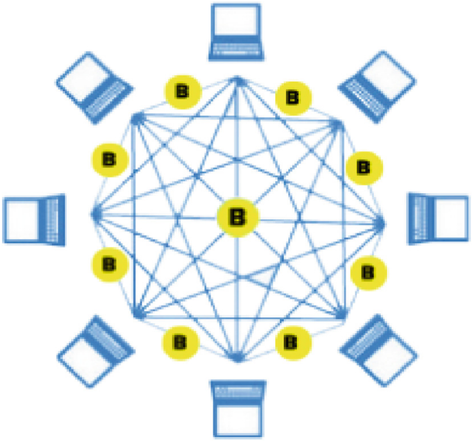


Fig. 1. Distributed ledger [6]

over the Blockchain network using Ethereum platform as it acts as a layer to prevent the malpractices in buying and selling of land. The current paper based method has proven itself to be a convenient and a primary method to deal with the transaction of physical entities. A written form, if irreplicable, can act as a token of ownership of land. But unfortunately it has many security, lack of traceability and transparency challenges. Mostly seller and buyer have to depend on an untrusted third party broker which results in many forgery cases. Also, registration process is entirely paper based which makes it time consuming. So with these motivations we have taken one application named "land registry" and proposed a theoretical model for seller and buyers protocol. We have also implemented the secure digital ledger of land related transactions on ethereum platform and shown the contract between two trusted parties which is faster, secure and immutable. It keeps tracks of all transactions in distributed manner and uses one shared database for records with some back-up facilities and ensures that there is always a single owner. Our proposed architecture is based on timestamp which keeps track of the creation time and modification time of a document or transaction. So no one can edit the (content of the) document or transaction once it has been recorded [6]. Even the owner of the document can't change the document once it is recorded on distributed ledger [7]. Also, proposed seller and buyer's protocol consists of all transactions as a part of distributed ledger which are traceable (validation). Anyone who would like to upload a transaction on the blockchain can do so (Fig. 2).

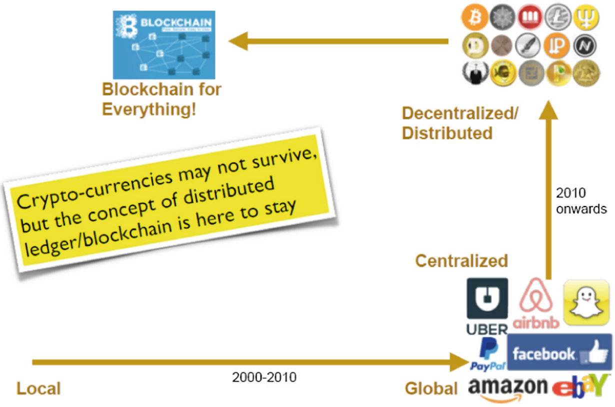

BLOCKCHAIN
Blockchain for Everything!
Decentralized/ Distributed
Crypto-currencies may not survive, but the concept of distributed ledger/blockchain is here to stay
2010 onwards
Centralized
UBER
airbnb
PayPal
facebook
Local
2000-2010
Global
amazon
ebay

Fig. 2. Ever changing landscape of communication [6]

Smart contracts are digital agreements between transacting parties that are written in computer to code and deployed to the blockchain, where they will self-execute when predetermined conditions are met [5]. They reduce complexity in land registry through automated verification and execution of the multiple business transactions involved [8, 9]. A decentralized, immutable record also ensures all sellers and buyers have equal access to information and helps build trust (Fig. 3).

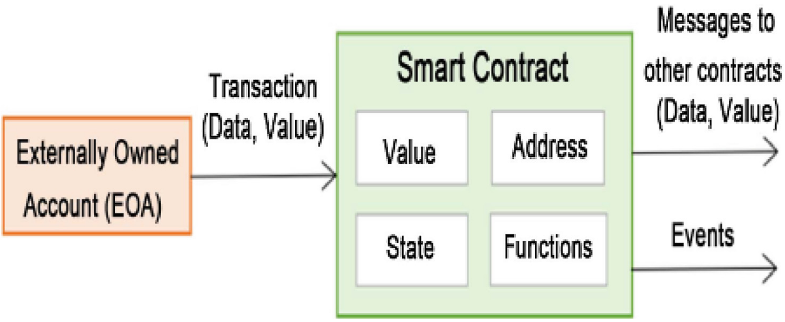

Externally Owned Account (EOA)
Transaction (Data, Value)
Smart Contract
Value
Address
State
Functions
Messages to other contracts (Data, Value)
Events

Fig. 3. A basic structure of smart contract [8]

Smart contracts improve the transparency, traceability and efficiency of seller and buyer protocol [11]. This ethereum platform are specially conceived for smart contract and decentralized application development [10, 12]. In our contribution part, preliminary transaction shows that the contracts are able to interact with one another.

Roadmap. The paper is organized as follows. We describe our proposed system model in Sect. 2. In Sect. 3, we formally show the implementations of our proposed model. In Sect. 4 we formally describe the results obtained so far and user validation and successfully retrieval of land record using blockchain technology. Finally we conclude in Sect. 5.

2 System Architecture

Our Proposed theoretical model is depicted below. To accomplish the task of processing land registrations using the help of blockchain, we use the smart contract ethereum platform and blockchain technology. Firstly, we create a file associated with land, called a land document. This land document holds the necessary fields that describe the land. The fields are as follows:

- Land_id: It is a unique id associated with the land.
- Land_coordinates: It holds the central land coordinates.
- Land_dimensions: It holds the land dimensions.
- Land_address: It holds the land location address.
- Land_owner: It holds the name of the landowner.
- Land_owner_id: It holds the unique id associated with the owner.
- Previous_owners: It holds a list of pairs of owner_name and owner_id who previously owned the land.
- Land_status: It holds a flag value of whether the land is available for sale or not.

New land documents can only be created by registrars (a person, who on behalf of the government, acknowledges the presence of the land and the correctness of the details of land). Every time the land is purchased by a new owner, Land_owner field of this document is updated to the current owner and the previous owner is added in the list of Previous_owners. This process is done by the use of Smart Contracts, which will be briefed in the later sections of this topic. Also, each verified land document (verification is the process done by a registrar) is digitally signed by the registrar and can be traced back to him/her who is present in the network for validation (it is the process by which the details in the network is checked of correctness). Now, any land with Land_status 'For Sale' can only be negotiated for purchase. We provide a separate portal which maintains the record of lands that are available for sale. Secondly, each customer (here, customers refer to both landowner and a new land buyer) holds an account in this platform. To introduce previously existing landowners into this system, we allow a registrar to update the Land_owner, Land_owner_id of the land to the corresponding customer. This account is also linked with the Ethereum crypto wallet (digital wallet which helps to store and maintain cryptocurrency) which helps the customers purchase land or earn money by selling the land. The lands owned by a customer is also stored in his/her account. Thirdly, to initiate a new land transaction (transfer of land from a landowner to a buyer), we use Smart Contracts. A buyer who wants to purchase land on sale initiates this Smart Contract by passing the land document address (an address that refers to the document) as an argument. This Smart Contract then performs the following actions:

- It fetches the Land_owner_id from the land document and sends a notification to the landowner's (or seller's) account that a buyer wishes to buy his/her land.
- If the seller agrees to the request, a notification is sent back to the buyer announcing the price of the land.
- If the buyer accepts the amount, the corresponding quantity of Ethers from the wallet of the buyer is transferred to the wallet of the seller

- Then the land_owner_id from the land document is changed to the buyer's ID also the seller's detail is moved into Previous_owners field.
- The land buyer's account is updated with the addition of the details about the land. A field that shows the possession of land in the previous landowner account is removed.

Model Diagram

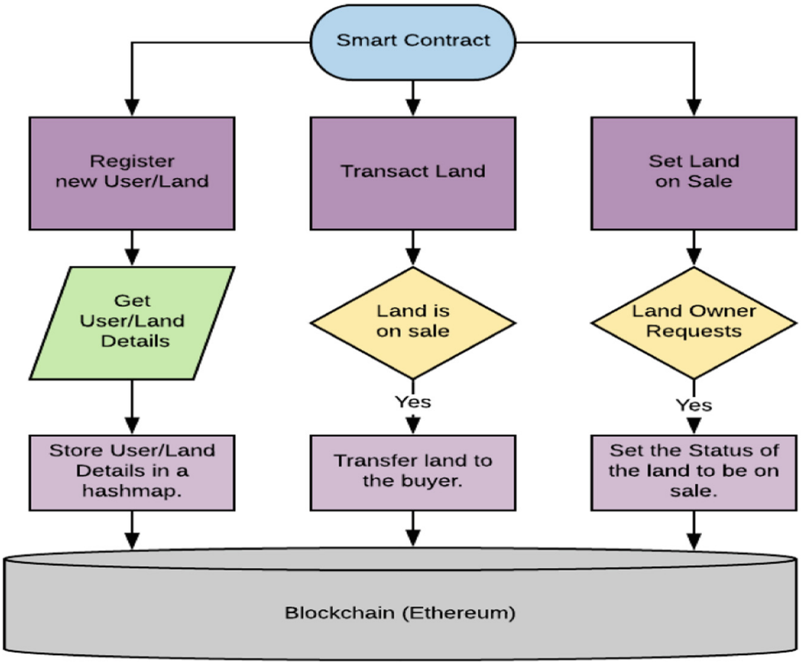
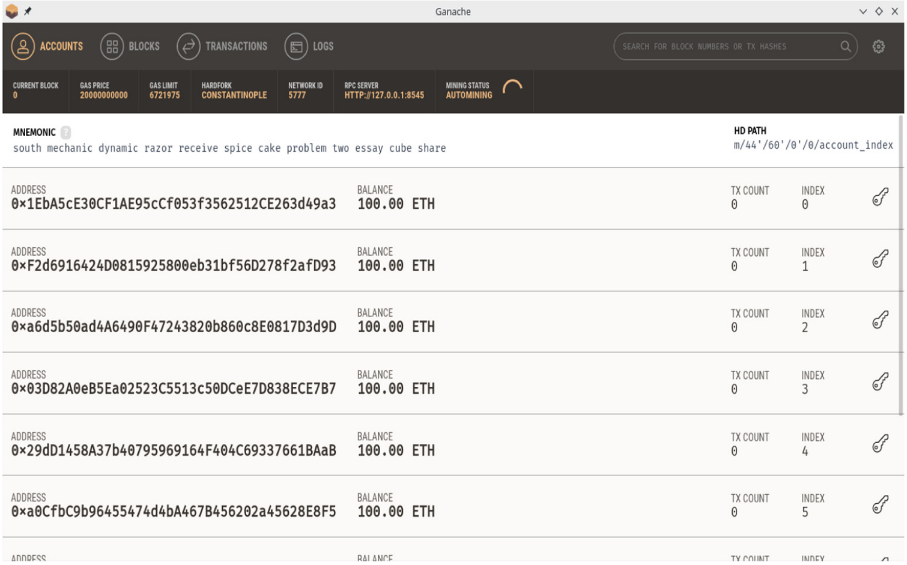


Fig. 4. Proposed architecture

Blockchain technology is immutable as well as distributed technology where all the data is been replicated to all the participating clients in a blockchain. So if we want to tamper a record, we need to tamper data on each client to whom data has been distributed which is nearly impossible because of the availability of computing power. Hence the blockchain is seen as one of the most secure technology ever the world has witnessed. Integrating blockchain to secure birth records will not just only save paperwork but it will be more secure and all the fake birth certificate rackets would shut down, no more duplicate birth certificate would be valid. Also we can easily share Birth certificate on blockchain which can be used for verification purpose such as passport verification, Aadhar card verification, pan card verification etc. The proposed model consists of the following parts (See Fig. 4).

3 Our Contribution: Implementation on Ethereum Platform

The particular application has been developed using Blockchain technologies provided by Ethereum. The core of this application is handled by smart contracts. All the programming logic required to register users, lands and handle land transactions are written in smart contracts using the programming language called Solidity. This language is similar to javascript but helps us prevent critical errors like memory leaks, type errors etc. These smart contracts are compiled and then run on the Ethereum Virtual Machine (EVM). This Virtual Machine provides the resources for our smart contract to execute and its specialty is that it runs on the Ethereum Blockchain. Each state of a program is recorded as a transaction in the blockchain. To test this application, we run it in our local network using tools like Ganache and Truffle. Ganache simulates a local blockchain in our test environment by hosting a local blockchain and providing a few accounts to perform transactions. Truffle is used to compile the written smart contract and migrate (deploy) the contract to a running instance of a blockchain network. In our test network, we provide the address as that of the localhost (//127.0.0.1) and the port as 8545 (we can set the port value manually). After setting up all this to deploy our contract in a blockchain network, we could interact with our smart contract using 'truffle console' command. This command starts a console in which we could interact with our deployed contract and call the functions linked with our contract (Fig. 5).



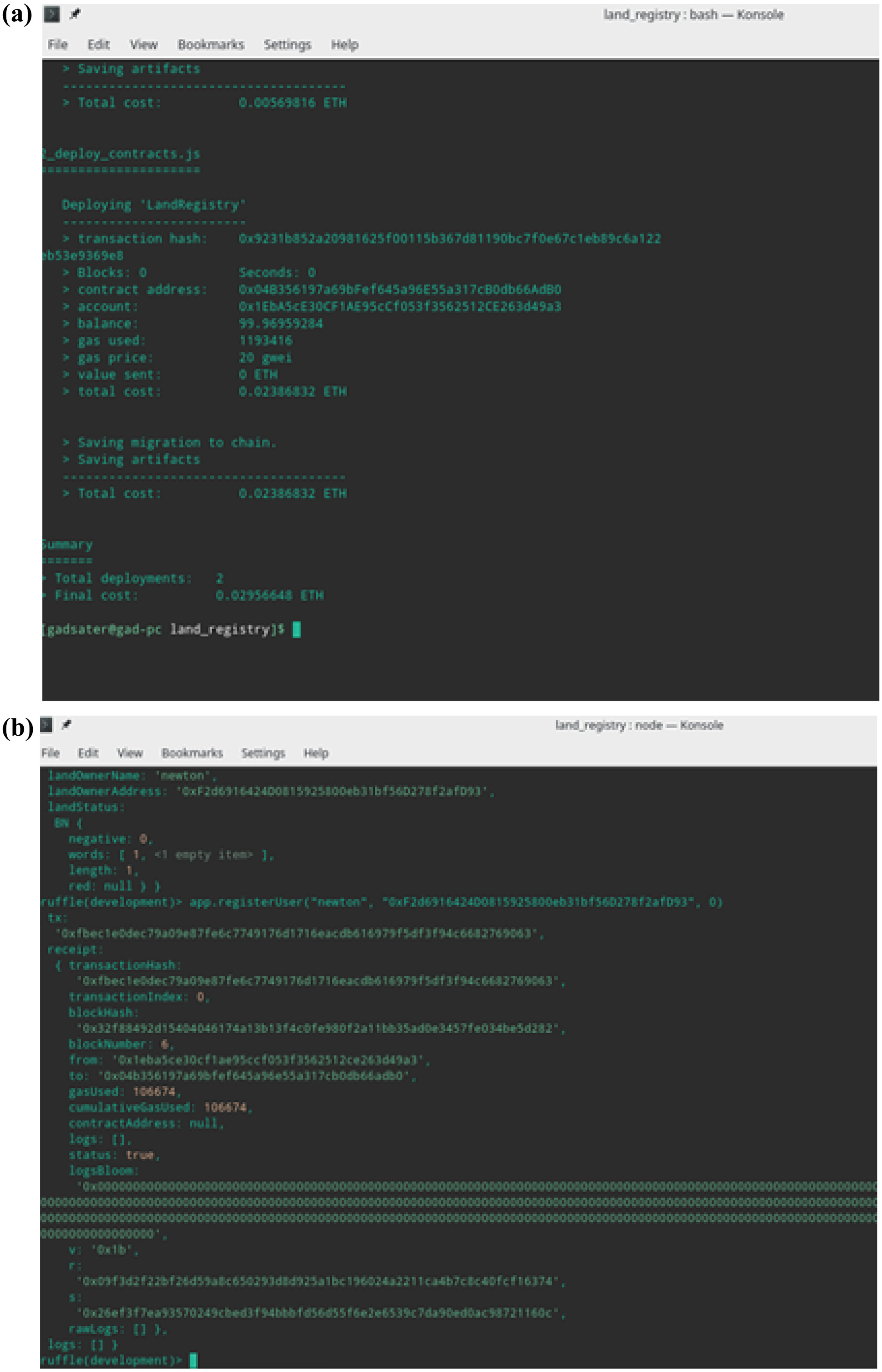
CURRENT BLOCK	GAS PRICE	GAS LIMIT	HARDFORK	NETWORK ID	RPC SERVER	MINING STATUS
0	20000000000	6721975	CONSTANTINOPLE	5777	HTTP://127.0.0.1:8545	AUTOMINING

MNEMONIC
south mechanic dynamic razor receive spice cake problem two essay cube share

HD PATH
m/44'/60'/0'/0/account_index

ADDRESS	BALANCE	TX COUNT	INDEX
0×1EbA5cE30CF1AE95cCf053f3562512CE263d49a3	100.00 ETH	0	0
0×F2d6916424D0815925800eb31bf56D278f2afD93	100.00 ETH	0	1
0×a6d5b50ad4A6490F47243820b860c8E0817D3d9D	100.00 ETH	0	2
0×03D82A0eB5Ea02523C5513c50DCeE7D838ECE7B7	100.00 ETH	0	3
0×29dD1458A37b40795969164F404C69337661BAaB	100.00 ETH	0	4
0×a0CfbC9b96455474d4bA467B456202a45628E8F5	100.00 ETH	0	5

Fig. 5. The interface of Ganache application



(a)

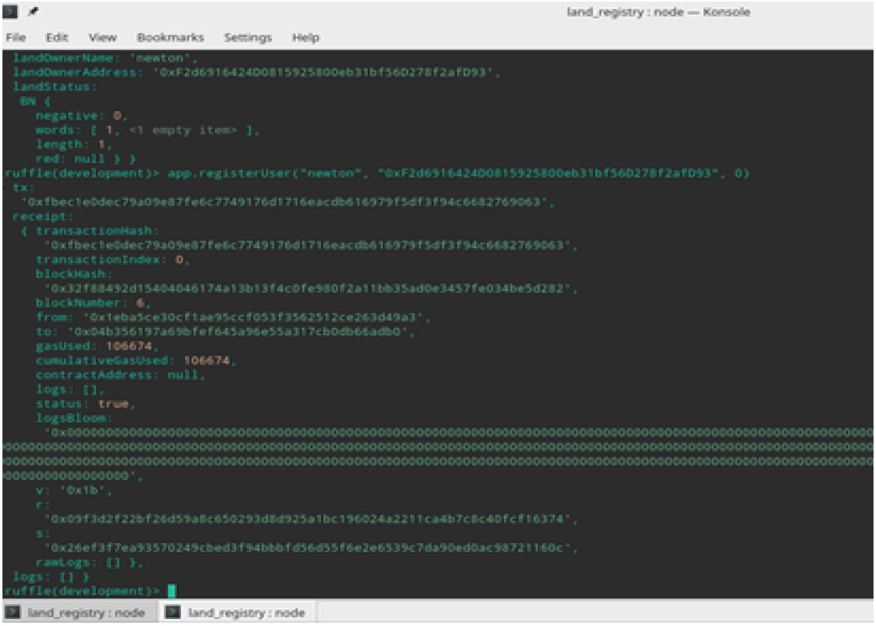
```
   > Saving artifacts
   -------------------------------------
   > Total cost:          0.00569816 ETH


2_deploy_contracts.js
=====================

   Deploying 'LandRegistry'
   ------------------------
   > transaction hash:    0x9231b852a20981625f00115b367d81190bc7f0e67c1eb89c6a122
eb53e9369e8
   > Blocks: 0            Seconds: 0
   > contract address:    0x04B356197a69bFef645a96E55a317cB0db66AdB0
   > account:             0x1EbA5cE30CF1AE95cCf053f3562512CE263d49a3
   > balance:             99.96959284
   > gas used:            1193416
   > gas price:           20 gwei
   > value sent:          0 ETH
   > total cost:          0.02386832 ETH


   > Saving migration to chain.
   > Saving artifacts
   -------------------------------------
   > Total cost:          0.02386832 ETH


Summary
=======
> Total deployments:   2
> Final cost:          0.02956648 ETH

[gadsater@gad-pc land_registry]$ 
```

(b)

```
 landOwnerName: 'newton',
 landOwnerAddress: '0xF2d6916424D0815925800eb31bf56D278f2afD93',
 landStatus:
  BN {
    negative: 0,
    words: [ 1, <1 empty item> ],
    length: 1,
    red: null } }
truffle(development)> app.registerUser("newton", "0xF2d6916424D0815925800eb31bf56D278f2afD93", 0)
 tx:
  '0xfbec1e0dec79a09e87fe6c7749176d1716eacdb616979f5df3f94c6682769063',
 receipt:
  { transactionHash:
     '0xfbec1e0dec79a09e87fe6c7749176d1716eacdb616979f5df3f94c6682769063',
    transactionIndex: 0,
    blockHash:
     '0x32f88492d15404046174a13b13f4c0fe980f2a11bb35ad0e3457fe034be5d282',
    blockNumber: 6,
    from: '0x1eba5ce30cf1ae95ccf053f3562512ce263d49a3',
    to: '0x04b356197a69bfef645a96e55a317cb0db66adb0',
    gasUsed: 106674,
    cumulativeGasUsed: 106674,
    contractAddress: null,
    logs: [],
    status: true,
    logsBloom:
     '0x00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000',
    v: '0x1b',
    r:
     '0x09f3d2f22bf26d59a8c650293d8d925a1bc196024a2211ca4b7c8c40fcf16374',
    s:
     '0x26ef3f7ea93570249cbed3f94bbbfd56d55f6e2e6539c7da90ed0ac98721160c',
    rawLogs: [] },
 logs: [] }
truffle(development)> 
```

Fig. 6. User Registration

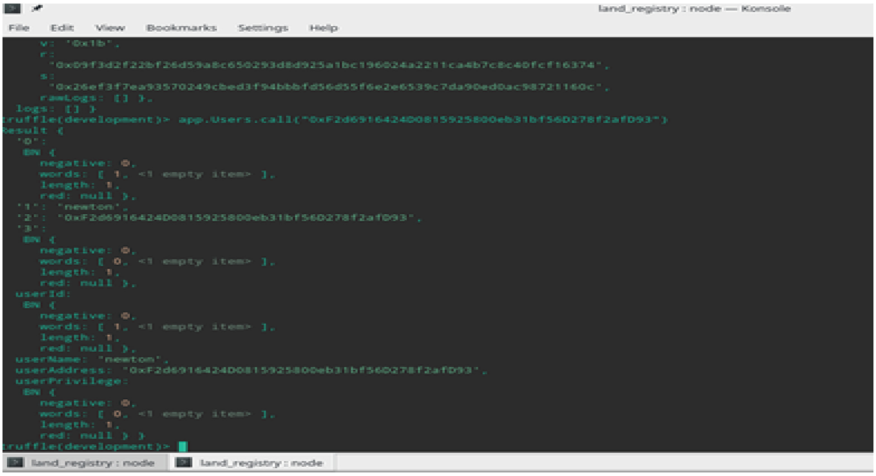
(c)



land_registry : node — Konsole
File Edit View Bookmarks Settings Help
land_registry : node
land_registry : node

```
  landOwnerName: 'newton',
  landOwnerAddress: '0xF2d6916424D0815925800eb31bf56D278f2afD93',
  landStatus:
   BN {
     negative: 0,
     words: [ 1, <1 empty item> ],
     length: 1,
     red: null } }
ruffle(development)> app.registerUser("newton", "0xF2d6916424D0815925800eb31bf56D278f2afD93", 0)
{ tx:
   '0xfbec1e0dec79a09e87fe6c7749176d1716eacdb616979f5df3f94c6682769063',
  receipt:
   { transactionHash:
      '0xfbec1e0dec79a09e87fe6c7749176d1716eacdb616979f5df3f94c6682769063',
     transactionIndex: 0,
     blockHash:
      '0x32f88492d15404046174a13b13f4c0fe980f2a11bb35ad0e3457fe034be5d282',
     blockNumber: 6,
     from: '0x1eba5ce30cf1ae95ccf053f3562512ce263d49a3',
     to: '0x04b356197a69bfef645a96e55a317cb0db66adb0',
     gasUsed: 106674,
     cumulativeGasUsed: 106674,
     contractAddress: null,
     logs: [],
     status: true,
     logsBloom:
      '0x0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000',
     v: '0x1b',
     r:
      '0x09f3d2f22bf26d59a8c650293d8d925a1bc196024a2211ca4b7c8c40fcf16374',
     s:
      '0x26ef3f7ea93570249cbed3f94bbbfd56d55f6e2e6539c7da90ed0ac98721160c',
     rawLogs: [] },
  logs: [] }
ruffle(development)>
```

(d)



land_registry : node — Konsole
File Edit View Bookmarks Settings Help
land_registry : node
land_registry : node

```
     v: '0x1b',
     r:
      '0x09f3d2f22bf26d59a8c650293d8d925a1bc196024a2211ca4b7c8c40fcf16374',
     s:
      '0x26ef3f7ea93570249cbed3f94bbbfd56d55f6e2e6539c7da90ed0ac98721160c',
     rawLogs: [] },
  logs: [] }
truffle(development)> app.Users.call("0xF2d6916424D0815925800eb31bf56D278f2afD93")
Result {
  '0':
   BN {
     negative: 0,
     words: [ 1, <1 empty item> ],
     length: 1,
     red: null },
  '1': 'newton',
  '2': '0xF2d6916424D0815925800eb31bf56D278f2afD93',
  '3':
   BN {
     negative: 0,
     words: [ 0, <1 empty item> ],
     length: 1,
     red: null },
  userId:
   BN {
     negative: 0,
     words: [ 1, <1 empty item> ],
     length: 1,
     red: null },
  userName: 'newton',
  userAddress: '0xF2d6916424D0815925800eb31bf56D278f2afD93',
  userPrivilege:
   BN {
     negative: 0,
     words: [ 0, <1 empty item> ],
     length: 1,
     red: null } }
truffle(development)>
```

Fig. 6. (*continued*)

3.1 User Registration

The following methods are used to perform user registration:

registerCustomer(): This method is used to register a user with Customer privilege. It receives Name of the user as a parameter, finds out the address of the user from the person who calls the smart contract and creates the user details structure associated with the customer. This structure is then stored in a User's mapping.

registerUser(): This method can be invoked only by a user with Admin privilege. An admin can specify both the user name and user address in order to register the user. Similar to the previous method, a structure is created and the details get stored in the Users mapping (Fig. 6).

3.2 RegisterLand()

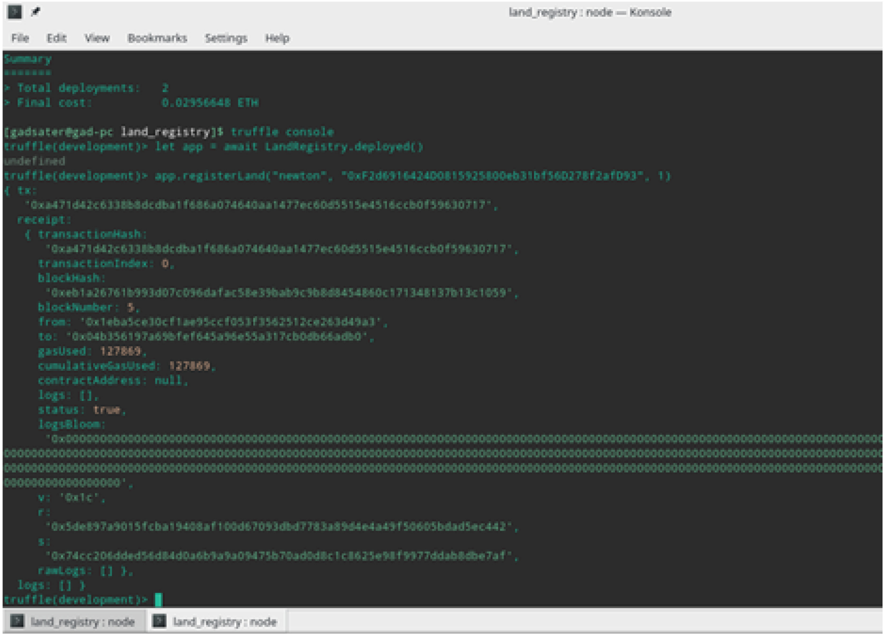
This method can be invoked only by a user with privilege level of Admin or Registrar. It is assumed that the presence of land is verified before the land is registered in this system. A unique tag and land owner address is fed in as parameters. This is used to create the land detail structure and store it in Lands mapping.

3.3 TransactLand()

This method is invoked to transfer land ownership from one user to another. The land tag associated with the land is fed in as parameter. The user who calls this method is the buyer and the land owner is the seller. If the land is in 'on sale' state and if the buyer has enough balance to purchase the land, the transaction is initiated. The ownership of the land changes and the balance deducted from buyer is added to the wallet of seller.

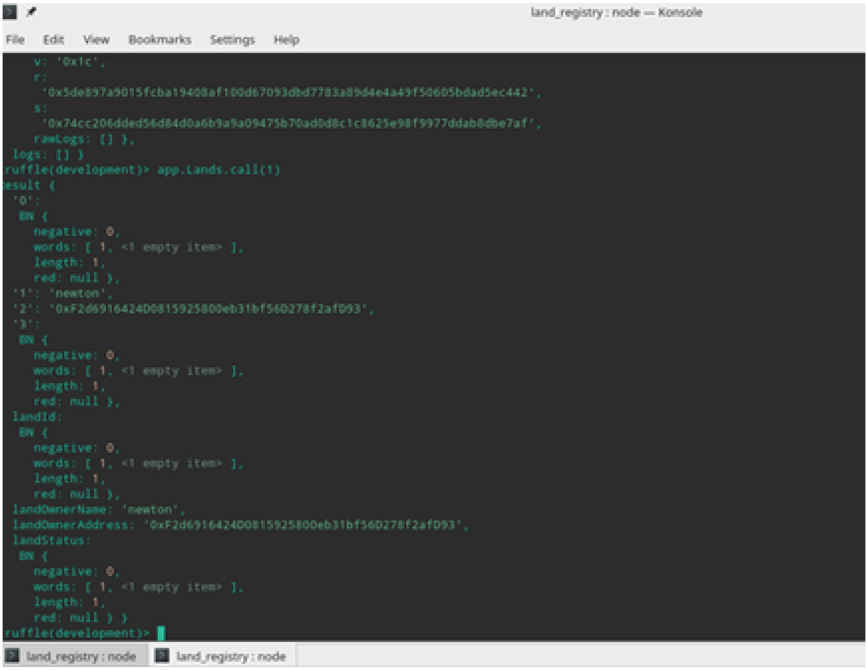
4 Results Obtained so Far

We have utilized the Ethereum platform to implement our Land Registry system. Ethereum, as a platform, provides us with the means to write smart contracts. The core logic of our transaction model is handled by smart contracts. It includes the methods to store details of lands, details of users and transact land from one customer to another. This can then be deployed in the Ethereum network. When the functions of smart contracts are called, the piece of code gets executed in EVM (Ethereum Virtual Machine). The changes of state are permanently recorded in the blockchain as transactions. Once these transactions are recorded, they cannot be reversed or removed. This ensures the immutability of data and hence the transaction records can be used to trace the history of the land. All the details of Users and Lands are stored as state variables (a mutable variable whose value becomes static for every smart contract state) in the blockchain (Figs. 7 and 8).


land_registry : node — Konsole
File Edit View Bookmarks Settings Help
land_registry : node
land_registry : node

```
Summary
=======
> Total deployments:   2
> Final cost:          0.02956648 ETH

[gadsater@gad-pc land_registry]$ truffle console
truffle(development)> let app = await LandRegistry.deployed()
undefined
truffle(development)> app.registerLand("newton", "0xF2d6916424D0815925800eb31bf56D278f2afD93", 1)
{ tx:
   '0xa471d42c6338b8dcdba1f686a074640aa1477ec60d5515e4516ccb0f59630717',
  receipt:
   { transactionHash:
      '0xa471d42c6338b8dcdba1f686a074640aa1477ec60d5515e4516ccb0f59630717',
     transactionIndex: 0,
     blockHash:
      '0xeb1a26761b993d07c096dafac58e39bab9c9b8d8454860c171348137b13c1059',
     blockNumber: 5,
     from: '0x1eba5ce30cf1ae95ccf053f3562512ce263d49a3',
     to: '0x04b356197a69bfef645a96e55a317cb0db66adb0',
     gasUsed: 127869,
     cumulativeGasUsed: 127869,
     contractAddress: null,
     logs: [],
     status: true,
     logsBloom:
      '0x0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000',
     v: '0x1c',
     r:
      '0x5de897a9015fcba19408af100d67093dbd7783a89d4e4a49f50605bdad5ec442',
     s:
      '0x74cc206dded56d84d0a6b9a9a09475b70ad0d8c1c8625e98f9977ddab8dbe7af',
     rawLogs: [] },
  logs: [] }
truffle(development)>
```

Fig. 7. The seller and buyer protocol on ethereum platform


land_registry : node — Konsole
File Edit View Bookmarks Settings Help
land_registry : node
land_registry : node

```
     v: '0x1c',
     r:
      '0x5de897a9015fcba19408af100d67093dbd7783a89d4e4a49f50605bdad5ec442',
     s:
      '0x74cc206dded56d84d0a6b9a9a09475b70ad0d8c1c8625e98f9977ddab8dbe7af',
     rawLogs: [] },
  logs: [] }
ruffle(development)> app.Lands.call(1)
esult {
  '0':
   BN {
     negative: 0,
     words: [ 1, <1 empty item> ],
     length: 1,
     red: null },
  '1': 'newton',
  '2': '0xF2d6916424D0815925800eb31bf56D278f2afD93',
  '3':
   BN {
     negative: 0,
     words: [ 1, <1 empty item> ],
     length: 1,
     red: null },
  landId:
   BN {
     negative: 0,
     words: [ 1, <1 empty item> ],
     length: 1,
     red: null },
  landOwnerName: 'newton',
  landOwnerAddress: '0xF2d6916424D0815925800eb31bf56D278f2afD93',
  landStatus:
   BN {
     negative: 0,
     words: [ 1, <1 empty item> ],
     length: 1,
     red: null } }
ruffle(development)>
```

Fig. 8. Smart contract between network participants in a decentralized network

5 Conclusion and Future Direction

In this work, we have implemented land registry application using blockchain technology and Smart contract. We have shown the results on ethereum platform where transaction of land has happened between sellers and buyers. For future work, We would like to explore more on dealing with multiple ownership of lands, store documents in a distributed file system like IPFS (InterPlanetary File System) and present an user interface instead of a command line application to make these transactions.

References

1. Nakamoto, S.: Bitcoin: a peer-to-peer electronic cash system (2008)
2. Tschorsch, F., Scheuermann, B.: Bitcoin and beyond: a technical survey on decentralized digital currencies. IEEE Commun. Surv. Tutorials **18**(3), 2084–2123 (2016)
3. Bahga, A., Madisetti, V.K.: Blockchain platform for industrial internet of things. J. Softw. Eng. Appl. **9**(10), 533 (2016). https://solidity.readthedocs.io/en/v0.4.24/
4. Sato, T., Himura, Y.: Smart-contract based system operations for permissioned blockchain. In: 2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS). IEEE (2018)
5. Mohanta, B.K., Panda, S.S., Jena, D.: An overview of smart contract and use cases in Blockchain technology. In: 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (2018)
6. Shah, M., Kumar, P.: Tamper proof birth certificate using blockchain technology. Int. J. Recent Technol. Eng. (IJRTE). **7**(5S3) (2019). E-ISSN 2277-3878
7. Buchmann, N., Rathgeb, C., Baier, H., Busch, C., Marian, M.: Enhancing breeder document long-term security using blockchain technology. In: IEEE 41st Annual Computer Software and Applications Conference (2017)
8. Chen, Y., Li, H., Li, K., Zhang, J.: An improved P2P file system scheme based on IPFS and blockchain. In: IEEE International Conference on Big Data (BIGDATA) (2017)
9. Lo, S.K., Xu, X., Chiam, Y.K., Lu, Q.: Evaluating suitability of applying blockchain. In: International Conference on Engineering of Complex Computer Systems (2017)
10. Nomura Research Institute: Survey on blockchain technologies and related services (2016)
11. Thompson, S.: The preservation of digital signatures on the blockchain - Thompson - See Also. Univ. Br. Columbia iSchool Student J. vol. 3, no. Spring (2017)
12. Schmidt, P.: Certificates, reputation, and the blockchain. MIT Media Lab, Cambridge (2015)

Author Index