



Automatic Identification Fingerprint Based on Machine Learning Method

Long The Nguyen¹ · Huong Thu Nguyen^{2,3} ·
Alexander Diomidovich Afanasiev¹ · Tao Van Nguyen³

Received: 23 September 2019 / Revised: 4 March 2020 / Accepted: 28 September 2020

Published online: 3 January 2021

© Operations Research Society of China, Periodicals Agency of Shanghai University, Science Press, and Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

The fingerprint identification technology has been developed and applied effectively to security systems in financial transactions, personal information security, national security, and other fields. In this paper, we proposed the development of a fingerprint identification system based on image processing methods that clarify fingerprint contours, using machine learning methods to increase processing speed and increase the accuracy of the fingerprint identification process. The identification system consists of the following main steps: improving image quality and image segmentation to identify the fingerprint area, extracting features, and matching the database. The accuracy of the system reached 97.75% on the mixed high- and low-quality fingerprint database.

Keywords Fingerprint identification · Feature extraction · Image segmentation · Wavelet transform · Neural network algorithm · Machine learning

This research was supported by the National Natural Science Foundation of China (Nos. 00001 and 00010) and Chongqing Municipal Education Commission (No. KJ120616).

✉ Long The Nguyen
thelongit88@gmail.com

Huong Thu Nguyen
thuhuonyb@gmail.com

Alexander Diomidovich Afanasiev
aad@istu.edu

Tao Van Nguyen
nvtao@ictu.edu.vn

¹ Laboratory of Artificial Intelligence and Machine Learning, Institute of Information Technology and Data Science, Irkutsk National Research Technical University, Irkutsk, Russia

² Baikal School of BRICS, Irkutsk National Research Technical University, Irkutsk, Russia

³ University of Information and Communication Technology, Thai Nguyen University, Thai Nguyen, Viet Nam

Mathematics Subject Classification 68T45 · 68Q32 · 68W40

1 Introduction

Fingerprints are one of the most reliable biometric features for human identification. Most of the fingerprint databases depend on the status of the finger's surface (humidity, dust, temperature, etc.), which can affect the accuracy of the identification system. Fingerprint identification is a big challenge for the research teams when faced with a low-quality fingerprint database. This makes the identification process more difficult, which affects the accuracy of the system. Our paper proposes a way to identify the problem that causes the low-quality images based on a combination of morphology method with image segmentation using move–expansion algorithm and then matching template by the artificial neural network. This paper is organized as follows. In the next section, the fingerprint identification system is described. Namely, in Sect. 3.1 preprocessing image to enhancement fingerprint image quality is discussed. In Sect. 3.2, fingerprint feature analysis and extraction are given. In Sect. 3.3, the matching with templates in system database is given. In Sect. 4, the result of experimentation is given. Finally, Sect. 5 concludes the article with the final remarks and future work plans.

2 Related Work

Fingerprint identification systems have the same structure (Fig. 1) which is pre-processing input data, feature extraction, and matching databases. The differences between the systems are the technical choice, the features selected to be used and the matching method.

Fingerprint confirmation [1, 2] requires comparing a registered or enrolled biometric sample besides a newly captured biometric sample. During enrollment, a sample of the fingerprint attribute is captured, which is processed through a computer and stored for later comparison and matching. The fingerprint identification technology involves the identification and verification of individuals by analyzing fingerprint characteristics and has been widely used in various aspects of life for different purposes. Despite the numerous advantages of the fingerprint identification system and its impact on various work sectors across the globe, most users of identification technology still face the challenge of defining the right and accurate technology system that will be cost-effective for solving particular problems in a specific environment [3].

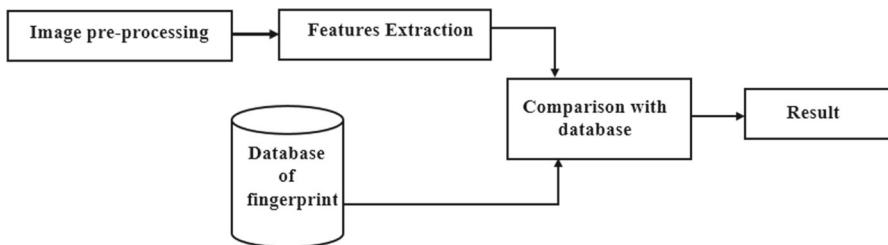


Fig. 1 Basic structure of the fingerprint identification system

The authors in [4] conducted a study entitled “A Wireless Fingerprint Attendance System” which combined fingerprint authentication with the process of attendance management, thus forming a novel “automatic attendance management technique”, a design method of a wireless fingerprint attendance system based on ZigBee. It achieves attendance management by fingerprint identification. It combines ZigBee wireless technology and attendance management. It has achieved low-cost, low-power and high-performance fingerprint information acquisition, transmission and recognition function. In [5], the authors focused their attention on the technical details and performance comparison of various available fingerprint sensors and explored the future direction and system development that state using similar techniques for chance or latent fingerprint enrollment. The process of scanning a fingerprint image consists of a sensor and human interaction, and the quality of image acquired is highly dependent on the fingerprint sensor used. Miguel et al. [6] introduced a short review on combination methodology in hand biometric technology based on hand traits. It shows a clear tendency on increasing the number of traits to be combined and the step to contactless devices. Besides, it is shown that the four biometrics improve the performance of the biometric scheme when they are combined. The purpose of authors in [7] is to develop an embedded system, which is used for security applications and includes a microcontroller-based prototype of an attendance system using a fingerprint sensor and face recognition module is implemented. They conducted a study entitled “Attendance Monitoring Systems of Students Based on Biometric and GPS Tracking System”. This system is a fingerprint recognition system based on minutiae-based fingerprint algorithms used in various techniques.

In [8], the authors investigated possibilities of incorporating artificial neural networks [9] into the fingerprint recognition process, implemented and documented for fingerprint identification based on neural networks [10] whose impact on feature extraction accuracy and overall recognition rate was evaluated. The result of this research is a fully functional software system for fingerprint recognition that consists of a fingerprint sensing module using the high-resolution sensor, an image enhancement module responsible for image quality restoration. An approach was proposed to generate very realistic fingerprints, which can be useful for performance evaluation, training and testing fingerprint-based systems [11]. It is a method for generating synthetic fingerprints on the basis of some mathematical models that describe the main features of real fingerprints. The synthetic images are randomly generated according to a few given parameters. SFinGe captures the variability which characterizes the acquisition of fingerprints through online sensors and uses a sequence of steps to derive a series of “impressions” of the same “artificial finger”.

3 System of Fingerprint Identification

3.1 Preprocessing Image to Enhancement Fingerprint Image Quality

Fingerprint image quality is really important. It is one of the factors that determine the success of an image fingerprint identification system. But unfortunately, almost all fingerprint photos are not of perfect quality. They can be deformed, damaged by noise,

such as the skin properties of each object (temperature, the humidity of the skin) are collected fingerprints, position on the finger (fingerprint angle, area fingerprints) and pressure (too strong or too light) when rolling out to sample. So to overcome the above errors, we performed image quality enhancement before segmenting and extracting features. The steps to improve image quality are as follows:

- Step 1 Filter noise image by Wiener [12]. The goal of the Wiener filter is to compute a statistical estimate of an unknown signal using a related signal as an input and filtering that known signal to produce the estimate as an output. The Wiener filter can be used to filter out the noise from the corrupted signal to provide an estimate of the underlying signal of interest. Therefore, it is very suitable for cases where the fingerprints fail during collection.
- Step 2 Histogram equalization and normalization to reduce the effect of non-uniform intensities and improve image quality by stretching its histogram
- Step 3 ROI extraction with morphology method. The morphology method has two operations: opening and closing. Opening is removing objects that are not large enough, separating them from large objects. Closing is connecting discrete samples, filling small holes. Here, we use both properties to localize and select the desired area.
- Step 4 Image segmentation using move-expansion algorithm [13, 14] image segmentation method is used to segment the images [15, 16]. We applied an efficient graph-based method (alg. 1) to find the optimal image segmentation. A graph $G = (V, E)$ is constructed by nodes V and edges E . Pixels of the image as nodes, two additional nodes S (source) for the object-fingerprints and T (not fingerprints) for the background are considered. To build a weighted graph from an image $d = (4, 8)$, neighborhood system is considered. For each pair of nodes, t -link is the edges that are connected.

Algorithm 1 Process of image segmentation based on the graph-based method

Data: The graph $G = (V, E)$ with a capacity c , the source S , sink T .

Result: Maximum flow from source S to sink T

1. for $i:=1$ to all edges (u,v) do
 - $f(u,v) \leftarrow 0$
 - end
2. if $cf(u,v) > 0$ to all edges $(u,v) \in p$
 - $cf(p) = \min\{cf(u,v) | (u,v) \in p\}$;
 - else todo 1
 - end
3. for $i:=1$ to all edge (u,v) do
 - $f(u,v) \leftarrow f(u,v) + cf(p)$
 - end

Let us briefly outline the approach we used. These segments are called “sites” and have a predefined orientation of 0° , 45° , 90° , or 135° . The separation between both cases is done with parameter $k \in \{0; 1\}$. Our goal will be to segment an image by constructing a graph such that the minimal cut of this graph will cut all the edges that connect the pixels of different objects with each other.

- An edge directed graph represents the size and dimension of the target segmenting image has to be created.
- Object and background seeds have to be distinguished properly with the formation of nodes—source s and sink t . Based on the object or the background labels, all seeds have to be connected with either source or sinks node.
- Each link of the formed graph is to be associated with suitable edge costs.
- Any minimum s - t cut method is to be used, which indicates the graph nodes represent image boundaries for object and background.

At the end of the image preprocessing process, we obtained a fingerprint image with a better quality than that of the original image, with a clear fingerprint line as shown in Fig. 2.

3.2 Fingerprint Feature Analysis and Extraction

3.2.1 Fingerprint Feature Analysis

Human fingerprints are completely different. Each person has a unique fingerprint, and the probability of coinciding with others is almost zero. The overall shape of a human fingerprint does not change over time. It has these irreversible characteristics: singularity (as shown in Fig. 3a) and minutiae (as shown in Fig. 3b). These special points will exist on the lines of the fingerprint so that we can extract and use them as a basis to accurately identify each object. The database of fingerprint used to research has a resolution of each image of 300 dpi and a size of 256×256 .

3.2.2 Feature Extraction of Singularity

On the fingerprint, there are areas with unusual structures compared with that of other areas. They often have a parallel structure called “singularity”. There are two types of singularity: core and delta. To extract singularity characteristics, we proceed as follows:



Fig. 2 Example of fingerprint image preprocessing

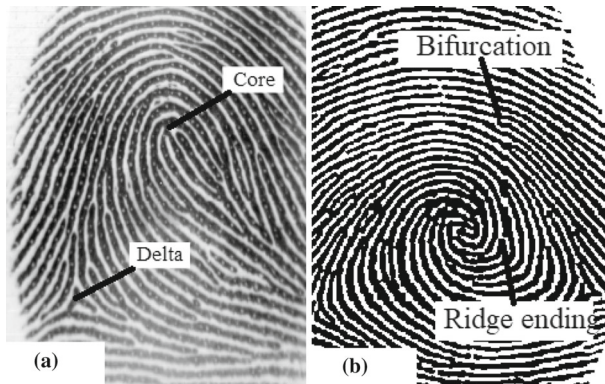


Fig. 3 Example of fingerprint features

Step 1 Divide the image into window cells of size $N * N$.

Step 2 At each pixel, gradients calculated in two directions x and y are G_x and G_y based on the formula:

$$\varphi = \frac{1}{2} \tan^{-1} \left[\frac{\sum_{i=1}^W \sum_{j=1}^W 2G_x(i, j)G_y(i, j)}{\sum_{i=1}^W \sum_{j=1}^W (G_x^2(i, j) - G_y^2(i, j))} \right].$$

Step 3 Identify singularity points using the Poincare index [17]. Poincare index at the pixel with coordinates (i, j) is the sum of the deviations of the direction of adjacent points, calculated as follows:

$$PC(i, j) = \sum_{k=0}^{N_p-1} \Delta(k),$$

$$\Delta(k) = \begin{cases} d(k), & |d(k)| < \frac{\pi}{2}, \\ d(k) + \pi, & d(k) \leq -\frac{\pi}{2}, \\ d(k) - \pi, & \end{cases}$$

$$d(k) = \varphi(x_{k+1}, y_{k+1}) - \varphi(x_k, y_k).$$

Based on the Poincare index, we can identify singularity points as follows:

$$PC(i, j) = \begin{cases} 0^\circ, & (i, j) \text{ is not singularity;} \\ 360^\circ, & (i, j) \text{ is whorl;} \\ 180^\circ, & (i, j) \text{ is loop;} \\ -180^\circ, & (i, j) \text{ is delta.} \end{cases}$$

3.2.3 Feature Extraction of Minutiae

There are points on the fingerprint line that is ridge ending or bifurcation. They are called minutiae features. To extract singularity characteristics, we proceed as follows:

- Step 1 Decompose a given textured image with 2D wavelet transform into 4 sub-images. The image is divided into four sub-bands after wavelet transforms: horizontal, vertical, diagonal, and low-resolution sub-images.
- Step 2 Extract wavelet transform-based extraction with coefficients and create vector:

$$\vec{I}_i, i = 1, \dots, N.$$

- Step 3 Calculate the mean absolute value feature for each decomposed image as follows:

$$I_N = \frac{1}{N} \sum_{i=1}^N I_i.$$

- Step 4 Calculate the standard deviation from each image as follows:

$$\vec{u}_i = \sum_{k=1}^N V_{ik} \theta_k, i = 1, \dots, N.$$

$\theta_k = \vec{I}_i - \vec{I}_N$: subtraction of the average image from each image. V_{ik} : vector of each matrix $W^t W$, $W = \{\vec{\theta}_1, \dots, \vec{\theta}_N\}$.

3.3 Matching with Templates in the System Database

Common algorithms used for fingerprint matching can be categorized into three groups: correlation-based matching, minutiae-based matching and ridge feature-based matching [18]. Correlation-based algorithms use cross-correlation as a measure of similarity between two image pairs. Ridge feature-based techniques refer to methods utilizing features belonging to ridge patterns [17]. These algorithms use global and/or local cross-correlation techniques. Both correlation and minutiae-based algorithms can be considered subsets of ridge feature-based algorithms since they are based on information from ridges.

The method of fingerprint identification by the artificial neural network [19, 20] will train the neural network based on data samples and the location of the fingerprint features. The trained neural network will then be used to match new fingerprints that need to be identified. Fig. 4 shows the flowchart.

Neural networks comprised layers of computational units called neurons, with connections in different layers (Fig. 5). These networks transform data until they can classify it as an output. Each neuron multiplies an initial value by some weight, sums results with other values coming into the same neuron, adjusts the resulting number by the neuron's bias, and then normalizes the output with an activation function.

A key feature of neural networks is an iterative learning process in which records (rows) are presented to the network one at a time, and the weights associated with the

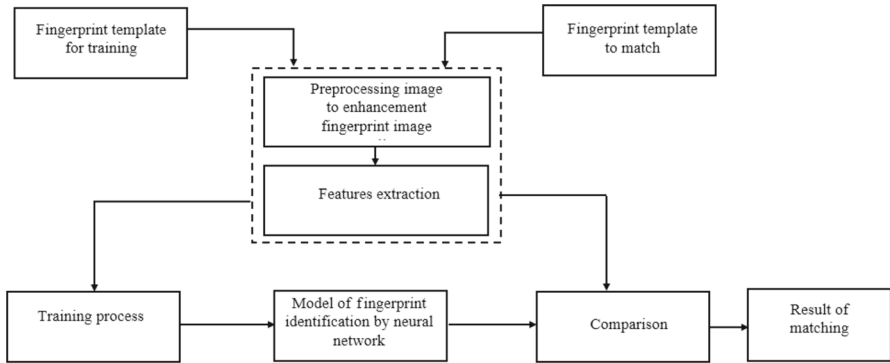


Fig. 4 Flowchart of the fingerprint identification system by neural network algorithm

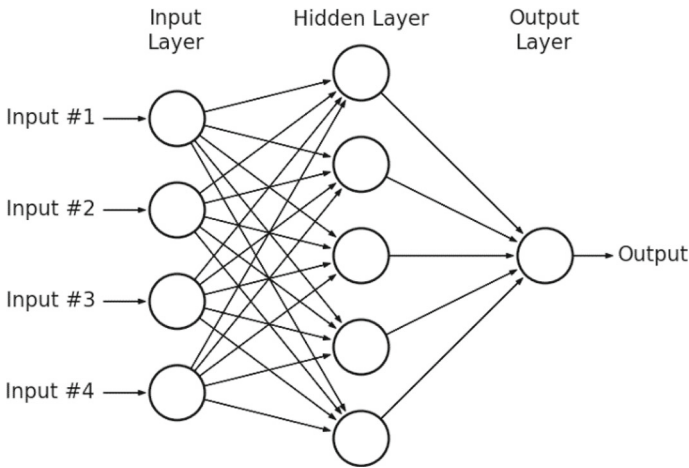


Fig. 5 Structure of neural network algorithm to training and testing templates

input values are adjusted each time. The advantages of neural networks include their high tolerance to noisy data, as well as their ability to classify patterns on which they have not been trained.

4 Result of Experimentation

The experiments were performed on two databases of fingerprint verification competition (FVC) group, and our database of 500 fingerprint images taken from 100 different fingers, each finger was taken 5 times. The resolution of each image is 300 dpi, and the size is 256 × 256. We divided these images to create training data of 100 and 400 images for matching. All experiments were conducted using a Win10 PC with Intel Core i3 CPU @ 3.00 GHz and 2.00 GB RAM.

The network input is the location of the feature points. To determine the location of a point, we must have a relatively fixed origin. Here, we choose the core point as the

origin, because the core point always exists and is relatively fixed in the fingerprint. In this paper, we selected the perceptron straight through network with one output, each network corresponding to a sample. Thus, when we need to match a pattern, we have to compare that pattern across all networks in the database. Because comparing the sample over simple networks is faster than the time it takes to train a large network, the method is feasible. Based on such network selection, we will activate the output as a linear function and will be trained to zero for each sample. The network processes the records in the “training set” one at a time, using the weights and functions in the hidden layers and then compares the resulting outputs against the desired outputs. Errors are then propagated back through the system, which causes the system to adjust the weights for application to the next record.

This process occurs repeatedly as the weights are tweaked. During the training of a network, the same set of data is processed many times as the connection weights are continually refined.

To evaluate the performance of a fingerprint identification system, we need to analyze two types of errors: false acceptance rate (FAR—the percent of unauthorized users incorrectly matched to a valid user’s biometric) and false rejection rate (FRR—the percent of incorrectly rejected valid users) [21]. FAR occurs when we accept a user whom we should actually have rejected. This type of issue is also referred to as a false positive. FRR is the problem of rejecting a legitimate user when we should have accepted it. This type of issue is commonly known outside the world of biometrics as a false negative. A false acceptance occurs when an unauthorized subject is accepted as valid. If an organization’s biometric control is producing a lot of false rejections, the overall control might have to lower the accuracy of the system by lessening the amount of data it collects when authenticating subjects. When the data points are lowered, the organization risks an increase in the false acceptance rate. The organization risks an unauthorized user gaining access. This type of error is also called a type II error.

In addition, the receiver operating characteristic (ROC) curve [22, 23] (Fig. 6) was used to analyze the effectiveness and accuracy of the methods neural network for fingerprint identification.

The result of Table 1 shows that the time to collect a sample of two-layer perceptron needs more than one-layer perceptron, but its matching results are much better and more stable than a one-layer perceptron.

To demonstrate the robust accessibility of the proposed method-neural network algorithm. We have experimented FVC group database with two algorithms: random forest [24] and support vector machine (SVM) [25]. Criteria are used to compare algorithms (shown in Table 2): training time and correct rate and mean square error (MSE) [26].

Table 1 Compare collection time and matching time between one-layer perceptron and two-layer perceptron

	Time to collect template/s	Time to match/s
One-layer perceptron	0 105	0 256
Two-layer perceptron	2 885	0 387

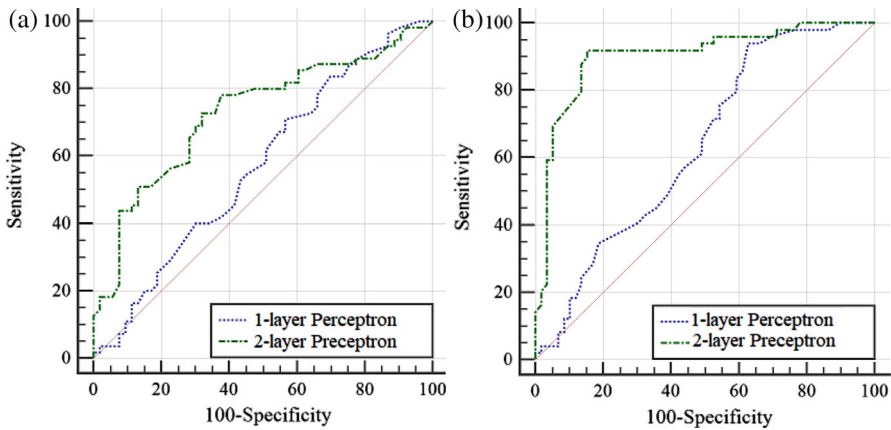


Fig. 6 Result of analysis ROC for fingerprint identification. **a** Result of the FVC group database, **b** result of our database

Table 2 Results of identification fingerprint of FVC database by algorithms: neural network one-layer perceptron, neural network one-layer perceptron, random forest, and support vector machine

	One-layer perceptron	Two-layer perceptron	Random forest	SVM
Training time/s	11 h 45 m	12 h 10 m	10 h 50 m	8 h 30 m
Correct rate/%	97.50	97.75	95.80	95.02
MSE	0.286	0.205	0.511	0.662

- Parameters of the random forest algorithm:
 - The number of trees to train model = 1 200.
 - The function to measure the quality of a split—Gini Impurity.
 - Bootstrap samples = True.
- Parameters of the support vector machine algorithm:
 - Penalty parameter to measure error term = 1.0.
 - Kernel: basis functions.
 - Shrinking heuristic = True.

5 Conclusions

5.1 Limitations

The fingerprint authentication generates a very high percent accuracy. However, it has been observed that the unrecognized fingerprint data input was mainly from the finger not being pressed hard on the quality of the scanner. So, we need to develop an algorithm for enhancement image to improve the quality of the image and reduce the time of the running process.

5.2 Advantages and Disadvantages of the Proposed Methods

We proposed an automated fingerprint identification system for image enhancement, feature extraction and matching. The method of neural networks represents the core of machine learning for comparison template. The accuracy of the system reached 97.75% on the mixed high- and low-quality fingerprint database. However, its accuracy can be further improved by extending the training database. In addition, the network's architecture can be easily modified to support more complex training process.

5.3 Future Research Works

Our interest is also aimed at deep learning [27–29] by means of neural networks that have proved their exceptional performance in a multitude of applications. One of the future prospects is to extend the proposed system to design a convolution neural network for extracting feature fingerprints (orientation field, density map, etc.), which is one of the top open research issues in modern fingerprint recognition, particularly in forensic and law enforcement applications.

Acknowledgements The authors are grateful to ICO'2018 Pattaya, Thailand, and ICO'2019 Koh Samui, Thailand, for giving them opportunity to present the research results in this special issue. This research was performed at Irkutsk National Research Technical University Irkutsk, Russia.

References

- [1] Bartumek J. S.: Fingerprint Image Enhancement, Segmentation and Minutiae Detection, Doctoral Dissertation, Blekinge Institute of Technology (2016)
- [2] Kulkarni, S.: Fingerprint feature extraction and classification by learning the characteristics of fingerprint patterns. *Neural Netw. World* **3**, 219–226 (2011)
- [3] Oloyede, M.O., Adedoyin, A.O., Adewole, K.S.: Fingerprint biometric authentication for enhancing staff attendance system. *Int. J. Appl. Inf. Syst.* **5**, 19–24 (2013)
- [4] Patil, P., Khachane, A., Purohit, V.: A wireless fingerprint attendance system. *Int. J. Secur. Privacy Trust Manag.* **5**(4), 11–17 (2016)
- [5] Barbadekar, A., Jadhav, S.D., Patil, S.P.: Performance analysis of fingerprint sensors. Vishwakarma Institute of Technology, Pune (2010). <https://doi.org/10.1109/ICMEE.2010.5558571>
- [6] Miguel A.F, Aythami M., Carlos M.T, Jesus B. A.: Combining hand biometric traits for personal identification, In 43rd Annual 2009 International Carnahan Conference on Security Technology, IEEE (2009). <https://doi.org/10.1109/CCST.2009.5335547>
- [7] Chandramohan, J., Nagarajan, R., Kumar, M., Dineshkumar, T., Kannan, G.: Attendance monitoring systems of students based on biometric and GPS tracking syste. *Int. J. Adv. Eng. Manage. Sci.* **3**, 3 (2017)
- [8] Pavol, M., Hambahl, A.: Fingerprint recognition system using artificial neural network as feature extractor: design and performance evaluation. *Math. Publ.* **67**, 117–134 (2016)
- [9] Uhrig R.E.: Introduction to artificial neural networks. In: Proceedings of IECON Annual Conference on IEEE Industrial Electronics (2002). <https://doi.org/10.1109/iecon.1995.483329>
- [10] Bartunek J. S., Nilsson M., Nordberg J., Claesson I.: Neural network based minutiae extraction from skeletonized fingerprints (2006). <https://doi.org/10.1109/TENCON.2006.344104>
- [11] Cappli R.: SFinGe: an approach to synthetic fingerprint generation, in: International Workshop on Biometric Technologies, Calgary, Canada, 147–154 (2004)

- [12] Kumar, N., Kalyan Singh, K.: Wiener filter using digital image restoration. *Int. J. Electron. Eng.* **3**(2), 345–348 (2011)
- [13] Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via Graph Cuts. *IEEE PAMI* **23**(11), 1222–1239 (2001)
- [14] Kolmogorov, V., Zabih, R.: Computing visual correspondence with occlusions via graph cuts. In: *International Conference on Computer Vision*, vol. 2, pp. 508–515 (2001)
- [15] Nguyen, T.H., Nguyen, T.L., Dreglea, A.I.: Machine learning algorithms application to road defects classification. *Intell. Decis. Technol.* **59–66**, 12 (2018)
- [16] Nguyen, T.H., Nguyen, T.L., Dreglea, A.I.: Robust approach to detection of bubbles based on images analysis. *Int. J. Artif. Intell.* **16**, 167–177 (2018)
- [17] Maltoni, D., Maio, D., Jain, A.K., Prabhakar, S.: *Singularity and Core Detection*” Extract from “Handbook of Fingerprint Recognition. Springer, New York (2003)
- [18] Jain, A.K., Flynn, P., Ross, A.A.: *Handbook of Biometrics*. Springer, Berlin (2008)
- [19] Maas, A.L., Hannun, A.Y., Ng, A.Y.: Rectifier nonlinearities improve neural network acoustic models. In: *Proceedings of icml*, vol. 30, no. 1 (2013)
- [20] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhut dinov, R.: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**(1), 1929–1958 (2014)
- [21] Vasseur, J.-P., Dunkels, A.: in *Interconnecting Smart Objects with IP* (2010). <https://doi.org/10.1016/C2009-0-20667-2>
- [22] Zweig, M.H., Campbell, G.: Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.* **39**, 561–577 (1993)
- [23] Nguyen, T.H., Nguyen, T.L.: ROC curve analysis for classification of road defects. *BRAIN Broad Res. Artif. Intell. Neurosci.* **10**(2), 65–73 (2019)
- [24] Breiman, L.: Random forests. *Mach. Learn.* **45**, 5–32 (2001). <https://doi.org/10.1023/A:1010933404324>
- [25] Awad, M., Khanna, R.: Support vector machines for classification, efficient learning machines, pp 39–66 (2015)
- [26] Liu, Y.: Mean square error of survey estimates, encyclopedia of quality of life and well-being research, 2014, ISBN: 978-94-007-0752-8
- [27] Thomas, J., PinarKaragoz, J., BazeerAhamed, B., Vasant, P.: Deep learning techniques and optimization strategies in big data analytics. *IGI Global* **13**, 1–355 (2020). <https://doi.org/10.4018/978-1-7998-1192-3>
- [28] Vasant, P., Zelinka, I.: Gerhard-Wilhelm Weber, *Intelligent Computing & Optimization*, 2019, Volume 866, ISBN : 978-3-030-00978-6, <https://www.springer.com/gp/book/9783030009786>
- [29] Vasant, P., Zelinka, I.: Gerhard-Wilhelm Weber, *Intelligent Computing & Optimization*, Proceedings of the 2nd International Conference on Intelligent Computing and Optimization 2019, <https://www.springer.com/gp/book/9783030335847>