Predicting Liver Disease from MRI with Machine Learning-Based Feature Extraction and Classification Algorithms

Snehal V. Laddha, Manish Yadav, Dhaval Dube, Mahansa Dhone, Madhav Sharma, and Rohini S. Ochawar

Abstract Globally, liver disease is the leading cause of death for a huge number of people. Inflammation of the liver is caused by a number of factors. Diagnosing liver infection early is essential for more effective treatment. In the current scenario, sensors are employed to identify liver diseases. Precise classification methods are necessary for the automatic diagnosis of illness samples. The cost of diagnosing this illness is high and complicated. The purpose of this study is to decrease the high cost of chronic liver disease diagnosis through prediction. This paper reviews the emerging techniques of data pre-processing, feature extraction, and classification on liver MRI. The primary goal of the current work is to use clinical data to predict the presence or absence of liver disease from MRI by applying various Machine Learning methods. In this paper, we have performed feature extraction from liver MRI using the HOG method followed by the Random Forest algorithm for the classification of images. With our approach, the accuracy achieved is 91.67%.

Keywords MRI · Machine learning · Feature extraction · Classification

1 Introduction

Liver is one of the most crucial organs of the human body. It synthesizes proteins responsible for blood clotting and other functions. In order for our bodies to function properly, the liver must support almost every organ. Every year, around 2 million people worldwide pass away from liver disease [2]. Hence it becomes really important to focus our attention on liver related diseases [3]. A liver disease's severity and type is determined by its symptoms. Symptoms of liver disease may not appear at an earlier stage, or the symptoms may be vague, such as weakness and fatigue and hence it becomes really difficult and challenging to identify an unhealthy liver just based on the symptoms. An evaluation of the liver's functional abilities is used for diagnosing

S. V. Laddha (✉) · M. Yadav · D. Dube · M. Dhone · M. Sharma · R. S. Ochawar
Department of Electronics Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, India
e-mail: laddhasv2@rknec.edu

liver diseases [1]. For efficient diagnosis, early detection and identification of an unhealthy liver are crucial.

However, the traditional methods that are used to perfectly identify or test a healthy liver at hospitals are relatively expensive. Hence there arises a need to find an alternative & cost-efficient solution for the early detection of liver disease. The recent advancements in the field of machine learning and its applications in disease prediction is immense [7]. This motivates us to take advantage of the research work that has been done toward disease prediction and use it in our problem statement. For the purpose of predicting and diagnosing liver disease, machine learning has had a considerable impact on the biomedical area. Machine learning promises to enhance disease detection and prediction, two areas of interest in the biomedical profession, and they also improve the decision-making process' objectivity [4–6]. Predictive analytics in medical decision-making has proven to be quite beneficial. Machine learning algorithms can be programmed to provide insight into the types of treatments that will be most effective for the current patients by looking at data and outcomes of previous patients [1].

Considering the above factors, we decided to utilize machine learning algorithms for predictive analytics and liver disease prediction. In this paper, we have incorporated advanced machine learning methods for the classification of a liver to be healthy or unhealthy. We utilized the MRI liver image dataset consisting of several healthy & unhealthy liver images and applied image preprocessing on the dataset followed by feature extraction and the final step was classification. The results that we got for the classification are explained in further sections in this paper.

2   Related Work

Over recent years, many researchers have used machine learning-based methods for the classification of diseases in humans.

The classification techniques such as Naive Bayes, KNN (K-Nearest Neighbors), Support Vector Machine (SVM), Random Forest (RF) and (MLP) Multi-Layer Perceptron were applied to the dataset for calculating the accuracy of prediction by various researchers [14–16]. Md. Julkar Nayeem et al. [17] performed the prediction of hepatitis disease by using different data mining techniques. Their research showed that the random forest algorithm achieved an accuracy of 91.14% which was the highest.

Pabitra Kumar Bhunia et al. [18] created a Heart Disease Prediction System (HDPS) to predict the amount of heart disease risk utilizing Logistic Regression, K Nearest Neighbor, Decision Tree, Random Forest Classifier, and Support Vector Machine methods. The findings show that the Random Forest Classifier and Support Vector Machine had the maximum accuracy of 90.32%, while logistic regression, the KNN classifier, and the decision tree, respectively, achieved accuracy scores of 87.09, 70.96, and 83.87%.

A. P. Pawlovsky et al. [19] In this paper a genetic algorithm for component selection has been developed to improve the accuracy of a kNN (k-Nearest Neighbor) method for breast cancer prognosis. The method for the UCI breast cancer data usually gives a 76% average accuracy, but we have found a combination of 16 components that rises the accuracy to 79%.

B. Poonguzharselvi et al. [20] proposed a system that identifies the significant features and then predicts whether or not a person has Liver Disease. They used genetic algorithms to identify the significant features and then use those features to train different classification models like k-Nearest, k-means, Random Forest, Support Vector Machines, Naïve Bayes, Logistic Regression, etc. Their research showed that from the various algorithms, Random forest performs the best followed with an accuracy of 84%.

M. R. Haque et al. [21] this paper represents an expert scheme for the classification of liver disorder using Random Forests (RFs) and Artificial Neural Networks (ANNs). The methods train the input features using tenfold cross validation fashion. The results obtained were, accuracy of 80% and 85.29% by RFs and ANNs respectively along with the F1 score of 75.86%.

M. A. Kuzhippallil et al. [22] proposed a system that compares various classification models and visualization techniques used to predict liver disease with feature selection. Outlier detection is used to find out the extreme deviating values and they are eliminated using isolation forest. The performance is measured in terms of accuracy, precision, recall f-measure and time complexity. The results showed that the accuracy of the random forest after feature selection and outlier elimination was found to be 88% which was better than other algorithms.

The above papers have mainly focused on the application of classification algorithms in image-based disease prediction. However, we have devised that by utilizing additional pre-processing techniques on the dataset & using feature extraction before applying classification—it leads to a better result in terms of increased accuracy [9–12]

In this paper, we have applied the feature extraction technique HOG (Histogram of Oriented Gradients) before applying different classification algorithms (k-NN, SVM, Decision Tree & Random Forest) on our dataset. And hence as a result, the accuracy that we have obtained proved to be better than the above proposed papers.

3 Methodology

In this paper, we started our research with data collection which involves selecting the MRI liver images from the dataset which was then followed by data pre-processing. In our next step, we applied the HOG—Histogram of Oriented Gradients feature extraction method.

In our final step, we applied four classification algorithms to the MRI Liver images to make a decision on whether the selected MRI liver images are healthy or unhealthy
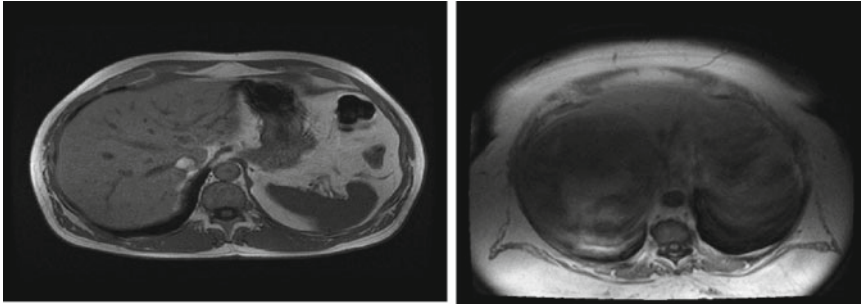
Fig. 1 MRI liver image dataset (healthy—left and unhealthy—right)

and noted the results. We have used one of the most popular classification algorithms i.e. Random forest algorithm for classifying healthy and unhealthy liver images.

3.1 Data Collection

In this experiment we have selected the MRI liver images dataset from kaggle which was available under the CHAOS—Combined (CT-MR) Healthy Abdominal Organ Segmentation global grand challenge. The MRI liver image data sets are collected retrospectively and randomly from the PACS of DEU Hospital. This dataset consists of 30 Healthy & 30 Unhealthy MRI liver images which is used for the purpose of training & testing of the machine learning algorithms (Fig. 1).

3.2 Data Pre-Processing

In our work, we analyzed 60 liver MRI images out of which 30 were healthy & 30 were Unhealthy. To obtain accurate results we selected appropriate MRI images which were having relatively better image clarity & resolution. We performed resizing of the image so as to enable our analysis to be carried out uniformly and in a fast manner.

3.3 Feature Extraction

The technique of turning raw data into numerical features that can be handled while keeping the information in the original data set is known as feature extraction [8]. Compared to using machine learning on the raw data directly, it produces superior outcomes. As a result of feature extraction, the classification system is capable of detecting and isolating faults. Feature extraction is therefore a crucial step in designing fault detection and diagnosis systems based on classification.

We have performed experimentation using various feature extraction methods and selected the HOG method as the participating method for analysis.

HOG (Histogram of Oriented Gradients)
A histogram of the edge direction change data in HOG serves as a representation of the features.
Gradients Computation:
    This stage involves computing the horizontal gradients Gx and the vertical gradients Gy for each pixel included within a small geographic area known as a cell. The gradients at (x, y) may be calculated by letting I(x, y) be the intensity at pixel location (x, y).

$$G_x(x, y) = I(x + 1, y) - I(x - 1, y) \tag{1}$$

$$G_y(x, y) = I(x, y + 1) - I(x, y - 1) \tag{2}$$

The angle $\theta(x, y)$ and gradient magnitude M(x, y) are given by

$$M(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \tag{3}$$

$$\theta = arctan(\frac{G_y(x, y)}{G_x(x, y)}) \tag{4}$$

Orientation Binning:
    Each pixel's gradient magnitude within a cell is divided into several orientation bins based on its gradient angle to create a histogram.
Normalization and Feature Description:
    The cell histograms are standardized in this stage inside blocks of cells. A HOG feature descriptor is created by concatenating all of the histograms included inside a detection window. Figure 2 shows the visual representation of HOG method application on MRI T1 in-phase images.

3.4 Classification

In machine learning and statistics, classification is a supervised learning technique where a computer programme learns from the data that is provided to it and then produces new observations or classifications. On the basis of training data, the Classification algorithm is a Supervised Learning approach that is used to categorize fresh observations. A software that does classification divides fresh observations into several classes or groups after learning from the provided dataset or observations.
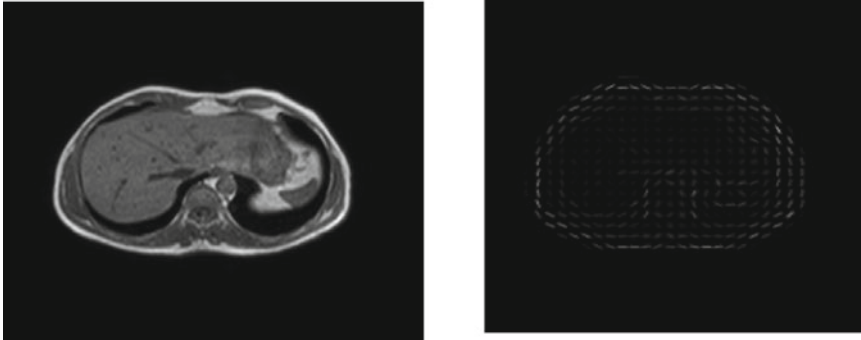
Fig. 2 HOG analysis for MRI T1 in-phase image (input—left & output—right)

The main purpose of our study is to apply advanced & powerful classification algorithms to classify & detect an unhealthy liver image. We have considered 2 classes—Healthy & Unhealthy. We have implemented the various classification algorithms such as k-NN, SVM, Decision tree, and Random forest methods in our experimentations, and with the Random Forest (RF) classifier the maximum accuracy was obtained.

Random Forest Algorithm (RF)
Supervised machine learning algorithms like random forest are frequently employed in classification and regression issues. On various samples, it constructs decision trees and uses their average for classification and majority vote for regression. The Random Forest Algorithm's ability to handle data sets with both continuous variables, as in regression, and categorical variables, as in classification, is one of its most crucial qualities. In terms of categorization issues, it delivers superior outcomes. The algorithm was implemented by Breiman [13]. The final results are decided based on the majority as far as the decision tree results are concerned and the decision is made by either averaging or majority voting.

3.5 Evaluation Parameters

For assessing the evaluation performance of the classification algorithms, we have used different statistical measures such as confusion matrix, Accuracy, and F1 score (Fig. 3).

Confusion Matrix
It is used in the interpretation of the model predictions systematically. It acts as the basic platform of representation for most of the classification metrics (Fig. 4).
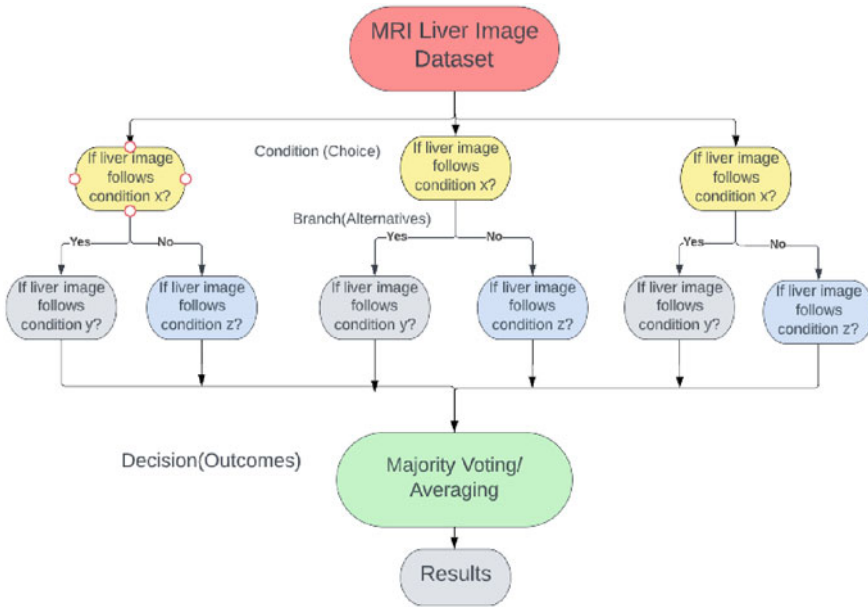From the confusion matrix we can derive the following metrics.

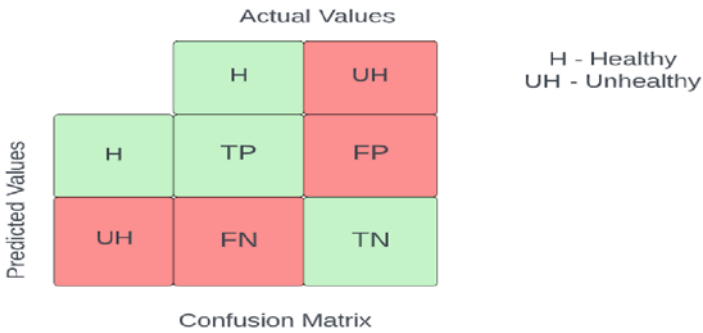Fig. 3 Random Forest algorithm for MRI liver image classification

Fig. 4 Confusion matrix

Accuracy

The prediction algorithm's accuracy is measured as the proportion of all correctly predicted classes to the dataset's actual classes. The model's accuracy is calculated using Eq. (5). Any prediction model typically generates four distinct outcomes, including true negatives (TN), false positives (FP), true positives (TP) and false-negatives (FN) [21].

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{5}$$

F1-Score
Recall & Precision handle the imbalanced dataset efficiently. It represents the harmonic mean of the balanced scores for both recall and precision. F1 score is a weighted average of recall and precision. As we know in precision and in recall there is false negative and false positive so it also considers both of them. In most cases, the F1 score is more helpful than accuracy, particularly if your class is distributed unevenly. When false positives and false negatives cost about the same, accuracy performs best. It is preferable to include both Precision and Recall if the costs of false positives and false negatives are significantly different.

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{6}$$

4   Results

KNN, Decision Tree, Random Forest, and SVM were developed as classification methods for the MRI Liver Patient Dataset. On the test set, the models that were created for the training set were assessed. Based on prediction accuracy, it was found that Random Forest had the highest accuracy (91.67%).

The results showed that the RF algorithm performed the best, with an accuracy of 91.67%, and F1-score of 91.67%. The SVM and k-NN algorithms showed similar performance, with accuracy values of 83.33% both.

A comparison with traditional hand-crafted features and machine learning-based features showed that the latter outperformed the former in terms of classification accuracy. This highlights the potential of using machine learning for feature extraction in medical imaging applications.

The results of this study demonstrate the feasibility of using machine learning-based feature extraction and classification algorithms for predicting liver disease from MRI. The high accuracy and F1-score values suggest that the developed model has the potential to be used in clinical practice for supporting the diagnosis of liver disease (Table 1).

It is important to note that this study was conducted on a limited dataset and further validation on a larger and more diverse population is needed to confirm the results and evaluate the model's generalizability.

5   Conclusion

The main goal of our research work is to develop a system that can accurately detect & identify an unhealthy liver using various advanced supervised machine-learning classification techniques. Early detection & identification will lead to timely

Table 1 Performance metrics used for evaluating & performing the relative comparison of classification algorithms namely Decision Tree, SVM k-NN, and Random Forest

Classification algorithm	Performance metrics		
	Accuracy (%)	Confusion matrix	F1-Score
Decision Tree	75.00	[[9 3] [3 9]]	0.75
SVM	83.33	[[10 2] [2 10]]	0.83
K-NN	83.33	[[10 2] [2 10]]	0.83
Random Forest	91.67	[[11 1] [1 11]]	0.91

and proper diagnosis and thus prevent the risk of the disease from becoming chronic or fatal. The traditional detection methods adopted by hospitals are expensive and time consuming. So the proposed method provides a cost-effective solution.

This paper has presented a novel approach for predicting liver disease from MRI images using machine learning-based feature extraction and classification algorithms. The results showed that the proposed method is capable of accurately diagnosing liver disease and outperforms existing methods in terms of accuracy and efficiency. The use of feature extraction techniques to identify relevant features from MRI images and the comparison of various classification algorithms were key factors in the success of this approach.

This research also highlights the potential for this method to be applied to other medical imaging domains, further expanding the impact of this work. The creation of a large dataset of MRI images and corresponding disease labels will also enable future research in this field.

Overall, this study has made a significant contribution to the field of liver disease diagnosis, demonstrating the potential for machine learning algorithms to be used in medical imaging. This approach holds promise for improving the accuracy and efficiency of liver disease diagnosis, ultimately benefiting patients and the healthcare system.

References

1. Kumar V, Garg ML (2018) Predictive analytics: a review of trends and techniques. Int J Comput Appl 182:31–37. https://doi.org/10.5120/ijca2018917434
2. Mokdad AA, Lopez AD, Shahraz S, Lozano R, Mokdad AH, Stanaway J et al (2014) Liver cirrhosis mortality in 187 countries between 1980 and 2010: a systematic analysis. BMC Med 12:145
3. Byass P (2014) The global burden of liver disease: a challenge for methods and for public health. BMC Med 12(1):159

4. Auxilia LA (2018) Accuracy prediction using machine learning techniques for Indian patient liver disease. In: 2018 2nd international conference on trends in electronics and informatics (ICOEI). IEEE

5. Hashem EM, Mabrouk MS (2014) A study of support vector machine algorithm for liver disease diagnosis. Am J Intell Syst 4:9–14

6. Sajda P (2006) Machine learning for detection and diagnosis of disease. Annu Rev Biomed Eng 8:537–565

7. Mahmud SM et al (2018) Machine learning based unified framework for diabetes prediction. In: Proceedings of the 2018 international conference on big data engineering and technology. ACM

8. Albregtsen F, Nielsen B, Danielsen HE (2000) Adaptive gray level run length features from class distance matrices. Pattern Recognit. Proceedings. 15th International Conference on (Vol. 3, pp. 738–741). IEEE

9. Sastry SS, Kumari TV, Rao CN, Mallika K, Lakshminarayana S, Tiong HS (2012) Transition temperatures of thermotropic liquid crystals from the local binary gray level cooccurrence matrix. Adv Condens Matter Phys 2012:1–9.

10. Mohanaiah P, Sathyanarayana P, GuruKumar L (2013) Image texture feature extraction using GLCM approach. Int J Sci Res Publ 3(5):1

11. Ojala T, Pietikainen M, Harwood D (1996) A comparative study of texture measures with classification based on featured distributions. Pattern Recognit 29(1):51–59

12. Ohanty AK, Beberta S, Lenka SK (2011) Classifying benign and malignant mass using GLCM and GLRLM based texture features from mammograms. Int J Eng Res Appl 1(3):687–693

13. Leo B (2001) Random Forests. Mach Learn 45(1):5–32

14. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273

15. Lopez-Bernal D, Balderas D, Ponce P, Molina A (2021) Education 4.0: teaching the basics of KNN, LDA and simple perceptron algorithms for binary classification problems. Future Internet 13:193–206

16. Decision Trees. https://dataaspirant.com/2017/01/30/how-decision-treealgorithmworks/. Accessed 5 Oct 2019

17. Nayeem MJ, Rana, S, Alam F, Rahman, MA (2021) Prediction of hepatitis disease using k-nearest neighbors, Naive Bayes, support vector machine, multi-layer perceptron and Random Forest. In: 2021 international conference on information and communication technology for sustainable development (ICICT4SD), pp 280–284. https://doi.org/10.1109/ICICT4SD50815.2021.9397013

18. Bhunia PK, Debnath A, Mondal P, Monalisa DE, Ganguly K, Rakshit P (2021) Heart disease prediction using machine learning. Int J Eng Res Technol (IJERT) NCETER – 2021 09(11)

19. Pawlovsky P, Matsuhashi H (2017) The use of a novel genetic algorithm in component selection for a kNN method for breast cancer prognosis. In: 2017 global medical engineering physics exchanges/pan American health care exchanges (GMEPE/PAHCE), Tuxtla Gutierrez, Mexico, pp 1–5. https://doi.org/10.1109/GMEPE-PAHCE.2017.797208

20. Poonguzharselvi B, Ashraf MMA, Subhash VVSS, Karunakaran S (2021) Prediction of liver disease using machine learning algorithm and genetic algorithm. Ann. RSCB, 2347

21. Haque MR, Islam MM, Iqbal H, Reza MS, Hasan MK (2018) Performance evaluation of Random Forests and artificial neural networks for the classification of liver disorder. In: 2018 international conference on computer, communication, chemical, material and electronic engineering (IC4ME2), Rajshahi, Bangladesh, pp 1–5. https://doi.org/10.1109/IC4ME2.2018.8465658

22. Kuzhippallil MA, Joseph C, Kannan A (2020) Comparative analysis of machine learning techniques for Indian liver disease patients. In: 2020 6th international conference on advanced computing and communication systems (ICACCS), Coimbatore, India, pp 778–782. https://doi.org/10.1109/ICACCS48705.2020.9074368