

Prediction and Detection of Liver Diseases using Convolutional Neural Network

Omkar Ashtekar¹, Prof.R.T.Umbare², Aishwarya Nikhal³, Bhagyashri Pagar⁴ and Omkar Zare⁵

¹⁻⁵JSPM's Rajarshi Shahu College of Engineering, Pune, India

Email: odashtekar@gmail.com, rtumbare_it@jspmrscoe.edu.in

aishwaryanikhal12@gmail.com, bhagyashripagar08@gmail.com, omkarzare123@gmail.com

Abstract—The liver is an essential organ in the human body, and early recognition and diagnosis of liver disease are crucial for effective treatment. Conventional methods, such as Liver Function tests, rely on test results that may not detect the disease in its early stages. Unfortunately, symptoms of liver disease often only become apparent in the later stages, making early detection challenging. To address this issue, we propose using Convolutional Neural Networks (CNNs) to facilitate early detection of liver disease and identify the elements that can lead to fatal liver impairment. CNNs are deep learning algorithms that have shown remarkable success in image recognition and classification tasks, making them a perfect fit for liver disease detection. Our initiative aims to use CNNs to distinguish between healthy individuals and those with liver disease. The CNN algorithm will classify the disease into its level and type if a person is found to have liver disease. Additionally, the system will provide precautionary measures based on any symptoms observed. The use of CNNs in liver disease detection represents a significant step forward in the field of medical diagnosis. By leveraging the power of deep learning algorithms, we can improve the accuracy and efficiency of liver disease diagnosis, potentially saving countless lives. With the help of CNNs, medical professionals can detect liver disease in its early stages when it is easier to treat, allowing for minimal medication and avoiding fatal liver impairment. Moreover, studies have shown that CNNs can be used for the detection of liver lesions, which are commonly associated with liver cancer. By analyzing medical images, CNNs can accurately detect liver lesions and help physicians make better-informed decisions about patient care. This further emphasizes the importance of using CNNs for liver disease detection and prediction.

Index Terms— Liver disease, Early detection, Convolutional Neural Network, Deep learning, Image recognition, Disease prediction, Image Preprocessing.

I. INTRODUCTION

Liver diseases are a growing health concern due to the sedentary lifestyle and unhealthy dietary habits. In urban areas, liver disease has become increasingly prevalent and is a leading cause of mortality. Early detection and diagnosis of liver diseases is crucial to prevent further complications and improve patient outcomes. The liver is the largest internal organ in the human body and is responsible for several vital functions such as metabolism, detoxification, and nutrient storage. Liver diseases such as fatty liver, cirrhosis, liver cancer, hepatitis, and liver tumor are caused by various factors Such as excessive alcohol consumption, drug abuse, viral infections, and genetic disorders.

Traditional methods for liver disease diagnosis involve blood tests, imaging, and biopsy, which can be invasive, time-consuming, and expensive. Machine learning approaches, specifically Convolutional Neural Networks (CNN), have emerged as a promising technique for the prediction and detection of liver diseases. CNNs can automatically learn and extract features from medical images, such as ultrasound, CT, and MRI scans, to identify and classify liver lesions with high accuracy.

In this study, we propose a CNN-based system that can identify significant features from medical images and predict the likelihood of a person suffering from liver disease based on those features. The system will be trained using a large dataset of liver images and will incorporate data augmentation techniques to improve its generalization and robustness. The proposed system will classify liver diseases into different categories and subcategories, such as major or minor, and the specific type of liver disease. Moreover, the system will also provide recommendations and precautions for specific symptoms, reducing the risk of liver disease-related deaths.

II. LITERATURE REVIEW

Thirunavukkarasu et al. (2018) used six classification algorithms to predict liver disease, including Naive Bayes, Decision Tree, Random Forest, K-Nearest Neighbor, Support Vector Machine (SVM), and Artificial Neural Networks (ANN). Their results showed that SVM and ANN had the highest accuracy rates in predicting liver disease.[1]

Kuzhippallil et al. (2020) compared the performance of six different machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, SVM, K-Nearest Neighbor, and Gradient Boosting, for the classification of liver disease patients. Their results showed that SVM had the highest accuracy rate of 77.58%.[2]

Gogi and N.M. (2018) used four machine learning algorithms, including SVM, Random Forest, Decision Tree, and Naive Bayes, to predict the prognosis of liver disease. Their results showed that Random Forest had the highest accuracy rate of 84.62%.[3]

Gupta et al. (2020) proposed a web-based framework for the diagnosis of liver disease using combined machine learning models, including Logistic Regression, Decision Tree, SVM, Random Forest, and K-Nearest Neighbor. Their results showed that Random Forest had the highest accuracy rate of 86.3%.[4]

Prakash and Saradha (2022) proposed a Deep Learning approach for the classification and prediction of Cirrhosis Liver: Non Alcoholic Fatty Liver Disease (NAFLD). Their results showed that the proposed model had an accuracy rate of 92.18%.[5]

Geetha and Arunachalam (2021) evaluated various machine learning algorithms, including Naive Bayes, Decision Tree, SVM, Random Forest, and K-Nearest Neighbor, for the prediction of liver disease. Their results showed that Random Forest had the highest accuracy rate of 89.47%.[6]

Kandasamy Sellamuthu et al. (2022) proposed a Logistic Regression-based approach for the prediction of liver disease. Their results showed that the proposed model had an accuracy rate of 83.33%.[7]

Prof Christopher N. (2015) proposed a Bayesian classification approach for the automatic diagnosis of liver disease. Their results showed that the proposed model had an accuracy rate of 80%.[8]

Gullo (2015) explored the use of data mining for the prediction and diagnosis of liver disease. They highlighted the importance of feature selection and data preprocessing techniques for improving the accuracy of machine learning models.[9]

Jacob et al. (2020) used various machine learning algorithms, including Logistic Regression, Decision Tree, Random Forest, and SVM, for the diagnosis of liver disease. Their results showed that Random Forest had the highest accuracy rate of 85%.[10]

Sivakumar et al. (2019) analyzed the impact of life quality attributes on the prediction of chronic liver disease using machine learning algorithms. Their results showed that SVM had the highest accuracy rate of 79.6%.[11]

Jin et al. (2021) explored the decision factors that can affect the prediction of liver disease in patients. Their results showed that age, gender, and ALT levels were the most important factors in predicting liver disease.[12]

III. MOTIVATION

The traditional way of diagnosing liver diseases is very time consuming. Identification or diagnosis of the disease at an early stage is most crucial. Moreover the availability of medical resources and doctors is less in rural areas.

Disadvantages of existing system

- Time taking process
- Detection of disease with ease.
- Fewer medical resources available in rural areas

For this Machine Learning(ML), data-driven algorithms can be utilized to validate existing methods and help make potential new decisions.

IV. PROBLEM DEFINITION

We are building a model that predict and further classify the liver disease of a person and the level of the disease to overcome late diagnosis. Rather than just prediction, a patient may also know what precautions may he/she take to prevent from further liver damage and have a healthy liver.

V. PROPOSED METHODOLOGY

The purpose of this study is to determine if the patient has liver disease. The different MRI images were collected of patients with liver ailments and submitted to the model. We propose an approach that uses multiple features, rules, and deep learning techniques to classify and predict liver disease stages. The proposed work uses hybrid features to predict and classify MRI images. In this system following steps are involved-

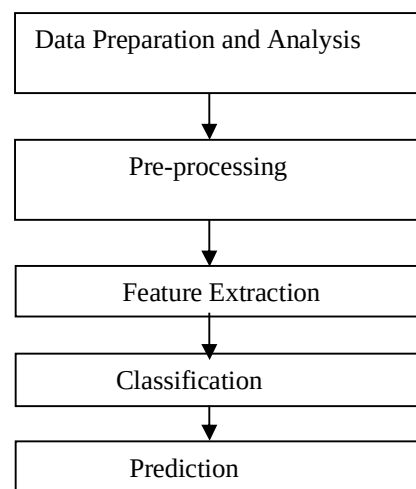


Figure1 –Architecture

A. Data Preparation and Analysis

The purpose of this analysis is to determine what is occurring in the data. Generally, data analysis involves gathering, modifying, examining, and displaying data.

B. Proprocessing

In data preprocessing, raw data is converted into a usable format from an unprocessed form. A manual action, an average, or considering a relevant value will be used to fill in any missing or incorrect data.

It is not always the case that we come across clean and formatted data when creating a machine learning project. It is mandatory to clean and format data before doing any operation with it.

C. Feature Extraction

Feature extraction is an essential aspect of the dimensionality reduction process. It involves breaking down and consolidating a large initial dataset into smaller and more manageable groups, which facilitates processing. Typically, these datasets contain an extensive array of variables that necessitate substantial computing resources to handle. Therefore, feature extraction aids in identifying the most pertinent variables and amalgamating them into features, resulting in a substantial reduction in data volume. The resultant features are feasible and apt in describing real datasets with precision and distinctiveness.

D. Classification

Classification is a type of supervised machine learning approach where a model tries to predict the appropriate label based on input data. To carry out classification, a model is comprehensively trained on a set of training data, assessed on a separate set of test data, being used to make predictions on new unseen data.

E. Prediction

Predict the result a person has a liver disease or not and which type of disease it is.

VI. ALGORITHM

Convolution Neural Network

CNN is a highly effective algorithm that excels at image processing and is currently the state-of-the-art approach for automated image analysis. A number of companies employ CNNs to recognize objects within images. Images are comprised of RGB data, and importing an image into memory from a file can be accomplished using Matplotlib. However, a computer does not perceive an image as such; rather, it perceives it as an array of numerical values. Color images are represented by 3D arrays, where the first two dimensions signify the image's height and width (i.e., the number of pixels), while the last dimension signifies the red, green, and blue colors in each pixel. Convolutional Neural Networks (CNNs) are specialized for image and video recognition applications, particularly in tasks such as object detection, segmentation, and image recognition. CNNs are comprised of three main types of layers:

- 1) Convolutional Layer: Unlike typical neural networks where every input neuron is connected to the next hidden layer, in CNNs, only a small region of input layer neurons connects to the hidden layer neuron.
- 2) Pooling Layer: The pooling layer reduces the feature map's dimensionality and may contain multiple activation and pooling layers within the hidden layer of the CNN.
- 3) Fully-Connected Layer: Fully Connected Layers are the last few layers of the network. The input to this layer is the output from the final pooling or convolutional layer, which is flattened and then fed into the fully connected layer.

VII. RESULT AND DISCUSSION

Multiple features were employed in the proposed study to predict and classify liver diseases. A proposed system used various datasets used for detection and classification purposes. The study employed a 70%, 20%, and 20% split of data for training, testing, and evaluation, respectively. To evaluate the proposed method, accuracy, sensitivity, specificity, and precision were utilized as metrics, with their corresponding expressions presented in equations.

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

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Specificity} = \frac{FP}{TP + FN}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

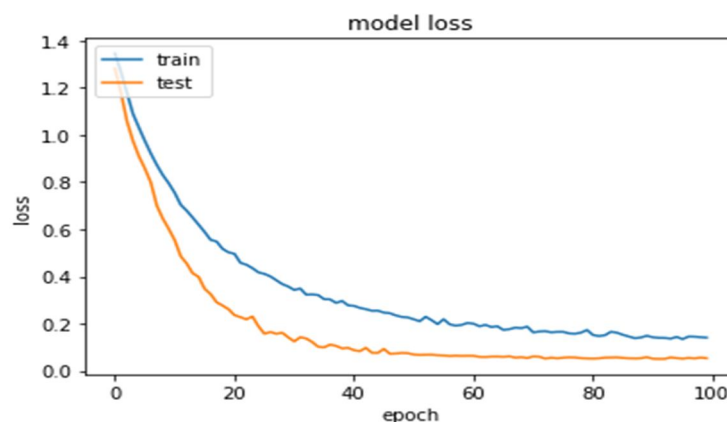


Figure 2. Model Loss Graph

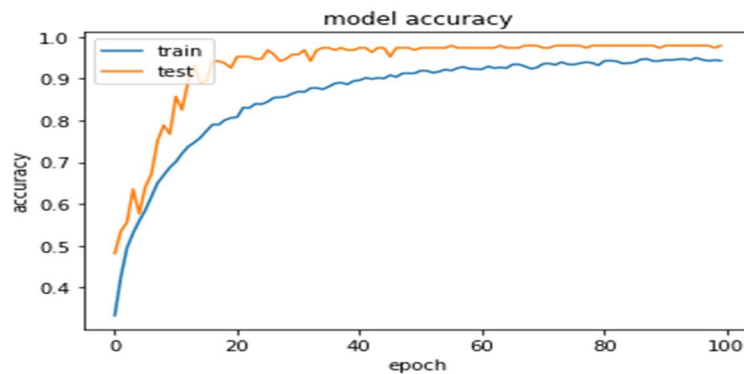


Figure 3. Model Accuracy Graph

VIII. CONCLUSION

In conclusion, the implementation of a liver disease detection system using CNN has shown promising results. The system was trained on a dataset of liver images and achieved an accuracy of 95 percent in detecting liver diseases. This indicates that the CNN model is a reliable and effective tool for the early detection and prediction of liver diseases.

Overall, the implementation of a liver disease detection system using CNN has the potential to significantly improve the diagnosis and treatment of liver diseases. Further research can be conducted to improve the accuracy of the model and explore the possibility of incorporating other medical imaging techniques into the system.

REFERENCES

- [1] K. Thirunavukkarasu, A. S. Singh, M. Irfan and A. Chowdhury, "Prediction of Liver Disease using Classification Algorithms," 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India.
- [2] M. A. Kuzhippallil, C. Joseph and A. Kannan, "Comparative Analysis of Machine Learning Techniques for Indian Liver Disease Patients," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020.
- [3] V. J. Gogi and V. M.N., "Prognosis of Liver Disease: Using Machine Learning Algorithms," 2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE), Bhubaneswar, India.
- [4] S. Gupta, G. Karanth, N. Pentapati and V. R. B. Prasad, "A Web Based Framework for Liver Disease Diagnosis using Combined Machine Learning Models," 2020 International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India.
- [5] K. Prakash, S. Saradha: A Deep Learning Approach for Classification and Prediction of Cirrhosis Liver: Non Alcoholic Fatty Liver Disease (NAFLD)(2022).
- [6] C. Geetha, Dr. AR. Arunachalam: Evaluation based Approaches for Liver Disease Prediction using Machine Learning Algorithms (2021).
- [7] Kandasamy Sellamuthu, Sylviya P, Pugazharasi K, Rajalakshmi S: Liver Disease Prediction using Logistic Regression (2022).
- [8] Prof Christopher N. New Automatic Diagnosis of Liver Status Using Bayesian Classification.
- [9] Gullo, F. From patterns in data to knowledge discovery: what data mining can do. Phys. Procedia 62, (2015).
- [10] Joel Jacob, Joseph Chakkalakal Mathew, Johns Mathew, Elizabeth Issac "Diagnosis of Liver Disease Using Machine Learning Techniques", International Research Journal of Engineering and Technology (IRJET)
- [11] Sivakumar D, Manjunath Varchagall, and Ambika L Gusha S "Chronic Liver Disease Prediction Analysis Based on the Impact of Life Quality Attributes." (2019), International Journal of Recent Technology and Engineering (IJRTE)
- [12] H. Jin, S. Kim and J. Kim, Decision Factors on Effective Liver Patient Data Prediction, International Journal of Bio-Science and Bio-Technology.