IJIREEICE
ISSN (O) 2321-2004, ISSN (P) 2321-5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.021 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 11, Issue 6, June 2023
DOI: 10.17148/IJIREEICE.2023.11609

Diagnosis of Liver Disease Using Machine Learning

Harish Gadhe[1], Nishigandha Kolapkar[2], Rushikesh Kakad[3], Ms. Nrupal Sankpal[4]

Student, Electronics and Telecommunications, P.E.S Modern College of Engineering, Pune, India[1,2,3]

Assistant Professor, Electronics and Telecommunications, P.E.S Modern College of Engineering, Pune, India[4]

Abstract: Liver diseases are becoming more prevalent and can cause significant morbidity and mortality. The identification of liver maladies is a pivotal medical procedure that can have a momentous impact on the final results of patients. The traditional method of diagnosis involves the use of medical imaging and biopsy, but these techniques can be intrusive, time-consuming, and expensive. The recent development of machine learning has opened up the possibility of a non-invasive and more efficient approach to the diagnosis of liver diseases. The use of convolutional neural network (CNN) models for the identification of liver maladies from computed tomography (CT) scan images is a promising method that has received a significant amount of attention from the medical community. These models can precisely detect and categorize different types of liver maladies, such as fatty liver disease and hepatitis, with an exceptional degree of accuracy. The CNN model is instructed using an extensive dataset of CT scan images of the liver, and the network is created to recognize the patterns and features that are peculiar to various liver diseases. Once trained, the model can categorize new CT scans into different groups based on their visual features, providing clinicians with a potent tool for precise and efficient diagnosis. The use of machine learning-based approaches for liver disease diagnosis has numerous advantages, including reduced invasiveness, improved accuracy, and lower costs. It has the potential to transform the way liver diseases are diagnosed and managed, resulting in improved patient outcomes and enhanced quality of care. Nevertheless, further research is necessary to substantiate these approaches and ensure their safety and effectiveness in clinical practice.

Keywords: Liver Diseases, Machine Learning, Convolutional Neural Network, Diagnosis, CT scan

I. INTRODUCTION

The liver, which is the most voluminous internal organ in the human anatomy, performs a variety of essential metabolic functions within the body. It transforms ingested nutrients into substances that are useful to the body and stores them for later use. Additionally, it neutralizes harmful substances and converts them into benign ones. Bile production, protein synthesis, glucose storage and release, haemoglobin processing, blood purification, immune factor production, and bilirubin clearance are among the liver's many critical functions. Given its crucial role, maintaining liver health is crucial for overall well-being. Unfortunately, many individuals ignore this important organ, resulting in a significant number of people worldwide experiencing various degrees of liver ailments due to unhealthy lifestyle habits.

Liver diseases are a significant global health issue, with millions of people being impacted annually. The precise and prompt identification of liver diseases is crucial for effective medical intervention and management. Traditional diagnostic approaches, including medical imaging and biopsy, are invasive, time-consuming, and costly. They also have potential dangers for patients. Recent developments in machine learning have led to the emergence of a novel and more efficient approach to diagnosing liver diseases. Convolutional neural network (CNN) models have gained significant attention from the medical community as a promising method for diagnosing liver diseases from CT scan images. These models can accurately classify different liver diseases, including fatty liver disease and hepatitis, with a high degree of precision.

A CNN model is trained using a vast dataset of CT scan images of the liver, enabling the network to learn the distinguishing features of various liver diseases. Once trained, the model can classify new CT scans into distinct categories based on their unique visual features, providing clinicians with a powerful and efficient tool for accurate diagnosis. Using machine learning-based approaches for liver disease diagnosis provides several advantages, including reduced invasiveness, improved accuracy, and lower costs. This development has the potential to revolutionize the way in which liver diseases are diagnosed and managed, leading to better health outcomes and improved quality of care. However, further research is necessary to validate the effectiveness and safety of these approaches in clinical settings.

II. LITERATURE REVIEW

Liver diseases are a significant global health issue, and early detection is crucial for successful treatment. Computed tomography (CT) scans are extensively utilized in the diagnosis of liver diseases due to their high specificity and sensitivity. In recent times, convolutional neural networks (CNN) have shown great potential in the automatic detection of liver diseases from CT scans. The purpose of this literature review is to provide an overview of the current state of the art in the use of CNNs for the diagnosis of liver diseases from CT scans. In the Paper "Prediction and Diagnosis of Liver Disease in Human Using Machine Learning" by Adekola Olubukola Daniel, Ekanem Edikan Uwem, Omidiran Daniel Tolulope, Owoade Samuel Jesupelumi (2020), the research endeavor, a cutting-edge machine learning framework emerges, unveiling its potential in prognosticating and discerning liver ailments through the analysis of clinical and demographic attributes. The ingenious system harnesses the prowess of decision tree and support vector machine algorithms, showcasing its supremacy when juxtaposed with conventional diagnostic methodologies. The inherent merits of this novel approach encompass heightened precision, diminished diagnostic fallacies, and the advent of economically viable and non-intrusive diagnostic techniques. Nonetheless, it is imperative to acknowledge certain limitations stemming from the restricted magnitude of the employed dataset, potentially impeding its representativeness, as well as the absence of imaging data, which could impact its comprehensive accuracy.[1]

In the article "A Comparative Study of Diagnosing Liver Disorder Disease Using Classification Algorithm" by A. Saranya and G. Seenuvasan (2017), it delves into the efficacy of employing machine learning models for the purpose of diagnosing liver disease, leveraging clinical and laboratory data as vital inputs. The authors of this study utilize classification algorithms to construct predictive models with the ability to discriminate between various liver disease types. Noteworthy advantages of this approach encompass precise diagnostic capabilities, early detection potential, the utilization of non-invasive techniques, and enhanced diagnostic efficiency. However, it is crucial to acknowledge certain limitations that arise, including the availability of data, the interpretability of intricate models, the risk of overfitting, and the reliance on feature selection, all of which necessitate careful consideration to ensure the reliability and robustness of the obtained predictions. [2]

The efficacy of employing machine learning models to diagnose liver disease using clinical and laboratory data is explored in the article "Diagnosis of Liver Disease using Machine Learning Models" by A. Sivasangari, Baddigam Jaya Krishna Reddy, Annamareddy Kiran, P. Ajitha. The authors employ classification algorithms to create predictive models that possess the capability to differentiate between various liver disease types. This methodology offers notable benefits, including precise diagnoses, early identification, non-intrusive procedures, and enhanced diagnostic efficiency. Nevertheless, it is imperative to acknowledge certain limitations that warrant consideration, such as the availability of data, the interpretability of intricate models, the risk of overfitting, and the reliance on feature selection. These factors must be carefully weighed to ensure dependable and resilient predictions. [3]

III. METHODOLOGY

A. Data Collection and Preprocessing:

During this section, we will delineate the methodology that was implemented for gathering data and pre-processing CT scan images related to liver diseases. The accumulation of an ample amount of annotated CT scan images is a crucial element that plays a vital role in training a deep learning model such as CNN.

- Data Collection: The process of data collection involved obtaining CT scan images of liver diseases from publicly available datasets, such as the Open Access Series of Imaging Studies (OASIS) and Cancer Imaging Archive (TCIA), which were sourced from various medical centers and hospitals. This ensured that the dataset exhibited diversity and heterogeneity, and was subsequently categorized into different types of liver diseases, such as fatty liver disease and hepatitis.

- Preprocessing: To ensure that the CNN model was accurately trained, the collected CT scan images underwent a series of preprocessing steps, these eliminates any undesired artifacts and enhance the contrast of the images. The following preprocessing methods were applied to the CT scan images:

Noise Removal: Due to multiple factors like patient motion and system noise, the CT scan images are prone to noise. To counteract this, we employed a median filter with a kernel size of 3x3 to remove any noise present in the images.
Contrast Enhancement: The contrast of the CT scan images is improved by utilizing histogram equalization. This amplified the visibility of the liver lesions and other anomalies present in the images.

DOI: 10.17148/IJIREEICE.2023.11609

Image Resizing: The CT scan images resized to a fixed size of 256x256 pixels. This ensured that all the images were of uniform size, which is a prerequisite for training a CNN model.

Data Augmentation: To augment the size of the dataset and to minimize overfitting, data augmentation methods such as rotation, flipping, and zooming on the CT scan images were employed.

B. Architecture of the CNN Model:

The design of the Convolutional Neural Network (CNN) model for detecting liver diseases from CT scans entails numerous layers. The primary layer is a convolutional layer that applies a group of filters to the input image to extract distinct features. Then, the result of this layer undergoes a rectified linear activation function (ReLU) to introduce non-linearity into the model. The succeeding layers are pooling layers that down-sample the output of the convolutional layers to minimize the dimensions of the data, thereby decreasing the computational cost of the model and avoiding overfitting. The last layer of the CNN model is a fully connected layer, which takes the output of the preceding layers and maps it to a distinct set of class labels. The result of this layer then proceeds through a SoftMax activation function, which converts the output into a set of probabilities for each class label. Various techniques such as dropout regularization, data augmentation, and batch normalization can be applied to optimize the performance of the CNN model. These techniques help to avoid overfitting and enhance the generalization capability of the model.

C. Training and Validation:

After the stage of gathering and processing the data, the subsequent phase involves training and validating the CNN model. The implementation of the CNN model in this research used Keras, a deep learning library, with TensorFlow as the backend. To carry out the training and validation, the subsequent steps were taken: The dataset was partitioned into training and validation subsets, with a random 80% allocated for the former and the remaining 20% for the latter.
The model was subjected to data augmentation techniques, including flipping, zooming, and rotation, to improve generalization and avoid overfitting. Training of the model was done on the training set with the use of binary cross-entropy, stochastic gradient descent as the optimizer, and accuracy as the performance metric. Validation of the model was done on the validation set to assess its performance. The evaluation metrics used were accuracy, precision, recall, F1 score, and the area under the ROC curve. Hyperparameters such as batch size, number of epochs, and learning rate were adjusted to optimize the model's performance. The training and validation process were repeated multiple times, each with varying hyperparameters, to determine the model's optimal performance. The final CNN model was then used to predict CT scan images for liver disease detection.

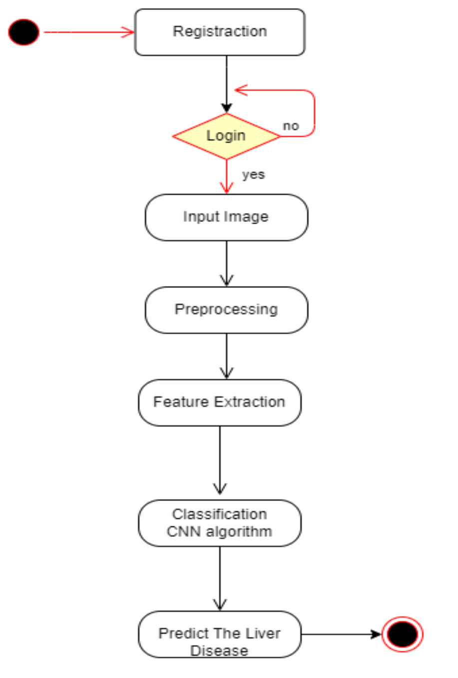

Figure 1: Flow Chart

DOI: 10.17148/IJIREEICE.2023.11609

Here is a description of the flowchart:

1. Registration: This component is used for user registration. It may include collecting user details such as name, email, and password, and storing them securely in a database for future login.

2. Login: This component is used for user authentication. Once registered, the user can log in using their credentials to access the system.

3. Input Image: This component allows the user to input an image of the liver, which is then used for diagnosis.

4. Preprocessing: This component is used to preprocess the input image, which may include resizing, normalization, filtering, or any other necessary image processing techniques to prepare the image for feature extraction.

5. Feature Extraction: This component extracts relevant features from the preprocessed image using techniques such as convolutional neural networks (CNN), which are commonly used for image recognition and analysis.

6. Classification: This component uses the extracted features to classify the input image as either "Liver Disease" or "No Liver Disease". The classification algorithm used here is a CNN algorithm.

7. Result: This component displays the final result of the classification process, indicating whether the input image shows signs of liver disease or not.

IV. RESULTS

The provided visuals showcase the outcomes of a diagnostic system for liver diseases.

This advanced system encompasses a range of features and pertinent information, empowering users with precise predictions and effective functionality. Below, you will find a comprehensive depiction of the achieved results, along with accompanying screenshots illustrating the windows integrated into the system.

Figure 2: Home Page

The home page serves as the main entry point to the system, offering users convenient options for login, registration, and exiting the application. By selecting the appropriate button, users can access the desired features seamlessly.

To utilize the system, users are required to register by providing their essential details.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified ⁑ Impact Factor 8.021 ⁑ Peer-reviewed / Refereed journal ⁑ Vol. 11, Issue 6, June 2023
DOI: 10.17148/IJIREEICE.2023.11609

\

Figure 3: Registration Page

The registration page ensures secure access and maintains user-specific data, enabling personalized experiences within the system.

Figure 4: Login Page

Once registered, users can log in to the system using their credentials. The login page verifies the user's identity and grants access to the system's core functionalities.

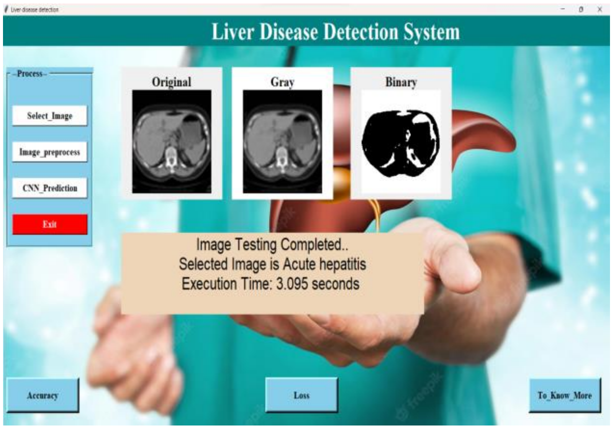

Figure 5: System and Detection Page

The system and detection page serve as the central hub, providing users with a comprehensive set of buttons and options for interaction. It comprises the following sections:

IJIREEICE
ISSN (O) 2321-2004, ISSN (P) 2321-5526
International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified ⁒ Impact Factor 8.021 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 11, Issue 6, June 2023
DOI: 10.17148/IJIREEICE.2023.11609

1. Select Image:
This section allows users to upload CT scan images of the liver for diagnosis. The system processes the uploaded images, extracting relevant features crucial for accurate predictions.
2. Image Processing:
Users can visualize the pre-processing steps applied to the CT scan images to enhance their quality and extract vital details. This section showcases the improvements made to the uploaded images.
3. Prediction:
Leveraging the power of the CNN algorithm, the system predicts the presence of liver diseases based on the processed CT scan images. The predicted disease and relevant information are displayed on this page.
4. Accuracy:
The accuracy page provides users with graphical representations, including accuracy graphs, illustrating the performance of the trained CNN algorithm. It presents the training and testing accuracy, enabling users to assess the reliability of the predictions.

5. Loss:
This section displays the loss function during the training process, offering insights into the model's optimization and convergence. Users can monitor the progress and stability of the CNN algorithm.
6. Information:
Users can access concise and informative descriptions about Fatty Liver, Acute Hepatitis, and Viral Hepatitis. This section provides a brief overview of these liver diseases, assisting users in understanding and interpretation.

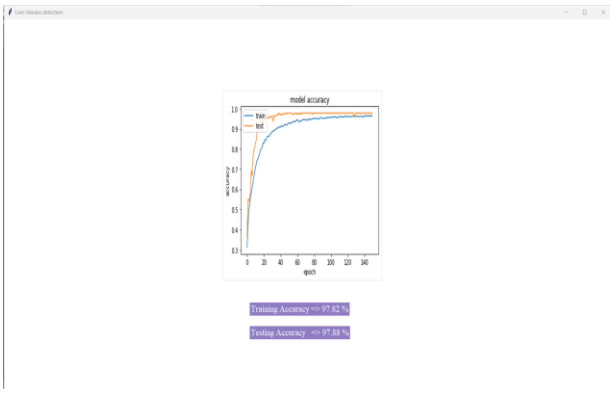

Figure 6: Accuracy Page

The accuracy page presents users with graphical representations, such as accuracy graphs, showcasing the performance of the CNN algorithm. Users can analyse the training and testing accuracy trends to evaluate the model's effectiveness.

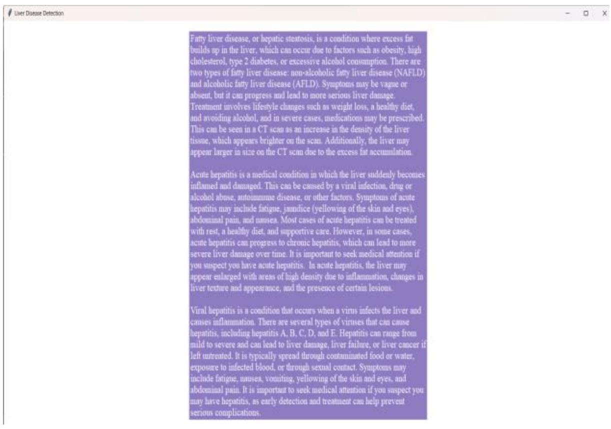

Figure 7: Information Page

DOI: 10.17148/IJIREEICE.2023.11609

The information page provides users with concise yet informative descriptions of Fatty Liver, Acute Hepatitis, and Viral Hepatitis diseases of the liver. This section enhances users' knowledge and awareness of these conditions.

V. CONCLUSION

The project has successfully developed a user-friendly diagnostic system for liver diseases, utilizing machine learning's Convolutional Neural Network (CNN) algorithm with CT scan images. The system's functionalities, which include image processing, disease prediction, and accuracy assessment, contribute to an efficient and precise diagnostic tool.

This project highlights the promising potential of machine learning algorithms in the field of medical diagnostics, specifically in the domain of liver disease diagnosis. By implementing advanced image processing techniques, the system enhances the quality of CT scan images, while the accuracy page allows users to evaluate the reliability of the predictions made. The outcomes of this project hold significant implications for the healthcare industry, providing a valuable tool for improving diagnostic processes and facilitating enhanced patient care.

REFERENCES

[1] Adekola Olubukola Daniel, Ekanem Edikan Uwem, Omidiran Daniel Tolulope, Owoade Samuel Jesupelumi, “Prediction and Diagnosis of Liver Disease in Human Using Machine Learning”, IJCTT Journal, Volume-68 Issue-8, 2020

[2] Saranya and G. Seenuvasan, “A Comparative Study of Diagnosing Liver Disorder Disease Using Classification Algorithm”, IJCSMC, Volume-6, pg.49 – 54, 2017

[3] A. Sivasangari, B. J. Krishna Reddy, A. Kiran and P. Ajitha, "Diagnosis of Liver Disease using Machine Learning Models," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), pp. 627-630, doi: 10.1109/I-SMAC49090.2020.9243375, 2020

[4] Golmei Shaheamlung, Harshpreet Kaur, “The Diagnosis of Chronic Liver Disease using Machine Learning Techniques”, ResearchGate, IT in Industry, vol. 9, no.2, 2021

[5] Selvalakshmi VM, S Nirmala Devi, “Segmentation and 3D visualization of liver, lesions and major blood vessels in abdomen CTA images”, Biomedical Research, Volume 28, Issue 16, 2017

[6] M. F. Rabbi, S. M. Mahedy Hasan, A. I. Champa, M. AsifZaman and M. K. Hasan, "Prediction of Liver Disorders using Machine Learning Algorithms: A Comparative Study”, 2nd International Conference on Advanced Information and Communication Technology (ICAICT), pp. 111-116, doi: 10.1109/ICAICT51780.2020.9333528, 2020

[7] Ahmed Gaber, Hassan A. Youness, Alaa Hamdy, “Automatic Classification of Fatty Liver Disease Based on Supervised Learning and Genetic Algorithm”, IEEE Access, 10.1109/ACCESS.2021.3049341, 2021

[8] S. Sontakke, J. Lohokare, and R. Dani, “Diagnosis of Liver Diseases using Machine Learning”, International Conference on Emerging Trends & Innovation in ICT (ICEI), 2017.

[9] Ramana, B. V., Babu, M. S. P., & Venkateswarlu, N. B., “A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis”, International Journal of Database Management Systems (IJDMS), 2011