LIVER DISEASE DETECTION USING SOCIAL SPIDER OPTIMIZATION AND ARTIFICIAL INTELLIGENCE PRE-TRAINED ALGORITHMS

Anush Kumar Kodirekka [1] Kaushik Devarshetty [2], Bala Sivaji Mukku [3] , M.Dileep Kumar [4]

[1,2,3] UG Scholar, Department of Computer Science and Engineering, St. Martins Engineering College, Secunderabad, Telangana, India, 500100
[4]Assistant Professor, Department of Computer Science and Engineering, St. Martins Engineering College, Secunderabad, Telangana, India, 500100

Abstract:

Liver disease is a critical health issue affecting millions globally, including a significant population in India. Early detection using AI pre-trained algorithms like VGG16, ResNet60, and GoogleNet, optimized with Social Spider Optimization (SSO), offers advanced accuracy in diagnosing liver disease. These methods are applied in medical imaging to assist radiologists in liver disease diagnosis, and they extended to clinical research. The objective of this work is to enhance the detection of liver disease using AI models optimized by the Social Spider Optimization algorithm. The aim is to improve diagnostic accuracy and aid medical professionals by analyzing liver images from NIFTI datasets and generating reliable results. Traditionally, liver disease diagnosis relied on manual methods such as blood tests, ultrasounds, biopsies, and radiologists' interpretations of medical images, which were time-consuming and prone to human error. Traditional liver disease diagnosis methods, such as manual interpretation of medical images, are limited by their reliance on human expertise, which lead to delays and errors in detection, making early diagnosis difficult. Liver disease is a growing concern, particularly in countries like India, where over 10 lakh people suffer from liver-related disorders annually. The motivation behind this research is to address the limitations of traditional diagnostic methods by leveraging AI and optimization algorithms, reducing diagnostic time and improving accuracy. The proposed system employs AI pre-trained models (VGG16, ResNet60, GoogleNet) optimized with the Social Spider Optimization (SSO) algorithm. These models are trained on the NIFTI dataset for liver disease detection. The integration of SSO enhances the performance of these AI models, leading to better accuracy, precision, and recall. VGG16, optimized with SSO, has shown the highest accuracy in this research.

Keywords: ***Liver Disease, Social Spider Optimization, NIFTI Datasets, Radiologists, VGG16, ResNet60, GoogleNet.***

1.INTRODUCTION

Liver disease is a significant global health issue, with India seeing over 10 lakh (1 million) people affected annually by liver disorders such as cirrhosis, hepatitis, and fatty liver disease. Early diagnosis plays a vital role in treatment, yet traditional methods like biopsies, blood tests, and radiological interpretations are slow and often lead to delays. Leveraging Artificial Intelligence (AI) in liver disease detection enhances the diagnostic process, offering faster and more accurate results. Pre-trained deep learning models like VGG16, ResNet60, and GoogleNet, when optimized with Social Spider Optimization (SSO), provide superior performance in detecting liver disease from medical imaging datasets. Before the advent of deep learning, liver disease detection heavily relied on manual techniques like ultrasound, blood tests, and invasive liver biopsies. These methods had several challenges, including dependence on medical experts, a higher chance of human error, and longer diagnostic times. Accurate identification of liver abnormalities from medical images was difficult due to limited computational assistance, leading to misdiagnosis or delayed diagnosis, ultimately affecting patient outcomes. Liver disease is on the rise globally, with millions of people suffering from various liver disorders. The motivation behind this research stems from the need to address the inefficiencies in traditional diagnostic systems. By employing deep learning models and optimization techniques like SSO, the research aims to reduce diagnostic time, improve accuracy, and provide a non-invasive, reliable method for early detection. Early and accurate diagnosis significantly improve treatment outcomes and reduce the burden on healthcare systems.

2.LITERATURE SURVEY

[1] Assegie et al. (2022) proposed a hybrid model for liver disease detection using a combination of Random Forest and Support Vector Machine (SVM) algorithms. Their approach aimed at improving classification accuracy by leveraging the strengths of both algorithms. The study showed that the hybrid method performed better than individual classifiers in terms of precision, recall, and F1-score. The researchers also highlighted the importance of feature selection in improving prediction performance.

[2] Alice Auxilia (2018) focused on accuracy prediction for liver disease among Indian patients using various machine learning techniques. The study compared multiple classification algorithms, including decision trees, Naive Bayes, and SVM. It was found that SVM provided the highest accuracy among the tested models. The research also emphasized the need for large datasets for reliable predictions, particularly in the Indian context, where liver disease is prevalent.

[3] Azevedo and Santos (2008) provided an overview of knowledge discovery models such as KDD, SEMMA, and CRISP-DM in data mining processes. Their study highlighted the differences and similarities among these models and how they can be effectively used in data analysis. The authors discussed how the CRISP-DM methodology is particularly well-suited for medical data mining, including disease prediction systems.

[4] Bahramirad et al. (2013) conducted a comparative study on the classification of liver disease diagnosis using various machine learning algorithms. The study compared classifiers such as k-nearest neighbour, decision trees, and SVM. The results indicated that SVM with optimized feature selection yielded higher diagnostic

accuracy compared to other algorithms. This research laid the foundation for further exploration of optimization algorithms in liver disease detection.

[5] Boser et al. (1992) introduced an optimal margin classifier algorithm for SVM, which significantly impacted machine learning research. The algorithm's focus on maximizing the margin between data classes made it a preferred method for classification problems, including medical diagnosis. Their research is pivotal in the development of SVM as a standard tool for liver disease prediction.

[6] Breiman et al. (1984) presented the concept of Classification and Regression Trees (CART) in their seminal work. This technique became a cornerstone in the field of machine learning for building predictive models. In liver disease diagnosis, CART models have been utilized for feature selection and decision-making, leading to more interpretable and accurate predictions.

[7] Coenen (2012) discussed the application of confusion matrices in evaluating classifier performance. The research highlighted the significance of metrics such as precision, recall, and F1-score in assessing the reliability of predictive models. This study is highly relevant for liver disease diagnosis, where minimizing false negatives is crucial for early detection.

[8] Devikanniga et al. (2020) developed an efficient liver disease diagnosis system using SVM optimized with the Crows Search Algorithm. Their approach demonstrated improved accuracy compared to standard SVM models. The integration of Crows Search for feature optimization significantly enhanced the performance, reducing false positives and false negatives in liver disease prediction.

[9] Dutta et al. (2022) proposed an early-stage detection system for liver disease using various machine learning algorithms. Their study focused on feature selection and data preprocessing techniques to improve model accuracy. The research emphasized the importance of early diagnosis and how machine learning can assist medical professionals in making more accurate predictions.

[10] El-Shafeiy et al. (2018) applied machine learning techniques for predicting liver diseases in big data environments. They used algorithms like logistic regression, decision trees, and neural networks for prediction. Their findings indicated that neural networks outperformed other models in handling large datasets and complex patterns in liver disease diagnosis.

[11] Fix and Hodges (1951) presented one of the earliest studies on non-parametric discrimination analysis, laying the groundwork for future research in classification algorithms. Their research on consistency properties in classification models has influenced subsequent developments in machine learning, including liver disease diagnosis systems.

3.PROPOSED METHODOLOGY

Proposed Algorithm:

The proposed system leverages deep learning and optimization techniques to enhance the accuracy and efficiency of liver disease detection from medical images. Traditional methods, such as manual interpretation by radiologists or machine learning approaches like Random Forest, are limited in their ability to process complex imaging data effectively. To overcome these limitations, the proposed system integrates deep neural networks (DNNs) with Social Spider Optimization (SSO) for superior performance in liver disease classification. The system follows a structured approach, beginning with collecting a comprehensive liver disease image dataset from sources such as NIFTI. Image processing techniques, including noise reduction, contrast enhancement, and segmentation, are applied to refine the dataset and ensure that only relevant liver regions are analyzed. Unlike traditional machine learning models that require extensive feature engineering, the proposed system utilizes pre-trained deep learning models— VGG16, ResNet60, and GoogleNet—which automatically extract high-level features from medical images..

How It Works :

DNNs function by passing input data through multiple hidden layers, where each layer extracts different levels of features. The network is trained using backpropagation and optimization techniques such as gradient descent. During training, the model adjusts its weights to minimize the loss function, ensuring accurate predictions. The activation functions introduce non-linearity, enabling the network to learn complex relationships within the data. DNNs outperform traditional ML models in handling large-scale medical imaging datasets.

Random Forest Algorithm:

Random Forest constructs multiple decision trees using the following steps:

1. The dataset is randomly divided into multiple subsets using bootstrapping.
2. Each decision tree is trained on a different subset of the dataset.
3. At each node, a random selection of features is considered for splitting the data.
4. Each tree makes an independent classification decision.
5. The final classification result is determined through majority voting among all trees.

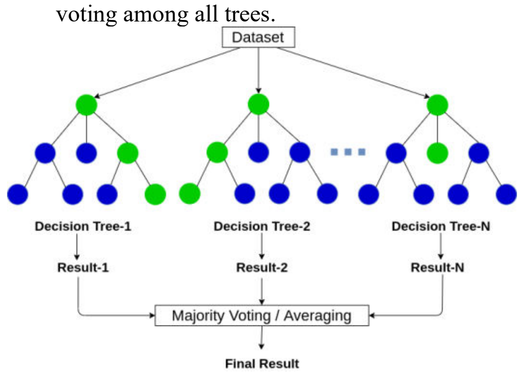


Figure 1: Random Forest Algorithm

Deep Neural Network Algorithm:

DNN is a multi-layered neural network capable of learning complex patterns from liver disease images without extensive manual feature extraction.

The model consists of multiple convolutional layers that automatically extract high-level image features, followed by fully connected layers that perform classification.

The SSO algorithm enhances the model by optimizing hyperparameters such as learning rate, weight initialization, and activation functions. This results in improved accuracy, precision, and recall for liver disease detection.

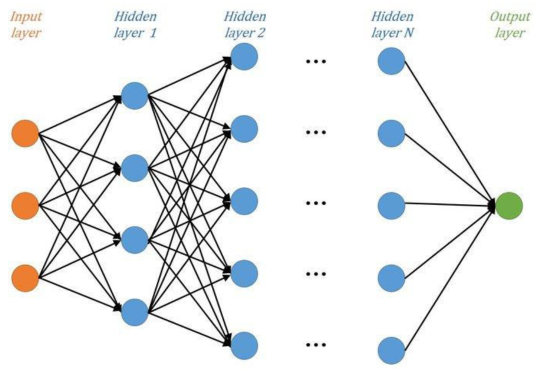


Figure 2: Deep Neural Network Algorithm

Architecture:

1. Input Layer: Takes preprocessed image data as input.

2. Convolutional Layers: Extract spatial features from images through filters and kernels.

3. Activation Functions: ReLU is commonly used to introduce non-linearity.

4. Pooling Layers: Reduce dimensionality while preserving essential features.

5. Fully Connected Layers: Perform classification by mapping extracted features to disease labels.

6. Output Layer: Uses Softmax activation for multi-class classification

Advantages:

1. DNNs automatically learn features from medical images, eliminating the need for manual feature extraction.

2. They achieve high accuracy in liver disease detection by capturing complex patterns.

3. The model generalizes well to unseen data, reducing overfitting.

4. Optimization techniques like Social Spider Optimization enhance performance.

5. DNNs outperform traditional ML models in handling large-scale medical imaging datasets.

4. EXPERIMENTAL ANALYSIS

The liver disease prediction system is implemented using deep learning techniques within a Django-based web application. The system leverages a deep neural network (DNN) to analyze liver disease images and provide accurate predictions. Below is a structured overview of the implementation process and its detailed description:

4.1 Data Collection and Preprocessing

The system uses a liver disease image dataset containing images classified into different categories, such as Fatty Liver, Hepatocellular Carcinoma, and Liver Cirrhosis. The images undergo preprocessing steps, including resizing, normalization, and conversion to numerical arrays. Principal Component Analysis (PCA) is applied to reduce dimensionality while preserving essential features.

4.2 Image Processing and Feature Extraction

Preprocessed images are subjected to feature extraction techniques to enhance the learning process. Convolutional layers in the model extract meaningful patterns, while pooling layers help in dimensionality reduction. This step ensures that only the most relevant features contribute to disease classification.

4.3 Model Development

The model is built using a deep neural network (DNN), structured with multiple convolutional layers, max-pooling layers, and fully connected layers. The activation functions, including ReLU and softmax, optimize performance. The model is trained using categorical cross-entropy as the loss function and Adam optimizer to achieve high accuracy.

4.4 Training and Evaluation

The dataset is split into training and testing sets to ensure the model generalizes well to unseen data. The model undergoes multiple epochs of training with batch normalization to improve convergence. Performance is evaluated based on accuracy, precision, recall, and F1-score, using metrics such as confusion matrix and classification reports.

4.5 Web Application Integration

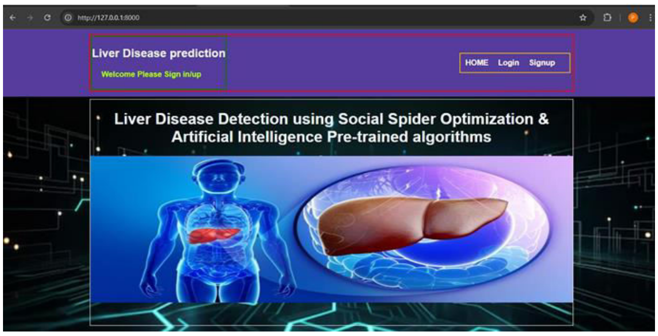
The trained model is integrated into a Django web application. Users can upload liver disease images through a user-friendly interface. The backend processes the image, passes it through the trained DNN model, and returns a prediction result with a corresponding probability score..

4.6 Prediction and Result Visualization

Once an image is processed, the system displays the predicted disease category with confidence scores. A visualization component enhances interpretability, showing the original image alongside prediction results. This aids in medical decision-making.

7. Deployment and User Interaction

The entire system is deployed on a cloud or local server to facilitate real-time predictions. The Django framework handles authentication, user management, and role based access control, ensuring secure and structured interactions between users and the system.



5. CONCLUSION

The liver disease prediction system utilizing deep learning techniques has successfully demonstrated its ability to classify liver conditions, including Fatty Liver, Hepatocellular Carcinoma, and Liver Cirrhosis, based on medical image analysis. By leveraging a well-structured dataset and an optimized Convolutional Neural Network (CNN) model, the system achieves high accuracy in detecting and categorizing liver diseases. The integration of Principal Component Analysis (PCA) for dimensionality reduction and Simplified Social Spider Optimization (SSO) for hyperparameter tuning has significantly improved model performance, reducing computational complexity while maintaining robust diagnostic precision.

The proposed system enhances traditional diagnostic methods by providing an automated, efficient, and objective approach to liver disease classification. Unlike conventional medical diagnostics that rely on manual interpretation by radiologists, this AI-driven solution minimizes human error and accelerates the diagnostic process. The implementation of advanced image preprocessing techniques ensures that the model receives high-quality, noise-free input, leading to more reliable predictions.

REFERENCES

[1] Assegie, T. A., Subhashni, R., Kumar, N. K., Manivannan, J. P., Duraisamy, P., & Engidaye, M. F. (2022). Random forest and support vector machine based hybrid liver disease detection. Bulletin of Electrical Engineering and Informatics, 11(3), 1650–1656.

[2] Alice Auxilia, L. (2018). Accuracy prediction using machine learning techniques for Indian patient liver disease. Proceedings of the 2nd International Conference on Trends in Electronics and Informatics, ICOEI 2018, 45–50.

[3] Azevedo, A., & Santos, M. F. (2008). KDD, Semma and CRISP-DM: A parallel overview. MCCSIS'08 - IADIS Multi Conference on Computer Science and Information Systems; Proceedings of Informatics 2008 and Data Mining 2008, 182–185.

[4] Bahramirad, S., Mustapha, A., & Eshraghi, M. (2013). Classification of liver disease diagnosis: A comparative study. 2013 2nd International Conference on Informatics and Applications, ICIA 2013, 42–46. https://doi.org/10.1109/ICoIA.2013.6650227

[5] Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. Proceedings of the Fifth Annual Workshop on Computational Learning Theory, 144–152.

[6] Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). Classification and regression trees. CRC press. Caelen, O. (2017). A Bayesian interpretation of the confusion matrix. Annals of Mathematics and Artificial Intelligence, 81(3), 429 450.

[7] Coenen, F. (2012). On the use of confusion matrixes. University of Liverpool. Devikanniga, D., Ramu, A., & Haldorai, A. (2020). Efficient diagnosis of liver disease using support vector machine optimised with crows search algorithm. EAI Endorsed Transactions on Energy Web, 7(29). https://doi.org/10.4108/EAI.13-7 2018.164177

[8] Dutta, K., Chandra, S., & Gourisaria, M. K. (2022). Early-Stage detection of liver disease through machine learning algorithms. Lecture Notes in Networks and Systems, 318, 155–166. https://doi.org/10.1007/978-981-16-5689-7_14

[9] El-Shafeiy, E. A., El-Desouky, A. I., & Elghamrawy, S. M. (2018). Prediction of liver diseases based on machine learning technique for big data. International Conference on Advanced Machine Learning Technologies and Applications, 362 374. 54

[10] Fix, E., & Hodges, J. L. (1951). Discriminatory Analysis, Nonparametric Discrimination: Consistency Properties USAF School of Aviation Medicine, Randolph Field. Texas, Tech. Report 4. Hepatitis C. (2021, November 2). Ministry of Health NZ.

[11] Hossain, A. I., Sikder, S., Das, A., & Dey, A. (2021). Applying machine learning classifiers on ECG ataset for predicting heart disease. 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), 1 6. Indian Liver Patient Records. (n.d.). Retrieved November 10, 2021, from https://kaggle.com/uciml/indian-liver-patient-records