Metaheuristic Optimization Review

♦ ISSN: 2692-4048 Vol: 03, No. 01, 2024 ♦ PP. 01–04

Article DOI: https://doi.org/10.54216/MOR.030101

REVIEW ARTICLE

Optimization Algorithms for Deep Learning Prediction of Liver Cirrhosis: A Survey

Aya Ebrahim[1] **Asmaa H. Rabie**[2] **El-Sayed M. El-Kenawy**[3] **Hossam El-Din Moustafa**[4]

[1] Department of Applied Health Sciences, Higher Technological Institute of Applied Health Sciences, Mansoura, Egypt
[2] Department of Computers Engineering and Control Systems, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt
[3] Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura 35111, Egypt
[4] Professor at the Department of Electronics and Communications Engineering, Faculty of Engineering, Mansoura University, Egypt

Emails: eng2008aya@gmail.com · asmaahamdy@mans.edu.eg · skenawy@ieee.org · hossam_moustafa@mans.edu.eg

Received: October 16, 2024 Revised: December 01, 2024 Accepted: December 13, 2024 ⋆ Corresponding author

ABSTRACT

Today, new Artificial Intelligence (AI) techniques are utilized to help doctors forecast the occurrence of diseases because of the necessity of sustaining public health and early disease diagnosis. One significant kind of liver damage is liver cirrhosis, which typically results from long-term liver damage brought on by a variety of liver conditions and diseases, including hepatitis, persistent alcoholism, or heredity. We created this review to provide an overview of liver cirrhosis since it is essential to identify it early and prevent the damage from spreading throughout the liver tissues. In order to identify liver cirrhosis from biomedical markers rather than images, this study has recently conducted nine studies overlaying it with various artificial intelligence deep learning techniques. Our suggested approach used various Machine Learning (ML) models to predict the signs of cirrhosis in conjunction with other illnesses. Because this condition is so important, it is important to summarize these studies based on the methodology and findings of detection accuracy and precision.

Keywords: Liver Cirrhosis ▪ Artificial intelligence ▪ Deep learning ▪ Optimization algorithms

1. INTRODUCTION

With over a million fatalities annually, liver cirrhosis ranks as the fourteenth most prevalent cause of death globally. Cirrhosis is a severe form of liver disease and typically results in irreversible liver damage. However, the extent of damage might be minimized if liver cirrhosis is identified early. In the modern world, a number of connected technologies, including artificial intelligence, deep learning, and machine learning, are crucial for identifying disease and offering medical support. Because of the participation of both healthy and infected individuals worldwide, researchers are able to gather a massive dataset of diseases. For this reason, prediction plays a very important role in correctly predicting diseases. Lately, early disease detection has decreased the disease's complications for patients. While some people with cirrhosis have severe signs of end-stage liver disease and a low likelihood of survival, others are completely asymptomatic and have a very typical life expectancy [1].

Improving patient outcomes and lowering the morbidity and mortality rates linked to liver cirrhosis depend heavily on the ability to predict and detect the disease in its early stages. Symptoms of liver cirrhosis show up much later in the progression of the disease. In the early stages of infection, over 80% of infected patients show no symptoms, which leads to increased liver damage and high mortality rates. The most effective method for detecting liver cirrhosis is liver biopsy; nevertheless, it is an expensive and intrusive technique that is impractical for recurrent sampling. Cirrhosis is detected by

biopsy techniques, imaging studies, and blood testing. The patients' treatment cannot begin right away due to the cost of these tests and the time it takes to get the test results [2, 3].

Utilizing test data and medical history to develop machine-learning-based early detection methods that are non-invasive and more accurate can help physicians diagnose and treat patients more quickly. Initially, the quantity of patient datasets is acquired. Second, the data pre-processing step combines data cleaning, transformation, and integration. Noise data and data that is not important to the study are eliminated from the obtained data during the data cleaning phase. Several data sources are merged into a single source during the data integration process. Thirdly, the pre-trained model is used and classification performance is displayed using ROC curves, confusion matrices, and performance metrics such as accuracy, precision, F1-score, and AUC. By comparing the outcomes with those of other studies, we can determine whether methods are accurate and able to generalize disease diagnosis [4].

Serum biomarkers have received little attention in previous research, which has mostly focused on comparative studies and the creation of machine learning models incorporating imaging modalities including ultrasound, magnetic resonance imaging (MRI), and elastography. Explainable AI helps close the gap between AI models and human understanding by offering insights into the complex decision-making process of the suggested machine learning model. This improves transparency and reliability in liver cirrhosis assessment and advances early detection using biomarkers [5].

2. RELATED WORK

In [6], the authors suggested a novel model known as LivMarX. LivMarX's non-invasiveness, excellent accuracy, and efficiency make it a promising prognostic model for liver cirrhosis in clinical settings. The model can offer useful insights into the stages of liver cirrhosis by utilizing easily accessible clinical and laboratory data, ensuring that healthcare providers make well-informed decisions about patient management and treatment regimens. Patient ID, hospital days, status, prescribed medications, age, gender, ascites, hepatomegaly, spider angiomas, edema, bilirubin, cholesterol, albumin, copper, SGOT, triglycerides, platelet counts, alkaline phosphatase, prothrombin time, and disease stage were included in the dataset.

The study used a range of machine learning methods. LivMarX's performance was improved by hyperparameter optimization, and the accuracy was further increased to 86%. After feature engineering and data preprocessing, the dataset was split into 70% training, 20% test, and 10% validation sets. Logistic Regression achieved 39% accuracy, Decision Tree achieved 73.49%, Random Forest achieved 84.33%, Gradient Boosting achieved 79.51%, and CatBoost achieved 78.31%.

In [7], the study suggested an ensemble machine learning classification method. Label encoding and min-max data normalization are performed during preprocessing. Characteristics such as age, gender, medical history, multiple medical conditions, and liver function tests are obtained using the ConvNeXt technique. Using the Improved Grasshopper Optimization Algorithm (IGOA), the key elements that further improve accuracy were chosen. Lastly, liver cirrhosis disease is classified using the optimized ensemble machine learning method known as optimized naïve Bayes and logistic regression (ONBLR). The hyperparameters are optimized using the Harris Hawks Optimization technique. The proposed method achieved 99.18% accuracy, 99.12% sensitivity, and 98.92% specificity.

In [8], machine learning algorithms such as Random Forest (RF), Extra Trees (ET), and Support Vector Machine (SVM) were used to create an intelligent automated system that can forecast the stages of cirrhosis. The public dataset includes 10,000 records and 70 attributes. Before prediction, the data was analyzed using data mining techniques. Because of the notable imbalance in the dataset's classifications, data balancing was required using the Synthetic Minority Oversampling Technique (SMOTE). In order to choose important features, feature selection methods were used, including Chi-Square, Mutual Information (MI), and Recursive Feature Elimination and Cross-Validation (RFECV) based on RF and SVM classifiers. The Extra-Trees model using the Chi-square feature selection approach reached the highest accuracy of 93.87%. Extra-Trees achieved 93.87%, SVM achieved 91.50%, and Random Forest achieved 93.08%.

In [9], machine learning methods were applied to liver cirrhosis prediction with an emphasis on improving accuracy rates. The dataset variables include age, gender, total and direct bilirubin, alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, total proteins, albumin, and the albumin/globulin ratio. The study investigated Random Forest, Gradient Boosting, MLP, Extra Trees, SVM, Decision Tree, KNN, and FNN classifiers. Random Forest creates many decision trees and returns the class mode or average prediction, Gradient Boosting progressively adds trees to reduce residual errors, MLP uses weighted inputs with activation functions such as sigmoid, ReLU, and tanh, and Extra Trees adds additional randomness during tree construction. The study reported precision of 0.68, recall of 0.5, and F1-score of 0.53.

In [10], four ensemble-based machine learning models—Random Forest, Gradient Boosting Machine, XGBoost, and Extra Trees—were developed to detect cirrhosis in hepatitis C patients. Before creating the models, pre-processing methods were used. Binarization was employed to predict only cirrhosis-affected HCV patients, so all other classes were changed to 0 and the cirrhosis class to 1. The min-max scaler was used to standardize the data, and random oversampling addressed the imbalance introduced by the binary target conversion. Grid search with stratified 10-fold cross-validation optimized hyperparameters. For the original data, RF achieved 74.22% accuracy, GBM achieved 73.93%, XGBoost achieved 73.93%, and ET achieved 74.22%. For oversampled data, RF achieved 96.48%, GBM achieved 95.70%, XGBoost achieved 90.99%, and ET achieved 96.82%. With only 16 of the 28 features, the Extra Trees model achieved 96.82% accuracy, 94.00% recall, 99.81% precision, and 96% AUC.

In [11], a deep learning model based on Multilayer Perceptron (MLP) was created to predict cirrhosis. DT, kNN, LR, NB, RF, and SVM were compared to the constructed model using accuracy, precision, recall, and F1-score. In the MLP

framework, neurons with non-linear activation functions are hierarchically connected through an input layer, hidden layers, and an output layer. Blood test results and demographic traits are displayed as input. The model uses two hidden layers, ReLU activation in the input and hidden layers, and sigmoid activation in the output layer for binary classification. GridSearchCV was used for hyperparameter analysis. DT achieved 0.6585 accuracy, kNN achieved 0.5365, LR achieved 0.6341, NB achieved 0.4146, RF achieved 0.7317, SVM achieved 0.7317, and the suggested MLP achieved 0.8048 accuracy with 0.8571 precision and 0.8571 recall.

In [12], the research provided a thorough empirical review of machine learning for early diagnosis of liver cirrhosis. The goal was to create a novel architecture for non-invasive early detection. The experiment included data acquisition, feature selection, class balancing, data preprocessing, categorical data transformation, and cross-validation. Logistic Regression and XGBoost classifiers were used. LR estimates the probability of cirrhotic and non-cirrhotic outcomes and achieved 68% accuracy. XGBoost, a scalable tree-boosting method that optimizes gradient boosting training by adding scores from decision trees, achieved 90.5% accuracy.

In [13], seven classification algorithms were used: ANNs-MLP, RF, DT, NB, KNN, LR, and SVM. The original dataset included 25 factors from 2000 actual patients. After removing variables with missing-data rates exceeding 40%, 20 variables remained. The Boruta method was used for feature selection, and density-based spatial clustering was used to detect anomalies. The accuracies of MLP-ANN, KNN, LR, NB, RF, SVM, and DT were 84.90%, 85.40%, 83.30%, 67.80%, 84.35%, 85.20%, and 87.75%, respectively. DT produced the best result.

In [14], Naïve Bayes, SVM with 10-fold cross-validation, and CART were used to predict liver cirrhosis. Model performance was assessed using F1-score, recall, accuracy, and precision. Naïve Bayes achieved 60% accuracy, 59% F1-score, 45% recall, and 98% precision. CART achieved 73% accuracy, 80% F1-score, 81% recall, and 79% precision. SVM achieved the best results with 75% accuracy, 73% precision, 100% recall, and 84% F1-score.

3. OPTIMIZATION AND LEARNING FORMULATION

The reviewed studies generally follow a common machine-learning pipeline: preprocessing, feature extraction or selection, optimization, classification, and evaluation. This process can be summarized as follows:

$$\mathscr{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{n}, \qquad \mathbf{x}_i \in \mathbb{R}^d,\ y_i \in \{1,\ldots,C\}. \tag{1}$$

After preprocessing and feature selection, the optimized representation is given by:

$$\mathbf{z}_i = \Phi(\mathbf{x}_i;\theta), \tag{2}$$

where Φ denotes the selected feature extraction or transformation method and θ denotes optimized parameters. A generic supervised learning objective is:

$$\min_{\mathbf{w}} \frac{1}{n}\sum_{i=1}^{n} \mathscr{L}\big(y_i, f(\mathbf{z}_i;\mathbf{w})\big) + \lambda\Omega(\mathbf{w}), \tag{3}$$

where $\mathscr{L}$ is the classification loss, Ω is a regularization term, and λ controls model complexity.

For binary classification, common metrics are computed as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}, \tag{4}$$

$$\text{Precision} = \frac{TP}{TP+FP}, \tag{5}$$

$$\text{Recall} = \frac{TP}{TP+FN}, \tag{6}$$

$$\text{F1-score} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}}. \tag{7}$$

4. CONCLUSION

As stated by the summary of these studies, data classification is crucial for achieving good illness detection performance. The data, which is divided into two classes—cirrhosis and normal persons—is the most crucial component in predicting the disease [15, 16]. In order to prevent overfitting and enable classifiers to detect diseases accurately, data optimization is also crucial. Many machine learning models were available, including RF, DT, SVM, ET, GBM, XGBoost, MLP, and KNN. These models assist in identifying illness at its earliest stage and preventing its spread. With the aid of optimiza-

Table 1. Different works in liver cirrhosis prediction summary.

Author	Year	AI Techniques	Optimization/Feature Selection	Metrics	Data Set	Approach	Cross Validation
Kamath et al. [6]	2024	LR, RF, DT, XGBoost, CatBoost	Genetic Algorithm, GridSearchCV	Accuracy: 86%	424 patients, 11 variables	2	5-fold
Badvath et al. [7]	2024	NB, LR	Improved Grasshopper Optimization Algorithm	Accuracy: 99.18%	–	2	–
Ali and Al-jabery [8]	2024	RF, ET, SVM	Chi-square feature selection	Accuracy: 93.87%	10000 records, 70 features	2	5-fold
Choi and Oh [9]	2024	RF, MLP, DT, KNN, ET, SVM, FNN	Stochastic Gradient Descent	Precision: 0.68; Recall: 0.5; F1-score: 0.53	616 patients, 13 variables	2	–
Alotaibi et al. [10]	2023	RF, GBM, XGBoost, ET	GridSearchCV	Accuracy: 96.82%	2038 patients, 28 variables	4	10-fold
Utku [11]	2023	DT, KNN, LR, RF, SVM	GridSearchCV	Accuracy: 80.48%	418 patients, 19 variables	2	10-fold
Arya et al. [12]	2023	LR, XGBoost	Genetic Algorithm	Accuracy: 90.5%	424 patients, 18 variables	4	20-fold
Güneş et al. [13]	2023	MLP-ANN, DT, RF, NB, SVM, KNN, LR	Boruta feature selection	Accuracy: 87.75%	2000 patients, 25 variables	2	10-fold
Ganty et al. [14]	2022	NB, SVM, CART	Swarm Optimization Algorithm	Accuracy: 75%	583 patients, 11 variables	2	10-fold

tion techniques, the results for disease identification were excellent.

REFERENCES

[1] E. Dickson, P. Grambsch, T. Fleming, L. Fisher, and A. Langworthy, "Cirrhosis patient survival prediction," UCI Machine Learning Repository, 2023.

[2] Y.-Y. Chen, C.-Y. Lin, H.-H. Yen, P.-Y. Su, Y.-H. Zeng, S.-P. Huang, and I.-L. Liu, "Machine-learning algorithm for predicting fatty liver disease in a taiwanese population," *Journal of Personalized Medicine*, vol. 12, no. 7, p. 1026, 2022.

[3] M. L. de Mello, J. C. L. de Oliveira, M. C. de Souza, and H. M. de Almeida, "Predictive modeling of fatty liver disease using machine learning techniques," *Journal of Biomedical Informatics*, vol. 83, pp. 188–196, 2018.

[4] R. Zemouri, N. Zerhouni, and D. Racoceanu, "Deep learning in the biomedical applications: Recent and future status," *Applied Sciences*, vol. 9, no. 15, p. 1526, 2019.

[5] S. Sun, Z. Cao, H. Zhu, and J. Zhao, "A survey of optimization methods from a machine learning perspective," *IEEE Transactions on Cybernetics*, vol. 50, no. 8, pp. 3668–3681, 2020.

[6] S. K. Kamath, S. K. Pendekanti, and D. Rao, "Livmarx: An optimized low-cost predictive model using biomarkers for interpretable liver cirrhosis stage classification," *IEEE Access*, vol. 12, pp. 92 506–92 522, 2024.

[7] D. Badvath, A. S. Miriyala, S. C. K. Gunupudi, and P. V. K. Kuricheti, "Onblr: An effective optimized ensemble ml approach for classifying liver cirrhosis disease," *Biomedical Signal Processing and Control*, vol. 89, p. 105882, 2024.

[8] D. S. Ali and M. A. Aljabery, "Predicting liver cirrhosis stages using extra trees, random forest, and svm with data mining techniques," *Informatica*, vol. 48, no. 21, pp. 15–26, 2024.

[9] Y.-S. Choi and E. Oh, "Investigating machine learning-based algorithms for liver cirrhosis prediction," *Advances in Engineering and Intelligence Systems*, vol. 3, no. 1, pp. 117–126, 2024.

[10] A. Alotaibi, L. Alnajrani, N. Alsheikh, A. Alanazy, S. Alshammasi, M. Almusairii, S. Alrassan, and A. Alansari, "Explainable ensemble-based machine learning models for detecting the presence of cirrhosis in hepatitis c patients," *Computation*, vol. 11, no. 6, p. 104, 2023.

[11] A. Utku, "Deep learning based cirrhosis detection," *Operational Research in Engineering Sciences: Theory and Applications*, vol. 6, no. 1, pp. 95–114, 2023.

[12] G. Arya, A. Bagwari, H. Saini, P. Thakur, C. Rodriguez, and P. Lezama, "Explainable ai for enhanced interpretation of liver cirrhosis biomarkers," *IEEE Access*, vol. 11, pp. 123 729–123 745, 2023.

[13] O. M. Güneş, P. Kasap, and B. S. Çorba Zorlu, "The comparison of machine learning classification algorithms used to diagnose liver cirrhosis disease and a brief review," *Concurrency and Computation: Practice and Experience*, vol. 35, p. e7628, 2023.

[14] J. Ganty, G. M. Wajiga, Y. M. Malgwi, and A. H. Maidabara, "A diagnostic model for the prediction of liver cirrhosis using machine learning techniques," *Computer Science & IT Research Journal*, vol. 3, no. 1, pp. 36–51, 2022.

[15] "Liver cirrhosis prediction input," Kaggle. [Online]. Available: https://www.kaggle.com/code/varshaakshir/liver-cirrhosis-prediction/input

[16] Fedesoriano, "Cirrhosis prediction dataset," Kaggle, 2021. [Online]. Available: https://www.kaggle.com/fedesoriano/cirrhosis-prediction-dataset