Artificial Intelligence–Driven Approaches For Liver Disease Diagnosis And Prediction: A Comprehensive Survey

M.S.Mahavarshini[1], S.Abarna[2], S.Rithika Sri[3], M.Tamilselvi[4], A.Sivaramakrishnan[5]
[1, 2, 3, 4]Dept of Artificial Intelligence and Data Science
[5]Assistant Professor, Dept of M Computer Science and Engineering
[1, 2, 3, 4, 5] Chettinad College of Engineering and Technology, Karur, India

Abstract- *Liver diseases, including liver cirrhosis and cholangiocarcinoma, are the composites formed in the binding process of buried solidic aggregates into light and porous preforms. Epidemiology analyses suggest increasing disease burden and long-term trends in progression internationally. Early diagnosis is essential, but the current diagnostic methods, such as biopsy and imaging interpretation, are invasive, expensive, and subjective to inter-observer variability. Recent developments in artificial intelligence (AI) and specifically the subfield of machine learning (ML), including deep learning (DL), have led to automated data-driven systems for better diagnostic accuracy and prognosis prediction of diseases. The genomic modeling based on LightGBM has a good performance to diagnose cholangiocarcinoma, and ensemble learning methods can stabilize prediction for structured clinical datasets. Very recently, methods based on deep learning networks that utilize EfficientNet-B7 and dual attention mechanisms have attained high performance for fibrosis staging. Additionally, explainable AI (XAI) frameworks like XAIHO advance interpretability and clinical trust, which is consistent with general healthcare AI transparency principles. However, the data imbalance, overfitting, and lack of interpretability problems remain. In this review, a comprehensive collection and synthesis of the state-of-the-art AI-based approaches is summarized, and methodological trends are reviewed alongside future research topics such as multimodal fusion or the development of explainable, clinically validated AI systems.*

Keywords- Liver Disease Diagnosis, Liver Cirrhosis, Cholangiocarcinoma, Machine Learning, Deep Learning, Explainable AI

I. INTRODUCTION

Liver-related diseases represent an enormous health burden globally, including cirrhosis being among the leading liver-associated causes of death worldwide [6], [15]. Cirrhosis is a condition of hepatic disease development by which fibrosis, architectural distortion, and hepatic dysfunction progress, resulting in portal hypertension, liver failure, and the risk of hepatocellular carcinoma [1]. Cholangiocarcinoma (CCA) is a bile duct malignancy that is characterized by aggressive biological behaviors and dismal survival because of the late diagnosis [2].

Conventional methods for diagnosis are liver biopsy along with ultrasonography, CT, and MRI. Although biopsy is still the gold standard, it is invasive, and sampling can be heterogeneous. Although imaging modalities are non-habitus, they rely on the experience of clinicians and their subjective interpretation [13]. These restrictions have led to the investigations on diagnostic computational methods.

With the advent of artificial intelligence (AI) as a paradigm for medical data analysis. Using structured clinical data, genomic databases, and imaging parameters, AI models set their sights on enhancing early detection and prognostication [3]. For example, ensemble learning methods on clinical data have shown better liver disease prediction results [9]. It was also able to achieve robust diagnostic performance for cholangiocarcinoma using genomic feature selection and LightGBM in [2].

Deep learning-based architectures have the potential to improve imaging-based assessment of liver disease. RefMulNet-101 with spatial and channel attention augmented has obtained the accuracy of liver fibrosis diagnosis at 98.5% [4]. Moreover, interpretable AI frameworks such as XAIHO incorporate optimization strategies and interpretability functions to improve explainability in the cirrhosis detection [5]. The significance of explainability in AI for healthcare has been systematically reviewed recently [10], [11].

This paper presents the state-of-the-art in AI-enabled diagnosis and prediction of liver diseases by summarizing recent developments, including proposed algorithms, dataset descriptions, evaluation methodologies, and future research directions.

II. PROBLEM FORMULATION AND TASK DEFINITION

In general, the AI-based work of liver disease revolves around two main tasks: diagnosing and assessing the severity of diseases.

A. Disease Diagnosis

Diagnostic models separate patients as diseased or non-diseased, or one type of liver disease from another. Zhang et al. constructed a LightGBM model for diagnostic cholangiocarcinoma based on WGCNA feature selection with AUC of 0.84 [2]. This work brings attention on how strong gradient boosting is for high-dimensional genomic data.
By using principles of ensemble learning on structured clinical data, stability in prediction is improved by integrating several base learners [9]. Moreover, genetic algorithm-optimized multi-layer perceptron (MLP) structures have been suggested in liver disease classification to enhance convergence and feature selection [12].

B. Prediction of Disease Severity and Prognosis

AI models also predict disease severity and progression rather than merely diagnose it. Imaging datasets with deep convolutional neural networks (CNNs) allow fibrosis staging and cirrhosis detection [4]. H/DAM-EN-B7 not only focuses on the fibrosis-related region of the lung but also achieves a better classification performance than EfficientNet-B7 [4].

In addition, spectral CT imaging methods quantify the degree of cirrhosis based on estimation of extracellular volume fraction in the liver and calculation of objective severity assessment indexes [13]. Such systems move beyond mere classification to predictive models of treatment response and disease progression.

III. DATASETS AND DATA MODALITIES

Diversity-data sources of AI in liver disease The AI system for liver disease uses two distinctive data sets.

A. Clinical and Laboratory Data

Clinical biomarkers, for example bilirubin, AST, ALT, and albumin, serve as structured parameters for predictive models [9], [12]. Cohort studies have been used to investigate long-term trends in cirrhosis burden at a population level using epidemiological data [6], [15].

B. Gene Expression Data

Molecular-level diagnosis based on high-throughput gene expression data can be realized. Genomic models based on LightGBM characterized DGE genes that are related to cholangiocarcinoma [2], indicating an application potential in the utilization of machine learning for molecular oncology.
C. Medical Imaging Data

Deep learning algorithm for the detection of fibrosis: Ultrasound, CT, and MRI imaging sequences are used as data sources in deep-learning-based fibrosis detection systems [4]. The estimation of severity based on spectral CT offers quantifiable imaging biomarkers for staging cirrhosis [13].

However, most studies use single-center datasets, which may reduce cross-institutional generalizability [3].

IV. FEATURE REPRESENTATION TECHNIQUES

A. Handcrafted and Structured Features

Conventional ML frameworks apply feature selection methods such as WGCNA and statistical filtering to obtain a low-dimensional space [2]. Model averaging uses some of the clinical markers to improve prediction stability [9].

B. Deep Feature Learning

Hierarchical representations are automatically learned by deep learning models using imaging data. The multi-scale spatial features through EfficientNets are inching out, and the dual attentive mechanisms directly concentrate on diagnostic regions [4]. Hybrid optimization methods can result in better performance and more interpretability [5].

V. RECOGNITION AND PREDICTION MODELS

A. Classical Machine Learning

LightGBM performs well with high-dimensional genomic datasets [2]. The ensemble learning enhances the robustness of classification and reduces the overfitting [9]. Genetic algorithm-constructed MLP systems work for parameter tuning more to support diagnosis [12].

B. Deep Learning and Explainable AI

The performance of efficientnet-b7 with dual attention is highly accurate in terms of fibrosis detection [4]. XAIHO combines deep learning techniques with

explainability and optimization approaches for the purpose of making cirrhosis detection more transparent [5].

Interpretability, as interpretability is a key feature in healthcare-focused explainability frameworks [10]. Explainable models based on echo state networks also contribute to human–machine trust calibration [11].

VI. EVALUATION PROTOCOLS

The performance metrics used are accuracy, precision, recall, F1-score, and AUC-ROC [2], [4], [9]. Cross-validation and independent validation sets minimize the overfitting bias. The imaging-based severity assessment also uses quantitative biomarkers based on spectral CT analysis [13].

VII. COMPARATIVE ANALYSIS

TABLE 1: Comparative Analysis of AI-Based Approaches for Liver Disease Diagnosis and Prediction

Author(s)	Year	Disease Focus	Dataset	Method	Perform-ance
He *et al.* [1]	2025	Liver Cirrhosis (Therapy Evaluation)	Clinical trial (9 patients) + Preclinical (Rat model)	Hepatocyte-derived liver progenitor-like cells (HepLPCs)	Safe & feasible; improved biochemical markers
Zhang *et al.* [2]	2026	Cholangiocarcinoma (CCA)	GEO gene expression datasets (307 tumor, 124 control)	WGCNA + Feature selection + LightGBM	AUC = 0.84, Accuracy = 80%
Farahi *et al.* [3]	2025	Multiple Liver Diseases	Clinical, laboratory & imaging datasets	ML & DL survey (CNN, SVM, RF, U-Net)	Comparative review (performance varies by task)
Lilhoree*t al.* [4]	2026	Liver Fibrosis	LiverFib imaging dataset	EfficientNet-B7 + Dual Attention (Spatial & Channel)	Accuracy = 98.5%, AUC=0.98
Mishra *et al.* [5]	2025	Liver Cirrhosis	Public liver disease dataset (clinical features)	XAIHO (DL + XAI + Optimizers)	Accuracy = 92.35%, Improved interpretability

VIII. CHALLENGES AND OPEN ISSUES

Despite the edge of AI in diagnosing and predicting liver disease, it still faces some critical bottlenecks to gain wide-scale clinical acceptance. One such issue is the less-than-ideal training data, which is chargeable and not easily diversified. Most prior works were conducted with region-specific or institution-specific data, which limit the generalizability of these models across populations, imaging protocols, and clinical settings.

A second major problem is that of model validation. Many AI systems are built and tested in controlled research environments, but only a small number receive rigorous multi-center validation. In the absence of external validation in varied clinical settings, the validity and generalizability of these models for actual clinical use are unknown.

Interpretability is also a serious impediment. Advanced deep learning methods provide strong predictive performance but can have inscrutable mechanisms of decision-making. In the clinic, physicians need justification for diagnostic decision-making. Inadequate explanation mechanisms lead to reduced clinician confidence and adoption by regulatory agencies.

In addition, heterogeneity in disease prevalence, healthcare resources, and patient demographics is another layer of complexity. Algorithms developed in one population may not be adequate for another, making the use of flexible AI systems that can adapt to different clinical settings important.

Overcoming such challenges is important for the transformation of AI from experimental research utilities to reliable clinical decision-support systems.

IX. FUTURE RESEARCH DIRECTIONS

Follow-up studies in AI-based diagnosis of liver diseases should focus on multi-modality data integration. Integration of the structured clinical data with the genomic biomarkers and imaging features in a joint model representation may facilitate a more holistic picture of disease patterns. It is expected that multimodal systems will increase predictive robustness and then enable the identification of PD in its earliest stages.

Furthermore, it is also important to have large-scale and multi-center cooperations to establish better sets of data. Standardized benchmarking procedures and agreed validation criteria will support guaranteeing a fair comparison of model performance and reproducibility between different works.

Explainability as a design principle should continue to be core in future diagnostic systems. Explicit modeling approaches where predictions can be understood by clinicians will be essential for building clinical confidence and obtaining regulatory approval.

Furthermore, AI may also help with follow-up therapy and individualized treatment planning. Smart systems could help assess treatment response, forecast disease progression, and recognize those patients likely to be helped with regenerative or targeted therapies.Finally, future AI systems will need to trade off accuracy, interpretability, scalability, and clinical practicality in order to realize meaningful impact in healthcare.

X. CONCLUSION

In this paper, we reviewed the changing function of AI in diagnosing and predicting liver disease. The field has evolved from classical machine learning to a sophisticated deep learning architecture and hybrid modeling approach, which is designed to improve both the performance and interpretability. These breakthrough technical developments greatly enhance the diagnosis power and enlarge prospective clinical applications of computational systems in hepatology.

Nevertheless, there are still practical issues to be faced, such as data quality, generalizability and interpretability, and validation standards. Connecting experimental performance to field deployment depends on strong validation, clear decision-making processes, and integration of multiple data sources. The future of smart model-based liver disease diagnostics relies on creating reliable, explainable, and clinically applicable AI systems that can stand behind the physicians to offer timely and accurate patient care.

REFERENCES

[1] K. He, X.-J. Zhu, Y.-P. Shi, W.-J. Huang, T.-H. Yang, Z.-F. Xi, Q.-G. Li, H.-Y. Sun, L.-J. Qian, X.-S. Chen, *et al.*, "Treatment of liver cirrhosis using hepatocyte-derived liver progenitor-like cells: A prospective, open-label, single-arm safety trial," *Cell Discovery*, vol. 11, no. 1, p. 831, 2025. [Online].Available:https://doi.org/10.1038/s41421-025-00831-y

[2] Z. Zhang, X. Geng, M. Yin, Y. Liang, and G. Zheng, "Establishment and validation of a diagnostic model for cholangiocarcinoma based on LightGBM machine-learning algorithm," *Scientific Reports*, vol. 16, p. 933, 2026.[Online].Available:https://doi.org/10.1038/s41598-025-30431-5

[3] R. Farahi and N. Derakhshanfard, "A comprehensive review of the methods of diagnosing and predicting liver diseases using smart methods," *Discover Artificial Intelligence*, vol. 5, p. 230, 2025. [Online].Available:https://doi.org/10.1007/s44163-025-00483-7

[4] U. K. Lilhore, S. Simaiya, R. Alroobaea, A. M. Baqasah, M. Alsafyani, A. Alhazmi, and M. M. Khan, "Advanced liver fibrosis detection and classification through deep learning-driven image analysis," *International Journal of Computational Intelligence Systems*,vol.19,p.9, 2026.[Online].Available: https://doi.org/10.1007/s44196-025-01065-2

[5] P. K. Mishra, B. K. Chaurasia, and M. M. Shukla, "XAIHO: Explainable AI leveraging hybrid optimized framework for liver cirrhosis detection," *Discover Artificial Intelligence*, vol. 5, p. 206, 2025.[Online]. Available: https://doi.org/10.1007/s44163-025-00470-y

[6] X.Luo, Y. He, Z. Jiang, and J. Liao, "Global burden of liver cirrhosis: trends from 1990–2021 and projectionsto2060," *Journal of Health, Population and Nutrition*, vol. 44, p. 370, 2025.[Online].Available: https://doi.org/10.1186/s41043-025-01109-5

[7] R. Ades and H. Archer-Dyer, "Nutritional interventions in liver cirrhosis: dietary management for improved outcomes," *Egyptian Liver Journal*, vol. 15, p. 56, 2025. [Online].Available: https://doi.org/10.1186/s43066-025-00454-8

[8] Y. Li, X. Wang, J. Chen, *et al.*, "Stem cell–based therapeutic strategies for liver regeneration and cirrhosis treatment," *Stem Cell Research & Therapy*, vol. 16, p. 4821, 2025. [Online].Available: https://doi.org/10.1186/s13287-025-04821-5

[9] S. M. Ganie, P. K. Dutta Pramanik, and Z. Zhao, "Improved liver disease prediction from clinical data through an evaluation of ensemble learning approaches," *BMC Medical Informatics and Decision Making*, vol. 24, p. 160, 2024. [Online].
Available: https://doi.org/10.1186/s12911-024-02550-y

[10] E. Kyrimi, S. McLachlan, J. M. Wohlgemut, *et al.*, "Explainable AI: definition and attributes of a good

explanation for health AI," *AI Ethics*, vol. 5, pp. 3883–3896, 2025. [Online]. Available: https://doi.org/10.1007/s43681-025-00668-x

[11] S. Hao, F. Teng, R. Hou, *et al.*, "Explainable AI and echo state networks calibrate trust in human–machineinteraction,"*Scientific Reports*, vol. 16, p.1189,2026.[Online]. Available: https://doi.org/10.1038/s41598-025-30899-1

[12] R. Farahi, N. Derakhshanfard, and A. Ghaffari, "Intelligent decision support system for liver disease diagnosis with MLP network optimized by genetic algorithm," *International Journal of Computational Intelligence Systems*, vol. 18, p. 241, 2025. [Online]. Available: https://doi.org/10.1007/s44196-025-01013-0

[13] H. Zhang, E. Hao, D. Xia, *et al.*, "Estimating liver cirrhosis severity with extracellular volume fraction by spectral CT," *Scientific Reports*, vol. 15, p. 18343, 2025. [Online]. Available: https://doi.org/10.1038/s41598-025-03717-x

[14] A. Hernández-Rubio, S. Ceballos, X. García-Calvo, *et al.*,"Diastolic dysfunction is associated with liver fibrosis and insulin resistance in alcohol use disorder," *Journal of General Internal Medicine*, 2026. [Online]. Available: https://doi.org/10.1007/s11606-025-10106-7

[15] Z. Sarkoohi, M. M. Bastan, M. Khajuei Gharaei, *et al.*, "Epidemiological trends and burden of metabolic dysfunction-associated steatotic liver disease in the Middle East and North Africa region: a 32-year analysis of healthimpact,"*Journal of Health, Population and Nutrition*, vol. 44, p. 207,2025.[Online]. Available: https://doi.org/10.1186/s41043-025-009735