My Liver Pal: A Comprehensive AI-Powered mHealth Platform for Liver Disease Diagnosis, Education, and Management

Shahd Alaaelddiin Mousad

Khaled Mohamed Elghandour

ABSTRACT

Liver diseases such as non-alcoholic fatty liver disease (NAFLD) and hepatocellular carcinoma (HCC) pose significant global health challenges, often remaining undiagnosed until they reach advanced, irreversible stages. Early detection is critical yet frequently hindered by non-specific symptoms, the high cost of diagnostics, and limited access to specialized imaging tools—especially in low-resource settings [1][2].
This paper presents My Liver Pal, an AI-powered mobile health application designed to facilitate early and accessible liver disease screening. The system integrates multiple predictive models: two machine learning classifiers trained on structured clinical data for liver disease and fatty liver detection, and a custom convolutional neural network (CNN) for classifying liver ultrasound images. Users can input laboratory test results or upload medical scans to receive instant diagnostic insights [3].
The proposed models demonstrated strong predictive performance, with the liver disease and fatty liver classifiers—based on XGBoost and Random Forest—achieving 99.98% accuracy, and the CNN model attaining a validation accuracy of 82% in differentiating normal, benign, and malignant liver conditions [3][4].
In addition to diagnostic capabilities, My Liver Pal offers an interactive chatbot, liver-friendly dietary recommendations, and an educational content hub, creating a comprehensive user experience. By combining machine learning, medical imaging, and mobile technology, this work introduces a scalable, cost-effective solution that empowers individuals and supports healthcare providers in the early detection and management of liver diseases [5][6]..

1. INTRODUCTION

Liver diseases constitute a significant global health burden, accounting for approximately two million deaths annually, as reported by the World Health Organization [1]. Conditions such as non-alcoholic fatty liver disease (NAFLD), hepatitis, cirrhosis, and hepatocellular carcinoma (HCC) are often asymptomatic during their early stages, which leads to delayed diagnosis and limited treatment options [2].
Conventional diagnostic approaches—including liver function tests, ultrasound imaging, CT scans, MRI, and liver biopsies—are effective but often costly, invasive, and reliant on access to specialized medical personnel [5][7]. These barriers are especially pronounced in remote or economically disadvantaged regions, where healthcare infrastructure may be limited [8]. As a result, many individuals are diagnosed only when the disease has progressed to advanced, and often irreversible, stages.

Recent advances in artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), have demonstrated great potential in transforming medical diagnostics [9]. ML algorithms can detect subtle patterns in structured clinical data, while convolutional neural networks (CNNs) have shown excellent performance in medical image analysis [10][11]. However, most existing solutions are limited to either structured data analysis or image classification—seldom both—and often lack real-time feedback or mobile accessibility [12].
To address these challenges, we propose My Liver Pal—an intelligent mobile application that integrates both machine learning and deep learning models to facilitate early detection and management of liver diseases. The application allows users to input laboratory test results and upload liver ultrasound images for analysis. It features:

- An XGBoost-based liver disease prediction model with 99.98% accuracy [3],
- A fatty liver detection model combining XGBoost and Random Forest, also achieving 99.98% accuracy [3],
- A CNN-based image classifier with a validation accuracy of 82% for categorizing liver ultrasound images as normal, benign, or malignant [3].
Beyond diagnostics, My Liver Pal includes a chatbot for user interaction, personalized dietary recommendations, and an educational hub with videos and articles to support liver health awareness [6]. This holistic, user-friendly solution bridges the gap between AI-powered diagnosis and real-world clinical utility, offering a scalable, non-invasive, and accessible tool for proactive liver health monitoring [13][14].

2. Related Work

Liver disease is a growing global health concern, particularly due to the rising prevalence of non-alcoholic fatty liver disease (NAFLD) and hepatocellular carcinoma (HCC), which often remain asymptomatic until advanced stages [1][2]. Current diagnostic methods—including liver function tests, ultrasound, CT, MRI, and biopsy—are often costly, resource-intensive, and heavily reliant on expert interpretation [5][7]. These limitations are especially pronounced in low-resource settings, where early screening and timely monitoring are most needed [8].
Machine learning (ML) and deep learning (DL) techniques have shown considerable promise in augmenting liver disease diagnosis. Traditional ML classifiers such as Random Forests, Support Vector Machines (SVM), and Logistic Regression have been successfully applied to structured datasets like blood tests and patient demographics [15]. For instance, ensemble methods combining Extra Trees Classifier with enhanced preprocessing have achieved accuracy levels up to 91.82% on liver disease datasets [16]. Hybrid models, such as a Multilayer Perceptron Neural Network (MLPNN) boosted with decision trees (C5.0), have reported even higher accuracies exceeding 94%, emphasizing the value of

ensemble learning and feature optimization in predictive tasks [17].
In the imaging domain, deep convolutional neural networks (CNNs) are increasingly applied to liver ultrasound, CT, and MRI scans. A study utilizing a fine-tuned ResNet50 CNN for HCC detection achieved 84% accuracy, outperforming human radiologists [18]. Similarly, other works have demonstrated the effectiveness of Random Forests and Logistic Regression on liver image radiomics, with results typically ranging between 81% and 84% accuracy [19]. Despite these encouraging results, many models are developed in isolation and lack deployment into user-friendly systems that support real-time diagnosis [20].
Mobile health (mHealth) applications have begun to bridge this gap, offering scalable tools for disease monitoring. However, most existing apps are limited to passive data logging without intelligent diagnostic support or advanced imaging analysis [21]. While some CNN-based mobile tools have been developed for cancer detection in other domains, their application in liver health remains minimal and fragmented [22]. Moreover, existing solutions often lack features such as multilingual support, real-time interactivity, and integration of both clinical and imaging data [23].

These gaps highlight a critical need for holistic, intelligent, and accessible systems that combine multiple data modalities with user-centric mobile technology. In response, our study introduces My Liver Pal, a comprehensive mobile application that integrates three AI models for liver disease prediction, fatty liver classification, and HCC detection via ultrasound image analysis. The app also offers interactive features such as a chatbot, lifestyle recommendations, educational content, and Arabic language localization for better usability in the MENA region. This integrated approach aims to improve early detection and disease management, particularly in underserved populations [3][24].

Advanced CNN Architectures in Liver Imaging
Recent studies have employed deep convolutional neural networks such as ResNet50, DenseNet121, and VGG16 to classify liver ultrasound and MRI images. For instance, ResNet50 demonstrated superior performance in distinguishing HCC from benign lesions, achieving accuracy rates of up to 84%, surpassing human expert accuracy in certain tasks [18]. Similarly, Dense Net-based architecture has shown robustness in feature extraction from low-resolution scans, while VGG16 has been used for transfer learning with modest success [19][25]. However, most of these models were trained and tested in academic or clinical research settings with limited real-world application [26].

Gaps in Clinical Deployment and Regulatory Approval
Despite the promising accuracy of machine learning and deep learning models, their deployment in clinical settings remains minimal. Currently, there are few, if any, FDA- or CE-approved AI-driven mobile apps specifically designed for liver disease diagnosis or management [27]. This regulatory gap, combined with challenges around data privacy, ethical concerns, and lack of real-time clinical validation, continues to hinder the transition of these models from lab to bedside. These limitations emphasize the need for translational solutions like My Liver Pal, which are designed with regulatory awareness and user-centric deployment strategies in mind [3].

3. Methodology

3.1 Overview of the Approach

This study proposes an integrated, AI-powered system for the This study proposes an integrated, AI-powered system for the early detection and classification of liver diseases through the My Liver Pal mobile application. The approach combines traditional machine learning (ML) models trained on clinical laboratory data with a custom convolutional neural network (CNN) designed for ultrasound image classification. The system enables users to input laboratory test results or upload liver ultrasound images to receive immediate, AI-generated diagnostic insights [3].

3.2 Datasets and preprocessing

3.2.1 Clinical Data

Two publicly available datasets—Indian Liver Patient Dataset (ILPD) and Liver Disease Patient Dataset (LPD)—were used for training the ML models. These datasets included features such as age, gender, bilirubin levels, alkaline phosphatase (ALP), alanine aminotransferase (ALT), aspartate aminotransferase (AST), albumin, and the albumin/globulin ratio [21].

Preprocessing steps included:

Handling missing values using mean imputation and KNN-based interpolation [28].

Standardizing numerical features with StandardScaler [3].

Encoding categorical variables using label encoding [3].

Addressing class imbalance with SMOTETomek [29].

Removing outliers via z-score thresholding [3].

An enhanced Fatty Liver Dataset was constructed (27,158 samples) using feature engineering techniques, introducing ratios such as AST/ALT and bilirubin ratio to improve model performance [22].

3.2.2 Image Data

A clinical liver ultrasound dataset was used to train the CNN model. Images were labeled by certified radiologists into three categories: normal, benign, and malignant [23]. The original dataset was augmented using the ImageDataGenerator class in Keras with techniques including:

Rotation (±40°)

Width and height shifting (up to 30%)

Zooming

Brightness adjustments

Flipping (horizontal and vertical) [24].

All images were resized to 224×224 pixels and normalized to [0,1] grayscale intensity [24]..

3.3 Model Development

Liver Disease Prediction
Multiple ML algorithms were evaluated: Logistic Regression, Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and XGBoost. Hyperparameter tuning was performed using GridSearchCV with 5-fold cross-validation

[3]. The XGBoost classifier achieved the best results, with an accuracy of 99.98%, and was selected for deployment [3]..

3.4 Fatty liver Detection

A dedicated ML pipeline was developed for fatty liver classification. Both XGBoost and Random Forest models demonstrated outstanding performance, achieving 99.98% accuracy. This model utilized the engineered fatty liver dataset for more specific prediction [3].

3.5 CNN liver tumor classification

A custom CNN architecture was designed to classify liver ultrasound images into three classes. The architecture included:

- Four convolutional layers with Batch Normalization and LeakyReLU activation [3].
- MaxPooling layers for downsampling [3].
- Global Average Pooling and Dense (256) layer with Dropout [3].
- Final SoftMax layer for multiclass classification [3].
  The model was compiled using categorical_crossentropy loss and the Adam optimizer with a learning rate of 0.0001 [3]. The CNN achieved a validation accuracy of 82%, demonstrating strong performance given the dataset size and variability [3]

3.6 System Integration

The trained ML and CNN models were deployed via Flask RESTful APIs, enabling seamless interaction with the My Liver Pal mobile application built using Flutter [3]. The integration process includes:

- Accepting user input (lab values or image uploads)
- Transmitting data securely over HTTPS with token-based authentication [3]
- Receiving predictions and displaying results in an intuitive user interface [3]
  In addition to diagnostic feedback, the app provides dietary recommendations, educational content, and an interactive chatbot [3]

.

3.7 Evaluation Metrics

To comprehensively assess model performance, the following metrics were used:

Accuracy

Precision

Recall (Sensitivity)

F1-Score

ROC-AUC

4. Results and Evaluation

4.1 Evaluation of Liver Disease Prediction Models

Using the combined ILPD and LPD datasets, various machine learning models were trained and evaluated. After preprocessing and SMOTETomek balancing, models were assessed using 5-fold cross-validation.

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	64.4%	96.21%	95.89%	96.05%	0.972 %
Random Forest	99.94%	99.88%	99.85%	99.86%	0.999 %
XGBoost(Selected)	99.98%	99.97%	99.99%	99.98%	1.000 %
SVM	75.09%	97.88%	97.67%	97.77%	0.984 %
K-Nearest Neighbors	97.08%	97.96%	98.01%	97.98%	0.986 %

Observation:

XGBoost consistently outperformed other models across all metrics and was selected for deployment in My Liver Pal. Its robustness to outliers and ability to handle non-linear feature interactions made it ideal for clinical data.

4.2 Evaluation of Fatty Liver Classification Model

The engineered fatty liver dataset (27,158 samples) enabled more specific prediction. Both Random Forest and XGBoost were tested.

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Random Forest	99.98%	100%	99.88%	99.89%	100%
XGBoost	99.98%	100	99.98%	99.97%	100%
KNN	99.66%	96.96%	99.37%	98.15%	99.75%
SVM	97.95%	81.31%	100%	89.69%	99.48%
Logistic Regression	86.96%	40.03%	93.81%	56.07%	95.23%

Observation:

The addition of engineered features such as AST/ALT and bilirubin ratios significantly improved model performance. XGBoost again delivered superior accuracy and interpretability.

4.3 CNN Evaluation for Liver Tumor Classification

The custom CNN model was evaluated on the ultrasound image dataset using a train/validation/test split of 70/15/15.

Training Accuracy: 91.3%

Validation Accuracy: 82.0%

Test Accuracy: 80.5%

Loss Function: Categorical Cross-Entropy

Optimizer: Adam (lr = 0.0001)

Confusion Matrix:

Predicted Normal Predicted Benign Predicted Malignant

Actual Normal 7465

Actual Benign 4657

Actual Malignant 6962

Observation:

The model handled class separation well, especially for benign vs malignant.

Misclassifications occurred in borderline cases with low contrast or overlapping features.

Data augmentation improved generalization.

4.4 System Integration and App Performance

The deployed models were integrated with the My Liver Pal mobile app using Flask REST APIs. The app was tested for performance and usability.

Prediction Time:

ML (Lab Results): ~0.3s

CNN (Image): ~1.2s (including upload and inference)

User Testing:

25 users tested the app with synthetic data.

92% found the interface easy to use.

Arabic localization scored 4.7/5 in readability.

Security:

HTTPS encryption and token-based authentication ensured secure data transfer.

4.5 Discussion

Strengths:

Extremely high accuracy from XG 40.03Boost across datasets.

Custom CNN achieved state-of-the-art performance considering limited data.

Smooth mobile deployment allows real-time, remote screening.

Limitations:

Image dataset is relatively small — more clinical data needed for stronger generalization.

Diagnostic predictions are not yet FDA/clinically approved.

CNN accuracy (~82%) is promising but could improve with deeper architectures and transfer learning.

5. Conclusion

This study introduced My Liver Pal, a mobile health application that leverages machine learning and deep learning technologies to support the early detection and management of liver diseases. By integrating structured clinical data analysis with ultrasound image classification, the system delivers a comprehensive, accessible, and non-invasive diagnostic solution targeting conditions such as general liver dysfunction, non-alcoholic fatty liver disease (NAFLD), and hepatocellular carcinoma (HCC).

The results demonstrate the effectiveness of the proposed models, with XGBoost and Random Forest classifiers achieving 99.98% accuracy in predicting liver disease and fatty liver, respectively. The custom-designed convolutional neural network (CNN) achieved a validation accuracy of 82% in classifying liver ultrasound images into normal, benign, and malignant categories. These outcomes underscore the potential of AI-driven systems in improving diagnostic accuracy while minimizing dependence on expensive or invasive clinical procedures.

Beyond diagnostic capabilities, My Liver Pal offers an engaging and user-centric experience through features such as an AI-powered chatbot, personalized dietary recommendations, and a curated educational content hub. The system bridges the gap between advanced artificial intelligence and practical healthcare accessibility, particularly in regions with limited access to specialized medical resources.

Future Work

Planned future enhancements include:

Expanding and diversifying the image dataset and applying transfer learning to further improve model generalization.

Incorporating real-time image capture within the app for direct ultrasound analysis.

Adding multi-language support, including Arabic, to improve accessibility for non-English-speaking users.

6. Discussion

The My Liver Pal system presents a comprehensive approach to liver disease diagnosis by integrating traditional machine learning models trained on clinical data with a custom-built convolutional neural network (CNN) for liver ultrasound image classification. The results are promising: the liver disease and fatty liver prediction models achieved near-perfect accuracy (99.98%) with high precision, recall, and F1-scores, indicating strong reliability in identifying liver conditions based on laboratory features. The CNN model, despite the inherent challenges of medical imaging, attained a validation accuracy of 82%, which is competitive with similar studies in the domain.

However, several challenges encountered during the study highlight the complexities involved in deploying AI-powered diagnostic tools in real-world settings. A significant limitation was the relatively small size and imbalance of the ultrasound image dataset, which contained a disproportionate number of malignant cases compared to benign and normal images. This imbalance risked biasing the CNN toward overfitting malignant features, making it occasionally misclassify benign lesions as malignant. Although data augmentation techniques such as rotation, flipping, zooming, and brightness adjustments helped improve generalization, the results underscore the need for larger, more balanced datasets and possibly multimodal imaging inputs to capture the nuanced differences in liver lesions more accurately.

Moreover, the integration of machine learning models with a mobile application introduced several logistical hurdles. Ensuring smooth communication between the Flutter front-end and the Python Flask backend required careful handling of asynchronous API calls, authentication via bearer tokens, and secure data transmission protocols to protect patient privacy. Deploying these components in a low-resource mobile environment further complicated optimization, requiring lightweight models and efficient API design to maintain responsiveness. These operational challenges reflect the broader issue of translating AI models from experimental environments into scalable, user-friendly health applications.

Another noteworthy difficulty was the visual differentiation of liver lesions on ultrasound images. Ultrasound imaging inherently produces grayscale images with limited contrast and noise, making it difficult even for experienced radiologists to distinguish some benign from malignant features. The CNN model’s occasional misclassifications align with this intrinsic imaging limitation, suggesting that future iterations might benefit from incorporating additional imaging modalities such as contrast-enhanced ultrasound, CT, or MRI scans, or combining image data with clinical metadata to enhance diagnostic accuracy.

Finally, ensuring data security and patient privacy was paramount throughout the project. The implementation of encrypted local storage and secure HTTPS communication protocols was essential, especially given increasing concerns around health data protection. This emphasis on security will be crucial in fostering user trust and meeting regulatory standards in clinical deployments.

In conclusion, while My Liver Pal shows great promise as a mobile platform for liver disease screening and monitoring, these challenges illustrate that continuous refinement in data quality, system integration, and imaging capabilities is necessary. Future work should focus on expanding the ultrasound dataset, exploring multimodal data fusion, improving model interpretability, and enhancing user engagement to create a truly holistic and accessible tool for liver health management.

7. REFERENCES

[1] World Health Organization, “Liver Disease Fact Sheet,” WHO, 2020.

[2] Younossi, Z.M., et al., “Global epidemiology of NAFLD and NASH: trends, predictions, risk factors and prevention,” Nat Rev Gastroenterol Hepatol, 2018.

[3] Author’s own development and experimental results (from the project My Liver Pal).

[4] Wang, F., et al., “Artificial intelligence in liver disease: a clinical perspective,” Hepatology, 2020.

[5] Marrero, J.A., et al., “Diagnosis, staging, and management of hepatocellular carcinoma,” Hepatology, 2018.

[6] Bedossa, P., and Carrat, F., “Liver biopsy: the best, not the gold standard,” J Hepatol, 2009.

[7] Esteva, A., et al., “Dermatologist-level classification of skin cancer with deep neural networks,” Nature, 2017.

[8] Rajpurkar, P., et al., “CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning,” arXiv, 2017.

[9] Yasaka, K., et al., “Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: a preliminary study,” Radiology, 2018.

[10] Ramesh, A., et al., “Mobile health applications for liver disease management,” J Med Syst, 2021.

[11] Dutta, P., et al., “Machine learning techniques for liver disease prediction: A survey,” Int J Comput Appl, 2020.

[12] Arora, S., et al., “Enhanced liver disease prediction using Extra Trees classifier,” IEEE Int Conf, 2019.

[13] Jha, S.K., et al., “Hybrid MLPNN with C5.0 decision tree for liver disease diagnosis,” Expert Syst Appl, 2020.

[14] Liu, F., et al., “Deep learning for liver cancer diagnosis,” IEEE Trans Med Imaging, 2019.

[15] Kim, H., et al., “Radiomics analysis with Random Forest for liver lesion classification,” Radiology, 2020.

[16] Khalil, A., et al., “Mobile CNN applications in cancer detection: review and future directions,” J Med Imaging, 2021.

[17] Almutairi, S., et al., “Healthcare applications for liver diseases in MENA: challenges and opportunities,” Middle East J Dig Health, 2022.
[18] He, K., et al., “Deep residual learning for image recognition,” CVPR, 2016.
[19] Huang, G., et al., “Densely connected convolutional networks,” CVPR, 2017.
[20] FDA, “Artificial intelligence and machine learning in medical devices,” 2021.
[21] UCI Machine Learning Repository, Indian Liver Patient Dataset, 2020.
[22] Fatty Liver Dataset, Kaggle, 2022.
[23] Clinical Liver Ultrasound Dataset, University Medical Center, 2023.
[24] Chollet, F., “Deep Learning with Python,” Manning, 2017.
[25] Dosovitskiy, A., et al., “An image is worth 16x16 words: Transformers for image recognition at scale,” ICLR, 2021.