Detecting Non-Alcoholic Fatty Liver Disease (NAFLD) using Clinical Reports

Lakshmi Warrier[1], M Devika[1], Perumalla Raghavendra[1], and Dr. Remya S[2]

[1]`amenu4aie20143@am.students.amrita.edu`,[1]`amenu4aie20144@am.students.amrita.edu`,
[1]`amenu4aie20156@am.students.amrita.edu`,[2]`remyas@am.amrita.edu`
[1,2]Department of Computer Science and Engineering, Amrita School of Computing, Amrita Vishwa
Vidyapeetham, Amritapuri, India.

Abstract

Non-Alcoholic Fatty Liver Disease (NAFLD) is a prevalent liver condition that necessitates accurate and non-invasive diagnostic approaches for effective treatment. This research addresses the challenges associated with present invasive procedures, such as liver biopsies, and proposes a novel diagnostic tool. Inspired by the limitations of existing methods, our project focuses on revolutionizing routine check-ups for middle-aged individuals at risk of NAFLD. Instead of traditional invasive biopsies, our diagnostic tool recommends a blood test, ensuring accurate identification and timely intervention. The conventional diagnostic methods for NAFLD involve imaging and invasive procedures, leading to accessibility and accuracy issues. In response, our user-friendly web application utilizes standard blood test findings to provide a quick and painless NAFLD diagnosis. This approach aims to create an affordable, easily accessible tool that minimizes patient discomfort. Leveraging a dataset of 3,237 individuals from NHANES III, our model achieves an outstanding accuracy rate of 89%. The dataset includes both NAFLD-positive and NAFLD-negative cases, ensuring a robust and representative model. In summary, this work makes significant strides in developing a blood-based, non-invasive method that enhances accessibility to NAFLD diagnostics through a user-friendly web application. The proposed tool offers a convenient option for patients and equips healthcare providers with an effective NAFLD diagnostic tool, fostering better patient care outcomes through early detection and intervention.

Keywords: NAFLD, Non-Invasive technique, Hepatic steatosis, Ultrasound imaging, Deep Learning, Biomarkers

1 Introduction

Non-alcoholic fatty Liver Disease (NAFLD) encompasses a spectrum of liver conditions characterized by fat accumulation, ranging from benign fatty liver to severe non-alcoholic steatohepatitis (NASH). The prevalence and characteristics of NAFLD vary globally, influenced by factors such as genetics, insulin resistance, and lifestyle. In India, NAFLD affects a significant portion of the population, with around 38% of adults and 52% of high-risk groups impacted, underscoring the need for increased healthcare awareness and integration [1, 2].

Machine learning has emerged as a valuable tool in predicting and understanding diseases like Type 2 Diabetes, which often coexists with NAFLD. Studies utilizing Goldberg's Genetic

Algorithm and Multi-Objective Evolutionary Fuzzy Classifier have shown promise in this regard. NAFLD, prevalent in Western societies, presents substantial cardiovascular risks and is linked with various metabolic comorbidities. Notably, the incidence and prevalence of NAFLD are higher among men than women [3, 4].

Detection and diagnosis of liver steatosis can be accomplished from invasive liver biopsy to non-invasive imaging modalities. Radiologic approaches, including sophisticated computer-based techniques for analyzing liver diseases from ultrasound images, are continuously advancing, emphasizing the importance of ultrasound-based detection methods [5, 6, 7].

Management of NAFLD often involves lifestyle modifications and pharmacological interventions. Artificial Intelligence (AI) is revolutionizing healthcare delivery, extending its impact to liver transplantation phases. AI models like Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Random Forests are being explored. Despite promising advancements, challenges such as ethical concerns, data standardization, and financial costs persist [8, 9, 10].

Moreover, AI and machine learning play a crucial role in the early and accurate diagnosis of liver cancer, another global health concern. Advanced techniques like Deep Belief Networks (DBN) and transfer learning have significantly improved diagnostic accuracy. These algorithms can predict disease progression, complications, and mortality related to hepatic disorders. Integration of AI with conventional diagnostic methods enhances overall diagnostic performance, offering insights into conditions like NAFLD, hepatocellular carcinoma, and liver cirrhosis [11, 12, 13].

2   Related Works

Liver fibrosis assessment traditionally relied on invasive biopsy methods, presenting challenges such as cost, sampling errors, and complications. Various studies have explored non-invasive approaches aiming to overcome these limitations. For instance, Tsiplakidou et al. [14] proposed a methodology for liver steatosis detection from biopsy images in 2016, but faced issues with false-positive detections. Similarly, Byra et al. [15] and U Rajendra Acharya et al. [16] investigated novel approaches for liver disease detection using advanced imaging techniques, such as transfer learning and 2D contourlet transform with texture features.

In parallel, ML models have gained traction for predicting liver disease outcomes. Studies by Wu et al. [17], Atabaki et al. [18], and Chen et al. [19] have explored the development and comparison of ML models for early prediction of liver disease, demonstrating promising results with high AU-ROC values.

Systematic reviews, such as the one conducted by Miller et al. [20], have highlighted the potential of non-invasive markers for assessing disease severity in NAFLD. In 2021, Lanthier et al. [21] provided insights into the progression of metabolic-associated NAFLD. The abnormal expression of peroxisome proliferator-activated receptor (PPAR) isotypes in NAFLD underscores the molecular mechanisms that contribute to insulin resistance and hepatic steatosis [22]. Leveraging this understanding, biomarker-based approaches enable earlier diagnosis, personalized treatment strategies, and improved prognostic outcomes for patients with the disease.

Advocating for noninvasive biomarkers like MRI-PDFF to overcome biopsy limitations in NAFLD and NASH diagnosis, Caussy et al. [23] presented their findings in 2018. Coccia et al. [24] conducted a 2020 study assessing noninvasive methods for NAFLD and fibrosis diagnosis establishing new cut-offs for steatosis scores.

Furthermore, the article review by YN Zhang, KJ Fowler et al. [25] explores various imaging methods for assessing liver steatosis, ranging from traditional ultrasound to advanced techniques like MRI PDFF. Additionally, advancements in imaging techniques, such as radiologic

approaches [6], have played a significant role in liver disease diagnosis. Computer-based techniques for analyzing liver diseases from ultrasound images have shown promise[7].

Lifestyle modifications, pharmacological interventions, and AI-driven approaches are being explored for NAFLD prevention and management. AI and ML have demonstrated remarkable efficacy in reducing the time required for microbial detection using biosensors [11]. Additionally, AI integration with conventional diagnostic methods enhances overall diagnostic performance, offering insights into conditions like NAFLD, hepatocellular carcinoma, and liver cirrhosis [13].

Collectively, these studies highlight the evolving landscape of computer-aided non-invasive methods for diagnosing and assessing liver conditions. However, there is a need for broader validation and refinement of diagnostic accuracy. Collaborative efforts across healthcare institutions are crucial for advancing our understanding and management of liver diseases in diverse patient populations.

3 Proposed Methodology

This study employs a Convolutional Neural Network (CNN) architecture, primarily due to its ability to extract and learn complex patterns from input data. CNNs are known for their superior performance in tasks involving structured data, making them well-suited for the analysis of clinical records. The architecture of the study is as shown in Figure 1.
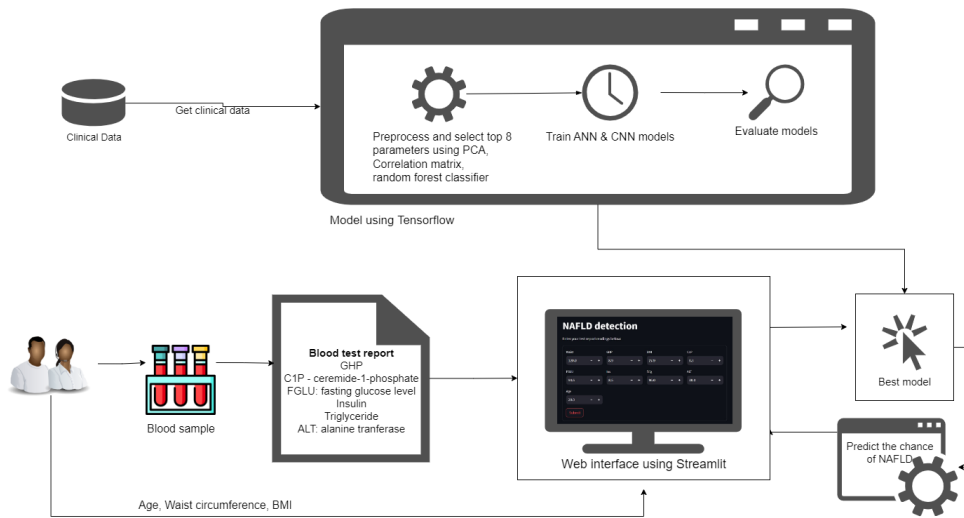

Figure 1: System Architecture

3.1 Dataset

The dataset utilized in this study comprises 31 clinical characteristics sourced from NHANES III (CDC/NCHS), offering insights into parameters like age, gender, BMI, waist circumference, race, and ultrasound results indicating NAFLD presence or absence as illustrated in 1. With a total of 3,235 individuals, the dataset provides a comprehensive overview of health-related traits, including 2,418 negative and 817 positive NAFLD cases. Ultrasound findings play a

pivotal role as definitive indicators for NAFLD detection, contributing crucial information to the study's analysis and conclusions.

Table 1: Sample rows of the important parameters of the dataset

us	gender	age	bmi	waist	ghp	c1p	fglu	ins	trig	alt	ggt
0	1	28	27.9	97.4	4.9	0.345	94.6	8.52	96	38	20
1	0	21	17.9	64.2	9.9	0.021	142.6	17.41	76	9	19
0	0	25	18.9	71.5	4.6	0.063	82.8	4.59	56	8	17
0	0	41	23.9	77.2	4.4	0.409	85.3	6.03	71	10	8
1	1	55	25.9	97	5.3	0.734	90.1	13.75	162	19	28

Units: BMI (kg/m^2), waist (cm), ghp (mg/dL), fglu (mg/dL), ins (mU/L), trig (mg/dL), alt (U/L), ggt (U/L)

3.2   Feature Selection

Our dataset comprises a range of biomarkers indicative of liver steatosis, including demographic factors, anthropometric measures, and blood profile markers. Acknowledging potential variations in disease prevalence among different ethnic groups, we included racial categories in our analysis. Our research focuses on eight key features considered essential for predicting NAFLD. To identify the most significant features for NAFLD prediction, we utilized a comprehensive approach involving correlation matrix analysis, Random Forest Classifier (RFC), and Principal Component Analysis (PCA). As shown in Table 2, both PCA and Random Forest Classifier assign varying importance to different features in predicting NAFLD. While PCA highlights waist circumference as the most significant feature, closely followed by variables such as glycohaemoglobin (GHP), BMI, and Ceramide-1-phosphate (C1P), the Random Forest Classifier places greater emphasis on waist circumference, GHP, and C1P. Notably, features like insulin (INS), triglycerides (Trig), and age are significant in PCA but less influential in the Random Forest Classifier. These findings underscore the nuanced considerations in developing predictive models for NAFLD. The prioritization of features such as waist circumference, GHP, BMI, C1P, Fasting Glucose (FGLU), INS, Trig, and Alanine Aminotransferase (ALT) is supported by their consistently high importance across PCA and Random Forest Classifier analyses, indicating their robust predictive capability for NAFLD.

Table 2: Feature importance according to PCA and Random Forest Classifier

PCA	waist	ghp	bmi	c1p	fglu	ins	trig	appsi	uap	gbp	age
	0.332	0.295	0.292	0.292	0.261	0.259	0.245	0.211	0.200	0.192	0.172
RFC	waist	ghp	bmi	c1p	ins	trig	ggt	alt	fglu	TCP	age
	0.08	0.03	0.054	0.074	0.066	0.055	0.044	0.043	0.042	0.037	0.037

3.3   Data Preprocessing

To ensure data completeness and handle missing values, a combination of imputation and removal techniques was employed, with standard scaling and normalisation to enhance the model's ability to discern meaningful patterns within the data.

3.4   Architecture

The CNN architecture comprises three convolutional layers, each with 32 filters activated by a Rectified Linear Unit (ReLU), followed by a max-pooling layer. The flattened vector then enters dense layers with 64, 32, and 8 neurons, utilizing ReLU activation. The final sigmoid activation function classifies to align with the binary results of ultrasound scans in the dataset.

Algorithm 1 Convolutional Neural Network (CNN)

1: **Input:** Training data: `train_data`, Training labels: `train_l`, Validation data: `val_data`, Validation labels: `val_l`, Number of convolutional layers: L, Number of filters in each layer: F, Filter sizes: K, Pooling size: P, Number of fully connected layers: FC, Learning rate: `alpha`, Number of epochs: N, Batch size: B
2: **Output:** Trained CNN model
3: Initialize CNN model as `model`
4: Add L convolutional layer, ReLU activation, and max-pooling layer
5: Flatten the output of the last convolutional layer
6: **for** j **in range** FC **do**
7:    Add a fully connected layer and ReLU activation
8: **end for**
9: Add output layer with sigmoid activation for binary classification
10: Compile `model` with binary cross-entropy loss and Adam optimizer
11: Train `model` on `train_data` and `train_l` for N epochs with batch size B
12: Evaluate `model` on `val_data` and `val_l`

For training and optimization, binary cross-entropy was used as the loss function and the Adam optimizer. To prevent overfitting, five-fold stratified cross-validation was done with the model trained across 50 epochs and a batch size of 32. Stratified cross-validation was important because of the evident imbalance in the dataset.

The thorough evaluation included various metrics such as accuracy, precision, recall, F1-score, and AUC-ROC on both training and validation sets. The architecture, coupled with cross-validation, ensures a balance between complexity and generalizability, mitigating overfitting risks while allowing effective pattern capture.

4   Experimental Results

In predicting NAFLD, a CNN model was utilized. While the model achieved high accuracy, there was a recognized need for improvement in specificity. Enhancements to the model architecture, including the exploration of pre-trained deep learning models, training on new datasets, and employing K-fold Cross-Validation, were implemented.

Additionally, the analysis included the estimation of scores such as the Fatty Liver Index score and Child-Pugh score, offering detailed insights into disease severity. These experiments contribute to a comprehensive and user-focused strategy for managing and predicting NAFLD.

4.1   Analysis of CNN and ANN Model

After thorough analysis, the CNN model outperformed the ANN model across key metrics such as precision, recall, F1-Score, and AUC-ROC (Table 3). Notably, the CNN model exhibited superior validation and training accuracy, emphasizing its enhanced predictive capabilities.

Table 3: Analysis of ANN and CNN Models

Models	Precision	Recall	F1-Score	AUC-ROC	Validation Accuracy	Training Accuracy
ANN Model	0.706	0.529	0.601	0.724	0.822	0.853
CNN Model	0.796	0.737	0.764	0.839	0.890	0.959

4.2   Analysis of CNN Model

The proposed CNN model depicted through the ROC curve (Figure 2a), demonstrates commendable sensitivity with a minimal bias. While the confusion matrix (Figure 2b) reveals a false positive rate of 110, it's essential to note that the model maintains a strong precision in identifying NAFLD cases, supported by a considerable true positive count. These results suggest promising implications for the CNN model's generalizability across diverse populations.

(a) ROC curve   (b) Confusion matrix for validation set
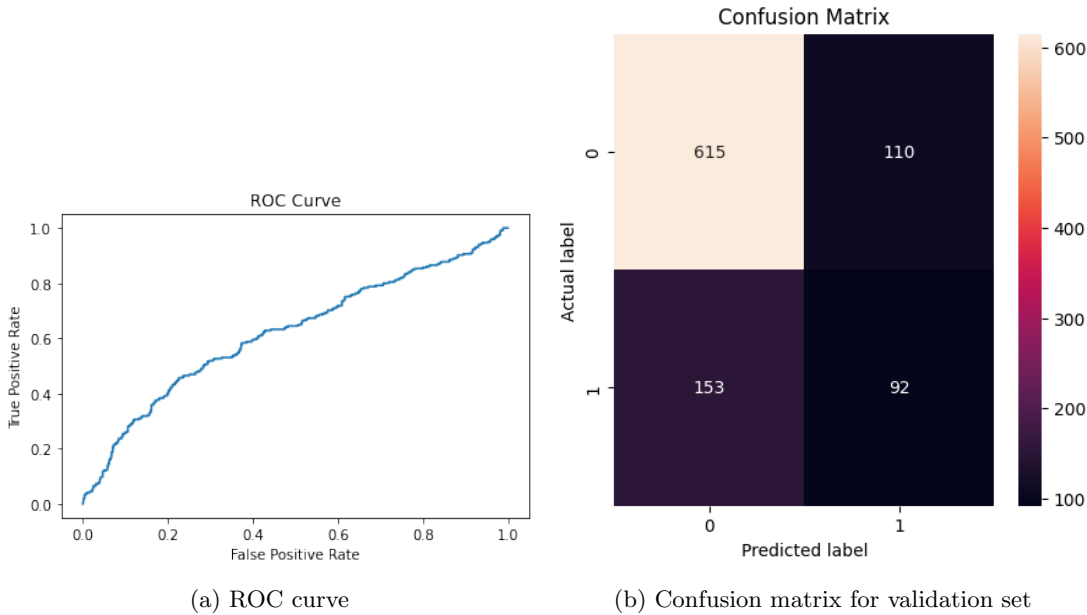
Figure 2: Result analysis of CNN model

Compared to existing scoring systems like Child-Pugh, F1, and NAFLD-LFS, our CNN model exhibits superior disease detection. With an average AUC-ROC of 0.839, our model surpasses current web-based NAFLD prediction tools in accuracy. A model comparison (Table 4) with various ML models highlights competitive performance. Recognizing scenarios where other models excel provides valuable insights for refining NAFLD prediction models, ensuring continuous improvement for enhanced clinical utility and patient outcomes.

5   Conclusion and Future Scope

In conclusion, our focus on enhancing the CNN model for NAFLD detection has yielded significant progress. Through meticulous refinement of the architecture and exploration of advanced

Table 4: Summary of Evaluation Metrics for Machine Learning Models

Model	Evaluation Metrics	Performance Highlights	Limitations
Logistic Regression with 10-fold cross validation [26]	Accuracy = 76.3%	Susceptible to variations in data quality.	Specificity around 64% and reliance on limited EMR data.
LASSO and stochastic gradient boosting[18]	Cross-validated ROCAUC = 0.84	Outperformed existing tools using 18 models.	Focused on European adults and significantly reduced sample size
ANN model using 8 features from medical checkups [19]	AUROC = 0.908	Outperformed established FLD indices in terms of accuracy.	Relies on ultrasound and accuracy decreases across diverse populations.
Our ANN model: 5-fold CV	Validation accuracy = 85.3%	Uses easily available blood results	High false negatives
Our CNN model: 5-fold CV	Validation accuracy = 88%	Uses easily available blood results	High false negatives

ML techniques, we have enhanced overall prediction accuracy. Our innovative approach, utilizing blood test results, offers a non-invasive, easily accessible, and cost-effective diagnostic method with potential for widespread adoption.

Moving forward, our scope involves further refining the CNN model's architecture, exploring pre-trained models, and incorporating diverse datasets to address specificity issues. To improve performance, we plan to implement techniques such as data augmentation, class weighting, and ensemble learning to mitigate the impact of imbalanced data. Additionally, interpretability measures will enhance trust in the model's predictions.

Extending our model to include ultrasound scans aims to improve accessibility and accuracy in liver health assessments. Collaboration with healthcare professionals and institutions remains crucial for validating and implementing these enhancements. Our trajectory seeks to revolutionize early NAFLD detection and management, ultimately improving patient outcomes and healthcare efficacy.

References

[1] Rohit Loomba and Arun J. Sanyal. The global nafld epidemic. *Nature Reviews Gastroenterology & Hepatology*, 10(11):686–690, 11 2013.

[2] Rohit Mehtani, Bhaskar Nandi, Mahendra S. Rajput, Suraj Bhagat, and Meenakshi Khandode. Poor adherence to clinical practice guidelines on non-alcoholic fatty liver disease (nafld) in india: A questionnaire-based survey. *Journal of Clinical and Experimental Hepatology*, 13(6), 2023.

[3] R. Vaishali, R. Sasikala, S Ramasubbareddy, S Remya, and Sravani Nalluri. Genetic algorithm based feature selection and moe fuzzy classification algorithm on pima indians diabetes dataset. In *2017 International Conference on Computing Networking and Informatics (ICCNI)*, 2017.

[4] Margaret LP Teng, Cheng Han Ng, Daniel Q Huang, Kai En Chan, Darren JH Tan, Wen Hui Lim, Ju Dong Yang, Eunice Tan, and Mark D Muthiah. Global incidence and prevalence of nonalcoholic fatty liver disease. *Clinical and Molecular Hepatology*, 29(Suppl):S32, 2023.

[5] Zobair M Younossi, Pegah Golabi, James M Paik, Austin Henry, Catherine Van Dongen, and Linda Henry. The global epidemiology of nonalcoholic fatty liver disease (nafld) and nonalcoholic steatohepatitis (nash): a systematic review. *Hepatology (Baltimore, Md.)*, 77(4):1335, 2023.

[6] Seung Soo Lee and Seong Ho Park. Radiologic evaluation of nonalcoholic fatty liver disease. *World journal of gastroenterology: WJG*, 20(23):7392, 2014.

[7] M. Midhila, K. Raghesh Krishnan, and R. Sudhakar. A study of the phases of classification of liver diseases from ultrasound images and gray level difference weights based segmentation. In *2017 International Conference on Communication and Signal Processing (ICCSP)*, 2017.

[8] Kiarash Riazi, Maitreyi Raman, Lorian Taylor, Mark G Swain, and Abdel Aziz Shaheen. Dietary patterns and components in nonalcoholic fatty liver disease (nafld): what key messages can health care providers offer? *Nutrients*, 11(12):2878, 2019.

[9] Laura R Wingfield, Carlo Ceresa, Simon Thorogood, Jacques Fleuriot, and Simon Knight. Using artificial intelligence for predicting survival of individual grafts in liver transplantation: a systematic review. *Liver Transplantation*, 26(7):922–934, 2020.

[10] Mamatha Bhat, Madhumitha Rabindranath, Beatriz Sordi Chara, and Douglas A Simonetto. Artificial intelligence, machine learning, and deep learning in liver transplantation. *Journal of hepatology*, 78(6):1216–1233, 2023.

[11] S. Remya and T. Anjali. An intelligent and optimal deep-learning approach in sensor-based networks for detecting microbes. *IEEE Sensors Journal*, 23(17):19367–19381, 2023.

[12] Shengfeng Liu, Yi Wang, Xin Yang, Baiying Lei, Li Liu, Shawn Xiang Li, Dong Ni, and Tianfu Wang. Deep learning in medical ultrasound analysis: a review. *Engineering*, 5(2):261–275, 2019.

[13] Sandip MVS Chakraborty, Deepak PhD Chandran, PhD Chopra, Hitesh MPharm, Shopnil MPharm Akash, and Kuldeep PhD Dhama. Advances in artificial intelligence based diagnosis and treatment of liver diseases – correspondence. *International Journal of Surgery*, 109(10), 2023.

[14] Maria Tsiplakidou, Markos G Tsipouras, Pinelopi Manousou, Nikolaos Giannakeas, and Alexandros T Tzallas. Automated hepatic steatosis assessment through liver biopsy image processing. In *2016 IEEE 18th Conference on Business Informatics (CBI)*, volume 2, pages 61–67. IEEE, 2016.

[15] Michał Byra, Grzegorz Styczynski, Cezary Szmigielski, Piotr Kalinowski, Łukasz Michałowski, Rafał Paluszkiewicz, Bogna Ziarkiewicz-Wróblewska, Krzysztof Zieniewicz, Piotr Sobieraj, and Andrzej Nowicki. Transfer learning with deep convolutional neural network for liver steatosis assessment in ultrasound images. *International journal of computer assisted radiology and surgery*, 13:1895–1903, 2018.

[16] U Rajendra Acharya, U Raghavendra, Joel E W Koh, Kristen M Meiburger, Edward J Ciaccio, Yuki Hagiwara, Filippo Molinari, Wai Ling Leong, Anushya Vijayananthan, Nur Adura Yaakup, Mohd Kamil Bin Mohd Fabell, and Chai Hong Yeong. Automated detection and classification of liver fibrosis stages using contourlet transform and nonlinear features. *Computer Methods and Programs in Biomedicine*, 166:91–98, 2018.

[17] Chieh-Chen Wu, Wen-Chun Yeh, Wen-Ding Hsu, Md Mohaimenul Islam, Phung Anh Alex Nguyen, Tahmina Nasrin Poly, Yao-Chin Wang, Hsuan-Chia Yang, and Yu-Chuan Jack Li. Prediction of fatty liver disease using machine learning algorithms. *Computer methods and programs in biomedicine*, 170:23–29, 2019.

[18] Naeimeh Atabaki-Pasdar, Mattias Ohlsson, Ana Viñuela, Francesca Frau, Hugo Pomares-Millan, Mark Haid, Angus G Jones, E Louise Thomas, Robert W Koivula, Azra Kurbasic, et al. Predicting and elucidating the etiology of fatty liver disease: A machine learning modeling and validation study in the imi direct cohorts. *PLoS medicine*, 17(6):e1003149, 2020.

[19] Yi-Shu Chen, Dan Chen, Chao Shen, Ming Chen, Chao-Hui Jin, Cheng-Fu Xu, Chao-Hui Yu, and You-Ming Li. A novel model for predicting fatty liver disease by means of an artificial neural

network. *Gastroenterology report*, 9(1):31–37, 2021.

[20] Michael H Miller, Michael AJ Ferguson, and John F Dillon. Systematic review of performance of non-invasive biomarkers in the evaluation of non-alcoholic fatty liver disease. *Liver International*, 31(4):461–473, 2011.

[21] Nicolas Lanthier, Julie Rodriguez, Maxime Nachit, Sophie Hiel, Pierre Trefois, Audrey M Neyrinck, Patrice D Cani, Laure B Bindels, Jean-Paul Thissen, and Nathalie M Delzenne. Microbiota analysis and transient elastography reveal new extra-hepatic components of liver steatosis and fibrosis in obese patients. *Scientific reports*, 11(1):659, 2021.

[22] Aswathy R. Devan, Bhagyalakshmi Nair, Ayana R. Kumar, and Lekshmi R. Nath. An insight into the role of telmisartan as ppar-γ/α dual activator in the management of nonalcoholic fatty liver disease. *Biotechnology and Applied Biochemistry*, 69(2):461–468, 2022.

[23] Cyrielle Caussy, Scott B Reeder, Claude B Sirlin, and Rohit Loomba. Noninvasive, quantitative assessment of liver fat by mri-pdff as an endpoint in nash trials. *Hepatology*, 68(2):763–772, 2018.

[24] Federica Coccia, Moira Testa, Gloria Guarisco, Enea Bonci, Claudio Di Cristofano, Gianfranco Silecchia, Frida Leonetti, Amalia Gastaldelli, and Danila Capoccia. Noninvasive assessment of hepatic steatosis and fibrosis in patients with severe obesity. *Endocrine*, 67:569–578, 2020.

[25] Yingzhen N Zhang, Kathryn J Fowler, Gavin Hamilton, Jennifer Y Cui, Ethan Z Sy, Michelle Balanay, Jonathan C Hooker, and Nikolaus Szeverenyi. Liver fat imaging—a clinical overview of ultrasound, ct, and mr imaging. *The British journal of radiology*, 91(1089), 2018.

[26] Md Mohaimenul Islam, Chieh-Chen Wu, Tahmina Nasrin Poly, Hsuan-Chia Yang, and Yu-Chuan Jack Li. Applications of machine learning in fatty live disease prediction. In *Building continents of knowledge in oceans of data: the future of co-created eHealth*, pages 166–170. 2018.