IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 3, March 2025

DOI: 10.17148/IJARCCE.2025.14301

# Using Generative Artificial Intelligence for Ultrasound Image-Based Liver Disease Diagnosis

**Frank Edughom Ekpar**[1, 2, 3]

Founder, Scholars University Ltd, Port Harcourt, Nigeria[1]

Associate Professor, Department of Computer Engineering, Rivers State University, Port Harcourt, Nigeria[2]

Principal Investigator, Department of Computer Engineering, Topfaith University, Mkpatak, Nigeria[3]

**Abstract**: This paper introduces a system that harnesses the recommendations of generative artificial intelligence (AI) and more specifically, large language models (LLMs) to develop machine learning (ML) models for the automated detection of non-alcoholic fatty liver disease (NAFLD) on the basis of liver B-mode ultrasound images. The image dataset is minimal so the option of utilizing convolutional neural networks (CNNs) and deep learning (DL) approaches built around artificial neural networks and comparable systems is not pursued. Rather, experiments are carried out with simpler machine learning algorithms and classifiers such as random forest classifier, logistic regression and decision tree classifier. Results indicate reasonable performance in light of the fact that the utilization of CNNs and comparable DL approaches could lead to overfitting of the data. The generative AI is prompted with tailored prompts engineered to elicit recommendations that account for the characteristics of the dataset.

**Keywords:** Generative Artificial Intelligence (AI), Large Language Model (LLM), Convolutional Neural Network (CNN), Deep Learning (DL), Machine Learning (ML), Healthcare System, Disease Diagnosis and Prediction, Non-alcoholic Fatty Liver Disease (NAFLD)**.**

## I. INTRODUCTION

As the most common type of liver disease, non-alcoholic fatty liver disease (NAFLD) affects a significant percentage of the global population in both developed and developing countries [1]. NAFLD can progress to a more severe form of liver disease known as non-alcoholic steatohepatitis (NASH).

One way to diagnose NAFLD is via the processing of liver B-mode ultrasound image datasets. This image processing could be automated using convolutional neural networks, comparable deep learning approaches as well a wide variety of other systems.

Early detection of NAFLD could offer valuable opportunities for lifestyle modifications and other beneficial interventions that could lead to improved health outcomes for those affected by the disease.

There has been previous research on the automated detection, prediction, and diagnosis of various diseases, including heart disease and epilepsy, using different systems, algorithms, and techniques [2] – [20]. These studies were carried out primarily in developed countries and although they have achieved varying levels of success and present a combination of advantages and disadvantages, they are susceptible to biases and limited global relevance.

As part of a drive to provide equitable application of artificial intelligence, machine learning and other advanced systems to enhance healthcare access for the highest possible number of beneficiaries, including those in low- and middle-income countries (LMICs) where the need is most acute and the available resources most constrained, Ekpar [21] – [24] created a comprehensive artificial intelligence-driven healthcare system. Featuring unique characteristics including direct support for the novel three-dimensional multilayer electroencephalography (Ekpar EEG) [25] – [27] paradigm as well as instructions for adaptation of conventional electroencephalography (EEG) systems into advanced three-dimensional multilayer electroencephalography (Ekpar EEG) mode of operation for enhanced insights into the brain and hitherto unattainable applications of EEG in domains ranging from medicine to computing and in between.

Generative artificial intelligence systems such as large language models (LLMs) are known to possess the ability to learn representations of data and draw inferences [28] – [29] and could be wielded as viable tools in the automated detection, prediction and diagnosis of diseases.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 3, March 2025
DOI: 10.17148/IJARCCE.2025.14301

This research introduces a system that harnesses recommendations extracted from generative artificial intelligence systems such as large language models (LLMs) via prompt engineering for the development of machine learning systems for the automated detection of NAFLD based on of liver B-mode ultrasound image sequences.

## II. MATERIALS AND METHODS

### Participant Recruitment

Individuals volunteered to participate in the studies that contributed to the development of the comprehensive AI-based healthcare system, with each participant providing informed consent for their involvement.

### Ethical Approval

Ethical clearance for the studies was granted by the Health Research Ethics Committee of the Rivers State University Teaching Hospital at Rivers State University. The studies adhered to all applicable ethical and regulatory standards. Publicly available data was used in compliance with the licensing terms set by the original creators.

### METHODOLOGY

Publicly accessible healthcare datasets can be enhanced by integrating data from local experiments and data collection initiatives. This combined dataset can then be used to train AI models that can make actionable predictions based on new data. Examples of public healthcare data sources include the Centers for Disease Control, the University of California Irvine Machine Learning Repository, the American Epilepsy Society, and Kaggle.

Adding local data improves model robustness, minimizes biases, and fosters inclusivity and global relevance. One key method in this project involves combining diagnostic data, such as electrocardiographic results, from local experiments with EEG data, which includes both traditional and innovative three-dimensional multilayer EEG systems.

For data collection, the project has obtained ethical approval from the appropriate research ethics committees in the regions where the experiments take place. Additionally, the project has partnered with licensed medical professionals who have direct access to patients and clinical teams. These doctors are providing anonymized clinical data to validate the AI models.

After training, the AI models will be integrated into a comprehensive healthcare system designed to support clinical decision-making for medical professionals and generate brain-computer interfaces (BCIs). The system will offer actionable predictions and insights based on new clinical data provided by healthcare providers, aiding in the early detection, diagnosis, treatment, prediction, and prevention of various conditions, including diabetes, cardiovascular diseases, stroke, autism, and epilepsy.

This project is dedicated to promoting open science, reproducibility, and collaboration. Therefore, all generated data will be made publicly accessible via repositories such as GitHub.

### SYSTEM DESIGN AND IMPLEMENTATION

The comprehensive healthcare system described in this paper follows a modular design, where each health condition (such as diabetes, heart disease, stroke, epilepsy, autism, etc.) is assigned to a separate module. This setup not only allows the system to be flexible in diagnosing and forecasting future conditions but also facilitates the efficient updating of modules by integrating new data. Furthermore, modules for Brain-Computer Interfaces (BCIs), such as those based on the motor imagery paradigm, can process EEG data to generate actionable commands and responses.

The system also includes guidelines for adapting traditional EEG systems to advanced three-dimensional multilayer EEG (Ekpar EEG) systems. These innovative systems, developed by Ekpar [25] - [27], are based on a conceptual approach that uses approximations of specific bio-signal features to characterize or manipulate the underlying biological systems.

For each module, robust AI models are created and trained using appropriately formatted data, as detailed in this paper. These AI models can incorporate various factors, including genetic, environmental, and lifestyle data, to provide more accurate representations of the participants' conditions.

 

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⋈ Peer-reviewed & Refereed journal ⋈ Vol. 14, Issue 3, March 2025

DOI: 10.17148/IJARCCE.2025.14301

Figure is a depiction of a conceptual representation of the design of the system.

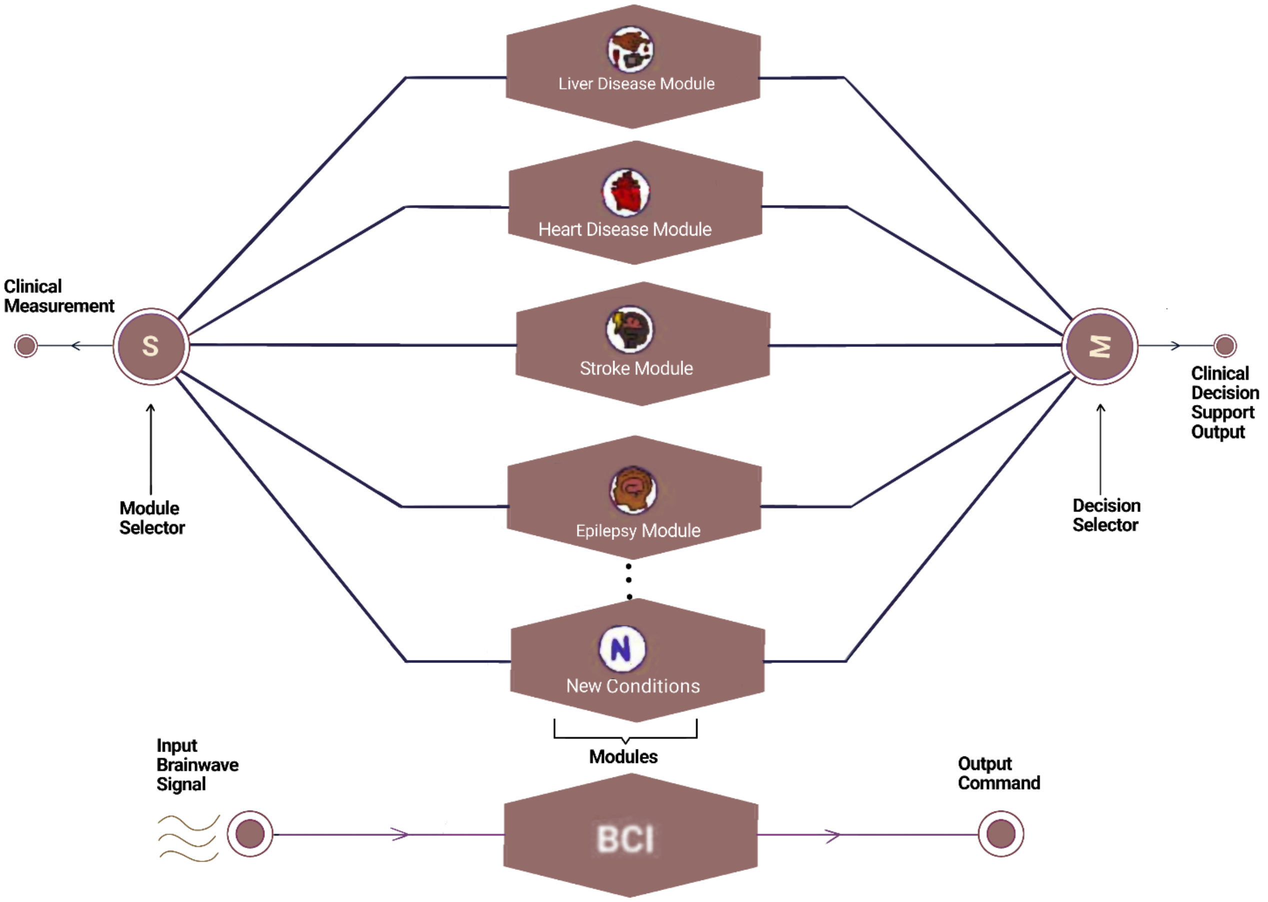


Fig. 1: System Schematic Design Diagram for the Comprehensive AI-Driven Healthcare Solution and Brain Computer Interface System. The New Conditions component represents additional health conditions that can be incorporated into the solution via new modules.

The AI models are created using four distinct approaches, as detailed below:

1. **Direct Application of LLMs:** Large Language Models (LLMs) like GPT-4 are utilized as inference engines, processing collected data formatted into multidimensional input vectors. This may also involve fine-tuning the LLM.

2. **Prompt Engineering with LLMs:** LLMs, such as Bard and GPT-4 (along with future enhanced versions), are applied through prompt engineering to define the necessary steps for developing the AI system. These steps are then executed by the developer, who applies advanced expertise in AI, neural networks, deep learning, and relevant tools like Python, TensorFlow, Keras, and other machine learning and visualization frameworks such as Scikit-learn and Matplotlib.

3. **Automated Model Creation:** Certain AI models are generated through an automated pipeline that utilizes the power of LLMs like Bard and GPT-4 (and their future versions).

4. **Manual AI Architecture Design:** The AI architecture is crafted directly by the developer using their in-depth knowledge of AI, neural networks, deep learning, and programming languages and frameworks like Python, TensorFlow, Keras, Scikit-learn, and Matplotlib.

All processes and tools involved in the system's development are thoroughly documented to ensure easy transfer and reuse of the solution.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2025

DOI: 10.17148/IJARCCE.2025.14301

The resulting AI models are evaluated and compared based on performance metrics such as specificity, sensitivity, and others, to determine their effectiveness in solving the present problem.

## AUTOMATED IMAGE-BASED NON-ALCOHOLIC FATTY LIVER DISEASE DIAGNOSIS MODULE

Prompt engineering of generative AI tools like LLMs, corresponding to the second approach outlined above was harnessed to build machine learning systems that after refinement could form part of an automated image-based non-alcoholic fatty liver disease (NAFLD) disease diagnosis module. Since the dataset contains a limited number of samples, alternatives to convolutional neural networks and deep learning approaches that could perform poorly and be susceptible to overfitting owing to the paucity of data were sought.

The generative AI tool or LLM utilized was ChatGPT.

An initial prompt was sent to ChatGPT to design an automated system for image-based detection of NAFLD using simpler alternative machine learning approaches such as Random Forest Classifier, Logistic Regression and Decision Tree Classifier. The response obtained from ChatGPT was analyzed and a follow-up prompt was generated accordingly and sent to ChatGPT to generate more specific instructions for the implementation of the system.

The recommendations of ChatGPT were then implemented in light of the characteristics of the dataset to obtain AI models that could be pressed into service for the automated image-based detection of NAFLD.

## DATASET

A dataset of liver B-mode ultrasound image sequences, provided by Byra et al. [30], was sourced from the publicly available Kaggle dataset repository. It consisted of 55 images in total, with 17 images belonging to class 0 (normal) and 38 images to class 1 (fatty). Each image had dimensions of 434 x 636 pixels.

Figure 2 illustrates the first three images from each of the two classes in the dataset.

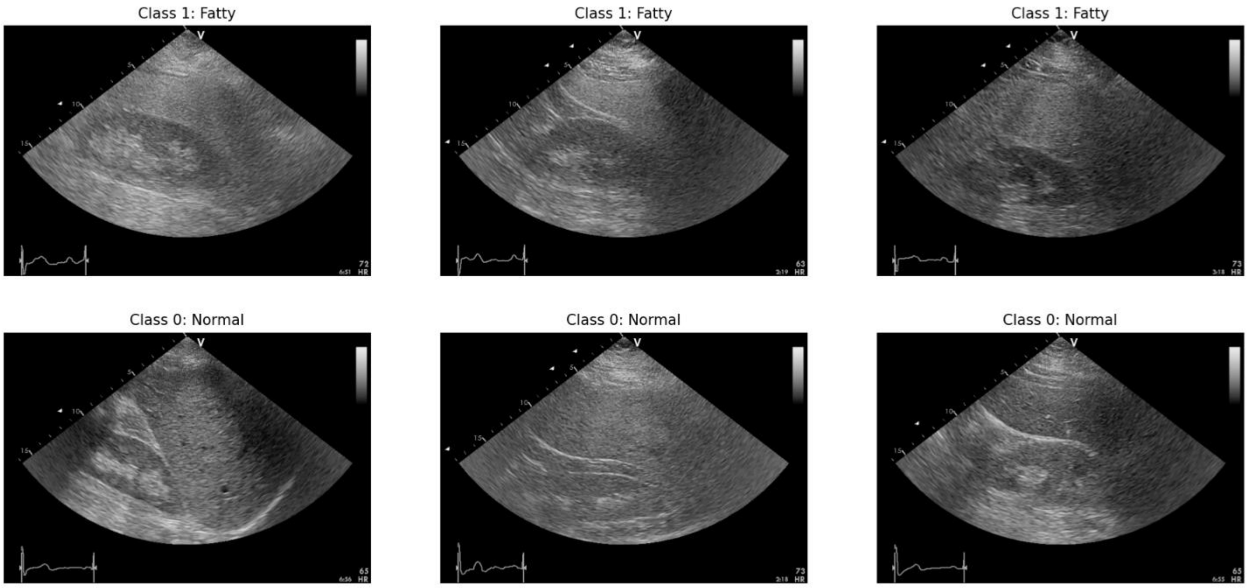


Fig. 2: First three images of each class in the source dataset.

.

## DATA AVAILABILITY

The Kaggle dataset repository served as the source for the liver B-mode liver ultrasound image sequences that were utilized in this study. The dataset is publicly available at https://www.kaggle.com/datasets/shanecandoit/dataset-of-bmode-fatty-liver-ultrasound-images.

 

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 14, Issue 3, March 2025

DOI: 10.17148/IJARCCE.2025.14301

The following section highlights the design of the system assisted by recommendations elicited from ChatGPT through careful prompt engineering.

INITIAL PROMPT

**Design a system for automated image-based detection of non-alcoholic fatty liver disease using machine learning algorithms like random forest classifier, logistic regression and decision tree classifier.**

ChatGPT RESPONSE

**Designing a system for the automated image-based detection of Non-Alcoholic Fatty Liver Disease (NAFLD) using machine learning algorithms like Random Forest Classifier, Logistic Regression, and Decision Tree Classifier involves several key steps. Here's an outline of the approach:**

- **1. Problem Definition and Objective**

**The objective is to develop a system that can classify medical images (e.g., ultrasound, MRI, CT scans) to detect the presence or absence of Non-Alcoholic Fatty Liver Disease (NAFLD). The system should use machine learning algorithms to make predictions based on image features extracted from the medical images.**

- **2. Data Collection**
- **Medical Image Data: Collect a dataset of labeled medical images, such as:**
    - **Ultrasound images**
    - **MRI or CT scans**
- **The images should be labeled for the presence or absence of NAFLD, with annotations if available (e.g., region of interest or severity).**
- **3. Preprocessing**
- **Image Preprocessing: Clean and normalize the medical images to ensure consistency.**
    - **Resizing images to a standard dimension.**
    - **Normalizing pixel values (e.g., scaling between 0 and 1).**
    - **Applying techniques like noise reduction (e.g., Gaussian blur, median filtering).**
- **Data Augmentation: Apply techniques to artificially increase the dataset size (rotation, flipping, zooming, and shifting), especially when working with limited datasets.**
- **Region-of-Interest (ROI) Extraction: Use algorithms like segmentation (e.g., thresholding or deep learning-based segmentation) to extract relevant portions of the liver from the images.**
- **4. Feature Extraction**

**Since most machine learning classifiers (like Random Forest, Logistic Regression, and Decision Trees) require structured data as input, features need to be extracted from the images.**

- **Manual Feature Engineering: Extract hand-crafted features from the images, such as:**
    - **Texture features (e.g., contrast, correlation, energy, homogeneity using Gray Level Co-occurrence Matrix (GLCM)).**
    - **Shape features (e.g., area, perimeter, and eccentricity).**
    - **Statistical features (mean, standard deviation of pixel intensity).**
- **Deep Feature Extraction: Use Convolutional Neural Networks (CNNs) to extract high-level features from images automatically. Pre-trained CNN models like ResNet, VGG, or Inception can be fine-tuned on the dataset for feature extraction.**
- **5. Data Splitting**
- **Split the dataset into training, validation, and test sets. Typically, 70% for training, 15% for validation, and 15% for testing.**
- **6. Model Building**
- **Random Forest Classifier:**
    - **Use Random Forest, an ensemble of decision trees, to classify based on the extracted features. This classifier works well with a large number of features and is robust against overfitting.**
    - **Parameters to tune: Number of trees, max depth, min samples split, etc.**
- **Decision Tree Classifier:**
    - **A decision tree classifier is a simple model that divides the feature space based on feature thresholds.**
    - **Hyperparameters to tune: max depth, min samples split, etc.**
- **Logistic Regression:**

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 3, March 2025
DOI: 10.17148/IJARCCE.2025.14301

- A logistic regression model can be used for binary classification (NAFLD present or not) based on the extracted features.
- Hyperparameters to tune: Regularization strength (C), solver type.

- 7. Model Evaluation
- Cross-validation: Use cross-validation to evaluate the performance of each model on different subsets of the training data.
- Performance Metrics:
  - Accuracy
  - Precision
  - Recall
  - F1-Score
  - AUC-ROC (Area Under the Curve for Receiver Operating Characteristics)
- 8. Model Comparison and Selection
- Compare the results of all three classifiers (Random Forest, Decision Tree, and Logistic Regression) based on the evaluation metrics.
- Select the model with the highest performance for deployment. Fine-tune the best-performing model using grid search or randomized search for hyperparameter optimization.
- 9. Post-Processing and Interpretation
- Feature Importance: Random Forest and Decision Tree classifiers offer feature importance metrics, which can help in identifying which features are most critical in detecting NAFLD.
- Model Interpretability: Tools like SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) can be used to explain the decisions made by the model for individual predictions.
- 10. Deployment
- System Integration: Integrate the machine learning model into a clinical system (e.g., a hospital's Electronic Health Record system) or a standalone software solution.
- Real-Time Prediction: Once deployed, the system should be able to take in a medical image (e.g., ultrasound or MRI), process it, and provide a prediction of NAFLD.
- Model Monitoring: Continuously monitor the system's performance to ensure it remains accurate over time. Perform regular re-training with updated datasets.
- 11. System Design Architecture
- Input: Medical images (ultrasound, CT, MRI scans).
- Preprocessing: Image resizing, normalization, noise reduction, and segmentation.
- Feature Extraction: Manual features and deep learning-based features from CNNs.
- Model: Random Forest, Decision Tree, or Logistic Regression classifier.
- Output: Classification result (presence/absence of NAFLD).
- Visualization: Display the prediction result with an option for clinicians to see the reasoning (e.g., features that led to the prediction).

---

- Tools and Libraries
- Python: Programming language for developing the system.
- Libraries:
  - OpenCV or PIL: For image processing.
  - scikit-learn: For building the machine learning models (Random Forest, Decision Trees, Logistic Regression).
  - TensorFlow/Keras: For deep learning-based feature extraction (CNNs).
  - XGBoost: For enhanced decision tree-based models.
  - SHAP/LIME: For model interpretability.

---

- Conclusion

This system involves the integration of medical image processing, feature extraction, machine learning models, and clinical deployment to automate the detection of Non-Alcoholic Fatty Liver Disease. The model's success depends on the quality of the dataset, appropriate feature extraction methods, and fine-tuning of machine learning models. By applying robust models like Random Forest, Decision Trees, and Logistic Regression, this system can significantly aid clinicians in diagnosing NAFLD with increased accuracy.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ∺ Peer-reviewed & Refereed journal ∺ Vol. 14, Issue 3, March 2025

DOI: 10.17148/IJARCCE.2025.14301

**FOLLOW-UP PROMPT**

**Design Random Forest, Decision Trees, and Logistic Regression classifiers for classification of images into two classes where each image is 434 pixels by 636 pixels.**

**ChatGPT RESPONSE**

**To design classifiers using Random Forest, Decision Trees, and Logistic Regression for classifying images into two classes, let's walk through the steps for preparing the data, preprocessing, and applying each classifier.**

**[1] Steps to build these classifiers:**

1. **Data Preprocessing**
   - **Reshape and Flatten: Each image has a size of 434x636 pixels. Since machine learning models typically expect 1D vectors as input, we need to flatten each image into a single vector of size 434×636=276,504434 \times 636 = 276,504.**
   - **Normalization: It's often beneficial to normalize the pixel values (usually scaling them between 0 and 1 or -1 and 1) to ensure the model performs well.**
   - **Train/Test Split: Split your dataset into a training set and a test set, commonly a 80/20 or 70/30 split.**
2. **Building Models**
   - **Random Forest Classifier: A Random Forest is an ensemble learning method that constructs a set of decision trees, typically used to improve the accuracy by averaging results.**
   - **Decision Tree Classifier: This is a simple model where the data is split based on feature values, forming a tree structure.**
   - **Logistic Regression: A simpler linear classifier that calculates the probability of an image belonging to one class based on a linear decision boundary.**
3. **Model Training and Evaluation**
   - **Train the classifiers using the training dataset.**
   - **Evaluate the classifiers using metrics like accuracy, precision, recall, and F1 score on the test dataset.**

**Let's break down each approach:**

**[2] 1. Random Forest Classifier**

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

# Assuming `X` is the flattened and normalized feature matrix and `y` is the target labels (0 or 1)
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

# Initialize the RandomForestClassifier
rf_classifier = RandomForestClassifier(n_estimators=100, random_state=42)

# Train the model
rf_classifier.fit(X_train, y_train)

# Predict on test data
y_pred_rf = rf_classifier.predict(X_test)

# Evaluate the performance
accuracy_rf = accuracy_score(y_test, y_pred_rf)
print(f"Random Forest Classifier Accuracy: {accuracy_rf}")
```

**[3] 2. Decision Tree Classifier**

```
from sklearn.tree import DecisionTreeClassifier

# Initialize the DecisionTreeClassifier
dt_classifier = DecisionTreeClassifier(random_state=42)

# Train the model
dt_classifier.fit(X_train, y_train)
```

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 3, March 2025

DOI: 10.17148/IJARCCE.2025.14301

```
# Predict on test data
y_pred_dt = dt_classifier.predict(X_test)

# Evaluate the performance
accuracy_dt = accuracy_score(y_test, y_pred_dt)
print(f"Decision Tree Classifier Accuracy: {accuracy_dt}")
```

**[4] 3. Logistic Regression Classifier**

```
from sklearn.linear_model import LogisticRegression

# Initialize the LogisticRegression classifier
log_reg_classifier = LogisticRegression(max_iter=1000)

# Train the model
log_reg_classifier.fit(X_train, y_train)

# Predict on test data
y_pred_log_reg = log_reg_classifier.predict(X_test)

# Evaluate the performance
accuracy_log_reg = accuracy_score(y_test, y_pred_log_reg)
print(f"Logistic Regression Classifier Accuracy: {accuracy_log_reg}")
```

**[5] Preprocessing Considerations:**

- **Flattening the Images: If the images are in the form of 2D arrays, flatten them into 1D arrays as follows:**

```
import numpy as np

# Assuming images is a list of 434x636 images
X = np.array([image.flatten() for image in images])  # Flatten each image
```

- **Normalization:**

```
X = X / 255.0  # Scale pixel values between 0 and 1 if they are originally in the range [0, 255]
```

**[6] Evaluation:**

**For each classifier, you'll evaluate them based on:**

- **Accuracy: Proportion of correctly classified images.**
- **Precision, Recall, F1-Score: In case of imbalanced classes, these metrics are important.**
- **Confusion Matrix: To understand the types of errors (false positives, false negatives).**

```
from sklearn.metrics import confusion_matrix, classification_report

# Confusion matrix and classification report for Random Forest
cm_rf = confusion_matrix(y_test, y_pred_rf)
print("Random Forest Confusion Matrix:\n", cm_rf)
print("Random Forest Classification Report:\n", classification_report(y_test, y_pred_rf))
```

**You can repeat this for each model (Decision Tree and Logistic Regression) to evaluate their performance in detail.**

**[7] Final Thoughts:**

- **Random Forest is generally preferred for image classification tasks due to its ensemble nature, which reduces overfitting and tends to work better for high-dimensional data.**
- **Decision Trees are simpler and interpretable but may overfit easily, so they are often less powerful without some tuning.**
- **Logistic Regression is very basic but can work if the data is linearly separable.**

**For better accuracy in image classification tasks, using more advanced techniques like Convolutional Neural Networks (CNNs) is typically recommended.**

Note that because of the paucity of samples in the dataset, simpler alternatives (Random Forest Classifier, Logistic Regression, Decision Tree Classifier) to more advanced techniques like convolutional neural networks (CNNs) have been adopted since CNNs and similar advanced deep learning approaches are likely to perform poorly and suffer from overfitting under these circumstances.

DOI: 10.17148/IJARCCE.2025.14301

## III. RESULTS

The recommendations made by ChatGPT, including source code in the Python programming language, were implemented faithfully using the Scikit-learn framework. Random Forest Classifier, Logistic Regression and Decision Tree Classifier were implemented as suggested. This approach was informed by the small number of samples in the dataset under consideration.

Results in terms of performance metrics (with a random state value of 42) appear in Table 1 to Table 3. An examination of the results indicates that Logistic Regression was the best performing model.

Table 1: Performance Metrics for Random Forest Classifier

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| **0** | 0.50 | 0.67 | 0.57 | 3 |
| **1** | 0.86 | 0.75 | 0.80 | 8 |
| **Accuracy** | | | 0.73 | 11 |
| **Macro Average** | 0.68 | 0.71 | 0.69 | 11 |
| **Weighted Average** | 0.76 | 0.73 | 0.74 | 11 |

Table 2: Performance Metrics for Logistic Regression

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| **0** | 0.67 | 0.67 | 0.67 | 3 |
| **1** | 0.88 | 0.88 | 0.88 | 8 |
| **Accuracy** | | | 0.82 | 11 |
| **Macro Average** | 0.77 | 0.77 | 0.77 | 11 |
| **Weighted Average** | 0.82 | 0.82 | 0.82 | 11 |

Table 3: Performance Metrics for Decision Tree Classifier

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| **0** | 0.50 | 0.67 | 0.57 | 3 |
| **1** | 0.86 | 0.75 | 0.80 | 8 |
| **Accuracy** | | | 0.73 | 11 |
| **Macro Average** | 0.68 | 0.71 | 0.69 | 11 |
| **Weighted Average** | 0.76 | 0.73 | 0.74 | 11 |

## IV. CONCLUSION

Suggestions gathered from generative artificial intelligence (AI) and more specifically, a large language model (LLM) – ChatGPT – were followed to develop a machine learning (ML) system for the automated detection of non-alcoholic fatty liver disease (NAFLD) using liver B-mode ultrasound image sequences. This approach was adopted as an alternative to the widely used convolutional neural networks (CNNs) and deep learning (DL) approaches to mitigate overfitting that could result from the small number of samples in the dataset. Availability of more data could be leveraged to refine and enhance the system for possible integration into a comprehensive artificial intelligence-driven healthcare system as a module for the automated detection of NAFLD.

## REFERENCES

[1]. NHS – Non-alcoholic Fatty Liver Disease (NAFLD): https://www.nhs.uk/conditions/non-alcoholic-fatty-liver-disease. Retrieved (2025).

[2]. Nomura, A., Noguchi, M., Kometani, M., Furukawa, K., Yoneda, T. Artificial Intelligence in Current Diabetes Management and Prediction, *Curr Diab Rep.* 21(12):61 (2021).

[3]. Kumar, Y., Koul, A., Singla, R., Ijaz, M. F. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda, *Journal of Ambient Intelligence and Humanized Computing* 14:8459–8486 (2023).

[4]. Ansari, S., Shafi, I., Ansari, A., Ahmad, J., Shah, S. I. Diagnosis of liver disease induced by hepatitis virus using artificial neural network, *IEEE Int Multitopic.* https://doi.org/10.1109/INMIC.2011.6151515 (2011).

[5]. Battineni, G., Sagaro, G. G., Chinatalapudi, N., Amenta, F. Applications of machine learning predictive models in

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁐ Peer-reviewed & Refereed journal ⁐ Vol. 14, Issue 3, March 2025
DOI: 10.17148/IJARCCE.2025.14301

the chronic disease diagnosis, *J Personal Med.* https://doi.org/10.3390/jpm10020021 (2020).

[6]. Abdar, M., Yen, N., Hung, J. Improving the diagnosis of liver disease using multilayer perceptron neural network and boosted decision tree, *J Med Biol Eng* 38:953–965 (2018).

[7]. Chaikijurajai, T., Laffin, L., Tang, W. Artificial intelligence and hypertension: recent advances and future outlook, *Am J Hypertens* 33:967–974 (2020).

[8]. Fujita, S., Hagiwara, A., Otsuka, Y., Hori, M., Takei, N., Hwang, K. P., Irie, R., Andica, C., Kamagata, K., Akashi, T., Kumamaru, K. K., Suzuki, M., Wada, A., Abe, O., Aoki, S. Deep Learning Approach for Generating MRA Images From 3D Quantitative Synthetic MRI Without Additional Scans, *Invest Radiol* 55:249–256 (2020).

[9]. Juarez-Chambi, R. M., Kut, C., Rico-Jimenez, J. J., Chaichana, L. K., Xi, J., Campos-Delgado, D. U., Rodriguez, F. J., Quinones-Hinojosa, A., Li, X., Jo, J. A. AI-Assisted *In Situ* Detection of Human Glioma Infiltration Using a Novel Computational Method for Optical Coherence Tomography, *Clin Cancer Res* 25(21):6329–6338 (2019).

[10]. Nashif, S., Raihan, R., Islam, R., Imam, M. H. Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System, *World Journal of Engineering and Technology* Vol 6, No. 4 (2018).

[11]. Chen, P. H. C., Gadepalli, K., MacDonald, R., Liu, Kadowaki, S., Nagpal, K., Kohlberger, T., Dean, J., Corrado, G. S., Hipp, J. D., Mermel, C. H., Stumpe, M. C. An augmented reality microscope with real time artificial intelligence integration for cancer diagnosis, *Nat Med* 25:1453–1457 (2019).

[12]. Gouda, W., Yasin, R. COVID-19 disease: CT Pneumonia Analysis prototype by using artificial intelligence, predicting the disease severity, *Egypt J Radiol Nucl Med* 51(1):196 (2020).

[13]. Han, Y., Han, Z., Wu, J., Yu, Y., Gao, S., Hua, D., Yang, A. Artificial Intelligence Recommendation System of Cancer Rehabilitation Scheme Based on IoT Technology, *IEEE Access* 8:44924–44935 (2020).

[14]. Chui, C. S., Lee, N. P., Adeoye, J., Thomson, P., Choi, S. W. Machine learning and treatment outcome prediction for oral cancer, *J Oral Pathol Med* 49(10):977–985 (2020).

[15]. Koshimizu, H., Kojima, R., Okuno, Y. Future possibilities for artificial intelligence in the practical management of hypertension, *Hypertens Res* 43:1327–1337 (2020).

[16]. Kather, J. N., Pearson, A. T., Halama, N., Jäger, D., Krause, J., Loosen, S. H., Marx, A., Boor, P., Tacke, F., Neumann, U. P., Grabsch, H. I., Yoshikawa, T., Brenner, H., Chang-Claude, J., Hoffmeister, M., Trautwein, C., Luedde, T. Deep learning microsatellite instability directly from histology in gastrointestinal cancer, *Nat Med* 25:1054–1056 (2019).

[17]. Kwon, J. M., Jeon, K. H., Kim, H. M., Kim, M. J., Lim, S. M., Kim, K. H., Song, P. S., Park, J., Choi, R. K., Oh, B. H. Comparing the performance of artificial intelligence and conventional diagnosis criteria for detecting left ventricular hypertrophy using electrocardiography, *EP Europace* 22(3):412–419 (2020).

[18]. Khan, M. A. An IoT Framework for Heart Disease Prediction Based on MDCNN Classifier, *IEEE Access* 8:34717–34727 (2020).

[19]. Oikonomou, E. K., Williams, M. C., Kotanidis, C. P., Desai, M. Y., Marwan, M., Antonopoulos, A. S., Thomas, K. E., Thomas, S., Akoumianakis, I., Fan, L. M., Kesavan, S., Herdman, L., Alashi, A., Centeno, E. H., Lyasheva, M., Griffin, B. P., Flamm, S. D., Shirodaria, C. Sabharwal, N., Kelion, A., Dweck, M. R., Van Beek, E. J. R., Deanfield, J., Hopewell, J. C., Neubauer, S., Channon, K. M., Achenbach, S., Newby, D. E., Antoniades, C. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography, *Eur Heart J* 40(43):3529–3543 (2019).

[20]. Sabottke, C. F., Spieler, B. M. The Effect of Image Resolution on Deep Learning in Radiography, *Radiology: Artificial Intelligence* Vol. 2. No. 1, 2:e190015 (2020).

[21]. Ekpar, F. E. A Comprehensive Artificial Intelligence-Driven Healthcare System, *European Journal of Electrical Engineering and Computer Science*, 8(3), Article 617. (2024).

[22]. Ekpar, F. E. Diagnosis of Chronic Kidney Disease Within a Comprehensive Artificial Intelligence-Driven Healthcare System, *International Journal of Advanced Research in Computer and Communication Engineering*, 13(9). (2024).

[23]. Ekpar, F. E. Image-based Chronic Disease Diagnosis Using 2D Convolutional Neural Networks in the Context of a Comprehensive Artificial Intelligence-Driven Healthcare System, *Molecular Sciences and Applications*, 4(13). (2024).

[24]. Ekpar, F. E. Leveraging Generative Artificial Intelligence Recommendations for Image-based Chronic Kidney Disease Diagnosis, *International Journal of Advanced Research in Computer and Communication Engineering*, 14(1). (2025).

[25]. Ekpar, F. E. A Novel Three-dimensional Multilayer Electroencephalography Paradigm, *Fortune Journal of Health Sciences*, 7(3). (2024).

[26]. Ekpar, F. E. System for Nature-Inspired Signal Processing: Principles and Practice, *European Journal of Electrical Engineering and Computer Science*, 3(6), pp. 1-10, (2019).

Impact Factor 8.102 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 14, Issue 3, March 2025

DOI: 10.17148/IJARCCE.2025.14301

[27]. Ekpar, F. E. Nature-inspired Signal Processing, *United States Patent and Trademark Office*, US Patent Application Number: 13/674,035 (Filed: November 11, 2012, Priority Date: December 24, 2011), Document ID: US 20140135642 A1: Published: (2014).

[28]. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D. Language Models are Few-Shot Learners, *Advances in Neural Information Processing Systems 33*, (2020).

[29]. Gurnee, W., Tegmark, M. Language Models Represent Space and Time, *arXiv*, DOI: https://doi.org/10.48550/arXiv.2310.02207 (2023).

[30]. Byra, M., Styczynski, G., Szmigielski, C., Kalinowski, P., Michałowski, L., Paluszkiewicz, R., Ziarkiewicz-Wróblewska, B., Zieniewicz, K., Sobieraj, P., Nowicki, A. 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).