International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

# Cardiovascular Disease Detection from ECG Images Using Deep Learning Techniques

***Prathamesh S Salokhe***

**Student, Rani Channamma University, Belagavi**

**ABSTRACT**

Cardiovascular diseases (CVDs) are the primary cause of death worldwide, underscoring the importance of early detection. This research utilizes Convolutional Neural Networks (CNNs) for the analysis of ECG images to classify and identify CVDs. Various machine learning models were trained and evaluated on accuracy, precision, recall, and F1-score using a extensive dataset of annotated ECG images. CNNs showed superior performance compared to traditional techniques, achieving high accuracy in identifying conditions such as arrhythmias and myocardial infarction. This showcases how machine learning has the potential to enhance the diagnosis of CVD and outcomes for patients.

**Keywords:** Convolutional Neural Networks (CNNs), Cardiovascular diseases (CVDs), ECG images, machine learning, patients

## 1. INTRODUCTION

The electrocardiogram (ECG) is a key diagnostic tool for identifying cardiovascular issues, assessing the heart's electrical and muscular functions. While the test itself is simple to conduct, interpreting ECG readings requires significant expertise. Traditionally, ECG records were maintained in paper form, making manual review labor-intensive. By digitizing these records, automated diagnosis and analysis can be improved. this project aims to use machine learning techniques to transform ECG paper records into a onedimensional signal, focusing on the P, QRS, and T waves that reflect heart activity. The process involves segmenting ECG reports into 13 leads, converting the data into signals, applying smoothing, and generating binary images through thresholding. Dimensionality reduction, such as Principal Component Analysis, will enhance data comprehension. Machine learning models, including KNN, Logistic Regression, SVM, and ensemble classifiers, will be implemented. The final model will diagnose conditions like Myocardial Infarction or arrhythmia, aiding in effective cardiac health assessment.

## 2. LITERATURE SURVEY

Detection of Cardiovascular Diseases in ECG Images Using Machine Learning and Deep Learning Methods by Mohammed B. Abubaker; Bilal Babayigit. Heart diseases are the top reason for death globally, according to statistics on cardiovascular diseases. Predicting and categorizing them sooner can lead to saving more lives. An electrocardiogram (ECG) is a widely used, affordable, and noninvasive method for assessing the heart's electrical function and identifying heart disease. This article utilized deep learning methods to forecast four main cardiac issues: irregular heartbeat, heart attack, past heart attack, and normal individuals categories from the public ECG images dataset of heart patients. Initially, the study explored transfer learning by utilizing small-scale pretrained deep neural networks such as SqueezeNet and AlexNet. Next, a different convolutional neural network (CNN) structure was suggested for predicting heart abnormalities. Next, we utilized both the pretrained models mentioned earlier and our new CNN model to extract features for traditional machine learning algorithms.[1]

Detection of Cardiovascular Diseases in ECG Images Using Machine Learning and Deep Learning Methods Mohammed B. Abubaker, Heart diseases, also known as cardiovascular diseases, are the leading cause of death worldwide. Predicting and categorizing these conditions promptly can greatly increase the likelihood of saving lives. The ECG is a widely used, inexpensive, and harmless way to evaluate the heart's electrical function, playing a vital role in identifying heart conditions. This research utilizes deep learning techniques to detect four main heart abnormalities - arrhythmia, heart attack, past heart attack, and healthy heart conditions, using a freely accessible dataset of ECG pictures from heart patients. At first, the study examined a transfer learning approach using small-scale pre-trained deep neural networks, specifically SqueezeNet and AlexNet. Later on, a new design for Convolutional Neural Network (CNN) was proposed to predict cardiac abnormalities. In addition, the pre-trained models mentioned earlier and the newly suggested CNN model were also employed for extracting features in traditional machine learning methods, such as SVM, K-NN, DT, RF, and NB. The results of the experiment show that the proposed CNN model outperforms previous studies in terms of performance metrics, with an accuracy of 98.23%, recall of

98.22%, precision of 98.31%, and F1 score of 98.21%. Furthermore, when utilized for feature extraction, the suggested CNN model achieved an outstanding accuracy of 99.79% with the Naïve Bayes algorithm.[2]

Enhancing Disease Prediction through Machine Learning Smriti Mukesh Singh1, Dr. Dinesh B. Hanchate2 – 2018, The application of Big Data in the fields of biomedical and healthcare is on the rise, as precise analysis of medical data facilitates early disease detection, patient care, and community services. Incomplete medical data can hinder the accuracy of analyses. This study proposes machine learning algorithms for the effective prediction of chronic diseases. To address the challenge of missing data, a Genetic Algorithm will be employed to reconstruct the absent information. The dataset includes both structured and unstructured data. The RNN algorithm will be utilized to extract features from the unstructured data. The framework suggests the use of SVM and Naive Bayesian algorithms for disease prediction, utilizing structured and unstructured data from hospital records, respectively. Additionally, a Community Question Answering (CQA) system is proposed, which will predict questions and answers, providing appropriate responses to users. For this purpose, two algorithms, KNN and SVM, are recommended. The KNN algorithm will classify the answers, while the SVM algorithm will also classify the responses, assisting users in finding the most relevant questions and answers related to diseases.[3]

Heart Disease Prediction System Utilizing Data Mining Techniques by Abhishiek Taneja – 2015, In the contemporary world, cardiovascular disease stands as the most deadly ailment. This condition can strike individuals with such rapidity that there is often insufficient time for effective treatment. Consequently, accurate and timely diagnosis poses a significant challenge for the medical community. An erroneous diagnosis can tarnish a hospital's reputation and lead to a loss of trust. Furthermore, the cost of treating this disease is prohibitively high for many patients, particularly in India. This paper aims to create a cost-effective treatment approach through the application of data mining technologies, thereby enhancing the decision support system for healthcare databases. Most hospitals employ some form of hospital management system to oversee patient care; however, substantial efforts are necessary to facilitate intelligent decisionmaking. Diagnosing this disease, based on various features or symptoms, is a complex endeavor. This paper endeavors to leverage diverse data mining technologies to aid in the diagnosis of the disease in question.[4]

[Enhanced Analysis of Heart Disease Prediction Systems Utilizing Data Mining Classification Techniques Chaitrali S. Dangare, Sulabha S. Apte – 2016, The healthcare industry often has a lot of information, but not all of it is used efficiently to find hidden patterns and help make informed decisions. This research utilizes advanced data mining methods to extract insights from databases, with a specific emphasis on forecasting heart disease. The study assesses forecasting systems that include a more extensive range of input characteristics. More precisely, the system uses 13 attributes including gender, blood pressure, and cholesterol levels to determine the likelihood of a patient getting heart disease. In this research, the prediction model has been updated to include two more factors: obesity and smoking. The classification methods investigated in data mining are Decision Trees, Naive Bayes, and Neural Networks, all utilized on a heart disease dataset. The accuracy is used to compare the performance of these techniques. The results show that Neural Networks have an accuracy rate of 100%, Decision Trees have an accuracy rate of 99.62%, and Naive Bayes has an accuracy rate of 90.74%. Our evaluation has determined that out of the three classification models, Neural Networks exhibit the best accuracy in forecasting heart disease.[5]

## 3. PROPOSED METHODOLOGY

In today's landscape, there is a growing focus on the development of systems designed for the automatic identification of cardiac-related conditions. These systems demonstrate high predictive accuracy based on one-dimensional ECG beat signals; however, they have yet to be widely implemented as tools within healthcare institutions. Several critical factors influence the effectiveness of these methodologies, including the selection of features, extraction techniques, classification

algorithm types, and notably, the challenge posed by imbalanced data during classification, which can adversely affect the recognition accuracy of the minority class.

### *3.1 System Architecture*

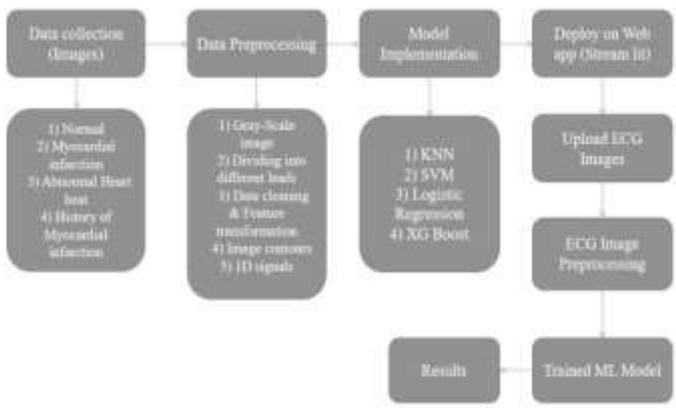


Fig.1 System architecture

**Step 1: Dataset**

A dataset, or data set, is a structured collection of data, typically presented in table format, where each column represents a specific variable and each row corresponds to a unique record. The dataset holds values for variables, such as the size or mass of an item, with each individual value referred to as a data, For this project, we are utilizing a publicly available healthcare dataset with a small number of inputs and cases. The organization of this data allows medical professionals to clearly see the relationship between traditional statistics and new machine learning methods. The small size of the dataset ensures that it can be processed efficiently on modern computers. Data collection, involving standardized methods for gathering, measuring, and analyzing information, aims to generate reliable, information-rich data for future analysis. After collection, the data undergoes a comprehensive cleaning and processing phase to improve its usefulness.

We classified our ECG images into four categories:

- We have utilized four categories for the classification of our ECG images.Normal
- Myocardial infarction
- Abnormal Heart beat
- History of Myocardial infarction

**Step 2: Data Preprocessing**

The `sklearn.preprocessing` module provides utilities to transform raw feature vectors into a suitable format for estimators. Standardized datasets often improve machine learning performance. If outliers are present, robust scalers or transformers should be used. Standardizing data, where features are centered with a mean of zero and a variance of one, is important for many models, as features with different distributions can negatively affect the learning algorithm. Rescaling features to a specific range, typically zero to one, can be achieved using `MinMaxScaler` or `MaxAbsScaler`, which are particularly effective for sparse datasets with low feature standard deviations.

**Step 3: Normalization**

Normalization is the process of modifying individual samples to achieve a normalized value. This technique is especially useful when employing quadratic forms, such as dot products or kernels, to evaluate similarity in tasks like text classification and clustering. During preprocessing, all ECG images are converted to grayscale, resized, and segmented into 12 sections to extract the 12-lead values. Grid lines are removed, and images are converted into binary format, with contour techniques isolating essential signals. Each lead's one-dimensional signal is saved in a `.csv` file, and the 12 leads are consolidated into a single CSV file with corresponding labels.

**Step 4: Deployment of Web Application**

The model file is created by training machine learning algorithms like SVM, KNN, Logistic Regression, and XGBoost. The model is deployed using the Streamlit framework.

**Step 5: Prediction**

The web application allows users to upload ECG images, which undergo preprocessing steps such as grayscaling, contour detection, Canny edge detection, and conversion to one-dimensional signals. After dimensionality reduction, the model predicts whether the ECG image shows a Normal heart, Myocardial Infarction, Abnormal Heartbeat, or History of Myocardial Infarction

## 4. IMPLEMENTATION AND RESULT

### *IMPLEMENTATION*

We tested multiple algorithms and applied an ensemble technique to stack models, enhancing performance for this classification problem. Classification was based on algorithms such as K-Nearest Neighbors (KNN), Decision Trees, and clustering. In this project, we implemented various supervised classification algorithms using CSV data: KNN, Logistic Regression, Support Vector Machine (SVM), and a Voting-Based Ensemble Classifier.

KNN Algorithm:

K-Nearest Neighbors (KNN) is a supervised machine learning algorithm used for both classification and regression. It's a non-parametric method that predicts based on the similarity between a new data point and its K nearest neighbors from the training data.

Support Vector Machines (SVM):

SVM is a versatile supervised machine learning algorithm that excels in binary classification but can also handle multi class classification. The main idea is to find the optimal hyperplane that separates data points from different classes. The best hyperplane maximizes the margin, the distance between the hyperplane and the nearest data points of each class, which helps in making better classification decisions.

Logistic Regression:

Logistic Regression is used for binary classification problems. It predicts the probability of an event by fitting data to a logistic curve. The outcome is binary, such as "yes" or "no," with features being either continuous or categorical.

XGBoost:

XGBoost (Extreme Gradient Boosting) is an efficient machine learning algorithm known for its performance in handling structured data. It builds a predictive model by sequentially combining weak learners, typically decision trees, to improve accuracy. Each new model focuses on correcting the errors of the previous models until a stopping criterion is met.

***RESULT***

A dataset, or data set, is a collection of information typically organized in a table format, where each column corresponds to a specific variable and each row represents a distinct record. For this project, we utilized a publicly accessible healthcare dataset that contains a small number of inputs and cases. This organization allows medical professionals to easily establish connections between traditional statistical methods and emerging machine learning techniques. Additionally, the compact size of the dataset facilitates quick processing on most modern computers.

Data collection involves gathering, quantifying, and analyzing various types of information using approved methodologies. The primary goal is to obtain reliable and informative data that can be analyzed for informed decision-making. Once the data is collected, it undergoes thorough cleansing and processing to enhance its quality and relevance, ensuring that it is valuable for subsequent analyses and applications in healthcare decision-making.

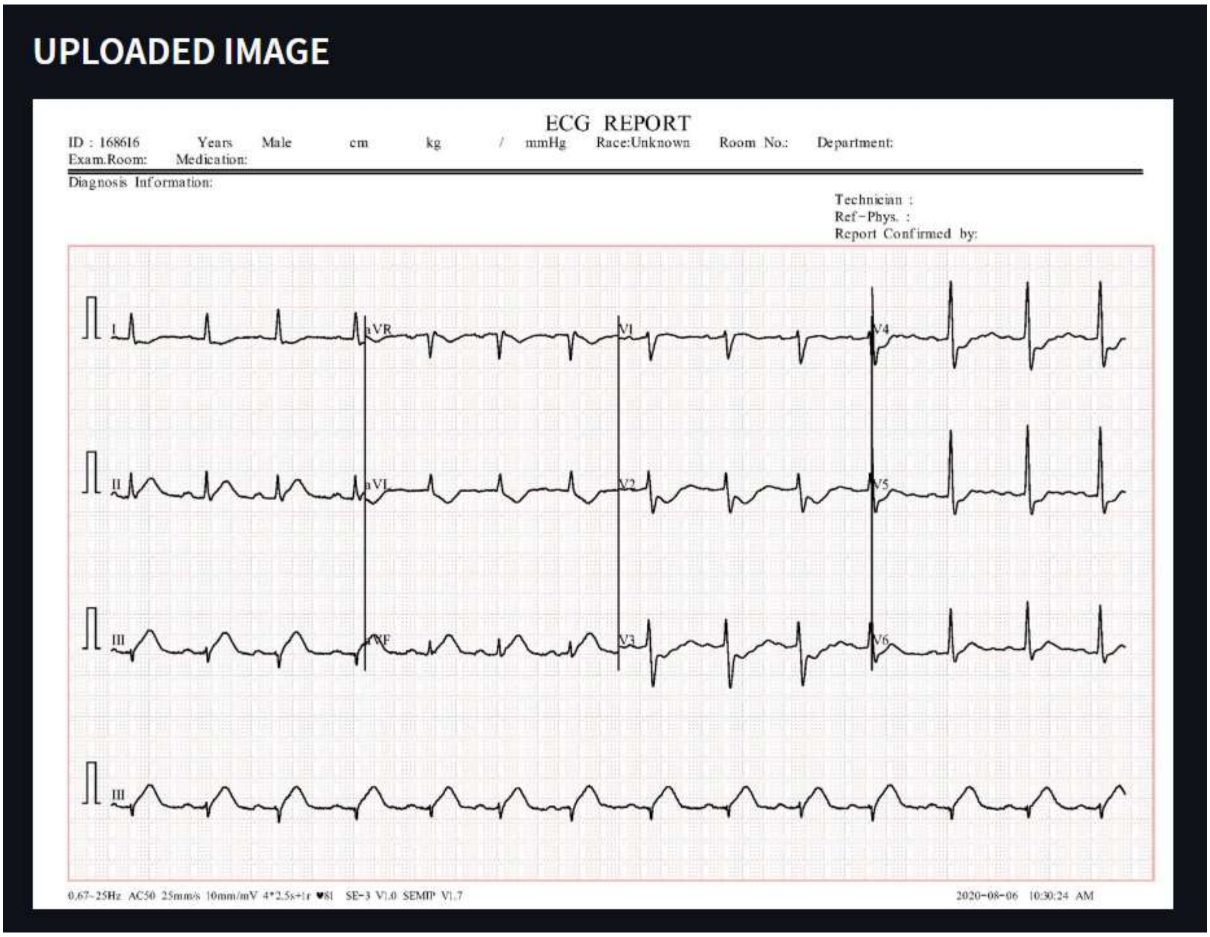


Fig 1. Uploading ECG image

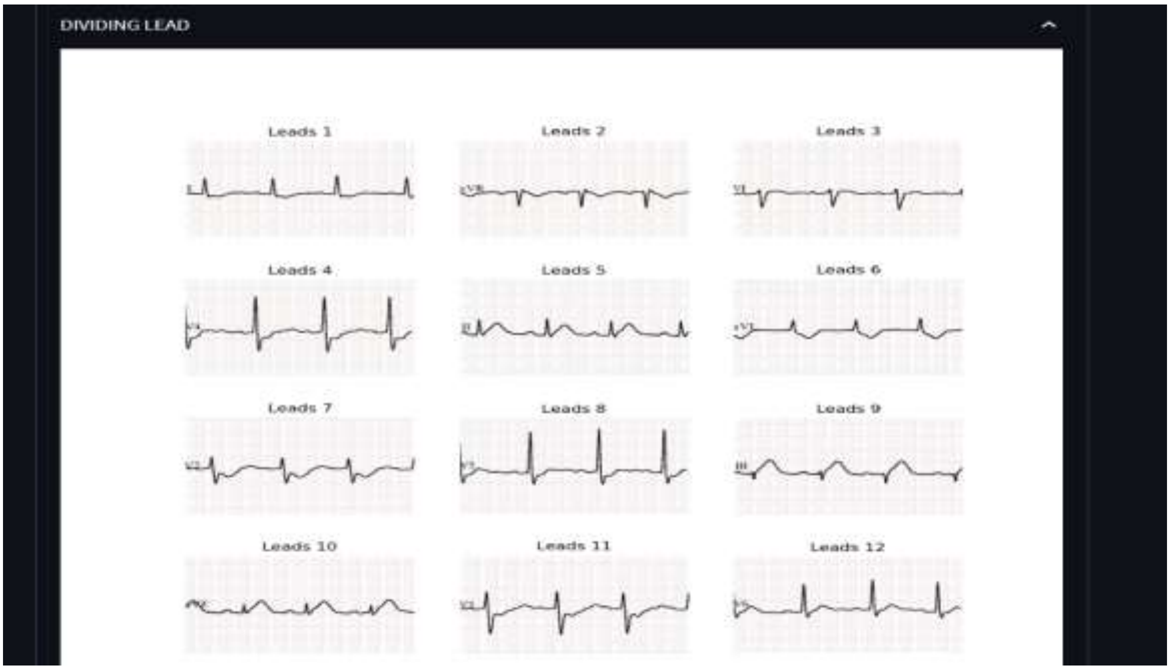


**Fig 2. Converting ECG image to Different Leads**

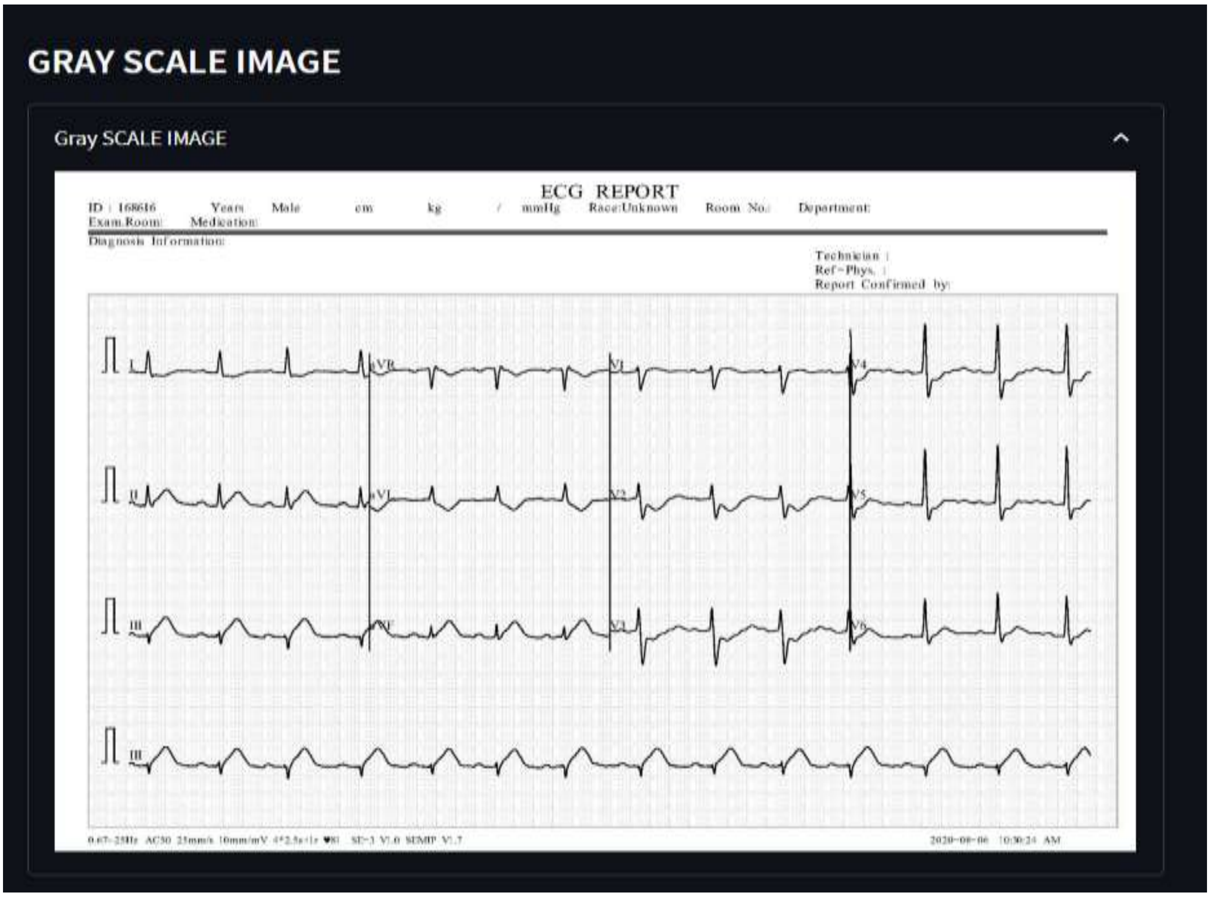


**Fig 3. Converting Colored ECG image to Gray Scale(Black and white).**

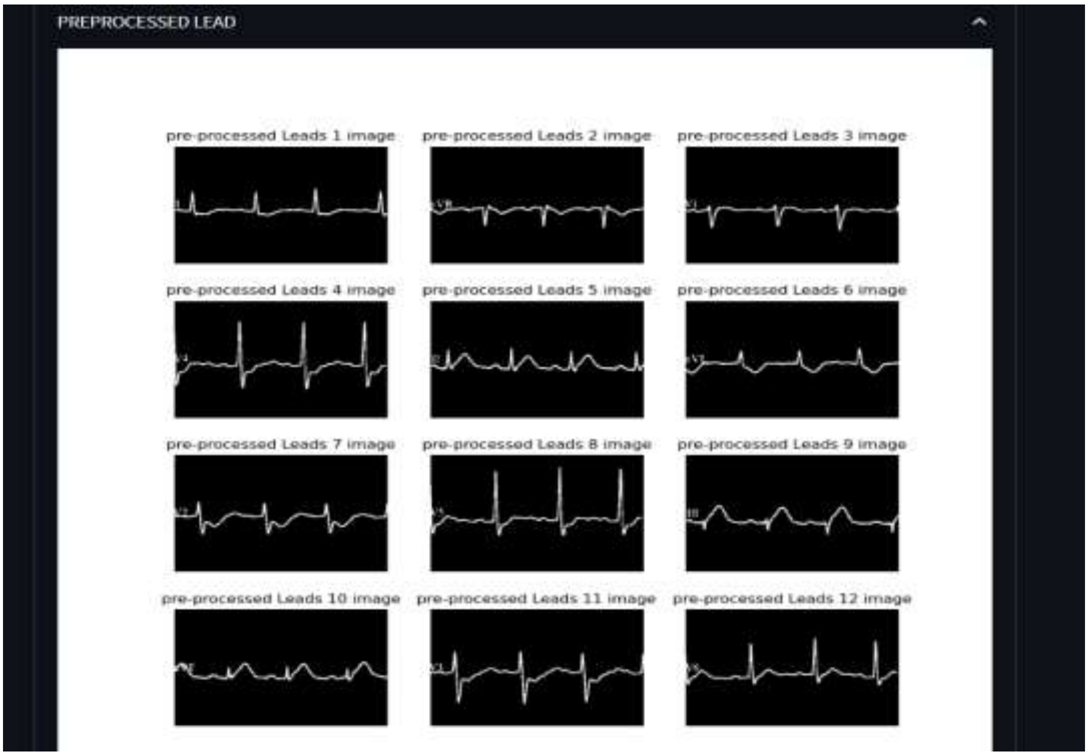


**Fig 4. PreProcessing Lead Image for Feature Extraction**

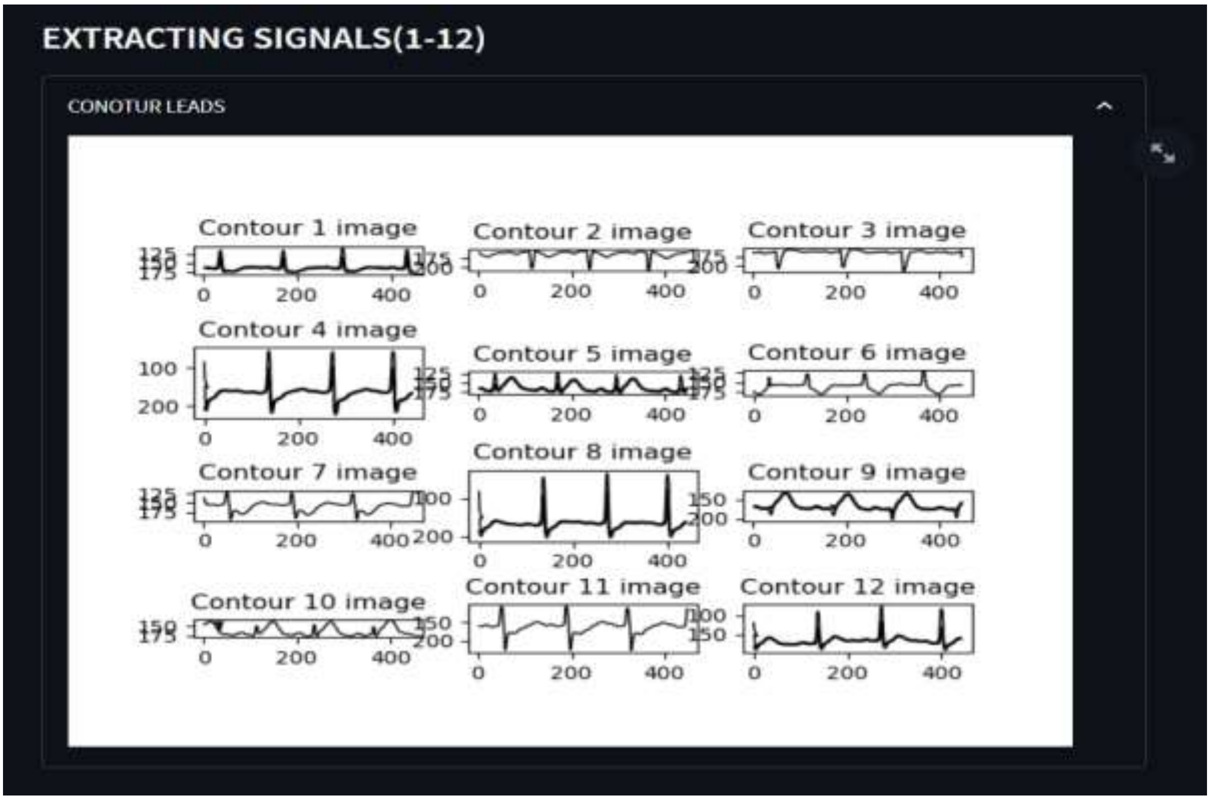


**Fig 5. Extracting Features from ECG Gray Scale Image.**

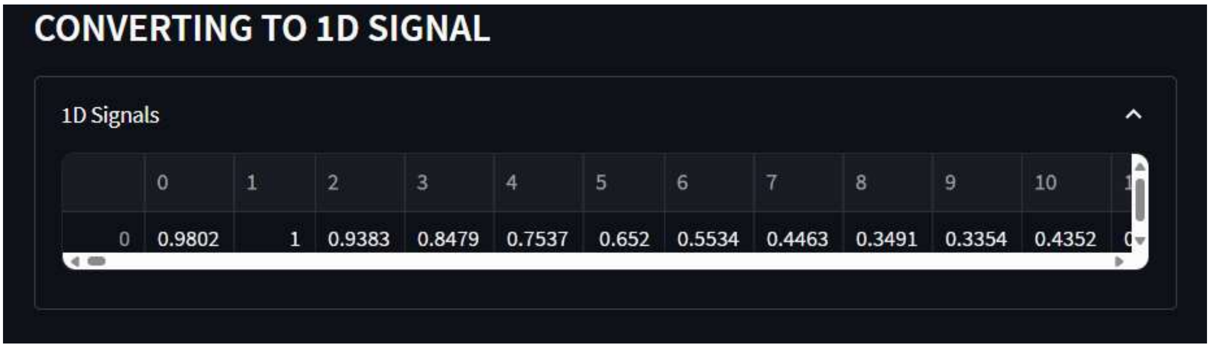


**Fig 6. Converting Features to 1 Dimensional Array.**


**Fig 7. Performing Dimensionality Reduction**

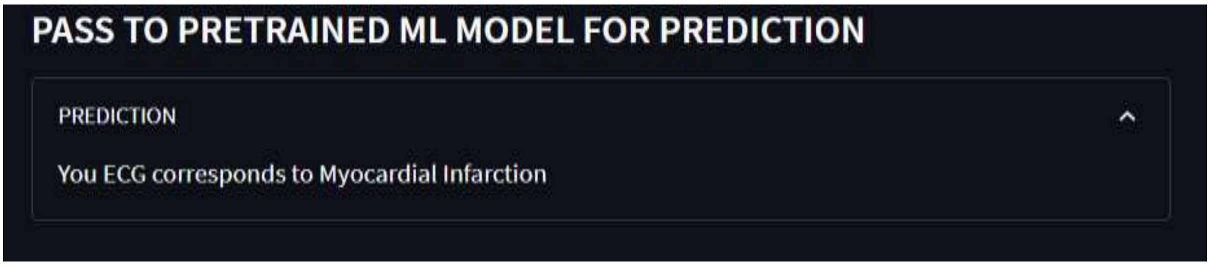


**Fig 8. Predicted Result.**

## 5. CONCLUSIONS AND FUTURE WORK

Initially, users upload ECG images to our web application, where various image processing techniques such as rgb2gray conversion, denoising, Gaussian filtering, thresholding, and contouring are employed to isolate signals and eliminate grid lines. The essential waves (P, QRS, T) are extracted and analyzed using a pre-trained model, enabling faster and more accurate predictions for heart conditions.

While this research can be expanded to cover various heart diseases, it is crucial to remember that ECG should complement other clinical assessments for a comprehensive diagnosis. Future efforts will focus on integrating multi-lead ECG data, implementing advanced deep learning architectures like Transformers, and utilizing federated learning for privacy-conscious collaborations. Emphasis will also be placed on real-time monitoring systems and explainable AI methodologies to enhance interpretability and efficacy, ensuring adherence to medical safety regulations through clinical trials and international collaboration.

**REFERENCES**

[1]. L. Wang, H. Zhang, K. C. Wong, H. Liu, and P. Shi, "Physiological-model-constrained noninvasive reconstruction of volumetric myocardial transmembrane potentials," IEEE Transactions on Bio Medical Engineering, vol. 57, no. 2, pp. 296–315, 2019

[2]. Q. Zhang, D. Zhou, and X. Zeng, "HeartID: a multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications," IEEE Access, vol. 5, pp. 11805–11816, 2017.

[3]. U. R. Acharya, S. L. Oh, Y. Hagiwara et al., "A deep convolutional neural network model to classify heartbeats," Computers in Biology and Medicine, vol. 89, pp. 389–396, 2017.

[4]. C. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds," in Proceedings of the 2016 Computing in Cardiology Conference (CinC), Vancouver, Canada, September 2016.

[5]. Baitharu, Tapas Ranjan, and Subhendu Kumar Pani. "Analysis of data mining techniques for healthcare decision support system using liver disorder dataset." Procedia Computer Science 85 (2016): 862-870.

[6]. Gavhane, Aditi, Gouthami Kokkula, Isha Pandya, and Kailas Devadkar. "Prediction of heart disease using machine learning." In 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 1275-1278. IEEE, 2018.

[7]. Singh, Smriti Mukesh, and Dinesh B. Hanchate. "Improving disease prediction by machine learning." Int J Res Eng Technol 5, no. 6 (2018): 1542-1548.

[8]. Taneja, Abhishek. "Heart disease prediction system using data mining techniques." Oriental Journal of Computer science and technology 6, no. 4 (2013): 457-466.

[9]. Dangare, Chaitrali S., and Sulabha S. Apte. "Improved study of heart disease prediction system using data mining classification techniques." International Journal of Computer Applications 47, no. 10 (2012): 44-48.

[10]. Thomas, J., and R. Theresa Princy. "Human heart disease prediction system using data mining techniques." In 2016 international conference on circuit, power and computing technologies (ICCPCT), pp. 1-5. IEEE, 2016.

[11]. Kaur, Beant, and Williamjeet Singh. "Review on heart disease prediction system using data mining techniques." International journal on recent and innovation trends in computing and communication 2, no. 10 (2014): 3003-3008.

[12]. Meyer, Alexander, Dina Zverinski, Boris Pfahringer, Jörg Kempfert, Titus Kuehne, Simon H. Sündermann, Christof Stamm, Thomas Hofmann, Volkmar Falk, and Carsten Eickhoff. "Machine learning for real-time prediction of complications in critical care: a retrospective study." The Lancet Respiratory Medicine 6, no. 12 (2018): 905-914.

[13]. Rajkomar, Alvin, Michaela Hardt, Michael D. Howell, Greg Corrado, and Marshall H. Chin. "Ensuring fairness in machine learning to advance health equity." Annals of internal medicine 169, no. 12 (2018): 866-872.

[14]. Rajamhoana, S. P., C. Akalya Devi, K. Umamaheswari, R. Kiruba, K. Karunya, and R. Deepika. "Analysis of neural networks based heart disease prediction system." In 2018 11th International Conference on Human System Interaction (HSI), pp. 233-239. IEEE, 2018.

[15]. Ramalingam, V. V., Ayantan Dandapath, and M. Karthik Raja. "Heart disease prediction using machine learning techniques: a survey." International Journal of Engineering & Technology 7, no. 2.8 (2018): 684-687.

[16]. Kohli, Pahulpreet Singh, and Shriya Arora. "Application of machine learning in disease prediction." In 2018 4th International conference on computing communication and automation (ICCCA), pp. 1-4. IEEE, 2018.

[17]. Marimuthu, M., M. Abinaya, K. S. Hariesh, K. Madhankumar, and V. Pavithra. "A review on heart disease prediction using machine learning and data analytics approach." International Journal of Computer Applications 181, no. 18 (2018): 20-25