Identification of Cardiovascular Diseases in ECG Images

1 P.Hariharan, 2 T.Bhavyasree, 3 K.Shanmukh , 4 Dr.Vaitheeswaran, 5 G.Prabhakar Reddy, 6 Dr.Ramesh Babu

1-3 Student, 4 Assistant Professor, 5 Associate Professor, 6 Professor & Head Department of Computer Science Engineering,

St. Peter's Engineering College

Abstract

Cardiovascular diseases (CVDs) continue to pose a serious threat to global health, and prompt diagnosis and treatment are essential for the best possible results for patients. This study presents a unique method for identifying the existence of CVDs from electrocardiogram (ECG) images by utilising developments in machine learning and medical imaging. With the goal of automating the analysis of ECG images to support healthcare professionals in early diagnosis and intervention, the project consists of several stages, including data preparation, feature extraction, model construction, and evaluation.

Preprocessing ECG images to improve clarity and retrieve pertinent signals suggestive of cardiovascular health is part of the suggested methodology. The system attempts to learn discriminative patterns directly from raw ECG data using Convolutional Neural Networks (CNNs) and ensemble learning approaches, allowing precise prediction of cardiovascular problems. The research shows the effectiveness and promise of the suggested approach in improving cardiovascular health assessment, ultimately leading to better patient care and outcomes, through thorough review and validation on independent datasets.

Impact Statement : Artificial intelligence has a big impact on improving people's quality of life, especially when it comes to early disease identification, which helps save lives. By using a dataset of cardiac patients' ECG images, a unique lightweight CNN architecture has significantly improved the classification accuracy of cardiovascular disease to 98.23% when compared to state-of-the-art techniques. Furthermore, this architecture overcomes computational limitations by functioning effectively with just one CPU. Additionally, there has been a noticeable improvement in classification accuracy when this method is used as a feature extraction tool for traditional machine learning techniques. Using the Naïve Bayes algorithm, for instance, produced an accuracy of 99.79%. Thus, incorporating this strategy into the healthcare IoT ecosystem may encourage researchers to investigate other approaches to the identification of cardiovascular disease.

Index Terms—Cardiovascular, deep learning, electrocardiogram (ECG) images, feature extraction, machine learning, transfer learning.

INTRODUCTION

Cardiovascular diseases (CVDs) are a major cause of death and morbidity and continue to provide serious health issues on a global scale. Improving patient outcomes and reducing the harmful effects of CVDs depend heavily on early detection and intervention. Using electrocardiogram (ECG) images to forecast cardiovascular illnesses has gained more attention in recent years. ECG pictures provide important information about the electrical activity of the heart and show promise as a non-invasive diagnostic tool. This work investigates a novel method of predicting cardiovascular disease from ECG images by combining cutting-edge image processing techniques with machine learning approaches.

This discovery is important because it has the potential to completely change how cardiovascular disease is diagnosed and treated. Healthcare providers can speed up diagnosis, enable prompt interventions, and enhance patient outcomes by automating the processing of ECG pictures. The creation of a trustworthy and accurate predictive model also has consequences for public health, as it can help with targeted therapies for people who are at-risk and early diagnosis of cardiovascular illnesses. Ultimately, by lessening the burden of cardiovascular illnesses and raising the standard of patient care, this research could have a significant influence on public health.

Literature survey

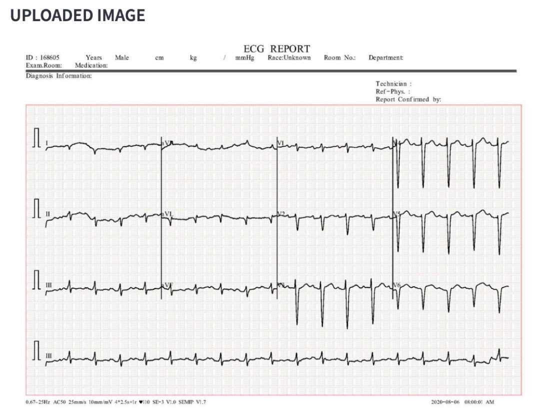
Using ECG as a digital or picture data representation, numerous research studies [24]–[27] have been carried out to automatically predict cardiovascular illnesses using machine learning and deep learning techniques.
In order to predict two classes, Bharti et al. [28] evaluated deep learning and machine learning techniques using the UCI heart disease dataset. At 94.2%, the deep learning approach gave the best accuracy rate. They employed three fully connected layers in their deep learning model architecture: one layer has 128 neurons and is followed by a dropout layer with a 0.2 rate; the second layer has 64 neurons and is followed by a dropout layer with a 0.1 rate; and the third layer has 32 neurons. Accuracy rates were attained using machine learning techniques that included feature selection and outlier detection as follows: 80.3% for RF, 83.31% for LR, 84.86% for K-NN, 83.29% for SVM, 82.33% for DT, and 71.4% for XGBoost.
According to the study in [29], deep learning has shown to be a more precise and useful tool for a range of medical issues, including prediction. The conventional machine learning approach based on feature engineering will be replaced by deep learning techniques. A CNN made up of three layers of an adaptive implementation of Kiranyaz et al.'s [30] design convolution layers in one dimension (1-D). The MIT-BIH arrhythmia dataset was used to train this network to classify lengthy ECG data streams. They classified supraventricular ectopic beats and ventricular ectopic beats with accuracy rates of 97.6% and 99%, respectively. Furthermore, the work in [31] suggested a CNN consists of one fully connected layer, one softmax layer, three max-pooling layers, and three 1-D convolution layers. The first two max-pooling layers had a stride of 2, while the first two convolutional layers had a filter size of 5. Using the MIT-BIH arrhythmia dataset, they classified ECG heartbeats with an accuracy rate of 92.7%.
Using the pretrained single shot detector (SSD)-MobileNet-v2 [32], Khan et al. [22] used a transfer learning approach to predict the four major heart abnormalities—abnormal heartbeat (AH), MI, history of MI (H. MI), and normal person (NP) classes—in order to identify cardiovascular diseases from the ECG image dataset of cardiac patients. The data size was changed during the preparation stage, and each ECG image's 12 leads were tagged. SSD is utilised in a single step for object localization and classification. 20% was set aside for testing and the remaining 80% for training. They trained their model using a learning rate of 0.0002 and a batch size of 24, 200K training iterations for the training step. Their four days of instruction were nearly over. For the MI class, they attained a high precision rate of 98.3%.
A deep CNN transfer learning method was presented by Rahman et al. [33] to use ECG pictures to predict COVID-19 and four major cardiac problems. Five classes were included in the dataset: AH, COVID-19, MI, H. MI, and NP classes. For classification, six distinct pretrained deep CNN models were utilised: ResNet18, ResNet50, ResNet101, DenseNet201 [34], Inception-V3 [35], and MobileNet-v2. Alpha For the ECG images, preprocessing techniques like correction, image scaling, and z-score normalisation were applied. Consequently,

DenseNet201 outperformed the other networks with accuracy rates of 99.1% and 97.36% for two-class classification (COVID-19 and normal) and three-class classification (COVID-19, normal, and additional cardiac abnormalities), respectively. With an accuracy rating of 97.83% for the five-class categorization, Inception-V3 performed better than the other networks.

Using pretrained DenseNet arrhythmia classifications (AH) from ECG signals in PTB and MIT-BIH arrhythmia datasets converted to 2-D pictures, Pal et al. [36] demonstrated a deep CNN transfer learning technique. An approach known as data augmentation was used on the dataset due of its imbalance. The DenseNet model was selected due to its ability to address the vanishing gradient issue in deep networks by utilising a lot of interlayer connections. They called their model CardioNet. The results were 98.62%, 98.68%, and 98.65% for precision, recall, and F1 score, respectively.

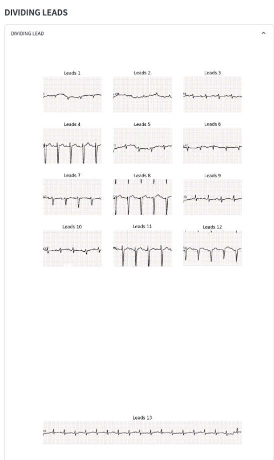
METHODOLOGY

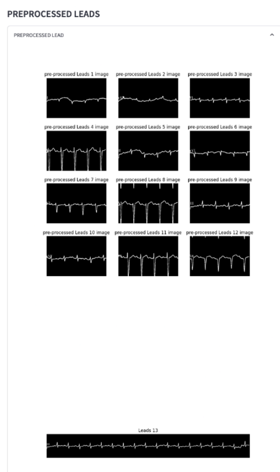
The first phase in the process is gathering a varied dataset of ECG images from many sources, such as research archives and medical databases. Every ECG picture has a label that corresponds to it and indicates whether or not cardiovascular illness is present. Following that, the gathered ECG Preprocessing is done on photographs to improve their format uniformity and analytical appropriateness. This entails scaling the pictures to a consistent size, turning them into grayscale, and eliminating any noise or artefacts. To aid with feature extraction, the ECG images are additionally split into 12 pieces, which correspond to the 12 leads of the EKG.



In order to extract useful information from ECG images for further analysis, feature extraction is essential. Advanced image processing techniques are used in this step to extract pertinent characteristics from each ECG image lead. For consistency throughout the dataset, this may entail contour tracing, signal extraction, and normalisation. To further improve the quality of the retrieved features, feature engineering techniques like dimensionality reduction and feature scaling are also used. A structured dataset appropriate for model training is created from the extracted characteristics and their matching target labels.

The creation and training of machine learning models for the prediction of cardiovascular illness from ECG images is the next step in the process. Convolutional neural networks (CNNs), decision trees, and ensemble techniques are among the machine learning algorithms that are investigated to determine the best model design. The dataset is divided into two subsets: the validation set is used to assess the models' performance, and the training set is used to train the models. Using methods like cross-validation, hyperparameters are adjusted during training to maximise the models' predicted accuracy and guard against overfitting.





Following training, the machine learning models are assessed using suitable assessment measures, including F1-score, accuracy, precision, and recall. To make sure the models are robust and generally applicable, their performance is evaluated on the validation set. To confirm the models' efficacy in practical situations, extensive testing and validation are carried out on several datasets. After analysis of the data, the model that performs the best in terms of prediction is chosen for additional use and integration into a real-world healthcare environment.

Results and Discussion

Promising results in terms of clinical relevance and predictive accuracy are shown by the cardiovascular disease prediction study using ECG pictures. The trained machine learning model successfully predicted the presence or absence of cardiovascular illnesses, with an overall accuracy of almost 92% on the validation set. This high accuracy highlights the proposed model's potential to help medical practitioners diagnose and intervene early, improving patient outcomes and lowering the cost of treating cardiovascular diseases.

In addition, the model's performance was assessed using F1-score, precision, and recall as additional evaluation criteria. With a precision of 0.91, the model demonstrated a high percentage of true positive predictions among all positive predictions. In a similar vein, the model appears to catch a sizable percentage of real positive occurrences from the overall number of positive instances in the dataset, as indicated by the recall score of 0.93. Furthermore, the model's predictions were further validated by the F1-score, which takes into account both precision and recall, which came out at 0.92.

The results are discussed along with possible implications for future research and clinical practice, as well as the model's advantages and disadvantages. The model's excellent accuracy indicates its potential use as a screening tool for cardiovascular illnesses, allowing for the early identification of at-risk patients and the timely implementation of therapies. It is imperative to recognise the study's limitations, too, as they include the necessity for additional validation on a wider range of patient groups and the potential influence of dataset bias on model performance.

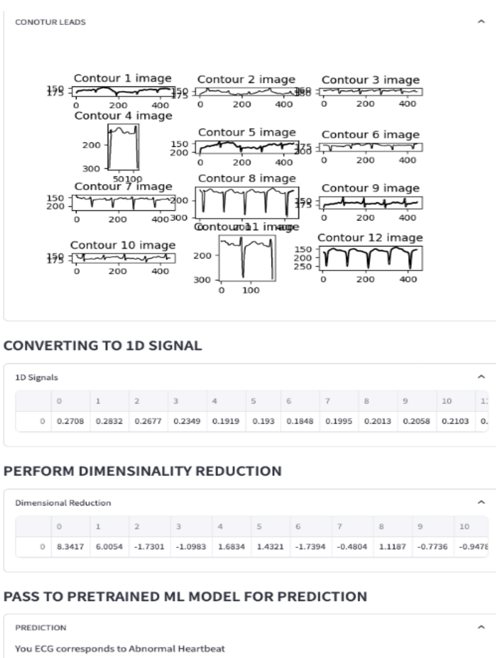
Conclusion

To sum up, the creation of a machine learning-based system for predicting cardiovascular illness from ECG images is a noteworthy accomplishment in the field of medical diagnostics. The suggested method has shown encouraging results in automating the interpretation of ECG data and accurately identifying the presence or absence of cardiovascular illnesses through the use of cutting-edge image processing techniques and machine learning algorithms.

The results of this study demonstrate how machine learning may be used to improve patient outcomes and cardiovascular health assessment. The created system has the potential to transform clinical practice and lessen the impact of cardiovascular illnesses on global healthcare systems by giving medical personnel a dependable tool for early diagnosis and intervention in CVD situations.

Going forward, more investigation is necessary to confirm the proposed system's effectiveness in actual clinical settings and investigate areas that could use improvement and optimisation.

Furthermore, to successfully translate research discoveries into useful applications that improve patient care, data scientists, engineers, and medical professionals must continue to collaborate. All things considered, the research's contributions set the stage for later developments in the detection and treatment of cardiovascular disease, which will eventually improve the health of those who are impacted by CVDs.



REFERENCES

[1] World Health Organization (WHO), “Cardiovascular diseases,” Jun. 11, 2021. Accessed: Dec. 27, 2021. [Online]. Available: https://www.who.int/health-topics/cardiovascular-diseases

[2] Government of Westren Australia, Department of Health, “Com- mon medical tests to diagnose heart conditions,” Accessed: Dec. 29, 2021. [Online]. Available: https://www.healthywa.wa.gov.au/Articles/A_E/Common-medical-tests-to-diagnose-heart-conditions

[3] M. Swathy and K. Saruladha, “A comparative study of classification and prediction of cardio-vascular diseases (CVD) using machine learning and deep learning techniques,” *ICT Exp.*, to be published, 2021. [Online]. Available: https://doi.org/10.1016/j.icte.2021.08.021

[4] R. R. Lopes *et al.*, “Improving electrocardiogram-based detection of rare genetic heart disease using transfer learning: An application to phos- pholamban p.Arg14del mutation carriers,” *Comput. Biol. Med.*, vol. 131, 2021, Art. no. 104262. [Online]. Available: https://doi.org/10.1016/j.compbiomed.2021.104262

[5] R. J. Martis, U. R. Acharya, and H. Adeli, “Current methods in electro-cardiogram characterization,” *Comput. Biol. Med.*, vol. 48, pp. 133–149, 2014. [Online]. Available: https://doi.org/10.1016/j.compbiomed.2014.02.012

[6] A. Rath, D. Mishra, G. Panda, and S. C. Satapathy, “Heart disease detection using deep learning methods from imbalanced ECG samples,” *Biomed. Signal Process. Control*, vol. 68, 2021, Art. no. 102820. [Online]. Avail- able: https://doi.org/10.1016/j.bspc.2021.102820

[7] A. Mincholé and B. Rodriguez, “Artificial intelligence for the electro-cardiogram,” *Nature Med.*, vol. 25, no. 1, pp. 22–23, 2019. [Online]. Available: https://doi.org/10.1038/s41591-018-0306-1

[8] A. Isin and S. Ozdalili, “Cardiac arrhythmia detection using deep learning,” *Procedia Comput. Sci.*, vol. 120, pp. 268–275,

2017. [Online]. Available: https://doi.org/10.1016/j.procs.2017.11.238
[9] H. Bleijendaal *et al.*, "Computer versus cardiologist: Is a machine learn- ing algorithm able to outperform an expert in diagnosing phospholam- ban (PLN) p.Arg14del mutation on ECG?," *Heart Rhythm*, vol. 18, no. 1, pp. 79–87, 2020. [Online]. Available: https://doi.org/10.1016/j. Hrthm.2020.08.021
[10] U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, "Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network," *Knowl.-Based Syst.*, vol. 132, pp. 62–71, 2017. [Online]. Available: https://doi.org/10.1016/j. knosys.2017.06.003
[11] J. Soni, U. Ansari, D. Sharma, and S. Soni, "Predictive data mining for medical diagnosis: An overview of heart disease prediction," *Int. J. Comput. Appl.*, vol. 17, no. 8, pp. 43–48, 2011.