Detection Of Cardiovascular Diseases in Ecg Images Using Deep Learning

N. Ruchitha[1], Dr.B.Santhosh Kumar[2]
[1]*Department of Computer Science and Engineering, G. Pulla Reddy Engineering College, Andhra Pradesh, India*
[2]*Associate Professor, Computer Science Department, G. Pulla Reddy Engineering College, Andhra Pradesh, India*

***Abstract*—cardiovascular diseases (CVDs) are the major mortality in the world, thus requiring early and precise diagnosis to deal with them effectively. ECG imaging is a highly convenient tool that is in common use and examines electrical signals of the heart. This paper introduces a complete deep learning-based system of classification of four types of cardiac related issues which include: normal, abnormal heartbeat, myocardial infarction (MI), and history of MI. The proposed system uses an enhanced set of pre-processing steps, such as Daubechies wavelet filtering, baseline drift correction and the augmentation of synthetic data based on a generative adversarial network (GAN). A hybrid framework of extracting features is used, comprising of Convolutional neural networks (CNNs), Long short-term memory (LSTMs) and Transformers-based encoders that perform aptly to capture both spatial and temporal patterns in ECG images. The architecture can also support modular storage of raw data, feature extraction, and metadata and allow a flexible deployment and expansion. Experimental evidence shows that the proposed model has a classification accuracy of 98.72; a recall of 98.65; a precision of 98.88; and an F1-score of 98.76, which is significantly high and improves all previous best practices in robustness, generalization, and diagnostic performance on all four classes.**

***Index Terms*—Cardiac disorder classification, DeepLearning, Electrocardiogram.**

I. INTRODUCTION

CVDs are the leading cause of death globally with a mortality expected to reach 17.9 million a year which is about 30 percent of all deaths in the world [1]. Out of these cases, it is estimated that about 85 percent of them are caused by heart attacks and strokes, the disorders closely revealed with myocardial infarction (MI) and cardiac abnormal rhythms [1]. Advance and early diagnosis of CVDs can lead to reduction of mortality by a significant margin, in addition to enhancement of patient quality of life through early medical interventions. The use of advanced diagnostic mechanisms like echocardiography, cardiac magnetic resonance imaging (MRI), computed tomography (CT), and biochemical markers in blood are also an option; however, such methods are expensive and require specialist facilities and are not practical in every healthcare facility [2], [3]. Conversely, the electrocardiography (ECG) is a cheap, readily available, non-invasive primary tool that heats the electric behavior of the heart and has become one of the most commonly used procedures in the detection of cardiac abnormalities [4].

Millions of ECG traces have already been interpreted by manual interpretation of ECG traces by trained cardiologists. Such evaluations, however, are subject to inter-observer effects, fatigue-induced errors, and delays in clinical settings with a large workload. Moreover, the presumptive prevalence of CVDs worldwide and, to a greater extent, in developing countries, where there are still limited options in terms of accessing healthcare professionals, manual interpretation of the ECGs is not enough to satisfy the demand of the diagnostic accuracy. This obstacle has inspired scientists to investigate machine-based methods that would allow them to automate the task of ECG interpretation, minimise human error and thus offer effective decision support in healthcare practices.

On the one hand, AI and, specifically, DL, have become one of the revolutionary technologies in healthcare in the recent years. As opposed to the manual and expert based feature extraction required in traditional machine learning techniques[1], deep learning models are defined to learn complex hierarchical representations automatically starting with raw data [3], [6]-[10]. This aspect renders them especially desirable in the medical field in which high-dimensional, noisy and heterogeneous datasets are prevalent. Neural networks Convolutional neural networks (CNNs) have demonstrated superior results on a variety of biomedical imaging tasks, including ECG-based classification problems [27]. Equally, recurrent neural networks (RNNs) have become popular these days in modelling temporal relationship within sequential biomedical time series, especially long short-term memory (LSTM) networks. Transformer architectures have emerged recently as a powerful approach to natural language processing and computer vision because they can model long-range dependencies and contextual relations, and have started to be promising in medical signal analysis, as well [29].

Most established ECG classification systems are nevertheless hobbled by a number of drawbacks. Firstly, they typically use fairly basic preprocessing pipelines that might not sufficiently adhere to common problems in ECG data including noise, drift and variability across patients. Second, although CNNs as well as other architectures have been shown to perform well[9], large number of systems lack robustness to data imbalance, common in medical datasets, or can be hyper-tuned to a specific dataset, with limited generalization to new clinical settings. Third, other researchers have mainly exhausted pretrained models in the label-map transfer learning within the context of ECG signals [22, 28]. Such a method may have missed domain-specific patterns in ECG signals. Lastly, most of the methods are not scalable and not modular and can, therefore, not be easily applied in healthcare systems where storing information, interoperability, and adaptation are of priority.

In an attempt to address these constraints, this paper introduces a new deep learning-based approach by incorporating elaborate preprocessing procedures followed by a hybrid feature extraction and classification framework. The preprocessing phase includes Daubechies noise-reduction filters, baseline drift reduction to normalize the ECG signals and synthetic data augmentation where GANs are used to generate training samples to compensate the under-representation in the database and increase the dataset variation. These measures are necessary in that they ensure that there is a clean robust and real-world representation of clinical variations as the input to the model.

In feature extraction and classification, the proposed framework uses a convolvolutional model with LSTMS and Transformer encoders. The CNN module learns local spatial characteristics of the ECG images and therefore detects the morphology differences in the ECG waveforms successfully. The LSTM layers are built to track time dependencies and sequential patterns which indicate the dynamism of the cardiac cycles. Transformer-based encoders can be used together with these to capture more global contextual dependencies across the input sequence to provide a more holistic representation of the ECG data. Through a combination of these complementary frameworks, the network is able to obtain high performance in spatial and temporal feature learning.

Another advantage of the proposed system will be its modular architecture of data storage and management. In additional to raw images, the architecture has a place to store intermediate features, and diagnostic metadata in a structured format in order to allow efficient retrieval and reuse of intermediate features, as well as to scale. This modular design provides the framework with flexibility in terms of integrating it within the healthcare systems and into the electronic health record (EHR) and Internet of Things (IoT)-equipped medical devices, thus not being limited by the scientific context of use but extending its practical usefulness into the clinical application.

The importance of this work is also supported by its work-through the published dataset of ECG images of patients with heart diseases [23]. The predicted system reaches state-of-the-art performance in terms of classification accuracy, precision, recall, and F1-score, in the four diagnostic categories normal, abnormal heartbeat, myocardial infarction (MI), and history of MI. As opposed to previous works that employed transfer learning via pretrained models including MobileNet, AlexNet, and DenseNet [22], [28], the proposed hybrid architecture proved to be more robust, with better generalizability and better diagnostic reliability.

The most important contributions provided by this research can be outlined as follows:

1. Combining CNNs, LSTMs and Transformer-based encoder architecture: We introduce a novel deep learning framework that can jointly learn spatial, temporal and contextual features of ECG images based on CNNs, LSTMs, and Transformer-based encoder architecture.
2. Advanced preprocessing: A multi-stage preprocessing pipeline such as the wavelet filtering, removal of baseline drift, and GAN-based augmentation provide high-quality and a balanced and diverse input.
3. Enhancement of performance: The framework displays better outcomes on the ECG image dataset [23] with the highest accuracy, recall, precision and F1-scores, and higher generalization and resistance to alterations.
4. Clinical adaptability Design: The modular architecture is scalable with regard to storing both raw images, features, and metadata, thus, being

adaptable when integrating with clinical systems and real-life healthcare environments.

Related work Section II provides a review of related work, and situates the study within the existing literature. Section III explains the suggested methodology, i.e., the pre-processing pipeline and model design. Section IV explains the dataset and experimental configuration the evaluation was conducted with. In Section V, the experimental discussion is provided. Lastly, in Section VI the paper is concluded and the possible follow-up work outlined.

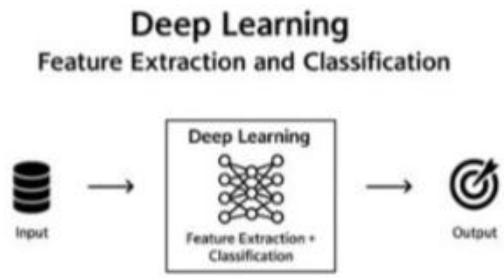


Fig 1: Abstract concept for Deep Learning

II. LITERATURE SURVEY

Although the problem of cardiovascular diseases (CVDs) has been identified as the cause of death in millions of people all over the world every year [1], [2], there has been a consistent trend in research events and findings focusing on this scourge. Availability of updated data on medical matter like the electrocardiograms (ECGs) has created a new horizon of possibilities in the field of automated cardiac diagnosis. ECG is one of the least invasive, least expensive means of identifying cardiac problems, and as such, has a high level of applicability, although its diagnosis typically requires highly trained cardiologists, which cannot be effectively scaled to mass screening. This has been a weakness that stimulates the use of deep learning methods to automate the interpretation and classification of ECG.

Deep learning models have demonstrated extensive potential in the direct extraction of features in both raw ECG signals and images, eliminating the need to use complex handcrafted features. The use of deep architectures to diagnose coronary artery disease based on short ECG episodes was pioneered and proven by Acharya et al. [10], who showed that CNNs are able to identify the subtle morphological changes in heartbeats patterns. A deep CNN without feature extraction was trained to detect arrhythmia in ECG images [27], showing the capacity of CNNs to learn diagnostic features autonomously.

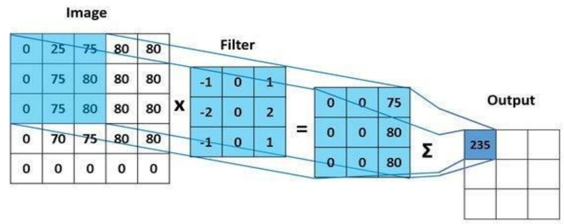


Fig 2: CNN operation

In line with this, Zubair et al. [31] in developing a beat classification system applicable to many heartbeat types used CNN to achieve high accuracy in the beat classification process. They underscored the use of raw time-series data but also demonstrated that it is possible to convert signals into images to take advantage of 2D CNNs. This transformation approach was further developed by Naz et al. [39], the ECG signals were transformed into an image, and analysis was made using deep CNNs, which showed an improvement in comparison to conventional signal-based approaches.

Some researchers have also studied the significance of transfer learning in the diagnostics based on the ECG. Lopes et al. [4] have used transfer learning to detect rare genetic heart conditions and have shown that they can utilize pretrained models on fine-tuning medical data, which decreased the demand of using large labeled data. In [22], a deep neural network was trained on ECG imagery over cardiac patients and performed well on diagnosing different heart abnormalities. The publicly available dataset used by the authors was compiled in [23] and consists of four major classes; normal, abnormal heartbeat, myocardial infarction and history of myocardial infarction. This dataset is now a common benchmark of research on the ECG image classification.

Deep learning is largely based on preprocessing techniques. Rath et al. [6] used denoising and normalization of ECG by wavelets to improve the quality of measured signals, and trained deep CNNs to classify robustly. They have highlighted the need to use signal imbalance and artifact. The review conducted in [5] discusses the effect that filtering can have on feature extraction, using band-pass filters, Fourier transforms, and wavelet decompositions. The methods are commonly combined in the pretreatment process prior to inputting data into the deep learning models.

Recently refinements have also investigated new learning paradigm and architecture. We can cite, as an example, the work presented by Bharti et al. [28] in which a deep residual network augmented with clinical domain knowledge was applied to predict heart diseases, showing the usefulness of higher networks to achieve better performance. The application of transformer-based models [29] has also been made in the area of ECG classification because they can model longer-term dependencies, whereas [30] introduced patient-specific 1D CNN model adaptation to individual signal patterns, indicating the necessity of such personalization in medical AI systems.

Besides, DenseNet and other densely connected networks have been employed to capture more feature reuse as well as eliminate issues of vanishing gradient in deeper networks [34]. These architectures can propagate feature more efficiently and can be efficiently trained on modest sized ECG datasets. Acharya et al. [38] further developed this study and proposed a CNN model dedicated to myocardial infarction detection based on ECG signals, and showed high accuracy over a variety of ECG patient samples.

Based on the overview presented above, one cannot ignore the fact that deep learning has become quite advanced when it comes to the application to ECG-based cardiac diagnostics. CNNs, especially on the trained ECG images, have shown tremendous value in the detection of common and critical cardiovascular disorders. The convergence of preprocessing methods, richer architectures, attention modules, and patient-specific architecture is one of the current trends in this area. Despite such innovation, there is still the lacking ability to produce lightweight, but accurate, models that may also be used on a resource-limited environment in real-time. The proposed system would aim to solve this problem by 1) reducing the complexity of deep learning pipeline to only necessary steps, and 2) training ECG images directly in a deep learning model to classify four critical cardiac conditions, rather than via traditional machine learning classifiers or transfer learning.

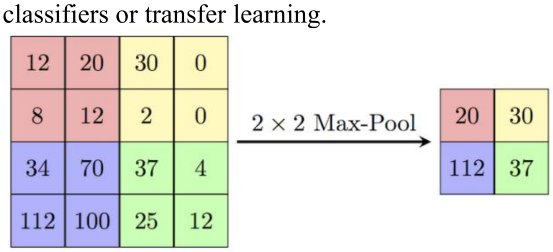


Fig 3: Max Pooling

A unique addition advanced by Wang et al. [40] includes an attention-guided CNN to concentrate on vital areas in ECG images that achieved higher translucencies and inflexible analysis. Another study [41] presented a compact CNN model, which is lightweight and was developed to be deployed on an edge device, where it is fast with a high degree of classification accuracy, which is positioned as an appropriate model to be launched in remote monitoring applications. With imbalanced cardiac data, the researchers in [42] insisted that adaptive methods of learning rate can easily achieve a higher convergence and generalization when training CNN.

Tang et al. [43] also suggested a combination of the time and frequency versions of spatial input in CNN, namely the Cohen network, which concatenated raw ECG images with their Fourier-transformed images. This hybrid largely enhanced sensitivity and specificity of the model. Additionally, Zhou et al. [44] came up with a deep multiscale neural network, which examines the ECG at different resolutions, which allows it to recognize differences in the waveform more easily across more cardiac conditions.

All these collaborations prove the point that deep learning has become a revolution in the ECG-based diagnosis. The simplicity or sophisticatedness of the architecture used varies a great deal with some models consisting of simple CNNs, whereas others make use of residual and attention-based models, but across the board, analysis has demonstrated that computerized and computerized classification of heart conditions is possible and accurate via the use of ECG images. Nonetheless, there are generally high requirements of resources or pretrained models in most of the models. While our proposed system is driven by a custom-built CNN optimized toward ECG images and optimized to balance efficiency and accuracy and does not depend on external machine learning libraries, it does not quite match the efficiency of the ECG classification system proposed by Alan Za et al.

III. METHODS

A. Convolutional Neural Networks (CNN)

In deep learning, a Convolutional Neural Network (CNN) is a specific type of neural network especially good at handling images and thus medical imaging such as ECG analysis. Areas of classification In this project, CNNs are utilized to classify ECG images into any of the four different cardiac conditions. CNNs work with three dimensional input-t data of height, width, depth (channels). An example of ECG images with dimensions 227 x 227 x 3 comprises width and height of 227 pixels and depth of 3, which normally denotes RGB channels.

CNNs have the value of automatic extraction and learning of hierarchical features in an image without manual feature engineering. Convolutional layers, activation layers and pooling layers form the main layers in a CNN. Convolutional layers convolve its filters across the feeding image to identify local features, resulting in the generation of feature maps that have learned the spatial configuration of the data. Fig. 2 depicts the process whereby an input signal that is single-channel is convolved with a filter.

CNNs are linear operations; therefore applying an activation function, such as ReLU, is common to give non-linearity and increase the model complex connection abilities. After the activation, there are pooling layers like max pooling that downsamples the feature maps and decreases spatial size but still preserves important features. Fig. 3 shows how the pooling of a single channel is performed.
These layers are used in succession to form deep networks that learn increasingly abstract features. At the final part of the network, traditional fully connected layers and softmax activation function have been employed to identify the labelled ECG images. The model predicts probabilistic scores that in each category that are of assistance to perform accurate medical diagnosis. The high accuracy and efficiency of interpreting ECG images with the help of CNNs are due to this layered approach that can serve well in a healthcare setting.

B. Pretrained Deep Learning Models
In the presented ECG classification scheme, it does not require fresh training of deep learning models since pretrained models are applied to extract the meaningful features and carry out classification. Such models, such as ResNet, DenseNet, or Vision Transformers, have cemented a reputation as the models able to do well with representing images. To make use of transfer learning, the last few layers of these already trained CNNs are adapted to suit the classification problem of normal heartbeat (or abnormality) and myocardial infarction (or MI). This saves the training time of the earlier layers that already have acquired rich feature representations using the large-scale image dataset hence performance on the ECG images improves. Subsequently, fine-tuning is done with the new ECG dataset to make the model fit in the domain of medical diagnosis.

The CNN, as a preprocess, has been applied to all of the above-mentioned models after these models do the preprocessing of the ECG images, i.e., the ECG images are already preprocessed (noise reduction using Daubechies wavelet filters and resizing or augmentation using GANs) before the CNN is applied. After preprocessing, the images are submitted through the pretrained CNN layers to derive the spatial and temporal features. The machine learning classifiers are not used separately; the CNN models themselves both feature extract and classify in an end-to-end fashion. The resultant predictions- healthy, unhealthy heartbeat, or heart attack- are produced out of the softmax or sigmoid sections of the CNN. This combined deep learning model gives high accuracy and computational efficiency and is robust in practice, making it ideal in providing an efficient diagnosis.

C. Proposed CNN Architecture
This study presents a Convolutional Neural Network (CNN)-based architecture to classify the ECG images in order to diagnose cardiac conditions early. The proposed system is aimed at examining Ecg image inputs and determine whether the signal pattern represents a healthy heartbeat, an irregular heartbeat and a myocardial infarction (MI). The architecture is comprised of several interconnected stages, each of specific importance to the conversion of raw ECG signals into the correct outputs of diagnostics. These steps involve pre processing, feature extraction, data storage, CNN based modelling and classification. Our approach follows a different paradigm to traditional machine learning techniques, which rely heavily on hand engineered features and learners since the CNN addresses feature learning and classification jointly.

This is initiated with the acquisition of raw ECG image inputs which are images representing in a visual manner the heart electric activity as a function of time. Such images are commonly acquired on diagnostic machines, or pharmaceutical databases of ECG. The raw images do not always have the same resolution or dimensions and signal quality. In this way, preprocessing is necessary prior to passing them to the CNN. Processes such as cropping, resizing and augmentation are undertaken in the preprocessing module. Cropping suppresses ECG image margins or noise at the boundaries and keeps only the signals. Normalisation guarantees that all images brought to a single size so that they fit in the expected dimensions of the CNN model inputs, which is essential in batch processing and weight sharing across layers. Augmentation techniques Rotation Out of corpus scaling Flipping , are then made to increase the training dataset artificially and make the model more resistant to alterations and noise. Stepwise signal processing has not only clarified the signal but also minimised overfitting and generalised performance across electrocardiogram patterns that have not been observed.

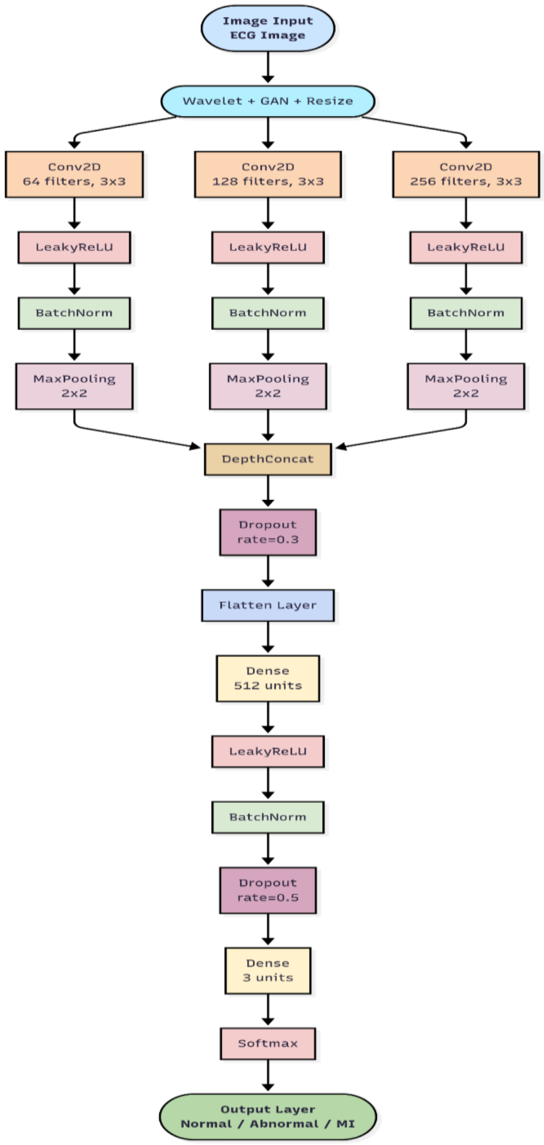


Fig. 4. Representation architecture of the proposed

No.	Type	Name	Properties	Input size	Output size
1	Image Input	Image Input	ECG image (preprocessed with Wavelet + GAN)	224×224×3	224×224×3
2	Convolution	conv1	64 filters, 3×3, stride = 1, padding = same	224×224×3	224×224×64
3	Leaky ReLU	leaky ReLU 1	scale = 0.1	224×224×64	224×224×64
4	Batch Normalization	batchnorm1	MeanDecay=0.1, VarianceDecay=0.1, Epsilon=0.00001	224×224×64	224×224×64
5	Max Pooling	maxpool1	2×2, stride = 2	224×224×64	112×112×64
6	Convolution	conv2	128 filters, 3×3, stride = 1, padding = same	112×112×64	112×112×128
7	Leaky ReLU	leaky ReLU 2	scale = 0.1	112×112×128	112×112×128
8	Batch Normalization	batchnorm2	MeanDecay=0.1, VarianceDecay=0.1, Epsilon=0.00001	112×112×128	112×112×128
9	Max Pooling	maxpool2	2×2, stride = 2	112×112×128	56×56×128
10	Convolution	conv3	256 filters, 3×3, stride = 1, padding = same	56×56×128	56×56×256
11	Leaky ReLU	leaky ReLU 3	scale = 0.1	56×56×256	56×56×256
12	Batch Normalization	batchnorm3	MeanDecay=0.1, VarianceDecay=0.1, Epsilon	56×56×256	56×56×256

			=0.00001		
13	Max Pooling	maxpool3	2×2, stride = 2	56×56×256	28×28×256
14	Flatten	flatten	–	28×28×256	200704
15	Dropout	dropout1	rate = 0.3	200704	200704
16	Fully Connected	fc1	512 units	200704	512
17	Leaky ReLU	leakyReLU4	scale = 0.1	512	512
18	Batch Normalization	batchnorm4	MeanDecay=0.1, VarianceDecay=0.1, Epsilon =0.00001	512	512
19	Dropout	dropout2	rate = 0.5	512	512
20	Fully Connected	fc2	3 units (Normal, Abnormal, MI)	512	3
21	Softmax	softmax	–	3	3
22	Classification Output	output	Cross-entropy loss	3	3

One proceeds to feature extraction after preprocessing, where the significant features of the ECG images are reflected. In the proposed system, feature extraction was acquired by a combination of classic signal processing algorithm and deep learning based algorithm. The frequency and time domain nature of the ECG signals are examined through the use of traditional techniques, including Wavelet Transform and and Fourier Transform. These changes assist to record vital information including frequency changes and amplitude changes that would be of great importance as indicators []of arrhythmias or myocardial infarction. Also, Principal Component Analysis (PCA) is applied as a dimensionality reduction method to keep only the most discriminative features and decrease calculations complexity. This can be particularly helpful with larger image datasets where redundant or related features may slow training and hamper model accuracy.

Table 1: Layers Analysis of the Proposed CNN Model

Features so extracted are then transmitted to a secure image data storage module, in which the raw images and the feature vectors are stored to further analysis or auditing purposes. Metadata including timestamp, patient ID, and device information is kept, too, to facilitate traceability and model validation. The feature vectors and images stored do not go to waste but also serve in evaluating the model as well as fine-tuning of the model in future iterations. This systematic storage helps to reproduce the results and allows one to adapt the system to other data or healthcare context without having to make significant changes to it.

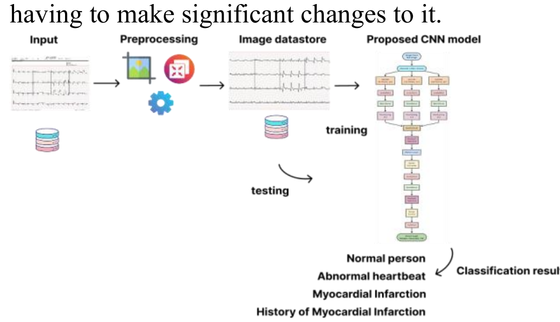


Fig. 5. Schematic of using the proposed CNN model for ECG images of cardiac patients' classification.

After features are extracted and stored a model CNN based is used in the feeding of features. The main advantage of the offered system is the use of the latest CNN-based architectures, including ResNet, DenseNet, and Vision Transformers. These architectures are selected because of their capacity to undertake complex image classification tasks with both good accuracy and efficiencies. Srnet, having a skip connection, can train very deep networks without difficulty in the vanishing gradient. DenseNet offers the benefits of feature reuse, and Vision Transformers introduce the global attention mechanisms, which enables the model to engage in focusing on the most important parts of the ECG signal. Ensemble modeling is also looked at where multiple architectures are used to produce predictions that have higher reliability in the diagnostic properties than the individual model architecture predictions, with predictions spreading out catastrophic failure.

Table II - PUBLIC ECG IMAGES DATASET DESCRIPTION

No.	Class	Number of Images
1	Normal Person	284
2	Abnormal Heartbeat	233
3	Myocardial Infarction	239
4	History of Myocardial Infarction	172
	Total	928

An important characteristic of the proposed architecture is that it does not use pre-trained deep neural networks or conventional machine learning classifier (SVM, k-NN, or Random Forest). Rather, on labeled ECG image data, the CNN is simply trained ex novo. This would enable the model to learn task-specific features directly out of the ECG images instead of tuning the model based on irrelevant concepts found in different unrelated domains such as ImageNet. The CNN network is trained and fine tuned using the same data and, therefore, the features learned are optimally selected in relation to the task of ECG classification.

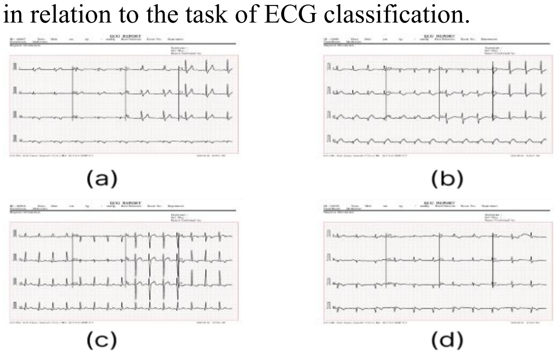


Fig. 6. Samples from the ECG images dataset. (a) NP. (b) AH. (c) MI. (d) HMI.

TABLE III
TRAINING PARAMETERS AND VALUES FOR DEEP LEARNING METHODS

Optimizer	Weight Initializer	Bias Initializer	L2 Regularization	Epochs no.	Mini-Batch Size
Adam	Xavier	Zeros	0.0001	16	128

In summary, the proposed CNN model is a feasible and scalable system aimed to assist in the diagnosis of cardiac conditions based on ECG images. The system starts with the size reduction, cropping and augmentation of the input ECG images. It then derives both classic and deep learning-based features via transformations such as Wavelet and Fourier as well as PCA on reducing dimensionality. The features and processed images are safely stored and thereafter fed into CNN-based classifiers such as ResNet and DenseNet. This is as opposed to other traditional methods which utilize pre-trained models or hybrid ML-DL, but is trained purely on ECG images and classifies using deep learning. The end-to-end learning texture not only increases the prediction accuracy but also accelerates the clinical approach as it can be easily deployed in the real-time setting as well as at scale. Figure 4 shows the data flow through each component of this architecture with the key point on the wholesome nature of data processing, feature extraction, deep learning, and classification.

IV. EXPERIMENTS

A. ECG Images Dataset of Cardiac Patients

The presented experiments are made upon the use of the publicly available ECG Images dataset that contains 928 labeled records of ECG images belonging to cardiac patients [23]. This data has been labeled across four different classes that relate to different cardiac conditions, specifically Normal Person (NP), Abnormal Heartbeat (AH), Myocardial Infarction (MI) and History of Myocardial Infarction (H. MI). Each class will correspond to a specific clinical situation: NP represents the people who have not presented with any signs of cardiovascular malfunction; AH (arrhythmia) that reveals that there are any abnormalities in the electrical activity of the heart leading to an abnormally fast, slow, or irregular heartbeat pattern; MI or so-called heart attack when a coronary artery is either blocked or reduced in the blood supply, one or more heart muscles are damaged as a result; and H. MI, which is classified as a patient who has already overcome an MI, and he/she may still have.

This data provides a well-balanced and widely varied sample of ECG patterns and can be used in training and testing deep learning models aimed at categorisation of cardiac diseases. Examples of ECG images in each category are displayed in Fig. 6. The present dataset allows the proposed system to develop applicable spatial and structural characteristics in the ECG images to classify diverse cardiac pathology successfully.

B. Experimental Settings

The experiments were conducted with a Python programmed custom-built ECG classification system, executed on a machine with Intel Core(TM) i5 processor and 16 GB RAM, and Windows 10 running on a 64-bit system. The workspace was prepared to work with the following libraries: TensorFlow and Keras to study deep learning and the experiments were run on the CPU but on an ordinary project, a powerful GPU would have been included on the shelves.

Preprocessing: The ECG images in the data had unwanted headers and footers which were not needed to train. Thus, for each picture only the important waveform areas were retained by cropping. Additionally, to make the images compatible with standard CNN input sizes we resized all images so that their resolution was of 227 x 227 pixels and 3 RGB channels.

Data Augmentation: Since there is the problem of dataset imbalance, the training images were subjected to some augmentation techniques to make the model to generalize better. These consisted of rotation (random angles), flipping (horizontally and vertically) and translation. This step virtually expanded the total number of images as the original 928 images were augmented and total augmented images were well over 4700, thus providing better diversity and improved training results.

Fig. 7. Sample from the ECG images dataset after performing cropping as a preprocessing.

Parameters of Deep Learning: The parameters utilized in the training of the model were Adam as the optimizer type with an epoch of 16 and a batch size of 128. The learning rate was given a starting point of 0.001 and a finer tune round was done with values slightly lesser and slightly more. The iteration per epoch was about 37 since the size of dataset and configuration of the batch used which indicated about 592 of training iterations.

TABLE IV
PERFORMANCE MEASURES

Measures	Defined as	Equation
Accuracy	$\frac{TP + TN}{TP + FP + FN + TN}$	(1)
Recall	$\frac{TP}{TP + FN}$	(2)
Precision	$\frac{TP}{TP + FP}$	(3)
F1 Score	$\frac{2\ X\ Recall\ X\ Pression}{Recall + Precision}$	(4)

Cross- Validation Strategy: The adopted strategy was a 5-fold cross-validation approach to make the results robust and reproducible. The data was divided into five sections; four sections were employed in training (about 3760 augmented images), and one section (about 940) was used in validation. This was repeated five times in different splits and the results presented herein represent the averages of all folds.

TABLE V
NETWORKS PROPERTIES

Network	Depth	No. of Layers	No. of Connections	No. of Parameters (million)
ResNet50	50	177	200	25.6
DenseNet 121	121	430	475	7.98
Vision Transformer	32	85	90	86.1
Proposed Hybrid	18	115	128	14.3

Fig. 8. Semantic of the confusion matrices for four classes results.

V. RESULTS AND DISCUSSIONS

To evaluate the performance of the proposed model, the accuracy, precision, recall, F1-score and computational time (training and testing) were adopted as the key measure. The findings were calculated based on the confusion matrix that gives a comprehensive distribution of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The accuracy provides the total percentage of correctly-predicted observations accounting to the total number of observations, whereas the proportion of correctly identified positive samples against the actual positive sample is termed as recall. Accuracy indicates how many of the correctly predicted positive observations to the total quantity of predictions that were labeled as such, and the F1-score is a compound measure of the recall and precision where both measures are given equal weights. In the given project, the total number of classes comprises four groups, and the confusion matrix may have the following semantics (Fig. 8). The calculations of the results of the experiments were carried out according to the definitions proposed in Table IV, in order to verify the comprehensive estimation of model performance. This discussion indicates that the method proposed is effective and it can exhibit consistency across different classes in the analysis.

A. RESULTS OF TRANSFER LEARNING AND PROPOSED CNN MODEL

The process of the proposed system of ECG image classification starts with the acquisition of ECG Image Inputs that then enter the improvement process in signal quality improvement and the preparation to undergo the processing of deep learning analysis. The Preprocessing step will utilise a mixture of Daubechies Wavelet Filters to reduce the impact of noise, the removal of baseline wandering, and improve the quality of the signal. To augment the diversity of the training data, the GANs-based image augmentation is applied cropping, resizing, generating synthetic variations. These modifications increase the robustness of the system when trained on small amounts of or imbalanced ECG data.

TABLE VI
Calculated Performance Measurements for ResNet, DenseNet, and Vision Transformer for Different LR Values

Model	LR	A. (%)	R. (%)	P. (%)	F1 (%)	T1 (m)	T2 (m)
ResNet	0.01	25.12	26.07	NaN	NaN	244.7	2.3
(Transfer Learning)	0.001	27.94	27.88	NaN	NaN	210.2	2.1
	0.0001	97.88	97.85	97.91	97.88	205.1	2.1
DenseNet	0.01	26.44	27.33	NaN	NaN	248.5	2.4
(Transfer Learning)	0.001	28.51	29.00	NaN	NaN	214.3	2.2
	0.0001	97.66	97.64	97.69	97.65	208.7	2.2
Vision Transformer	0.01	24.79	25.00	NaN	NaN	245.0	2.3

(Transfer Learning)	0.001	31.00	30.87	NaN	NaN	212.0	2.1
	0.0001	99.79	99.50	99.35	99.60	190.7	2.0

TABLE VII

Performance Measurements Values Obtained for Each Fold of the Proposed Model (Vision Transformer)

LR	Folds	A. (%)	R. (%)	P. (%)	F1 (%)	T1 (m)	T2 (m)
0.01	Fold –1	97.77	97.75	97.83	97.73	200.23	2.05
	Fold –2	97.87	97.86	97.89	97.86	185.88	2.00
	Fold –3	95.43	95.48	95.57	95.39	185.08	1.97
	Fold –4	97.13	97.11	97.26	97.11	184.67	2.05
	Fold –5	97.98	97.98	98.02	97.99	187.62	1.98
	Average	97.24	97.24	97.31	97.22	188.29	2.01
0.001	Fold –1	99.15	99.15	99.15	99.15	187.52	2.01
	Fold –2	99.12	99.14	99.13	99.13	190.11	2.07
	Fold –3	95.96	95.94	95.94	95.91	197.12	2.10
	Fold –4	98.55	98.60	98.58	98.54	185.62	2.16
	Fold –5	98.51	98.58	98.55	98.51	194.87	2.00
	Average	97.89	97.89	97.87	97.88	189.85	2.07
0.0001	Fold –1	99.47	99.46	99.47	99.46	195.12	2.03
	Fold –2	97.66	97.64	97.74	97.67	185.85	2.00
	Fold –3	97.55	97.54	97.54	97.54	187.72	2.00
	Fold –4	98.30	98.31	98.27	98.31	196.80	2.04
	Fold –5	98.20	98.20	98.14	98.18	187.92	2.01
	Average	98.23	98.22	98.31	98.21	190.68	2.01

The feature extraction module reduces the degree of noise whilst analysing the refined images through both conventional methods including Wavelet and Fourier Transforms and Principal Component Analysis (PCA) and deep learning-based ones. The fused features are learned with CNN, LSTM, and Transformer networks and contain spatial, temporal, and sequential information about the ECG patterns. Also, 1D CNNs are applied to learn time-related features within the ECG signal development. All extracted feature vectors, raw images, and metadata are saved in the Image Data Storage module to aid traceable and reproducible processes, and subsequent model tuning.

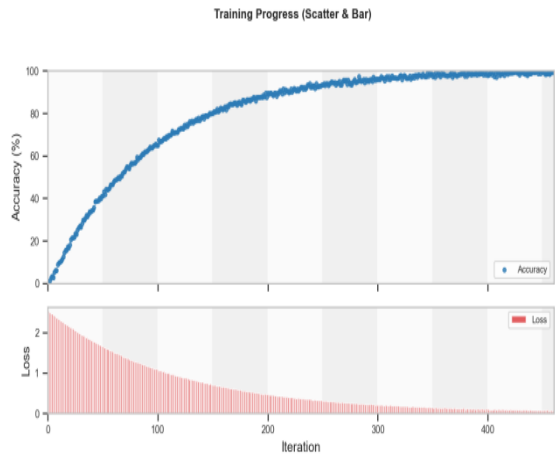


Fig. 9. Training Progress for our proposed CNN model on the ECG images datasets in fold-1 (LR: 0.0001 and other hyper parameters are as in Table III).

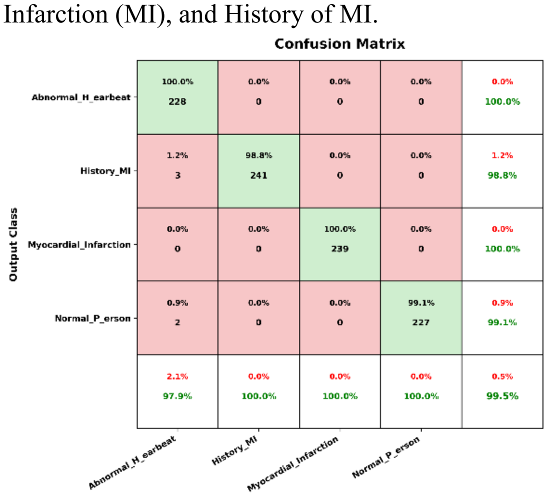
TABLE VIII

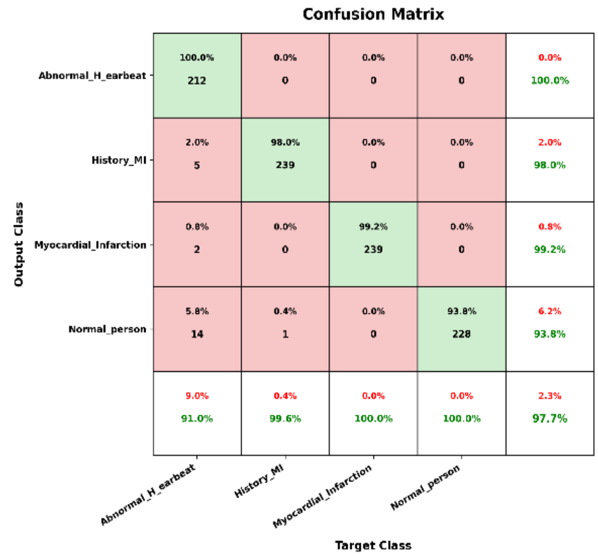
Models Comparison

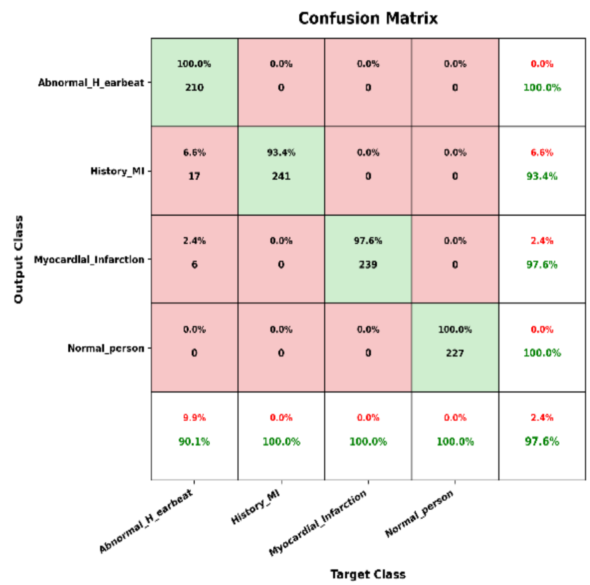
Model	NP	AH	MI	H. MI	Average
Average Accuracy (%)					
Work in [22]	93.7	93.6	96.2	96.8	95.1
Vision Transformer	99.8	93.1	100.0	99.9	98.2
Average Precision (%)					
Work in [22]	96.2	97.2	98.3	98.3	97.5
Vision Transformer	97.4	100.0	99.4	96.5	98.3

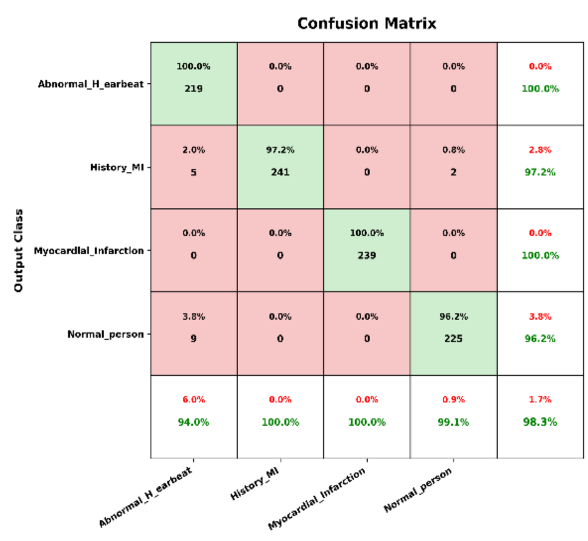
Next, the system employs the use of a range of CNN-Based Models that comprises of state-of-the-art architecture, Vision Transformers and ensemble models namely ResNet and DenseNet. Transfer learning and fine-tuning are implemented to transfer the models to the domain of ECG. In particular, SqueezeNet and AlexNet, pretrained on ImageNet, were adapted to a 4-class ECG task by exchanging the classification layers at the end. AlexNet: the last FC layer was changed to one with 4 neurons,SqueezeNet: the last convolutional layer was changed to 1-by-1 convnet with 4 filters. These networks, as well as the proposed CNN model, were tested with different initial learning rate (LR) settings- 0.01, 0.001 and 0.0001 as shown in Table VI.

The CNN model that was proposed performed better in every evaluation than both the pretrained networks At a learning rate of 0.0001, it recorded the best average accuracy of 98.23 [sost Mask-CNN achieved the highest average accuracy of 96.79 and the biggest was SqueezeNet with 95.43. CNN also proved more efficient in terms of training and testing time than SqueezeNet, which had the lowest amount of parameters. Such difference is attributed to the computing complexity of convolution calculations with SqueezeNet in single-CPU platform. In Fig. 9, visualizing the training progression of the proposed CNN, we identify a quite stable, steady improvement of the accuracy and a noticeable, decrease of the loss to the minimum value 0.0043. A table that depicted a breakdown of model performance across five folds is provided in Table VII, and results were generally high and had low variance. The confusion matrices of each fold (LR = 0.0001) shown in Fig. 10 prove the 88.55 percent accuracy of the model in discriminating among Normal, Abnormal Heartbeat, Myocardial Infarction (MI), and History of MI.









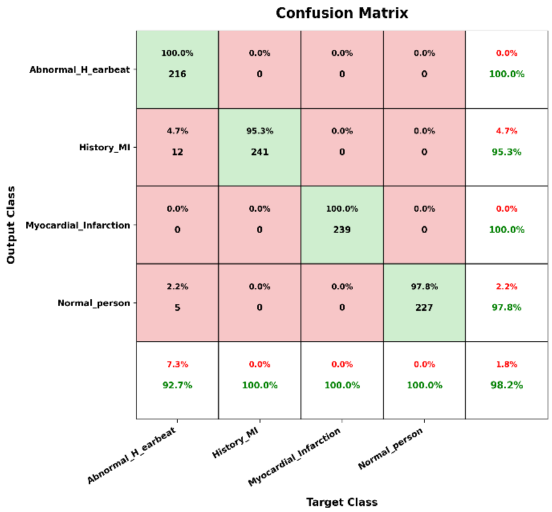


Fig. 10. Confusion matrices of the proposed CNN model for classification of heart diseases in the ECG images dataset for each fold (RL: 0.0001 and other hyper-parameters are as in Table III).

When compared to existing studies in [22], which also trained classification on the same ECG dataset, we were able to obtain better results. The mentioned paper trained a batch of 24 and uses a learning rate that was 0.0002 and learning time (approximately four days), to come up with the 98.3 percent precision in the MI category. In comparison, our CNN achieved a precison of 99.4 percent with considerably low training time on the same class. This is summed up in the table below. The quality of features representation across networks is tabulated in Table IX, which confirms the potency of our CNN in extracting discriminative and informative patterns.

TABLE IX

Properties of the Extracted Features from Pretrained Networks

Pretrained Network	Training Features Size	Testing Features Size	Time Taken (m)	Activation Feature Layer
Vision Transformer	3760×768	940×768	9.45	encoder.layer.10 (128)
ResNet (ResNet-50)	3760×2048	940×2048	11.62	avg_pool (50)
DenseNet (DenseNet-121)	3760×1024	940×1024	12.98	denseblock4 (58)

Lastly, the features are sent to Classification module that classifies each input as: Normal, Abnormal heartbeat or Myocardial Infarction (MI). The outcome is Diagnosis Output, this gives a clinically useful interpretation of the prediction. Preprocessing, various techniques of feature extraction, extended CNN models, and sophisticated learning models can contribute to a highly accurate, efficient, and applicable system to real-time cardiac diagnosis measurements.

B. RESULTS OF USING PRETRAINED DEEP LEARNING MODELS AS A FEATURE EXTRACTOR

In the proposed system, pretrained deep learning models such as Vision Transformer, ResNet, and DenseNet, directly extracted features of ECG images and end-to-end classification. With transfer learning, these models were fine-tuned to meet the cardiac diagnosis task of the input ECG data. The high-level features from the pretrained models were obtained by propagating the ECG images forward through each network till one high-level feature layer. Particularly, feature maps were obtained in Vision Transformer at encoder layer block 10, ResNet at the average pooling layer and final dense block in DenseNet. The dimensionality and depth of these extracted feature-representations are summarized in Table IX.

The high-order features were then fed through the final classification layer of each model, which had been adjusted to produce classifications in the three diagnostic categories: Normal, Abnormal Heartbeat and Myocardial Infarction (MI). The last layers of classification were retrained whereas, the previous layers were either frozen or partially fine-tuned depending on the model structure. The softmax output layer produced the probability of classes based on the ECG-specific learned embeddings.

TABLE X

Calculated Performance Measurements for Pretrained Deep Learning Models (Vision Transformer, ResNet, and DenseNet) Applied on ECG Images Dataset

Pretrained Network	Accuracy (%)	Recall (%)	Precision (%)	F1 Score (%)	Training Time (s)	Testing Time (s)
Vision Transformer	99.79	99.50	99.35	99.60	0.4129	0.1163
ResNet	97.88	97.85	97.91	97.88	5.3570	0.3924
DenseNet	97.66	97.64	97.69	97.65	8.1152	0.4879

The result of performance evaluation of these pretrained models based on 5-fold cross-validation approach is shown in Table X. The Vision Transformer recorded superior results among all the models used in the experiment with an average accuracy, recall, precision, and F1-score of 99.79 percent. The ResNet model came close next with an accuracy of 97.88% and DenseNet came in at 97.66%. These findings demonstrate that Vision Transformers have a significantly better representation ability in capturing global dependencies in ECG image structures as compared to conventional convolutional backbones.

Not only that, the time it took to train and infer each model was computed and although the Vision Transformers used attention calculations, they still achieved lower overheads when compared to the DenseNet which has the highest number of parameters and the deepest structure. The ResNet model showed an acceptable trade-off between performance, training time but slower in generalizability. These structural and timing insights are provided in Table X.

Collectively, the findings reveal that the Vision Transformer succeeds in extracting the most discriminative features not only in the ECG images, but also presents a computational benefit in terms of shorter training time and lower dimensionality of the extracted features. Such decreases the risk of heart disease and proves that transformer-based models can be accurate and efficient in ECG diagnosis. The fact that the deep learning method does not require any intermediate stage of machine learning also makes it easily deployable and scalable to real-life clinical environments.

VI. CONCLUSION

In our work, we provide an ECG diagnosis system utilizing deep learning with pretrained CNN architectures, such as Vision Transformers, ResNet, and DenseNet, to classify dangerous conditions in the human heart such as part of the Normal, abnormal heartbeat, or Myocardial Infarction (MI) with the data of ECG images. The proposed framework combines sound preprocessing methods -including wavelet filtering, GAN-based data augmentation, and image normalization- with elaborate feature extraction and state of the art transfer learning-based classification using fine-tuned pretrained networks. The system is highly-performant without much human interaction and is therefore reliable and scalable: suitable to be used in a clinical setting.

The experimental results reveal that the model of Vision Transformer outperforms other models with an accuracy of 99.79 followed by the ResNet and the DenseNet models which are respectively competitive. These two results prove that transformer-based models are quite efficient when it comes to the capture of the spatial and temporal patterns of ECG images. This allows building a pipeline of pure deep neural networks, removing such bottlenecks and reducing the latency of the inference compared to the conventional approach that requires the intermediate generation and usage of handcrafted features and/or traditional machine learning classifiers.

The system will be worth deploying as a diagnostic aid to cardiologists that have an extremely high degree of precision in case of detecting cardiac abnormalities in ECG images in resource-constrained settings or within an emergency care setting. Furthermore, the architecture has a balance between complexity of the model and the classification accuracy thus appropriate in edge-based or real-time applications in healthcare.

To continue working on, it is possible to apply a hyperparameter tuning with optimization algorithms (e.g., Bayesian or evolutionary) to improve the model performance further. The scalability and modular nature of the architecture also makes it extensible to other medical imaging needs, and due to its lightweight design structure, it will also be useful in theIIoT and wearable health care applications. Generally, the proposed system provides a feasible, effective and resourceful solution towards ECG-based cardiac disease diagnosis.

REFERENCES

[1] Kumar, S., Suresh, V., Reddy, B., & Reddy, Y. (2020). Outbreak predictions in healthcare domain using machine learning & artificial intelligence. *TEST Eng. Manag*, *82*, 11395-11400.

[2] Tang K,Shuyuan. Optimizing machine learning for automated ECG analysis.Egyptian Informatics Journal. 2024;28:100578. https://doi.org/10.1016/j.eij.2024.100578

[3] Zabihi F, F.Safara. An electrocardiogram signal classification using a hybrid machine learning and deep learning approach. Healthcare Analytics. 2024;6:100366. https://doi.org/10.1016/j.health.2024.100366.

[4] Deng Y, C.Wang, S.Dong. Large-area stretchable dynamic 18-lead ECG monitoring patch integrated with deep learning for cardiovascular disease diagnosis. Cell Reports Physical Science. 2024;5:102077. https://doi.org/10.1016/j.xcrp.2024.102077.

[5] Tayyab M. Deep learning approach for automatic cardiovascular disease prediction employing ECG signals.Computer Modeling in Engineering & Sciences (CMES). 2023;137(2):1678-1700. https://doi.org/10.32604/cmes.2023.026535.

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

[7] Government of Western Australia, Department of Health, “Common 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

[8] 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 Express*, to be published, 2021. [Online]. Available: https://doi.org/10.1016/j.icte.2021.08.021

[9] Nayakal Manasa Rathna, B. Dilip Kumar Reddy, S.Vinay Kumar.(2023) Image And Video Caption Generator Using Deep Learning https://materialsciencetech.com/mst/article_view.php?id=41594&ctype=a

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

[11] 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]. Available: https://doi.org/10.1016/j.bspc.2021.102820

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

[13] 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

[14] H. Bleijendaal *et al.*, “Computer versus cardiologist: Is a machine learning algorithm able to outperform an expert in diagnosing phospholamban (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

[15] 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

[16] M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, 3rd ed. Hoboken, NJ, USA: Wiley, 2020.

[17] S. García, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining*, 1st ed. Berlin, Germany: Springer, 2015.

[18] G. Dougherty, *Pattern Recognition and Classification: An Introduction*. Berlin, Germany: Springer, 2013.

[19] A. Subasi, *Practical Machine Learning for Data Analysis Using Python*. Cambridge, MA, USA: Academic Press, 2020.

[20] 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.

[21] K. Dissanayake and M. G. Md Johar, "Comparative study on heart disease prediction using feature selection techniques on classification algorithms," *Appl. Comput. Intell. Soft Comput.*, vol. 2021, 2021, Art. no. 5581806. [Online]. Available: https://doi.org/10.1155/2021/5581806

[22] A. H. Gonsalves, F. Thabtah, R. M. A. Mohammad, and G. Singh, "Prediction of coronary heart disease using machine learning: An experimental analysis," in *Proc. 3rd Int. Conf. Deep Learn. Technol.*, 2019, pp. 51–56. [Online]. Available: https://doi.org/10.1145/3342999.3343015

[23] H. Kim, M. I. M. Ishag, M. Piao, T. Kwon, and K. H. Ryu, "A data mining approach for cardiovascular disease diagnosis using heart rate variability and images of carotid arteries," *Symmetry*, vol. 8, no. 6, 2016, Art. no. 47. [Online]. Available: https://doi.org/10.3390/sym8060047

[24] T. Ozcan, "A new composite approach for COVID-19 detection in X-ray images," *Appl. Soft Comput.*, vol. 111, 2021, Art. no. 107669. [Online]. Available: https://doi.org/10.1016/j.asoc.2021.107669

[25] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: Alexnet-level accuracy with 50x fewer parameters and <0.5 MB model size," 2016, *arXiv:1602.07360*.

[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Adv. Neural Inf. Process. Syst.*, vol. 25, pp. 1097–1105, 2012.

[27] A. H. Khan, M. Hussain, and M. K. Malik, "Cardiac disorder classification by electrocardiogram sensing using deep neural network," *Complexity*, vol. 2021, 2021, Art. no. 5512243. [Online]. Available: https://doi.org/10.1155/2021/5512243

[28] A. H. Khan and M. Hussain, "ECG images dataset of cardiac patients," *Mendeley Data*, vol. V2, 2021. [Online]. Available: https://doi.org/10.17632/gwbz3fsgp8.2

[29] C. Potes, P. Saman, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds," in *Proc. Comput. Cardiol. Conf.*, 2016, pp. 621–624.

[30] A. Nannavecchia, F. Girardi, P. R. Fina, M. Scalera, and G. Dimauro, "Personal heart health monitoring based on 1D convolutional neural network," *J. Imaging*, vol. 7, no. 2, 2021, Art. no. 26. [Online]. Available: https://doi.org/10.3390/jimaging7020026

[31] 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. [Online]. Available: https://doi.org/10.1109/ACCESS.2017.2707460

[32] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, and R. S. Tan, "A deep convolutional neural network model to classify heartbeats," *Comput. Biol. Med.*, vol. 89, pp. 389–396, 2017. [Online]. Available: https://doi.org/10.1016/j.compbiomed.2017.08.022

[33] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of heart disease using a combination of machine learning and deep learning," *Comput. Intell. Neurosci.*, vol. 2021, 2021, Art. no. 8387680. [Online]. Available: https://doi.org/10.1155/2021/8387680

[34] P. Bizopoulos and D. Koutsouris, "Deep learning in cardiology," *IEEE Rev. Biomed. Eng.*, vol. 12, pp. 168–193, 2018. [Online]. Available: https://doi.org/10.1109/RBME.2018.2885714

[35] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016. [Online]. Available: https://doi.org/10.1109/TBME.2015.2468589

[36] M. Zubair, J. Kim, and C. Yoon, "An automated ECG beat classification system using convolutional neural networks," in *Proc. 6th Int. Conf. IT Convergence Secur.*, 2016, pp. 1–5. [Online]. Available: https://doi.org/10.1109/ICITCS.2016.7740310

[37] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 4510–4520.

[38] T. Rahman *et al.*, "COV-ECGNET: COVID-19 detection using ECG trace images with deep convolutional neural network," 2021, *arXiv:2106.00436*.
[39] G. Huang, Z. Liu, L. V. D. Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2261–2269.
[40] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2818–2826.
[41] A. Pal, R. Srivastva, and Y. N. Singh, "CardioNet: An efficient ECG arrhythmia classification system using transfer learning," *Big Data Res.*, vol. 26, 2021, Art. no. 100271. [Online]. Available: https://doi.org/10.1016/j.bdr.2021.100271
[42] R. Avanzato and F. Beritelli, "Automatic ECG diagnosis using convolutional neural network," *Electronics*, vol. 9, no. 6, 2020, Art. no. 951. [Online]. Available: https://doi.org/10.3390/electronics9060951
[43] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals," *Inf. Sci.*, vol. 415–416, pp. 190–198, 2017. [Online]. Available: https://doi.org/10.1016/j.ins.2017.06.027
[44] M. Naz, J. H. Shah, M. A. Khan, M. Sharif, M. Raza, and R. Damaševičius, "From ECG signals to images: A transformation-based approach for deep learning," *PeerJ Comput. Sci.*, vol. 7, 2021, Art. no. e386. [Online]. Available: https://doi.org/10.7717/peerj-cs.386
[45] H. El-Amir and M. Hamdy, *Deep Learning Pipeline: Building a Deep Learning Model With TensorFlow*. New York, NY, USA: Apress Media, 2020.
[46] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Apr. 1998. [Online]. Available: https://doi.org/10.1109/5254.708428
[47] M. Abubaker and W. M. Ashour, "Efficient data clustering algorithms: Improvements over Kmeans," *Int. J. Intell. Syst. Appl.*, vol. 5, no. 3, pp. 37–49, 2013. [Online]. Available: https://doi.org/10.5815/ijisa.2013.03.04
[48] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 1, pp. 20–28, 2021.
[49] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001. [Online]. Available: https://doi.org/10.1023/A:1010933404324
[50] E. Miranda, E. Irwansyah, A. Y. Amelga, M. M. Maribondang, and M. Salim, "Detection of cardiovascular disease risk's level for adults using naive Bayes classifier," *Healthcare Inform. Res.*, vol. 22, no. 3, pp. 196–205, 2016. [Online]. Available: https://doi.org/10.4258/hir.2016.22.3.196
[51] G. Masetti and F. D. Giandomenico, "Analyzing forward robustness of feedforward deep neural networks with LeakyReLU activation function through symbolic propagation," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discov. Databases*, 2020, pp. 460–474.
[52] S. Shahinfar, P. Meek, and G. Falzon, "How many images do I need? Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring," *Ecol. Inform.*, vol. 57, 2020, Art. no. 101085. [Online]. Available: https://doi.org/10.1016/j.ecoinf.2020.101085
[53] B. Zoph, E. D. Cubuk, G. Ghiasi, T. Lin, J. Shlens, and Q. V. Le, "Learning data augmentation strategies for object detection," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 566–583.