Architecture Image Studies

Vol.7, Issue 1, pp.1484-1498, 2026
DOI: https://doi.org/10.62754/ais.v7i1.1053

https://journals.ap2.pt/index.php/ais/index

A Custom Dilated-Separable CNN for Automated Cardiovascular Disease Detection Using Electrocardiogram Images

Abdulrahman Hamid Mahmood[1], Taha Mohammed Hasan[2]

Abstract

Cardiovascular diseases are a leading cause of millions of deaths worldwide, placing a huge burden on healthcare systems. Therefore, early and accurate diagnosis of cardiovascular diseases is essential in determining efficient treatment and preventing life-threatening complications. Traditional diagnostic approaches that rely on manual interpretation of electrocardiograms (ECGs) are often subject to inter-observer variability and are time-consuming, highlighting the urgent requirement for effective, accurate, and automated diagnostic support systems. In this article, a customized Convolutional Neural Network (CNN) based diagnostic system is proposed for automated detection of cardiovascular diseases using a publicly available 12-lead ECG images dataset. Extensive pre-processing and data augmentation were applied to the dataset to improve signal diversity and minimize overfitting. This proposed CNN incorporates several building blocks, involving dilated temporal and separable wide convolutional layers, aimed to possess fine-grained morphological patterns and wider spatial dependencies in ECG images. Additionally, several layers of batch normalization and dropout were utilized for stabilizing training and improving generalization. Experimental evaluations revealed a superior classification accuracy of 99% for the proposed system, outperforming state-of-the-art pre-trained CNNs and existing related systems. Moreover, a diagnostic support tool has evolved to facilitate real-time implementation in clinical environments, providing an effective and easy-to-interpret framework for detecting cardiovascular diseases.

Keywords: *Automated Cardiovascular Disease Detection, Dilated-Separable CNN, Data Augmentation, Electrocardiogram Images Dataset.*

Introduction

With an aging population and unhealthy lifestyles such as a deleterious diet, obesity, excessive alcohol consumption, drug abuse, smoking, some diseases such as diabetes, high blood pressure, and high cholesterol, and even poor mental health (such as despair or anxiety), the risk of cardiovascular disease is gradually increasing [1]. The World Health Organization estimates that more than 23 million people worldwide will die from this disease by 2030. Therefore, early diagnosis is crucial for the survival of patients with cardiovascular disease [2].

In the detection of cardiovascular diseases, electrocardiograms (ECGs) and phonocardiograms (PCGs) provide important, albeit different, diagnostic information. While PCG signals are beneficial for detecting valvular defects and murmurs, they are highly susceptible to interference and lack support from high-quality and large public datasets, limiting their efficiency in robust deep learning [3]. However, ECGs remain the preferred method for identifying cardiac arrhythmias, conduction disturbances, ischemic changes, and other electrophysiological abnormalities. This is because ECGs directly capture the heart's electrical activity, provide abundant datasets, and offer clearer clinical classification [4]. Most cardiologists utilize paper-based ECG images to manually analyze diverse parameters (such as rhythm, heart rate, potential abnormalities, etc.) to assess cardiac function. However, conventional manual ECG image analysis is time-consuming and laborious. The process is also prone to errors and can lead to life-threatening consequences. This requires the development of a fast, accurate, reliable, and automated detection system using ECG images to assist physicians and potentially reduce mortality [5].

[1] Department of Computer Science, College of Science, University of Diyala, Iraq, Email: scicompms232406@uodiyala.edu.iq
[2] Department of Computer Science, College of Science, University of Diyala, Iraq, Email: dr.tahamh@uodiyala.edu.iq

Deep learning architectures, along with advances in image analysis and computer vision, have revolutionized the interpretation of medical images by extracting essential quantitative information and utilizing it to predict risks in patients [6-8]. Over the past few years, the detection of heart abnormalities in ECG data has undergone significant changes, following the emergence of new technologies that have further enhanced deep learning capabilities beyond former, less effective models [9]. As a specialized type of deep learning model, Convolutional Neural Networks (CNNs) are utilized to provide faster and more efficient predictions of heart abnormalities based on ECG data [10].

This article utilizes ECG images as its fundamental input modality. These two-dimensional representations of ECG signals are well-suited for CNNs, offering the advantages of increased ECG diagnostic sensitivity, improved dataset availability, and accurate and effective performance in detecting cardiovascular disease. Leading contributions of the proposed cardiovascular disease detection system are as follows:

1. Developing a customized CNN of dilated, separable, and wide kernel convolutional layers for improving temporal context extraction from ECG images and decreasing computational complexity.
2. Implementing an advanced pipeline of image enhancement, normalization, standardization, and augmentation on the publicly accessible ECG images dataset. Accordingly, this improvement led to more balanced class distribution, minimized overfitting, and made the model robust and generalizable across diverse ECG image morphologies.
3. Incorporating a hybrid block structure of dilated and separable convolutions for wider temporal reception domains and feature learning decomposition, thus enhancing sensitivity to subtle morphological ECG changes associated with cardiovascular abnormalities.
4. Utilizing several layers of batch normalization, global average pooling, and spatial dropouts for inhibiting overfitting and stabilizing training over limited ECG images.
5. Providing comparative experiments with state-of-the-art pre-trained networks, demonstrating the superior results of the proposed CNN across all evaluated metrics.
6. Developing a diagnostic support system to facilitate immediate prediction of cardiovascular diseases from ECG image inputs in a clinical setting.

Related Work

Numerous recent works have shown the effectiveness of utilizing ECG images as time-amplitude graphs instead of less versatile raw ECG signals in training deep learning models for detecting cardiovascular diseases, for instance, Mangaraj et al. [11] presented an arrhythmias detection system, which transformed ECG signals into two-dimensional grayscale images to highlight irregularities and patterns that might be overlooked in the one-dimensional signals and trained a deep CNN model to classify ECG beats from the MIT-BIH arrhythmia dataset. Other works have utilized spectrogram-based ECG images to acquire frequency and temporal features for the automated classification of arrhythmias. For instance, Lilda and Jayaparvathy [12] employed a continuous wavelet transform to transform ECG signals into two-dimensional ECG spectrograms, and then applied a multi-branch CNN to classify ECGs from the MIT-BIH arrhythmia dataset. However, the transformation to spectrogram images has critical limitations, including poor time–frequency resolution, reduced clinical interpretability, and high computational complexity [13]. Accordingly, analyzing ECG signals by directly transforming them into ECG images rather than relying on spectrograms is more effective in deep learning based-cardiovascular disease detection.

Recent developments in deep learning models have considerably improved the detection of cardiovascular diseases from ECG images. Furthermore, various techniques have been advanced to handle class imbalances and computational efficiency issues, further enhancing classification performance. Table 1 demonstrates a comparative summary of the latest cardiovascular disease detection systems. Khan et al. [14] proposed a system for detecting cardiovascular disease using a 12-lead ECG images dataset in non-standardized formats, manually collected from diverse ECG devices for patients with Cardiac. In this system, 928 ECG images (acquired from Ch. Pervaiz Elahi Institute of Cardiology) were first manually resized and labeled. Subsequently, the lightweight MobileNet-V2 single-shot detection architecture was utilized to automatically extract features and classify four cardiac abnormalities (normal heartbeats, abnormal, myocardial infarction, and prior history of myocardial infarction). This system provided a generalized, image-based pipeline that achieved remarkable accuracy (98%) across diverse ECG print formats. Despite this high accuracy, imbalance might affect the generalization of the minority category. Furthermore, training the system on a relatively small dataset could increase the risk of overfitting. Mhamdi et al. [15] presented a deep learning-based system for cardiovascular disease detection using a 12-lead ECG images dataset. In this system, data augmentation was employed to increase the quantity and variety of limited data, and before being fed into deep learning models, these images underwent resizing and normalization. Subsequently, transfer learning and fine-tuning techniques were used in the VGG16 and MobileNetV2 models, with their final layers adapted to produce four classes. Both models achieved a high validation accuracy (about 95%) on the validation set. However, this presented system lacked an independent test set for robust validation and explicitly warned that it should not replace a clinician's handiwork. Fatema et al. [16] presented a deep learning-based detection system for classifying cardiovascular disease into five classes (normal, abnormal heartbeat, myocardial infarction, prior history of myocardial infarction, and COVID-19) using a highly imbalanced and limited in size, 12-lead ECG images dataset for patients with Cardiac and COVID-19 conditions. In this system, 1682 ECG images (acquired from various cardiac institutions across Pakistan) were first pre-processed by removing artifacts and enhancing the quality of the ECG images. Subsequently, in addition to implementing several core deep learning models (VGG19, InceptionV3, MobileNetV2, DenseNet201, and ResNet50), a combination of InceptionV3 and ResNet50 (designated InRes-106) was implemented. The proposed hybrid CNN model successfully extracts hidden and high-level features from pre-processed images, superimposing other implemented models by achieving a best-in-class performance with a test accuracy of 98.34%. However, the performance of the hybrid model on real-time image data has not been explored. Prashant et al. [17] also utilized the 12-lead ECG images dataset for patients with Cardiac and COVID-19 conditions that were first preprocessed by cropping only the ECG plots, removing gridlines, and resizing. Three pre-trained CNNs (VGG-19, EfficientNet-B4, and DenseNet-121) were implemented individually, without an ensemble approach for binary classification and with an optimized weighted ensemble approach for multi-class classification. The achieved binary and multi-class classification accuracies were 100% and 95.29%, respectively. Achieving these results, especially concerning binary classification, may be questionable given the unbalanced and non-augmented dataset. Aversano et al. [18] presented a CNN-based detection system that utilized a 12-lead ECG images dataset for a binary and multi-class classification of cardiovascular diseases. In this system, each ECG image was first segmented into three anatomical bands for isolating various lead groups and resolution-reduced before being fed into a two-dimensional CNN model of several layers (convolution with ReLU, pooling, and fully-connected) to learn ECG patterns and classify the ECG images. Although the strategy of three bands can be considered innovative, it is not medically approved and requires further studies to confirm its reliability. Furthermore, the aggressive resizing applied for reduced resolution can cause information loss, which potentially leads to inaccurate results. The accuracy achieved by this system was 91% for binary classification, where ill patients were the best classified, and 80% for multi-class classification of cardiovascular diseases. Ashtaiwi et al. [19] assessed three cardiovascular disease detection systems that vary in their utilization of data augmentation, pre-processing, essential feature extraction, and classification models using a 12-lead ECG images dataset for patients with Cardiac. In the first system, several augmentation techniques (including image resizing and rotation, vertical and horizontal flipping, and RGB normalization) were applied without any preprocessing fed to the baseline CNN for classification. The second system utilized image cropping, followed by VGG16 to extract essential features, which were then fed to an artificial neural network (ANN) for classification. The third proposed system implemented comprehensive pre-processing (including cropping, background and grid-line removal, and binary pixel assignment), and it converted every ECG image into a one-dimensional summed feature vector using a developed image vectorization approach for training various ANN classifiers. This proposed system was capable of overcoming severe imbalances in features, significantly reducing feature dimensions, and improving classification accuracy. However, its ability to generalize is limited, and its performance is weaker in distinguishing between multi-classes compared to binary classification. Hasan et al. [20] presented a hybrid cardiovascular disease detection system that integrates both deep and machine learning models. This system combined the two versions of the 12-lead ECG images datasets (of 4 classes and 5 classes) to create a new balanced ECG images dataset of 4 classes. The 1406 ECG images created were first pre-processed to remove artifacts and background line and enhance image quality. A deep CNN model was employed to extract essential and complex ECG features, and several machine learning models (XGBoost, Support Vector Machine, Gaussian Naïve Bayes, Random Forest, Logistic Regression, and Decision Tree) with an ensemble model were utilized for classification. This system achieved an accuracy of 99.29%; however, it faced challenges with overfitting, primarily due to the limited dataset size. However, it has faced issues with overfitting, primarily due to the dataset's limited size.

Table 1: Comparison of the Latest Cardiovascular Disease Detection Systems Using the ECG Images Dataset.

Author/(s), Year, Ref.	Dataset Used / No. of Classes	Pre-processing Techniques	Augmentation Techniques	Feature Extraction and Deep Learning Models	Accuracy Results
Khan et al., 2021, [14]	ECG images Dataset / 4-Classes	ECG Images Resizing and Labeling	✕	MobileNet-V2 Single-Shot Detection	98%
Mhamdi et al., 2022, [15]	ECG images Dataset / 4-Classes	ECG Images Resizing and Normalization	✓	VGG16 or MobileNetV2	≈95%
Fatema et al., 2022, [16]	ECG images Dataset / 5-Classes	Automated Cropping System, Graph Line Removal, and Histogram Equalization	✕	InRes-106	98.34%
				InceptionV3	90.56%
				ResNet50	89.63%
Prashant et al., 2022, [17]	ECG images Dataset / 5-Classes	Image Cropping, Gridlines Removal, and Image Resizing	✕	(VGG-19, EfficientNet-B4, and DenseNet-121) with Optimized Weighted Average Ensemble	95.29%
Aversano et al., 2023, [18]	ECG images Dataset / 4-Classes	Cropping ECG images into three bands and reducing their resolution	✕	Two-dimensional CNN	91% for binary and 80% for multi-class classification
Ashtaiwi et al., 2024, [19]	ECG images Dataset / 4-Classes	Image Cropping, Background and Grid-line Removal, and Binary Pixel Assignment	✓	ANN	≈98% for binary and ≈88% for multi-class classification
Hasan et al., 2025, [20]	New Combined ECG Images Dataset / 4-Classes	Image Cropping, Canny Edge Detection, Morphological Operations, Histogram Equalization	✕	CNN Model and An Optimized Weighted Average Ensemble Machine Learning	99.29%

Proposed Methodology

In this article, the proposed system presents a developed deep learning model that enables individuals to recognize cardiac abnormalities. The primary aim of this system is to create a dilated-separable CNN model that accurately and precisely detects cardiovascular diseases. This section, which includes ECG images input, pre-processing and data augmentation, feature extraction, regularization, and classification, describes the proposed methodology employed to achieve the system's aim. All the structural elements of the proposed methodology are demonstrated in Figure 1.

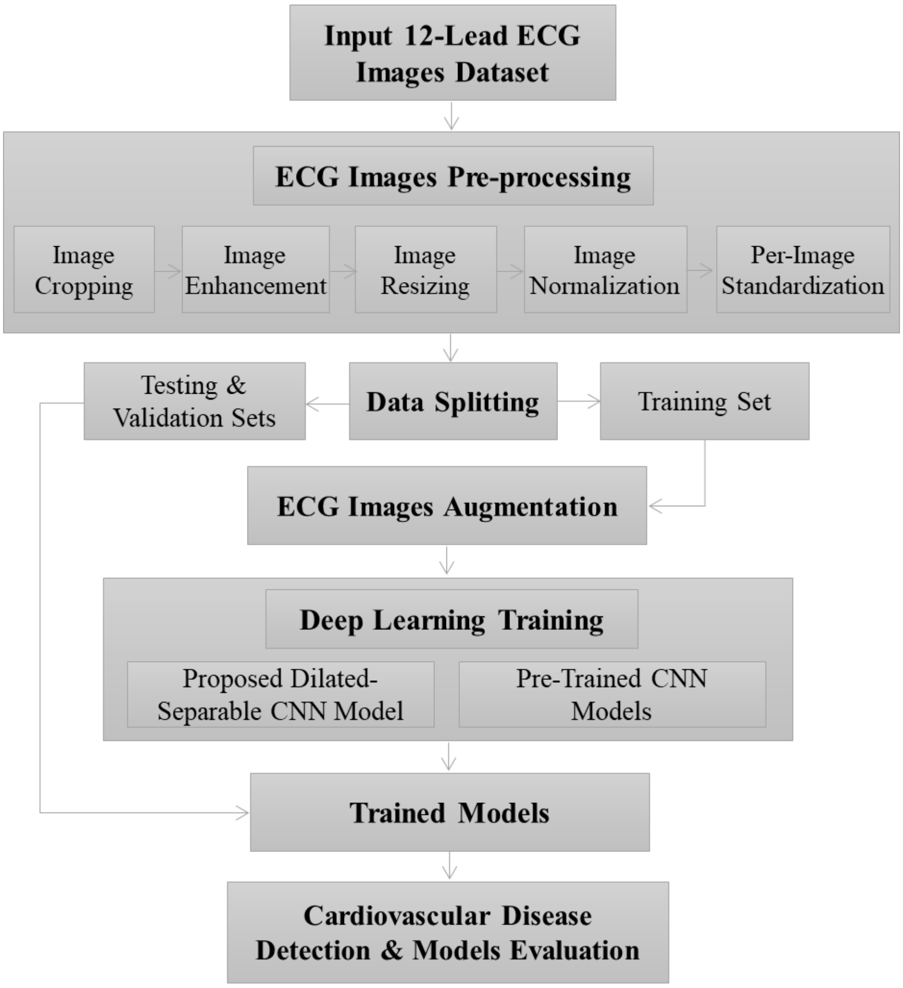


Figure 1: Structural Elements of the Proposed Methodology.

Dataset Description

In this proposed system, a standard 12-lead ECG images dataset for patients with Cardiac conditions was utilized. This dataset was manually collected from diverse ECG devices with the support of Ch. Pervaiz Elahi Institute of Cardiology, and acquired from the Mendeley repository [21]. It includes 928 images of four classes. A description of the dataset used is presented in Figure 2. The following are the classes of heart conditions as seen in following Table:

Table 2: Description of the Classes of Heart Conditions

Heart Conditions	Description
"Normal heartbeats"	It refers to healthy people who do not suffer from any heart disease, and an adult's heart rate typically ranges between 60 and 100 beats per minute.
"Abnormal heartbeats"	It is involving structural issues in the heart, changes in the heart's electrical conduction system, and underlying medical conditions like sleep apnea, diabetes, and high blood pressure.
"Myocardial infarction"	It is caused by a blockage of blood flow to a portion of the heart. This can lead to the death of the heart muscle, causing severe heart damage and complications capable of ending the patient's life.
"A prior history of myocardial infarction"	It is attributed to the patient having suffered a prior heart attack.

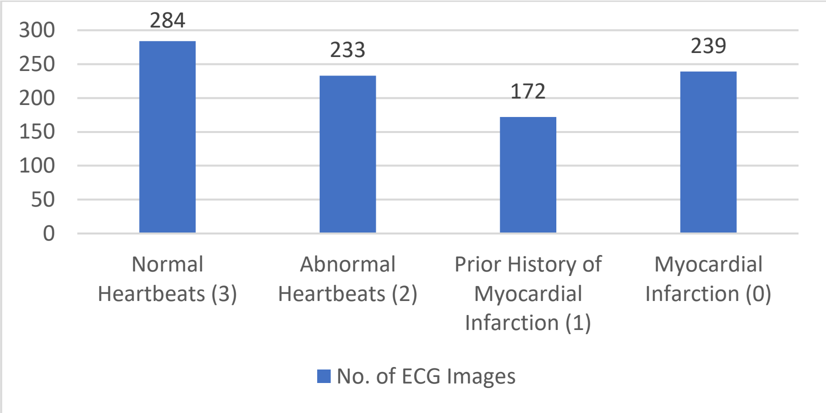


Figure 2: The Description of the 12-Lead ECG Images Dataset.

ECG Images Pre-Processing

Pre-processing ECG images has a significant impact on training the proposed deep learning model, thereby improving the model's performance. ECG images from the Mendeley dataset are diverse in size, contrast, and illumination; therefore, multiple processes are implemented before the feature extraction and classification stages in this proposed system. To preserve the integrity of the ECG visual signals and ensure compatibility with the proposed deep learning architecture, the ECG images are initially loaded and then transformed into three-channel RGB, so that the grayscale information is preserved by repeating the same channel across the three channels without alteration. The following are details of the pre-processes:

- **Image Cropping:** All ECG images involve white margins and dispensable text information (including patient and hospital data) that appear at the top and bottom of each image. These portions represent distraction (noise), which can affect the accuracy of the model classification. Therefore, dispensable portions are removed by cropping the ECG images. This is accomplished by selecting appropriate pixel dimensions that do not affect the area of interest while removing noise. Accordingly, dimensions of 2025×1175 pixels were chosen to reduce the size of the ECG images from 2213×1572 pixels.

- **Image Enhancement:** This process enhances the visibility and discriminative quality of ECG images by applying several techniques, including increasing sharpness by a factor of 1.5, contrast by a factor of 1.2, and slightly improving brightness by a factor of 1.1.

- **Image Resizing:** In this preliminary process, the ECG images are resized to a width of 512 pixels and a height of 256 pixels to obtain adequate chronological sequences and amplitude detail, and to guarantee consistency of input dimensions for the proposed deep learning model. Here, the ECG images are scaled while maintaining the original aspect ratio, and symmetrical padding is added where necessary. This prohibits geometric distortions (warping or stretching of waveforms), which could negatively affect the model's capability of distinguishing cardiovascular patterns.

- **Image Normalization:** This process aims to minimize the variation in input intensity, introduce more stabilized gradient behaviour throughout the backpropagation process, and thus accelerate convergence during training. Here, pixel values are converted from eight-bit integer numerical values (0-255) to values within the range of zero and one.

- **Per-image Standardization:** This process standardizes every ECG image separately (subtracts the image's mean pixel value and divides the result by standard deviation) to enhance the network's robustness against changes in background noise, contrast, and illumination in ECG recordings. Standardization focuses on clinically relevant waveform characteristics while minimizing image inconsistencies.

These processes combined guarantee consistent input quality, reduce noise, and improve feature clarity, leading to improved performance of the proposed system and its ability to generalize in detecting cardiovascular diseases. After ECG images pre-processing, the dataset is split into 70% training set and 30% testing and validation sets, as depicted in Figure 3.

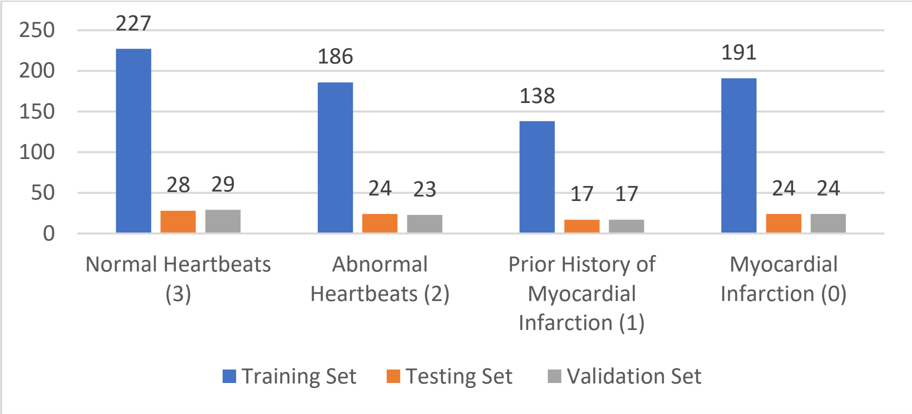


Figure 3: ECG Images Dataset Splitting.

ECG Images Augmentation

Augmentation techniques can be implemented to increase the diversity of datasets and enhance the generalizability of CNN models. Because ECG images vary depending on the equipment, imaging conditions, and patient movement, augmentation assists in simulating these potential variations and inhibiting the model from overfitting.

Although ECG images are structured, they may show slight variations in scale, direction, and location. Chosen augmentation techniques provide realistic transformations without changing the fundamental shape of the ECG waveform, ensuring that clinical features of ST-segment, QRS complex, and P-wave remain intact while improving the model's capability in abnormality recognition under various imaging situations. In this proposed system, selected augmentation techniques were implemented during training to imitate real-world variability in ECG signal morphology to ensure that each epoch obtained a unique and varied set of ECG images, thus enhancing the performance of the proposed CNN model. After this stage, the total number of training samples becomes 1484, and the testing and validation samples remain unchanged, as shown in Figure 4.

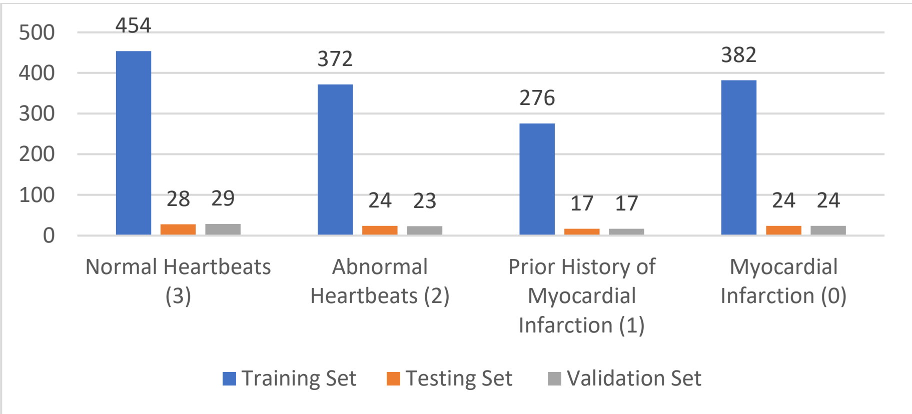


Figure 4: ECG Images Dataset Splitting After Augmentation.

ECG Images Feature Extraction

The proposed dilated-separable CNN model structure includes several convolutional units or blocks for the gradual extraction of higher-level distinctive features, accompanied by a regularization and classification stage. Dilated and wide separable convolutions with multiple-scale receptive fields are integrated to enhance the extraction of temporal patterns and maintain computational effectiveness.

- **Extracting Primary Features (Block 1):** This block extracts the low-level (basic) temporal and spatial features, including wave contours and edges, from the input ECG images. It incorporates two successive layers of two-dimensional convolutions, each containing 32 filters and a 3×15 kernel size, along with batch normalization for stabilizing training and accelerating convergence.

Furthermore, a pooling layer of size "2×2" is implemented to decrease dimensionality and complexity of computation.

- **Dilated Temporal Feature Learning (Block 2):** This block utilizes dilated convolutions to improve the field of temporal receptive fields and obtain inter-beat ECG relationships without increasing computational cost. It incorporates two successive layers of two-dimensional convolutions, each containing 64 filters and a 3×21 kernel size, with a (1, 2) dilation rate, along with batch normalization after each convolution. Finally, a pooling layer is applied to obtain miniaturized feature maps. This block improves the ability of the CNN model in detecting arrhythmia patterns that appear over longer timescales.
- **Wide Separable Feature Learning (Block 3):** This block utilized separable convolutional layers to enhance model effectiveness and increase its ability to detect complex spatiotemporal features. It incorporates two wide separable convolution layers, each containing 128 filters and a 3×31 kernel size to obtain wider waveform structures. After each convolution, batch normalization is applied. The feature map resolution is further decreased by a pooling layer of size "2×2".
- **Deeper Dilated Feature Extraction (Block 4):** This block introduces a deeper convolution containing 192 filters and a 3×41 kernel size, utilizing a (1, 4) dilation rate to broaden the receptive field. This enables the model to incorporate long-range dependencies and subtle abnormalities spanning large time segments of the ECG signal.

Regularization and Classification

Following the final block of feature extraction, regulatory layers are applied to minimize overfitting and improve model performance. The first is a 0.15 two-dimensional spatial dropout layer that decorrelates the spatial feature maps. Additionally, global average pooling is utilized to convert the feature maps into a simplified representation while preserving channel information.

Considering the final classification, two fully-connected layers are employed. The first is a dense layer comprising 256 neurons and Rectified Linear Unit (ReLU), accompanied by a 0.3 dropout. The second is a final output layer (Softmax) representing four disease classes. This incorporation provides efficient generalization and suitable class-probability predictions to support diagnostic decision-making.

The proposed model is trained utilizing a loss function “categorical cross-entropy” with an adaptive Adam optimizer. Training process is carried out with 30 epochs and 32 batch size, with augmented ECG images created during runtime. During training, a learning rate schedule was applied to stabilize convergence. Additionally, early stopping was employed to halt training when the validation loss stopped improving for a specific number of successive epochs, thus minimizing the risk of disposable calculations and preventing overfitting. This stage and the previous stage are depicted in Figure 5.

	Layer (Type)	Output Shape	Parameters
	Input layer	(None, 256, 512, 3)	0
Block 1	Conv2D	(None, 256, 512, 32)	4,352
	Batch Normalization	(None, 256, 512, 32)	128
	Conv2D	(None, 256, 512, 32)	46,112
	Batch Normalization	(None, 256, 512, 32)	128
	Average Pooling2D	(None, 128, 256, 32)	0
Block 2 (Dilated Temporal)	Conv2D	(None, 128, 256, 64)	129,088
	Batch Normalization	(None, 128, 256, 64)	256
	Conv2D	(None, 128, 256, 64)	258,112
	Batch Normalization	(None, 128, 256, 64)	256
	Average Pooling2D	(None, 64, 128, 64)	0
Block 3 (Wide Separable)	Separable Conv2D	(None, 64, 128, 128)	14,272
	Batch Normalization	(None, 64, 128, 128)	512
	Separable Conv2D	(None, 64, 128, 128)	28,416
	Batch Normalization	(None, 64, 128, 128)	512
	Average Pooling2D	(None, 32, 64, 128)	0
Block 4 (Deeper + Dilation)	Conv2D	(None, 32, 64, 192)	3,023,040
	Batch Normalization	(None, 32, 64, 192)	768
Regularization and Classification	Spatial Dropout2D	(None, 32, 64, 192)	0
	Global Average Pooling2D	(None, 192)	0
	Dropout	(None, 192)	0
	Dense	(None, 256)	49,408
	Dropout	(None, 256)	0
	Dense	(None, 4)	1,028

Figure 5. Layers of proposed dilated-separable CNN model

Diagnostic Support Tool Deployment

To ensure the proposed system can be implemented in real-time, it has been incorporated into a diagnostic support tool that receives ECG images as input and produces probabilistic disease predictions. This promotes clinical use and supplies a practical means of implementing the model in healthcare settings. The software for cardiovascular disease diagnosis primarily includes three key functions: uploading a clear ECG image (with PNG, JPG, and JPEG supported formats), training a deep learning model, and performing data analysis and prediction. Figure 6 shows a brief overview of the software.

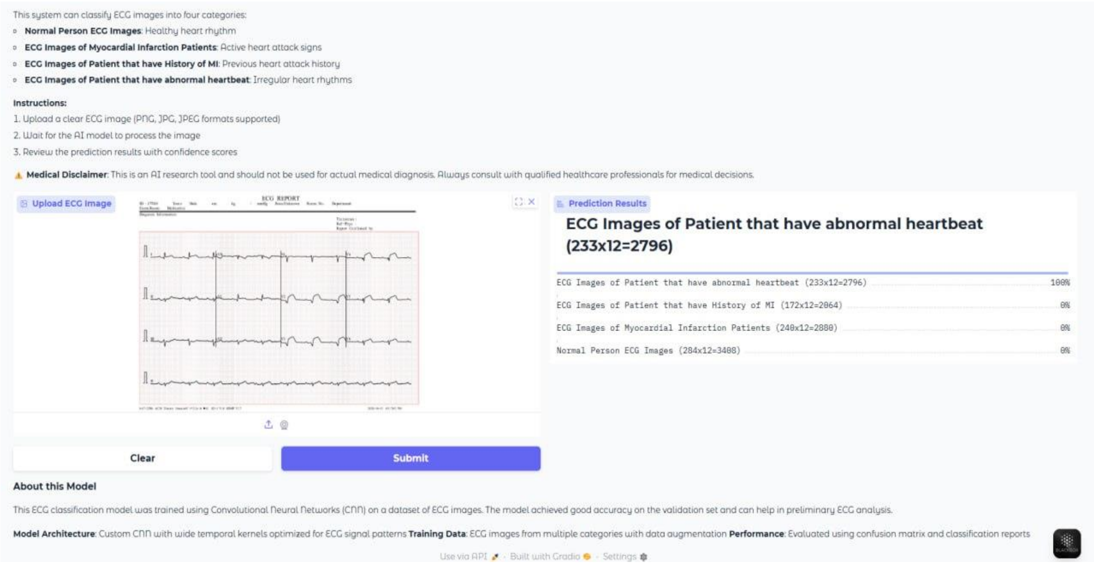


Figure 6: Diagnostic Support Tool Interface

Experimental Result Analysis

To assess the performance of the proposed cardiovascular disease detection system, various vital assessment measures are employed. The proposed system's task is to classify ECG images into four distinct classes; therefore, measures were used that precisely reflect the model's capability to classify each of these classes. Besides the proposed dilated-separable CNN model, the performance of three pre-trained CNN models (InceptionV3, EfficientNetB0, and ResNet50) was also assessed on the ECG images dataset to determine whether the proposed model could exceed common baseline models in the task of detecting cardiovascular disease. The measures used are Accuracy (Acc), Precision (Pre), Sensitivity (Sen), and F1 score ($F1$). The proposed CNN model achieved results of 99% in all assessment measures across all four classes, thus outperforming other baseline models applied. The results of these assessment measures are depicted in Table 3.

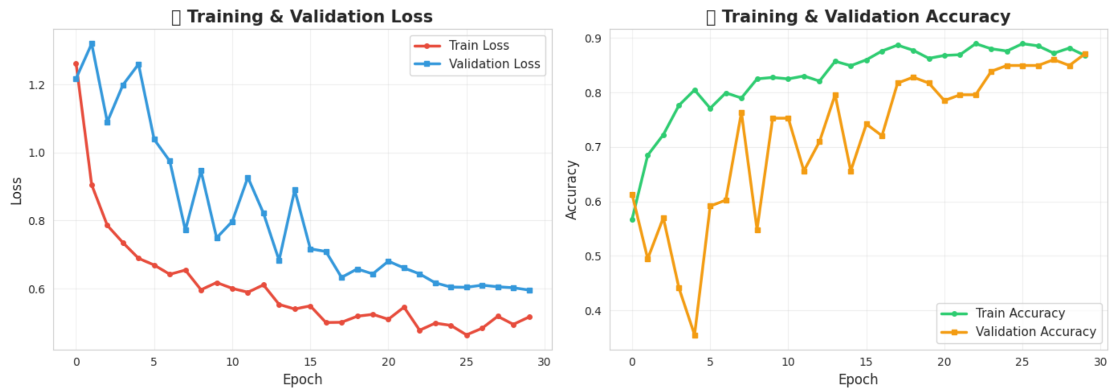
Table 3: Measures Assessment for the Proposed and Baseline CNN Models

	Proposed CNN Model			InceptionV3 Model			EfficientNetB0 Model			ResNet50 Model			Supported
	Pre	Sen	$F1$	Pre	Sen	$F1$	Pre	Sen	$F1$	Pre	Sen	$F1$	
0	1.00	1.00	1.00	0.96	1.00	0.98	0.81	0.92	0.86	0.87	0.83	0.85	24
1	1.00	1.00	1.00	0.78	0.82	0.80	0.93	0.76	0.84	0.76	0.76	0.76	17
2	1.00	0.96	0.98	0.96	0.96	0.96	0.95	0.83	0.89	0.85	0.96	0.90	24
3	0.97	1.00	0.98	0.92	0.86	0.89	0.74	0.82	0.78	0.85	0.79	0.81	28
Acc	0.99			0.91			0.84			0.84			93
M_{avg}	0.99	0.99	0.99	0.90	0.91	0.91	0.86	0.83	0.84	0.83	0.84	0.83	93
W_{avg}	0.99	0.99	0.99	0.92	0.91	0.91	0.85	0.84	0.84	0.84	0.84	0.84	93

The results of all assessment measures showed that the proposed dilated-separable CNN model exceeded other pre-trained models, performing outstanding classification accuracy and better processing of ECG classes. Furthermore, Pre, Sen, and $F1$ measures reflect the efficiency of the proposed model in detecting all cardiac conditions, especially abnormal heartbeats and myocardial infarction, which are crucial for early detection and intervention in medical settings.

To understand the learning behavior of the proposed model and baseline models, and to assess their efficiency in the training and validation processes, Figure 7 comprises the curves of training and validation losses and accuracies for these models. For the proposed CNN model, the realized training loss was 0.2601 and training accuracy was 0.9933, and the realized validation loss was 0.2600 and validation accuracy was 0.9892.

(a)



(b)

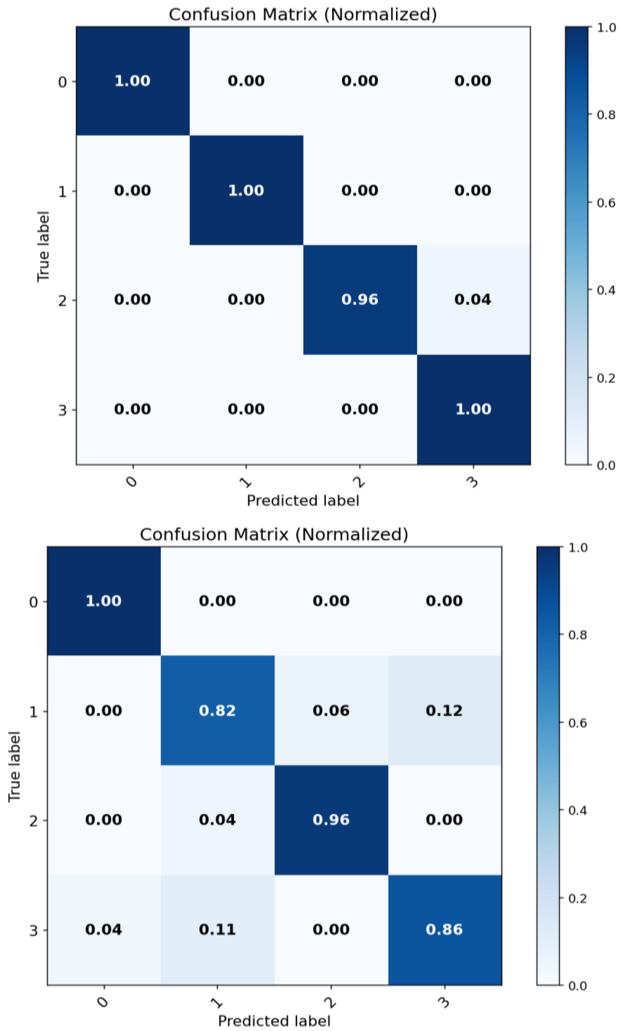
(c)

(d)

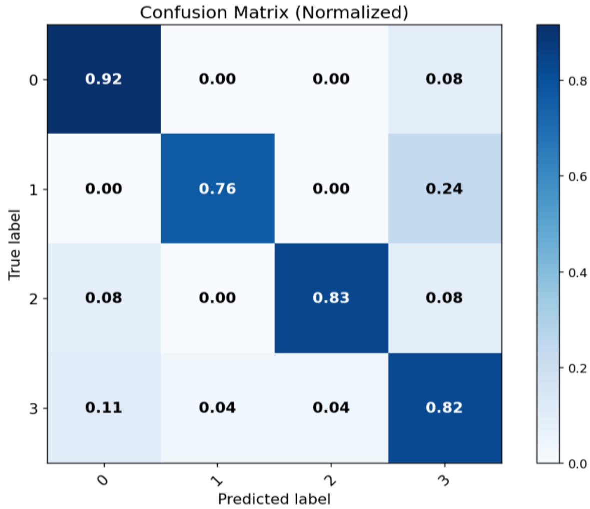
Figure 7. The learning behavior of (a) The proposed dilated-separable CNN, (b) InceptionV3, (c) EfficientNetB0, and (d) ResNet50 models, through the curves of training and validation losses and accuracies.

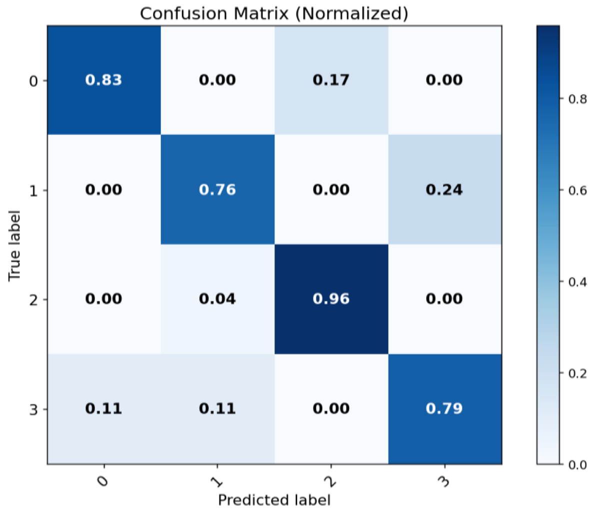
Compared to other baseline pre-trained models, the InceptionV3 model performed better in learning, achieving 87.10% validation accuracy and 88.95% training accuracy, with a lower validation loss of 0.5959 and a training loss of 0.4640. This reflects strong feature extraction capabilities, but with significant fluctuations in validation loss. The EfficientNetB0 model yielded 83.87% validation accuracy and 80.46% training accuracy, with a validation loss of 0.5573 and training loss of 0.6536. This indicates moderate learning ability but limited generalization performance. The ResNet50 model achieved 82.80% validation accuracy and 85.71% training accuracy, with validation loss of 0.5710 and training loss of 0.5496. ResNet50 achieved the training loss of 0.5496 and a verification loss of 0.5710, with training and verification accuracy of 85.71% and 82.80%, respectively. This indicates stable training but lower validation accuracy compared to the InceptionV3 model. On the contrary, the proposed CNN model significantly surpassed the pre-trained models, yielding the highest classification accuracy, with substantially low loss values and near-perfect convergence on validation and training sets. This outstanding performance can be attributed to the task-specific design of the proposed system, improved preprocessing and augmentation techniques, and the incorporation of early stopping, which effectively prevented overfitting and enhanced generalization.

The proposed CNN model offers superior classification performance, as illustrated in Figure 8, where high values along the matrix diagonal correctly classify most ECG instances.



(a) (b)





(c) (d)

Figure 8: Confusion Matrix results for (a) The proposed dilated-separable CNN, (b) InceptionV3, (c) EfficientNetB0, and (d) ResNet50 Models.

Conclusion

This article presents a novel CNN architecture specifically customized for detecting cardiovascular diseases utilizing ECG images. The proposed CNN model efficiently acquires spatial and temporal information included in ECG visual patterns by utilizing dilated, separable, and wide kernel convolutional layers. It achieved superior classification performance (exceeding 99% on all metrics), surpassing even the most advanced pre-trained CNNs and existing related architectures. These outcomes demonstrate that leveraging ECG image representations with an optimized deep learning model can allow an efficient and reliable foundation for computer-aided diagnosis of cardiovascular diseases.

The implementation of sophisticated pre-processing and augmentation methods has considerably enhanced the generalizability and robustness of the proposed model, guaranteeing consistent performance throughout various ECG classes. Moreover, the incorporation of a diagnostic assistance tool has confirmed the feasibility of utilizing such models in real-time clinical settings, sustaining the link between research and clinical application.

In future work, we will expand the dataset to incorporate multiple lead ECG images with comprehensive demographic diversity, thereby enhancing the system's generalizability to diverse clinical contexts. Furthermore, incorporating CNN with transformer-based models may improve the comprehension of temporal dependencies among ECG image sequences. Ultimately, combining the proposed system with wearable electrocardiogram imaging devices or the Internet of Things and expanding interpretable artificial intelligence visualizations can provide additional clinical confidence and facilitate real-world applications in cardiovascular healthcare.

References

[1] Hang Qiu, Ping Yang, Liya Wang, "DCNeT: A disease comorbidity network-based temporal deep learning framework to predict cardiovascular risk in patients with mental disorders," Expert Systems with Applications, vol. 254, 124312, 2024. https://doi.org/10.1016/j.eswa.2024.124312.

[2] Yuanhu Sun, Junqi Mao, Liang Cao, Xianjie Zheng, Qingjiang Meng, Song Zhao, Zhongkun Wang, Yuanzheng Zhang, Daxiang Cui, Haiwu Zheng, "Intelligent cardiovascular disease diagnosis system combined piezoelectric nanogenerator based on 2D Bi2O2Se with deep learning technique," Nano Energy, vol. 128, Part A, 109878, 2024. https://doi.org/10.1016/j.nanoen.2024.109878.

[3] Hongwen Ma, Wenhao Zhao, Ni Jin, Yineng Zheng, Xingming Guo, "Assessment for coronary heart disease severity by STFTs from synchronized PCG and ECG," Applied Acoustics, Vol. 240, 110940, 2025. https://doi.org/10.1016/j.apacoust.2025.110940.

[4] Mohamed Hammad, Mohammed ElAffendi, Ahmed A. Abd El-Latif, "CardioECGNet: A novel deep learning architecture for accurate and automated ECG signal classification across diverse cardiac conditions", Biomedical Signal Processing and Control, Vol. 106, 107720, 2025. https://doi.org/10.1016/j.bspc.2025.107720.

[5] Mehmet Akif Ozdemir, Gizem Dilara Ozdemir, Onan Guren, "Classification of COVID-19 electrocardiograms by using hexaxial feature mapping and deep learning", BMC Medical Informatics and Decision Making, Vol. 21, No. 170, 2021. https://doi.org/10.1186/s12911-021-01521-x.

[6] Taha Mohammed Hasan, Sahab Dheyaa Mohammed, Jumana Waleed, "Development of Breast Cancer Diagnosis System Based on Fuzzy Logic and Probabilistic Neural Network", Eastern-European Journal of Enterprise Technologies, Vol.4, No. (9 (106), pp. 6-13, 2020. doi: 10.15587/1729-4061.2020.202820.

[7] Jumana Waleed, Ahmad Taher Azar, Saad Albawi, Waleed Khaild Al-Azzawi, Ibraheem Kasim Ibraheem, Ahmed Alkhayyat, Ibrahim A. Hameed, Nashwa Ahmad Kamal, "An Effective Deep Learning Model to Discriminate Coronavirus Disease From Typical Pneumonia," International Journal of Service Science, Management, Engineering, and Technology, Vol. 13, No. 1, 2022. https://doi.org/10.4018/IJSSMET.313175.

[8] Saad Albawi, Muhanad Hameed Arif, Jumana Waleed, "Skin cancer classification dermatologist-level based on deep learning model", Acta Scientiarum. Technology, Vol. 45, No. 1, 2022. https://doi.org/10.4025/actascitechnol.v45i1.61531.

[9] Muhammad Faisal Abrar, Muhammad Saqib, Sikandar Ali, It Ee Lee, Arshad Khan, Salman Jan, Ahmed Alkhayyat, "Automated heart disease detection using Swin Transformer and ECG signal processing: a high-accuracy approach", Scientific Reports, Vol. 15, No. 39338, 2025. https://doi.org/10.1038/s41598-025-23097-6.

[10] Avinash L. Golande, T. Pavankumar, "Electrocardiogram-based heart disease prediction using hybrid deep feature engineering with sequential deep classifier", Multimedia Tools and Applications, Vol. 84, pp. 7443-7475, 2025. https://doi.org/10.1007/s11042-024-19155-2.

[11] Soumyashree Mangaraj, Kamalakanta Mahapatra, Samit Ari, "HLS-compiled PYNQ-based cardiac arrhythmia detection system leveraging quantized ECG beat images," Biomedical Signal Processing and Control, Vol. 109, 108063, 2025. https://doi.org/10.1016/j.bspc.2025.108063.

[12] S. Daphin Lilda, R. Jayaparvathy, "Enhancing cardiovascular disease classification in ECG spectrograms by using multi-branch CNN," Computers in Biology and Medicine, Vol. 186, 109737, 2025. https://doi.org/10.1016/j.compbiomed.2025.109737.

[13] Fatma Murat Duranay, Ender Murat, Oğuzhan Katar, Yakup Demir, Ru-San Tan, Özal Yildirim, U. Rajendra Acharya, "Arrhythmia detection in multi-channel ECG images: vision transformer and explainable approaches," Knowledge-Based Systems, Vol. 331, 114889, 2025. https://doi.org/10.1016/j.knosys.2025.114889.

[14] Ali Haider Khan, Muzammil Hussain, Muhammad Kamran Malik, "Cardiac Disorder Classification by Electrocardiogram Sensing Using Deep Neural Network", Complexity, Vol. 2021, Article ID 5512243, pp. 1-8, 2021. https://doi.org/10.1155/2021/5512243.

[15] Lotfi Mhamdi, Oussama Dammak, François Cottin, Imed Ben Dhaou, "Artificial Intelligence for Cardiac Diseases Diagnosis and Prediction Using ECG Images on Embedded Systems", Biomedicines, Vol. 10, No. 8, 2022. https://doi.org/10.3390/biomedicines10082013.

[16] Kaniz Fatema, Sidratul Montaha, Md. Awlad Hossen Rony, Sami Azam, Md. Zahid Hasan, Mirjam Jonkman, "A Robust Framework Combining Image Processing and Deep Learning Hybrid Model to Classify Cardiovascular Diseases Using a Limited Number of Paper-Based Complex ECG Images", Biomedicines, Vol. 10, No.11, 2022. https://doi.org/10.3390/biomedicines10112835.

[17] Kunwar Prashant, Prakash Choudhary, Tarun Agrawal, Evam Kaushik, "OWAE-Net: COVID-19 detection from ECG images using deep learning and optimized weighted average ensemble technique," Intelligent Systems with Applications, Vol. 16, 200154, 2022. https://doi.org/10.1016/j.iswa.2022.200154.

[18] Lerina Aversano, Mario Luca Bernardi, Marta Cimitile, Debora Montano, Riccardo Pecori, "Early Diagnosis of Cardiac Diseases using ECG Images and CNN-2D," Procedia Computer Science, Vol. 225, pp. 2866-2875, 2023. https://doi.org/10.1016/j.procs.2023.10.279.

[19] AbdulAdhim Ashtaiwi, Tarek Khalifa, Omar Alirr, "Enhancing heart disease diagnosis through ECG image vectorization-based classification", Heliyon, Vol. 10, No. 18, e37574, 2024. https://doi.org/10.1016/j.heliyon.2024.e37574.

[20] Md Nahid Hasan, Md Ali Hossain, Md Anisur Rahman, "An ensemble based lightweight deep learning model for the prediction of cardiovascular diseases from electrocardiogram images," Engineering Applications of Artificial Intelligence, Vol. 141, 109782, 2025. https://doi.org/10.1016/j.engappai.2024.109782.

[21] Ali Haider Khan, Muzammil Hussain, “ECG Images dataset of Cardiac Patients”, Mendeley Data, V2, 2021. doi: 10.17632/gwbz3fsgp8.2.