

## Enhancing Diagnostics: A Novel CNN-Based Method for Categorizing ECG Images with Attention Mechanism and Enhanced Data Augmentation



May S. Khorsheed\*<sup>ORCID</sup>, Abdulmir A. Karim<sup>ORCID</sup>

Computer Science Department, University of Technology, Baghdad 10011, Iraq

Corresponding Author Email: [cs.22.22@grad.uotechnology.edu.iq](mailto:cs.22.22@grad.uotechnology.edu.iq)

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.290532>

### ABSTRACT

**Received:** 16 April 2024

**Revised:** 20 September 2024

**Accepted:** 30 September 2024

**Available online:** 24 October 2024

#### Keywords:

*convolution neural networks, electrocardiogram, transfer learning, VGG16*

The paper showcases a consistent diagnostic process that fully automates the categorization of ECG images for cardiac diseases, including potential COVID-19-related complications. We propose a CNN-based model as the central component of this system. It has a customized attention mechanism and also uses advanced data augmentation and pre-processing techniques such as adaptive brightness, accurate resizing, and selective cropping. In this approach, we were concerned about the great variability in clinical ECG images, which can have adverse effects on data classification. This led us to design augmentation methods for this problem. We demonstrated the validity of the model by applying it with the help of the dataset, which has 1937 ECG images showing different heart abnormalities and a satisfactory classification score of 98.73%. Putting cardiovascular conditions at the core of AI applications demonstrates its ability to provide accurate treatment decisions. A well-proven automated system would be a milestone in the cardiovascular diagnostics community that would improve the efficiency and accuracy of disease diagnosis.

## 1. INTRODUCTION

An ECG, short for electrocardiogram, is a recording of the electrical activity of the heart [1]. ECG is a graphical representation of the heart's electrical activity during the depolarization and repolarization of its atrial and ventricular chambers. Depolarization refers to the abrupt entry of positively charged ions when the membrane becomes permeable. At the same time, repolarization is the subsequent phase in which the ion concentrations return to their usual levels [2]. Electrocardiography graphic districting has helped physicians tremendously diagnose and manage heart conditions before invasive procedures. The role of ECG analysis in cardiology is pivotal; it plays a significant role in both the detection and management of any heart-related problems. Ad Covid-19 disease represented difficulties for the physician because the virus can pathologically harm the cardiac muscle, as some ECG examinations show. The nature of the ECG report presents challenges for both manual and automated reading approaches. This can lead to less-perfect results from the ECG data analysis. To raise the accuracy of ECG images, especially in the COVID-19 situations that are standing in the way of other illnesses, we suggest using a system architecture that is based on the concept of convolutional neural networks (CNN). The model that we have chosen involves the application of a custom attention mechanism with a focus on the features of the ECG data and incorporates methods of data augmentation and preprocessing to address the problems of varying image quality. These techniques mostly comprise adaptive brightness, accurate

resizing, and selective cropping. The goal is to achieve high accuracy and reliability in ECG image analysis, enabling applications in cardiology research and enhancing the quality of clinical practice decision-making.

Using state-of-the-art machine learning tools, this system builds on the work from the aforementioned investigations [3-5] to address the issues discussed in these earlier works [6]. This system brings several noticeable changes in the precision of the ECG image classification, thus using them for COVID-19 diagnosing purposes. We were able to partially preprocess and improve ECG data by using a fully connected network's condition-focused attention mechanism and better training methods to help people learn. Upon careful assessment, our model demonstrates a significant advantage in accuracy, thereby significantly enhancing its practical application.

## 2. RELATED WORK

The success of up-and-coming diagnostics due to the deep learning approach improves the efficiency of electrocardiography analysis. The study uses CNN models with data augmentation and transfer learning, which is the newest technology from previous years. The objective is to find solutions for issues related to ECG signal classification and COVID-19 cardiac disease diagnosis.

Nonaka and Seita's [7] study, as well as research done by Gajendran et al. [8] explored the effectiveness of learning models when combined with transfer learning, and they came to the point that they can produce results close to those of

medical experts. These research works demonstrate the potential of using pre-trained models for *ECG* data, thereby addressing the issue of limited labeled sets.

Nonaka and Seita [7] carried out a study and produced an *ECG* signal classifier with the introduction of "*ECG* Augment." This method aims to expand the training data sets of the model, thereby enhancing its robustness. They applied the CinC/Challenge 2017 dataset, which contains *ECG* readings that devices have taken. They pointed out that the CNN of 15 ConvBlocks used in *ECG* analysis worked better with the ResNet method for feature propagation. As a result of their experiments, the researchers from the corresponding study demonstrated that the process of data augmentation helped to obtain higher classification accuracy. They also had a baseline F1 score improvement rate of up to 3.47 percent. Therefore, their method showed strong superiority. This implies that the learning model could perform better for the *ECG* classification tasks under investigation with the addition of more data.

Gajendran et al. [8], who performed a study, found that deep transfer learning could process *ECG* signals into different types of rhythm signals like sinus, cardiac failure, and arrhythmia with a good level of accuracy. Using 162 *ECG* records from the PhysioNet databases, the system transformed the acquired tridimensional *ECG* signals into two-dimensional scalogram images. This enables networks within the CNN family to perform classification tasks when trained using ImageNet. They experimented with CNN architectures and found that ResNet 50 trained to an accuracy rate of 94.12%, demonstrating good transfer learning capabilities for simplifying and automating medical diagnosis.

The research by Attallah [9] exhibits an automated tool that uses electrical cardiogram data as a clue to detect COVID-19. By considering *ECG* images from COVID-19 patients in addition to other health issues such as heart attacks, irregular heart rhythms, and normal cases, the tool can run ten different deep-learning models with different structures to get the important findings. The research validated the significance of deep learning algorithms as essential tools for precision, resulting in a high accuracy rate of 98.2% in binary classification and 91.6% in multiclass classifications. Moreover, the use of various convolutional neural networks (CNNs) for medical diagnosis presents the possibility for further AI research throughout the disease detection process.

Recent research has gone further and better than earlier studies that looked at *ECG* data by creating new models. For example, Fatema et al. [10] created a hybrid CNN and LSTM model for finding arrhythmias. The study used InceptionV3, ResNet50, MobileNetV2, VGG19, DenseNet201, and a combination model constructively known as InRes106. The dataset included 1932 paper-based *ECG* images. Among all the models presented, the InRes-106 has shown the highest performance, which is 98.34% in testing accuracy. The InceptionV3 model gave the highest test accuracy, i.e., 90.56%, ResNet50-89.63%, followed by DenseNet201-88.94%, VGG19-87.87%, and MobileNetV2-80.56.

More recently, Ahmed et al. [11] applied a 1D CNN model for arrhythmia classification, achieving an accuracy of 97.15% on the PhysioNet MIT-BIH Arrhythmia dataset. Elmir et al. [12] used Gramian Angular Fields (GAF) with CNNs for *ECG* classification, achieving 97.47% accuracy on the Arrhythmia dataset.

While these previous studies achieved impressive results, they focused primarily on standard CNN architectures and

conventional data augmentation techniques. Our approach builds on these advancements by introducing novel data augmentation techniques, such as adaptive brightness adjustment and selective cropping, which help address the significant variability observed in clinical *ECG* data, particularly in cases related to COVID-19 complications.

Additionally, the attention mechanism plays a crucial role in enhancing the model's stability and its ability to focus on key regions of the *ECG* signals. This focus helps the model better handle variability in clinical *ECG* data, ensuring consistent and reliable predictions in challenging scenarios.

Our model's classification accuracy of 98.73% surpasses the accuracy reported by Ahmed et al. [11] (97.15%), Elmir et al. [12] (97.47%), and earlier studies. This improvement is largely driven by our custom data preprocessing methods, which enhance the model's ability to handle complex clinical *ECG* data and improve diagnostic performance across diverse scenarios.

### 3. METHODOLOGY

#### 3.1 Dataset

The dataset used in this study, which consists of 1937 *ECG* images, was sourced from the openly available "*ECG* Images Dataset of Cardiac and COVID-19 Patients" by Khan et al. [13], published on Mendeley Data. This dataset contains *ECG* images of patients from different demographic backgrounds and is divided into five distinct classes, which are described below:

**COVID-19:** Most affected individuals experience certain shortness of breath and respiratory illness, and they can recover naturally or with medical help [14].

**Healthy:** A healthy person is one whose physiological activities and functions are not abnormally intercepted, and who does not suffer from any apparent weaknesses or deficiencies.

**Myocardial Infarction: MI,** also referred to as a heart attack is a form of acute coronary syndrome that results from a sudden or brief interception of the supply of blood to the heart, which makes the person feel many symptoms including chest pains and shortness of breath [15-17].

**History of Myocardial Infarction (HMI):** This class represents patients who have already suffered an MI and may be in the recovery or management phase of post-MI conditions.

**Abnormal Heartbeats (AHB):** This class includes *ECG* images from patients who have just recovered either from COVID-19 or from a myocardial infarction, showing symptoms of shortness of breath and respiratory illness.

The dataset spans *ECG* images between dimensions of  $952 \times 1232$  pixels and  $2213 \times 1572$  pixels. Being both small and imbalanced, some classes may be underrepresented, which would thus be conducive to biased model performance. Therefore, balancing and augmentation of the dataset were performed using appropriate techniques that could improve the accuracy of the classification results for better generalization.

#### 3.2 Validation strategy

In this work, an 80-20 train-test split was employed, with 60% of the data allocated for training, 20% for validation, and 20% for testing. This split was considered necessary to ensure that a good enough-sized dataset was available for training,

and at the same time, there was a dedicated test set used for independent testing.

K-fold cross-validation was not necessary in this respect because the fixed train-validation-test split is enough for a clear and consistent evaluation of the performance of the model. The dataset is augmented first by the brightness variation technique so that it increases variability in training effectively, hence reducing the possibility of network overfitting. Early stopping was used, stopping the training in the case of a non-improving validation loss for more than 5 consecutive epochs.

This indeed provided the right balance between training efficiency and robust model validation; therefore, this study did not require k-fold cross-validation.

### 3.3 Preprocessing and augmentation

As for our model, it was very important to spend enough time on the data preprocessing and enhancement procedures. We aimed to identify the most suitable procedures for our classifier to achieve the highest accuracy and reliability values. Table 1 gives a clear and detailed discussion of how preprocessing and augmentation affect the data. For more insight into the particular procedures involved, see Algorithm 2, which contains the different steps involved in the preprocessing and the data augmentation procedures. The first step of the preparation procedure was cropping the *ECG* images, which progressed from dimensions of 6% to 97% horizontally and 19% to 96% vertically to enhance the data points. Next, we made all images of the same size: 224x224 pixels, which is essential for the *CNN* structure.

**Table 1.** Overview of data preprocessing and augmentation techniques

Technique	Description	Purpose
Strategic Cropping	Cropping images within 6-97% of the horizontal range and 19-96% of the vertical range.	Focus on the most informative regions of <i>ECG</i> images, improving model training on relevant features.
Image Resizing	Resizing all images to 224x224 pixels.	Standardize input size for <i>CNN</i> , ensuring uniformity across the dataset.
Brightness Adjustment	Adjusting image brightness by +/- 5 units.	Simulate variations in recording conditions and enhance feature visibility.

### 3.4 Class imbalance mitigation techniques

To handle the class imbalance, oversampling and under sampling techniques were used such that all classes, other than the MI class, were equal in 400 images. The MI class was oversampled to 368 images. These numbers were chosen in order not to make the model overfit for each class. However, this prevents one class from dominating the model's learning due to overrepresentation [18], while underrepresented classes are oversampled to make sure such categories are not left out. The balanced approach ensures that the model does not take any biased turn toward majority classes and hence results in a more accurate classification outcome.

### 3.5 Data augmentation method

Class balancing was followed by augmentation of the data:

from 400 images per class to 800 and, for the MI class, from 368 to 736 augmented images. This included a  $\pm 5\%$  variation in brightness/darkness; this is an important augmentation in *ECG* image analysis, as it would simulate the natural variations occurring in the real world due to different equipment and environmental lighting. These slight variations provided by the augmentation method would help in much better generalization across different input conditions, preventing overfitting by providing a diverse range of examples to train upon.

---

#### Algorithm 1: Data Preprocessing and Augmentation

---

Input: A set of raw *ECG* images

Output: pre-processed and augmented dataset

Start

Step 1: For each image in the dataset, perform the preprocessing operations:

1.1. Static Crop: 6–96% horizontal region, 21–93% vertical region

1.2. Resize the image to 224 x 224 pixels.

End for

Step 2: For each image in the dataset, perform image augmentation:

Adjust the brightness between -5% and +5%.

End for

Return the pre-processed and augmented dataset.

End.

---

We carefully crafted the methods before running the *ECG* image classification model to address the challenges it faces in enhancing its ability to identify patterns in heart conditions, including those associated with COVID-19. The thoughtful implementation of these approaches highlights the thoroughness of our strategy to improve our *CNN* model's abilities.

### 3.6 CNN architecture

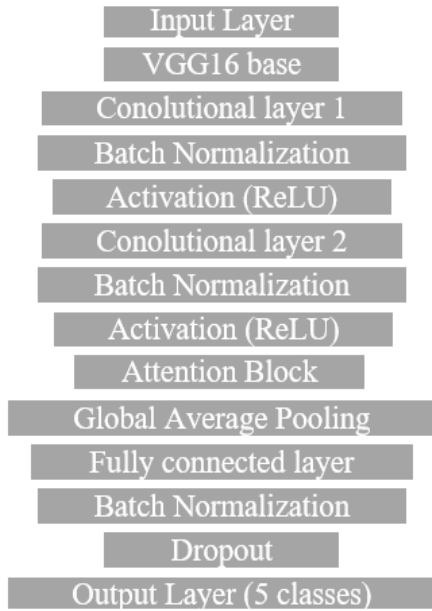
*CNN* is a multistage trainable neural network architecture sophisticated for classification tasks, and it is chosen for classification tasks [19, 20]. When developing our network (*CNN*) for classifying *ECG* images, we integrate the VGG16, which is a lightweight and straightforward model [21], which includes only 13 convolutional layers as a key component [22], using its pre-trained weights to extract advanced features. Known for its depth and consistent performance in image recognition tasks, VGG16 offers a set of features for identifying patterns in *ECG* images.

To customize VGG16 for our purpose, we start by locking all layers of the trained model. This method maintains the integrity of the features learned from the ImageNet dataset, ensuring that these generalized representations remain unchanged during our layer training. By keeping these layers fixed, we effectively employ VGG16 as a feature extractor, where the initial layers capture image traits crucial for our focused analysis.

After integrating VGG16, our design expands with customized layers designed to enhance and interpret the extracted features within the *ECG* classification. Below, we outline these unique layers, as depicted in Figure 1: *CNN* Architecture and Custom Layers. For a detailed breakdown of each layer's type, configuration, and specific purpose within our architecture, refer to Table 2.

- Following the VGG16 model, we add layers tailored to

detect subtle patterns specific to *ECG* images. These new layers consist of 512 filters, and the dimensions of the receptive field—the region from which information is taken—



**Figure 1.** CNN architecture and custom layers

**Table 2.** CNN architecture and custom layers

Layer Type	Configuration	Purpose
Input	Shape: 224x224x3	Receive standardized <i>ECG</i> images as input.
VGG16 Base (Frozen)	Pre-trained on ImageNet	Extract foundational features from <i>ECG</i> images.
Conv2D custom	512 filters, 3x3, ReLU activation	Further, refine features specific to cardiac conditions.
Batch Normalization	-	Normalize the activations from Conv2D layers.
Attention Mechanism	-	Weight feature importance, focusing on salient parts of the image.
Global Average Pooling	-	Reduce feature maps to a vector, summarizing important features.
Dense	256 units, ReLU activation	Interpret the summarized features for classification.
Dropout	Rate: 0.4	Prevent overfitting by randomly omitting units from the dense layer during training.
Output	5 units, SoftMax activation	Classify the <i>ECG</i> image into one of five categories.

- Following each custom layer is a batch normalization step that standardizes the activations from the layer, enhancing training stability and speed.
- Using the ReLU activation function on batch normalization to add non-linearities that help the model see complex patterns in *ECG* data.
- Adding an attention mechanism to help the model focus

are directly determined by the diameters of the convolution kernels in *CNNs*, which are 3x3 [23], aligning with VGG16's structure while honing in on the intricacies for cardiac analysis. on features by giving them weighted importance based on earlier layer extractions. This lets us look at the parts of the *ECG* that are most likely to show heart problems or COVID-19 effects.

- Before reaching the classification stage, a global average pooling layer is used to streamline feature map dimensions and distill information into a format conducive to classification. The architectural design wraps up with layers that are responsible for interpreting the aggregated features and carrying out the final classification task, where the model distinguishes among five distinct classes. These layers incorporate L2 regularization to prevent overfitting and are then complemented by dropouts for regularization.

By keeping the VGG16 layers unchanged and incorporating our customized layers, the model benefits from a mix of general and specific feature extraction. This strategy ensures a foundation for recognizing patterns thanks to VGG16, while our enhancements fine-tune the models to focus on the unique characteristics of cardiac conditions seen in *ECG* images.

---

**Algorithm 2: CNN with Attention Mechanism**

---

Input: The pre-processed and augmented dataset

Output: Classification predictions

Start

Step 1: Load the VGG16 pre-trained model, excluding the top layer.

Step 2: Freeze the layers of the VGG16 model to prevent weights from being updated.

Step 3: Append custom convolutional layers on top of VGG16 for further feature extraction.

Step 4: Implement the attention mechanism:

4.1. Apply Global Average Pooling (*GAP*) to the feature maps from the last convolutional layer.

4.2. Use a dense layer to predict attention scores from the *GAP* output.

4.3. Multiply the original feature maps by the attention scores to focus on important features.

Step 5: Flatten the output and pass-through dense layers for classification.

Step 6: Compile the model with Categorical Cross-Entropy as the loss function and Adam as the optimizer.

Step 7: Train the model on the *ECG* dataset using defined training protocols.

Return: Model capable of classifying *ECG* images into 5 classes with an attention mechanism

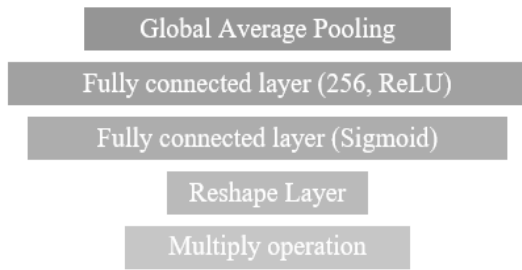
End.

---

**3.7 Incorporation of the attention mechanism**

The attention mechanism shown in Figure 2 is included in the model, allowing the network to concentrate on the most important aspects of the *ECG* images. This mechanism functions by assigning weights to sections of the feature maps created by the layers, thereby highlighting areas that are more pertinent to the classification task. To implement this mechanism, we start with a pooling layer, then move on to dense layers that generate a context vector, which we then multiply to adjust the feature maps. The decision to incorporate this attention mechanism serves two purposes: firstly, to enhance the model's interpretability by defining its focal points, and secondly, to boost classification accuracy by

minimizing the impact of uninformative or noisy regions within the *ECG* images.



**Figure 2.** Attention mechanism component

### 3.8 Improved training method

When it comes to training, we utilize improved training methods to enhance the effectiveness of learning, hence we opt for the Adam optimizer due to its ability to adjust the learning rate dynamically, making it easier to navigate the optimization landscape of neural networks. We set the starting learning rate to 0.001. Every 10 epochs, the step decay function decreases by 0.1. This gradual decline enables the model to make weight updates for quicker convergence, followed by smaller, more precise updates in later phases.

The concept of stopping involves monitoring the validation loss with a patience setting of 5 epochs. This implies that if there is no improvement in the validation loss for five epochs, the training process stops and reverts to the model weights from the epoch with the validation loss. To prevent overfitting, we ensure that the model does not continue learning from data noise beyond a point where it is beneficial. For a detailed overview of the specific training parameters employed, including their values and a description of each parameter's role in the training process, refer to Table 3.

This detailed explanation aims to offer insight into the *CNN* architecture, how we implement the attention mechanism specifically, and why we've chosen these training strategies. It demonstrates how all these elements come together seamlessly to enhance our model for *ECG* image classification.

**Table 3.** Training protocols

Parameter	Value	Description
Initial Learning Rate	0.001	Starting learning rate for the Adam optimizer.
Learning Rate Schedule	Step Decay	Reduce the learning rate by a factor of 0.1 every 10 epochs.
Early Stopping Patience	5 epochs	Stop training if validation loss does not improve for 5 consecutive epochs.
Batch Size	32	Number of samples per gradient update.
Epochs	200 (max)	Maximum number of epochs to run if early stopping criterion is not met.

## 4. RESULTS

On the test set, the *CNN* model demonstrated an accuracy

of 98.73%. It's worth mentioning that the model's precision, recall, and F1 score consistently outperformed the standards in all categories as outlined in Table 4.

**Table 4.** Precision, recall, and F1-score of the *CNN* model across classes

	F1-score	Recall	Precision
<b>COVID-19</b>	100%	100%	100%
<b>AHB</b>	97.44%	95.60%	99.35%
<b>Normal</b>	98.19%	98.19%	98.19%
<b>MI</b>	99.24%	100%	98.48%
<b>HMI</b>	98.68%	100%	97.40%

Evaluation for different categories of *ECG* data was performed, and it turned out that most classes had pretty consistent accuracy. However, a deeper analysis of each category's unique characteristics and the model's handling of those differences is as follows:

**COVID-19 Complications:** With 100% precision and recall, COVID-19 complications were identified. The abnormalities of COVID-19 complications can easily be highlighted due to their pronounced nature in the *ECG*. Adaptive brightness augmentation and attention mechanism allowed the model to effectively highlight the critical features that distinguished COVID-19 complications from the rest.

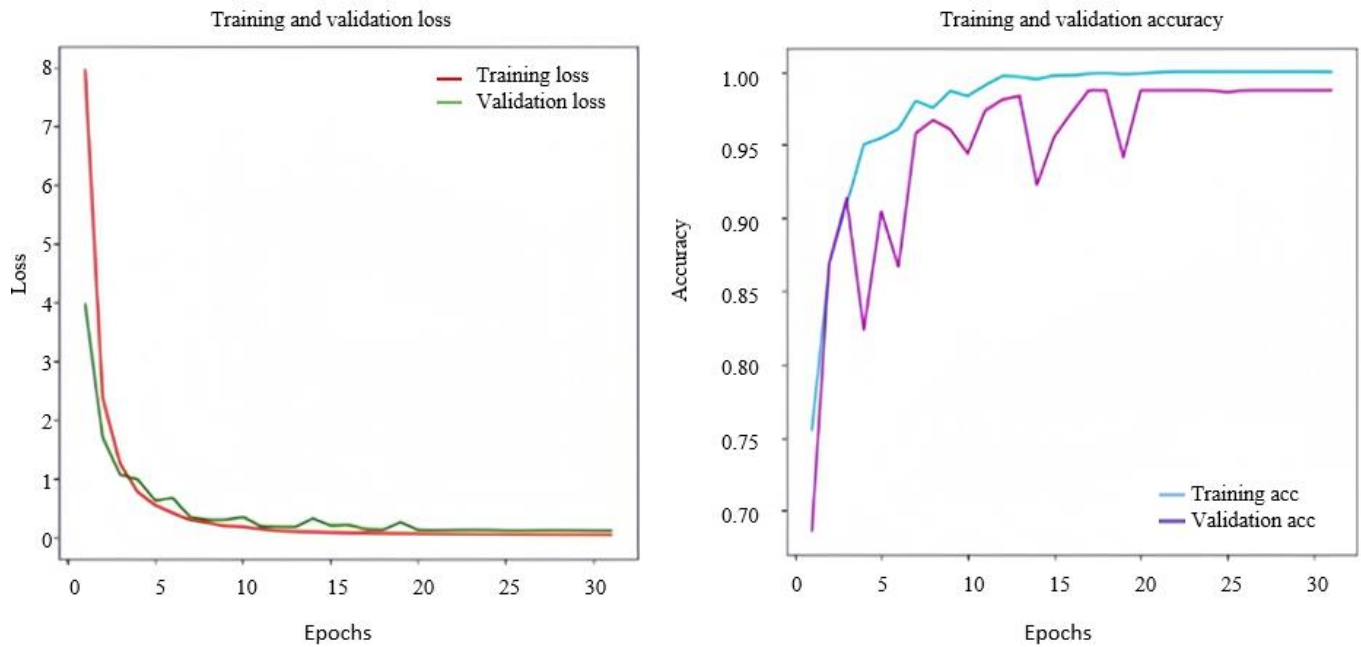
**MI:** The model is very accurate in detecting MI, with an accuracy of about 99.24%, due to its good feature capture for well-defined features such as ST-elevations marking the presence of MI. Application of tuned image preprocessing, especially resizing and selective cropping, made this model concentrate on critical regions of the *ECG*; hence, it was highly reliable in identifying this life-threatening condition.

**AHB:** Although the model did very well (97.44%) in this category, the subtlety of waveform variations that characterize AHB made this class a little more difficult to classify correctly. With a somewhat lower recall of 95.60%, it would appear the model at times missed these subtler abnormalities. This could be further improved by refinement of feature extraction or the addition of more nuanced examples to make the model sensitive at a higher degree toward such less obvious patterns.

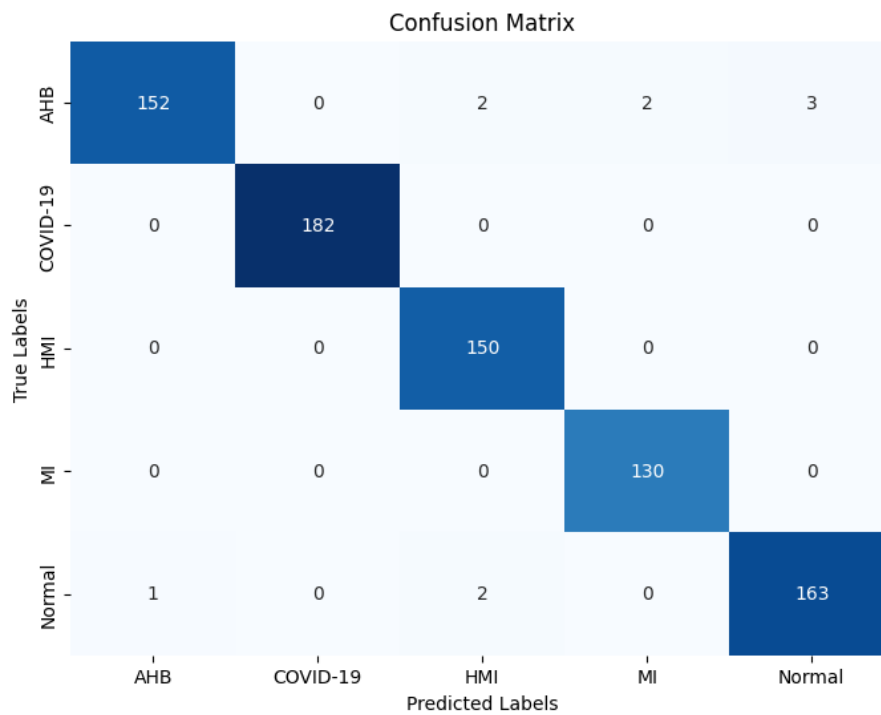
**Healthy cases:** It is observed that the model performed very well in classifying healthy subjects with 100% for both precision and recall. This may always affirm that the model really is much more robust and sure in separating normal from abnormal patterns for *ECG*; thus, this will be of high applicability in clinical decision-making.

**HMI:** The model performed quite well in tracing the history of myocardial infarction by detecting 98.68% precision. Myocardial infarction patients usually have residual signs in their *ECGs*, and the attention mechanism of the model was very helpful to identify these post-event, somewhat subtle variations.

Visual aids provide a picture of how the model is doing. Figure 3, illustrates the trends in loss and accuracy over time, showing how the model is learning. Also, Figure 4 shows a chart that helps us understand how well the model can predict outcomes in spotting COVID-19 cases. Together, these visuals highlight just how effective the *CNN* model is when it comes to classifying *ECG* images.



**Figure 3.** Loss and accuracy curves over epochs



**Figure 4.** Confusion Matrix for CNN model predictions

**Table 5.** Precision, recall, and F1-score of the traditional CNN model across classes

	F1-Score	Recall	Precision
<b>COVID-19</b>	100%	100%	100%
<b>AHB</b>	95.53%	95.00%	96.07%
<b>Normal</b>	96.51%	96.51%	96.51%
<b>MI</b>	100%	100%	100%
<b>HMI</b>	98.67%	99.33%	98.01%

The attentional CNN model proposed was compared with the traditional CNN without an attention mechanism in these experiments to show its effectiveness in improving the performance by the mechanism of attention. These

experiments involved using identical architectural settings, identical datasets, and identical conditions but varying in the presence of an attention mechanism.

The performance metrics of a traditional CNN model, devoid of attention, against the five classes are presented in the Table 5.

Although the attention mechanism yields only a modest improvement in numerical values between the CNN model with and without the attention mechanism for some classes, precisely COVID-19 and MI, where the performance for both models is perfect, what really plays a very important role is the attention mechanism's improved recall and F1-score for the more ambiguous classes like AHB and Normal.

**Table 6.** Precision, recall, and F1-score of the proposed CNN model without data augmentation

	<b>F1-Score</b>	<b>Recall</b>	<b>Precision</b>
<b>COVID-19</b>	98.78%	100%	97.59%
<b>AHB</b>	75.50%	67.06%	86.36%
<b>Normal</b>	81.66%	86.25%	77.53%
<b>MI</b>	91.43%	96.97%	86.49%
<b>HMI</b>	78.05%	78.05%	78.05%

For instance, the recall for the Normal class increased from 96.51% to 99.35%, and the F1-score for AHB was enhanced from 95.53% to 97.44%. This is significant in the medical context, where accurately detecting the presence of abnormal heartbeats—AHB—ensuring a minimum number of false negatives in normal cases is critical for patient safety and diagnosis accuracy.

Attention mechanisms will force the model to focus more on the salient features of an ECG image in a manner that it handles noisy or less-informative parts much better. This results in enhanced robustness for situations where data could be ambiguous or of varying quality, which usually holds with real-world clinical data.

To expand the results analysis part, an analysis was performed with respect to the model's performance without the application of augmentation methodologies. The results for this are underlined within Table 6, which sketches out the variance in precision, recall, and F1-scores of models with no augmentation applied.

Comparing the results, it is clear that most performance classes have increased significantly by augmentation among the CNN model. When no augmentation technique has been used, there is a sharp drop in recall for the classes AHB and HMI, indicating the remarkable loss in the model's ability to identify cases correctly for these classes. Thus, with no augmentation, precision and F1-scores also drop substantially for most of the classes, especially for the classes AHB, HMI, and Normal. By contrast, the COVID-19 class still maintains high recall without augmentation, although with a slight gain in precision when using augmentation. These findings highlight that augmentation improves the generalization of the model, especially those classes for which detection is more challenging, like AHB and HMI.

## 5. COMPLEXITY AND COMPUTATIONAL COST ANALYSIS

The proposed CNN model combines several state-of-the-art techniques, including a customized attention mechanism and data augmentation strategies, which are the probable reasons for its high classification accuracy. Nevertheless, there are other issues, such as the complexity of this model and its computational cost in practical application, especially in real-time or resource-constrained environments.

### A. Model complexity:

The complexity of the model can be understood from several perspectives: the number of layers, parameters, and operations involved. Our proposed model has an architecture comprising a total of 14 layers, of which 2 are convolutional backbone layers, followed by a custom-designed attention mechanism with dense layers. In all, it comprises 5.38 million trainable parameters. These layers help in providing fine details from ECG images for better classification of multiple cardiac conditions.

### B. Computational cost (Training and inference):

The training of this model on this dataset, which contained 3935 ECG images, required 696.86 seconds with the T4 GPU used for the training, or approximately 11.6 minutes. This model occupied a maximum of up to 8.3 GB of 15 GB GPU memory and system RAM of approximately 6.3 GB out of 12.7 GB at any moment in time during its training. This model is suitable for use in clinical real-time applications because the inference time was approximately 0.16 seconds (159.87 milliseconds) for an ECG image.

### C. Memory usage:

The model consumes about 8.3 GB of GPU RAM and about 6.3 GB of system RAM during training; this suggests that the model is quite well-optimized to run with, such as the T4 GPU with 15 GB of memory. For storage, a trained model occupies 76.57 MB on disk. These levels of memory usage are at manageable levels in modern clinical systems.

### D. Comparisons to existing models:

Compared to the existing state-of-the-art models, such as VGG16 or ResNet50, which are used for ECG classification, our model introduces a moderate increase in computational cost due to the attention mechanism and enhanced data augmentation. However, this addition boosts accuracy as high as 98.73% and robustness—which is critical in clinical decision-making. This trade-off between increased complexity and improved diagnostic performance justifies itself in high-accuracy applications such as cardiac health diagnostics.

## 6. DISCUSSION

This study demonstrates how a combination of preprocessing and data augmentation methods with a convolutional neural network that has an attention mechanism leads to effectiveness gains. As a result, the integration of modern methods has improved the precision and categorization of *ECG* images, which is an important diagnostic criterion for heart health. Furthermore, we aim to explore the significance of these findings and identify potential avenues for improving this model.

The proposed CNN model demonstrates high performance in classifying ECG images, particularly its strong capability in detecting MI with an accuracy of 99.24%. The model's robustness is further exemplified by its consistent performance across other conditions, including COVID-19-related complications and abnormal heartbeats. However, the slightly lower recall for certain conditions, such as AHB, suggests that while the model excels at identifying severe conditions, improvements are needed for recognizing subtle ECG abnormalities. These variations could be attributed to the inherent complexity of ECG patterns in those cases.

Below are discussed the strengths and weaknesses of the model, and how performance improvement may be achieved to enhance its applicability in the clinical setting.

### Strengths of the model:

A. Enhancing Model Accuracy via Customized Data Processing and Focused Image Enhancement: Cropping and resizing play a role in refining *ECG* image processing and aiding in feature extraction. Our method has notably boosted the model's capacity to detect patterns linked to heart conditions, including COVID-19, through fine-tuning image dimensions. This ensures that all data is precise, leading to accuracy across scenarios.

B. Moreover, by adjusting the brightness levels in our data

augmentation approach, we have equipped the model to handle the range of variations seen in clinical *ECG* recordings. Focusing on tweaking image brightness within a  $\pm 5\%$  range replicates real-world recording conditions, pushing the model to identify and grasp essential, consistent features linked to heart conditions. Proves effectiveness in improving the model's adaptability and ability to accurately classify cardiac problems across different patient groups.

C. When we add the attention mechanism to the *CNN* model, it can focus on specific regions of the *ECG* image for differentiation, potentially gaining attention: In addition, the narrow-focus model can boost its accuracy. Between removing the noise and filtering out the unnecessary features, the attention mechanism has remarkably increased the accuracy of the model's performance. This narrow practice enables the determination of the primary pathological factors and the identification of potential heart-related conditions that the COVID-19 infection may trigger.

D. Important features from *ECG* images are automatically extracted by the *CNN* model architecture with greater efficiency. It greatly simplifies the process of extracting critical features without manually engineered features, thereby enhancing scalability and boosting generalization across a wide range of datasets. It focuses on the most relevant aspects of the *ECG* data to give an accurate classification, and the capturing of key patterns independently of the mechanisms for external interpretability underlines its robustness and reliability in clinical settings.

### 6.1 Weaknesses of the model

(1) Detection of Subtle Conditions: Regarding the more subtle cardiac conditions, such as abnormal heartbeats, the model performed slightly poorer. This is probably due to the subtlety of the waveform that could only be fully expressed after further refinement among the feature extraction processes or core training on more nuanced examples.

(2) Evaluation on a Single Dataset: Although the performance of the proposed approach has been quite good, it has been evaluated on this single dataset. The study may thus be less generalizable regarding different *ECG* data or clinical settings. Additional validation of more datasets could be done to give a better picture of the robustness of the model.

## 7. COMPARISON WITH STATE-OF-THE-ART METHODS

In the following, we compare our approach against state-of-the-art approaches that have benefited from the *ECG* Images dataset of cardiac and COVID-19 patients for *ECG* classification. Several recent works, based on *CNN*-based architectures on this dataset, have focused on the classification of COVID-19, myocardial infarction, and other cardiac conditions. As such, they provide a good basis for the performance evaluation of our proposed model.

Shahin et al. [23] presented the results of different *CNN* architectures: VGG16, VGG19, InceptionResNetV2, and DenseNet201. Their study resource was directed at classifying multiple types of *ECG* images, including COVID-19, and other cardiac conditions. Among them, VGG16 yielded the best performance in the COVID-19 classification task with an

accuracy of 85.92%. While this provides a reasonable benchmark, the model performs much lower than the even higher accuracies achieved by more advanced models and approaches. The rather poor accuracy reflects the limitation of the *CNN* models, which have not been further optimized with state-of-the-art augmentation techniques or attention mechanisms.

In this regard, Hassan et al. [24] proposed a deep learning-based diagnosis model in 2023 by incorporating transfer learning along with ensemble learning for the classification of COVID-19 cases specifically among heart patients. After increasing the dataset to give a better balance of classes, this model yielded very good results in classifying heart patients with or without COVID-19. The proposed heart patient-based model presented the power of data augmentation and transfer learning techniques for improving model performance. Still, it was narrowly tuned to the discrimination between COVID-19 and non-COVID heart patients; thus, its generalization over the broad multi-class classification task is not possible.

Irmak [25] also performed classification of COVID-19 and myocardial infarction cases, using a *CNN* model, and reported remarkable accuracy in diagnosing COVID-19 equal to 98.57%. The high value of accuracy underlined the good perspective of how *CNN* models can perform on this data. However, the paper focused primarily on binary classification tasks rather than more general multi-class classification problems including several cardiac conditions.

In contrast, our proposed model leveraged attention mechanisms and advanced data augmentation techniques, driving superior performance for various cardiac conditions. Concretely, our model achieved 99.24% in terms of MI and 100% in COVID-19, outperforming the model proposed by Shahin et al. and showing very competitive performance compared to the model proposed by Irmak. Moreover, our approach has yielded more balanced performances between multiple classes, assuring much more suitability for diverse classification tasks. In the following Table 7, we summarize our model's performance and the performance of other state-of-the-art approaches.

This performance comparison underlines the robustness of our approach for handling multi-class classification tasks, in particular with MI and COVID-19, while offering state-of-the-art performance across all classes.

**Table 7.** Comparison with state-of-the-art methods

Study	Best Model	Accuracy (COVID-19)	Accuracy (MI)	Key Strengths
Shahin et al. [23]	VGG16	85.92%	N/A	Application of various <i>CNN</i> architectures
Hassan et al. [24]	Ensemble Learning (VGG19)	99.1%	N/A	Transfer learning and ensemble for heart patients
Irmak [25]	<i>CNN</i> model	98.57%	N/A	Focus on binary classification for COVID-19
Our Method	Attention-Based <i>CNN</i>	100%	99.24%	Superior accuracy with multi-class classification



## 8. CONCLUSION

Overall, the research demonstrates the power of *CNN* for the classification of *ECG* images, which in turn has improved accuracy in condition diagnosis and COVID-19 detection as well. We used VGG16, whose function is to introduce hidden features that generate medical picture concepts. This fundamental system serves as the foundation for constructing custom layers based on specific *ECG* image classifications.

The attention mechanism we come up with is targeted feature analysis. This mechanism condenses the information while accentuating aspects that are informative while at the same time leaving out less related ones, enabling the model to focus more on the more evident sections of the heartbeat rhythm.

Moreover, our proposed training technique utilizes a learning rate schedule associated with step decay, which linearly reduces the learning rate by 0.1 every 10 epochs. A key role in this technique is that it corrects the convergence of the model, which, on its way, enhances learning efficiency. Another important aspect to consider is the early stopping process and the patience level, which allows the model to stop training when the validation loss does not decrease further, thereby preventing overfitting and accelerating its process.

The research recognizes constraints. Wants to make sure all experiments are carried out effectively and proposes ways to improve this, such as widening the dataset scope, and training in a real-world clinical environment to verify the practical usefulness of the enhanced *CNN*.

## 9. FUTURE WORK

The proposed *CNN* model has been developed to improve its applicability for *ECG* analysis. Several avenues exist for further improvements and adaptations in the future, particularly optimizing the model for real-time use in the clinics and integrating it into the existing healthcare systems.

(1) Expanding the dataset to include a wider range of categories.

(2) Validation: Using *ECG* data from a variety of real-life situations for validation is necessary to give the model the empirical corrections that make it work much better and prove that it works in clinical settings.

(3) Improving interpretability: Developing the model's interpretability with graphs and other relevant information increases doctors' comfort zone around its diagnosis results.

(4) Enhanced Feature Extraction for the Subtle Conditions: The process of feature extraction in improving subtle cardiac condition detection can be refined. The inclusion of different preprocessing techniques in this regard, like advanced filtering might help the model catch minor variations in *ECG* waveforms.

(5) Tuning the Model for Real-Time *ECG* Analysis: Real-time *ECG* monitoring is vitally important clinically, in scenarios where immediate diagnosis is necessary, such as during surgery or in any critical care unit. In these real-time uses of the model, optimization for speed may be necessary. Techniques such as quantization and model pruning can be used to ensure faster processing time with not too much sacrifice of accuracy.

(6) Integration with Current Medical Systems: The model can be altered to be incorporated easily into a wide variety of hospitals and clinical environments by utilizing pre-existing

healthcare information systems and electronic health record platforms.

- Interoperability Standards: By adhering to the standard of communication in healthcare such as HL7 or FHIR, the model will easily be able to send and receive *ECG* data with hospital systems. This would then allow the model to be used as a plug-and-play diagnostic tool with other devices.

- API Development: Creating APIs that would enable other medical software systems to request *ECG* analysis with the availability of predictions could make the model more applicable in healthcare environments.

## REFERENCES

- [1] Omer, O.A., Salah, M., Hassan, A.M., Mubarak, A.S. (2022). Beat-by-beat *ECG* monitoring from photoplethysmography based on scattering wavelet transform. *Traitement du Signal*, 39(5): 1483-1488. <https://doi.org/10.18280/ts.390504>
- [2] Sadiq, A., Shukr, N. (2013). Classification of cardiac arrhythmia using ID3 classifier based on wavelet transform. *Iraqi Journal of Science*, 54(S4): 1167-1175.
- [3] Ebrahimi, Z., Loni, M., Daneshdab, M., Gharehbaghi, A. (2020). A review on deep learning methods for *ECG* arrhythmia classification. *Expert Systems with Applications*: X, 7: 100033. <https://doi.org/10.1016/j.eswax.2020.100033>
- [4] Al Rahhal, M.M., Bazi, Y., Al Zuair, M., Othman, E., BenJdira, B. (2018). Convolutional neural networks for electrocardiogram classification. *Journal of Medical and Biological Engineering*, 38: 1014-1025. <https://doi.org/10.1007/s40846-018-0389-7>
- [5] Liu, X., Wang, H., Li, Z., Qin, L. (2021). Deep learning in *ECG* diagnosis: A review. *Knowledge-Based Systems*, 227: 107187. <https://doi.org/10.1016/j.knsys.2021.107187>
- [6] Choi, S., Seo, H.C., Cho, M.S., Joo, S., Nam, G.B. (2023). Performance improvement of deep learning based multi-class *ECG* classification model using limited medical dataset. *IEEE Access*, 11: 53185-53194. <https://doi.org/10.1109/ACCESS.2023.3280565>
- [7] Nonaka, N., Seita, J. (2020). Data augmentation for electrocardiogram classification with deep neural network. *Electrical Engineering and Systems Science*. arXiv:2009.04398. <https://doi.org/10.48550/arXiv.2009.04398>
- [8] Gajendran, M.K., Khan, M.Z., Khattak, M.A.K. (2021). *ECG* classification using deep transfer learning. In 2021 4th International Conference on Information and Computer Technologies (ICICT), HI, USA, pp. 1-5. <http://doi.org/10.1109/ICICT52872.2021.00008>
- [9] Attallah, O. (2022). An intelligent *ECG*-based tool for diagnosing COVID-19 via ensemble deep learning techniques. *Biosensors*, 12(5): 299. <https://doi.org/10.3390/bios12050299>
- [10] Fatema, K., Montaha, S., Rony, M.A.H., Azam, S., Hasan, M.Z., Jonkman, M. (2022). 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*, 10(11): 2835. <https://doi.org/10.3390/biomedicines10112835>
- [11] Ahmed, A.A., Ali, W., Abdullah, T.A., Malebary, S.J.

- (2023). Classifying cardiac arrhythmia from ECG signal using 1D CNN deep learning model. *Mathematics*, 11(3): 562. <https://doi.org/10.3390/math11030562>
- [12] Elmir, Y., Himeur, Y., Amira, A. (2023). ECG classification using deep CNN and Gramian angular field. In 2023 IEEE Ninth International Conference on Big Data Computing Service and Applications (BigDataService), Athens, Greece, pp. 137-141. <http://doi.org/10.1109/BigDataService58306.2023.00026>
- [13] Khan, A.H., Hussain, M., Malik, M.K. (2021). ECG Images dataset of Cardiac and COVID-19 Patients. *Data in Brief*, 34: 106762. <https://doi.org/10.1016/j.dib.2021.106762>
- [14] Ciotti, M., Ciccozzi, M., Terrinoni, A., Jiang, W. C., Wang, C.B., Bernardini, S. (2020). The COVID-19 pandemic. *Critical reviews in clinical laboratory sciences*, 57(6): 365-388. <http://doi.org/10.1080/10408363.2020.1783198>
- [15] Saleh, M., Ambrose, J.A. (2018). Understanding myocardial infarction. *F1000Research*, 7: PMC6124376. <http://doi.org/10.12688/f1000research.15096.1>
- [16] Indira, D.N., Lakshmi, V.S.M., Markapudi, B.R., Yannam, A., Prasad, M.B., Babu, C.S., Rao, K.K. (2022). Detection of cardiac arrhythmia using multi-perspective convolutional neural network for ECG heartbeat classification. *Revue d'Intelligence Artificielle*, 36(4): 629-634. <http://doi.org/10.18280/ria.360416>
- [17] Fradi, M., Lazhar, K., Zahzah, E.H., Machhout, M. (2024). FPGA implementation of a CNN application for ECG class detection. *Traitement du Signal*, 41(1): 179-188. <http://doi.org/10.18280/ts.410114>
- [18] Wongvorachan, T., He, S., Bulut, O. (2023). A comparison of undersampling, oversampling, and SMOTE methods for dealing with imbalanced classification in educational data mining. *Information*, 14(1): 54. <http://doi.org/10.3390/info14010054>
- [19] Ahmed, W.S. (2020). The impact of filter size and number of filters on classification accuracy in CNN. In 2020 International conference on computer science and software engineering (CSASE), Duhok, Iraq, pp. 88-93. <http://doi.org/10.1109/CSASE48920.2020.9142089>
- [20] Aalaa Abdulwahab, H.A., Ali, Y.H. (2020). Documents classification based on deep learning. *International Journal of Scientific & Technology Research*, 9(2): 62-66.
- [21] Abdulhadi, M.T., Abbas, A.R. (2023). Human action behavior recognition in still images with proposed frames selection using transfer learning. *International Journal of Online & Biomedical Engineering*, 19(6): 47-65. <https://doi.org/10.3991/ijoe.v19i06.38463>
- [22] Ali, A.S., Abdulmunem, M. (2020). Image classification with deep convolutional neural network using tensorflow and transfer of learning. *Journal of the College of Education for Women*, 31(2): 156-171. <https://doi.org/10.36231/coedw/vol31no2.1>
- [23] Shahin, I., Nassif, A.B., Alsabek, M.B. (2021). COVID-19 electrocardiograms classification using CNN models. In 2021 14th International Conference on Developments in eSystems Engineering (DeSE), Sharjah, United Arab Emirates, pp. 448-452. <http://doi.org/10.1109/DeSE54285.2021.9719358>
- [24] Hassan, A., Elhoseny, M., Kayed, M. (2023). A novel and accurate deep learning-based Covid-19 diagnostic model for heart patients. *Signal, Image and Video Processing*, 17(7): 3397-3404. <http://doi.org/10.1007/s11760-023-02561-8>
- [25] Irmak, E. (2022). COVID-19 disease diagnosis from paper-based ECG trace image data using a novel convolutional neural network model. *Physical and Engineering Sciences in Medicine*, 45(1): 167-179. <http://doi.org/10.1007/s13246-022-01102-w>

## NOMENCLATURE

AHB	abnormal heartbeats
CNN	convolutional neural network
ECG	Electrocardiogram
GAP	Global Average Pooling
HMI	history of myocardial infarction
MI	myocardial infarction