Evolution in Electrical and Electronic Engineering Vol. 3 No. 2 (2022) 720-727
© Universiti Tun Hussein Onn Malaysia Publisher's Office

**EEEE**

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/eeee
e-ISSN : 2756-8458

# A CNN System for Automatic Classification of ECG Signal Abnormalities

**Nor Misha Safia Mohd Ariff**[1], **Audrey Huong**[1]*

[1]Faculty of Electrical and Electronic Engineering.
Universiti Tun Hussein Onn Malaysia, 86400 Parit Raja, Johor, MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/eeee.2022.03.02.085
Received 12 July 2022; Accepted 28 September 2022; Available online 30 October 2022

**Abstract**: The objective of this research is to build a CNN system that can classify ECG data automatically. The creation of a CNN system for automatic classification of ECG signal abnormalities, which automatically identifies and classifies ECG signal abnormalities utilising, is one effort to build a medical technology concept. CNN is highly sought after to reduce these ECG classification mistakes that happen in this medical field. The medical expert uses a CNN classification system from Matlab to aid in the diagnosis of a heart issue. The development of the project must be monitored closely to make sure everything runs smoothly and according to plan. As a result, the objective of this research is to automatically categorise and describe abnormalities of the ECG signal patterns. The outcome demonstrates the reliability of the findings and information that may be utilised to classify and identify ECG abnormalities.

**Keywords**: Convolutional Neural Network, Electrocardiography

## 1. Introduction

For many years, cardiovascular diseases are a collection of ailments that have long been recognised by medical professionals as one of the major killers. Myocardial infarction is a type of cardiovascular illness (MI). Myocardial infarction is the prolonged inability of the cardiac muscles to contract which is normally referred to as a heart attack. The mortality rate of someone having a heart attack can be decreased if appropriate medication is given within an hour of the attack starting. Electrocardiograms (ECGs), the primary method for identifying the cardiovascular disease, are performed as the initial diagnostic procedure when a heart condition manifests. The electrocardiograph records the heart's electrical activity during the test, and the results are then displayed on a visual diagram that depicts the cyclical electrophysiological processes that occur in the cardiac muscle. The ECG record can help doctors determine whether a myocardial infarction has occurred. However, it is crucial to remember that manual identification of acute myocardial infarction has a sensitivity and specificity of 91% and 51%, respectively. Cardiologists could make better decisions if a computer-aided system was created

*Corresponding author: audrey@uthm.edu.my
2022 UTHM Publisher. All rights reserved.

to automatically detect MI. As a result, many studies on automatic MI detection have recently been done.

Recently, methods based on neural networks have been used because the categorization of heart abnormalities is not linear. An effective technique for the early identification of cardiovascular disease was developed in a recent study using a radial basis probabilistic neural network. The suggested technique has been tested for ECG interpretation and the identification of irregular heartbeats that the network has categorised as related disorders. Convolutional neural networks (CNN) and audio biometrics techniques are examples of machine learning and deep learning methodologies that researchers have successfully experimented with throughout the years. In this domain, CNN has been applied in this field to identify arrhythmias, identify coronary artery disease, and to categorise beats.

Several works in the past that are worth highlighting include ECG signal quality classification using a deep neural network in [1]. In addition, to identify MI, several researchers in [2] used up to an 11-layer CNN. It was demonstrated in [2] whether shallow convolutional neural networks can be applied to the treatment of inferior myocardial infarction. Other researchers [2] stated that by applying different filter widths within a similar convolution layer, the system can access to the right features from the signal patches of different lengths. The authors in [3] suggested a cardiovascular illness classification system based on the MLP (Multilayer perceptron) network and the CNN network in it. They compared the outcomes between the models, which used various classes but a similar data set, in particular. The MLP network used two classes: "arrhythmias" and "normal," while the 4-layer CNN used nine classes.

PhysioBank.com and kaggle.com provided the ECG data for the training/validation and test datasets, respectively. There is also other research that uses deep learning algorithms based on convolutional neural networks to classify cardiac illness based on the ECG output [4]-[7]. This research presented a method for automatically detecting heart illness by converting ECG data straight from the time domain to other domains using a CNN-based classification network.

## 2. Materials and Methods

The data used in the CNN training originates from the public (https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset). The best model is then put to use in the real world. The important steps involved in the completion of this project are shown in Figure 1. From the data from the public sources, the system will tune the data to help in pre-trained a network of some different datasets to boost the performance of the network. Then after that, the data can be classified throughout the CNN classifier to get the result data of ECG signal abnormalities.

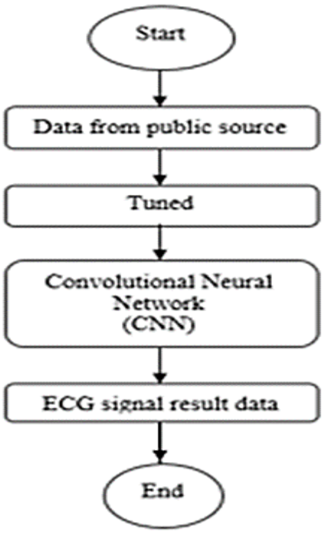


**Figure 1: The flowchart of the system**

As previously indicated, the proposed study used the Kaggle databases for performance evaluation, which are commonly used in the computerised classification of heart disorders based on ECG signals as a database. In order to evaluate classification accuracy, the data related to the neural network's training and testing were taken from this dataset. The network's accuracy suggested that it did a decent job classifying the two classifications connected to cardiac disease as well as the one relating to good health. It was able to test the proposed strategy using statistical classification functions based on the findings received from the confusion matrix.

The sensitivity was also stated as the true positive ratio (TPR), the specificity, also a term that is used to describe the true negative ratio (TNR), the fall-out, also known as the false positive ratio (FPR), and the test accuracy. As a result, each of the statistical classification parameters stated above might be defined that the classifier's specificity was determined by how frequently this might categorize the ECG recordings that weren't in that category; ECG records were accurately and widely accepted as belonging to a particular category, according to Fall-Out; sensitivity represented the proportion of ECG recordings successfully assigned to a particular group and related to that category; False discovery ratio showed that ECG recordings were really included in a class that was thought to be unrelated to them; ECG recordings were revealed to be a component of a category that was previously not thought to include them; false discovery ratio showed that ECG recordings were included in the class even though they were not thought to; The F1 score was calculated by dividing the number of true positives (TP) by the total number of positive results, which included both true positives (TP) and false positives (FP), and recovery was calculated by dividing the number of true positives (TP) by the total number of tests that should have resulted in a positive result, which included both true positives (TP) and false negatives (FN).

## 2.1 Training and Validation Set

A collection of data called "validation data" is kept apart from the training data. It is used to see how the network performs on data that hasn't been explicitly trained during the training phase. The "training data" is what is used in the backpropagation process to update the layer weights. Every iteration, training data is supplied into the network, the loss is computed, and the layer weights are adjusted using backpropagation to lower the loss. If the training is going well, the network's weights are updated with each iteration, making it better and better at predicting training data.

However, there is a risk that the network will grow too good at predicting from training data. It can learn extremely particular properties from the training set, rather than more generic traits that would help it forecast on fresh data. This is referred to as "overfitting." The "validation data" can be utilized to verify the data. The validation data is not used to train the network directly. Instead, it is utilized to monitor the network's performance.

The validation data results are not used to update the network weights. The validation accuracy is slightly lower than the training accuracy, as shown in the training plot - the network is better at predicting data it has been directly trained on. This is a common occurrence. More indication for overfitting would be if the validation accuracy started to decline significantly as it has been trained further the network would be learning specific features of the training set rather than general features also helpful for the validation set.

Text descriptions are susceptible to errors by both practising professional cardiologists and the labeling approach, and automatic categorization has a limited degree of accuracy. Following this initial stage, there are two situations in which the diagnosis must be accepted or rejected. They are: (i) both classifiers indicate the anomaly but not the expert; or (ii) only the expert indicates the abnormality but none of the classifiers do. Figure 2 shows the deep learning network analyzer.

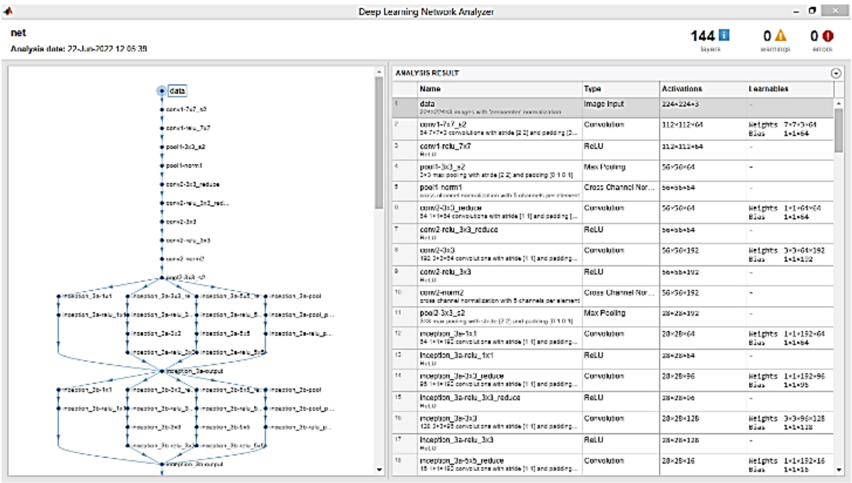


**Figure 2: Deep Learning Network Analyzer**

### 2.2 Convolution Neural Network

Google inception net has 144 one-dimensional convolutional layers with ReLU one layer that is fully connected and uses the SoftMax activation function and activation functions. The system generates a class distribution vector that has been standardized using the SoftMax function from inputs of ECG signals with 1000 samples per channel. The network locates the point of the vector's largest value to determine the class to which the input signal belongs. The class of this index is acknowledged as a category for input signals.

The architecture of a convolutional network. With the ReLU activation function, a twelve-channel ECG signal is processed through seven one-dimensional convolutional layers. The computation's results are flattened into a one-dimensional vector representation. A full-linked layer with a SoftMax activation function processes the calculation results. The result value is a one-dimensional vector that describes the input signal's probability distribution for each of the preset classes.

### 2.3 Test Set

The dataset used to evaluate the DNN came from the data ECG system as well. Two trained cardiologists with ECG experience separately annotated it. The dataset utilised to evaluate the DNN came from the ECG system as well. Two trained cardiologists with electrocardiography experience separately annotated it. When they came to an agreement, the common diagnosis became the ground truth. In cases in which there was a dispute, a third senior doctor who has been aware of the remarks from the other two made the diagnosis. The American Heart Association's standardisation served as the foundation for the classification.

## 3. Results and Discussion

### 3.1 Dataset Summary (Model Specification and Training)

A dataset was gathered from1988 that consisted of four databases containing information from 303 people from four different countries: Cleveland, Hungary, Switzerland and Long Beach V. Despite the fact that there are 76 trials in all, including the expected attribute, all published trials only use 14 of them. If the person has a heart illness, the “target” field indicates that. The methods for acquiring and annotation this dataset are given in Methods. A training set and a validation set were created from this dataset. 80% of the data is in the training set. The validation set is made up of the remaining 10% of the dataset, while the test set is made up of the remaining 10%.

### 3.2 Loss Function

The loss function is a metric for evaluating the prediction model's ability to anticipate expected outcomes. This study presents a TMSE loss function for the estimation of ECG data loss, as shown in Formula, based on the mean square error loss function (MSE). If the error value has an outlier, the loss model will give the outlier a higher weight in the mean square error function. The value of the loss function's computed error will skyrocket. Singular points and outliers will frequently arise in ECG data due to unpredictable noise interference during collection. We increase the tan function to control the error when taking the mean square error is too large or too small floats, and we can effectively suppress the influence of outliers on the entire model, get a more stable loss calculation, and adjust the direction and magnitude of gradient descent through the mapping of the tan function.

The suggested loss function will change the model to decrease outlier data points and improve and increase robustness to outliers. The proposed loss function has a considerable influence on enhancing the accuracy of ECG classification, as demonstrated by experimental comparison. A training and test loss function diagram is shown in Figure 3. There are outliers and unique points in the ECG data after the loss function, we proposed to control. The loss function curve eventually stabilises as a result of the mapping and management of the loss function, we provided, and the error float also tends to be stable. Figure 4 shows the result analysis of the process while Table 1 shows the accuracy chart.

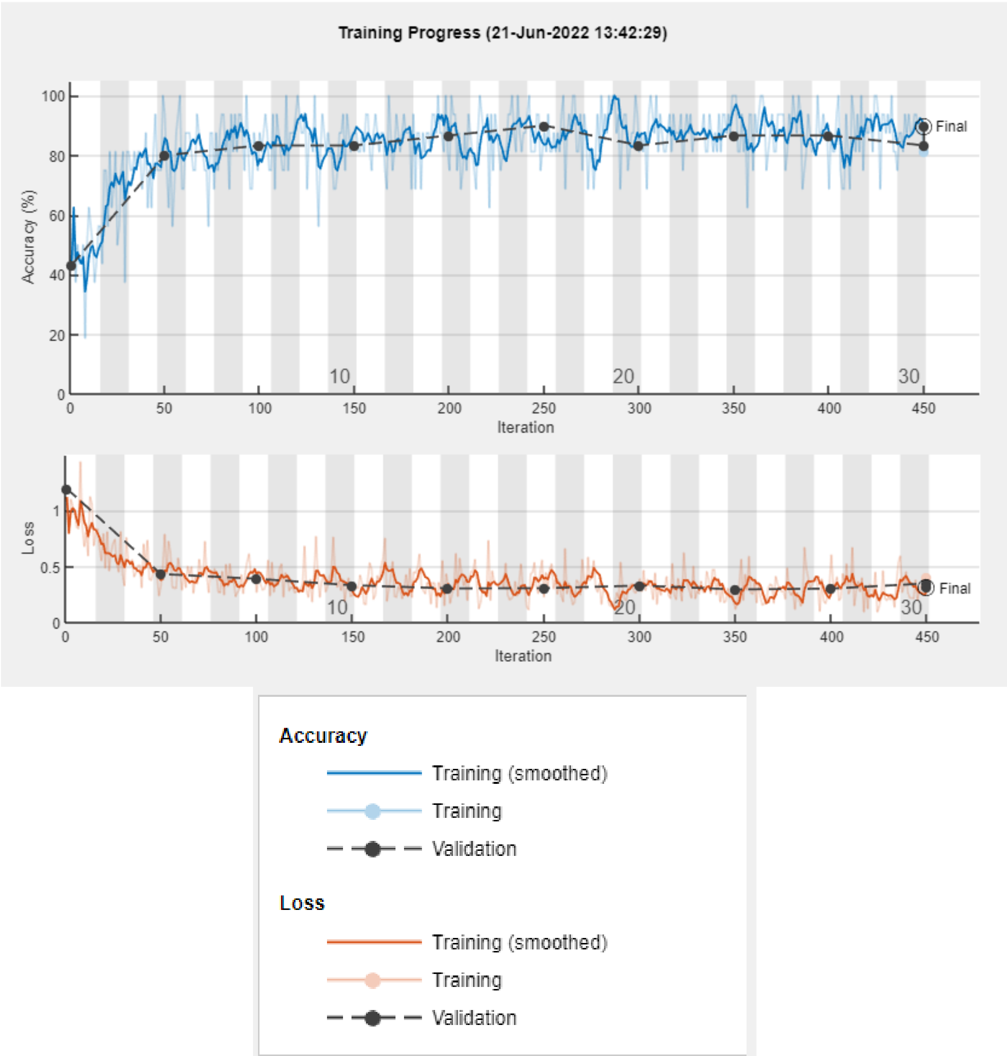


**Figure 3: Training Progress Graph**


**Figure 4: Result Analysis**

**Table 1: Accuracy Chart**

| Classification Method | Total Sample | Training Data | Testing Data | Accuracy |
|---|---|---|---|---|
| Deep Neural Network | 303 | 80% | 10% | 90% |

### 3.3 Evaluation Standard

In this experiment, three-level assessment indicators were used to assess the model's classification effect. The classification effect is displayed in the first-level evaluation using a confusion matrix (also known as an error matrix or Confusion Matrix). The confusion matrix is based on four parameters: the real value is positive, and the ECG classification model takes into account the quantity of positives (True Positive = TP). Although the true value is positive, the ECG classification model is negative (False Negative = FN). Although the true value is negative, the ECG classification model treats it as a positive (False Positive = FP). Because the true value is negative, the ECG classification model takes the number of negatives into account (True Negative = TN).

As shown in Figure 5, the four indicators are displayed in a table with a confusion matrix. The number is counted using the confusion matrix. Counting the numbers can make it difficult to assess the model's benefits and drawbacks. As a result, the confusion matrix in the fundamental statistical data extends the secondary index accuracy (Accuracy). The outcome of the quantity in the confusion matrix can be translated into a ratio between 0 and 1 using the secondary index. The accuracy rate in the second-level assessment index is used to evaluate the entire model, as shown in Equation (1), which is useful for standardised measurement.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

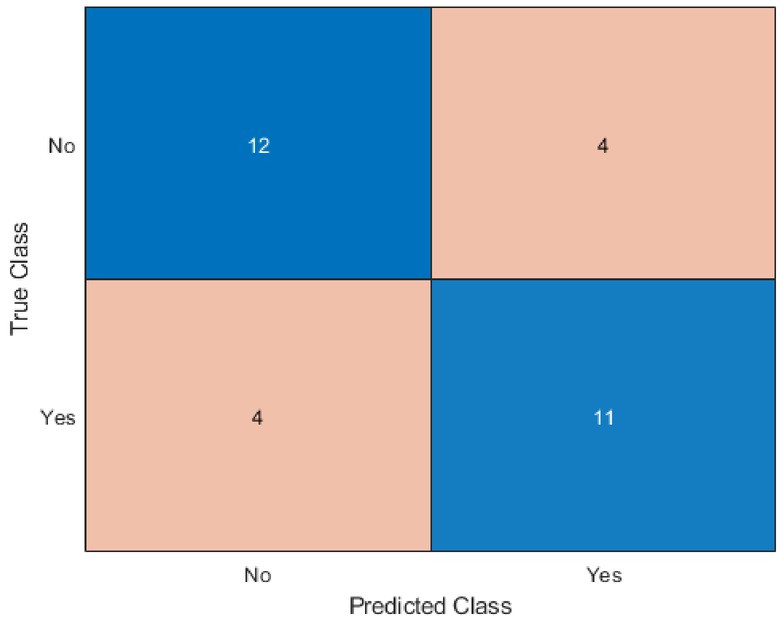


**Figure 5: Confusion Matrix**

The classification of the number of people with sickness which is heart disease as a result of the classifier. The number of prediction classes and true classes obtained from the CNN classifier result and result can be seen. The true negative (TN) demonstrates that 12 people were expected to be disease-free and they were found to be disease-free and healthy. In terms of false positives (FP), it appears that roughly four people have been diagnosed with a condition when actually they don't. The false negative (FN) shows that four people are expected to have an illness but they really have it, and the true positive (TP) shows that eleven people are predicted to have a condition and actually have it. From this confusion matrix, we can get the value of the sensitivity, specificity, accuracy, negative predictive value and also precision by using Equations (2) – (6). Table 2 shows that the accuracy of the manual calculation and the accuracy of the data is the same.

$$\text{Sensitivity} = \frac{TP}{(TP+FN)} \tag{2}$$

$$\text{Specificity} = \frac{TN}{(TN+FP)} \tag{3}$$

$$\text{Accuracy} = \frac{TP+TN}{(TP+TN+FP+FN)} \tag{4}$$

$$\text{Negative Predictive Value} = \frac{TN}{(TN+FN)} \tag{5}$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \tag{6}$$

**Table 2: Calculation of the Confusion Matrix**

| Sensitivity | Specificity | Negative Predictive Value | Precision | Accuracy |
|---|---|---|---|---|
| 0.733 | 0.75 | 0.75 | 0.733 | 0.7419 |

## 4. Conclusion

In the past, neural networks have been used to classify ECGs in both a traditional feature-based and an end-to-end learning setting. Hybrid approaches that combine the two paradigms are also available: classification can be done using a combination of handcrafted and learned features, or by utilising a two-stage training, with one neural network learning the features and the other classifying the exam based on the learned features. The shift in thinking toward end-to-end learning had a big impact on the quantity of datasets needed to train the models. Many traditional methods train their models on datasets with few samples, such as the MIT-BIH arrhythmia database, which has only 47 distinct patients. End-to-end deep learning or blended techniques are used in the most convincing articles.

**Acknowledgement**

The authors would like to thank the Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia for its support.

## References

[1] M. M. al Rahhal, Y. Bazi, M. al Zuair, E. Othman, and B. BenJdira, "Convolutional Neural Networks for Electrocardiogram Classification," Journal of Medical and Biological Engineering, vol. 38, no. 6, pp. 1014–1025, Dec. 2018, doi: 10.1007/s40846-018-0389-7.

[2] J. Li, Y. Si, T. Xu, and S. Jiang, "Deep Convolutional Neural Network Based ECG Classification System Using Information Fusion and One-Hot Encoding Techniques," Mathematical Problems in Engineering, vol. 2018, 2018, doi: 10.1155/2018/7354081.

[3] S. Azhar et al., "A SMRT-RPMS-APP: A Novel Remote Patient Monitoring System Featuring Vital Signs Via Smart Phones," Sci.Int.(Lahore), vol. 29, no. 1, pp. 65–71, 2017.

[4] J. Cheng, Q. Zou, and Y. Zhao, "ECG signal classification based on deep CNN and BiLSTM," BMC Medical Informatics and Decision Making, vol. 21, no. 1, Dec. 2021, doi: 10.1186/s12911-021-01736-y.

[5] S. Śmigiel, K. Pałczyński, and D. Ledziński, "ECG signalclassification using deep learningtechniques based on the PTB-XL dataset," Entropy, vol. 23, no. 9, Sep. 2021, doi: 10.3390/e23091121.

[6] R. Avanzato and F. Beritelli, "Automatic ecg diagnosis using convolutional neural network," Electronics (Switzerland), vol. 9, no. 6, pp. 1–14, Jun. 2020, doi: 10.3390/electronics9060951.

[7] S. Kusuma and K. R. Jothi, "ECG signals-based automated diagnosis of congestive heart failure using Deep CNN and LSTM architecture," Biocybernetics and Biomedical Engineering, vol. 42, no. 1, pp. 247–257, Jan. 2022, doi: 10.1016/j.bbe.2022.02.003.