iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** IX **Month of publication:** September 2021

DOI: https://doi.org/10.22214/ijraset.2021..38230

Heart Disease Prediction System

Anjana K. S
Department of Computer Science, APJ Abdul Kalam Technological University, Thrissur, India

Abstract: Heart diseases are the one of the primary reasons of human death today. There are many recent technologies are used to assist the medical professionals and doctors in the prediction of heart disease in the early stage. Prediction of heart disease is a critical challenge in the area of clinical data analysis. This paper introduces a technique to detect arrhythmia, which is a representative type of cardio vascular diseases. Arrhythmia refers to any irregular change from the normal heart rhythms, means that your heart beats too quickly, too slowly, or with an irregular pattern. The Electro Cardiogram (ECG) is used as an input for the arrhythmia detection. It displays the rhythm and status of the heart. This paper propose an effective ECG arrhythmia classification approach based on a deep convolutional neural network (CNN), which has lately demonstrated remarkable performance in the field of machine learning. It perform the classification without any manual pre-processing of the ECG signals such as noise filtering, feature extraction, and feature reduction.
Keywords: Arrhythmia, ECG, deep learning, CNN, ResNet

I. INTRODUCTION

Cardiovascular diseases (CVDs) are the leading cause of health globally, taking an estimated 17.9 millions lives each year. It is a group of disorders of the heart and blood vessels and include coronary heart disease, cardiovascular disease etc. The most important behavioural risk factors of heart disease and stroke are unhealthy diet, physical inactivity, tobacco use and harmful use of alchohol. Identifying those at highest risk of CVDs and ensuring they receive appropriate treatment can prevent premature deaths. This paper introduces a technique to detect arrhythmia which is a representative type of heart diseases. An arrhythmia describes an irregular heartbeat. With this condition, a person's heart may beat too quickly, too slowly, too early, or with an irregular rhythm.

There are several type of arrhythmia including Premature Ventricular Contraction (PVC), Right Bundle Branch Block Beat (RBB), Left Bundle Branch Block Beat (LBB), Atrial Premature Contraction (APC), Ventricular Flutter Wave (VFW), Ventricular Escape Beat (VEB) etc. Electro Cardigram (ECG) is used here as input. An ECG records the electrical signal in the heart. An ECG is a non-invasive medical tool. As a result, in the field of cardiology, automatic detection of abnormal cardiac rhythms from ECG readings is a major task. Two dimensional convolutional neural network(Conv2D) with gray scale ECG images is used here. It is usually used for image data. It is called 2D CNN because the kernel slides along the data in two dimensions. CNN can detect edges, color distribution, etc. in an image. The detection of arrhythmia disease is necessary. Therefore, numerous ways has been adopted for the purpose.

Jagdeep Singh, Amit Kamra, Harbhag Singh [1], in their work comprises the development of a framework based on associative classification techniques on heart dataset for early diagnosis of heart based diseases. This work is developed on the Cleveland heart diseases dataset from the University of California Irvine to test on different data mining techniques. There are many risk factors that causes the arrhythmia including age, gender, pulse rate, diabetics, blood pressure, etc. Various data mining algorithms are used in this paper. Examples of such algorithms are Apriori, FP-Growth, Naive bayes, ZeroR, OneR, J48 and k-nearest neighbor. These algorithms are applied in their study for prediction of heart diseases. Data mining techniques has been successfully used in different areas for knowledge discovery. This paper discusses different association and classification techniques are used to detect the type of arrhythmia.

S.P Rajamohana, C Akalya devi, K.Umamaheswari [2], proposed a method to decrease the expense of prediction and to predict the heart disease in an easy way. The main aim of this paper is to analyze various papers of research that done on heart disease prediction and classification using various machine learning and deep learning techniques. They are: data mining, Machine learning, neural network, and deep learning. Also they found some limitations of the previous works. For example, the major issue of data mining technique is method overloading. Abhishek Rairikar, Vedant Kulkarni [3], presented a robust methodology that uses a precise and scalable approach by using KNN, Data Mining and Genetic Algorithm. Because of the healthcare industry collects large amounts of Healthcare data, but not all the data are mined which is required for discovering hidden patterns and effective decision making they propose a method of efficient genetic algorithm with the back propagation technique approach for heart disease prediction.

Divya Krishnani, Anjali Kumari, [4] presented a method that would overcome these challenges which performs extensive experiments with multiple architectures which could deal with various diseases. Machine learning and Big data analytics methods are used to implement this work. Machine learning and Big data analytics methods have been proven to provide promising solutions to biomedical communities, healthcare problems, and patient care. They propose a extensive method to predict the Coronary heart disease (CHD). Aditya Baksh, MamtaRani [5], proposed a robust method for Heart disease prediction using artificial neural network and image processing. This work explores various algorithms and evaluates them for the task of identifying the disease in a dataset. Different image processing techniques are proposed in this paper. Image processing techniques are used for getting useful information from medical reports of patients using artificial neural network.

Sena Yagmur and Nalan Ozkurt [6], proposed a method using the electrocardiography (ECG). The ECG arrhythmias have been classified by using the deep neural networks in order to features information. This technique operate with a large volume of raw ECG time-series data and ECG signal spectrograms. And these ECG data and ECG signal spectrograms are used as inputs to a deep convolutional neural network (CNN). Heartbeats are classified as normal (N), premature ventricular contractions (PVC), right bundle branch block (RBBB) rhythm by using ECG signals obtained from MIT-BIH arrhythmia database. Only four type of heartbeat is considered here. The work presented by Ibrahim Abdel-Motaleb and Rohit Akula [7], on detecting the heart disease named Artificial Intelligence Algorithm for Heart Disease Diagnosis using Phono cardiogram Signals. In this work an Artificial intelligent system has been developed using the artificial neural network to diagnose heart disease from Phonocardiogram (PCG) signals.

Sushmita Manikandan [9] , introduced a paper for Heart attack prediction system. This system introduced a technology to reduce the efforts and time put in by the doctor. The main impression this work by Abder rahmane Ed-daoudy and Khalil Maalmi [8] is that this work is the streaming big data analytics and machine learning technology that can have a significant impact in healthcare field especially early detection of heart disease. Amin Ul Haq, Jianping Li, Jianping Li [10], presented a paper namely Heart disease prediction system using model of machine learning and sequential backward selection algorithm for features selection. The heart disease treatment and recovery is easy if detected the disease at early stages.

II. METHODOLOGY

With the advancement of machine learning techniques, detection of ECG arrhythmia has become more easier. Taking into account various approaches identified from the literature, an approach is proposed which detect the seven types of arrhythmia using deep two dimensional convolutional neural networks. The development of the model has been divided into different sections based on context. In this section, the overall system architecture of the proposed technique is described.

A. *Overall Architecture*

The figure below shows the overall architecture of the system.

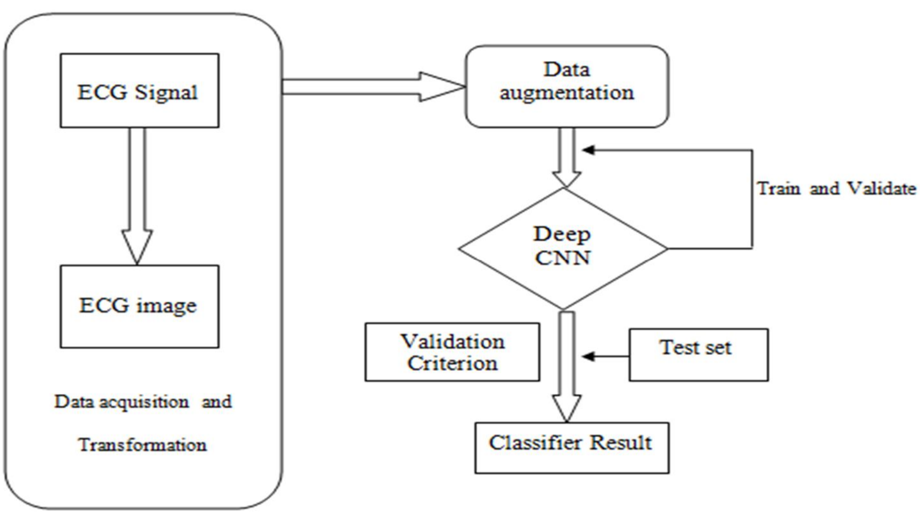


Figure 1. Overall Architecture

The figure 1 shows the main steps of the system. They are, ECG data preprocessing and ECG Arrhythmia classifier.

1) *ECG Data Preprocessing:* In this paper, MIT-BIH arrhythmia database is used for the CNN model training and testing. MIT-BIH Arrhythmia dataset contains the ECG signals as the input data. It is difficult to deal with that kind of data. Because of the CNN model handles two-dimensional image as an input data. So, here these ECG signals are converted into ECG images. This conversion is done in the ECG data preprocessing step. These ECG recordings with every single ECG beat are transformed into grayscale image. This transformation is done by plotting each ECG beat as an individual grayscale image. The reason why the CNN model is used, this model can automatically ignore the noise data when learning the model. This work considered seven ECG beat types including the normal beat.
2) *Data Augmentation:* Since the data was highly imbalanced because we have majority of normal heart beat and minority affected or non normal heart beat. So, if we don't do data augmentation, the model will be biased. Since we are using the image as the input data, the data augmentation is one of the key benefits of this system. The majority of previous ECG arrhythmia studies were unable to manually add augmented data to the training set since the distortion of a single ECG signal value could degrade the test set performance. Here the data augmentation can reduce the overfitting and that provide a balanced distribution between the classes. It is possible because of the CNN model is used as the classifier. The overfitting happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. The augmented images are produced inside the model. The Resnet50 model is used to implement the convolutional neural network. This model has 50 layers deep.
3) *ECG Arrhythmia Classifier:* The convolutional neural network is used as the ECG arrhythmia classifier. 2D convolutional and pooling layers are more suitable for filtering the spatial locality of the ECG images. That is why the CNN model is used here. There are many CNN models are used. Here the Resnet50 model is used. The Figure 2 shows the proposed CNN architecture.

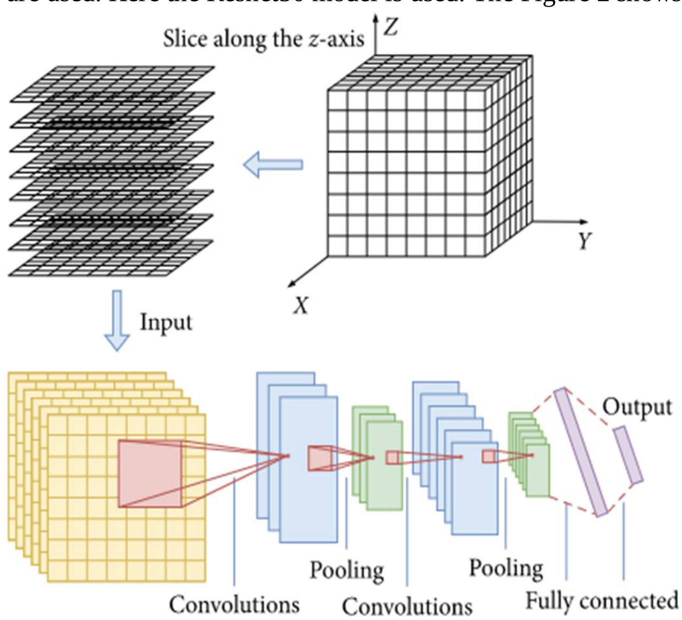


Figure 2. The Proposed CNN Architecture

The process of extracting valuable features from an image is performed in the convolution step. It has several filters that perform the convolution operation. Every ECG image is considered as a matrix of pixel value. After the convolution operation it applies the activation function. The ReLU (Rectified Linear Unit) activation function is applied here. ReLU is linear for all positive values, and zero for all negative values.

The next step in the CNN is pooling. The pooling is a down sampling operation that reduces the dimensionality of the feature map. The rectified feature map now goes through a pooling layer to generate a pooled feature map. The next step in the process is called the flattening operation. Flattening is used to convert all the resultant two dimensional arrays from pooled feature maps into a single long continuous linear vector. The flattened matrix is fed as input to the fully connected layer to classify the ECG image.

Resnet is short for Residual Network, is a classic neural network used as a backbone for many computer vision tasks. It consists of 5 stages each with a convolution and identity block. Each convolution block has 3 convolution layers and each identity block also has 3 convolution layers. The Resnet50 has over 23 million trainable parameters. The Resnet is a powerful backbone model.

III. RESULT AND ANALYSIS

After the model has been built, the test dataset was used to classify them into respective classes. The implementation results and the respective performance analysis is displayed using screenshots in the following section.

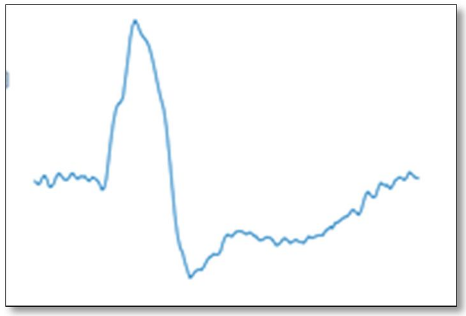


Figure 3. Example Of Image Dataset Creation Output - LBB

The Figure 3 shows an example of output that produce after the image dataset creation. It is the output after converting the ECG beat signals in the MIT-BIH arrhythmia dataset into the ECG images .Here shows the example of LBB arrhythmia.

Model: "resnet50"

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 128, 128, 3)	0	
conv1_pad (ZeroPadding2D)	(None, 134, 134, 3)	0	input_1[0][0]
conv1_conv (Conv2D)	(None, 64, 64, 64)	9472	conv1_pad[0][0]
conv1_bn (BatchNormalization)	(None, 64, 64, 64)	256	conv1_conv[0][0]
conv1_relu (Activation)	(None, 64, 64, 64)	0	conv1_bn[0][0]
pool1_pad (ZeroPadding2D)	(None, 66, 66, 64)	0	conv1_relu[0][0]
pool1_pool (MaxPooling2D)	(None, 32, 32, 64)	0	pool1_pad[0][0]
conv2_block1_1_conv (Conv2D)	(None, 32, 32, 64)	4160	pool1_pool[0][0]

Figure 4. Resnet50 Model

The Figure 4 shows the Resnet50 model. It consists of 50 layers. Only few of them are shown in this figure.

Model: "sequential_1"

Layer (type)	Output Shape	Param #
resnet50 (Functional)	(None, 4, 4, 2048)	23587712
flatten_1 (Flatten)	(None, 32768)	0
dropout_1 (Dropout)	(None, 32768)	0
dense_2 (Dense)	(None, 1024)	33555456
dense_3 (Dense)	(None, 7)	7175

Total params: 57,150,343
Trainable params: 33,562,631
Non-trainable params: 23,587,712

Figure 5. Sequential Model

Figure 5 shows the output after performing the flattening and dropout functions. It is performed to convert all the resultant 2d arrays into a single continuous linear vector.

A. Performance Evaluation

The model that was built for predicting the ECG arrhythmia was evaluated with various performance measures like accuracy and loss.

Accuracy: To estimate the accuracy we should calculate the proportion of True Positive(TP) and True Negative(TN) in all evaluated cases. The training data can be enlarged by augmenting the ECG images which results in higher classification

accuracy. To calculate the accuracy here using the following equation,

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

This model achieved 97% accuracy.

Loss: The loss value is used to optimize a machine learning algorithm. Loss value implies how poorly or well a model behaves after the each iteration of optimization. The loss is calculated on the training and validation and its interpretation is based on how well the model is doing in these two sets. It is the sum of errors that made for each example in the training or validation set. This model results the loss value 0.096.

IV. CONCLUSION

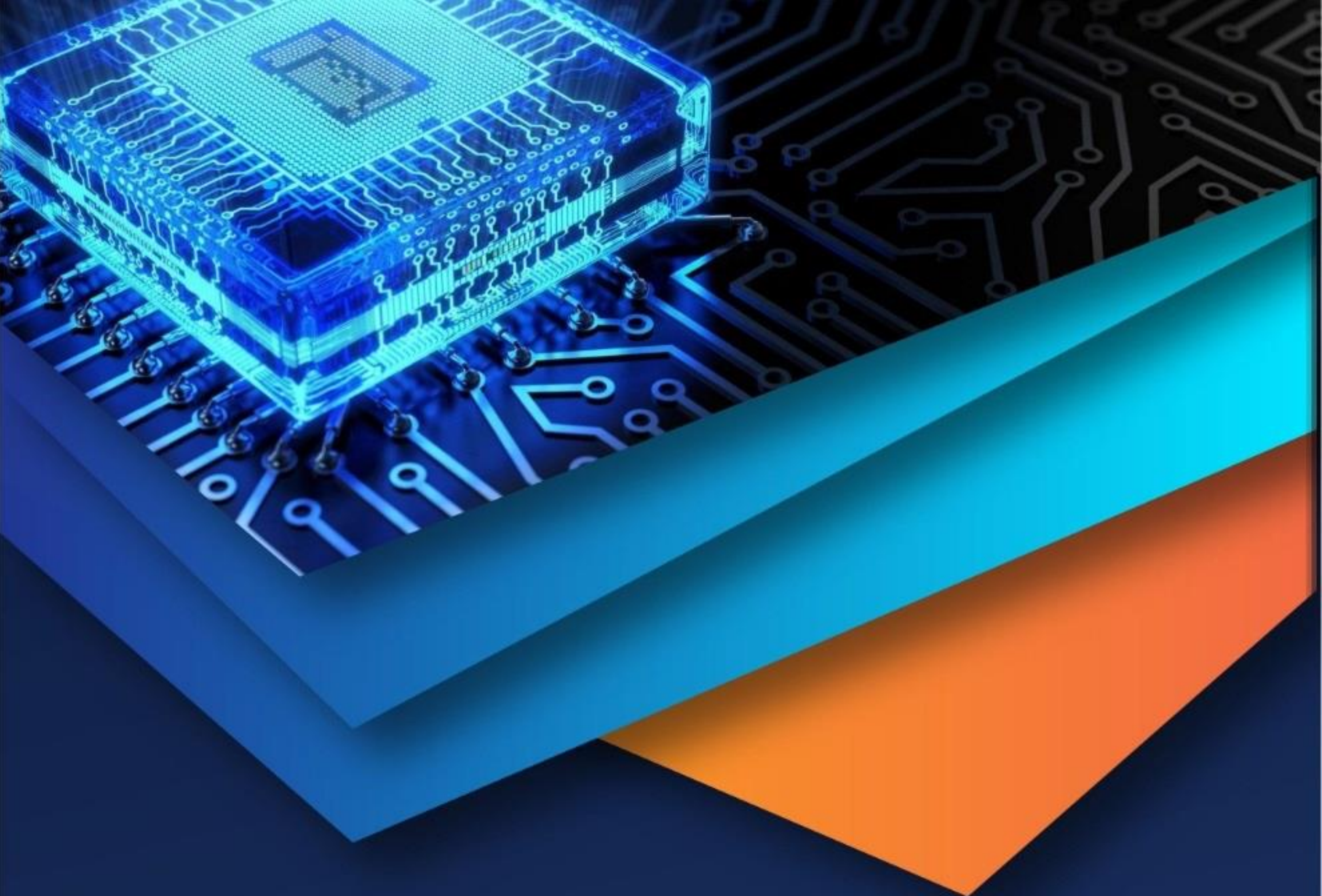
This paper proposes an effective ECG arrhythmia classification method using two dimensional Convolutional neural network ECG image as an input. There are eight types of ECG beats including normal beat and seven arrhythmia beats are obtained. CNN model is an effective approach to diagnose heart diseases which can be seen from ECG signals. Optimized CNN model is designed with considering important concepts such as data augmentation, regularization, and K-fold cross-validation. As a result, this proposed scheme achieved 96.99% average accuracy and 0.96% loss. Furthermore, proposed ECG arrhythmia classification method can be applied to the medical robot or the scanner that can monitors the ECG signals and helps the medical experts to identify ECG arrhythmia more precisely and easily.

V. ACKNOWLEDGEMENT

I would like to express my special thanks of gratitude to my project coordinator, Mr Valanto Alappatt, Assistant Professor of Computer Science and Engineering Department as well as our project guide Mrs.Vineetha A.V, Assistant Professor and Head of Department of Computer Science and Engineering for providing all the necessary facilities. I sincerely express my gratitude to Dr. K. Vijayakumar, Principal ofThejus Engineering College, Vellarakkad, who gave me the golden opportunity to do this project.

REFERENCES

[1] Divya Krishnani, Anjali Kumari, “ Prediction of Coronary Heart Disease using Supervised Machine Learning Algorithms ”, 2019.
[2] Senthilkumar Mohan, Chandrasegar Thirumalai, “Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques”, 2019.
[3] Dinesh Kumar G, Santhosh Kumar D, “Prediction of Cardiovascular Disease Using Machine Learning Algorithms ”, 2018.
[4] Abderrahmane Ed-daoudy, Khalil Maalmi, “Real-time machine learning for early detection of heart disease using big data approach ”, 2019.
[5] Elif IZCI, Aydin AKAN , “Cardiac Arrhythmia Detection from 2D ECG Images by Using Deep Learning Technique ”, 2019.
[6] Abhishek Rairikar, Vedant Kulkarni, Vikas Sabale, “Heart disease prediction using data mining techniques”, 2019.
[7] SP Rajamohana, K Karunya, “ Analysis of neural networks based heart disease prediction system ”, 2018.
[8] Jagdeep Singh, Amit Kamra, Harbhag Singh, “Prediction of Heart Diseases Using Associative Classification ”, 2017.
[9] Arif M, Akram MU, Afsar FA. “Arrhythmia beat classification using pruned fuzzy k-nearest neighbor classifier”, International Conference of Soft Computing and Pattern Recognition pp, 2019.
[10] Chetlur S, Woolley C, Vandermersch P, ”Efficient primitives for deep learning”,2014.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY