iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** XII **Month of publication:** December 2020

DOI: https://doi.org/10.22214/ijraset.2020.32399

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue XII Dec 2020- Available at www.ijraset.com

A Deep Learning Approach for Cardiac Arrhythmia Detection

Yashwanth Telukuntla[1], Varun Totakura[2], Vedha Sai Telukuntla[3], S. Naveen Kumar[4], V. Sai Charan[5]
[1, 2, 4, 5]Guru Nanak Institutions Technical Campus, Hyderabad, India,
[3]Amritha School of Engineering, Bangalore, India,

Abstract: Cardiac arrhythmia specifies uncommon electrical impulses of the heart that may be a major threat to humans. It should be reported for clinical evaluation and care. Electrocardiogram monitoring (ECG) measurements perform a significant part in the treatment of heart failure. Due to heartrate differences between individual patients and unknown disturbances in the ECG readings it is difficult for doctors to identify the type of arrhythmia. Classification plays an important role in health protection and computational biology. In this work, we aim to classify the heartbeats extracted from an ECG using deep learning, based only on the line shape (morphology) of the individual heartbeats. The goal would be to develop a method that automatically detects anomalies and help for prompt diagnosis of arrhythmia. A trained neural feed-forward network was chosen for this study. Experimental findings suggest that deep-learning models are more reliable than traditional cardiac diagnosis methods. The details used for the study of ECG signals were from the MIT-BIH database

Keywords: Electrocardiogram (ECG), Feed-forward, Deep learning, MIT-BIH Database

I. INTRODUCTION

A. About Cardiac Arrhythmia

With the growing scope of Database Management and Network Technology, people want to get information within seconds. It is now possible to solve medical issues with emerging technologies. Data is everything, in the healthcare industry. One of the costliest diagnosis is Heart failure. Electrocardiograms (ECG) play a vital role in the diagnosis of Cardiovascular Diseases. The heartbeat of normal humans varies between 60 and 90 beats per minute. Cardiologists generally interpret ECG reports to identify different types of irregularities in heartbeat. An ECG signal has P wave, QRS complex and T wave. For normal humans these waves occur with specific amplitudes at the duration of 250 to 1300ms. The duration of a P wave, QRS complex and T wave is 80ms, 80-150ms and 160ms respectively. Any variation in the defined timings is considered as an irregular heartbeat or Arrhythmia. This may lead to a wrong diagnosis for patients who may be in the earlier stages of cardiac disorders. Many feature extraction techniques from ECG signals were suggested in literature such as linear and non-linear features of HRV Signal, K Means clustering algorithm, Fuzzy neural networks, Dual-slope method, Dynamic features of ECG signals, Associative Petri net, Complexity analysis.

To have the classification with high accuracy and f1 score, we need to detect the QRS complexes of the ECG signals. PQ interval, QRS complex & T wave represent atrial contraction, ventricular depolarization and contraction and repolarization of ventricles

The QRS projection of ECG Signal is shown below (fig 1):

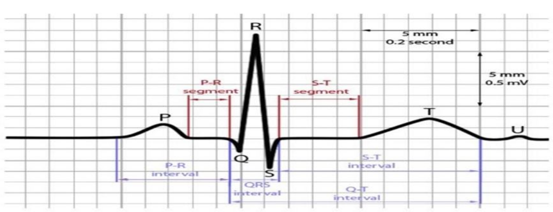


Fig 1 Structure of ECG Signal

Here in this work we have used the database which is already available in Kaggle by de Chazal where all the values are extracted using the Pam-Tompkins algorithm with every row on this dataset is a QRS Complex. Here we proposed a simple algorithm to identify four different cardiac arrhythmias using the ECG signal. These arrhythmias are namely the Supraventricular Ectopic Beat (SEB), Ventricular Ectopic Beat (VEB), Fusion Beat (FB) and the Normal Beat (NB). To classify ECG signal we are using this Deep Learning NN method which gives the classification accuracy and F1 score.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue XII Dec 2020- Available at www.ijraset.com

B. Deep Learning

Deep Learning emerged from machine learning family and is a subset of it. Supervised and Unsupervised are 2 learning methods. In supervised learning, the classification of arrhythmia is defined by a particular set of classes possible and train the model with those classes to identify those different arrhythmias. The model classifies different types of arrhythmias by calculating the values from the available R-R peak values from the dataset. Different types of arrhythmias can be identified by implementing a CNN model. Trained data is sent to the CNN model to train and test data is used to classify different types of arrhythmias. Any CNN model processes the data and classifies arrhythmia with the help of I/O layers, hidden layers and activation functions. The structure of a deep learning model is shown below (fig 2)

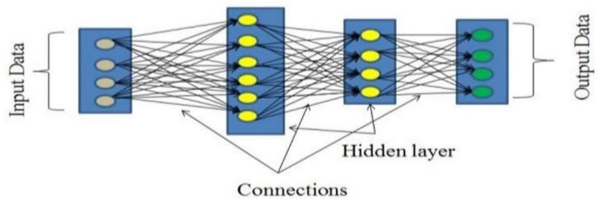


Fig 2 Structure of any deep learning model

Implementation

To implement the above-proposed process we have used Windows 10 Pro, Python 3.7.3, Kaggle as software.

From the above figure, the steps involved in the implementation are:

1) Step 1: Pre-processing

The collected dataset from the MIT-BIH database is already pre-processed and made available by de Chazal [8]. Out of 51,002 recordings available, 89.8% belong to Class 0, 7.4% belong to Class 2, 1.9% belong to Class 2 and 0.08% belong to Class 3. The figure below shows the data in each class. (fig 4)

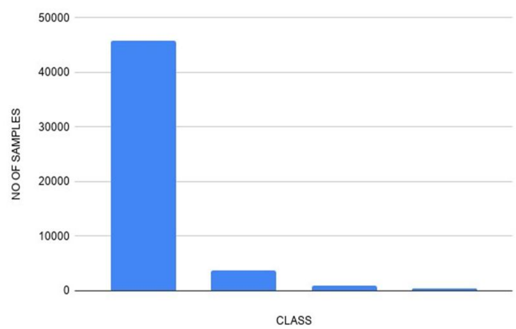


Fig 4 percentage of data of each class

2) Step 2: Stratified Shuffle Split of data

Split the data to train and test from the dataset. Since there are not many classes with 1 to 3, the data cannot be split randomly. To keep data with these classes in the train and test data, a stratified data split is performed with test size 0.2. From 51,002 samples, 40,801 are train data samples and 10,201 are test data samples. The percentage of classes present in the test data split will be the same as that of the present in the complete dataset using the stratified data split.

3) Step 3: Data standardization

Before running the model, standardization has to be implemented on the data. Standardization is to center the available data around 0 and to scale to the standard deviation

$$x_{standardized} = \frac{x - \mu}{\sigma},$$

where μ is mean and σ is the standard deviation of the dataset.

4) *Step 4:* Build and train the neural model

A Neural model is built using TensorFlow and Kera's, which has an input layer, 2 hidden layers, 1 with 200 and the other with 100 neurons. For the output layer 4 neurons are used, one for each class. A ReLu activation function for the hidden layer and a SoftMax function for the output layer is used. This model is trained with 40,801 train samples which were previously divided using *Stratified Shuffle Split.*

5) *Step 5:* Test the model

In the testing phase, the remaining 10,201 samples are tested and metrics such as Precision, Accuracy, f1 Score, Confusion Matrix, Support and Recall are measured.

C. Algorithm

1) Take ECG dataset from the MIT-BIH database.
2) Perform stratified shuffle split on the data so that the percentage of classes present in the trai & testdata will be the same as that of complete dataset.
a) Use StratifiedShuffleSplit() split data into train and test data.
b) Label each split as trainfeatures, train_labels and test_features, test_labels

3) Using StandardScaler() standardise the data.
a) Use fit_transform() on train_features and label as std_features_train.
b) Use trasform() on test_features and label as std_features_test.

4) Using Sequential() build and train the neural model on train data (trainfeatures, train_labels)
a) Create a model with 200 neurons for input layer, 100 neurons for hidden layer and 4 neurons for output layer with relu and softmax activation functions. 4.2 Using to_categorical() convert the labels to categorical.
b) Use SGD() as optimizer. 4.3.1 Fixing parameters in SGD() lr = 0.01 decay = le-6 momentum = 0.9 nesterov = Trule
c) Using compile() to compile
- Fixing parameters in compile() loss = categorical_crossentropy optimizer = sgd metrics= accuracy
d) Use evaluate() to evaluate to find the scores using train data and batch size.
5) Using predict() test the model on test data 5.1 Print accuracy, f1 score, confusion_matrix, recall, precision and classification_report

II. RESULTS

The classified ECG signals are provided as input to Deep Neural Network. The 40,801 sample recordings of ECG signals are used for training the DNN and 10,201 samples are used for testing. The network thus classifies the signals into the normal heartbeat and three kinds of cardiac arrhythmia. We conclude that the proposed solution dependably differentiates the four kinds of heartbeats based on the R-R interval of the ECG signals which has been validated over the entire MIT-BIH arrhythmia's database and yields an accuracy of 98.7% and 0.9 f1 score. Experimental results are tabulated within the below tables. We investigated the general outline of the model and derived the classification report and metrics for individual classes (fig 5) and also metrics such as confusion matrix, precision, accuracy and f1 score. (fig 6)

	Precision	Recall	F1-Score	Support
NB	0.99	1.00	0.99	9165
SEB	0.87	0.71	0.78	195
VEB	0.97	0.98	0.97	758
FB	0.89	0.82	0.86	83

NB - Class 0

SEB - Class 1

VEB - Class 2

FB - Class 3

Figure 5 Metric values for individual classes.

Metric	Value
Accuracy	0.987
F1 Score	0.90
Recall	0.87
Precision	0.93

Fig 6 Metric values for Overall classes

PREDICTED \ ACTUAL	Normal Heartbeat	Supraventricular Ectopic Beat	Ventricular Ectopic Beat	Fusion Beat
Normal Heartbeat	9112	36	12	5
Supraventricular Ectopic Beat	41	153	1	0
Ventricular Ectopic Beat	15	7	725	11
Fusion Beat	12	0	3	68

Fig 7 Confusion Matrix

Precision-Recall metric to estimate classifier result quality. It is a useful measure of success of prediction when the classes are very imbalanced. A high area under the curve represents both high recall and high precision, where high precision relates to a low false-positive rate, and high recall relates to a low false-negative rate. High scores for both show that the classifier is returning accurate results (high precision), as well as returning a majority of all positive results (high recall). The following graphs (fig 8 & 9) depict P-R curves for each class and the average precision graph micro averaged over all classes

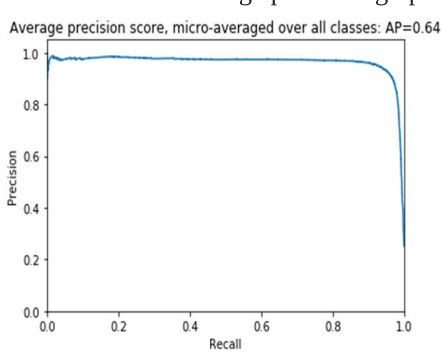


Fig 8 P-R curves over all classes

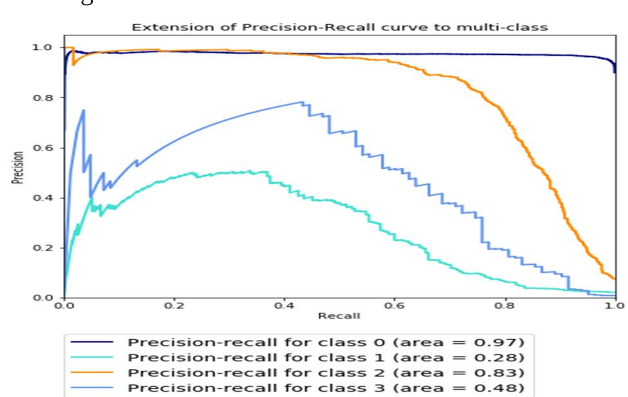


Fig 9 P-R curves for each class

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue XII Dec 2020- Available at www.ijraset.com

III. CONCLUSIONS

The irregularity within the heartbeat is known as cardiac arrhythmia. Cardiac arrhythmia leads to issues like heart failure, stroke and additionally cardiovascular arrest that ends up passing away. Therefore, recognizing and treating the heart abnormality at the right time is necessary. If the cardiac arrhythmia is recognized, then the detection of the type of arrhythmia is very feasible. We have developed a model using deep neural networks that classify the ECG signal recordings into 4 types including normal and Supraventricular Ectopic Beat, Ventricular Ectopic Beat and Hybrid Heartbeat by using MIT-BIH cardiac arrhythmia database

REFERENCES

[1] Sunita Soni, Jyoti Soni, Ujma Ansari, Dipesh Sharma, “Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction”, International Journal of Computer Application (IJCA, 0975 8887) Volume 17 No.8, March 2011.

[2] D. Mendes, S. Paredes, T. Rocha, P. Carvalho, J. Henriques, R. Cabiddu, J. Morais” Assessment of Cardiovascular Risk based on a Data-driven Knowledge Discovery Approach”,2015 IEEE.

[3] FrancisS. Collins, Mobile Technology and Healthcare, Available at http://www.nlm.nih.gov/medlineplus/magazine/issues/winter11.

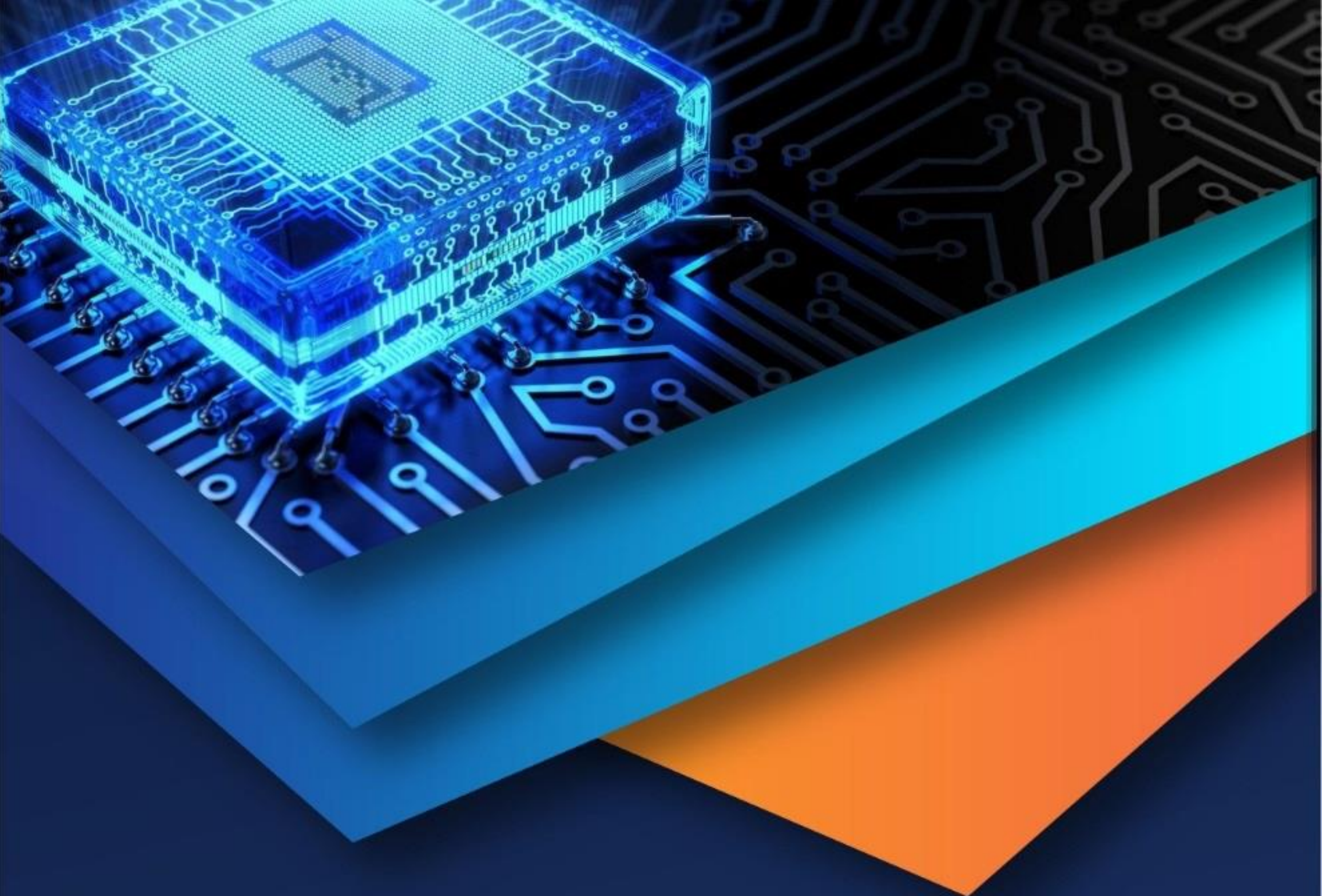
[4] Aieshwarya B. ChavanPatil, Prof. S. S. Sonawane, “To Predict Heart Disease Risk and Medication Using Data Mining Techniques with an IoT Based Monitoring System for Post-Operative Heart Disease Patients”, Sixth Post Graduate Conference for Computer Engineering (cPGCON 2017) Procedi International Journal on Emerging Trends in Technology (IJETT).

[5] Dimitra Azariadi; Vasileios Tsoutsouras; SotiriosXydis; DimitriosSoudris, "ECG signal analysis and arrhythmia detection on IoT wearable medical devices”, 5th International Conference on Modern Circuits and Systems Technologies (MOCAST) 2016.

[6] P.K. Anooj "Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules”, Production and hosting by Elsevier.

[7] Abdul Quaiyum Ansari, “Automated diagnosis of coronary heart disease using the neuro-fuzzy integrated system”, Information and Communication Technologies (WICT), 2011.

[8] De Chazal, “Automatic classification of heartbeats using ECG morphology and heartbeat interval features.”, at https://www.ncbi.nlm.nih.gov/pubmed/15248536



10.22214/IJRASET

INDEX COPERNICUS

45.98

ISRA JIF

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY