iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.44671

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

Cardiac Arrhythmia Prediction and Prevention of Heart failure using PCG(PhonoCardioGram) and CNN

Srinivasa Raghuram[1], Tanziya Niyaz[2], Harika Purma[3], Dr. Shruti Bhargava Choubey[4], Dr Y Sreenivasulu[5]
[1, 2, 3]Student, [4]Professor, Department of Electronics & Communications Engineering, Sreenidhi Institute of Science & Technology

Abstract: Arrhythmias are abnormal cardiac rhythms. According to WHO, in today's world 31% of deaths occur due to Cardiac Arrhythmia. Life-threatening arrhythmias including Ventricular Tachycardia (VT) and Ventricular Fibrillation are frequent causes of cardiac arrest (VF). The sinus node controls the heart's rhythm by triggering an electrical signal that goes through the heart, causing the heart to beat and circulate blood around throughout. The heart does not pump efficiently if there is too much electrical activity in the top or bottom chambers. Shortness of breath, fainting, an abrupt loss of heart function, and unconsciousness are the most common signs of Arrhythmia, which can result in death within minutes unless the victim receives emergency medical treatment to restart the heart.

The purpose of this research is to use the CNN and VGG16 models in conjunction with data augmentation and picture pixel creation to diagnose cardiac arrhythmias using PCG signals. For greater efficiency, phonocardiography (PCG) is also investigated. The majority of arrhythmia detection and classification methods rely solely on surface ECG analysis. So, to improve the efficiency of heart diagnostics, an algorithm is devised that relates to wavelet analysis at several resolutions combining temporal and wavelet properties of Electrocardiogram and Phonocardiogram, as well as Electrocardiogram-Phonocardiogram interactions. We want to be able to classify phonocardiograms (PCGs) or heartbeat recordings. as "normal" or "abnormal" in order to identify individuals who will require further diagnosis. The main concept is to transform each cardiac sound recording (wav file) into a spectrogram image and train a CNN model on that picture. We will then be able to categorize a fresh PCG recording as normal or abnormal.

Keywords: Cardiac Arrhythmia, Convolutional Neural Network (CNN), phonocardiogram (PCG), wav audio file.

I. LITERATURE SURVEY

After years of comprehensive research in the discipline of detecting cardiac arrhythmia in earlier stages, in the discipline of both medical and technologies such as AI and ML, we can prevent a lot of health diseases by predicting them in earlier stages.

The authors of [1] used ECG as a source to detect this disease and have achieved an accuracy of 94.74%. The authors of [2] used nonlinear parameters such as Discrete Wavelet Transform(DWT) on ECG and have attained an accuracy of 92.8%. The author of [3] used CNN to detect minor cardiac arrhythmia and got an accuracy score of 84.54%. The researchers of [4] classified the MIT Arrhythmia ECG database into normal and abnormal using ANN model and have attained an accuracy of 89.66%. The authors of [5] used RNN and feedback loops while the authors of [6] used the LSTM model which is the advanced version of RNN to predict the outcome. The authors of [7] used the GRU model to work on the ECG data. The authors of [8] produced a comparative study between all the deep learning models.

The authors of [9] used image recognition techniques and pre-trained models to detect cardiac arrhythmia. The authors of [10] used SVM algorithm and PCG signals to get the best possible disease classification.

II. EXISTING METHODOLOGY

A. DATA SET

The training dataset has 800 images (of which 400 belong to abnormal and 600 to normal class) and the validation set contains around 225 images(80 of abnormal and rest belong to normal class). We convert the .wav file audio recordings into spectrogram images in order for us to operate using our deep learning model.

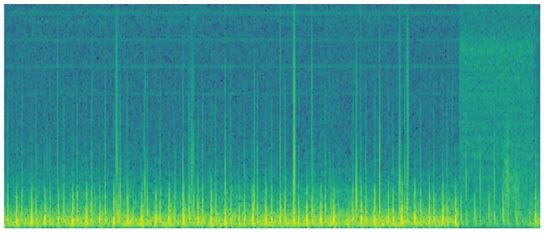


Fig.1 Spectrogram image of a normal heart beat.

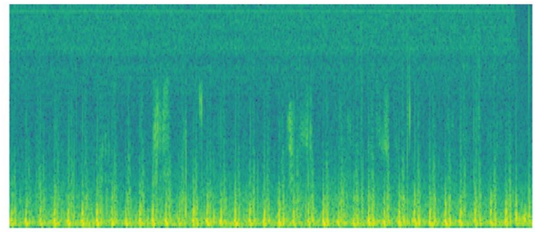


Fig.2 Spectrogram image of abnormal heart beat.

We basically perform **FFT.**

Fig.1 shows a spectrogram image of a person having normal heart beat and Fig.2 shows the spectrogram image of a person facing abnormal heart beat.

B. CNN

CNN Is a deep learning model which takes an input image and assigns different weights to the items in the image based on the relevance of the features to differentiate from the other image. We use CNN as the pre-processing required for this algorithm is very low compared to any other classification algorithm.

For all our experiments, we used Keras and TensorFlow as backend with a graphics processing unit (GPU). The existing methodologies used models like ANN, CNN, CNN-LSTM, CNN-GRU, etc. The pre-existing models used deep neural network algorithms such as ANN and CNN to classify the PCG images as normal or abnormal.

A cnn model contains four layers

1) *Convolutional Layer:* The convolutional layer is the most important component of a CNN since it is where the majority of the processing takes place. To generate a projected output, it requires an input data layer, a filter layer, and a feature map. Assume the input is a colour picture composed of a 3D matrix of pixels. This implies the input will have three dimensions: height, width, and depth, which match to the RGB colour space of a picture. A feature detector, commonly named a kernel or a filter, will scan the image's receptive fields for the existence of the feature.
2) *Pooling Layer:* Pooling layers are down sampling layers which reduces the quantity of parameters in the input by doing dimensionality reduction. Similar to the convolutional layer, the pooling method sweeps a filter across the whole input, but this filter does not include any weights. Instead, the kernel populates the output array with values from the receptive field using an aggregation function. They are two types of pooling.
   - Max Pooling
   - Average Pooling
3) *Padding Layer:* When the filters don't fit the input image, zero-padding is extensively employed. This reduces the size of all elements outside the input matrix to zero, resulting in a bigger or more evenly proportioned output.
4) *Fully-connected (FC) Layer:* This layer performs classification tasks based on the characteristics retrieved by the preceding input layers and their numerous filters. While ReLu functions are commonly used to detect and categorise inputs in convolutional and pooling layers, FC layers generally utilise a softmax activation function to offer a probability from 0 to 1 for prediction.

C. CNN Architecture

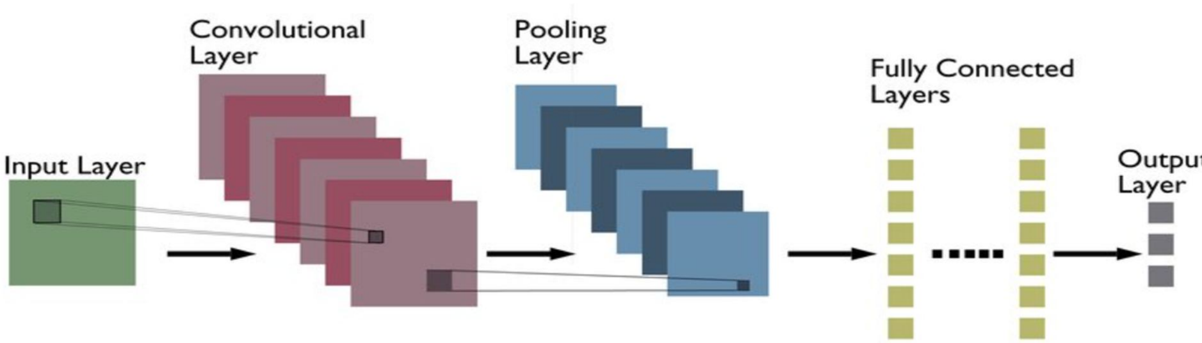


Fig.3 CNN Architecture

- **Statistical equation of CNN**

$$n_{out} = \left\lfloor \frac{n_{in} + 2p - k}{s} \right\rfloor + 1$$

n_{in}: number of input features
n_{out}: number of output features
k: convolution kernel size
p: convolution padding size
s: convolution stride size

To keep the size of the feature maps from shrinking at each layer, we need padding layers to be 2 times the input layers. By this, the dimensions of the input layer picture and feature map will stay unchanged as a result of this.

III. PROPOSED METHODOLOGY

A. *Transfer Learning VGG16 Over CNN*

1) *VGG16 and It's Architecture:* VGG stands for Visual Geometry Group which is a specific type of convolutional network designed for classification and localization whereas we use a cnn algorithm for basic semantics of the program.

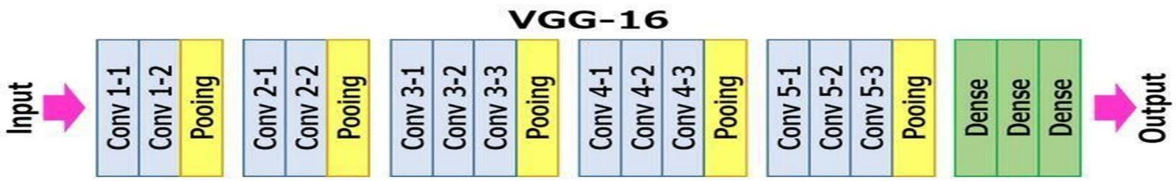


Fig 4: VGG16 Architecture

The 16 in VGG16 stands for 16 weighted layers. VGG16 comprises 13 convolutional layers, 5 Max Pooling layers, and 3 Dense layers, for a total of 21 layers, but only 16 weight layers, or learnable parameters layers are usable.The RGB picture 224x244 is used as VGG's input. On the training set picture, the average RGB value is determined for all images, and the image is then used as an input to the VGG convolutional network. The convolution step is fixed and a 3x3 or 1x1 filter is employed.VGG completely connected layers range from VGG11 to VGG19, depending on the total number of convolutional layers Plus fully connected layers. VGG11 includes 8 convolutional layers and 3 fully connected layers at the very least. There are 16 convolutional layers in the VGG19 at its maximum. +3 layers that are all related.

2) *Work Flow Diagram*

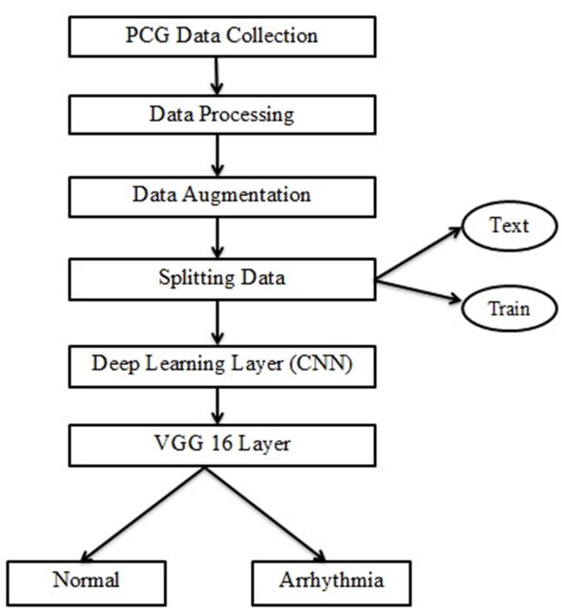


Fig 5: Flow Diagram

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

The first process is to collect the PCG data i.e, collecting the spectrogram data by performing FFT(Fast Fourier Transform) on the audio files. The data processing and data augmentation process is very important for our model as it never sees the exact same picture twice since different spectrograms look similar. This helps in preventing overfitting and the model generalizes better. For images, this could be done by rotating the original image, changing lighting conditions, cropping it differently, so for one image we can generate different sub-samples. We perform operations such as resizing, normalization, shear transformation, horizontal flipping, zooming and re-scaling the image. We then split the data into training and validation data and apply the CNN layer and perform transfer learning using the VGG16 layer on top of the CNN layer for a better prediction with utmost accuracy and precision.

3) *VGG Loss Equation:* The Perceptual Losses for Real-Time Transfer Learning framework introduces VGG Loss as a sort of content loss.

This loss is based on the ReLU(Rectified Linear Unit) activation layer of pre pre-trained model.

$$l^{SR}_{VGG/i.j} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2$$

Here,

$\phi_{i,j}$ - Indicates the feature extraction map that we get by j^{th} convolutional layer and i^{th} max pooling layer.

I^{HR} - Image reference.

$G_{\theta_G}(I^{LR})$ - Reconstructed image feature.

$W_{i,j}H_{i,j}$ - Dimensions of the feature map.

4) *VGG Loss Graph*

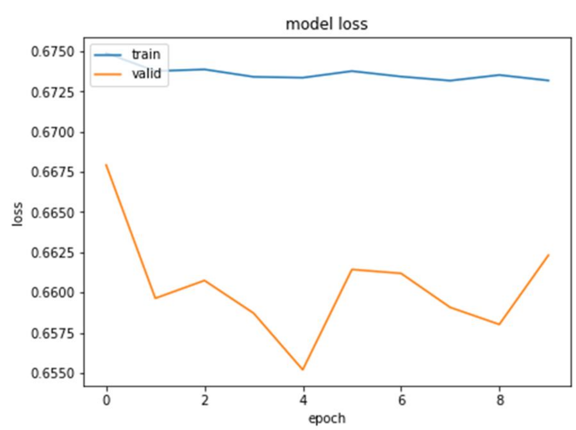


Fig.6: VGG Loss Graph

This VGG loss graph shows the loss occurring in training and validation data due to transfer learning of the VGG model over the CNN model.

- The perceptual VGG loss for our model turned out to be 0.6733 which is better than the previously existing model which had a loss of about 0.7423.

5) CNN Results Of Our Research Work

- Our model has achieved an accuracy of 0.9618, precision of 0.9621, recall of 0.9672, F-beta score of 0.9572 and loss of 0.1279 which is much less than the previous models.

IV. RESULTS

A. Results Of Various Parameters From The Reference Papers

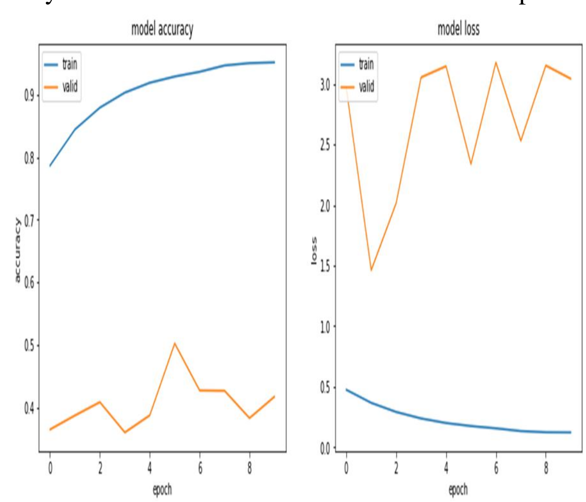
Algorithm	Accuracy	Loss
DWT	94.74%	0.822
ANN	89.6%	0.673
RNN With feedback loops	67.8%	0.712
GRU	88.62%	0.323
CNN	92.8%	0.55
VGG-16	69.6%	0.179

1) As we can see from the results formulated by previous works, DWT(Discrete Wavelet Transformation) had a better accuracy score but also had high loss values.
2) CNN outperformed any others models providing better accuracy and lower losses for the prediction.

B. Results of CNN from our Proposed Model

CNN	Accuracy	Precision	Recall	F-beta score	Loss
	0.9618	0.9621	0.9672	0.9578	0.1279

1) We have used various parameters such as accuracy, precision, recall score, f-beta score to validate the performance of prediction and also used binary cross-entropy loss for loss measurement.
2) Our proposed model got an accuracy of 96.18% with negligible loss of 0.1279 outperforming all the previous models.
3) The graph below shows the model accuracy and model loss with increase in the no.of epoch layers.



C. VGG16 Transfer Learning Results

CNN+VGG16 (transfer learning)	Accuracy	Precision	Recall	F-beta score	Perpetual Loss
	0.7201	0.7212	0.7204	0.7488	0.6733

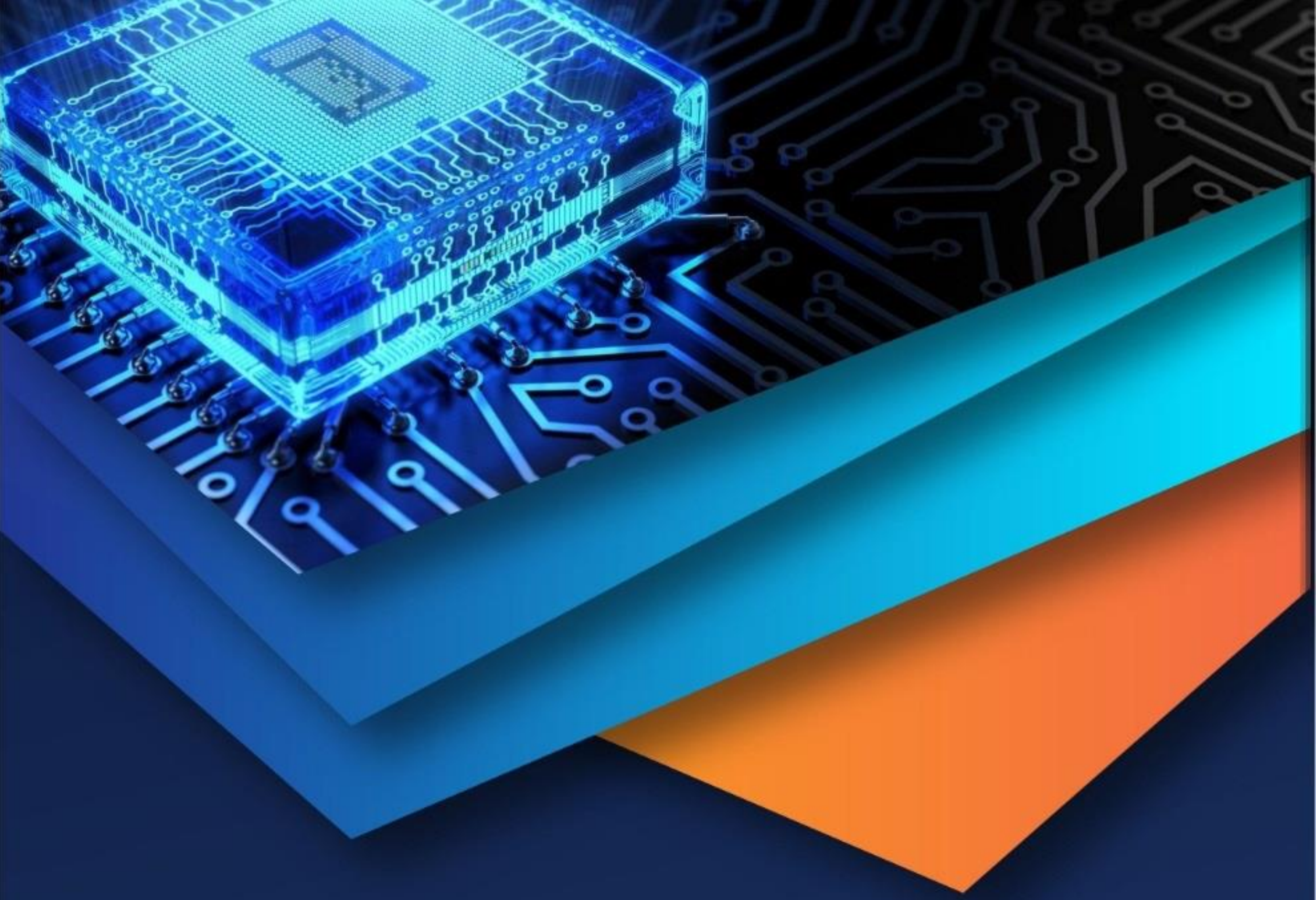
We can observe from the table above that our model has a better accuracy of 72.01% and lesser loss compared to previous research works.

V. CONCLUSION

Cardiac arrhythmia is a disorder in which the heart beats in an abnormal pattern.. Sudden cardiac death can be caused by certain types of cardiac arrhythmias. As a result, early detection and diagnosis of arrhythmia is critical. Once an arrhythmia has been diagnosed, the following step is to determine the kind of arrhythmia. We compared the performance of CNN, CNN-RNN, CNN-LSTM, and CNN-GRU deep learning architectures and found an accuracy of 0.834. With concern on computational cost, we used PCG instead of ECG. PCG testing is very cost economical compared to ECG. Using a complicated deep learning architecture, the given results can be enhanced even further. Complex network architectures can be learned employing sophisticated hardware and a distributed training technique that we are unable to implement.

REFERENCES

[1] Naser Safdarian, Nader Jafarnia, and Gholamreza Attarodi. (2014) “A new pattern recognition method for detection and localization of myocardial infarction using T-wave integral and total integral as extracted features from one cycle of ECG signal.” Journal of Biomedical Science and Engineering 7(10):818-824.
[2] Rajendra Acharya U., Hamildo Fujita, Vidya K Sudarshan, Shu Lih Oh, Muhammad Adam, et al. (2016) “ Automated detection and localizationofmyocardialinfarctionusingelectrocardiogram:Acomparativestudyofdifferentleads.” Knowledge-BasedSystems99:146-156.
[3] Tahsin Reasat and Celia Shahnaz. (2017) “Detection of inferior myocardial infarction using shallow convolutional neural networks.”arXiv preprint arXiv:1710.01115.
[4] Abhinav-Vishwa, Lal M.KDixit S, and Vardwaj P. (2011) “Classification of arrhythmic ECG data using machine learning techniques.” International Journal of Interactive Multimedia and Artificial Intelligence1(4):67-70
[5] Jeffrey L. Elman. (1990) “Finding structure in time.”Cognitive Science 14(2): 179-211.
[6] Sepp Hochreiter, and Jürgen Schmidhuber. (1997) “ Long short-term memory.” Neural Computation 9(8): 1735-1780.
[7] Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. (2014) “Learning phrase representations using RNN encoder-decoder for statistical machine translation.” arXiv preprint arXiv:1406.1078.
[8] V. Sai Krishna, A. Nithya Kalyani (2019, June) Prediction of cardiac arrhythmia using artificial neural network. Int J Recent Technol Eng (IJRTE 8(1S4): ISSN 2277-3878
[9] Kiranyaz S, Ince T, Gabbouj M (2016). “Real-time patient-specific ECG classification by 1-D convolutional neural networks.” IEE Transactions on Biomedical Engineering, Volume 63, pp. 664-675
[10] "UCI machine learning repository: Arrhythmia data set," 1998. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Arrhythmia. Accessed:Feb. 10, 2018.
[11] Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. IEEE Eng Med Biol Mag 2001;20:45–50.
[12] Nolle F, Badura F, Catlett J, Bowser R, Sketch M. CREI-GARD, a new concept in com- puterized arrhythmia monitoring systems. Computers in Cardiology 1986;13:515–8.
[13] Moody G. A new method for detecting atrial fibrillation using RR intervals. Com- puters in Cardiology 1983:227–30.
[14] Clifford G, Liu C, Moody B, Silva I, Li Q, Johnson A, et al. AF classification from a short single lead ECG recording: the PhysioNet computing in cardiology challenge 2017 presented at the Computing in Cardiology Rennes-France ; 2017.
[15] https://neurohive.io/en/popular-networks/vgg16/



10.22214/IJRASET

INDEX COPERNICUS
45.98

IMPACT FACTOR:
7.129

TOGETHER WE REACH THE GOAL
IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY