

AN ENHANCED METHOD FOR THE CLASSIFICATION OF ECGS THROUGH DEEP LEARNING

Abstract

The functionality of the cardiovascular system can be successfully evaluated through the use of electrocardiograms. In recent years, there has been a rise in the necessity of making a good diagnosis of arrhythmia. This is due to the fact that there are many commonalities between the various ECG limitations. They are unable to perform a more precise analysis of ECG graphs because rural areas do not have access to the most modern medical equipment and do not have sufficient numbers of trained medical professionals. Numerous people are losing their lives as a direct consequence of this. In this research, we describe a novel method for the categorization of arrhythmias and myocardial infarctions that is based on a convolution network of neurons that are deep in nature. In accordance with the AAMI EC57 standard, the model categorizes five distinct types of arrhythmias. In addition, we classified cases of myocardial infarction (MI) using the information collected from this work. For the purpose of evaluating our proposed model, we used diagnostic datasets provided by PhysionNet's MIT-BH and PTB. The findings demonstrated a significant increase in the accuracy of the classification of arrhythmia as well as MI. We are able to embed this software in a chip and make this product more readily available at more affordable prices in order to circumvent the resource scarcity that plagues medical treatment in rural or remote areas of any country.

Keywords : ECG, Classification, Healthcare, Deep Convolutional Network, Arrhythmia, Myocardial Infarction;

Authors

P. V. Ravi Kumar

Associate Professor,
Dept. of CSE,
Krishna Chaitanya Institute of Technology
& Sciences,
Markapur.
putta.msc@gmail.com

G N R Prasad

Associate Professor,
Dept. of MCA,
Chaitanya Bharathi Institute of Technology,
Hyderabad
gnrp@cbit.ac.in

Venkata Pavan Kumar Savala

Associate Professor & Head,
Dept. of CSE-DS,
Siddhartha Institute of Engineering &
Technology,
Ibrahimpattam, Hyderabad.
venkatapavankumarsavala@gmail.com

Ponnaboyina Ranganath

Research Scholar,
Acharya Nagarjuna University,
Guntur .
ranganathponnaboyina@gmail.com

P M Yohan

Professor & Principal,
CSI Wesley Institute of Technology and
Sciences,
Secunderabad.
pmyohan@rediff.com

SK Althaf Hussain Basha

Professor and R&D Coordinator,
Krishna Chaitanya Institute of Technology
& Sciences,
Markapur.
althafbashacse@gmail.com

I. INTRODUCTION

An electrocardiogram is one of the tests that is used to monitor the cardiovascular system in the human body. The fact that an ECG graph is made up of a variety of waveforms and curvatures in the signal, much like time-series data, presents the most significant challenge when analyzing the graph. As a result of this issue, the doctor might not be able to provide an appropriate diagnosis of the patient's heartbeat. In addition to this, incorrect diagnoses are the cause of one third of all deaths that occur around the globe. Therefore, a diagnostic of arrhythmia and MI that is both reliable and reasonably priced is essential.

The research that has been done up to this point in literature has studied utilized machine learning techniques to accurately categorize the signal in order to suggest solutions for the difficulties that have been outlined above. As part of the preprocessing step, the signal is run through a band-pass filter in the vast majority of these approaches. After that, the features that are extracted and given to the final classification task are those that represent the statistical figures of the signal windows. Examples of these methods that may be found in the relevant literature include decision trees, support vector machines, and multi-layer perceptrons, amongst others.

Recent developments in deep learning have made it possible to accurately forecast outcomes, analyze final characteristics in an effective and scalable manner, and evaluate them effectively. Deep learning has made it possible to train features that incorporate a huge number of variables in a relatively straightforward manner. The management of enormous amounts of data can be made much simpler by implementing the concept of knowledge transfer across two activities. For instance, the skills learned to complete one task of image processing can be applied to the completion of an other picture comprehension task, as demonstrated in M.Kachuee et al.[1].

This chapter presents a novel conceptual framework for the building of neural networks that is both deep and capable of comprehending the electrocardiogram (ECG) signals, particularly those that are related to the form properties of the ECG. In addition, we have a sizable amount of data that has been tagged and provided to the model in the form of training data. The skills acquired in identifying arrhythmias are successfully applied to the problem of determining the type of myocardial infarction a patient has had.

II. LITERATURE COLLECTION

1. Network Neural Convolutional Convolutional: Recent developments in the form of convolutional networks, as described by K.He et al. [5], marked the beginning of a revolutionary era. The process of convolutional networks makes the otherwise difficult task of processing photos of a high resolution more manageable. As a result of the development of GPUs (Graphic Processing Units), which simplified the process of computing parameters on a large scale in convolutional networks, deep convolutional networks came into existence. The following is an example of the typical architecture of a convolutional network.

A convolutional neural network is nothing more than a learning algorithm. This algorithm may take the pixel values of an image or any data that we wish to process,

learn the weights of the neural network through training, and then classify the data pattern or input image. In the context of this discussion, we have utilized the patient's heartbeats as an input for the deep CNN. Visual cortex served as the model for the architecture of ConvNet, which is based on the idea that it is analogous to the connection patterns of neurons in the human brain. The term "receptive field" refers to a specific region of the visual field in which the individual neurons fire with a higher frequency. The confluence of these receptive fields will make it possible for us to recognize a pattern in the data that is being entered or in an image.

A CNN is able to extract temporal and spatial correlations through the application of suitable filters. An image consists of pixels that constitute a 3×3 matrix, and thus provides the network with 9 input values. The initial layer of the CNN is known as the convolutional neural network. Fig.2 is the graphic that illustrates the filters that are used by the ConvNet. Filters, also called kernels, are another name for filters. A matrix with the dimensions $[h_2 * w_2 * d_1]$, which can be found in Fig. 2. For each layer of convolution, there is a large number of kernels that are stacked on top of one another, which ultimately manifests a three-dimensional matrix with the dimensions $[h_2 * w_2 * d_2]$, where d_2 is the total number of kernels. We have a bias that is a scalar value for each kernel.

The next layer is an output layer, and it is represented by a green matrix with the dimensions $[h_3 * w_3 * d_2]$.

Two things are clear to see in this situation. 1) The depth of input, denoted by d_1 , and the size of a single kernel are same. 2) The depth of output, denoted by d_2 , has the same number of kernels. Scanning the input matrix (color is blue) with the size of the kernel matrix, multiplying the corresponding pixel positions, and adding the products to generate the resultant output matrix (color is in green) is how the output calculation of 2D-convolution is performed.

2. **A Layer that is Pooling:** There are two different kinds of pooling: the average pooling and the maximum pooling. The Max Pooling algorithm determines the maximum of all of the pixel values that are occupied by the size of the kernel. On the other hand, when you use average pooling, the value of the kernel is used to determine the average of all of the pixel values that it occupies. One of the benefits of using max pooling is the suppression of noise. In addition to noise reduction, it also does dimensionality reduction on the data. On the other hand, using an average for everything contributes some form of noise. Therefore, we are able to draw the conclusion that maximum pooling is superior to average pooling. The concept of maximum pooling is illustrated in Figure 4.
3. **Block of Residual Material:** An architecture of brain connections known as a residual block occurs when the activation output of one layer is sent to the next layer after that, thus skipping the layer below it. The following is the rationale behind acting in this manner. One of the most important benefits of a deep network is that it can accurately categorize intricate data patterns. However, while training is taking place, there is a problem with the gradient disappearing in the process of sending feedback of mistakes at output levels to input layer of the network. This causes a difficulty during the training process. As the errors or gradients in output layers get multiplied to weights, if these

errors are very small, the gradients can approach very close to zero, which occurs exponentially and quickly. This can happen because of the nature of the multiplication.

When the network begins training, the accuracy eventually reaches a plateau, after which it rapidly declines. Therefore, utilizing residual blocks is the only option available.

The thought process behind this is to let the network discover the residual mapping rather than have the layers learn the implied mapping. Therefore, the result is $Z(x)$, and the network needs to conform to it. $Z(x)$ is equivalent to $F(x)+x$ by default, and the output of residual mapping is therefore $F(x)=Z(x)-x$.

The plan is to insert a skip link between the n th and $(n+2)$ th layers in the structure. There are essentially two distinct varieties of leftover blocks. The first block is called the identity block, and the second block is called the convolutional block. In the identity block, the skip connection feature is implemented just as we have been going over it. In the convolutional block, all that has to be done is the addition of a 2D-convolution that is performed across the skip link. Figure 5 provides an illustration of the structure of the residual block.

As a result, following the addition of new layers and the addition of skip or residual connections, the training error starts to gradually increase.

- 4. Rectified Linear Unit (abbreviated as ReLU):** The rectified linear input is a type of activation function that may be found in neural networks. This activation function has an output that is equal to the input if the input is larger than or equal to zero, and it has an output that is equal to zero if the input is negative. According to the findings of V.Nair et al.[4], the most effective method for activation functions in neural networks is to employ either the sigmoid or the ReLU.

III. A BRIEF INTRODUCTION TO THE T-SNE METHODOLOGY

t-SNE is an acronym for t-Distributed Stochastic Neighbor Embedding, which describes the method. This method is used primarily for dimensionality reduction and for visualizing or dimensionality reduction that is particularly well suited for the visualizing output of data sets that are having a large number of dimensions. Dimensionality reduction is utilized in the primary sense. In this chapter, we utilized this method for visually representing the newly gained knowledge in the classification process.

Different datasets Arrhythmia. In addition to this, we demonstrated that all of the knowledge obtained from the first database can be successfully transferred to the second database. During the course of the experiment that was carried out, we made use of an ECG lead-II input that was transformed to 125Hz frequency.

IV. METHODOLOGY

1. The Preliminary Processing Step

The following is the procedure that should be followed in order to process the beats of the ECG data.

- To begin, divide the signal from the ECG into 10-second windows, and then assign a value to each window.
- The values of the amplitude are then normalized into the interval $[0,1]$.
- Calculate all of the local maximums by setting the first derivative to an equal value of zero.
- Compute the group of peak R points by applying a cut-off value of 0.9 to the values of the normalized maximum values at the local level.
- Use the median of the R-R time intervals to calculate the heartbeat period (T) of the window in question.
- We have to choose a signal component of length that has a value that is equal to 1.2 times T for each and every R value.
- At this point, we need to pad each component with zeros in order to get a total length that is equal to a predefined fixed length. Now, the length of each of the extracted beats is the same, which is necessary so that it may be utilized as an input for the further stages of processing.

2. **Arrhythmia Trainer and Classifier in Training** : In this part of the tutorial, the convolutional layer is trained to categorize the different types of heart beats using the MIT-BIH dataset. The network receives its inputs in the form of the beats that were extracted in the previous section. Each CONV layer in this network performs 1D convolution on time, and each layer has 32 kernels of size 5, making the total number of kernels in the network 132. In each pooling layer, we take five maximum pool sizes and assign each one a stride value of 2.

The architecture consists of five blocks of the residual type, which are connected to two complete layers of connected neurons, each of which contains 32 neurons, and an additional layer of the softmax type at the very end. Every single residual block consists of one convolutional layer, followed by a single ReLU, a convolutional layer that is then connected to a ReLU as specified in K.He.et.al.[5,] residual skipped connection, and pooling layer. The end result is a deep neural network made up of 13 weight layers. The values in the cells of Fig.8 are the normalized values of the items that were classified into each category with two digits of precision.

- 3. The Training network for the MI Prediction:** In order to make a prediction of MI, the outputs of the last convolutional layer are used as input for a further two complete connected layers. The accuracy of the model's prediction of MI is improved by training the model's final two fully linked layers.

The analogy of the classification of arrhythmias can be found in table II.

- Dataset from the PTB, electrocardiogram lead II (TABLE III).
- The 12-lead electrocardiogram that the authors used (TABLE III)
- PTB dataset, 12-lead electrocardiogram (TABLE III)

- 4. The Results of the Experiment:** Tensorflow library is utilized in the experiment to facilitate the training of the network as well as the evaluation of it. The function of cross entropy is updated with the loss of cross entropy that occurs on the outputs. In order to train the network, we utilized the Adam optimization approach that was described in D.Kingma et al.[15] with the parameters beta-1, beta-2, and learning rates of 0.9, 0.999, and 0.001 correspondingly. It has taken fewer than two hours to train on the 1080Ti processor of GeForce GTX. The learning rate has decreased exponentially as a result of a decaying factor of 0.75 for every 10000 iterations.

V. RESULTS

It allows us to deduce that the model is able to differentiate between classes and produce predictions with a high level of accuracy. Table-II displays the average levels of accuracy achieved by various methods. According to it, the accuracy levels that were achieved through the experiment are comparable to those achieved through older methods. The success of this method can be attributed to the utilization of residual networks for the training of deeper networks.

- 1. The Classification of the MI:** We have trained our MI predictor, and the training set consisted of eighty percent of the PTB dataset, while the remaining twenty percent served as the test set. The recall, average accuracy, and precision of preceding methods are compared in Table III, along with the new method. The findings indicate that the proposed method is superior to those used in the past in terms of effectiveness. In addition, it is essential to take note that employs a 12-lead numbered electrocardiogram, which is in direct opposition to the fact that we use only 2-leads, in terms of the result's precision and correctness.
- 2. A representation of the Information that has been Learned:** In order to view the information that was gained, we mapped the high-dimensional data from the final convolutional layer to the 2D plane using t-SNE visualization. Fig. 10a displays a visualization of the information that was gained regarding the heart beat using the MIT-BIH dataset. Given this number, we are able to deduce that the various classes are, in fact, distinct from one another. The visualization of knowledge regarding the MI classification task on PTB samples using the MIT-BIH dataset is shown in figure 10b. From this image, we are able to deduce that the representation used for the arrhythmia

classification task is capable of delivering an accurate separation for the MI prediction test. In the figure, various categories are denoted by corresponding color designations.

VI. SUMMARY AND CONCLUSION

In this chapter, we proposed an innovative approach to the classification of arrhythmias that is based on the concept of transferable knowledge. In particular, a convolutional neural network that consists of residual blocks is trained for heart beat classification, and it is demonstrated that the state obtained in this phase may be used for the MI classification task. When compared to other methods described in the literature, the results of the suggested method are able to provide good predictions on both the classification of arrhythmias and the prediction of MI. In addition to this, we demonstrated that the suggested strategy is effective by projecting the knowledge using the t-SNE methodology.

REFERENCES

- [1] B Pang and L Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval* 2(1-2), 2008, pp 1–135.
- [2] M Hu and B Liu, "Mining and summarizing customer reviews," *Proceedings of the tenth ACM international conference on Knowledge discovery and data mining*, Seattle, 2004, pp 168-177.
- [3] B Pang, L Lee, and S Vaithyanathan, "Thumbs up? sentiment classification using machine learning techniques," *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, vol10, 2002, pp 79-86.
- [4] Jie Yang University of Wollongong, Australia "Mining Chinese social media UGC- a big-data framework for analyzing Douban movie reviews", *Journal of Big Data Springer*, 2016.
- [5] Kia Dashtipour Scotland, United Kingdom "Multilingual Sentiment Analysis: State of the Art and Independent Comparison of Techniques", *Springer*, 2016.
- [6] Kigon Lyu Korea University, Korea "Sentiment Analysis Using Word Polarity of Social Media", *Springer*, 2016.
- [7] Monu Kumar Thapar University, Patiala "Analyzing Twitter sentiments through big data", *IEEE*, 2016.
- [8] Minhoe Hur Seoul National University "Box-office forecasting based on sentiments of movie reviews and Independent subspace method", *Information Sciences*, 2016
- [9] Jorge A Balazs University of Chile "Opinion Mining and Information Fusion- A survey", 2015.
- [10] Donglin Cao Xiamen University, China "A cross-media public sentiment analysis system for microblog", *Springer*, 2014.
- [11] Ashraf I, Hur S, Shafiq M & Park Y. Catastrophic factors involved in road accidents: Underlying causes and descriptive analysis. *PLoS one* 2019 14(10), e0223473.
- [12] Singh H, Kushwaha V, Agarwal AD & Sandhu SS. Fatal road traffic accidents: Causes and factors responsible. *Journal of Indian Academy of Forensic Medicine* 2016), 38(1), 52-54.