Classifying electrocardiograph waveforms using trained deep learning neural network based on wavelet representation

Noor Yahya Jawad¹, Ahmed Mohammed Merza², Hussein Tami Sim³
¹Department of Computer Techniques Engineering, Al-Safwa University College, Karbala, Iraq
²Department of Biomedical Engineering, College of Engineering, University of Warith Al-Anbiya, Karbala, Iraq
³Department of Prosthetic Dental Technology, Hilla University College, Babylon, Iraq

ABSTRACT

Due to the rise in cardiac patients, an automated system that can identify different heart disorders has been created to lighten and distribute the duty of physicians. This research uses three different electrocardiograph (ECG) signals as indicators of a person’s cardiac problems: Normal sinus rhythm (NSR), arrhythmia (ARR), and congestive heart failure (CHF). The continuous wavelet transform (CWT) provides the mechanism for classifying the 190 individual cases of ECG data into a 2-dimensional time-frequency representation. In this paper, the modified GoogLeNet is used for ECG data classification. Using a transfer learning approach and adjustments to parts of the output layers, ECG classification was conducted and the effectiveness of convolutional neural network (CNN) designs was tested. By comparing the results that the optimized neural network and GoogLeNet both had classification accuracy about of 80% and 100%, respectively. The GoogLeNet provide the best result in term of accuracy and training time.

This is an open access article under the CC BY-SA license.

1. INTRODUCTION

The top five causes of mortality worldwide, including the majority of deaths, are cardiovascular disorders. The accompanying health issues, notably heart-related issues, have developed by increasing the sedentary lifestyles among vast populations [1]. The electrocardiogram is the main instrument used by medical professionals to examine and analyze a patient’s heart condition. Only a few of the distinctive components that may be utilized to determine the heart activity of a healthy person are the depolarization of the atria (P-wave), time taken from atrial depolarization to ventricular depolarization (PR interval), depolarization of the ventricles (QRS complex), time for ventricular depolarization and repolarization to take place (QT interval), and repolarization of ventricles (T-wave) [2]. Any cardiac ailment may be diagnosed and identified by looking at different patterns and anomalies in these waves. Since deep learning approaches have advanced over the past ten years, it is becoming common to diagnose various heart disorders utilizing computer-based methods [3]. We can reach a 100% outcome if we can keep getting better at it and tweak the datasets and technique layers. In order to do that, we will try K-fold validation and other comparable approaches in the future together with a federated learning strategy. The cardiac ailment known as arrhythmia (ARR) affects the pace at which the heart beats [4]. Inappropriate electrical impulse production, which controls the heartbeat, is one of the primary causes. The result is that the heart either beats extremely slowly, very fast or altogether randomly as a result of these erroneous electrical impulses. As a result, if the proper medical care is not given, it may result in a heart attack, heart failure, or sudden cardiac arrest. A healthy person's heart's regular electrical activity is represented
by the normal sinus rhythm (NSR). It is proof that the sinus node properly generates and transmits the electrical pulse [5]. When a person has chronic congestive heart failure (CHF), their heart's capacity to pump blood is significantly impaired. Heart artery narrowing and high blood pressure are the key contributors to the heart's ineffective pumping ability, which over time causes the heart to become very weak [6]. Three distinct cardiac states are represented by the electrocardiograph (ECG) signal. It is crucial to identify and categorize any anomalies in the ECG signal. Therefore, it will be useful for doctors and other health experts to find specific ECG anomalies so they can take the necessary medical action. Neural networks with numerous layers of neurons are used in deep learning techniques [7]. Deep learning has multiple layers functioning as processing units to achieve feature extraction. Each layer that extracts a particular feature takes its input from the output of the layer before it [8]. Deep learning approaches have surpassed previously used techniques in a number of pattern recognition competitions, which has inspired the scientific community to apply these cutting-edge methods to the analysis of biological images [9]. A deep bidirectional network with a unique wavelet-based layer was presented for the categorization of electrocardiogram signals [10]. The ECG signals were separated into frequency sub-bands at different scales using a wavelet-based layer. For the input of networks, these sub-bands functioned as sequences. By using the MIT-BIH arrhythmia database, five different cardiac kinds were discovered. A 144-layer convolutional neural network (CNN) has been created with the aim of recognizing different heart rhythms from electrocardiograms recorded using single-lead wearable monitors. In this study, a sizable dataset of 70,325 ECG recordings from 190 individuals was employed. Using the CNN model, a set of ECG data is converted into a set of rhythm classes. When the CNN model's performance with six different cardiologists was evaluated, the proposed model outperformed the performance of the typical cardiologist in terms of sensitivity and accuracy. The primary advantage of deep neural networks (DNNs) is the automated extraction and identification of complex and minute details in images, which eliminates the need for conventional feature extraction as in conventional machine learning (ML) methods. To extract complicated ECG properties and learn from a variety of available inputs, however, a sizable amount of data must be provided. Only then can rising levels of accuracy and human-level output be achieved. The modified GoogLeNet is employed in this study to classify ECG data. Using a transfer learning approach and adjustments to some of the output layers, the ECG classifications were performed and the efficacy of CNN designs was tested. Several sections are presented in this research as: In section 2 shows the literature review, and section 3 represents the methods used in this paper. Section 4 is the results and discussion of the ECG. Section 5 presents the conclusions.

2. LITERATURE REVIEW

In this section, we represent some of the previous studies: Moskalenko et al. [11] In 2019 The method they suggested generates a list of the onsets and offsets of the P and T waves as well as the QRS complexes and accepts as input any ECG data with any sampling rate. Our segmentation method is better than others since it can be used for a wide range of sampling rates and ECG monitor types and is rapid, only needs a few parameters, and has excellent generalizability. The recommended method performs better in terms of quality than current state-of-the-art segmentation methods. F1 measures are particularly effective in detecting the start and end of P and T waves as well as QRS-complexes, with 97.8%, 99.5%, and 99.9% accuracy, respectively. Murat et al. [12] in 2020 The examination of deep learning methods was then expanded upon using a five-class ECG dataset with 100,022 beats. Results from using this dataset to test the built models are shown. Therefore, this paper gives details on deep learning techniques utilized for classifying arrhythmias as well as recommendations for future research in this field. Naz et al. [13] in 2021 in this paper, a brand-new deep learning method for Veterans Affairs (VA) identification is proposed. The ECG impulses are then converted into hitherto unheard-of pictures. These photos are later normalized and used to train the deep learning models AlexNet, visual geometry group-16 (VGG-16) which is a type of CNN, and Inception-v3. To train a model and retrieve the deep information from several output layers, transfer learning is used. Then, using a heuristic entropy calculation technique, the best features are chosen after the features are concatenated and fused. For the final feature classification, supervised learning classifiers are used. The accuracy of the results, which were assessed using the MIT-BIH dataset, was 97.6% (using Cubic support vector machine as a final stage classifier). Jawad et al. [14] in 2022 the wavelet transform is used in this study to extract features. The electrocardiogram (ECG) signal is classified using an optimal neural network with eight classes using information from two ECG signals (ST-T and MIT-BIH database). The artificial neural network's training method uses the wavelet transform coefficients, which are then tuned using the invasive weed optimization (IWO) algorithm. The proposed approach offers over 70% sensitivity, over 94% specificity, over 65% positive predictive, over 93% negative predictive, and over 80% classification accuracy. The classifier performs better when the buried layer's number of neurons is increased. By comparing our work with the above studies we used the continuous wavelet transform (CWT) for feature extraction and the modified GoogLeNet is used for ECG data classification, the accuracy exceeded all the above studies and reach 100%.
3. METHOD

The suggested approach, which is different from the ECG domain, uses transfer learning from several classes and is an ECG multiple classification technique. Specifically, architectures that have already been trained on data related to image classification and object identification (1,000 classes, such as a chair, mouse, and table lamp) are used in place of training the CNNs from scratch using ECG data. Large datasets are readily available in these fields, allowing for effective feature map extraction and training that captures intricate characteristics and patterns in the images [15]. Only after converting 1-dimensional ECG signal samples into a 2-dimensional image of the ECG signal, known as a scalogram, using continuous wavelet transform, can such learned and accessible feature maps be transmitted for ECG classification applications. The GoogLeNet structure was chosen for this work because it can successfully extract features from a relatively small collection of ECG signal scalograms.

3.1. Google deep learning network

The existing depth learning approach may achieve high accuracy by improving the neural network's performance by introducing extra layers. This technique has a severe problem in that the computing cost grows exponentially with layer depth. At the 2014 ImageNet large-scale visual recognition challenge, Google unveiled GoogLeNet as the model with the greatest performance (ILSVRC14). In order to get the most performance out of each layer of the neural network, a range of probability distributions with strong correlations to the input data were acquired, therefore the inner layer of the neural network was enlarged to output a variety of correlation distributions. The basic inception v1 module aggregates the outputs into a single set after feeding the input data into four separate layers (1 1, 3 3, 5 5 convolution layers, and 3 3 maxpooling layer) [16]. By reducing the channel and size of the input data, the maximum pooling layer performs the function of extracting various features, and the convolutional layers provide various spatial information from the input data [17]. By expanding the layer of the neuron network, which is now just made up of the depth neurons, the inception module provides a method for fitting more data into a smaller layer.

3.2. Training and testing of GoogLeNet

To identify the ECG rhythm using our recommended deep learning architecture, the arrhythmia rhythm was divided into normal sinus rhythm (NSR), premature ventricular contraction (PVC), left bundle branch block (LBBB), right bundle branch block (RBBB), PACE, and atrial premature contraction (APC) as input data for training and testing. The MIT-BIH arrhythmia database area under consideration contains a large number of arrhythmias, and a reference beat is employed to classify them [18]. There are 190 patients from the MIT-BIH arrhythmia data that were picked at random. The total number of beats in the MIT-BIH arrhythmia data, together with the volume of data used to train and test the GoogLeNet model.

3.3. Modify Google neural network

Google neural network is a deep CNN that was first created to categorize photos into different categories. It may be used to classify ECG signals using deep neural networks and time series data turned into images using the CWT [19]. The three ECG status information is used for training the CNN. Google Net needs 224 by 224 by 3 red, green, blue (RGB) images which are considered the first component of the network layers properties. The layers of network architecture are operated as filters [20]. Edges, points, and colors are some of the more common visual components that the primary layers can identify. The following layers concentrate on more exact indicators for categorical differentiation. GoogLeNet network was programmed to recognize images into 1,000 various object categories [21]. Therefore, this network can be retrained for ECG categorization issues. A dropout layer is used for preventing overfitting. With a defined probability, this layer randomly sets input elements to zero [22]. Where this layer defined 0.5 as the default probability. The last dropout layer is configured to have a probability of 0.6. Convolutional layers extract the picture sign, and the final learnable layer and classification layer are applied to classify the input image [23]. The output from these two layers, which go by the name loss3-classifier, shows instructions on how to incorporate the traits that the network can extract into projected labels, a loss value, and class probabilities. These two layers are replaced with additional layers that are tailored to the data for retraining GoogLeNet to distinguish RGB images. These newly connected layers have the same number of filters as classes that must be employed instead of the loss3-classifier layer. For the fully linked layer, the learning rate parameter is taken into account to ensure that new layers learn more quickly than transferred layers. The modified network google net was analyzed as shown in Figure 1. The convolutional layers can convolute with learnable factors. GoogLeNet learns to recognize useful features with one feature for each channel. The first convolutional layer has 64 channels. The image inputs to the layer with a specific input size. the input image can resize before passing through the network. However, the network can pass larger images and the network activations become larger. Since the network was trained on a specific size of images, it cannot be trained to recognize features larger than the specified size.
Classifying electrocardiograph waveforms using trained deep learning neural networks

3.4. GoogLeNet training

Iteratively minimizing a loss function is a key step in the training of neural networks. It must utilize a gradient descent approach to minimize the loss function [24]. The descent method weights are updated and the gradient of the loss function is assessed in each iteration [25]. In the GoogLeNet model, which only has 144 layers and completes each epoch in around 1 minute, therefore, the training time of GoogLeNet takes over 30 minutes to complete. The training options are set iteration per epoch is 6, maximum epochs to 30, and the learning rate is 0.0001. The training progress is shown in Figure 2, the network converges in about 25 epochs with around 100% accuracy.
4. RESULTS AND DISCUSSION

The GoogLeNet model is evaluated in terms of performance using publicly accessible physiological data. To extract samples from different ECG recordings, the MIT-BIH normal sinus rhythm, MIT-BIH arrhythmia, and congestive heart failure databases (BIDMC) were used. For this study, a total of 190 ECG recordings were used. These recordings consist of 100 cases are ARR patients, 50 cases are NSR patients, and 40 cases are CHF patients. Each recording contained 70,325 samples in total. Figure 3 shows plots of three classes chosen at random from 1,000 samples for each class. ECG recording of each class was partitioned into 500 samples for expanding the size of the training dataset for the GoogLeNet network. The same scalograms were created from an equal number of ECG recordings for each class (CHF, ARR, and NSR), assuring an equal distribution of various labels as shown in Figure 3.

![Image](image_url)

Figure 3. Representation of the ECG recordings for each class

4.1. CWT representations

CWT representations of the ECG data are produced which are the absolute value of a signal’s CWT coefficients cleared using a scalogram. Precomputing a CWT filter bank is utilized after producing the scalograms. It is advised to use this CWT filter bank to obtain the CWT of multiple signals while maintaining the same settings. The CWT of the signal’s first 1,000 samples is obtained from the filter bank, and the coefficients are utilized to create the scalogram as shown in Figure 4.

![Image](image_url)

Figure 4. Producing the scalograms for the ECG classes

4.2. GoogLeNet activations

Each layer of a CNN is created via activation in response to an input image. It is possible to obtain visual information from a few layers. The earliest layers capture simple visual features. Each activation is scaled so that it has a minimum value of 0 and a maximum value of 1, accordingly. The activation consists of...
64 pictures, one photo for each channel in the layer. The activation is represented in an 8 by 8 grid as shown in Figure 5. Features can be examined by noting whether regions of a photo's convolutional layers are active and contrasting them with the corresponding regions of the original pictures.

The activations may be assessed and identified by the elements by comparing areas of activation with the original image GoogLeNet Learns. Convolutional layer regions on an image from the ARR, CHF, and NSR classes are active as shown in Figure 6, Figure 7, and Figure 8. The regions can be compared with their counterparts in the original picture. Convolutional layers are made of several channels, which are 2-D arrays. The output activations can be examined in the first convolutional layer after running the image through the network. The output of a channel in the convolutional layer is represented by each tile in the grid of activations. White pixels denote strongly positive activations, whilst black pixels denote considerable negative activations. A channel that is mostly grayscale responds to the incoming image less forcefully. The place of a pixel during channel activation coincides with that pixel's place during the original image. When a channel has a white pixel there, it means that the channel is very active there. The channel activations are resized to the size of the original image to display the activations. Where Figure 6(a) is the activation layer representation of ARR and Figure 6(b) is the classification result for one case of ARR. Where Figure 7(a) is the scalograms for the activation layer representation of CHF and Figure 7(b) is the scalograms classification result for one case of CHF. Where Figure 8(a) is the Scalograms for the activation layer representation of NSR and Figure 8(b) is the scalograms classification result for one case of NSR.

4.3. GoogLeNet confusion chart
A confusion chart is created to show the trained network's performance on the validation set. The confusion chart has columns and rows, which summarize the class's accuracy and recall. The accuracy of each
class is 100% as shown at the bottom of the confusion matrix in Figure 9. The recall values are shown at the right of the confusion chart.

![Figure 7](image_url)

Figure 7. Scalograms of; (a) CHF activation of GoogLeNet and (b) classified image for CHF heart rhythms

![Figure 8](image_url)

Figure 8. Scalograms of; (a) NSR activation of GoogLeNet and (b) classified image for NSR heart rhythms

![Figure 9](image_url)

Figure 9. Confusion matrix for GoogLeNet

5. CONCLUSION

This study shows the use of continuous wavelet analysis and transfer learning for classifying three categories of ECG signals by benefitting the pretrained GoogLe neural network. Wavelet representations are used for creating the scalograms of ECG signals. Therefore, the scalograms are created for each RGB image.
These scalograms of images are used for fine tuning the deep GoogLeNet. Activations were also shown for various network layers. The obtained results are strongly affected by the period of ECG signals, samples per signal, and the wavelet transformation parameters. This paper illustrates using a modified GoogLeNet model for enhancing the classification of the ECG signals. GoogLeNet neural network is a pretrained network for subsets of the Images database. The scalograms also underwent data reduction to fit with the GoogleNet network architecture. High accuracy was obtained for 190 individual cases of ECG data which was equal to the accuracy of the human level. In comparison to the 80% accuracy achieved by the optimized neural network, the GoogLeNet architecture achieves 100% accuracy for the multi-class ECG signal classification challenge utilizing a small dataset of scalogram images.

REFERENCES


BIOGRAPHIES OF AUTHORS

Noor Yahya Jawad was born in Karbala, Iraq, on June 11, 1990. He received a BSc in Electrical and Electronic Engineering from Karbala University. He completed and received a master's degree in Communications Engineering from the University of Technology Baghdad, Iraq. Now, affiliation is a Lecturer in the Department of Computer Techniques Engineering, Al-Safwa University College, Karbala, Iraq. His research interests include image/signal processing, communications, fiber optics, and artificial intelligence. He can be contacted at email: noor.yahya90@alsafwa.edu.iq.

Ahmed Mohammed Merza was born in Babylon, Iraq. He received the B.S and M.S degrees from University of Babylon, College of Engineering, Electrical Engineering, Iraq, in 2013 and 2016 respectively. Now he is a Ph.D. student in research stage at Islamic Azad University, Isfahan, Iran in electrical engineering. He currently a Lecturer at the Department of Biomedical engineering, college of engineering, University of Warith Al-Anbiyaa, Karbala, Iraq. His research interest including wireless sensor network, wireless communication, soft switching interleaved DC/DC converters. He can be contacted at email: ahmed.merza@uowa.edu.iq.

Hussein Tami Sim was born in Babylon, Iraq, on June 5, 1986. He received a BSc in General Physics from Babylon University. He completed and received a master's degree in Nuclear Physics from the University of Babylon, Iraq. Now currently a Ph.D. student (Medical Physics) at Erciyes University, Turkey. Now, affiliation is a Lecturer in the Department of Dentistry and Department of Prosthetic Dental Technology, Hilla University College, Babylon, Iraq. His research interests include the using of laser (Low Level Laser Therapy (LLLT)) in the treatment of oral problems (ulcers, cleft lip, gum disease, sensitivity of fillings). In addition to theoretical physics. He can be contacted at email: husseint@hilla-unc.edu.iq.