Deep classifier on the electrocardiogram interpretation system

Siti Nurmaini, Radiyati Umi Partan, Muhammad Naufal Rachmatullah

1 Intelligent System Research Group, Universitas Sriwijaya, Palembang, Indonesia
2 Faculty of Medicine, Universitas Sriwijaya, Palembang, Indonesia

E-mail: sitinurmaini@gmail.com

Abstract. Electrocardiogram (ECG) is a primary diagnostic tool for cardiovascular diseases. A higher accuracy of heart diseases needs an automatic classification for intelligent interpretation of cardiac arrhythmia. The classification process consists of following stages: detection of QRS complex in ECG signal, feature extraction from detected QRS using R-R interval, segmentation of rhythms using extracted feature set, learning system by using Deep Neural Networks (DNNs). The performance is analyzed as a rhythm of arrhythmia classifier and MIT-BIH arrhythmia database uses to validate the method. To benchmark, the performance of DNNs algorithm is compared to, MLP and SVM algorithm in terms of accuracy. The result obtained show that the proposed method provides good accuracy about 97.7% with less expert interaction.

1. Introduction
Cardiovascular diseases are leading causes of death and disability in the world. Moreover, all death-causing diseases in Indonesia, 5.1% due to Ischemic heart disease and 4.6% due to heart disease [1,2]. Unfortunately, the number of cardiologists in Indonesia is limited. They are only about 600 people for 262 million people, who not evenly distributed ac-cross the country, even though, they are needed. It needs a tool that has work automatically based on patient data to make a medical information system especially, for heart disease in order to overcome its lacks cardiologists. In solving various problems in medical diagnosis, the application of machine learning is needed to produce ECG signal quality for accurate and reliable interpretation and learning [3].

The state of cardiac heart is generally reflected in the shape of ECG waveform and heart rate. It may contain important pointers to the nature of diseases afflicting the heart. However, bio-signals being non-stationary signals, the reflection may occur at random in the time-scale (that is, the disease symptoms may not show up all the time, but would manifest at certain irregular intervals during the day). Therefore, for effective diagnostics, ECG pattern and heart rate variability may have to be observed over several hours. Thus, the volume of the data being enormous, the study is tedious and time-consuming. Naturally, the possibility of the analyst missing (or misreading) vital information is high. Therefore, computer-based analysis and classification of diseases can be very helpful in diagnostics [4]. Several algorithms have been developed in the literature for detection and classification of ECG beats. Most of them use either time or frequency domain representation of the ECG waveforms, on the basis of which many specific features are defined, allowing the recognition between of the beats belonging to different classes. The most difficult problem faced by today’s automatic ECG analysis is the large variation in the morphologies of ECG waveforms, not only of different patients or patient groups but also within the same patient. The ECG waveforms may differ for the same patient to such an extent that they are dissimilar to each other and at the same time they are similar for different types of beats. This is main reason that the beat classifier, performing well on the training data, generalizes poorly when presented with different patients’ ECG waveforms [5].
One of the methods of ECG beat recognition is neural network classification method [6-8]. Multi-layer perceptron (MLP) [6-11], which can be called “conventional back-propagation neural networks (BPNN)”, has been shown to be able to recognize and classify ECG signals more accurately. However, conventional BPNN suffers from slow convergence to local and global minima and from random settings of initial values of weights, which may make the neural networks have very poor mappings from inputs to an output. In [6], trained neural network correctly distinguished between normal heartbeats and premature ventricular contractions in 92% of the cases presented.

Overcoming the drawback of conventional BPNN in ECG classification, this paper proposes a novel classification of ECG signals based on Deep Learning by using Deep Neural Network algorithm [12]. Compared to the state of the art methods, this approach has several desirable properties: (i) it learns automatically an appropriate representation from the raw ECG to plot into the beat, which is each beat can be separated into a series of waves known as the P, Q, R, S, and T waves; (ii) simple feature extraction using R-R interval from QRS complex; (iii) selecting the most valuable ECG beats for inducing the DNN classifier. The idea of deep learning also known as feature learning is about learning a good feature representation automatically from the input data [13]. Recently, deep learning has shown outstanding results in medical applications and produce good accuracy [14-16].

2. Deep Neural Networks
The neural network has a very simple architecture and concept. One of the neural network (NN) technique is Feed-forward neural networks with many hidden layers, which are often referred to as deep neural networks (DNNs). DNNs has the same weakness as NNs, with BP training often resulted in poor performance, due to the network was not properly trained [16]. If the network gets deeper (Deep neural network), the local optimum happens along with the increase of hidden layer [16]. However, if learning parameters are trapped into the local optimum, the network can still work well because the probability of having a low local optimum is lower than when a small number of neurons are used in the network [17,18].

The neural network structure contains the following points, units are organized as layers, with every layer containing one or more units. The last layer is referred to as the output layer and all layers before the output layers are referred to as hidden layers. The number of units in a layer is referred to as the width of the layer. The width of each layer need not be the same, but the dimension should be aligned. Neural Network can be thought of as a function $f_\theta: x \rightarrow y$ which takes an input $x \in \mathbb{R}^n$, and produces an output $y \in \mathbb{R}^p$ and whose behavior is parameter by $\theta \in \mathbb{R}^n$. Therefore, for instance, $f_\theta$ could be simply $y = f_\theta(x) = \theta \cdot x$. A unit is a parameter by a weight vector $w$ and a bias term denoted by $b$. The first layer has $p^1$ units then each of the units has $w \in \mathbb{R}^n$ weights associated with them. The first layer produces an output $o_1 \in \mathbb{R}^{p^1}$. The output of the unit can be described as, $o_i = f[\sum_{k=1}^{n} x_k \cdot w_k + b_i]$. The index $k$ corresponds to each of the inputs/weights (from 1 to n) and the index i corresponds to the unit in the first layer (from 1 to $p_1$). We can assume that our data has the form $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$. For a single data point, the output of the NNs, which denote as $\hat{y}$.

In the BP algorithm is needed to compute how good the prediction of Neural Network $\hat{y}$ is as compared to $y$, in the notion of a loss function. A loss function measures the disagreement between $\hat{y}$ and $y$ which denote by $l$. There are a number of loss functions appropriate for the task at hand: binary classification, multi-classification, or regression, (typically derived using Maximum Likelihood) as a function,

$$l(\hat{y}, y) = l(f_{NN}(x, \theta), y)$$  \hspace{1cm} (1)

$$\nabla l(f_{NN}(x, \theta), y)$$  \hspace{1cm} (2)

Therefore, cross entropy function is given by the expression,

$$-\sum_{i=1}^{n} y_i \log f(x_i, \theta)$$  \hspace{1cm} (3)
Most of the DNNs training approach employs the cross entropy-driven learning rules. This is due to their superior learning rate and performance [12,16]. To overcome over-fitting is making the model as simple as possible using regularization.

3. Methods
The main objective of this proposed work was to construct intelligent interpretation system based DNN to analyze the ECG signal to detect abnormality of human heart condition. The proposed works was implemented using the steps of research aid to ease works shown in Figure 1.

3.1. Data Preparation
This study uses ECG signal database published by Harvard-MIT Division of Health Sciences and Technology [22]. The number of data about 48 instances consists of 25 data from male patients with age range between 32-89 years and 23 data from female patients with age range 23-89 years. The Arrhythmia taken from 2 modified limb lead II (MLII) signal types for upper signal and modified lead V1 (sometimes V2 or V5, and there is 1 instance V4) for a lower signal. The signal contains 650,000 records for 30 minutes. The attribute of Arrhythmia database can be defined as Table 1.

| Arrhythmia Data Set | Data Set Characteristics: | Number of Instances: | Area: | Life |
|---------------------|---------------------------|-----------------------|-------|------|
| Categorical Integer, Real | 47 | Date Donated | Between 1975 and 1979 |
| Nominal Attribute | 3 | Associated Tasks: | Detection |
| 2 Type per Groups (MLII, (V1, V2, V3, V4, V5) | 650,000 per 30 minutes | Signal Sample: | 360 samples per second per channel |
| Yes | Available | Total Set: | 1.805 |
| 4 for beat, 2 for rhythm | Normal: N | Abnormal: A, F, V for beat, VT for rhythm |

Table 1. MIT-BIH arrhythmia data set [22].
All 650,000 data records divided into 2167 sample rate, the duration of each sample records is sampled about 6 seconds. The total data used in the classification process about 14,352, with the data distribution such as 7826 data for normal condition, 2101 data for bradycardia condition, 1694 data for tachycardia condition and 3271 data for an irregular condition. The sample data are plotted into the ECG diagram as presents in Figure 2. Each beat can be separated into a series of waves known as the P, Q, R, S, and T waves [12].

![Figure 2](image_url)

**Figure 2.** The sample of ECG signal.

### 3.2. Feature Extraction

Simple feature extraction on Arrhythmia data classification is selected by using R-R Interval approach of ECG signal and the Waveform Database (WFDB) Software Package is used for viewing, analyzing, creating, recording and validation of physiologic signal [20]. For each ECG record, R peaks are detected and an annotation file containing their positions are generated. These positions compared to the reference annotations from the database. Standard comparison options are used: the comparison started five minutes after the beginning of the record and the match window, a maximum absolute difference in annotation times allowable for matching annotations, is set about 0.15 second. The input signal from Arrhythmia database is fed into the system for extracting, therefore it produces the following vector values. To find the R-R interval from QRS complex, all vectors are transformed into ECG graphic to assign the mark in a peak of R graph. From the R-R interval, heart rate is classified based on data Beats per Minute (BPM) from each ECG wave.
3.3. Deep Neural Networks (DNNs)

In this research DNNs with BP algorithm is used to process classification of Arrhythmia database. However, one of the major challenges in such method is the large computational load [16]. Therefore, the feature extraction must be a simple technique, in order to reduce computational load of the process, but still, produce good accuracy. In this study, the classification of the rhythm of ECG data will be divided into four classes, namely normal sinus, bradycardia sinus, ventricular tachycardia, and sinus arrhythmia. The stages of the classification process, as follows: data collection, data preparation, feature extraction, feature segmentation, learning process, and validation. The data are sharing about 70% for training and 30% for testing and the performance of DNN classifier compare to Support Vector Machine (SVM) in terms of accuracy and processing time [21,22]. In general, the process of DNN classification using backpropagation algorithm is well known. Determine the number of inputs (input patterns), hidden layer, and output (training target) and provides a random initial value for all weight between an input-hidden layer and hidden layer output. Perform the process repeated until the minimum error value that allows the DNN to run properly. Each input unit ($X_i$) receives an input signal and the signal is transmitted to all hidden layer units. The DNNs flowchart to process ECG Arrhythmia data classification use six stages such as data collection, data pre-processing including data preparation, feature extraction, feature classification, learning processing using DNN algorithm, validation, and analysis. All stages can be described in Figure 3.

4. Results and Discussion

The arrhythmia database contains 48 half-hour records, sampled at 360 Hz, and only four types of heartbeats are classified and labeled in this initial research to simplify the computation. The goal of our research is to use minimal information about the physiology of the heart to identify and classify beats. The classes are typically diagnosed by detecting into its constituent beats as, normal sinus, bradycardia sinus, ventricular tachycardia, and sinus arrhythmia. Our goal is to use a minimal amount of information
about the physiology of the heart to identify and classify beats using DNNs with BP algorithm, rather than building a complicated model of each beat. In order to analyze the classification performance, two simple metrics are used, accuracy and error. Accuracy is defined as the ratio of total beats correctly classified and the number of total beats.

![Figure 4](image1.png)

**Figure 4.** Classification accuracy in training and testing based on 18 cases.

![Figure 5](image2.png)

**Figure 5.** Classification accuracy.

The experiments are conducted by using 13 cases with several combination activation function, objective function, the number of neuron and the number of hidden layer. Figure 5, shows the training results obtained for each configuration on DNN structure against SVM. Here, we clearly notice that using a DNN representation (case 2.8) leads to better classification results compared to others structure. The number of hidden layers is 8, by using 100 hidden nodes. All the parameters are a reasonable choice for our initial DNNs model. However, SVM (1 vs 1) it can be seen that the largest classification error is obtained in normal and irregular data. This is probably because both types of signals (normal and irregular) have nearly the same number of R-R intervals. Therefore, the features generated by both types of signals are mutually exclusive. The highest accuracy of testing data classification is obtained by DNN algorithm of 2-8 cases with an accuracy of 97.7%. While the highest accuracy of testing data classification use SVM algorithm is 79.51% with a configuration of linear SVM (1 vs 1). However,
SVM with the polyline kernel was not successfully modeled because of the long-lasting training process (more than eight hours of training process not completed). From the experimental results, the greatest declassification occurs in normal and irregular classes, due to both signals have overlapping features. A comparison graph of the accuracy of data testing against the correct number of classifications data in Figure 4.

From the all experimental results are obtained that the highest accuracy in DNN structure with 8 layers with activation function at hidden layers is ReLU function, activation function in an output layer is soft-max function and loss function is categorical-cross entropy, it produces accuracy about 97.7%. But, in general, the change number of hidden layers not very influential only + 2%. In addition, the time required for training tends to increase as the hidden layer increases. Therefore, the computational cost is increased too. The results are shown in Figure 5. Confirm clearly the superiority of DNNs over standard NNs and SVM. From the classification result using SVM algorithm, the best data testing accuracy is obtained by using SVM -Linear with one-against-one setting with an accuracy of 79.19. RBF kernel has 100% accuracy in training by using SVM with, but it decreases to 51.6 % in testing. However, the SVM performance is not good when compared to NNs 1 hidden layer because with NNs 1 layer accuracy reach 95.56%. The biggest classification error is obtained in both normal and irregular data, due to both types of signals have nearly the same number of R-R intervals, it affects the accuracy of the beat rate. Moreover, the data is classified without the noise cancelation, to analyze all data segmentation based-on heart rate and beats rate value.

5. Conclusion

We have proposed and evaluated an intelligent system for the cardiac activity. This system classifies an ECG waveform in terms of features using a Deep Neural Networks with 18 cases against SVM algorithm. Our proposed design achieved a high-level of accuracy when evaluating data from the MIT-BIH Arrhythmia Database about 97.7 %, 95.6 % and 79.2 % for DNNs, NNs, and SVM respectively. This form of test is appropriate for a project at an early stage. In the future, we intend to implement and evaluate our system based on unsupervised approach and the data obtained directly from a patient, by adding a "live mode" to our system, which will allow it to be combined with Holter monitoring. We also plan to improve the accuracy of classification by reformulating the method of P wave detection.

6. References

[1] MOH, National Institute of Health Research and Development, Center for Community Empowerment, Health Policy and Humanities 2015. Sample Registration System 2014. NIHRD Library Cataloguing in Publication Data. Indonesia.
[2] Badan Penelitian dan Pengembangan Kesehatan 2014 Jakarta. Pokok-Pokok Hasil Riskesdas Indonesia tahun 2013. Lembar Penerbit Balitbangkes.
[3] Arsanjani R, Dey D, Khachatryan T, Shalev A, Hayes S W, Fish M, et al 2014 Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population. J Nucl Cardiol. 22(5) 877–84.
[4] R. Acharya, P.S. Bhat, S.S. Iyengar, A. Roo, S. Dua, Classification of heart rate data using artificial neural network and fuzzy equivalence relation, J. Pattern Recognition Soc. (2002).
[5] S. Osowski, T.H. Linh, ECG beat recognition using fuzzy hybrid neural network, IEEE Trans. Biomed. Eng. 48 (11) (2001) 1265–1271.
[6] S.Y. Foo, G. Stuart, B. Harvey, A. Meyer-Baese, Neural network-based ECG pattern recognition, Eng. Appl. Artif. Intell. 15 (2002) 253–260.
[7] M. Engin, S. Demirag, Fuzzy-hybrid neural network based ECG beat recognition using three different types of feature set, Cardiovasc. Eng. Int. J. 3 (2) (2003) 71–80.
[8] B. Karlik, M.O. Tokhi, M. Alek, A fuzzy clustering neural network architecture for multifunction upper-limb prosthesis, IEEE Trans. Biomed. Eng. 50 (11) (2003) 1255–1261.
[9] R.K. De, J. Basak, S.K. Pal, Unsupervised feature extraction using neuro-fuzzy approach, Fuzzy Sets Syst. 126 (2002) 277–291.
[10] W. Pedrycz, Conditional fuzzy clustering in the design of radial basis function neural networks, IEEE Trans. Neural Networks 9 (4) (1998) 601–612.
[11] P. Meesad, G.G. Yen, Pattern classification by a neuro fuzzy network application to vibration monitoring, ISA Trans. 39 (2000) 293–308.
[12] Y. Bengio, A. Courville, P. Vincent, Representation Learning: A review and new perspectives, IEEE Trans. Pattern Anal. Mach. Intell. 35 (2013) 1798–1828.
[13] G.E. Hinton, Reducing the dimensionality of data with neural networks, Science 313 (2006) 504–507.
[14] J.-S. Wang, W.-C. Chiang, Y.-L. Hsu, Y.-T.C. Yang, ECG arrhythmia classification using a probabilistic neural network with a feature reduction method, Neurocomputing 116 (2013) 38–45.
[15] Deng L and Yu D. 2014 Deep Learning Methods and Applications. Foundations and Trends in Signal Processing, 7:(3–4) 197–387.
[16] Min S, Lee B, Yoon S 2012 Deep learning in bioinformatics. Briefings in bioinformatics. 18(5) 851-869.
[17] Arsanjani R, Xu Y, Dev D, Vahistha V, Shalev A, Nakanishi R, et al 2013 Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population. J Nucl Cardiol. 20(4) 553–62.
[18] Medina E L, Mesquita C T, and Filho O L 2016 Healthcare social networks for patients with cardiovascular diseases and recommendation systems. International Journal Cardiovascular Sci. 29(1) 80-5.
[19] Deng L. and Yu D. 2014 Deep Learning Methods and Applications. Foundations and Trends in Signal Processing, 7:(3–4) 197–387.
[20] Yu D and Deng L 2016 Automatic Speech Recognition a Deep Learning Approach. Signals and Communication Technology. Springer
[21] Min S, Lee B, Yoon S 2012 Deep learning in bioinformatics. Briefings in bioinformatics. 18(5) 851-869.
[22] Massachusetts Institute of Technology, “MIT-BIH ECG database” : http://ecg.mit.edu/

Acknowledgments
Authors would like to thank Universitas Sriwijaya and Ministry of Research Technology and Higher Education of Indonesia, for their support in our research work, with the funded of Hibah Penelitian Profesi.