Int. J.Adv.Sig.Img.Sci, Vol. 2, No.1, 2016

ECG SIGNAL CLASSIFICATION BASED ON STATISTICAL FEATURES WITH SVM CLASSIFICATION

Vijaya Arjunan R,
School of Engineering and IT,
Manipal University, Academic city,
Dubai, UAE
vijaya.arjunan@manipaldubai.com

Abstract: An important diagnostic technique to detect the abnormalities in the human heart is Electrocardiogram (ECG). The growing number of heart patients increases the physicians work load. To reduce their work load, a computerized automated detection system is required. In this paper, a computerized system is presented to categorize the ECG signals. MIT-BIH ECG arrhythmia database is used for analysis purpose. After de-noising the ECG signal in the preprocessing stage, the following features; mean, variance, standard deviation, and skewness are extracted in the feature extraction stage and Support Vector Machine (SVM) is developed to classify the ECG signal into two categories; normal or abnormal. Results show that the system classifies the given ECG signal with 90% of sensitivity and specificity as well.

Keywords: ECG Signal, Arrhythmia, Statistical Features, SVM.

I. INTRODUCTION

In the last two decades, the biomedical engineering community gives specific attention for the classification of ECG signals. The functioning of the heart can be easily monitored by cardiologists using ECG signal. Also, it provides various information's about the rhythm. Figure 1 shows a sample ECG recording. Discrete Wavelet Transform (DWT) and Multi-Layer Perceptron Neural Networks (MLP-NN) based classification of ECG arrhythmias is presented in [1]. DWT is used for processing ECG recordings, and extracting features, MLP-NN performs the classification task. Morphological and dynamic features based heartbeat classification is described in [2]. Morphological features are extracted using DWT and independent component analysis is applied to each heartbeat. SVM classifier is utilized for the classification.

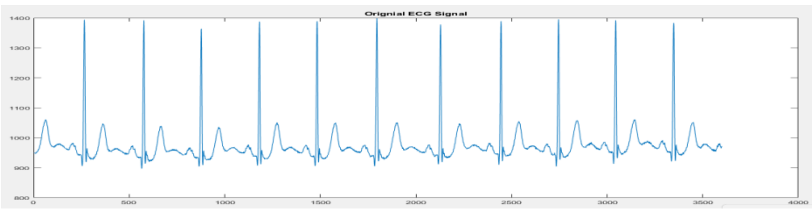


Fig. 1 ECG Signal

DWT and geometrical features based cardiac arrhythmia recognition using classifiers of MLP- Back Propagation (BP), SVM and Probabilistic-NN (PNN) is discussed in [3]. DWT is used to reject the noise and artifacts. Also, it is used to extract features from the segmented QRS complexes of a signal. MLP-BP, SVM and PNN are used to classify the signals after the extraction of features from the segmented QRS signal.

Multi view learning based automatic subject adaptable heartbeat classifier is explained in [4]. Multi view learning is used to address the significant inter person variations in ECG signals and automate subject adaptation is used which does not require manual labelling.1D Convolutional-NN (CNN) for ECG classification is illustrated in [5]. It uses real time patient specific signals for the analysis. Two major blocks; feature extraction and classification of the system are fused by CNN into a single learning body.

A study about the extraction of feature and classification of ECG signal is discussed in [6]. The extracted wavelet coefficients from the DWT process are used as features. S-transform, Genetic Algorithm (GA), and NN based ECG signal classification is described in [7]. S-transform is used to extract features in the first stage. The extracted features are optimized by GA so that the major information's are retained which is useful for effective classification by NN.

A detailed survey about ECG signals classification is presented in [8]. It consists of available ECG databases, various preprocessing and feature extraction approaches. Also, Neural Networks (NN) based ECG classification is discussed. Hjorth Descriptor based ECG signal classification is explained in [9]. Hjorth descriptor is used as a method for feature extraction. K-nearest neighbour and MLP are used as a classifier.

Neighborhood feature extraction based ECG signals classification is discussed in [10]. Non-linear dimension reduction methods are applied to ECG signals. Also, segmentation of data through neighbourhood feature extraction method is enabled by transiting from high dimensioned space to low dimension space by considering the longitudinal combination of ECG signals.

Block-Based NN (BBNN) and Particle Swarm Optimization (PSO) is described in [11] for ECG signal classification. From ECG signals, temporal and DWT based features are extracted. Then, the input vector of BBNN is created by the extracted features. The BBNN parameters have been optimized by PSO algorithm. DWT, time, and frequency domain features based ECG signals classification is presented in [12]. DWT and Fast Fourier transform (FFT) are used to transform ECG signals. Then, the signals are changed into the frequency domain using FFT and DWT.

II. METHODS AND MATERIALS

A. Methodology

Figure 2 shows the overall ECG signal classification system. The ECG signal classification system consists of three stages. First is the pre-processing stage. It is a common name for operations with signals at the lowest level of abstractions and it is an improvement of the signal data that suppresses unwanted distortions or enhances some signal features. In this stage, signal denoising is employed using the median filter.

The second stage is feature extraction which uses statistical features for the classification. The interesting parts of a signal are represented with lower dimension in this stage. It is a type of dimensional reduction as the whole signal is represented by a compact feature vector and this step is very useful in a situation where the dimension of an input signal is too large. In order to complete the classification task quickly in signal retrieval and matching, the input features should be in a compact form. The following statistical measures; mean, standard deviation, variance, and skewness are computed.

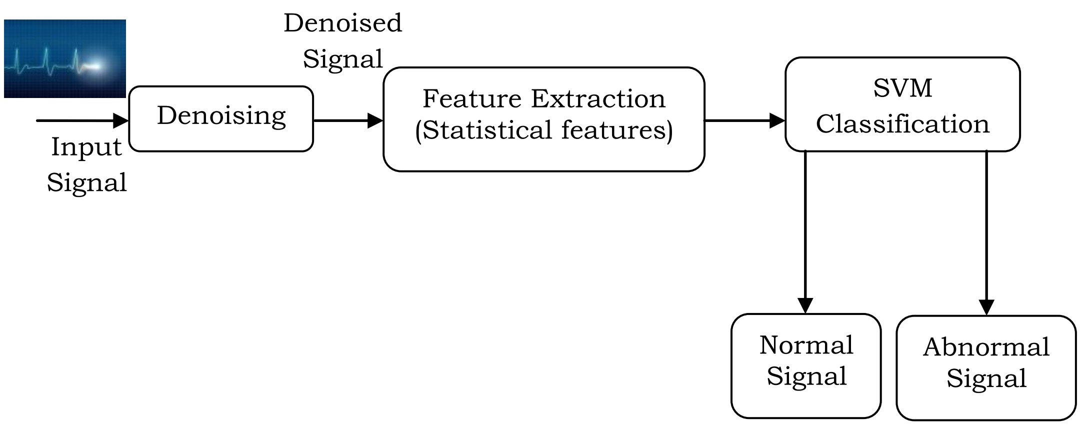


Fig. 2 ECG Signal Classification System

The final stage of the system is classification. It analyzes the numerical properties of unknown signal features in order to categorize them into different classes using training features of known classes. It typically employs two phases: training and testing. In training, the features of known samples and their classes are fed into the SVM classifier. Then, the trained classifier is tested with the test sample. Based on the test outcomes with ground truth data, the performance of the system is analyzed.

B. Statistical Features

The statistical features extracted from the given input ECG signal are as follows:

- **Mean:**
  It is the average of a set of values and defined as the ratio between the sum of these values and the number of elements in the set.
- **Standard Deviation:**
  The amount of dispersion or variation of a given dataset is quantified by using this measure.
- **Skewness:**
  It's the sum of the values in the data distribution divided by the number of values in the distribution.
- **Variance:**
  Let us consider a set of random numbers X and its mean $\overline{X}$. Variance measures how far X are spread out from $\overline{X}$. Also, it is computed from the squared deviation of X from its $\overline{X}$.

C. SVM Classifier

SVM is a learning algorithm which has many good properties. It is associated with data analysis and recognizes pattern. It uses a linear discriminate function for classification [13]. However, non-linear classification can also be done if a non linear kernel is used [14]. SVM performs well in real time situation, robust, easy to understand. While compared to other classifiers, it has a global solution. A classification task typically requires the knowledge about the data to be classified.

Hence, the classifier must be trained before classifying any data. One of the main advantages of SVM classifier is that it automatically finds the Support Vectors (SVs) for better classification. The performance of SVM depends on the kernel function and more information about SVM is found in [15].

III. RESULTS AND DISCUSSION

In this study, MIT-BIH arrhythmia dataset [16-17] is employed for the analysis. A sample normal and abnormal recording in the MIT-BIH database is shown in Figure 3 (a) and (b) respectively. The database contains two-channel ambulatory recording for 48 half-hour excerpts from 47 subjects. They are digitized at 360 samples per second.

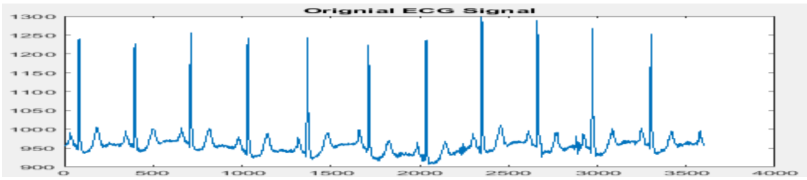


Fig. 3 (a) Normal Signal

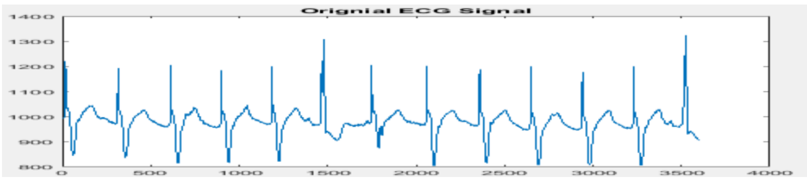


Fig. 3 (b) Abnormal Signal

Before extracting signal features, all the ECG signals are de-noised using simple median filtering. The window size used in this study is 3. Figure 4 (a) and (b) show the de-noised version of normal recordings in Figure 3 (a) and abnormal recordings in Figure 3 (b) respectively.

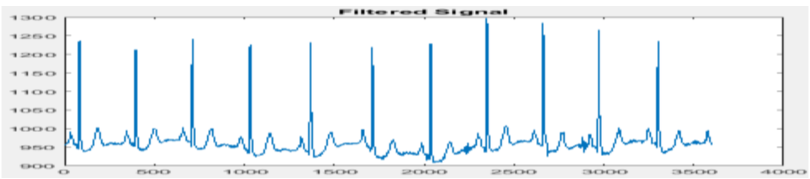
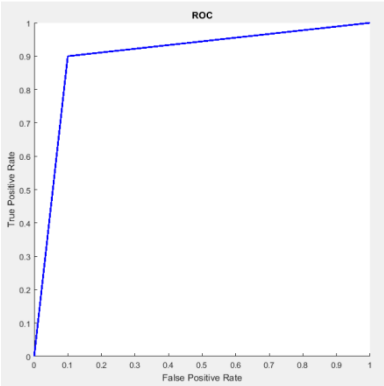


Fig. 4 (a) De-noised Normal Signal

Fig. 4 (b) De-noised Abnormal Signal

Finally, the classification of a signal using SVM classifier is analyzed with the help of statistical features extracted from the de-noised version of input ECG recordings. k-fold cross validation is used. Figure 5 (a) and (b) shows the ROC curve and confusion matrix of the analyzed ECG signal classification. It is observed from the Figure 5 (a) and (b) that the ECG system provides 90% accuracy in terms of sensitivity, specificity and accuracy as well.



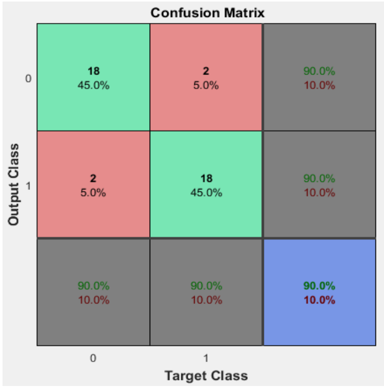


Fig. 5 (a) ROC Curve ***Fig. 5 (b) Confusion Matrix***

IV. CONCLUSION

In this paper, ECG signal classification using statistical features is analyzed. MIT–BIH arrhythmia database records are employed for the classification task. First, preprocessing is done using the median filter, and statistical features are extracted. Finally, SVM-RBF based classification is employed to classify the signals. Experimental results show the performance of ECG signal classification system with promising results. SVM-RBF classifier classifies 90% of the given ECG signal correctly with simple statistical features.

REFERENCES

[1]. M.K. Sarkaleh, and A. Shahbahrami, "Classification of ECG arrhythmias using discrete wavelet transform and neural networks", International Journal of Computer Science, Engineering and Applications, Vol. 2, No. 1, 2012, pp. 1-5.

[2]. C. Ye, B.V. Kumar, and M.T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals", IEEE Transactions on Biomedical Engineering, Vol. 59, No.10, 2012, pp.2930-2941.

[3]. F. Asadi, M.J. Mollakazemi, S.A. Atyabi, I.L.I.J.A. Uzelac, and A. Ghaffari, "Cardiac arrhythmia recognition with robust discrete wavelet-based and geometrical feature extraction via classifiers of SVM and MLP-BP and PNN neural networks", IEEE Computing in Cardiology Conference, 2015, pp. 933-936.

[4]. C. Ye, B.V. Kumar, and M.T. Coimbra, "An Automatic Subject-Adaptable Heartbeat Classifier Based on Multiview Learning", IEEE journal of biomedical and health informatics, Vol. 20, No. 6, 2015, pp. 1485-1492.

[5]. S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks", IEEE Transactions on Biomedical Engineering, Vol. 63, No. 3, 2015, pp. 664-675.

[6]. S. Basu, and Y.U. Khan, "On the aspect of feature extraction and classification of the ECG signal", IEEE Communication, Control and Intelligent Systems, 2015, pp. 190-193.

[7]. M.K. Das, D.K. Ghosh, and S. Ari, "Electrocardiogram (ECG) signal classification using s-transform, genetic algorithm and neural network", IEEE 1st International Conference on Condition Assessment Techniques in Electrical Systems, 2013, pp. 353-357.

[8]. S.H. Jambukia, V.K. Dabhi, and H.B. Prajapati, "Classification of ECG signals using machine learning techniques: A survey", IEEE International Conference on Advances in Computer Engineering and Applications, 2015, pp. 714-721.

[9]. A. Rizal, and S. Hadiyoso, "ECG signal classification using Hjorth Descriptor", IEEE International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology, 2015, pp. 87-90.

[10]. C. Bakır, "ECG signals classification with neighborhood feature extraction method", IEEE Medical Technologies National Conference, 2015, pp. 1-4.

[11]. S. Shadmand, and B. Mashoufi, "Personalized ECG signal classification using block-based neural-network and particle swarm optimization", IEEE 20th Iranian Conference on Biomedical Engineering, 2013, pp. 203-208.

[12]. S.A. Shufni, and M.Y. Mashor, "ECG signals classification based on discrete wavelet transform, time domain and frequency domain features", IEEE 2nd International Conference on Biomedical Engineering, 2015, pp. 1-6.

[13]. V. Vapnik, and A. Lerner, "Pattern recognition using generalized portrait method", Automation and Remote Control, Vol. 24, 1963, pp. 774-780.

[14]. B.E. Boser, I.M. Guyon, and V.N. Vapnik, "A training algorithm for optimal margin classifiers", 5th annual workshop on computational learning theory, 1992, pp. 144-152.

[15]. A.J. Smola, B. Scholkopf, and K.R. Muller, "The connections between regularization operators and support vector kernels", Neural Networks, Vol. 11, 1998, pp. 637-649.

[16]. G.B. Moody, and R.G. Mark, "The impact of the MIT-BIH Arrhythmia Database", IEEE Engineering in Medicine and Biology Magazine, Vol. 20, No. 3, 2001, pp. 45-50.

[17]. A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C.H. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, and H.E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals", Circulation, Vol. 101, No. 23, pp. e215-e220