ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Design of Lead II ECG Waveform and Classification Performance for Morphological features using Different Classifiers on Lead II.

Akanksha Dhiman*, Ambesh Singh, Shwetanjali Dubey, Shruti Jain.

Department of Electronics and Communication Engineering, Jaypee University of Information Technology, Solan, Himachal Pradesh. 173234. India.

ABSTRACT

The aim of thispaper is hardware optimization of the electronic circuit of electrocardiogram (ECG) by using a bio-potential amplifier in such a manner that it reduces noise, RF interference, DC offset value, and common voltage from the existing circuit.We have also made a CAD system which helps the doctor in analyzing whether patient waveform is normal or abnormal. We have used the knowledge of automatic classification of ECG signals with feature selection and extraction techniques.83 samples of 12 lead systems were collected from various doctors out of whom we have used the lead 2 samples of the ECG waveforms. In this paper, we have accurately classified and differentiated normal and abnormal waveforms. Experiments reveal that the overall classification accuracy is 71.43% using SSVM classifier.

Keywords: ECG, instrumentation amplifier, lead II, extraction, classification

Corresponding author

INTRODUCTION

The electrocardiogram (EKG or ECG) is a graphical recording of the heart generated electric potentials [1-3]. We detect the signals with the help of metal electrodes placed at the extremities and chest wall these signals are amplified and recorded by the ECG. Instantaneous differences occurringin the potential between these electrodes are displayed by the leads.

Here, an instrumentation amplifier is used as a bio-potential amplifier. Instrumentation amplifier possesses several characteristics such as highcommon mode rejection ratio (CMRR), along with higher value of input impedance, high amplification and rejection of electrical interference, hence contributing to an accurate measurement of bio-potentials.

The importance of ECG has grown many folds in last few decades, as we now have an accurate measurement of heart signals and the correct interpretation along with simplicity has resulted in saving millions of lives. In ECG, heart's electrical activity versus time is analyzed as a signal and then we interpret the heart functioning. Now a days we have the computerized analysis of ECG waveforms which has significantly reduced the workload of doctors.

In this paper we are aiming to simplify the existing ECG circuitry as simple as possible and developing a system that will automatically categorise the ECG signals in to various classes as per the abnormalities for the easier interpretation of results by the doctor.

MATERIALS AND METHODS

HARDWARE DESCRIPTION

Heart signals are picked up by an ECG machine.We place electrodes externally to specific body locations i.e. thelegs and thearms and then heart signals of very small amplitude (few mill volts). To view electrical activities of the heart from different anglesthe electrodes are placed at specific locations on the body and each one is represented as a lead on the ECG printout. We have 12 leads in ECG and these 12-leadECG are recorded using left arm (LA), right arm (RA), left leg (LL), right leg (RL), and chest (C) electrodes.Categorising our ECG on basis of lead systemit can be divided into two standard planes i.e. transverse plane and the frontal plane. The leads can be further divided into the chest leads the bipolar limb leads and the unipolar leads.In case of Bipolar limb leads signals are derived from electrodes on the limbs, and we can designate them as lead I (RA to LA), lead II (RA to LL), andlead III (LA to LL).Coming to the unipolar leads we can designate them as aVR, aVL, and aVF, which we can design by connecting RA, LA and LLrespectively to the non-inverting terminal.The remaining two electrodes are connected tothe inverting terminal of the Instrumentation amplifier. Now the chest leadscomprises of the six remaining leads i.e V_1,V_2,V_3,V_4,V_5 andV_6[2-4].

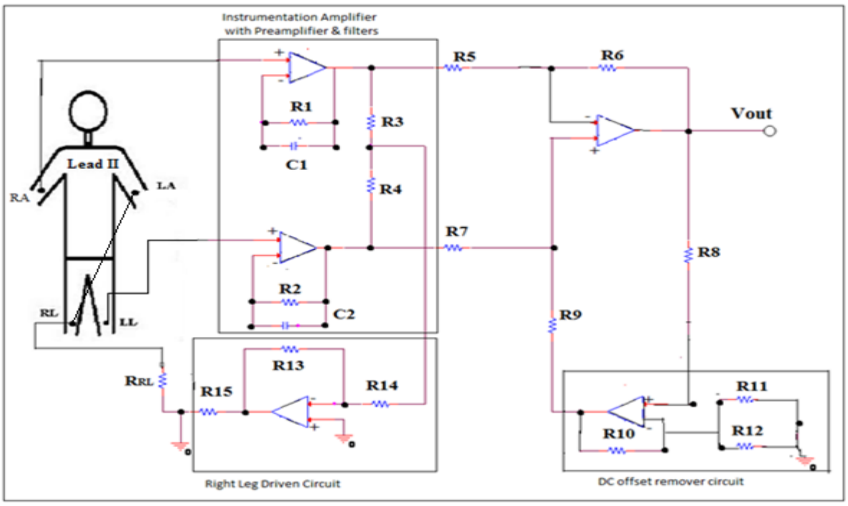

Figure1:Proposed Circuit Diagram

In the Fig. 1 shown below we have proposed the circuit diagram of ECG system. In the circuit diagram shown below we have used instrumentation amplifier as a bio potential amplifier, then there is a high-pass filter using an inverter op-amp stage as a feedback which helps in blocking the DC offset value and then there is a right leg driven circuit which helps in reducing the common mode value and noise. This proposed circuit diagram is relatively smaller and is more compact in size than the existing circuit [4].

Proposed CAD System Design

For the design of CAD system, a dataset of 83 samples of ECG waveform of various patients were taken. Our analysis consists of three stages which are Feature selection, Feature extraction and Feature classification [5].

Feature Selection

In the selection process, we have used around 83 samples of patients collected from different doctors and selected the appropriate lead 2 waveforms out of the 12 lead system.

Feature Extraction Module

The process of transforming the visually non-extractable and extractable features into mathematical descriptors is known as feature extraction. These descriptors are taken as shape-based (morphological features) and the intensity distribution based (textural features). The intensity distribution methods include (*a*)**T**ransform domain methods (*b*) Signal Processing Based Methods (*c*) Statistical Methods.In this paper we are using the morphological extraction method.

Shape based properties like Area, Perimeter, Convexity, Eccentricity, Extent, Hole Area Ratio (HAR) and Solidity are calculated over the entire class of normal and abnormal waveforms.

1. *Convexity:* It is the ratio of the perimeter of the convex hull to the overall contour.
2. *Eccentricity:* It is the ratio of the minor axis to the major axis. Its value always lies between the 0 and 1.
3. *Extent*: It is the ratio of the area to the bounding area.

$$\text{Extent} = \text{Area}/\text{Bounding Area} \tag{1}$$

4. *Hole Area Ratio (HAR):* It is the ratio of the area of the holes in a shape to the area of the whole shape.

$$\text{HAR} = \text{Area of holes in shape} / \text{Area of the shape} \tag{2}$$

5. *Solidity:* It gives the extent to which the given shape is convex or concave.

$$\text{Solidity} = \text{Area} / \text{Convex Area} \tag{3}$$

Feature Classification

Feature classification is a technique that is mainly used for predicting the class of any unknown data according to the set of known classes of data. In this module we will be using three different types of classifiers that are k-NN, SSVM and SVM .These classifiers will be usedto classify unknown data. We will try to find out the unknown instances of data in testing set with the help of known data instances in training set. To avoid biasing due to uneven featured values the extracted features obtained are normalized with the use of min-max normalization process in the range [0, 1] [10-18].

***k*-NearestNeighbor (*k*-NN) Classifier**

This classifier helps in determining the class of an unknown data with the help of its nearest neighbours. It also attempts to group the data containing feature vectors into disjoint classes by assuming that the data placed in the neighbourhood of each other in the feature space represent instances which will belong to the same class. The object present in the set is given that particular class which is the most common among its k nearest neighbours. The procedure is carried between testing and training set.

Support Vector MachineClassifier (SVM)

This classifier belongs to a class of machine learning algorithms of high level . In SVM, kernel functions are used to map the non-linear training data from input space to a high dimensionality feature space. It is based on the concept of decision planes that define the decision boundary. The important step to obtain a good generalized performance is the correct choice of parameter C called the regularization parameter and the kernel parameter .The parameter C attempts to keeping the training error low and maximize the margin [16-18].

Smooth Support Vector Machine Classifier (SSVM)

To explain significant mathematical problems linked tosmoothing and programming methods are widely used. It works on the idea of even unrestrained optimization reformulation based on the conventional quadratic program. Just like in SVM implementation in SSVM also, on the training data for each combination, ten-fold cross justification is carried out. This grid search procedure in parameter space gives the optimum values of C and for which training precision is maximum.

RESULTS & DISCUSSIONS

Figure 2 shows the block diagram of the proposed CAD system for two-class classification using morphological features.

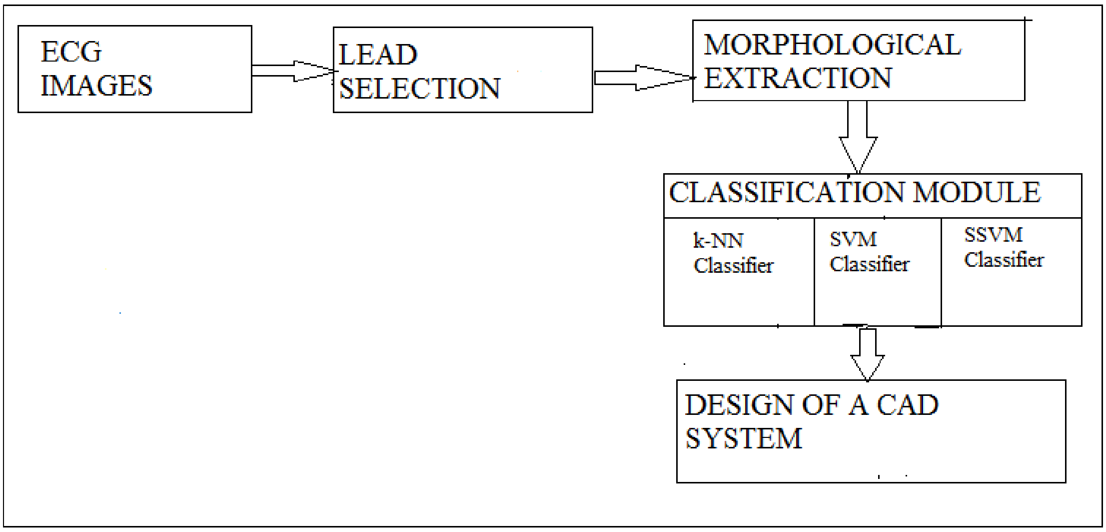

Figure 2: Proposed CAD system design using statistical features

The approach is implemented on the 93 samples collected from doctors. The samples of lead 2 were selected. Training and testing data was calculated using different morphological features are calculated like Area, Perimeter, Solidity, Convexity, Extent etc. Different classifiers like *k*-NN, SVM and SSVM were used on 48 training images (20 normal and 28 abnormal) and 35 testing images (15 normal and 20 abnormal) and confusion matrix was made.

The confusion matrix is a 2x2 matrix containing four components. These are:-

Confusion Matrix		Predicted	
		Normal	Abnormal
Actual	Normal	**p**	**q**
	Abnormal	**r**	**s**

- *p* which gives the **correct** predictions that the value is **normal**,
- *q* which gives the **incorrect** predictions that the value is **abnormal**,
- *r* which gives the **incorrect** of predictions that the value is**normal**, and

- *s* which gives the **correct** predictions that the value is **abnormal**

Several standard terms are defined in this 2 class matrix:

- The ***accuracy*(*AC*)** is the ratio of the total number of valueswhich are correct to the total number of values. It is given by the equation:

$$AC=(p+s)/(p+q+r+s) \quad (4)$$

- The ***Specificity* or *true abnormal rate* (*TAN*)**is given by equation 5:

$$TAN = s/(r+s) \quad (5)$$

- The***false abnormal rate* (*FAN*)** is given by equation 6:

$$FAN = q/(p+q) \quad (6)$$

- The ***Sensitivity*** **or *true normal rate* (*TN*)** is defined as the proportion of normal cases that were classified correctly.

$$TN = p/(p+q) \quad (7)$$

- The***false normal rate (FN)*** is the proportion of abnormal cases that were incorrectly classified as normal, as calculated by the equation:

$$FN = r/(r+s) \quad (8)$$

- Finally, ***precision* (*P*)**is given by equation 9:

$$P = s/(q+s) \quad (9)$$

The calculations of these features help us to classify the collected waveforms as normal or abnormal. The table 1 show us how we try to get the result of testing and training set by using different classifiers . The main idea is to divide them into four different categories using confusion matrix which gives us different results accordingly.

We define two categories which are Normal and Abnormal. Normal gives us the waveform results which are familiar to the normal ECG results obtained. The abnormal gives us the waveforms which are varied than the normal ECG due to some disorders caused by various reasons.

Table 1: Classification analysis of morphological features using *k*-NN, SVM and SSVMclassifiers

Classifier		**CM**		**OCA (%)**	**ICA_N(%)**	**ICA_{AN} (%)**
		N	AN			
k-NN	N	11	4	54.29	73.33	40
	AN	12	8			
SVM	N	9	6	68.57	60	75
	AN	5	15			
SSVM	N	5	10	71.43	33.33	100
	AN	0	20			

Note: CM: Confusion matrix, ICA_{AN}: Individual class accuracy for abnormal class; OCA: Overall classification accuracy;ICA_N: Individual class accuracy for normal class.

Then 83 samples of 12 lead systems were collected from various doctors out of whom lead 2 samples of the ECG waveforms are used to accurately classify and differentiate normal and abnormal waveforms using a CAD system. The CAD system produces a confusion matrix which shows that the highest OCA of 71.43 % is obtained from SSVM classifier. The highest individual accuracy obtained for normal class is 73.33% from k-NN classifier and for abnormal class is 100% from SSVM classifier.

In the table below are the values obtained with these three classifiers:

Table 2: Values of standard terms using k-NN,SVM and SSVM classifiers.

Terms/Classifier	k-NN	SVM	SSVM
AC	0.54	0.68	0.71
TAN(Specificity)	0.4	0.75	1
FAN	0.26	0.4	0.66
TN(Sensitivity)	0.73	0.6	0.33
FN	0.6	0.25	0
P	0.66	0.71	0.66

CONCLUSION

This work indicates the performance of ECG circuit in terms of electronics using an instrumentation amplifier. It further explains several procedures carried out to reduce the circuitry to increase CMRR and noise with the help of right leg drive circuit. We can diminish DC offset with the help of high pass filter and by filtering we can reduce RF interference.Lastly, we have integrated all the parts and made the existing circuit compact in size. Then 83 samples of 12 lead systems were collected from various doctors out of whom lead 2 samples of the ECG waveforms are used to accurately classify and differentiate normal and abnormal waveforms using a CAD system. The CAD system produces a confusion matrix which shows that the highest OCA of 71.43 % is obtained from SSVM classifier.

REFERENCES

[1] N. V. Thakor, Encyclopedia of Medical Devices and Instrumentation.
[2] Carr, J.J.; Brown, J.M. Introduction to Biomedical Equipment Technology; Pearson Education, Inc.: edition 7th; Chapter 8.
[3] S. Franco, Design with Operational Amplifiers, New York: McGraw-Hill, 1988.
[4] Gupta R, Singh S, Garg K, Jain S, International Journal of Innovative Research in Science, Engineering and Technology, May 2014; 3: 12138-12145.
[5] Bhusri S , Jain S, Virmani J, International Journal of Pharma and BioSciences (IJPBS), 2016 ;7 : 617-624.
[6] Bhusri S, Jain S, Virmani J, 10th INDIACom: 3rd 2016 International Conference on Computing for Sustainable Global Development, New Delhi, 2016, pp 2523-2527.
[7] Rana S , Jain S, Virmani J, 10th INDIACom: 3rd 2016 International Conference on Computing for Sustainable Global Development, BVICAM, New Delhi, 2016, pp 2528-2532,.
[8] Weszka JS, Dyer CR, Rosenfeld A. IEEE Transactions on Systems, Man and Cybernetics, 1976; 6 : 269-285.
[9] Castellano G, Bonilha L, Li LM, Cendes, 2004; 59: pp. 1061-1069.
[10] Jain S, Naik PK, Bhooshan SV, International Conference on Computational Intelligence and Communication Networks (CICN2011), Gwalior, India ,Oct 07-09, 2011, pp 565-568.
[11] Raja BK, Madheswaran M, Thyagarajah K, J Med Syst, 2008; 1: 65–83.
[12] Jain S, Naik PK,International Journal of Pharma and BioSciences (IJPBS), 2012; 4: 358-373.
[13] Jain S, Chauhan DS,International Journal of Pharma and Bio Sciences (IJPBS), 2015;3: 164-176.
[14] Jain S. Control System and Power Electronics – CSPE 2015, Bangalore ,Aug 1-2, 2015.
[15] Jain S, Chauhan DS. International Conference on Information and Communication Technology for Sustainable Development (ICT4SD - 2015), Ahmedabad, India.
[16] Jain S. PhD Thesis, Jaypee University of Information Technology, Solan, Himachal Pradesh, India, 2012.
[17] http://www.csie.ntu.edu.tw/~cjlin /libsvm, Accessed 15 Jan 2016.
[18] Virmani J, Kumar V, Kalra N, Khandelwal N. Journal of Digital Imaging, 2013; 3: 530-543.