International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-6 Issue-2, December 2016

ANN based Multilevel Classification Technique with Optimum Measurement Period for Accurate Diagnosis using Biomedical Signals

Paul Thomas, R.S. Moni

Abstract: ***Biomedical signals are representations of the mechanical and electrical activities within the human body. These signals contain a lot of information on the state of health of a person and their analysis have a significant role in the diagnosis of various health disorders and medical abnormalities, such as activation levels and the biomechanics of the muscles and other human organs. Of the many Biomedical signals, focus of this work is on Electro-cardiogram (ECG) and Electro-myogram (EMG). ECG provides information on the rhythm and functioning of the heart. EMG is the recording of human muscular activity. ECG signals used in this work are taken from the standard MIT-BIH, and CU data bases of PhysioNet database and EMG signals are taken from the EMGLab and PhysioNet database. Automated analysis of Biomedical signals can largely assist the physicians in their diagnostic process. The extracted spectral and temporal features represent the diverse characteristics of a Biomedical signal. In this work, more emphasis is given to spectral features since a lot of critical information on the health of a person are hidden in the spectral content of the signal. A subset from a larger set of available features is experimentally selected for optimum performance. The feature vector has a size of 11 for ECG signal analysis and a size of 9 for EMG signal analysis. Accuracy of detecting a health disorder depends on the quality of the features extracted from a Biomedical signal. A few techniques are proposed to achieve improved quality for the features. Also a method is developed to arrive at the optimum length of the Biomedical signal to be used for analysis. Accordingly, the length of the ECG signal used in this work is 10 s and the length of the EMG signal is 11 s. It is observed that the variance of the features is minimum when the signal for analysis is taken from the mid portion of the whole Biomedical signal. To make the value of a feature close to its true value, each feature value is taken as the average of the values of the feature extracted from 20 consecutive signal segments. A technique is also proposed to reduce the effect of wild points in the computation of spectral parameters. It is observed that classification accuracy also depends on the sampling rate of the Biomedical signal. The sampling rate of ECG signal in this work is 128 Hz and that of EMG signal is 750 Hz. Classifying a Biomedical signal is the process of attaching the signal to a disease state or healthy state. The work proposes a Multi level classification approach for Biomedical signals. Each classifier is a cascade of two ANN classifiers, the first ANN has a linear transfer function and the second ANN has a sigmoid transfer function. First level classification is to the broad categories of the disorders. In the second level, these disorders are drilled down to more specific categories. This concept can be extended further to achieve finer classification of Biomedical signals. In this work the classification is demonstrated to two levels for ECG signals and one level for EMG signals. The performance of the proposed method is evaluated using the standard parameters of specificity, sensitivity and classification accuracy (CA). The performance is found to be better than the reported figures in the case of both ECG and EMG signals.***

Index Terms: ***ECG, EMG, FFT, DWT, Pattern recognition ANN, Feature extraction, Multilevel classification, Wavelet, PhysioNet database, CA, Atrial arrhythmias, Ventricular arrhythmias, NSR, MI, MUAP, Myopathy, ALS.***

I. INTRODUCTION

Biomedical signals are non stationary signals [1]. Among the many Biomedical signals, only ECG and EMG are considered in this work. ECG contains a plethora of information on the normal and pathological conditions of the heart. EMG is the electrical record of the activity of the skeletal muscles during different levels of muscular contractions and they contain a train of Motor unit action potentials (MUAP). It is tiresome and time consuming for Physicians to examine these long records of Biomedical signals and arrive at a conclusion. Moreover, it is also likely for them to miss out some minutiae details in the visual examination process. Some of the health disorders like Ventricular fibrillation and Amyotrophic lateral sclerosis (ALS) may turn out to be fatal [2], if they are not diagnosed at an early stage. The features extracted from a Biomedical signal are representative of the information hidden in the signal. Automated analysis of Biomedical signals have an important role to play in assisting a Physician in his diagnostic process [3].

Biomedical signals are monitored either non-invasively or invasively, the former being the most common method where electrodes placed on the surface of the skin capture the signal. Examples of such signals are ECG and surface EMG. Intramuscular EMG and Intracranial EEG are examples of Biomedical signals monitored invasively. Intramuscular EMG is monitored with needles inserted into the muscles. Invasive methods provide better quality signals leading to better quality features. But they are not available for all types of Biomedical signals and for all categories of health disorders.

Manuscript published on 30 December 2016.
***** Correspondence Author (s)
Paul Thomas*, Department of Electronics and Communication Engineering, Mar Baselios College of Engineering and Technology, Nalanchira, Trivandrum, India.
Dr. Moni R.S., Department of Electronics and Communication Engineering, Marian Engineering College, Menamkulam, Trivandrum, India.

Moreover, the general requirement is to detect health disorders from non invasively recorded Biomedical signals. This work uses signals that are recorded by non invasive methods.ECG signals from MIT-BIH and CU databases in PhysioNet database are used in this work. EMG signals are taken from EMGLab and PhysioNet database. The ECG signals correspond to Normal sinus rhythm (NSR) and six diseased conditions of Myocardial infarction(MI), Premature ventricular contraction (PVC), Ventricular tachycardia (VT), Malignant ventricular ectopic (MVE), Atrial fibrillation (AF) and Supra ventricular arrhythmia (SVA). The EMG signals correspond to Normal and four diseased conditions of Myopathy, Amyotrophic lateral sclerosis (ALS), Huntington's disease and Parkinson's disease.

Different techniques are reported in literature for feature extraction and classification of ECG signals. The techniques for ECG include Principal component analysis (PCA), Linear discriminant analysis (LDA) and Independent component analysis (ICA), Fuzzy C-means clustering (FCM), Discrete wavelet transform (DWT), Support vector machine (SVM), Artificial neural network (ANN), Fuzzy logic and a combination of these techniques. The techniques for feature extraction and classification of EMG signals include ICA, DWT, k Nearest neighbor (k-NN), SVM, soft computing techniques, dominant MUAP based techniques, Radial basis function neural network (RBFNN), and weight visibility algorithm with Multi layer perceptron (MLP) neural network.

Each of the reported classification methods use only a small set of features of ECG and EMG signals. In this work a larger feature set is used, which is a combination of the different types of features encompassing the diverse characteristics of the Biomedical signal. Use of different types of features is seen to provide better classification accuracy than one type of feature alone. These features include temporal and spectral/ wavelet/ statistical features with the highest emphasis on wavelet features. High detection accuracy of 99% [4] are reported only up to three disease conditions and normal. Wherever more disease conditions are addressed, the detection accuracy is relatively low, up to 96.8% [5]. The proposed technique works on a bigger feature set, and can detect multiple health disorders, at the same time provide accuracy figures on par with the highest levels reported. This technique can be extended to include more categories of Biomedical signals and more number of health disorders.

A modular approach is proposed for the classification of the extracted features. A module is a cascade of two ANN classifiers. The first ANN has a linear transfer function and the second ANN has a sigmoid transfer function. Each ANN is trained with a subset of the features extracted from the signals. Once training and validation are completed, the neural network architecture is formed. The second ANN classifies the first ANN output, for improved classification accuracy. First level classification of ECG signals is into four classes and first level classification of EMG signals is into five classes. This work also proposes a multi level classification approach for finer classification of Biomedical signals. Second level classification of ECG signals is to the finer levels of Atrial arrhythmias and Ventricular arrhythmias. For EMG signals, second level classification is not done due to non availability of respective signals.

This paper is organized as follows. Materials and methods used in this work are discussed in section 2, the results are discussed in section 3 and conclusions are made in section 4.

II. Materials and Methods

A. Materials Used in this Work

a. Sources of Biomedical Signals

ECG signals for this work are taken from CU and MIT-BIH databases in the PhysioNet database. These are 18 signals of NSR at 128 Hz, 128 signals of MI at 1000 Hz , 105 signals of Ventricular Arrhythmias (48 signals of PVC at 360 Hz, 35 signals of VT at 250 Hz, and 22 signals of MVE at 250 Hz), and 184 signals belonging to Atrial arrhythmias (22 signals of AF at 250 Hz, 78 signals of SVA at 128 Hz, and 84 signals of Long term AF at 128 Hz). EMG Signals for this work are taken from EMGLab and PhysioNet database. The signals from EMGLab are 300 Normal signals, 289 Myopathy signals and 178 ALS signals all at 23438 Hz. The oversampling is justified as a method to achieve adequate temporal precision. 15 signals of Parkinson's disease and 20 signals of Huntington's disease each at 300 Hz are taken from PhysioNet database. These signals are segmented to make the number of signals as 150 and 160 respectively. More details of the ECG signals and the EMG signals are available in the corresponding database.

b. Tools Used in this Work

The two stages in Biomedical signal analysis are feature extraction and classification of the extracted features. These tasks are carried out in MATLAB R2014a development environment, with specific use of Signal processing tool box, Wavelet tool box and Neural network toolbox.

B. Methods Used in this Work

The Accuracy of classification depends, to a large extent, on the quality of the features extracted from the Biomedical signal and also on the method of classification. Good quality for the extracted feature is very important in automated analysis of a signal to detect health disorders. The quality of a feature becomes better and better as it approaches its true value. Following four steps are adapted to improve the quality of the features derived from a Biomedical signal.

a. Optimum Length (Time duration) of Biomedical Signal for Analysis

The length of the Biomedical signal used for analysis is found to play an important role on the quality of the features extracted from the signal. There is not much information available in the literature on how to arrive at an optimum length of the signal for analysis. Researchers have been using different lengths of signals in their work, such as 1, 3 and 8 seconds for ECG signal, and 2 and 3 seconds for EMG signal. A novel method is proposed here, to arrive at the optimum length of the signal to be used for analysis.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The Biomedical signal is segmented to durations of 1s each. Features are extracted from 50 consecutive segments taken first from the mid portion of the signal, then from the beginning, and finally from the last portion of the signal. The variance of the extracted features is computed for each of these three sets. This process is repeated up to a segment duration of 25 s, each time incrementing the duration by 1s. This exercise is done for all available ECG and EMG signals. For signals of shorter duration, the number of consecutive segments is less than 50. It is observed that the variance of features is minimum for segments taken from the mid portion of the Biomedical signal. Accordingly the classification performance is better with features extracted from the mid portion segments. It is also observed that the variance of the extracted features are different for different segment durations. The observations are tabulated in TABLE I for features extracted from the mid portion segments.

TABLE I. VARIANCE OF FEATURES VS SEGMENT DURATION

ECG segment duration (sec)	EMG segment duration (sec)	Variance (%)
20	20	< 2%
15	15	< 3%
10	11	< 4%
5	6	< 5%

For both ECG and EMG signals, it is observed that there is no change in the classification accuracy, if the variance of the features is between 2% and 4%. But there is a decrease in the classification accuracy when the variance is 5%. Since the computational complexity increases with the size of the signal segment, 4% variance is taken as the acceptable level. In this work, the duration of the ECG segments is 10 s and that of the EMG segments is 11 s. The optimum duration of the segment for any type of Biomedical signal analysis can be found out in a similar manner.

b. Averaging of Feature Values

This is another technique that is proposed to make the value of a feature close to its true value. The features are extracted from segments in the mid portion of the Biomedical signal. The size of each segment is 10 seconds for ECG signal and 11 seconds for EMG signal. It is observed that the mean value of the features extracted from the various segments stabilizes at around 20 consecutive segments. In this work, each feature value is taken as the mean of the features extracted from 20 consecutive segments taken from the mid portion of the Biomedical signal.

c. Standardizing the Signal Sampling Frequency

Biomedical signals are recorded at different places using different recorders. The signal sampling rates are not the same in the various recordings. It is observed that there is a variation in the value of the features when the sampling rate is changed. This has an impact on the classification accuracy. The optimum sampling rate for a type of Biomedical signal is arrived at by observing the classification accuracy for the signal at different sampling rates. The sampling rate at which the classification accuracy is best is taken as the optimum sampling rate. Accordingly, the sampling rate used in this work is 128 Hz for ECG signal analysis and 750 Hz for EMG signal analysis. The sampling rate is converted to this optimum value before features are extracted from the signal. The optimum sampling rate is seen to slightly vary if a different feature set is used for classification. In real life situations, the optimum sampling rate will not change because the feature set is fixed and the Biomedical signals are captured by machines operating at fixed sampling rates. This technique provides a method to set the sampling rate to its optimum value, for improved classification performance.

d. Reducing the Effect of Wild Points in the Computation of Spectral Parameters

The spectral features of PF and MF are computed from FFT samples. Some wild points are observed in the FFT samples, predominantly towards the edges. In Biomedical signals, the frequency of interest starts from low frequency of the order of 0.2 Hz. Thus we cannot eliminate these FFT samples. But, we can expect a reasonable correlation between adjacent FFT samples. In order to reduce the effect of wild points, the ratio of the amplitudes of adjacent FFT samples is restricted to three. FFT samples are then averaged in durations of 0.5 Hz each and the coarse value of the feature is computed. Fine computation within the coarse band is done using the samples in that 0.5 Hz band.

e. Selection of the Lead for ECG Signal Analysis

It is observed that spectral features extracted from different leads in a 12 lead ECG system is more or less consistent. Temporal and statistical features are more pronounced in lead II. Thus signal from lead II is used for ECG signal analysis.

f. Features and Extraction of Features

The various temporal/morphological and spectral/ statistical/ wavelet features of ECG and EMG signals, used by various researchers are identified. Morphological and temporal features are signal specific. The morphological features specific to ECG signal are S-T interval, T-T interval, Q-T interval, R-R interval, P-R interval, P-P interval, Peak value of QRS, Peak value of P, Peak value of T, and Height to length ratio of S-T segment. The morphological features specific to EMG signal are Peak to peak amplitude, Rise time, Duration of MUAP, Spike duration and Area under the MUAP. The temporal features of the ECG signal are Heart rate, Threshold crossing interval, Threshold crossing count (TCC), Standard exponential, and Modified exponential. Temporal features of the EMG signal are Number of samples between peak to peak values of MUAP, and Integrated EMG..

The spectral features common to ECG and EMG signals are Peak frequency (PF), Mean frequency (MF), Median frequency, Mean power (MP) of the spectrum, First spectral moment (SM), Second, Third and Fourth spectral moments. The spectral features specific to ECG signal are Spectral band amplitude (SBA), VF leakage (VFL). The wavelet features common to ECG and EMG signals are Maximum and minimum values of approximation coefficients, Maximum and minimum values of detail coefficients, Wavelet mean (WM), Wavelet standard deviation (WSD), Wavelet average power (WAP),

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Mean absolute deviation of approximation coefficients, Skewness (SKW), and Kurtosis (KURT). The statistical features common to ECG and EMG signals are Mean, Variance (Var), Simple square integral (SSI), and Mean absolute value (MAV).

i. *Features for the First Level Classification*

Different feature sets are tried out in the classification of ECG signals and EMG signals. Since the number of possible combinations of the features is enormous, an exhaustive analysis to identify the optimum feature set is not feasible. The choice of the feature set is the rationale of the researcher. Accordingly, feature sets are chosen for the first level classification of ECG signals and for the first level classification of EMG signals. Morphological features are not used in this work. Evolving a technique to mathematically arrive at an optimum feature set for a category of Biomedical signals is an area for future study.

The feature vector for the first level classification of ECG signals contain 11 features: Temporal feature (TCC) [6], Statistical feature (MAV) [6], Spectral features (PF, MF, SM, SBA and VFL) [7] and Wavelet features of (WM, WSD, SKW and KURT) [8]. The feature vector for the first level classification of EMG signals contain 9 features: Statistical features (SSI, MAV, and Var) [9], Spectral features (PF, MF and MP) [9] and Wavelet features (WM, WSD and WAP) [8]. In this work, the first level category disorders for ECG are Ventricular arrhythmias, Atrial arrhythmias, MI, and NSR. The first level category disorders for EMG are Myopathy, ALS, Huntington's disease, Parkinson's disease, and Normal.

ii. *Features for the Finer Classification*

Feature sets for the higher level (finer level) classifications are sub sets of the feature set for the first level classification. In the second level classification, each output of the first level classification is drilled down to the next level categories. In this work, second level classification of ECG is carried out for Ventricular arrhythmias and Atrial arrhythmias. Ventricular arrhythmias are classified into VT, PVC and MVE. Atrial arrhythmias are classified into AF and SVA. MI and NSR are not further classified due to non availability of signals. In a similar way, it is possible to carry out finer classification of EMG signals. The possible sub classes of ALS are Progressive muscular atrophy, Primary lateral sclerosis, and Juvenile amyotrophic lateral sclerosis.

The optimum feature set for second level classification can be selected in a similar way as it is done for the first level classification. Accordingly, the set of features for the classification of Ventricular arrhythmias are TCC, PF, MF, SM, VFL, WM, WSD, KURT and SKW. The set of features for the classification of Atrial arrhythmias are MF, SM, MAV, WM, WSD and KURT. Many of the features have their values overlapping with the same features extracted from signals belonging to other category of health disorders. This limits our ability to minimize the number of features required to achieve good classification accuracy.

g. Feature Extraction Methods

The various features that are extracted from the ECG and EMG signals, for their classification are listed below.

i. *Peak Frequency (PF)*

PF is the frequency at the maximum sample amplitude of an N-point FFT of the signal.

$$PF = f_T$$

T : FFT sample number with highest amplitude

a_i : *Amplitude of* i^{th} *FFT sample,* $1 \le i \le N$

ii. *Mean Frequency (MF)*

MF is the central frequency of the power spectrum.

$$MF = \frac{\sum_{i=1}^{N} f_i \times (a_i)^2}{\sum_{i=1}^{N} (a_i)^2}$$

iii. *Wavelet Mean (WM)*

WM is computed from the approximation coefficients after a four level Daub 4 wavelet decomposition.

$$WM = \frac{1}{L}\sum_{i=1}^{L} y_i$$

y_i : Amplitude of the i^{th} aproximation coefficient

L : Number of approximation coefficients

iv. *Wavelet Standard Deviation (WSD)*

WSD is the standard deviation of the approximation coefficients

$$WSD = \sqrt{\frac{1}{L}\sum_{i=1}^{L} (y_i - WM)^2}$$

v. *Mean Absolute Value (MAV)*

MAV is mean of the absolute value of the signal samples.

$$MAV = \frac{1}{N}\sum_{i=1}^{N} |e_i|$$

e_i : Amplitude of the signal samples

vi. *VF Leakage(VFL)*

$$VFL = \frac{\sum_{i=T/2}^{N} |(a_i + a_{i-T/2})|}{\sum_{i=T/2}^{N} (|a_i| + |a_{i-T/2}|)}$$

T is the FFT sample number at the Peak frequency.

vii. *Spectral Moment(SM)*

$$SM = \frac{\sum_{i=1}^{N} f_i \times a_i}{\sum_{i=1}^{N} a_i}$$

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

viii. Spectral Band Amplitude (SBA)

$$SBA = \left\{ \sum_{i=0.7PF}^{1.4PF} a_i \right\} \div \left\{ \sum_{i=sample\#\ at\ 0.5Hz}^{\min(20PF\ and\ F_s/2)} a_i \right\}$$

F_s: Sampling frequency of the ECG signal.

ix. Skewness (SKW)

SKW is a measure of the asymmetry of the wavelet approximation coefficients.

$$SKW = \frac{1}{L} \frac{\sum_{i=1}^{L} (y_i - WM)^3}{(WSD)^3}$$

x. Kurtosis(KURT)

KURT is a measure of the peakiness of the wavelet approximation coefficients with respect to normal distribution.

$$KURT = \frac{1}{L} \frac{\sum_{i=1}^{L} (y_i - WM)^4}{(WSD)^4}$$

xi. Threshold Crossing Count(TCC)

TCC is the number of crossings of the ECG signal with a threshold set at 80% of maximum value of the ECG samples.

xii. Mean Power (MP)

$$MP = \frac{1}{N} \sum_{i=1}^{N} (a_i)^2$$

xiii. Wavelet Average Power (WAP)

$$WAP = \frac{1}{L} \sum_{i=1}^{L} (y_i)^2$$

xiv. Simple Square Integral (SSI)

$$SSI = \frac{1}{N} \sum_{i=1}^{N} (e_i)^2$$

xv. Variance (Var)

$$Var = \frac{1}{N} \sum_{i=1}^{N} (e_i - MAV)^2$$

h. Classification of the Biomedical Signals

ANN classifier is trained using supervised learning rule. ANN consists of three layers, input layer, hidden layer and output layer. The feature vector is fed to the input layer. Number of neurons in the input layer is equal to the size of the feature vector. All computations are carried out in the hidden layer which does not have any contact with the outside world. Number of neurons in the output layer is equal to the number of classes.

i. Modular Approach to Classification

This work proposes a modular approach for the classification of Biomedical signals. Each classifier is a module with a cascade of two ANN classifiers. The first ANN classifier has a linear transfer function and the second ANN has a sigmoid transfer function. The sigmoid function is configured with two parameters as

$$z_i = \frac{1}{1 + \exp\{-c(x_i + t_i)\}}$$

t is the time parameter and c is a measure for the steepness. z is the output and x is the input. c has a positive value and for larger values of c the sigmoid function approximates a Heaviside function. The value of c is around 2 and is experimentally set for the best classification accuracy at each level of classification.

ii. Selection of Neural Network

Classification of the health disorders using the extracted features is attempted with Pattern recognition network and also with Linear vector quantization network. Pattern recognition network is observed to give higher classification accuracy for ECG and EMG signals, by about 3%. The pattern recognition network architecture used in this work is obtained after training and validation of the network. The architecture of the Neural networks in the first level classification of ECG signals is shown in FIG 1 and FIG 2.

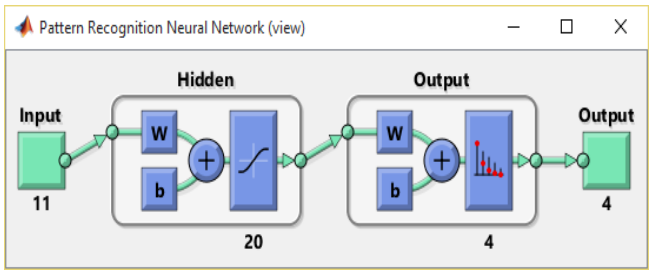

FIG 1. FIRST ANN IN THE FIRST LEVEL ECG SIGNAL CLASSIFICATION

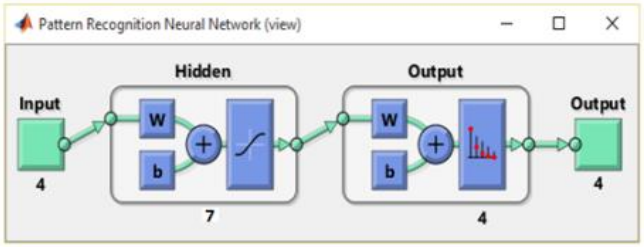

FIG 2. SECOND ANN IN THE FIRST LEVEL ECG SIGNAL CLASSIFICATION

The number of neurons in the input layer (I) is equal to the size of the feature vector and the number of neurons in the output layer (O) is equal to the number of classes. The number of neurons in the hidden layer (H) is obtained using the empirical relation H = [(0.67*I + O), 2*I]. The size of H is trimmed experimentally for optimum performance in terms of classification accuracy. The number of hidden neurons is 20 for the first classifier in the first level classification of ECG. The number of neurons in the input layer, hidden layer and output layer, used in this work, are tabulated in TABLE II. The first row in the second level classification corresponds to Ventricular arrhythmias and the second row corresponds to Atrial arrhythmias.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

TABLE II. NEURONS IN THE LAYERS OF ANN CLASSIFIERS

Signal	1st level classification						2nd level classification					
	1st classifier			2nd classifier			1st classifier			2nd classifier		
	I	H	O	I	H	O	I	H	O	I	H	O
ECG	11	20	4	4	7	4	9	16	3	3	5	3
							6	10	2	2	4	2
EMG	9	17	5	5	9	5	Not done					

iii. Multilevel Classification

This work proposes a multilevel classification technique to detect health disorders from Biomedical signals. This technique allows detection of health disorders to any fine level without imposing any restriction on the number of disorders for classification. The limitation is the availability of sufficient number of signals of each health disorder, to configure the classifier for finer levels of classification. This classification technique can be extended to any Biomedical signal. Block diagram of the multilevel classification scheme is shown in FIG 3.

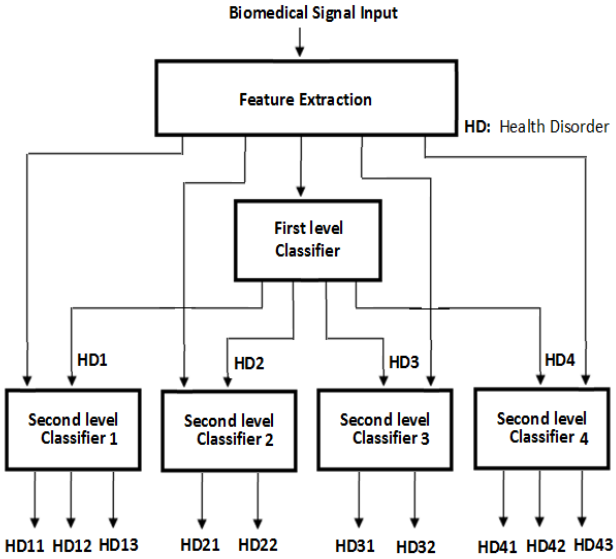

FIG 3. MULTILEVEL CLASSIFICATION SCHEME

III. RESULTS AND DISCUSSIONS

The ECG signals and EMG signals used in this work correspond to normal and various diseased conditions. The ECG signals belong to NSR, AF, SVA, MI, PVC, VT and MVE. The EMG signals belong to Normal, Myopathy, ALS, Parkinson's disease and Huntington's disease. The identified features are extracted from the ECG signals of length 10s and from the EMG signals of length 11s.

Typical values of the features extracted from ECG signals are listed in TABLE III and typical values of the features extracted from EMG signals is listed in TABLE IV.

TABLE III. TYPICAL FEATURE VALUES OF ECG SIGNALS

	NSR	MI	AF	SVA	PVC	VT	MVE
PF	19.6	33.5	25.7	13.8	37.9	26.1	26.4
MF	19.3	5.3	27.8	22.4	26.3	22.1	39.1
WM	39.4	-0.2	-13	6.4	3.9	0.1	-25.1
WSD	116	1.4	286	127	139	1	111
MAV	0.33	0.79	1.2	0.48	0.82	0.2	0.25
TCC	12	35	13.3	11	14.7	20	10.3
VFL	0.88	1	0.97	0.98	1	0.98	0.85
SM	1.1	0	1.5	1.1	0	0	0.5
SBA	0.04	0.18	0.09	0.03	0.15	0.09	0.22
SKW	0.8	1.3	-1.8	0.01	-0.44	-0.1	1.96
KURT	2.38	5.5	5.75	2.69	4.75	3.09	2.73

TABLE IV. TYPICAL FEATURE VALUES OF EMG SIGNALS

	Myo-pathy	ALS	Normal	Parkin-son's	Huntin-gton's
PF	147.3	91.2	178.1	4.6	8.7
MF	196.5	228.7	212.6	89.1	102.8
WM	0.1	0.2	0.3	19.7	28.1
WSD	39.1	26.7	34	3.2	4.9
MAV	63.4	68.8	56.8	28	44.6
MP	3162.1	3743	2743.7	743.3	1617.9
WAP	9.7	5.1	81.9	0.2	0.1
SSI	304	325.4	291.5	17.5	13.4
Var	1.1	1.3	1.2	38.1	56.5

A. Performance Evaluation

The classification is carried out with two different types of classifiers and the results are reported. In the first case, the classification is done using a single ANN with a sigmoid transfer function. In the second case, which is the proposed method, the classification is done with a module that consists of two ANNs in cascade, the first one with a linear transfer function and the second one with a sigmoid transfer function.

The performance of the classifier is evaluated against the three standard parameters of sensitivity, specificity and classification accuracy. Sensitivity is the percentage of correct detection of a category of diseased signals from among all the diseased signals belonging to that category. This parameter is specific to a disease condition. Specificity is the percentage of correct detection of healthy signals from among all healthy signals. Classification accuracy is the percentage of correct detection from among all the available signals. The sensitivity, specificity and classification accuracy values are measured for ECG and EMG signals for both classification methods. The results of ECG signals are tabulated in TABLE V and TABLE VI, and the results of EMG signals are tabulated in TABLE VII.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

TABLE V. SPECIFICITY, SENSITIVITY & CA IN THE ECG SIGNAL FIRST LEVEL CLASSIFICATION

Measurement	Disorder category (# of signals)	ANN	Cascade of two ANN
Sensitivity (%)	Atrial Arrhythmias (262)	95	98.9
	Ventricular Arrhythmias (416)	94.7	98.8
	MI (128)	97.7	100
Specificity (%)	NSR (126)	94.4	99.1
Classification accuracy (%)		95.2	99

TABLE VI. SPECIFICITY, SENSITIVITY & CA IN THE ECG SIGNAL SECOND LEVEL CLASSIFICATION

Measu-rement	Disorder category	Disorder (No. of signals)	ANN	Cascade of two ANN
Sensiti-vity (%)	Atrial Arrhythmias	AF (106)	94.3	99.1
		SVA (156)	93.6	98.7
	Ventricular Arrhythmias	PVC (144)	93.1	98.6
		VT (140)	93.6	98.6
		MVE (132)	94.7	98.5
	MI (128)		97.7	100
Specifi-city (%)	NSR (126)		94.4	99.1
Classification accuracy (%)			94.4	98.9

TABLE VII. SPECIFICITY, SENSITIVITY & CA IN THE EMG SIGNAL FIRST LEVEL CLASSIFICATION

Measurement	Disorder (# of signals)	ANN	Cascade of two ANN
Sensitivity (%)	Myopathy (289)	95.8	99
	ALS (178)	96.6	99.4
	Parkinson's (150)	94	97.3
	Huntington's (160)	94.4	97.5
Specificity (%)	Normal (300)	95.3	99
Classification accuracy (%)		95.4	98.6

Sensitivity, specificity and classification accuracy are higher when the classification is done with a module of two ANNs in cascade. The classification accuracy is improved by 3.8% for the ECG signals and it is improved by 3.2% for EMG signals. Hence it is inferred that the proposed modular classification with the classification module comprising of the cascade of two ANNs provide improved classification accuracy. The other major contributor to this improved classification accuracy is the multilevel classification technique. The achieved classification accuracy is 2.1% higher than the 96.8% that is reported for the detection of five cardiac health disorders [5]. The classification accuracy achieved in this work is on par with the highest accuracy levels reported for the detection of up to three health disorders. This scheme for classification of Biomedical signals can be extended to other signals like EEG and EOG. The sensitivity of Parkinson's and Huntington's diseases are lower than that of other disorders. Further studies can be carried out on improving the sensitivity of these health disorders to improve the overall classification accuracy of EMG signals.

IV. CONCLUSIONS

This paper discusses the extraction of temporal and spectral/ statistical/ wavelet features from the two Biomedical signals, ECG and EMG, and their classification to detect health disorders. It is observed that for an ECG signal segment of length 10s, the extracted feature values are within a variance of 4%. Corresponding length for EMG signal segment is 11s. The values of the extracted features are observed to be near its true value, if the features are extracted from signal segments from around the mid portion of the whole signal. A feature value is taken as the average of the values extracted from 20 consecutive segments from the mid portion of the Biomedical signal. This technique ensures that the feature value is close to its true value and guarantees an optimum performance.

The sampling frequency of the Biomedical signal is observed to have an influence on the classification accuracy. The optimum sampling frequency found out for this work is 128 Hz for ECG signal and 750 Hz for EMG signal. FFT and DWT are used for feature extraction. A modular approach to classification is proposed based on a cascade of two ANN structures. It is observed that Pattern recognition network provides improved classification accuracy compared to Learning vector quantization network, for the Biomedical signal classification. The multi level classification approach enables classification of health disorders successively to finer levels. It is observed that, with this technique the classification accuracy achieved for large number of health disorders is on par with the classification accuracy achieved for small number of health disorders. This scheme can be extended to other Biomedical signals and also to more number of health disorders. The multilevel classification technique is not tried with EMG signals due to the non-availability of signals belonging to different EMG related health disorders. The proposed technique can be extended to the automated analysis of real time Biomedical signals. This calls for appropriate pre-processing of signals and is an area for future work.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Many of the features extracted from signals belonging to one category of health disorder have their values overlapping with the values of features extracted from signals belonging to other category of health disorders. This limits our ability to minimize the number of features to be used for classification. Also, the classification accuracy will not be the best, with large number of features. Devising a method to find out the optimum feature set for a Biomedical signal classification is an area for future work.

REFERENCES

1. Martis Roshan Joy, U. Rajendra Acharya, Lim Choo Min, "ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform," Biomedical Signal Processing and Control, vol 8, issue 5, pp. 437-448, 2013.
2. Naik Ganesh S. Selvan, and Hung Nguyen. "Single-Channel EMG Classification With EnsembleEmpirical-Mode-Decomposition-Based ICA for Diagnosing Neuromuscular Disorders.",pp.1-11,2015.
3. Rahime Ceylan, Yuksel ozbay, "Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network," Expert Systems with Applications,vol 33, issue 2, pp. 286-295, 2007.
4. Swati Banerjee, Madhuchanda Mitra. "A classification approach for Myocardial infarction using voltage features extracted from four standard ECG leads" IEEE International conference on recent trends in information systems, 2011, pp.325-330.
5. Sambhu D, Umesh A. C, "Automatic Classification of ECG Signals with Feature Extraction using Wavelet Transform and Support Vector Machine," IJAREEIE, vol.2, special issue 1, pp. 235-241, 2013.
6. Felipe Alonso Atienza, Eduardo Morado, Lornea Fernandez Martinez, Arcadi Garcia Alberola, Jose Luis Rojo Alvarez, "Detection of life-threatening arrhythmias using feature selection and support vector machines." Biomedical Engineering, IEEE Transactions vol. 61, no.3, pp.832-840, 2014.
7. Roger Dzwonczyk, Charles G. Brown, H. A. Werman, "The median frequency of the ECG during ventricular fibrillation: its use in an algorithm for estimating the duration of cardiac arrest," IEEE Transactions on Biomedical Engineering, vol. 37, no. 6, pp. 640-646, 1990.
8. Subasi Abdulhamit. "Classification of EMG signals using combined features and soft computing techniques." Applied soft computing, Vol.12, Issue 08, pp. 2188-2198, 2012.
9. Phinyomark Angkoon, Pornchai Phukpattaranont, Chusak Limsakul. "Feature reduction and selection for EMG signal classification." Expert Systems with Applications Vol.39, Issue 08, pp. 7420-7431, 2012.