

Application of Empirical Mode Decomposition And Support Vector Machine Classifiers for Prediction of Sudden Cardiac Arrest in Susceptible Cases

Mr. Arun Raj P.D.

M.Tech Student,

Department of Electronics and Instrumentation,
R V College of Engineering,
Bengaluru, India

Prof. S.V. Venkatesh

Associate Professor,

Department of Electronics and Instrumentation,
R V College of Engineering,
Bengaluru, India

Abstract— One of the primary causes of Sudden Cardiac Arrest (SCA) is Ventricular Fibrillation (V.F.). It is a serious arrhythmic condition wherein the heart beats chaotically rather than pumping blood. It is essential for the subject to get instant medical attention when ventricular fibrillation occurs. In this paper, a novel approach is presented where the Hilbert Huang Transform is used for prediction of such an event. The ECG signals are pre-processed for removal of noises, artifacts. The empirical mode decomposition (EMD), which is a part of Hilbert Huang Transform is applied to the de-noised ECG signal. Three features were extracted from the first IMF (Intrinsic Mode Function) obtained after EMD, viz; spectral entropy, Lyapunov exponent and weighted frequency. Support Vector Machine (SVM) classifier was used to classify and predict the normal people and patients prone to sudden cardiac arrest, based on the extracted features from both the cases. By applying this technique, the results show that average classification accuracy of 95% was achieved in predicting an unknown signal, labelled as 'Normal' or 'Abnormal ECG signal'. A second SVM classifier was used for predicting the occurrence of V.F. in ECG signals prone to cardiac arrest within 15 minutes, which resulted in an accuracy of 55.5 %.

Keywords— Ventricular Arrhythmia; Empirical Mode Decomposition; Support vector machine;

I. INTRODUCTION

Dramatic increase in knowledge and exponential increase in technology has led to lot of sophisticated diagnostic tests. The basic idea of diagnostic tests is to increase the suspicion that a patient has a particular disease to the extent that medical crisis management decision could be made. Electrocardiogram (ECG) is one such diagnostic test. The ECG is painless, non-invasive, easily accessible and widely accepted method for heart condition diagnosis. By viewing the ECG recording, physicians are able to suspect the presence of any heart disease and anticipate the occurrence of any cardiac related threats or sudden death. Subjects experience sudden cardiac death (SCD) within an hour of onset of, acute symptoms mostly due to cardiac causes. [1]. Ventricular fibrillation is one of the major causes for SCA, leading to death. More than 90% of the subjects die because of sudden cardiac arrest as they don't get defibrillation, within 4 to 6 minutes, eventually leading to death of the brain and sudden death of the person. [2]. Therefore a

technique that predicts their event even a few minutes ahead, may possibly save lives.

Many cardiac related events for detection and prediction methods have evolved and recorded in the literature. Shen et al. [3] had devised a ECG acquisition system to acquire the ECG recordings of 20 normal people. Along with this, MIT-SCD database of SCA affected patients was used as abnormal data. They applied Wavelet analysis to normal as well as abnormal ECG to detect SCD events which resulted in an accuracy of 87.5 %. They used artificial neural networks like least mean squares technique to predict SCD events with 67.44% accuracy. Other neural networks like decision based and back propagation neural network were also used which produced an accuracy of 58.14% and 55.81 % respectively.

Elias Ebrahimzadeh et al. [4] had extracted features from HRV of ECG signals, viz., linear, time-frequency, nonlinear features to predict the risk of SCD using K-nearest neighbor and multi-layer perceptron neural network techniques. Their results show that when time-frequency features are combined with nonlinear features, better accuracy is obtained in detecting the SCD events. Yongquin Li et.al [5] have used wavelet transformation to predict cardiac arrest with accuracy of 78%.

Sheela et.al [6] have used HRV signals and support vector machines for predicting SCA by classifying subjects as normal or abnormal. They obtained an accuracy of 88%. Mitra, A.K. et al [7] used a time-frequency analysis approach to predict prenatal abnormalities, from the spectrograms of normal and abnormal fPCG, obtained using Short Time Fourier Transform (STFT). STFT assumes that the signal being considered is piecewise stationary and is also limited by window length with regard to the resolutions of time and frequency.

Karimi Moridani, M et al. [8] extracted non-linear features from return map by plotting it from HRV signals, and showcased the possibility of predicting the mortality of cardiovascular patients admitted in ICU. In this paper, a method to predict the onset of ventricular fibrillation, leading to sudden cardiac arrest, by applying empirical mode decomposition directly on ECG signals is presented.

Almost all the natural, physiological signals are chaotic and are characterized by non-linear and non-stationary features. ECG is one such bioelectric signal having both of these

mentioned features. Researchers, [9-12] have studied the dynamics of the human heart, ECG signals, along with Lyapunov exponents and shown that heart is a chaotic system. Therefore the ECG signals generated from this system must also be chaotic in nature which has been tested using the Lyapunov exponent values. Hilbert-Huang transform is one such novel technique for analysing non-linear as well as, non-stationary signals, developed by Dr. Huang et al. [13, 14].

The upcoming section of this paper discusses about the Hilbert Huang transform technique of signal synthesis. Section III portrays about the data used for analysis, the methodology and the results obtained. Finally Section IV summarizes and gives the conclusion.

II. HILBERT-HUANG TRANSFORM

Dr.Huang et.al [13] had developed a mathematical technique to analyse signals that are nonlinear as well as nonstationary. This technique is called as Hilbert Huang Transform. It generates a time-frequency plot of the signal under analysis, by applying the Empirical Mode decomposition (EMD) and Hilbert Spectral analysis. When EMD is applied to the input signal, a set of signals called as IMFs are obtained. The resultant number of IMFs depends on the input signal. Once these IMFs are obtained, time-frequency plot is constructed by applying Hilbert Transform on every obtained IMF. The details of the two steps are elaborated in the upcoming subsections.

A. Empirical Mode Decomposition (EMD)

In the process of obtaining a time-frequency plot via HHT, EMD procedure is first applied to the input signal that breaks down the input signal into a set of signals known as Intrinsic Mode Functions (IMFs).The resultant decomposed signal, obtained via EMD can be said to be an IMF if they satisfy certain criteria viz, [13]

1. The amount of zero crossing and extrema's may be the same or may differ at most by one.
2. The mean should be zero for the upper and lower envelope.

The process of decomposing the input signal into a set of IMFs is an empirical procedure, commonly called as sifting process.

The procedure is listed below along with flow chart in figure 1.

1. Positions of the local maxima's and local minima's in the given input signal $X(t)$, are identified.
2. Cubic spline interpolation method is used to produce an envelope by interpolating the local maxima's to produce an upper envelope. Similarly the lower envelope is generated. From these two envelopes, a mean envelope is also calculated.
3. The mean envelope is subtracted from the input signal, to obtain an intermediate signal say $h_1(t)$ and this intermediate signal may sometimes satisfy the IMF criteria. However, the iteration is still continued since the mean enveloped may not be the actual mean, when the input signal is noisy with inherent overshoots and undershoots. Therefore, $h_1(t)$ is considered as input signal and the iteration is continued k number of times, until IMF criteria is satisfied.
4. During the iteration process, the K th iteration may satisfy the IMF criteria, and this is assigned as first IMF, $C_1(t) = h_{1k}(t)$.

5. The residue is calculated from the first IMF, by subtracting the first IMF from input signal, i.e. residue, $r(t) = X(t) - C_1(t)$.
6. Now the residue becomes the new input signal and the sifting process is continued, to get next set of IMFs.
7. This procedure is continued until the stoppage criterion is met or it is stopped when no further IMFs can be extracted from the residue.

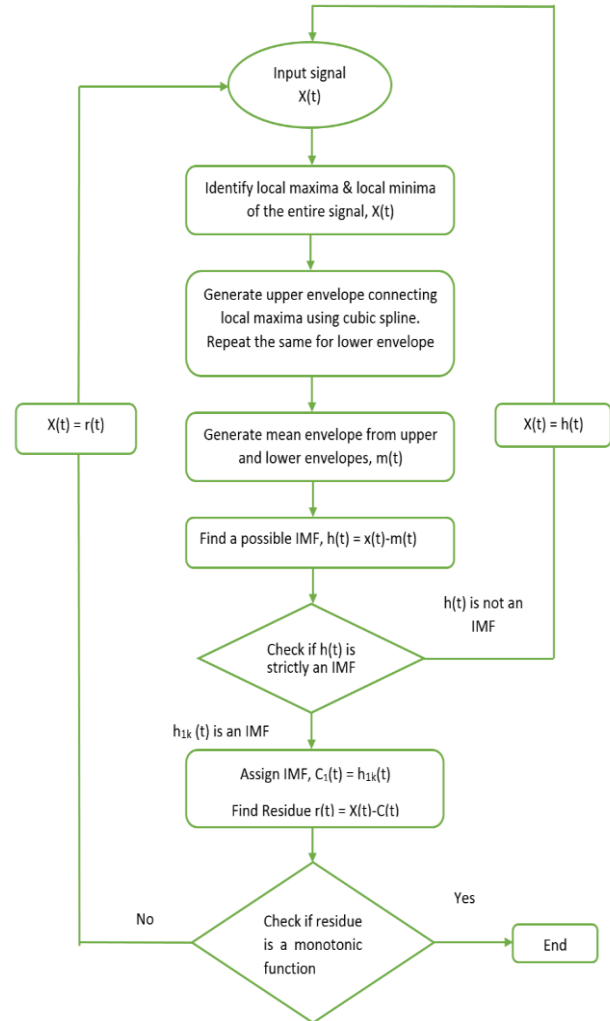


Figure 1 Flow chart of EMD sifting process

Since the EMD process, is a decomposition of the input signal, it can always be reconstructed back. Mathematically it can be written as shown in the below equation.

$$X(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (1)$$

$C_1(t), C_2(t) \dots C_n(t)$ are the IMF's and $r_n(t)$ is the residual component.

B. Hilbert Spectral Analysis

Once the EMD procedure decomposes the input signal in to a set of IMFs which are named as, $C_1(t), C_2(t) \dots C_n(t)$ and residue, Hilbert Transform is used and applied on every IMF. Mathematically this can be written as,

$$Z_i(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{C_i(\tau) d\tau}{t - \tau} \quad (2)$$

With the help of every IMF and the result of application of Hilbert Transform, an analytical signal can be formed as shown below:

$$V_i(t) = [C_i(t) + Z_i(t)] = a(t) e^{-i\theta t} \quad (3)$$

The property of Hilbert Transform helps in producing an orthogonal signal from which an analytical signal is formed. The corresponding amplitude and its phase can be written as,

$$a(t) = \sqrt{C_i(t)^2 + Z_i(t)^2};$$

$$\theta_i(t) = \tan^{-1} \frac{Z_i(t)}{C_i(t)}$$

And the Instantaneous frequency,

$$\omega_i(t) = \frac{d\theta_i(t)}{dt} \quad (4)$$

Once the instantaneous frequency is obtained, with the help of time and amplitude of the input signal, the time-frequency graph is plotted and the dynamics of the non-stationary and nonlinear signal are observed.

III. METHODOLOGY

The overall flow followed in this paper for the prior prediction of onset of the ventricular fibrillation leading to sudden cardiac arrest is shown in the figure 2 and figure 3. The analysis and results are obtained using MATLAB.

A. Data

A set of normal ECG signals are taken from MIT-BIH Normal Sinus rhythm database [15]. This database has the ECG recordings of 18 patients who did not have any arrhythmic events. This database had a total of 18 normal ECG signals, out of which 9 were taken as training data and the rest 9 as test data.

Similarly the ECG signals of the patients prone to sudden cardiac death are taken from MIT-Sudden cardiac Death Holter Monitor database [16] which totally has 23 Holter monitored recordings. Most of the patients in this database had an actual cardiac arrest with all the patients experiencing sustained ventricular tachyarrhythmia [16]. There were 19 patients with a Ventricular fibrillation (VF), and their VF onset time were available.

Therefore the first 18 ECG signals were used as the data for training and prediction of sudden cardiac arrest. Among the 18 SCA prone ECG signals, 9 were taken as training data, and remaining 9 as test data. Therefore, 36 ECG signals (18 from normal case and 18 from abnormal case) were used for analysis via HHT and SVM classifier.

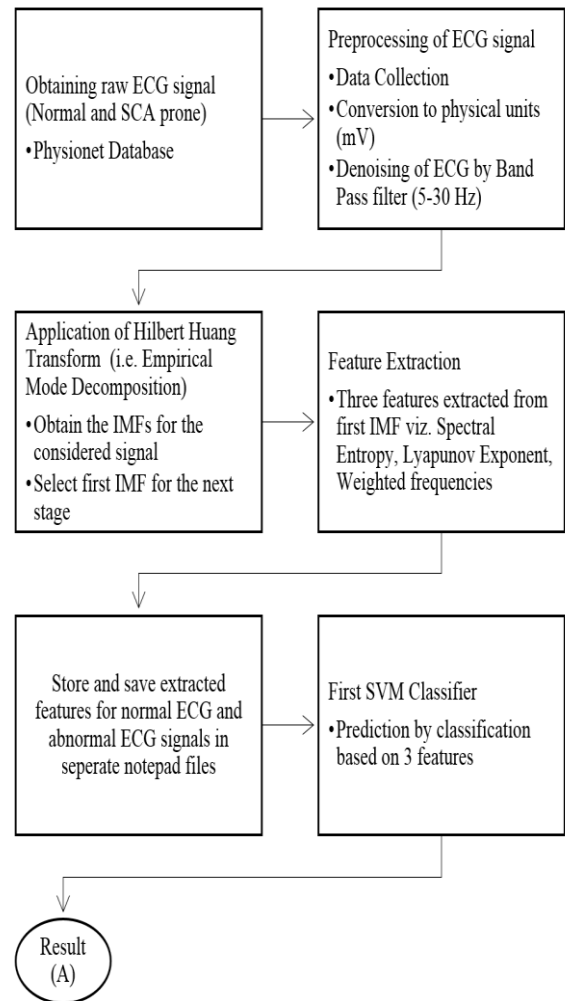


Figure 2 Methodology for prediction of sudden cardiac arrest

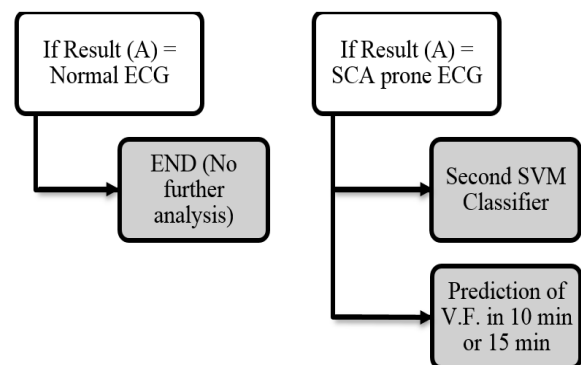


Figure 3 Methodology for prediction by SVM classification

B. Pre-Processing of the signal

The MIT-SCD database has ECG recordings of the subjects who experienced SCA. Figure 4 indicates the occurrence of sudden cardiac arrest in the mid portion of the 10 second ECG interval. As the PHYSIONET database has the onset time of Ventricular fibrillation of the subjects, five such sets of one minute ECG intervals are collected, prior to 10 minutes, before the onset of ventricular fibrillation (i.e. the first set being, the eleventh minute, second set being the twelfth minute and so on up to fifteenth minute with every set duration being only one minute).

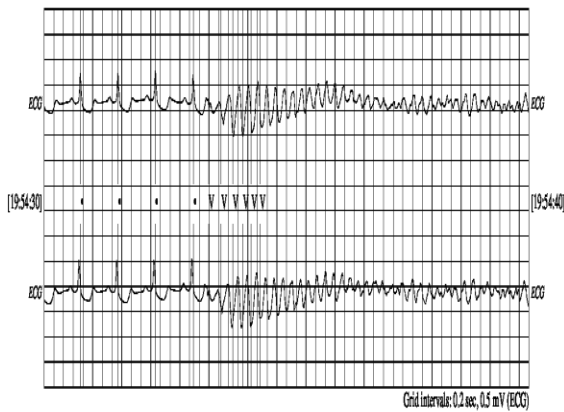


Figure 4: ECG of patient at the occurrence of cardiac arrest

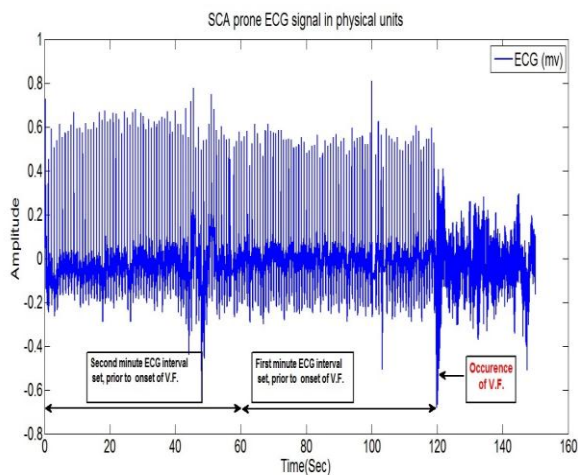


Figure 5: First minute and second minute ECG interval prior to onset of V.F.

Figure 5 shows an example of first and second minute ECG, prior to the occurrence of V.F. In similar fashion, the five sets of ECG mentioned earlier (eleventh minute to fifteenth minute), is collected for pre-processing them.

In case of normal ECG signals, five sets of random, one minute ECG signals are collected. These raw ECG signals (normal and abnormal) acquired from their respective databases are not in physical units (mV) and hence they are converted into the physical units by subtracting base and diving by gain value. These values are given in the respective PHYSIONET databases.

Once they are converted to physical units, it is de-noised using a bandpass filter (5-30 Hz) to remove any high frequency motion artifacts, base line wandering, noise due to power line interference. The de-noised ECG signal of an abnormal subject is shown in figure 6.

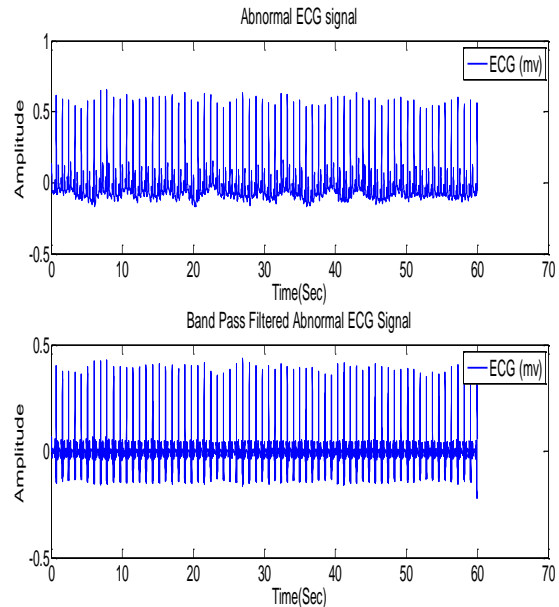


Figure 6 De-noising of ECG signal

C. IMFS Extraction From ECG By Empirical Mode Decomposition

The de-noised ECG signal of both the normal and abnormal subjects are broken down by using the EMD into a set of IMFs. The lower IMFs extract the fast changing high frequency content of the signal and higher IMFs extract the slow changing low frequency content of the signal. This is evident from the IMFs obtained from the normal ECG sample as shown in figure 7. Empirical mode decomposition was applied to a normal ECG signal and 11 IMFs were obtained, but only the first 8 are shown in figure 7. Since the first IMF captures the most vivid properties in the signal, it is used for further analysis.

D. Feature Extraction and Classification

For a good classification accuracy, selection of proper features from the processed ECG signal plays a vital role. Elias Ebrahimzadeh et.al [4] had shown that better accuracy was obtained by using both the time-frequency and nonlinear features. Hence in this paper, two non-linear features are chosen, namely 'Spectral Entropy' [17] and 'Lyapunov Exponent' [18] to extract the most distinctive properties of the signal.

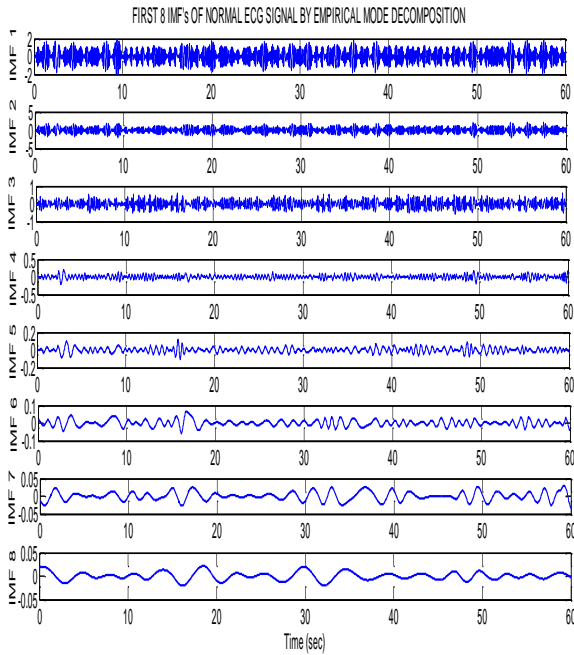


Figure 7: First 8 IMF's of Normal ECG Signal Obtained By Empirical Mode Decomposition

A positive Lyapunov exponent points that the system could be chaotic [8]. Assuming that, if we consider, two close points in phase plane, at time $t=0$, and at time, t ; and the distances between the two points in i^{th} direction, being $\|\delta x_i(0)\|$ & $\|\delta x_i(t)\|$ respectively [8].

Then, Lyapunov exponent ε , [8], is the rate of growth, of the initial considered points.

$$\varepsilon = \lim_{t \rightarrow \infty} \frac{1}{t} \log_2 \frac{\|\delta x_i(0)\|}{\|\delta x_i(t)\|} \quad (4)$$

Suppose, we have a finite discrete random variable X , with their probability mass function as $P(X_i)$, then, spectral entropy [18] equation is represented as

$$SpecEn = - \sum_{i=1}^n P(x_i) \log P(x_i) \quad (5)$$

R.J.Owens et.al [19] had used a feature namely 'weighted frequency' to classify seizures in EEG signals with 94% accuracy. The weighted frequencies [19] is defined as,

$$f = \frac{\sum_{i=1}^k a(i) f^2(i)}{\sum_{i=1}^k a(i) f(i)} \quad (6)$$

In this paper, the effectiveness of predicting SCA, based on these features, that are extracted from the first IMF of both the normal ECG and the Sudden cardiac arrest prone ECG are studied. Once these three features are extracted, SVM classifier is used to train and classify the test data.

IV. RESULTS AND DISCUSSION

The ECG signals for normal and cardiac arrest prone cases are taken from the PHYGONET database. One minute intervals of ECG were extracted from these signals and three features (Lyapunov exponent, spectral entropy, weighted frequency) were extracted from training ECG data for five sets. The first three sets are listed in Figure 8, Figure 9 and Figure 10. First SVM classifier was used to train with help of these data. Similarly, the above three features were also extracted from the test data of both the normal and SCA prone ECG. The first SVM classifier was used to classify the unknown ECG signals taken from test data, in order to predict which ECG signal are prone to SCA. The classification accuracy, sensitivity & specificity of the first SVM classifier is shown in the Table 1.

Once the first stage classification is over, depending on the output of the first SVM classifier, the second stage of SVM classifier runs. If the first stage SVM classifier is abnormal, then the second SVM classifier classifies and predicts the approximate time of occurrence of V.F., based on the training data set. The second SVM classifier is trained in prior with the help of training features collected from both the normal and abnormal ECG signals. For the case of second SVM classifier, it is trained only for two sets viz., features extracted from ECG signal taken 10 minutes prior to V.F. and 15 minutes prior to V.F. The results of the second SVM classifier, for the two cases of test data, namely 10 minutes prior to V.F. and 15 minutes prior to V.F. is shown in Table 2.

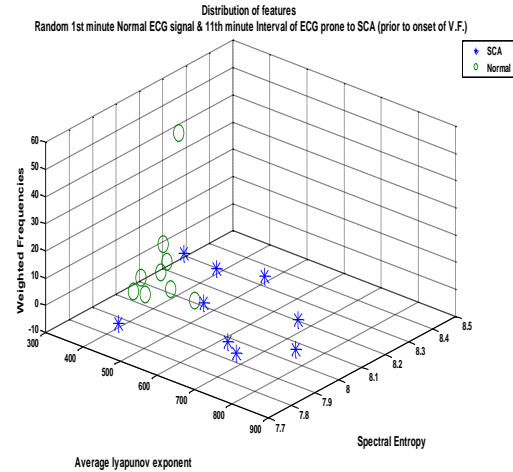


Figure 8 Feature distribution of normal and SCA prone ECG (First Set)

Also a comparison of the method used in this paper with the other techniques followed in the literature using the same MIT/BIH SCD database, is done in Table 3.

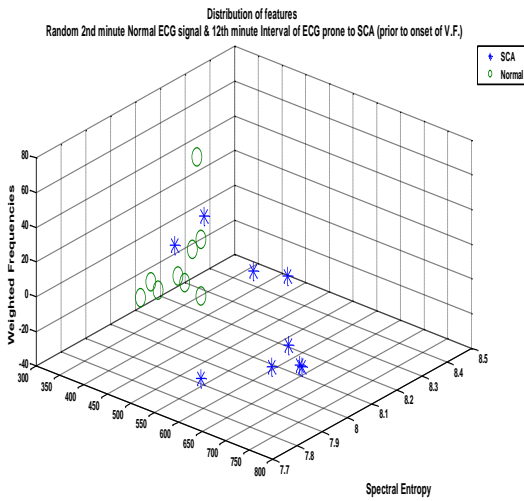


Figure 9 Feature distribution of normal and SCA prone ECG (Second Set)

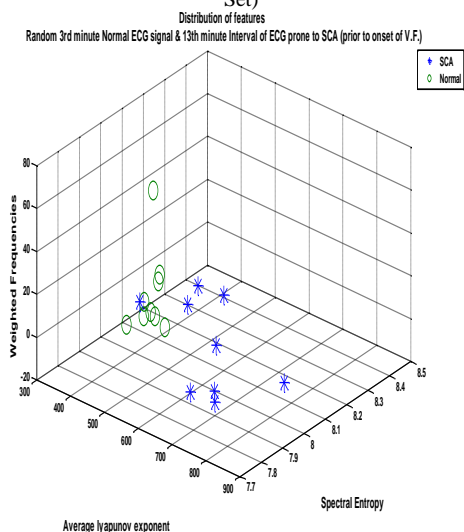


Figure 10 Feature distribution of normal and SCA prone ECG (Third Set)

Shen et al. [3] had used the same MIT-SCD database from PHYSIONET to predict the SCD events using artificial neural networks. The classification accuracy obtained were 67.44 % using least mean square (LMS) algorithm. Another researcher Ebrahimzadeh et al [4] had used k-nearest neighbours and Multilayer perceptron (MLP) neural network, for predicting the occurrence of V.F. Results obtained were 83.93% using MLP before 4 minutes of occurrence of V.F. In both the above mentioned papers, the occurrence of V.F. of SCD prone ECG were known and the same were being fed to the neural networks. In this paper, using Empirical Mode decomposition, it was possible to use ECG signals directly and with the help of the decomposed first IMF, prediction of SCD events by classifying the ECG signals as Normal or Abnormal was achieved. The ECG signals prone to SCD collected were 15 minutes or 10 minutes before occurrence of V.F. The classification accuracy of 95% was obtained using first SVM classifier. Once the ECG signal was classified as Abnormal, another SVM classifier was used in an attempt to predict the exact occurrence of V.F. using a second SVM classifier (i.e. within 10 minutes or 15 minutes to V.F.).

Data Set	Classification Accuracy	Sensitivity	Specificity
11 minutes before SCA	100%	100%	100%
12 minutes before SCA	88.88%	88.80%	88.80%
13 minutes before SCA	94.44%	100%	90%
14 minutes before SCA	100%	100%	100%
15 minutes before SCA	94.44%	100%	90%

Data Set	Classification Accuracy
10 minutes before SCA	33.30%
15 minutes before SCA	55.50%

Method	Technique followed	Accuracy obtained	Time of Prediction (Minutes before occurrence of SCA)
Shen et. al.	LMS algorithm based Neural network	67.44%	Two minutes before SCA
Ebrahimzadeh et al.	MLP and Combinational features space	83.96%	Four minutes before SCA
Proposed Method	EMD and SVM classifier	94.4%	15 minutes before SCA

V. CONCLUSION

This paper proposed the use of Hilbert-Huang transform for prediction of sudden cardiac arrest using the ECG signals. The IMF's of normal and cardiac arrest prone ECG signals were extracted by empirical mode decomposition. The first IMF was selected and 3 novel features were extracted from it. Based on these extracted features, first SVM classifier was able to classify the given unknown ECG signal as with 95 % average accuracy rate. The second SVM classifier is capable of predicting the approximately the occurrence of SCA for two cases viz., 10 minutes or 15 minutes prior to onset of V.F. Hence these results are encouraging in preventive treatment of such cardiac arrest cases, and could be used for real time testing of a person's health condition and potentially save many lives.

ACKNOWLEDGMENT

The authors would wish to thank and acknowledge Gabriel et.al [20] [21] for providing the MATLAB codes for EMD by making it open to the research community.

REFERENCES

- [1] Myerburg, Robert J. "Cardiac arrest and sudden cardiac death." Heart disease: a textbook of cardiovascular medicine (2005).
- [2] Honnekeri, B.S., Lokhandwala, D., Panicker, G.K. and Lokhandwala, Y., 2014. Sudden cardiac death in India: a growing concern. *J Assoc Physicians India*, 62(12), pp.36-40.
- [3] Shen, T.W., Shen, H.P., Lin, C.H. and Ou, Y.L., 2007, August. Detection and prediction of sudden cardiac death (SCD) for personal healthcare. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 2575-2578). IEEE.
- [4] Ebrahimzadeh, E., Pooyan, M. and Bijar, A., 2014. A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals. *PLoS one*, 9(2), p.e81896.
- [5] Li, Y., Bisera, J., Weil, M.H. and Tang, W., 2012. An algorithm used for ventricular fibrillation detection without interrupting chest compression. *IEEE Transactions on Biomedical Engineering*, 59(1), pp.78-86.
- [6] Sheela, C.J. and Vanitha, L., 2014, March. Prediction of Sudden Cardiac Death using support vector machine. In *Circuit, Power and Computing Technologies (ICCPCT), 2014 International Conference on* (pp. 377-381). IEEE.
- [7] Mitra, A.K. and Choudhari, N.K., 2009. Time-frequency analysis of foetal heart sound signal for the prediction of prenatal anomalies. *Journal of medical engineering & technology*, 33(4), pp.296-302.
- [8] Karimi Moridani, M., Setarehdan, S.K., Motie Nasrabadi, A. and Hajinasrollah, E., 2016. Non-linear feature extraction from HRV signal for mortality prediction of ICU cardiovascular patient. *Journal of medical engineering & technology*, 40(3), pp.87-98.
- [9] Hundewale, N., 2012. The application of methods of nonlinear dynamics for ECG in Normal Sinus Rhythm. *Int. J. of Computer Science*, 9, pp.458-467.
- [10] Übeyli, E.D., 2009. Detecting variabilities of ECG signals by Lyapunov exponents. *Neural computing and applications*, 18(7), pp.653-662.
- [11] Babloyantz, A. and Destexhe, A., 1988. Is the normal heart a periodic oscillator?. *Biological cybernetics*, 58(3), pp.203-211.
- [12] Casaleggio, A., Braiotta, S. and Corana, A., 1995, September. Study of the Lyapunov exponents of ECG signals from MIT-BIH database. In *Computers in Cardiology 1995* (pp. 697-700). IEEE.
- [13] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C. and Liu, H.H., 1998, March. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 454, No. 1971, pp. 903-995). The Royal Society.
- [14] Huang, N.E., 2006. An adaptive data analysis method for nonlinear and nonstationary time series: the empirical mode decomposition and Hilbert spectral analysis. In *Wavelet Analysis and Applications* (pp. 363-376). Birkhäuser Basel.
- [15] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K. and Stanley, H.E., 2000. Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. *Circulation*, 101(23), pp.e215-e220. <https://physionet.org/physiobank/database/nsrdb/>
- [16] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K. and Stanley, H.E., 2000. Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals. *Circulation*, 101(23), pp.e215-e220. <https://physionet.org/physiobank/database/sddb/>
- [17] Bai, B. and Wang, Y., 2011, May. Ventricular fibrillation detection based on empirical mode decomposition. In *Bioinformatics and Biomedical Engineering, (iCBBE) 2011 5th International Conference on* (pp. 1-4). IEEE.
- [18] <http://www.math.tamu.edu/~Michael.Pilant/math442/Matlab/examples.html>
- [19] Oweis, R.J. and Abdulhay, E.W., 2011. Seizure classification in EEG signals utilizing Hilbert-Huang transform. *Biomedical engineering online*, 10(1), p.1.
- [20] Rilling, G., Flandrin, P. and Goncalves, P., 2003, June. On empirical mode decomposition and its algorithms. In *IEEE-EURASIP workshop on nonlinear signal and image processing* (Vol. 3, pp. 8-11). IEEE.
- [21] <http://perso.ens-lyon.fr/patrick.flandrin/emd.html>
- [22] Owis, M.I., Abou-Zied, A.H., Youssef, A.B. and Kadah, Y.M., 2002. Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification. *IEEE transactions on Biomedical Engineering*, 49(7), pp.733-736.
- [23] Jeong, J., Kim, S.Y. and Han, S.H., 1998. Non-linear dynamical analysis of the EEG in Alzheimer's disease with optimal embedding dimension. *Electroencephalography and clinical Neurophysiology*, 106(3), pp.220-228.