

Detection of Congestive Heart Failure Based on Spectral Features and Extreme Learning Machine

Sulekha Saxena, P. N. Hrisheeksha, Vijay Kumar Gupta, Ram Sewak Singh

Abstract: In this paper we proposed a novel approach to evaluate the classification performance of features derived from various spectral investigation methods for congestive Heart Failure (CHF) analysis using ranking methods, Kernel Principal Component Analysis (KPCA) and binary classifier as 1-norm linear programming extreme learning machine (1-NLPELM). For this study, thirty different features are extracted from heart rate variability (HRV) signal by using spectral methods like multiscale Wavelet packet (MSWP), higher order spectra (HOS) and auto regression (AR) model. Top ten features were extracted by ranking methods and then reduced to only one feature by KPCA having kernel function as radial basis function (RBF) which was further applied to 1-NLPELM binary classifier. For this purpose, the HRV data were taken from standard database of Normal sinus rhythm (NSR), elderly (ELY) and Congestive heart failure (CHF) subjects. Numerical experiments were being done on the combination of database sets as NSR-CHF, NSR-ELY, and ELY-CHF subjects. The numerical results show that features at third level of decomposition of HRV data sets MSWP shows lowest p-value (< 0.001). Thus, third level of MSWP features are better than other features extracted by auto regression (AR) model and higher order spectra (HOS) spectral methods.

Index terms: 1-norm linear programming extreme learning machine (1-NLPELM), higher order spectra (HOS), Kernel Principal Component Analysis (KPCA), ranking methods

I. INTRODUCTION

Cardiovascular disease (CVD) is a noteworthy and regularly developing issue everywhere throughout the world, bringing about almost 33% of all deaths and prompting huge dreariness. It is a noteworthy problem in creating nations which experience an adjustment in way of life that presents novel hazard factors for cardiovascular disease, prompting a significant increase in cardiovascular disease throughout the developing world. The World Health Organization (WHO) assessed that over 75% of early stage CVD is preventable and by diminishing the risk factors, the developing CVD load on both people and healthcare professionals can be decreased [1]. Deaths due to congestive heart failure (CHF) are higher than any other type of cardiac heart disease [2]. CHF is a chronic progressive condition that affects the pumping of heart

Revised Manuscript Received on June 15, 2019.

Sulekha Saxena, Electrical & Electronics Engineering, IMS Engineering College, Ghaziabad, UP, (201009), INDIA.

P. N. Hrisheeksha, Campus Director, Chandigarh Engineering College, Mohali, INDIA.

Vijay Kumar Gupta, Electronics & Communication Engineering, Inderprastha Engineering College, Ghaziabad, INDIA.

Ram Sewak Singh, Electronics & Communication Engineering, IMS Engineering College, Ghaziabad, UP, (201009), INDIA.

muscles which often leads to heart failure [3, 4]. Thus, early detection of CHF is utmost important and for this, various differentiable features (attribute), efficient classifier and Heart Rate Variability (HRV) analysis methods are used.

The Heart rate changeability (HRV) investigation has turned into a well known non-invasive method in cardiovascular disease detection to evaluate the exercises of the autonomic nervous system (ANS). The HRV is controlled by figuring time intervals between progressive R pinnacles focuses on the QRS complex of the electrocardiogram (ECG) which demonstrate a connection between parasympathetic, sympathetic and, henceforward, sympathovagal balance [5-7]. Various methods like time domain, frequency domain and non-linear domain methods has been proposed for HRV analysis showing dissimilarities between ECGs of healthy young (YNG), elderly (ELY) subjects and cardiac patients leading to possible early prognostic or diagnostic information [8-10].

Computer-aided diagnosis is used for disease detection which is an economic method to enhance the detection speed of cardiac disease diagnosis. Various linear and non linear processing methods are used to extract information related to disease detection. [11,12] Nonlinear techniques are more helpful than linear techniques since the heart is a chaotic oscillator under normal cardiovascular activities. [13,14]

Machine learning methods are broadly utilized for the binary and multi-class classification and prediction of cardiac disease. Recently, a new learning method called extreme learning machine (ELM) has been widely utilized in the field of bio-medical signal examination [15, 16]. Despite the fact that the learning speed of ELM can be multiple times faster than traditional feed forward algorithms [15] yet because of arbitrary nature of the hidden layer of ELM, the learning accuracy might be decreased [17-20]. Thus, 1-norm regularization or penalty has been used in [21], as 1-norm is utilized to sort certain coefficients of the ELM model by making them zero and consequently gives sparse ELM models which are just interpretable. A linear programming of 1-norm ELM (1-NLPELM) method is used in this work. The main advantage of 1-NLPELM over ELM is that it leads to reduced number of hidden nodes [22].

In this work, initially thirty spectral features were derived from HRV signal by using spectral investigation methods. Out of thirty features, five spectral features were extracted by auto regression (AR) model, fifteen spectral features extracted by multiscale Wavelet packet



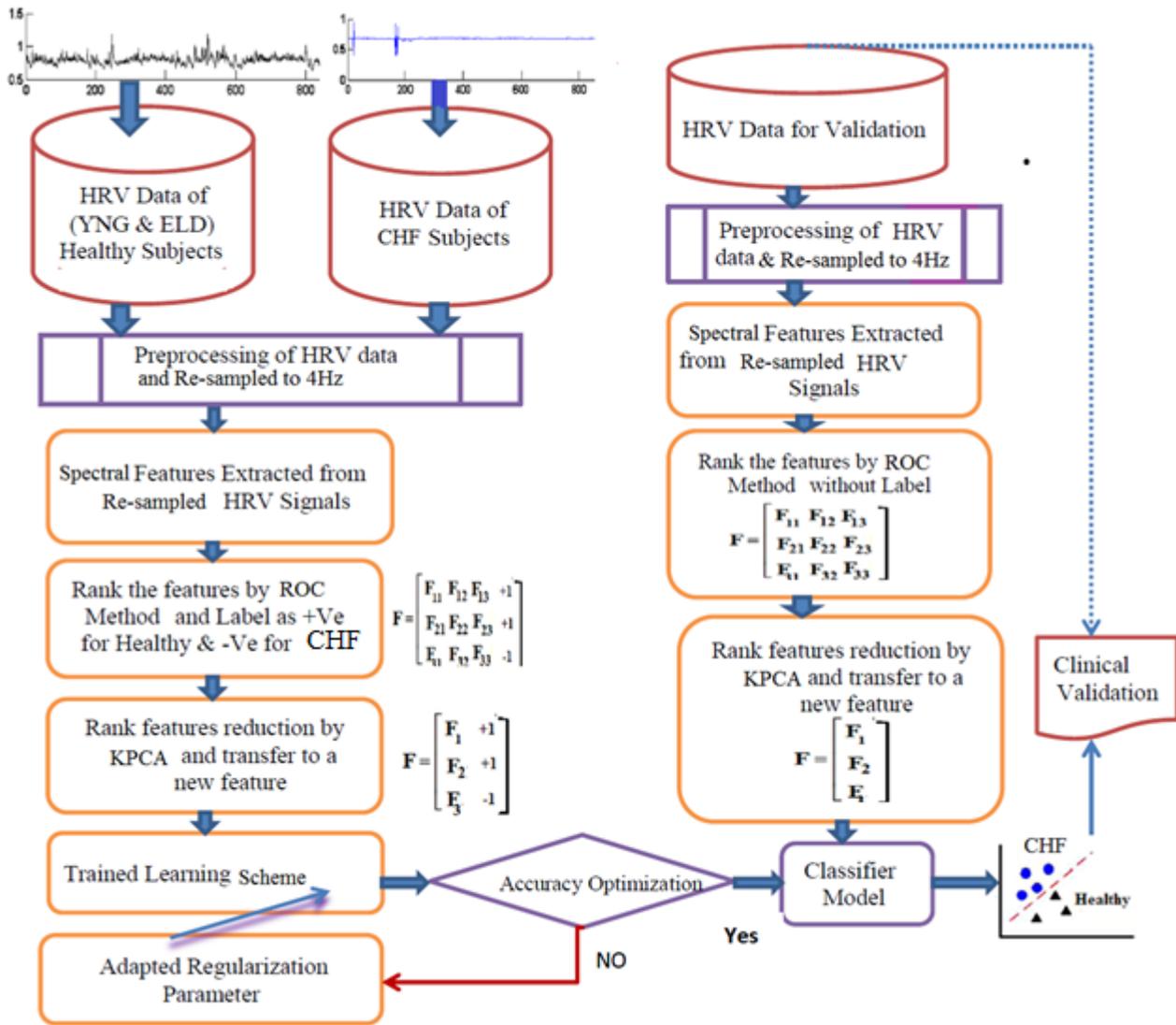


Fig.1 Generalized block diagram of HRV analysis and classification

(MSWP) decomposition and ten spectral features were extracted by higher order spectra (HOS). The flow chart for detection and analysis of CHF subjects is Fig. 1. In detection process, first, the pre-processed HRV data was fed to the feature extraction step. Not all features distinguish well between healthy and cardiac disease subjects. Therefore, feature ranking algorithms are used such as receiver operating characteristics (ROC) method to rank the features based on their clinical significance. These ranked top ten features (highest to lowest) are fed to dimension reduction technique as Kernel principle component analysis (KPCA). The reduced dimension of features are first normalized between -1 to 1 and then fed to 1-NLPELM binary classifier. The results of the study show that spectral methods yields approximately $97.19 \pm 0.61\%$ validation accuracy after using KPCA+1-NLPELM classifier. It has additionally been seen that features at the third level of decomposition of HRV data sets by MSWP have lowest p-value (< 0.001). As, lowest p-value features has highest discrimination ability. Thus, third level of MSWP features are better than considered other features extracted by AR and HOS spectral methods for HRV data set analysis.

The following section explains the HRV signal

extraction and detection method. The database and pre-processing are described in Section II, feature extraction methods are detailed in Section III.A, feature ranking methods are explained in Section III.B, Kernel principle component analysis (KPCA) in section III.C, background of ELM and 1-NLPELM in section III.D, parameters and activation function used in simulation are discussed in section III.E, the results and discussion in section IV and conclusion in Section V.

II. MATERIALS AND PRE-PROCESSING

A. Database

The R-R interval (HRV) data used in this work has been taken from the ECG signal given by the healthy young (YNG) and elderly/older (ELY) from Fantasia database, CHF from St. Petersburg Institute of Cardiological Technics database [23]. The Fantasia database involves 30 YNG (21 - 35 years of age) and 30 ELY (70 - 85 years of age) [24]. The database of St. Petersburg Institute of cardiological technics comprises of 75 illustrated recordings got from 32 Holter records. Each record is 30



minutes in length and comprises of standard ECG leads. Just thirteen subjects (9 men and 4 women, aged 18-80; mean age: 58) were suffered from CHF [23].

B. Pre-Processing

Pre-Processing of R-R interim time arrangement information is vital before examination of HRV signal to diminish error and upgrades the sensitivity of time series signal. Initially the ectopic beat or interim recognition and correction was done before HRV analysis. In this work, the ectopic beats are detected using standard deviation filter method which marks outliers as being interims that lie outside the general mean R-R interim by a user defined value of standard deviation. The user defined value was three times of standard deviation [25, 26]. A cubic spline interpolation method is utilized to supplant ectopic interims located during the detection procedure. Subsequent to supplanting R-R interims, it is known as Normal to Normal interims (NN interims). The NN interims were sampled at 4Hz (just for spectral investigation). For training and validation of proposed method, 3,000 sample of NN interim each subject of database was selected. To increase the training and testing estimate, the 3000 samples of NN interim are separated into segments of 600 NN interims. At last, the segments of NN interims are utilized for the feature extraction by spectral and chaos examination techniques.

III. METHODOLOGY

A. Feature extraction by spectral methods

Autoregressive Model

In Autoregressive (AR) spectral methods stab to model HRV data instead of assessing the power spectral density (PSD) directly. Numerous modeling methods are used for AR spectrum estimation, but the Burg technique is the most popular in HRV spectrum analysis [27]. The PSD of an q^{th} order autoregressive process is given by

$$P(F) = \frac{\epsilon_p}{F_s \left| 1 + \sum_{k=1}^{k=q} A_q(k) e^{-\frac{i2\pi k F}{F_s}} \right|^2} \quad (1)$$

Where F_s is the sample rate, ϵ_p is the total least square error and the A_q Burg AR model parameters [28]. Boardman, et al. [29] had been suggested that a model order of $p = 16$ and points in $PSD = 512$ points are a sound choice for HRV in human, if NN intervals sampled at 4 Hz. Five features are extracted from PSD of AR model [30] and expressed as (a) Power in low frequency (pLF): It reflects the sympathetic and parasympathetic (vagal) tone activity of heart in frequency band 0.04 to 0.15 Hz of HRV. (b) Power in high frequency (pHF): High frequency heart rate oscillations are related with respiratory sinus arrhythmia and reflect parasympathetic activity of heart in frequency band 0.15 to 0.4Hz of HRV. (c) Sympatho-vagal balance (LF/HF ratio): The power spectral analysis is performed on the HRV data to show vagal tone and the sympatho-vagal

balance as pLF/pHF ratio at any point of instant. (d) Normalized power in low frequency (nLF) defined as: $nLF = pLF/pLF + pHF$ (e) Normalized power in high frequency (nHF) defined as: $nHF = pHF/pLF + pHF$.

Multiscale Wavelet Packet Decomposition

The multiscale Wavelet Packet (MSWP) transform simplifies the linkage between wavelets and multiresolution approximations [31]. In which, the HRV data was decomposed by the MSWP using the Haar wavelet whose mother wavelet function is simply a step function. The dilation and translating for the MSWP are built on powers of 2 or dyadic chunks, e.g., $2^0, 2^1, 2^2$ etc. The dilation function is repeatedly represented as a tree of high and low pass filters. The first level of the tree decomposes the HRV original data into detail (high frequency, indicated by bold line in Fig. 2) and approximation (low frequency, indicated by dotted line in Fig. 2) components. Both branches of the tree are split into finer components as $\Omega_{0,0}$ into $\Omega_{1,0}$ and $\Omega_{1,1}$. Fig. 2 shows the tree for MSWP for 3 levels of decomposition. The feature as Log Root Mean Square of the each detail and approximation components of the signal at each scale of analysis is computed. A complete description of the method is explained in [32]. The MSWP generated 2^{level} features at each level of decomposition. Hence total numbers of features are fifteen for 3 level decomposition. Value of features repeated after 3 level of decomposition so, we have not taken 4, 5, ... -n level decomposition.

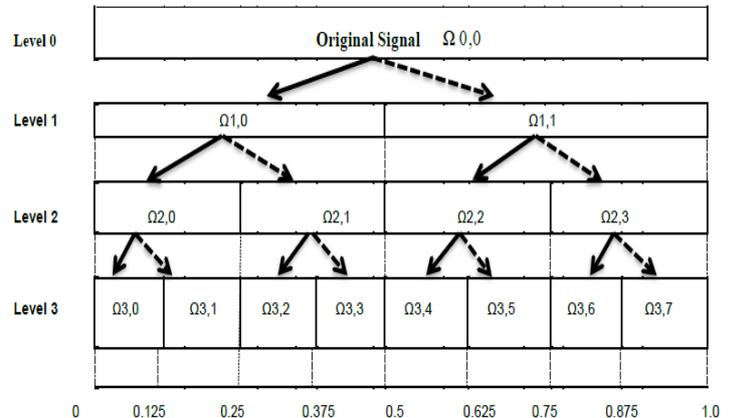


Fig.2 Example of decomposition of MSWP for 3 levels, vertical axis

and horizontal axis shows level of decomposition and frequency variety as a fraction of the Nyquist frequency. The $\Omega_{1,0}, \Omega_{1,1}, \dots, \Omega_{1,1}$ represents detail and approximation components of HRV signals.

Higher order spectral method

The higher order spectral HOS consist of moment and cumulant spectra. A mathematical expression of HOS described in [33, 34]. In this paper features related to the third order statistics of the HRV data are studied, namely bi-spectrum. For deterministic sampled HRV data, the bi-spectrum can be calculated directly from the



discrete Fourier transform (DFT) of N realization of HRV data $X(n)$ and it is defined as $B(F_1, F_2) = E\{X(F_1)X(F_2)X^*(F_1 + F_2)\}$. Where F_1, F_2 are the frequency indices, X^* is the complex conjugate of X and $X(F)$ is the DFT of the discrete signal $X(n)$ and $E\{\cdot\}$ is a mean over a group of realizations of HRV data.

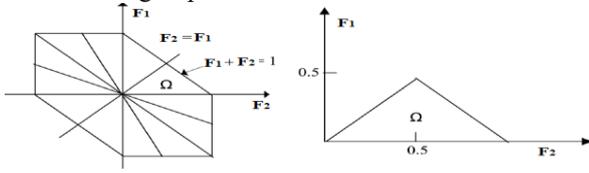


Fig. 3 Non-redundant region of computation of the bi-spectrum for HRV real data. The features are calculated from triangular Ω region. The vertical and horizontal axis represents the frequencies which are normalized by the Nyquist frequency.

The bi-spectrum of HRV real signals has a proportion of symmetries, is shown in **Fig. 3** (left of triangle) and thus redundant information. In which, focus is concentrated only on the positive frequencies and obviously only on the principal domain Ω , shown in **Fig. 3** (triangle region), for this the condition given as

$$\Omega = \{(F_1, F_2) : 0 \leq F_2 \leq F_1 \leq \frac{F_s}{2}, F_2 \leq -2F_1 + F_s\} \quad (2)$$

Here F_s is the sampling frequency.

We have extracted 10 types of features from bi-spectrum method of HRV data defined as

- (a) Mean of magnitude (MM) = $\frac{1}{L} \sum_{\Omega} |B(F_1, F_2)|$.
- (b) Sum of logarithmic amplitude of bi-spectrum (H_1) = $\sum_{F_1, F_2 \in \Omega} \log(|B(F_1, F_2)|)$.
- (c) Sum of logarithmic amplitudes of diagonal elements in the bi-spectrum (H_2) = $\sum_{F_k \in \Omega} \log(|B(F_k, F_k)|)$.
- (d) First-order spectral moment of amplitudes of diagonal elements in the bi-spectrum (H_3) = $\sum_{F_k \in \Omega} k \log(|B(F_k, F_k)|)$.
- (e) Normalized bi-spectral entropy (P_1) = $-\sum_m P_m \log(P_m)$, where $P_m = \frac{|B(F_1, F_2)|}{\sum_{F_1, F_2 \in \Omega} |B(F_1, F_2)|}$.
- (f) Normalized bi-spectral squared entropy (P_2) = $-\sum_n P_n \log(P_n)$, where $P_n = \frac{|B(F_1, F_2)|^2}{\sum_{F_1, F_2 \in \Omega} |B(F_1, F_2)|^2}$.
- (g) Bi-spectral phase entropy (PHE) = $\sum_n p(\phi_n) \log p(\phi_n)$, where $p(\phi_n) = \frac{1}{L} \sum_{\Omega} 1(\phi(B(F_1, F_2)))$ and $\phi_n = \{\phi : -\pi + \frac{2\pi n}{M} \leq \phi \leq -\pi + \frac{2\pi(n+1)}{M}\}$, $n = 0, 1, 2, \dots, M-1$. L is the number of points in the interior of the triangle (non-redundant region), condition given by Eq. (2), ϕ denotes to the phase angle of the bi-spectrum, and $1(\cdot)$ is a sign function which provides a value of 1 when the ϕ is in the range of bin ϕ_n .
- (h) Weighted center of bi-spectrum as WCOB1 = $\frac{\sum_{\Omega} i B(i, j)}{\sum_{\Omega} B(i, j)}$ and WCOB2 = $\frac{\sum_{\Omega} j B(i, j)}{\sum_{\Omega} B(i, j)}$, where i and j denotes the frequency bin indices in the interior of the triangle expressed in Eq. (2) and shown in Fig.3 (triangle region).
- (g) Spectral Flatness: The spectral flatness is ratio of geometric mean of the power spectrum by the arithmetic mean of the power spectrum, i.e. =

$\frac{\text{Geometrical Mean } (\sum_{F_1, F_2 \in \Omega} |B(F_1, F_2)|)}{\text{Airthmetical Mean } (\sum_{F_1, F_2 \in \Omega} |B(F_1, F_2)|)}$. A high spectral flatness tends to 1.0 show that the HRV spectrum has a comparable amount of power in entirely spectral range and the chart of the HRV spectrum would appear to be very flat and smooth. A low spectral flatness tends to 0 indicates that the HRV spectral power is contented in a quite small number of frequency bands. This usually occurs when HRV is a combination of sine waves, and the HRV spectrum would seem spiky [35]. For investigation of these features, blocks of 128 samples, equivalent to 32 seconds at the sampled rate of 4 samples/sec were used. These blocks were taken from each HRV data record with an overlap of 64 points (i.e 50%).

B. Features Ranking Methods

Most weighty features can be chosen using features ranking approach. The ranking method allocates rank to the considered large number of features and arrays them according to their statistical significance. Further, the lower ranked features can be ignored and higher ranked features can be considered for classification [36, 37]. These approaches reduce the intricacy of the features dimension and significantly decreases the time for processing the data without affecting the binary classification enactment. In this work, we have used five methods of features ranking for binary classification as (a) Fisher score method: It is based on filter methods, rank the features as a pre-processing step previous to the learning algorithm, and choose those features having high ranking Fisher score [38]. (b) Wilcoxon method: It ranks based on non-parametric test [39] (c) Entropy method: Entropy method employs the divergence approach (d) Receiver Operating Characteristic (ROC) method: This method ranked the features according to the area under the ROC curve and the random classifier slope, and (e) Bhattacharya space algorithm: Uses the Bhattacharya distance to measure the separation between the two classes [37].

C. Kernel Principal Component Analysis

There is a small variation between patterns of HRV of some subjects like NSR-CHF class in the non-linear feature space. It is challenge to differentiate between the two. In this condition, a feature dimension transformation technique will be extremely valuable [40]. Different techniques have been developed to reduce the features data size for classification [41-44]. In this paper, we have applied KPCA for reducing the features dimension. The Kernel principal component analysis (KPCA) is a dimension reduction technique which is also based on nonlinear kernel function [45], but its discriminating capability is higher as compared to GDA for NSR-CHF group subject.

D. 1-Norm Linear Programming Extreme Learning Machine

Background of extreme learning machine: Before the formulation of the extreme learning machine (ELM) method some significant point should be kept in mind, we assumed that all



the time series data vectors are column vectors. For any two data vectors X, T in the n -dimensional real space R^n the inner product of the data vectors will be denoted by $X^t T$ where X^t is the transpose of the data vector of X , the 1-norm and 2-norm of a data vector X will be indicated by $\|X\|_1$ and $\|X\|_2$.

Assume that $\{X_k, T_k\}_{k=1,2,\dots,m}$ is the data set of training samples, where for the input features $X_k = \{X_{k1}, X_{k2}, \dots, X_{kn}\}^t \in R^n$ are applied to the input nodes and its corresponding target vector value (class label) $T_k \in R$ or $\in \{-1, 1\}$. For one of the learning machine method viz. Single hidden layer feed forward neural network (SLFNs), ELM arbitrarily assigned value of weights vector $A_s = \{A_{s1}, A_{s2}, \dots, A_{sn}\}^t$ and the bias $B_s \in R$ containing the input layer to s^{th} hidden node. The weight vector A_s connecting the input layers to the hidden nodes. The SLFNs with L number of hidden nodes approximate the input features with zero error if there exists analytically determines the output weight vector $W = \{W_1, W_2, \dots, W_L\}^t \in R^L$ connecting between hidden nodes and output nodes. The target value of SLFNs or ELM is formulated as $T_k = \sum_{s=1}^L W_s \phi(A_s, B_s, X_k)$ for $k= 1, 2, \dots, m$. Where $\phi(\cdot) = \phi(A_s, B_s, X_k)$ is activation function, which represents the output of the s^{th} hidden nodes for the input features samples and parameters. The target vector (generated from output nodes) is linearly related with W and ϕ , hence, it can be formulated in the matrix form as

$$HW = T(3)$$

Where

$$H = \begin{bmatrix} \phi(A_1, B_1, X_1) & \dots & \phi(A_L, B_L, X_1) \\ \vdots & \dots & \vdots \\ \phi(A_1, B_1, X_m) & \dots & \phi(A_L, B_L, X_m) \end{bmatrix}_{m \times L} \quad (4)$$

is the output matrix of hidden layer neural network and $T = \{T_1, T_2, \dots, T_m\}^t \in R^m$ is the target vector. To obtain a minimum norm least square solution of W for the linear system of Equation (3), this can be explicitly achieved as in [16]. $W = H^\dagger T$. Where H^\dagger is the inverse of H matrix, this is known as name of Moore-Penrose generalized inverse [46]. Finally, by accomplishing the solution of $W = \{W_1, W_2, \dots, W_L\}^t \in R^L$, a decision function $F(\cdot)$ is determined for any input of features sample $X = \{X_{k1}, X_{k2}, \dots, X_{kn}\}^t \in R^n$. Mathematical expression of $F(\cdot)$ is represented as

$$F(X) = (\phi(A_{11}, B_{11}, X), \dots, \phi(A_{L1}, B_{L1}, X)) W \quad (5)$$

Though, intended for binary classification problem, the decision function is based on signum function, and defined as $F(X) = \text{sign}\{(\phi(A_{11}, B_{11}, X), \dots, \phi(A_{L1}, B_{L1}, X)) W\}$ (6) It is important to note that, one time the values of the weights vector $A_s \in R^n$ and the bias $B_s \in R$ are arbitrarily assigned at the establishment of the learning algorithm. These values remain permanent and so the elements of matrix H remain constant.

1- Norm linear programming extreme learning machine

For sparse model depiction, a 1-norm ELM combined with absolute loss has been developed by [22, 47] for binary classification. But this proposed model has optimization problem means its loss function is not differentiable. So in this paper, a methodology of changing over the underlying issue into another issue with differentiable and requirement capacities is considered. Indeed, encouraged by the literature of 1-norm SVM [48], it is promoted to solve 1-norm ELM (1-NELM) by expressing it as a linear program problem

(LPP). This technique leads to a fast converging iterative and simple method for estimation of binary classification problems. For given learning algorithms of SLFN having L number of hidden nodes in neural network, the ELM learning machine chooses the unknown output weight $W = \{W_1, W_2, \dots, W_L\}^t \in R^L$ vector connecting the hidden nodes to the output node having least minimum training error property and norm [16], that is to say the minimum norm least squares (MNLS) problem is W presented as:

$$\min \|HW - T\|_2 \text{ and } \min \|W\|_2 \quad (7)$$

Consider the MNLS problem (8) expressed in 1-norm defined by

$$\min_{W \in R^L} \|W\|_1 + \gamma \|HW - T\|_1 \quad (8)$$

Where $\gamma > 0$ is constant. Using the technique of [66], the 1-NELM learning problem (7) is expressed into a LPP in this way: $U, V \in R^L$ and $P, Q \in R^m$

$$\text{Assume that } W = U - V \text{ and } HW - T = P - Q \quad (9)$$

Such that $U, V \geq 0$ and $P, Q \geq 0$ hold. Then, using (8) in (9) can be obtained the 1-norm linear programming ELM (1-NLPELM) solution in primitive of the formula:

$$\min_{U, V, P, Q} E_L^t (U + V) + \gamma \min_{U, V, P, Q} E_m^t (P + Q)$$

Condition on $H(U - V) - P + Q = T$, such that $U, V, P, Q \geq 0$, (10)

Where E_L and E_m are the column vectors of ones of dimension L and m respectively. Hence, equation (10) indicates that the linear problem is probable, solvable and additionally its objective function is constrained below by zero.

E. Parameters Used for Simulation of 1-NLPELM

In this section, the performance of 1-NLPELM has been tested on RBF and additive hidden nodes. For this test, the activation function $\phi(A_s, B_s, X)$ was chosen as the sigmoid function and defined as $\phi(A_s, B_s, X) = 1 / (1 + \exp(-(A^t X + B)))$. For additive nodes, multiquadric function described by [49] as $\phi(A_s, B_s, X) = \sqrt{(|X - A||_2^2 + B^2)}$, the function is contemplated for RBF hidden nodes. Proposed for multi-quadric and sigmoid activation functions, the hidden node parameters were chosen arbitrarily with uniform distribution in $[-0.6, 0.6]$. The biases of hidden layer and input weights of the hidden nodes are selected arbitrarily at the beginning of the learning for 1-NLPELM and these values remain constant in every test of simulation. The optimal values of the regularization parameter γ were determined by executing 10-fold cross-validation.

If number of values of L is large, it prompts increment estimation time. It was depicted by [22] that improved generalization performance could be acquired for number of moderate/small values as $L = 100$. The estimations of γ from $\{2^{-4}, \dots, 2^{12}\}$ by 10-fold cross-validation strategy was determined, with these values, the average test accuracy for every dataset was determined by conducting 50 independent trials. To analyze the performance of classifier for datasets, we calculated



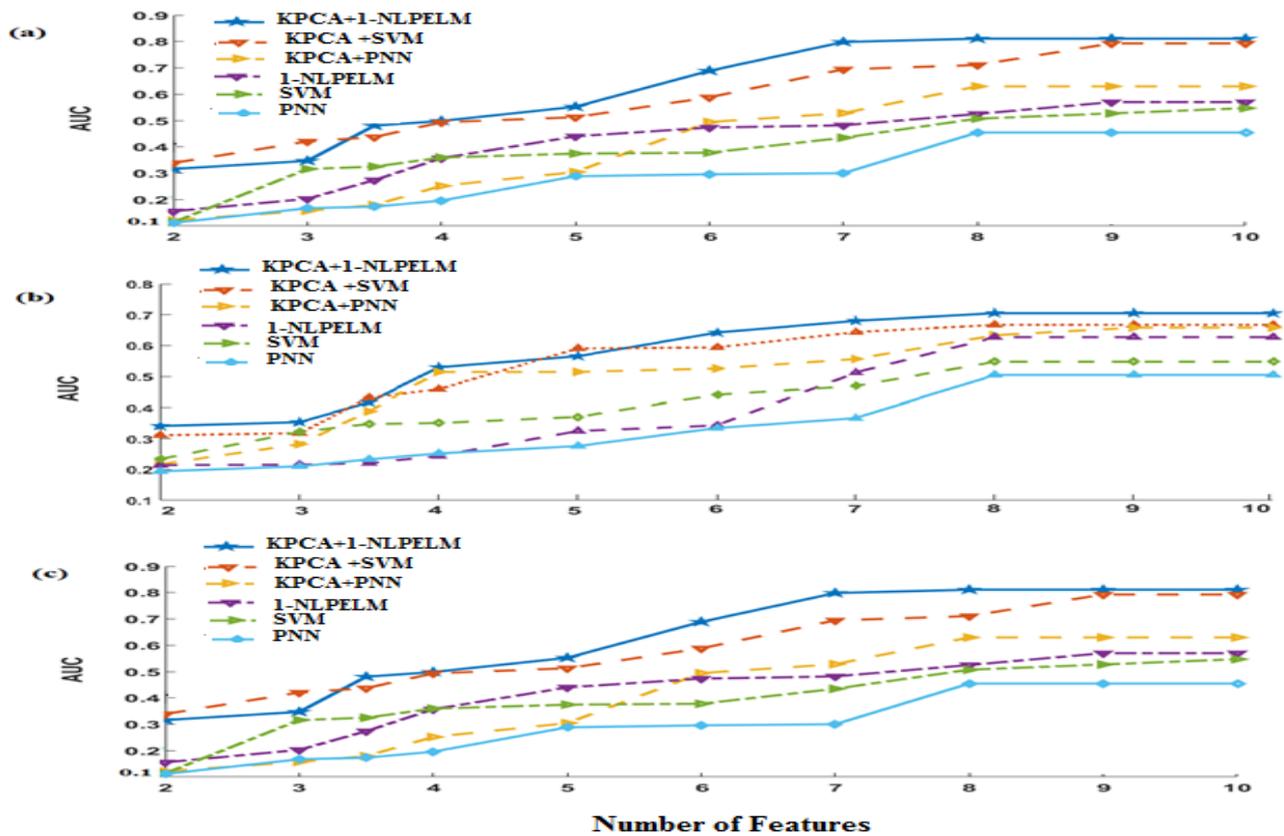


Fig. 4 Presents graph between AUC and top ten features for (a) NSR-CHF (b) NSR-ELY (c) ELY-CHF datasets. The priority of top ten features assigned using Bhattacharya ranking method

performance parameters as Accuracy (A_c) = (True Positive + True Negative) / (True Positive + True Negative + False Negative + False Positive), Sensitivity (Se) or True Positive Rate (TPR) = True Positive / (True Positive + False Negative) and specificity (Sp) or True Negative Rate (TNR) = True Negative / (False Positive + True Negative) and False Positive Rate (FPR) = (1 - Sp) using confusion matrix. The

area under the curve (AUC) for receiver operating characteristic (ROC) was calculated using formula $AUC = \int_0^1 ROC(\tau) d\tau \cong \frac{1}{2}(Se + Sp)$. Where $\tau = (1 - Sp)$ and ROC (τ) is sensitivity.

Table 1 represents the rank order by Bhattacharya method like X (Features), here X indicates rank of features and feature in parenthesis. Example for ranking order of features for NSR-CHF: index 20 (MSWPF15) at 1st rank, index 19 (MSWPF14) at 2nd rank,

Feature s	LF/HF Ratio	nLF	nHF	pLF	pHF	MSWPF1	MSWPF2	MSWPF3	MSWPF4	MSWPF5
Index	1	2	3	4	5	6	7	8	9	10
Feature s	MSWPF6	MSWPF7	MSWPF8	MSWPF9	MSWPF10	MSWPF11	MSWPF12	MSWPF13	MSWPF14	MSWPF15
Index	11	12	13	14	15	16	17	18	19	20
Feature s	MM	H1	H2	H3	P1	P2	PHE	WCOB1	WCOB2	FLATNESS
Index	21	22	23	24	25	26	27	28	29	30
NSR-CHF	20 MSWPF15	19 MSWPF14	18 MSWPF13	17 MSWPF12	16 MSWPF11	15 MSWPF10	14 MSWPF9	13 MSWPF8	12 MSWPF7	11 MSWPF6
NSR-ELY	16 MSWPF12	10 MSWPF7	15 MSWPF13	8 MSWPF4	12 MSWPF10	14 MSWPF9	18 MSWPF13	11 MSWPF6	19 MSWPF14	23 H2
ELD-CHF	13 MSWPF8	9 MSWPF4	7 MSWPF2	6 MSWPF1	29 WCOB2	14 MSWPF9	30 FLATNESS	25 P1	26 P2	28 WCOB1

IV. RESULTS AND DISCUSSION

A. Comparative Performance Using Area Under ROC Curve

In order to compare the performance in terms of area under ROC curve(AUC) achieved by the proposed method KPCA +1-NLPELM ; and KPCA +SVM ,KPCA +PNN, 1-NLPELM, SVM and PNN, a graph between AUC and top ten features ranked byBhattacharyaranking methods was drawn which are shown in **Fig. 4**for NSR-CHF(**Fig.4(a)**),for NSR-ELY (**Fig. 4(b)**) and forELY-CHF (**Fig.4(c)**).

Ranking order of top ten features based on Bhattacharya ranking is shown **Table.1**. Ranked top ten features were fed to the classifier one after another till the maximum AUC was reached. Note that KPCA reduced to one new feature in each feed and required initially minimum two features for this. In this simulation, additive hidden nodesfor 1-NLPELM and SVM were implemented using amultiquadric activation function and RBF.

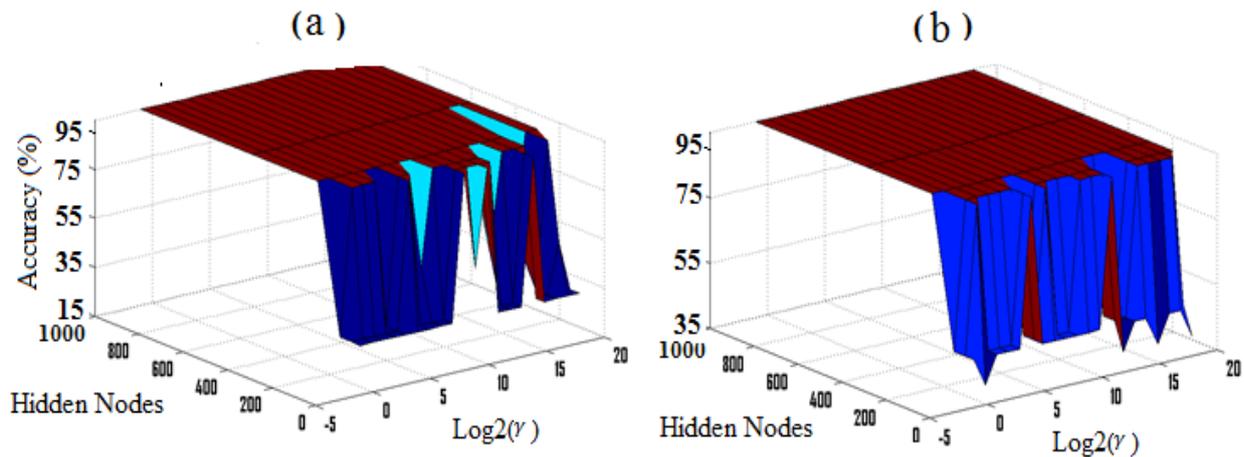


Fig. 5. Insensitivity of accuracy performance of proposed method KPCA with RBF kernel function + 1-NLPELM to the user specified parameters (γ, L) on two classification datasets (a) KPCA with RBF kernel function + 1-NLPELM with Sigmoid additive node for NSR-CHF database (b) KPCA with RBF kernel function + 1-NLPELM with multiquadric RBF node for NSR- CHF database.

Fig.4 (a) illustrates that the best AUC value was acquired by proposed method as KPCA +1-NLPELM for NSR-CHF dataset from 7 to 10 features. After KPCA +1-LPELM, the KPCA +SVM achieved highest AUC compared to KPCA +PNN, 1-NLPELM, SVM and PNN for each datasets from 8 to 10 features. This graph also reflects that AUC value decreased if number of features is less than 8. **Fig.4 (b)** shows that KPCA +1-NLPELM attained highest AUC after 8 features compared to other considered methods for NSR-ELY dataset. **Fig.4 (c)** demonstrates that the proposed technique achieved constant AUC value around 0.75 for ELY-NSR from 7 to 10 features while KPCA + SVM achieved constant AUC around 0.72 after 9 features. Results of **Fig. 4** shows that the proposed method performs better as compared to other considered approaches for each dataset. From results of graph, we also observed that upsurge in the AUC with KPCA compared to without KPCA, a clear result since KPCA with RBF kernel function was used as a feature space transformation method and convert to a new feature space by picking the best discriminating of ten features on the basis of Eigenvector. Hence, results of graph reveal that KPCA +1-NLPELM are an appropriate classifier or detector for CHF subjects.

B. Insensitivity Performance of Proposed Method

In order to validate whether the 1-NLPELM and proposed technique as KPCA with RBF kernel +1-NLPELM classifier results in a sparse model representation and least number of hidden nodes in the grit of the decision function for classification, NSR-CHF dataset is taken. The user specified parametric strength γ varies between $\{2^{-4}, \dots, 2^{18}\}$ and L from $\{10, 30, 60, \dots, 100, 300, 400, 800$ and $1000\}$.

Fig. 5 illustrates for every pair of user specified parametric strength (γ, L), with validation accuracies of proposed method (KPCA with RBF kernel function +1-NLPELM) for sigmoid additive node and multiquadric RBF node. **Fig.5(a)** shows that γ ranges from 2^5 to 2^{20} and L varies between 100 to 1000 and corresponding validation accuracy increases from 20% to 100% for NSR-CHF data set. This Figure also shows that for value of L between 200 to 1000, the validation accuracy is 100%. **Fig.5(b)** illustrates that proposed scheme achieved 100% validation accuracy for NSR-CHF data set, when γ varies from 2^5 to 2^{20} and L ranges from 200 to 1000. **Fig.5** reveals that proposed method is less sensitive to user specified parameter values and shows excellent generalization performance with very good accuracy. Hence proposed method is suitable for classification of cardiac diseases.

Table 2: Indicates P-values of features computed for the spectral methods. For clinical significant if $p > 0.05$: not significant, $p \leq 0.05$: significant and $p < 0.001$: very significant.

Features	LF/HF Ratio	nLF	nHF	pLF	pHF	MSWPF1	MSWPF2	MSWPF3	MSWPF4	MSWPF5
NSR-CHF (p-Value)	0.182	1.2 E-05	2.51E-05	1.58E-05	0.0005	1.83E-05	1.80E-05	4.21E-12	1.72E-05	4.31E-13
	MSWPF6	MSWPF7	MSWPF8	MSWPF9	MSWPF10	MSWPF11	MSWPF12	MSWPF13	MSWPF14	MSWPF15
	8.46E-12	7.85E-12	1.61E-05	2.93E-10	1.08E-11	9.89E-13	1.16E-10	1.99E-11	4.52E-11	1.60E-11
	MM	H1	H2	H3	P1	P2	PHE	WCOB1	WCOB2	FLATNESS
	2.60E-09	1.29E-09	8.67E-10	5.40E-09	0.0614	0.883	5.24E-01	3.89E-04	7.01E-03	2.21E-02
Features	LF/HF Ratio	nLF	nHF	pLF	pHF	MSWPF1	MSWPF2	MSWPF3	MSWPF4	MSWPF5
NSR-ELY (p-Value)	0.936	0.936	0.936	0.0019	0.0531	5.89E-07	3.20E-07	0.0029	2.96E-07	0.16
	MSWPF6	MSWPF7	MSWPF8	MSWPF9	MSWPF10	MSWPF11	MSWPF12	MSWPF13	MSWPF14	MSWPF15
	0.0003	0.023	2.96E-07	0.791	0.112	0.293	0.0004	0.00046	0.143	0.0041
	MM	H1	H2	H3	P1	P2	PHE	WCOB1	WCOB2	FLATNESS
	0.189	0.1067	0.4333	0.002	3.69E-05	7.76E-05	0.2415	0.00363	6.35E-07	0.046
Features	LF/HF Ratio	nLF	nHF	pLF	pHF	MSWPF1	MSWPF2	MSWPF3	MSWPF4	MSWPF5
ELY-CHF (p-Value)	0.3385	0.3385	0.3385	0.0001	0.0384	3.46E-08	2.45E-08	0.01021	2.20E-08	0.1494
	MSWPF6	MSWPF7	MSWPF8	MSWPF9	MSWPF10	MSWPF11	MSWPF12	MSWPF13	MSWPF14	MSWPF15
	0.0067	0.0308	2.20E-08	0.649	0.0978	0.216	0.00368	0.0079	0.0628	0.0085
	MM	H1	H2	H3	P1	P2	PHE	WCOB1	WCOB2	FLATNESS
	0.067	0.0615	0.7781	0.0011	0.00017	0.0009	0.0998	0.0922	2.86E-05	0.0803

C. Statistical Comparison of Features

A Student’s t-test one-tailed was applied to statistically compare the features extracted by considered non linear methods from HRV datasets. The Student’s t-test provides a p-value which is useful to predict the significance of features for detection of CHF. In the detection of CHF consequence, features should be dissimilar hence a low p-value of test is preferred [44]. Thus, a low p-value shows there is high probability that the individual classes of HRV data sets are detachable [50-51]. In a physiological diagnostic setting, a p-value less than 0.05 specifies clinical significance and less than 0.001 very significance. **Table 2** depicts the p-values of five features obtained from AR spectral method, fifteen from MSWP and ten from HOS method for each dataset. The results of **Table 3** indicate that p-value attained by features of AR spectra, MSWP and HOS are very significant ($p < 0.001$) for NSR-CHF dataset. But, features P1 and P2 are not significant for this dataset. While p-value obtained by features of AR spectra are not significant except pLF feature for NSR-ELY and ELY-CHF datasets. The features like MSWPF1, MSWPF2, MSWPF4, MSWPF8, MSWPF12 MSWPF13, P1, P2 and WCOB2 are significant ($p < 0.001$) for NSR-ELY set. So, these features can be used for clinical setting. The feature DFA- α_2 and IMPE are significant ($p < 0.05$) only for YNG-ELD dataset. Only few features as MSWPF1, MSWPF2, MSWPF4, MSWPF8, P1, H3, WCOB2 and P2 are significant for ELY-CHF dataset. The results of **Table 2** reveal that MSWPF are more appropriate than AR and HOS features for CHF detection.

D. Performance of Proposed Method with Ranking Methods

In order to demonstrate the performance, the classification accuracy was calculated for ranking methods+1-NLPELM and ranking methods + KPCA+1-NLPELM. The classification performance was measured using a 15 trials 10-fold cross validation scheme to

test robustness and ensure generalization performance of our proposed scheme. For each trial of the 10-fold cross-validation, the data was randomly divided as 90 out of 160 from NSR-ELY, 80 out of 140 from NSR-CHF and 80 out of 140 from ELY-CHF data sets for training and rest data set for validation are shown in **Table 3**. The error rate in terms of standard deviation (\pm S.d) was then evaluated on the 10th fold. Hence, the learning process was performed a total of 15 times on different training and validation sets. This resulted in the learning procedure being implemented 150 times on datasets. The average validation accuracy was evaluated for each dataset and 10 error estimates were averaged to produce a total error evaluation.

Table 3 summarizes the training and validation sample sizes of data sets used for simulation, and accuracy achieved by ranking methods+1-NLPELM and ranking methods+KPCA+ 1-NLPELM classifier using sigmoidal, multiquadric activation functions and RBF kernel function for KPCA. Result of **Table 3** shows that the 1-NLPELM classification with KPCA and ranking methods performed better than 1-NLPELM with ranking methods. The Fisher+KPCA +1-NLPELM with Sigmoid and Multiquadric activation function produced an accuracy of $97.32 \pm 1.15\%$ and 97.06 ± 1.42 for NSR-CHF dataset. For this dataset and activation functions, the Bhattacharya+KPCA + 1-NLPELM achieved an accuracy of 96.93 ± 1.12 and 97.44 ± 1.6 . While for same activation functions and dataset, the Fisher+ 1-NLPELM and Bhattacharya +1-NLPELM yielded an accuracy of 60.54 ± 11.04 and 66.67 ± 4.7 , and 72.41 ± 10.92 and 68.54 ± 4.77 . The Wilcoxon +KPCA +1-NLPELM and Entropy+ KPCA +1-NLPELM produced an accuracy of 85.06 ± 22.17 and 97.32 ± 1.29 for NSR-CHF dataset using sigmoidal activation function. The ROC+KPCA+1-NLPELM attained an accuracy of 97.19 ± 0.61 and 96.81 ± 0.96 for NSR-CHF dataset using Sigmoid and Multiquadric activation function. However ROC+1-NLPELM achieved an accuracy of 69.35 ± 10.29 using Sigmoid and an accuracy of 74.53 ± 5.42 using multiquadric activation function.

Table 3: presents classification performance in terms of accuracy achieved by proposed method with ranking methods for considered datasets. The binary classification of data sets

are carried out by using ranking methods+1-NLPEL and ranking methods+KPCA having RBF kernel function + 1-NLPELM having both sigmoid and multiquadric hidden nodes. The validation accuracy (%) ± S.D is shown for the optimal parameter values $L = 100$. The best result is indicated by boldface in table.

Database (Train size, Test size)	Fisher+ 1-NLPELM		Fisher+ KPCAA+1-NLPELM	
	<i>Sigmoid</i>	<i>Multiquadric</i>	<i>Sigmoid</i>	<i>Multiquadric</i>
NSR-CHF (80×10,60×10)	60.54±11.04	66.67±4.7	97.32±1.15	97.06±1.42
NSR-ELY (90×10, 70×10)	61.36±4.4	70.71±8.31	82.07±4.01	83.84±8.07
ELY-CHF (80×10, 60×10)	75.64±2.89	75±7.51	76.92±9.12	83.76±7.94
Database (Train size, Test size)	Wilcoxon+ 1-NLPELM		Wilcoxon+ KPCA+1-NLPELM	
	<i>Sigmoid</i>	<i>Multiquadric</i>	<i>Sigmoid</i>	<i>Multiquadric</i>
NSR-CHF (80×10,60×10)	74.58±8.06	69.41±7.21	85.06±22.17	83.91±21.43
NSR-ELY (90×10, 70×10)	63.64±7.62	82.83±2.3	78.79±3.77	82.32±4.79
ELY-CHF (80×10, 60×10)	68.8±8.48	74.36±9.02	79.7±4.31	79.49±2.54
Database (Train size, Test size)	Bhattacharya+1-NLPELM		Bhattacharya+ KPCA+1-NLPELM	
	<i>Sigmoid</i>	<i>Multiquadric</i>	<i>Sigmoid</i>	<i>Multiquadric</i>
NSR-CHF (80×10,60×10)	72.41±10.92	68.54±4.77	96.93±1.12	97.44±1.6
NSR-ELY (90×10, 70×10)	82.07±3.49	78.03±7.79	86.87±4.37	90.4±7.85
ELY-CHF (80×10, 60×10)	71.58±10.87	78.2±4.19	79.06±5.39	87.82±5.85
Database (Train size, Test size)	Entropy+ 1-NLPELM		Entropy+KPCA+1-NLPELM	
	<i>Sigmoid</i>	<i>Multiquadric</i>	<i>Sigmoid</i>	<i>Multiquadric</i>
NSR-CHF (80×10,60×10)	75.35±16.12	69.66±4.17	97.32±1.29	96.55±1.52
NSR-ELY (90×10, 70×10)	77.78±3.55	79.55±4.95	78.54±5.58	84.34±4.17
ELY-CHF (80×10, 60×10)	61.11±9.12	77.35±7.94	79.06±2.89	86.11±7.51
Database (Train size, Test size)	ROC+ 1-NLPELM		ROC+ KPCA+1-NLPELM	
	<i>Sigmoid</i>	<i>Multiquadric</i>	<i>Sigmoid</i>	<i>Multiquadric</i>
NSR-CHF (80×10,60×10)	69.35±10.29	74.53±5.42	97.19±0.61	96.81±0.96
NSR-ELY (90×10, 70×10)	78.53±6.34	80.81±3.03	89.55±7.96	85.86±4.92
ELY-CHF (80×10, 60×10)	78.42±4.38	76.71±8.52	86.96±7.37	91.88±7.5

Overall results shows that there was an increase in the accuracy with KPCA combined with classifier compared to without KPCA, since KPCA was used as a feature space reduction method and reduced the feature space to one new feature vector by selecting best discerning of rank features. Numerical results also shows that proposed ranking methods +KPCA + 1-NLPELM classification system produced high accuracy for all considered datasets when using multiquadric activation function compared to using Sigmoid activation function.

V. CONCLUSION

In the present work, a novel approach is presented for detection of CHF. This approach is based on ranking methods, ranked feature dimension reduction scheme as KPCA and 1-NLPELM as binary classifier. For Several ranking techniques, only one ranked feature dimension reduction scheme are used to enhance the classification performance. Detection of CHF is based on the features extracted by spectral methods. It has also been apparent that features at third level of decomposition of HRV data sets by MSWP have lowest p-value (< 0.001). As, lowest p-value features has highest discrimination ability, so third level of MSWP features are better than considered other features extracted by AR and HOS spectral methods for HRV data set analysis. Therefore, the proposed scheme can be used for diagnosis of CHF cardiac diseases.

Conflict of interest: None.

REFERENCES

- [1] World Health Organization, "Global status report on non-communicable diseases", 2010, Available online at: http://www.who.int/cardiovascular_diseases/about_cvd/en/ [last accessed December 2016].
- [2] N. D. Wong, "Epidemiological studies of CHD and the evolution of preventive cardiology", *Nature Reviews Cardiology*, vol. 11, 2014, pp. 276–289.
- [3] National Heart Lung and Blood Institute, "What is coronary heart disease?" Accessed 2015-11-25 <http://www.nhlbi.nih.gov/health/health-topics/topics/cad/>.
- [4] WHO (World Health Organization), "Definition and Diagnosis of cardiac disease and Intermediate Hyperglycemia", 2006, http://whqlibdoc.who.int/publications/2006/9241594934_eng.pdf, (Last accessed December 2016.).
- [5] M. G. Poddar, V. Kumar and Y. P. Sharma, "Automated diagnosis of coronary artery disease patients by heart rate variability analysis using



- linear and non-linear methods”, *J Med Eng Technol*, vol. 39, 2015, pp. 331-341.
- [6] E.Silber and L.N. Katz, “Heart Disease”, *New York: Macmillan Publishing Co.*, 1975, pp. 498.
- [7] U.R. Acharya, N.Kannathal and S.M.Krishnan, “Comprehensive analysis of cardiac health using heart rate signals”, *Physiological Measurement*, vol. 25, 2004, pp. 1139–1151.
- [8] S.Nikolopoulos, A.Alexandridi, S.Nikolakeas and G.Manis, “Experimental analysis of heart rate variability of long-recording electrocardiograms in normal subjects and coronary patients”, *J. Biomed. Inform*, Vol 36, 2003, pp. 202–217.
- [9] G.Manis, S.Nikolopoulos, A. Alexandridi and C. Davos, “Assessment of the classification capability of prediction and approximation methods for HRV analysis”, *Computers in Biology and Medicine*, vol. 37, 2007, pp. 642 – 654.
- [10] A.L.Goldberger, “Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside”, *Lancet*, vol 11, 1996, pp.1312–1314.
- [11] Y. Sun, K. L. Chan and S. M. Krishnan, “Arrhythmia detection and recognition in ECG signals using nonlinear techniques”, *Ann. Biomed. Eng*, Vol. 28, 2000, pp. 37-40.
- [12] S.M.Pincus, “Approximate entropy as a measure of system complexity”, *Proc. Nat. Acad. Science*, vol 88, 1991, pp. 2297–2301.
- [13] U.R.Acharya, N.Kannathal and S.M.Krishnan, “Comprehensive analysis of cardiac health using heart rate signals”, *Physiol. Measurement*, vol 25, 2004, pp. 1139–1151.
- [14] A.L.Goldberger and B.J.West, “Applications of nonlinear dynamics to clinical cardiology”, *Ann. N.Y. Acad. Science*, vol. 50, 1987, pp. 195–213.
- [15] G. B.Huang, Q. Yu. Zhu and C.K.Siew, “Extreme learning machine: Theory and applications”, *Neurocomputing*, vol 70, 2006, pp. 489–501.
- [16] G. B.Huang, D.H.Wang and Y.Lan, “Extreme learning machines: a survey”, *Int. J. Mach. Learn. Cybernet*, vol. 2, 2011, pp. 107–122.
- [17] Z.Chen, H. Zhu And Y.Wang, “A modified extreme learning machine with sigmoidal activation functions”, *Neural Comput. Appl.*, vol. 22, 2013, pp. 541–550.
- [18] Y.Lan, C.Soh, and G. B.Huang, “Two-stage extreme learning machine for regression”, *Neurocomputing*, vol 73, 2010, pp. 3028–3038.
- [19] Q. Y.Zhu, A.K.Qin, P.N.Suganthan and G.B.Huang, “Evolutionary extreme learning machine”, *Pattern. Recognit.*, vol 38, 2005, pp. 1759–1763.
- [20] Y.Yuan and Y. F.Wang, “Optimization approximation solution for regression problem based on extreme learning machine”, *Neurocomputing*, vol 74, 2011, pp. 2475–2482.
- [21] L.Zhang and W.Zhou, “On the sparseness of 1-norm support vector machines”, *Neural Networks*, vol. 23, 2010, pp. 373-385.
- [22] S.Balasundaram, D.Gupta, and Kapil, “1-Norm extreme learning machine for regression and multiclass classification using Newton method”, *Neurocomputing*, vol 128, 2014, pp. 4–14.
- [23] A. L.Goldberger, L. A.N.Amaral, L.Glass and J. M. Hausdorff, “PhysioBank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals”, *Circulation*, vol. 101, 2000, pp. 215-220, (<http://www.physionet.org>).
- [24] N.Iyengar, C. K.Peng, A. L. Goldberger, R. Morin and L. A. Lipsitz, “Age-related alterations in the fractal scaling of cardiac inter beat interval dynamics”, *Am J Physiol*, vol. 271, 1996, pp. 1078-1084.
- [25] I. P.Mitove, “A method for assessment and processing of bio medical signals containing trend and periodic Components”, *Med Eng Phys*, vol 20, 1998, pp. 660-668.
- [26] U. R.Acharya, K. P. Joseph, N. Kannathal, C.M. Lim and J. S. Suri, “Heart rate variability: a review”, *Med Bio Eng Comput*, vol. 44, 2006, pp. 1031–1051.
- [27] J. T.Ramshur, “Design, evaluation, and application of heart rate variability analysis software (HRVAS)”, thesis, vol. 23, 2010.
- [28] A.Subasi, “Comparison of subspace-based methods with AR parametric methods in epileptic seizure detection”, *Computers in Biology and Medicine*, vol 36, 2006, pp. 195-208.
- [29] A.Boardman, “A study on the optimum order of autoregressive models for heart rate variability”, *Physiological Measurement*, vol 23, 2002, pp.325-328.
- [30] R.K.Sunkaria, V.Kumar, and S.C.Saxena, “Aging effects on dynamics: A comparative study with FFT and AR models”, *International Journal of Signal and Imaging Systems Engineering*, vol 6, 2013, pp. 240–249.
- [31] R. R.Coifman, Y.Meyer, S.Quake and V.Wickerhauser, “Wavelet analysis and Signal processing”, *Wavelets and Their Applications*, Sudbury, MA: Jones and Barlett, vol 2, 1992, pp. 153–178.
- [32] R. N.Khushaba, S. Kodagoda, G.Dissanayake and S. Lal, “Driver Drowsiness Classification Using Fuzzy Wavelet-Packet-Based Feature-Extraction Algorithm”, *IEEE Transactions On Biomedical Engineering*, vol 58, 2011, pp. 512-518.
- [33] K. Chua, V. Chandran, U. Acharya and C. Lim, “Cardiac state diagnosis using higher order spectra of heart rate variability”, *Journal of*

- medical engineering & technology*, vol 32, 2008, pp. 145-155.
- [34] K. C.Chua, V. Chandran, U. R. Acharya and C. M. Lim, "Application of higher order statistics spectra in biomedical signals-A review", *Medical engineering & physics*, vol 32, 2010, pp. 679–89.
- [35] N. Madhu, "Note on measures for spectral flatness", *ELECTRONICS LETTERS*, vol 45, 2009, pp. 1195 -1196.
- [36] R. O.Duda, P. E. Hart, and D. G.Stork, *Pattern classification*, John Willey & Sons, 2000, ch. 2.
- [37] M.Kumar, R.B.Pachori and U. R. Acharya, "An efficient automated technique for CAD diagnosis using flexible analytic wavelet transform and entropy features extracted from HRV signals", *Expert Systems with Applications*, vol 63, 2016, pp. 165–172.
- [38] Gu Quanquan, Li Zhenhui, Han Jiawei, "Generalized Fisher Score for Feature Selection", *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*, Barcelona, Spain, 2011, July 14-17.
- [39] B. D. R.Derry, S. B.Schou, and W. J.Conover, "Teaching rank based tests by emphasizing structural similarities to corresponding parametric tests", *Journal of Statistics Education*, vol. 18, 2010, pp. 1–19.
- [40] A. Kampouraki, G. Manis and C. Nikou, "Heartbeat Time Series Classification with Support Vector Machines", *IEEE Transactions on Information Technology In Biomedicine*, vol 13, 2009, pp. 512-518.
- [41] S. Dua, X. Du, S. V. Sree and T. A., "Novel classification of coronary artery disease using heart rate variability analysis", *Journal of Mechanics in Medicine and Biology*, vol 12, 2012, pp. 124-129.
- [42] S. Patidar, R.B. Pachori and U. Rajendra Acharya, "Automated diagnosis of coronary artery disease using tunable-Q wavelet transform applied on heart rate signals", *Knowledge-Based Systems*, vol 82, 2015, pp. 1–10.
- [43] I. Babaoglu, O Findik and M. Bayrak, "Effects of principle component analysis on assessment of coronary artery diseases using support vector machine", *Expert Systems with Applications* vol 37, 2010, pp.2182–2185.
- [44] M. G.Poddar, V.Kumar, and Y. P.Sharma, "Automated diagnosis of coronary artery disease patients by heart rate variability analysis using linear and non-linear methods", *Journal of Medical Engineering & Technology*, vol 39, 2015, pp. 331–341.
- [45] D.Giri, U. R. Acharya, R. J. Martis, S. V Sree., T.C.Lim, and J. S. Suri, "Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform", *Knowledge-Based System*, vol 37, 2013, pp. 274–282.
- [46] C. R.Rao and S. K.Mitra, *Generalized Inverse of Matrices and its Applications*, John Wiley, New York, 1971.
- [47] S. S.Shwartz and A.Tewari, "Stochastic methods for ℓ_1 -regularized loss minimization", *In Proceedings of the 26th International Conference on Machine Learning, Montreal, Canada*, 2009, pp. 929–936.
- [48] O. L.Mangasarian, "Exact 1-norm support vector machines via unconstrained convex differential minimization", *J. Mach. Learn. Res.*, vol 7, 2006 ,pp. 1517–1530.
- [49] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification", *IEEE Trans. on Syst. Man Cybern. Part B: Cybern*, vol. 42, 2012, pp. 513–528.
- [50] R.S.Singh, B.S.Saini and R. K.Sunkaria, "Classification of cardiac heart disease using reduced chaos features and 1-norm linear programming extreme learning machine", *International Journal for Multiscale computational engineering*, vol 16, 2018, pp. 465-486.
- [51] R.Alizadehsani, M. J. Hosseini, A. Khosravi and F.Khozeimeh, "Non invasive detection of coronary artery disease in high risk patients based on the stenosis prediction of separate coronary arteries", *Computer methods and programs in Biomedicine*, vol 162, 2018, pp. 119-127.

AUTHORS PROFILE



Sulekha Saxena was born on 15 Apr 1983, in Allahabad. Obtained her Bachelor's degree in Electronics Engineering from Institute of Engineering and Rural Technology, Allahabad in 2005 and M.Tech Degree in VLSI System Design from Uttar Pradesh Technical University, Lucknow in 2013. She is pursuing her PhD from Dr. A P J Abdul Kalam Technical University, Lucknow. Currently she is working as Assistant Professor in IMS Engineering College with more than 13 years of technical teaching experience. Published over ten papers in National and International Journals of repute. Her current research interests include Biomedical Signal Processing. She is a member of the ICEIT Society.



Dr. P.N. Hrisheeksha is a B. E, M. Tech, Ph. D from IIT Roorke. Currently he is working as campus director at CGC landran. Has got more than 50 publications in reputed international and national journals and conferences. He has completed so many



funded projects. He is the editor of few reputed journals. He has Guided more than 50 PG and few Ph.Ds. He has few patents and books to his credit. He has established many incubation centers, Center of innovation, new labs too. His area of research and interest are renewable energy sources, power systems, distribution system automation, image processing, energy conservation, audit, printed electronics, AI, machine learning etc.



Dr. Vijay Kumar Gupta has born in 1976. He has received B.Tech degree from MJP Rohilkhand University, Bareilly, India in 2000, M.Tech degree in Instrumentation from DAVV, Indore, India in 2003 and PhD from Birla Institute of Technology, Mesra, Ranchi, India in 2013. His research interest includes signal, speech and image processing. Currently, he is working as Professor and Head, Department of Electronics and Communication Engineering in Inderprastha Engineering College, Ghaziabad (UP), India. He has published more than 30 research papers in the area of speech and image processing at National and International journals and conferences of repute. He is a life member of ISTE and member of IEEE.



Dr. Ram Sewak Singh was born on 15 Feb 1978, in Patna. Obtained his Bachelor's degree in Electronics and Tele Communication Engineering from College of Engineering, Roorkee in 2003 and M.Tech Degree in Electronics & Communication Engineering from National Institute of Technology (NIT) Kurukshetra in 2009. He has completed his PhD from National Institute of Technology (NIT) Jalandhar in 2019. Currently, he is working as Assistant Professor in IMS Engineering College and has more than 13 years of technical teaching experience. He was awarded the Haryana Young scientist award-2008 for research paper on "Handover Decision Algorithm for Next Generation hybrid wireless network". He has published more than 15 research papers International Journals and 9 in national and international conferences. His current research interests include Wireless communication, Wireless Sensors, and bio signal processing. He is a member of the ICEIT Society. He is also reviewer of Springer, IEEE access, and world scientific journals and, editorial board member of International Journal of Engineering and Applied Sciences (EAS), Science Publishing Group 548, FASHION AVENUE NEW YORK, NY 10018 U.S.A.