Research Article

Deep Layer Kernel Sparse Representation Network for the Detection of Heart Valve Ailments from the Time-Frequency Representation of PCG Recordings

Samit Kumar Ghosh [ID],[1] **R. N. Ponnalagu** [ID],[1] **R. K. Tripathy** [ID],[1] **and U. Rajendra Acharya** [ID][2]

[1]*Department of Electrical and Electronics Engineering, BITS-Pilani, Hyderabad Campus, Hyderabad 500078, India*
[2]*Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore*

Correspondence should be addressed to R. K. Tripathy; tripathyrk@hyderabad.bits-pilani.ac.in

The heart valve ailments (HVAs) are due to the defects in the valves of the heart and if untreated may cause heart failure, clots, and even sudden cardiac death. Automated early detection of HVAs is necessary in the hospitals for proper diagnosis of pathological cases, to provide timely treatment, and to reduce the mortality rate. The heart valve abnormalities will alter the heart sound and murmurs which can be faithfully captured by phonocardiogram (PCG) recordings. In this paper, a time-frequency based deep layer kernel sparse representation network (DLKSRN) is proposed for the detection of various HVAs using PCG signals. Spline kernel-based Chirplet transform (SCT) is used to evaluate the time-frequency representation of PCG recording, and the features like L1-norm (LN), sample entropy (SEN), and permutation entropy (PEN) are extracted from the different frequency components of the time-frequency representation of PCG recording. The DLKSRN formulated using the hidden layers of extreme learning machine- (ELM-) autoencoders and kernel sparse representation (KSR) is used for the classification of PCG recordings as normal, and pathology cases such as mitral valve prolapse (MVP), mitral regurgitation (MR), aortic stenosis (AS), and mitral stenosis (MS). The proposed approach has been evaluated using PCG recordings from both public and private databases, and the results demonstrated that an average sensitivity of 100%, 97.51%, 99.00%, 98.72%, and 99.13% are obtained for normal, MVP, MR, AS, and MS cases using the hold-out cross-validation (CV) method. The proposed approach is applicable for the Internet of Things- (IoT-) driven smart healthcare system for the accurate detection of HVAs.

1. Introduction

The heart valve ailments (HVAs) are cardiovascular abnormalities, and these ailments occur due to the defect in any of the valves (tricuspid, pulmonary, mitral, and aortic) of the heart [1, 2]. The valves of the heart prevent the backward flow of the blood, and for the proper functioning of the heart, the valve should be effectively closed or opened during the flow of blood from one chamber to another chamber of the heart [3]. The HVAs are classified as mitral stenosis (MS), mitral valve prolapse (MVP), mitral regurgitation (MR), and aortic stenosis (AS) based on the defect in the heart valves [4]. The MR ailments occur due to the improper clos-

ing of the mitral valve, which further causes the reverse flow of blood from the left ventricle to the left atrium [5]. Similarly, the AR refers to the improper closing of the aortic valve; as a result, the backward flow of blood from the aorta to the right ventricle may occur [5]. Moreover, the MS is termed as the problem in the opening of the mitral valve, where the left ventricle is not getting a sufficient amount of blood from the left atrium [6]. Similarly, the AS pathology refers to the improper opening of the aortic valve, which prevents the flow of blood from the left ventricle to the aorta of the heart [5] [6]. For the diagnosis of these pathologies, different imaging techniques such as computed tomography scan, magnetic resonance imaging (MRI), cardiac echocardiography, and

ultrasonic devices have been used [7–10]. It has been reported from the literature that various quantitative parameters such as transvalvular velocity, average value area, and mean value of transvalvular gradient have been considered to determine the progression of HVAs [11]. The aforementioned imaging modalities have limitations, such as the selection of tuning parameters in ultrasonic devices to obtain better resolution images of heart chambers and valves for the diagnosis of HVAs [10, 12]. Also, these imaging techniques are costly and require trained medical staff for the accurate assessment of HVAs [13]. The phonocardiography (PCG) is a noninvasive and low-cost diagnostic test used for the detection of HVAs [14, 15]. The diagnostic features such as the duration of both the systolic segment and diastolic segment, morphologies of both S1 and S2 components, and the appearance of murmurs have been investigated for the diagnosis of HVAs [14, 16]. To assist the clinicians in the diagnosis of HVAs, an automatic diagnosis system (ADS) will be helpful especially while treating patients admitted in the intensive care unit where continuous recording and monitoring of PCG signal is done 24 hours [17]. The ADS comprises the evaluation of various diagnostic features from the PCG recording and automated classification of HVAs using the PCG signal features [13]. For smart healthcare and the Internet of healthcare things (IoHT) applications [18, 19], the automated diagnosis of HVAs from the PCG signal is a challenging area of research. Therefore, the development of new methods for the extraction of PCG signal features and the classification of HVAs is required.

In the last two decades, various algorithms have been used for the automated detection of HVAs using PCG signals. These algorithms have considered different feature extraction methodologies to extract the features from the PCG signal and used various machine learning classifiers for the categorization of HVAs. A review of various automated methods for the detection of HVAs has been reported in [20, 21]. The time, frequency, time scale, and time-frequency domain-based features from PCG signal have been used for the detection of HVAs. The time-domain features from the PCG signals have been used in [22–26], for the categorization of both normal and abnormal heart sounds. Similarly, in [27–30], the frequency domain features from the PCG signals have been considered for the discrimination of normal and abnormal cardiac sounds. The time-scale-based methods such as discrete wavelet transform (DWT) [31, 32], empirical mode decomposition (EMD) [31, 32], and tunable Q-wavelet transforms (TQWT) [33] of PCG signals have also been used for the detection of HVAs. Moreover, the time-frequency analysis-based approaches such as the short-time Fourier transform (STFT) [34, 35], synchrosqueezing transform (SST) [36], and other time-frequency decomposition-based approaches [37–39] of PCG signals are used for the categorization of HVAs. The machine learning techniques such as the support vector machines (SVM) [40], random forest (RF) [41], convolutional neural network (CNN) [42], and hidden Markov model (HMM) [43] have been used for the classification of HVAs. It is evident from the literature that time-frequency and time-scale analysis-based approaches have demonstrated higher classification

performance for the detection of HVAs using PCG signals. Son et al. [44] have combined the Mel frequency cepstral coefficients (MFCC) and DWT-based features from the PCG signals and used these features for the detection of HVAs. They have considered various machine learning classifiers for HVA detection. In [45], the authors have applied a novel algorithm based on wavelet fractal dimension and a twin support vector machine (TWSVM) for the classification of HVAs using PCG signals. Moreover, Ghosh et al. [36] have extracted the magnitude and phase features from the time-frequency representation of the segmented PCG cycles for the discrimination of HVAs. They have used synchrosqueezing transform (SST) for the evaluation of the time-frequency matrix from the PCG signal. The SST-based method has drawbacks such as it has poor time-frequency resolution for PCG signals as it uses the coefficient reassignment in the time-frequency plane based on the instantaneous frequency of the PCG signal [36, 46]. Also, the SST method has shown less performance for the detection of HVAs. The methods reported in the literature have segmented the PCG signal into cardiac cycles and then extracted features from the segmented cardiac heart sound cycles for the detection of HVAs. The PCG signal with multiple cardiac heart sound cycles effectively captures the variations in the amplitudes and shapes of S1 and S2 sound components and the duration of systolic and diastolic segments [14]. The existing approaches have not considered the PCG signals from all HVA classes to design the automated diagnosis frameworks. Therefore, an intelligent system which uses PCG signal with multiple cardiac heart sound cycles and classifies all HVAs is required for healthcare applications.

The PCG signal is nonstationary, and the components of this signal such as S1, S2, and murmurs are nonlinear and time-varying [47, 48]. In our previous work, we have analyzed the PCG signal using Chirplet transform (CT) for the detection of HVAs [44]. The CT works well for chirp-like signals with linearly time-varying components [49, 50]. But the CT fails to capture the transition from S1 component to systolic murmur, and from S2 component diastolic murmur in the time-frequency plot of the pathological PCG signals [13]. In this work, we have considered the spline CT (SCT) as the extension of CT for the evaluation of the time-frequency matrix from the PCG signal. The SCT has advantages such as it has better time-frequency localization for the nonlinearly time-varying components of the nonstationary signal as compared to CT [51]. Therefore, we can expect that the time-frequency matrix computed using SCT of the PCG signal can effectively capture the pathological variations and provide better resolution in the time-frequency domain of the PCG signal. Recently, the convolutional neural network (CNN) and stacked autoencoder- (SAE-) based deep neural network (DNN) methods have been used for the automated assessment of HVAs using PCG signals [44, 52]. In order to obtain the optimal parameters in CNN and SAE networks, rigorous training based on the gradient descent algorithm is used [53]. Also, these networks require more instances during the training process for obtaining the optimal model parameters [54]. The DNN based on extreme learning machine- (ELM-) autoencoder has advantages such
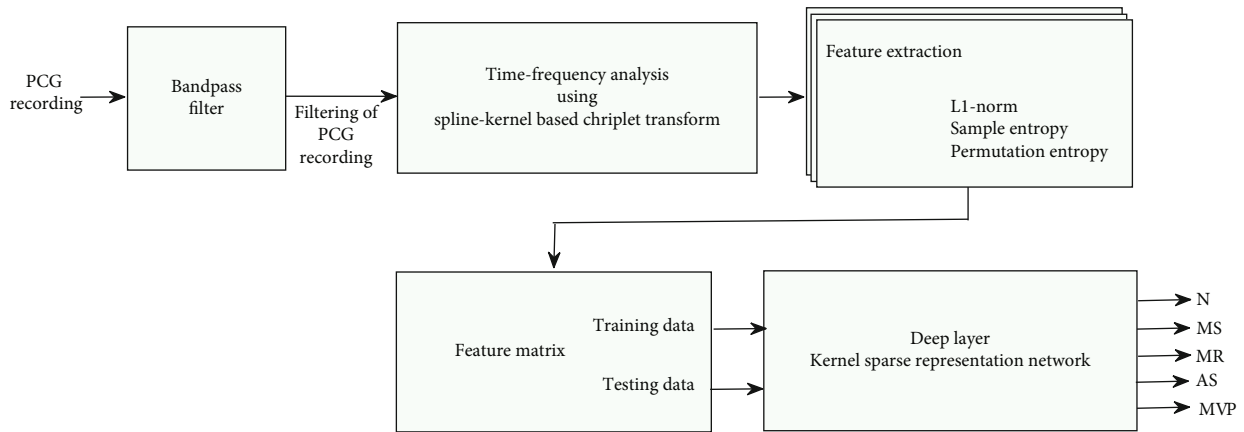
FIGURE 1: Flow diagram of the proposed approach for HVAs detection.

as it requires less training time for the evaluation of the model parameters [55], and the ELM-autoencoder model can be efficiently implemented in real-time for the dimension reduction [56]. The sparse representation-driven classification methods have been widely used for various biomedical applications [13, 57–59]. These methods require fewer features for training instances and also have fewer training parameters for the prediction of class labels from the test feature vectors [59]. The SRC has shown better performances as compared to other machine learning approaches for the classification of HVAs from PCG signal features [13]. The kernel sparse representation classifier (KSRC) uses the kernel trick to map the feature instances to the higher dimensional space, and the SRC is applied in the higher dimensional space for the classification [60, 61]. The KSRC has shown better classification performance for the dataset which consists of nonlinearly separable feature instances as compared to SRC [57, 62]. Therefore, the DNN developed based on the ELM-autoencoder, and KSRC will be effective for the automated detection of HVAs using the time-frequency representation of the PCG signal. The contributions of this paper are written as follows:

(i) The SCT-based time-frequency analysis is used for the evaluation of time-frequency representation of PCG recording

(ii) The nonlinear features such as the L1-norm (LN), sample entropy (SEN), and permutation entropy (PEN) are computed from different frequency components of the SCT-based time-frequency matrix of PCG signals

(iii) The deep layer kernel sparse representation network (DLKSRN) is proposed for the detection of HVAs using the time-frequency domain features of the PCG signal

The remaining sections of this manuscript are written as follows. In Section 2, the proposed method for the detection and classification of HVAs is described. The results obtained from the proposed work are discussed in Section 3, and conclusions are presented in section 4.

2. Proposed Method

The flow diagram of the proposed HVA detection approach is depicted in Figure 1, and the details of the various steps involved in the proposed approach are explained in detail in the following subsection.

2.1. PCG Signal Collection and Filtering. In this work, we have collected the PCG recordings from a public database available in (https://github.com/yaseen21khan/Classification-of-Heart-Sound-Signal-Using-Multiple-Features-). The detailed description of the PCG signals database is given in [44]. The database contains a total of 1000 PCG recordings of different classes. Out of those 1000 recordings, each class (normal or pathological) contains 200 PCG recordings. The annotations for the PCG signals for normal (N) and pathological (MS, MR, AS, and MVP) classes are given in the database. The PCG recordings are given in wav file format, and these signals were recorded from the subjects with different time durations [44]. The resolution of each PCG recording in the database is 16 bits, and the sampling frequency is 8 kHz. In this work, the collected PCG recordings are downsampled to 4 kHz for the time-frequency analysis. Moreover, we have also evaluated the performance of the proposed approach using the database available in 15 recorded PCG signals. These 15 PCG signals were recorded from 15 different subjects (12 males and 3 females with the age group of 27 ± 5 years) using Thinklab digital stethoscope (https://www.thinklabs.com/). The subjects have given written consent before recording the PCG signal in a noninvasive way [36]. The sampling frequency of each recorded signal is 4 kHz. In this work, we have also considered the Michigan heart sound and murmur database (MHSMD) (http://www.med.umich.edu/lrc/psb_open/html/repo/primer_heartsound/primer_heartsound.html) to evaluate the performance of the proposed method. The MHSMD contains both normal and abnormal (AS, MS, MR, and MVP) PCG signals with a sampling frequency of 44.1 kHz [63]. Each PCG signal from MHSMD has been downsampled to 4 kHz. For each database PCG recording, a Butterworth bandpass filter with a lower and upper cutoff frequency of 25 Hz and 900 Hz is used [64]. After filtering, the amplitude normalization is performed with respect to

the maximum amplitude value of the PCG recording. The normalized PCG recording, $x(n)$, is evaluated as follows [65]:

$$x(n) = \frac{x(n)}{\max\left[|x(1)|, |x(2)|, \cdots, |x(N)|\right]}, \quad (1)$$

where $|x(n)|$, $n = 1, 2, \cdots, N$ is the absolute value of the amplitude of nth sample of the PCG recording, and N is the total number of samples. After normalization, the time-frequency representation of each PCG recording is computed using SCT. The following subsection describes the spline kernel-based CT for the extraction of the time-frequency matrix from PCG recording.

2.2. Spline Kernel-Based Chirplet Transform (SCT). The spline kernel-based CT is the CT with a modified kernel function [51]. This modified kernel function uses different frequency rotate and frequency shift operators for the time-frequency representation of the nonstationary signal. For a PCG signal, $x(n)$ containing N samples, the discrete SCT is evaluated as follows [51]:

$$\mathbf{T}(\tilde{n}, k) = \sum_{n=0}^{N-1} \bar{x}(n) w_\sigma(n - \tilde{n}) e^{-j(2\pi nk/N)} \text{ for } \tilde{n} \in (n_i, n_{i+1}), \quad (2)$$

with $\bar{x}(n) = x(n) \cdot \Psi^R(n, Q) \cdot \Psi^S(n, \tilde{n}, Q)$. $\mathbf{T}$ represents the time-frequency matrix, where $\Psi^R(n, Q)$ and $\Psi^S(n, \tilde{n}, Q)$ are the frequency-rotate and frequency-shift operators, respectively. The window function is given by [49, 50],

$$w_\sigma(n) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\left(n^2/2\sigma^2\right)} \quad (3)$$

The frequency-rotate operator is expressed as in (4), and the frequency shift operator is as in (5) [51]:

$$\Psi^R(n, Q) = e^{\left(-j \sum_{l=1}^{L} q_l^i (n - n_i)^l + \gamma_i\right)}, \quad (4)$$

$$\Psi^S(n, \tilde{n}, Q) = e^{\left(j \sum_{l=1}^{L} q_l^i (\tilde{n} - n_i)^{l-1} n\right)}, \quad (5)$$

where $Q(i, l) = q_l^i$ represents the local polynomial coefficient matrix for the spline kernel. The parameter L is denoted as the order of the spline. The parameter γ_i in SCT should satisfy the following conditions as [51],

$$\gamma_i - \gamma_{i+1} = \sum_{l=1}^{L} \frac{q_l^{i+1}}{l} (n - n_i)^l, \quad (6)$$

with initial value $\gamma_1 = 0$. The factor $i = 1, 2, \cdots, I$ is the ith piece, where the spline is defined in a piecewise polynomial form and I is the total number of pieces [51]. For a pathological PCG signal, we have compared the time-frequency representations that are obtained using both CT and SCT methods. The AS pathological PCG recording is shown in

Figure 2(a). The time-frequency contour plots of pathological PCG signal computed using CT and SCT are shown in Figure 2(b) and Figure 2(c), respectively. It can be observed from the figure that the time-frequency plot obtained using CT has an energy distribution between 25 Hz and 300 Hz. However, the murmurs are high-frequency sounds produced during the recording of the PCG signal [66, 67]. It is clearly observed from the time-frequency plot of the PCG recording obtained using SCT that the murmur energies are distributed between 100 Hz and 780 Hz. This shows that the information regarding the murmurs is not effectively captured in the CT-based time-frequency representation and the SCT provides better time-frequency localization for PCG recording as compared to CT.

The PCG signals for normal (N) and pathological classes such as MR, MS, AS, and MVP are depicted in Figure 3(a), Figure 3(c), Figure 3(e), Figure 3(g), and Figure 3(i), respectively, and the time-frequency plots for these signals were obtained using SCT are shown in Figure 3(b), Figure 3(d), Figure 3(f), Figure 3(h), and Figure 3(j), respectively. It can be observed that the pattern associated with the pathological PCG signal has different morphology for each type of HVA as compared to the normal PCG signal. The energies in the S1 and S2 components of the normal PCG signals are grossly distributed from 25 Hz to 300 Hz (as shown in Figure 3(b)). However, during HVA, the energy is distributed above 300 Hz in the time-frequency plot of the PCG signal. Each frequency component in the time-frequency matrix of the PCG recording has different characteristics for normal and pathological PCG signals. Therefore, the features computed from each frequency component of the PCG recording in the time-frequency domain will be helpful for the accurate detection of HVAs. In this study, we have extracted three types of nonlinear features, namely, L1-norm, sample entropy, and permutation entropy from the first 400 frequency atoms or components of the time-frequency representation of the PCG recording. The L1-norm (LN) features for the kth frequency component is evaluated as [68]

$$LN^k = \sum_{\tilde{n}=1}^{N} |\mathbf{T}(\tilde{n}, k)|. \quad (7)$$

Moreover, we have also evaluated the sample entropy (SEN) [69] and permutation entropy (PEN) [70] features from the kth frequency atom of the matrix $\mathbf{T}$. The features are denoted as SEN^k and PEN^k. A 1200-dimensional feature vector based on the combination of 400 LN, 400 SEN, and 400 PEN features is formulated for each PCG recording obtained from the database and 15 recorded PCG signals. The KSRC classifier is used to detect HVAs from the 1200-dimensional feature vector. In the following subsection, the descriptions of DLKSRN for the classification of HVAs are presented.

2.3. Deep Layer KSRC. In this work, the DLKSRN is proposed for the classification of HVAs using PCG signal features. The architecture of DLKSRN is shown in Figure 4. It consists of an input layer, first ELM-autoencoder hidden layer, second ELM-autoencoder hidden layer, and an output layer. In this
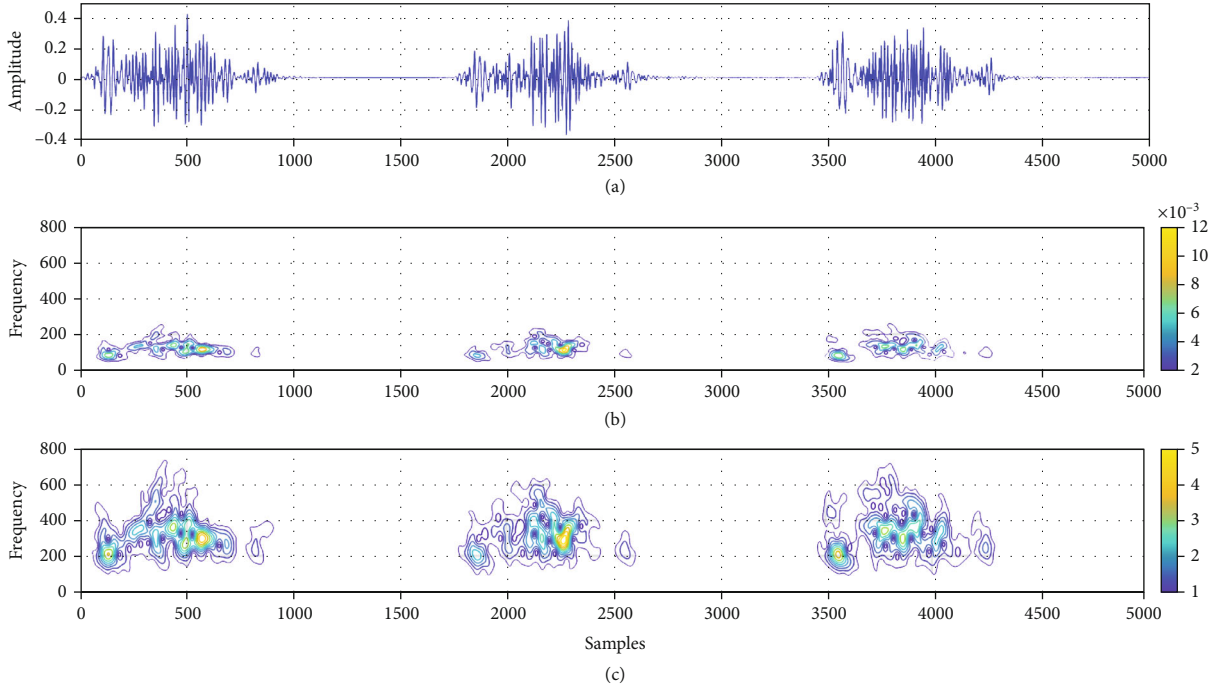
FIGURE 2: (a) Pathological PCG signal (AS pathology with murmurs present between S1-component and S2-component of each cardiac cycle). (b) Time-frequency representation of the pathological PCG signal obtained using CT. (c) Time-frequency representation of the pathological PCG signal obtained using SCT.

work, the hold-out and 10-fold cross-validation (CV) techniques are used to select the training and test PCG recordings. The feature matrix which comprises of the training feature vectors of the PCG recordings and the class labels are given as, $\{z_i, y_i\}_{i=1}^{m}$ with, $z_i \in \mathbb{R}^p$ and $y_i \in 1, 2, 3, 4, 5$, where 1, 2, 3, 4, and 5 are class label representations for normal, MVP, AS, MR, and MS classes. p is the size of the feature vector obtained from each PCG recording, and m is the number of PCG recordings considered during the training of the DLKSRN. The hidden layer matrix in DLKSRN is evaluated by solving the following optimization problem as,

$$J = \min_{\mathbf{W}_i} \left[\frac{\gamma}{2} \left\| \mathbf{H}_i \mathbf{W}_i - \tilde{\mathbf{Z}} \right\|_2^2 + \|\mathbf{W}_i\|_1 \right], \qquad (8)$$

where $\mathbf{W}_i$ is the ith hidden layer weight matrix and $\tilde{\mathbf{Z}}$ is the input feature matrix for the ELM-autoencoder. For first ELM-AE, $\tilde{\mathbf{Z}}$ is the feature matrix ($\mathbf{Z}$) containing PCG instances and time-frequency-based features. Similarly, for the second AE, the feature matrix ($\tilde{\mathbf{Z}}$) is the hidden layer matrix ($\mathbf{H}_1$) obtained from the first ELM-autoencoder. The weight matrix evaluation for each ELM-autoencoder is given by,

$$\mathbf{W}_i = \left(\mathbf{H}_i^{\mathrm{T}} \mathbf{H}_i + \frac{\mathbf{I}}{\gamma} \right)^{-1} \mathbf{H}_i^{\mathrm{T}} \tilde{\mathbf{Z}}. \qquad (9)$$

The feature matrix obtained in the second hidden layer of ELM-autoencoder is given as follows:

$$\mathbf{Z}^* = f(f(\mathbf{W}_1 \mathbf{Z}) \mathbf{W}_2). \qquad (10)$$

The new feature matrix, $\mathbf{Z}^*$, is used as the input to the KSRC layer of the proposed DLKSRN for the classification. KSRC is a kernel-based sparse representation technique, and it does not require rigorous training like deep neural networks (DNNs) to evaluate the class labels of the test feature vectors [60, 61]. It consists of four steps to estimate the class label of test PCG feature vectors. These steps are (i) mapping of the feature vectors of PCG signal into higher dimension space using kernel function, (ii) use of kernel-based dimension reduction for feature reduction, (iii) evaluation of coefficient vector and residual to test PCG feature vector by solving L1-norm optimization problem, and (iv) assignment of the class label to test PCG vector based on finding the minimum distance for all classes [60, 61]. The SRC has less performance when the feature vectors are not linearly separable. To overcome this limitation, KSRC maps the input PCG feature vector to a higher dimension space and performs the SRC in that new space.

The mapping function $\Psi(\mathbf{z}^*)$ projects each training feature vector to a higher dimensional space, and it is given as $\Psi(\mathbf{z}^*) = [\Psi_1(\mathbf{z}^*), \Psi_2(\mathbf{z}^*), \cdots \cdots \Psi_r(\mathbf{z}^*)]^T$, where $\Psi(\mathbf{z}^*) \in \mathbb{R}^r$ with $r \gg p$ is the dimension of new feature space or higher-dimension space. In new feature space, we can represent the mapped feature vector, $\Psi(\mathbf{z}_t^*)$, as the linear combination of the mapped training feature vectors of the PCG recordings, and it is given by $\Psi(\mathbf{z}_t^*) = \sum_{i=1}^{p} \gamma_i \Psi(\mathbf{z}_i^*) = \mathbf{\Psi}\gamma$, where $\gamma = [\gamma_1, \gamma_2, \cdots \cdots \gamma_m]^T$ is the coefficient vector, and it can be evaluated based on the solution of the following optimization problem as [60, 61]

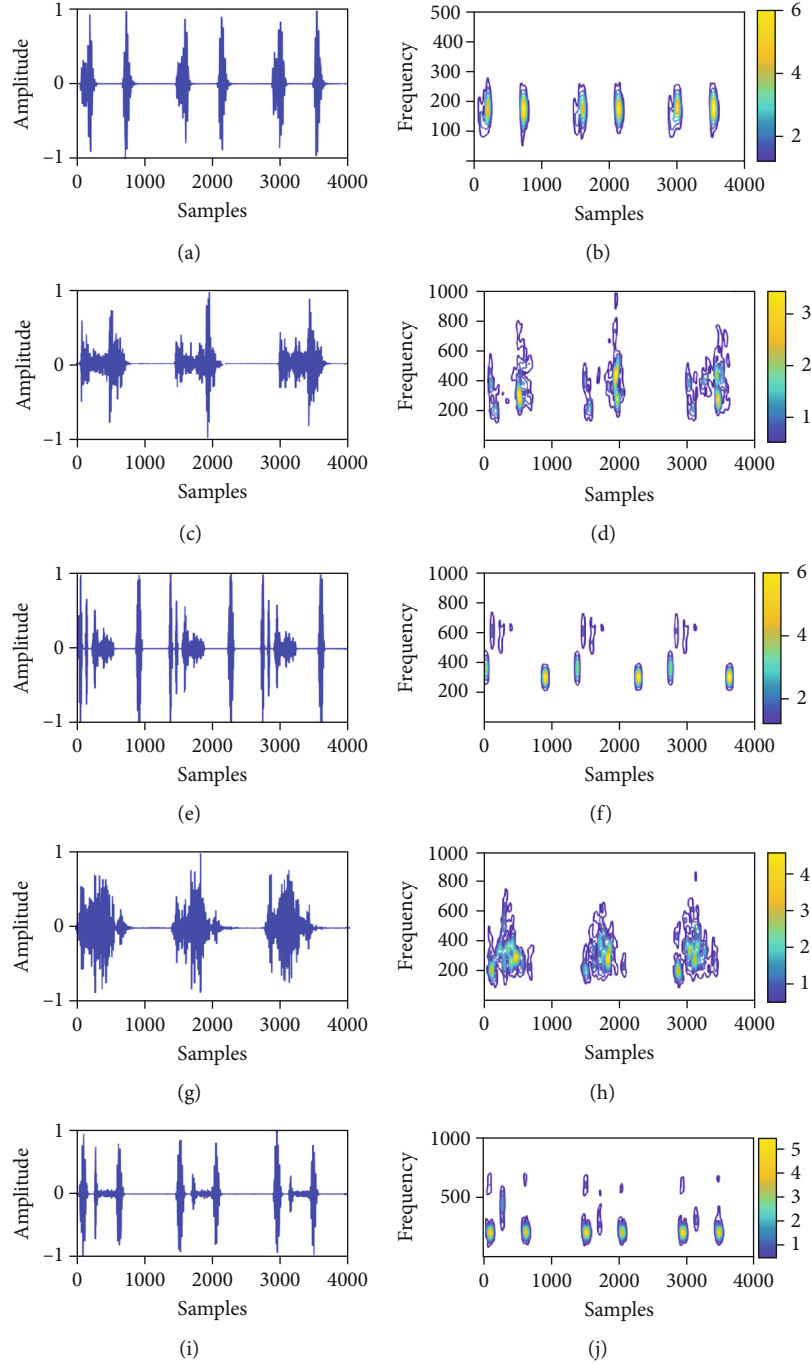$$\min_{\gamma} \|\gamma\|_1, \qquad (11)$$

FIGURE 3: (a) Typical PCG recording of the normal class. (b) Time-frequency representation of normal PCG signal obtained using SCT. (c) PCG recording of the MR class. (d) Time-frequency representation of MR PCG signal obtained using SCT. (e) PCG recording of the MS class. (f) Time-frequency representation of MS PCG signal obtained using SCT. (g) PCG recording of the AS class. (h) Time-frequency representation of AS PCG signal obtained using SCT. (i) PCG recording of the MVP class. (j) Time-frequency representation of MVP PCG signal obtained using SCT.

subjected to $\Psi(\mathbf{z}_t^*) = \mathbf{\Psi}\gamma$. In KSRC, the dimension of kernel space r is higher as compared to the second hidden layer space $\tilde{p}$, and also, it can be higher than the number of training instances m. Therefore, for getting a sparse solution of γ in (11), the dimension reduction step is used in the kernel space. The dimension reduction is performed based on the use of the transformation matrix $\mathbf{A}$. The constraint for the optimization problem in (11) is modified as follows:

$$\mathbf{A}^T\Psi(\mathbf{z}_t^*) = \mathbf{A}^T\mathbf{\Psi}\gamma, \qquad (12)$$
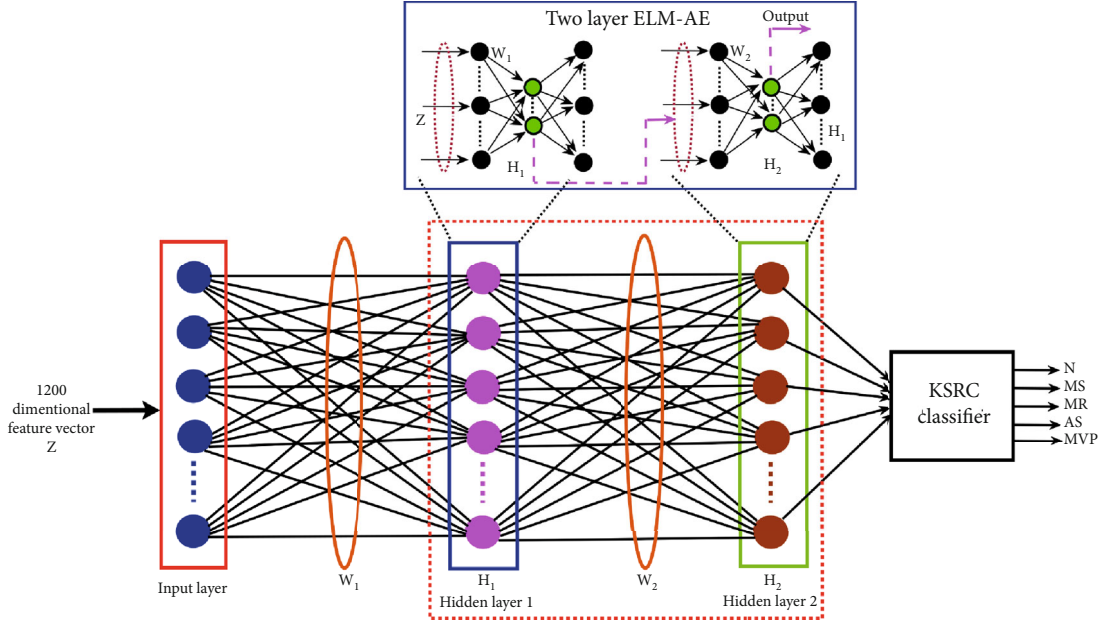
FIGURE 4: Architecture of the proposed DLKSRN for the classification of HVAs.

where the transformation matrix can be evaluated as follows:

$$\mathbf{A} = \mathbf{\Psi S}. \tag{13}$$

The matrix $\mathbf{S}$ is the pseudotransformation matrix, and it is evaluated using any one of the dimension reduction techniques (random projection, kernel principal component analysis (KPCA), and kernel linear discriminant analysis (KLDA)) [60, 61]. The expression in (12) can be simplified as follows:

$$\mathbf{\Psi}^T \Psi(\mathbf{z}_t^*)\mathbf{S}^T = \mathbf{\Psi}^T \mathbf{\Psi}\gamma\mathbf{S}. \tag{14}$$

The above equation can also be written as, $\mathbf{S}^T k(\mathbf{z}_t^*, \mathbf{z}_i^*) = \mathbf{K}\gamma\mathbf{S}$, where $k(\mathbf{z}_t^*, \mathbf{z}_i^*)$ and $\mathbf{K}$ are the kernel function and the kernel matrix, respectively. The original optimization problem in KSRC is modified as follows:

$$\min_{\gamma}\|\gamma\|_1, \tag{15}$$

subject to $\mathbf{S}^T k(\mathbf{z}_t^*, \mathbf{z}_t) = \mathbf{S}^T \mathbf{K}\gamma$. The residual for the test instance $\mathbf{z}_t^*$ for the cth class is obtained as follows [60]:

$$\mathrm{rs}_c(\mathbf{z}_t^*) = \left\|\mathbf{S}^T k(\mathbf{z}_t^*, \mathbf{z}_i^*) - \mathbf{S}^T \mathbf{K}\delta_c\right\|_2, \tag{16}$$

where $\delta_c = [\delta_c(\gamma_1), \delta_c(\gamma_2), \cdots, \delta_c(\gamma_m)]$, and $\delta_c(\gamma_i)$ is the characteristic function for the cth class. This function is defined as follows [60, 61]:

$$\delta_c(\gamma_i) = \begin{Bmatrix} \gamma_i & \text{if } y_i = c \\ 0 & \text{elsewhere} \end{Bmatrix}. \tag{17}$$

The residual for each class is computed, and the final class label for the second hidden layer feature vector of test PCG recording is given by

$$\mathbf{y}_t = \arg \min_{c=1,2,\cdots C} \mathrm{rs}_c(\mathbf{z}_t^*). \tag{18}$$

In this study, the number of neurons used in the first and second hidden layers of the proposed DLKSRN is 800 and 600, respectively. Moreover, we have also considered the random forest (RF) [36] and K-nearest neighbour (KNN) [65] classifiers for the classification of HVAs from the feature vectors of test PCG recordings. The optimal parameters of the RF classifier [71] such as the number of trees, number of splits for each decision tree, and depth of each decision tree obtained using the grid-search technique are 20, 20, and 15, respectively. For the KNN classifier, we have considered the number of the nearest neighbours as 3 and used Euclidean as the distance metric [72]. The performance of the 1200 dimensional SCT-based time-frequency features of PCG recordings is evaluated using DLKSRN, KSRC, RF, and KNN classifiers with the hold-out cross-validation (CV) strategy. For hold-out CV, 60%, 10%, and 30% of PCG signal instances are considered for training, validation, and testing of the DLKSRN classifier. Similarly, for the 10-fold CV case, 90% of PCG signal instances from the feature matrix are used to train the DLKSRN classifier. The remaining 10% PCG signal instances are evaluated during the testing phase of the DLKSRN classifier in each fold. The metrics, namely, the sensitivity, specificity, precision, F-score, and overall accuracy (OA), are used to evaluate the performance of DLKSRN, KSRC, RF, and KNN classifiers [72]. In the following section, the results obtained using the proposed approach are discussed in detail.
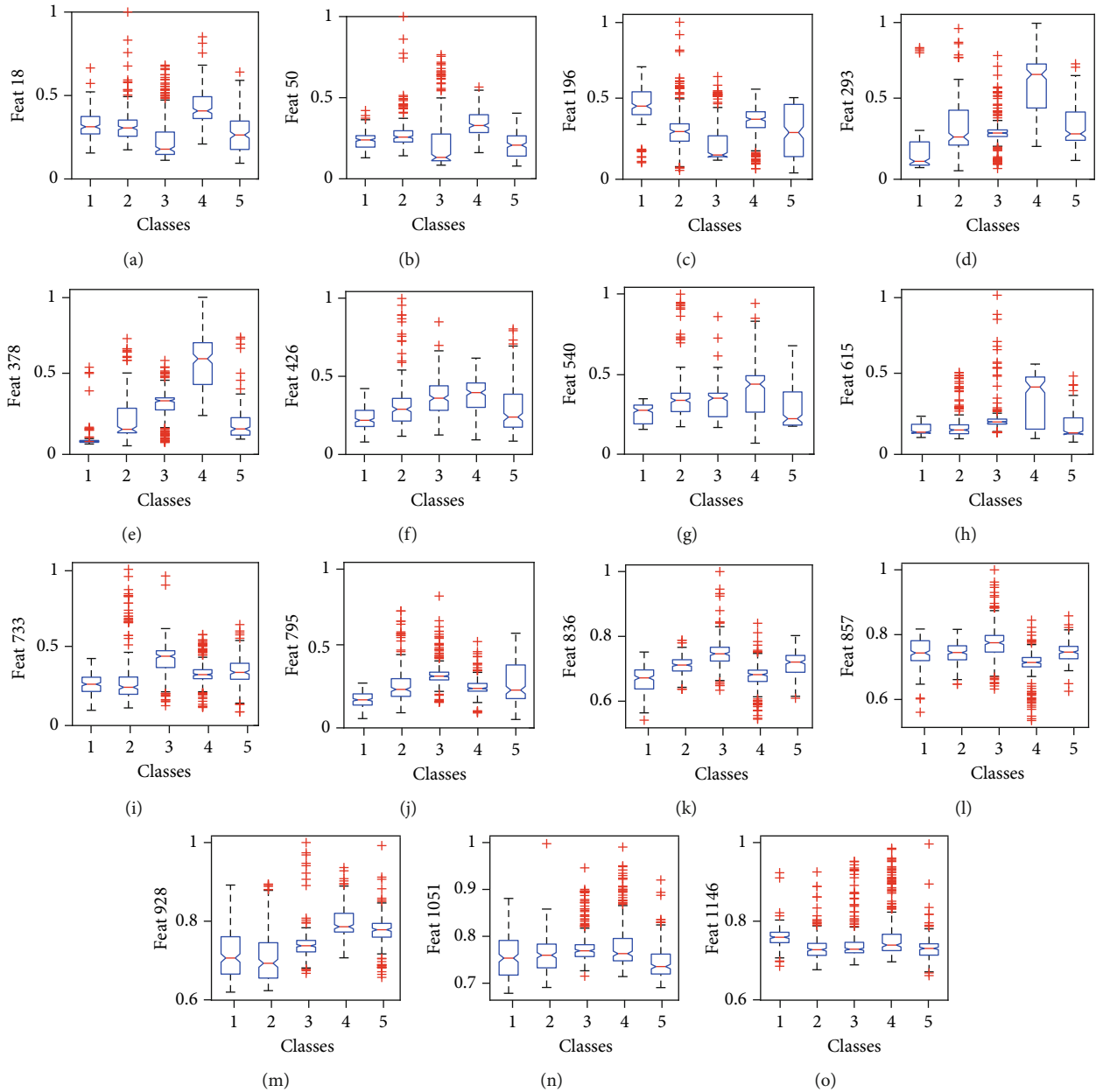
FIGURE 5: (a) Boxplot of 18th L1 norm feature or attribute (Feat 18) for all classes. (b) Boxplot of 50th L1 norm attribute (Feat 50) for all classes. (c) Boxplot of 196th L1 norm attribute (Feat 196) for all classes. (d) Boxplot of the 293rd L1 norm attribute (Feat 293) for all classes. (e) Boxplot of the 378th L1 norm attribute (Feat 378) for all classes. (f) Boxplot of the 26th SENT attribute (Feat 26) for all classes. (g) Boxplot of 140th SENT attribute (Feat 140) for all classes. (h) Boxplot of the 215th SENT attribute (Feat 215) for all classes. (i) Boxplot of the 333th SENT attribute (Feat 333) for all classes. (j) Boxplot of the 395th SENT attribute (Feat 395) for all classes. (k) Boxplot of the 36th PENT attribute (Feat 36) for all classes. (l) Boxplot of the 57th PENT attribute (Feat 57) for all classes. (m) Boxplot of the 128th PENT attribute (Feat 128) for all classes. (n) Boxplot of the 251st PENT attribute (Feat 251) for all classes. (o) Boxplot of the 346th PENT attribute (Feat 346) for all classes.

3. Results and Discussion

In the first part of this section, the statistical analysis results of SCT-based features of PCG recordings are presented. In the second part, the classification results using RF, KNN, KSRC, and the proposed DLKSRN models are shown. The third part of this section describes the comparison and advantages of the proposed approach for HVA detection. In this study, we have conducted a statistical analysis of all

1200 SCT-based features of the PCG recording. The results are shown for 15 different features out of 1200 features. The intraclass variations of the LN features for 18th, 50th, 196th, 293th, and 378th frequency components for all N, MVP, AS, MR, and MS categories are depicted in Figures 5(a)–5(e), respectively. Similarly, the within-class variations of the SEN features for the 26th, 140th, 250th, 333th, and 395th frequency components for all classes are shown in Figures 5(f)–5(j), respectively. Moreover, in Figures 5(k)–5(o), we have shown

TABLE 1: Mean ± SD of the selected features extracted from the time-frequency matrix for the five pathological classes.

Feature	Feature number	N	MS	MR	AS	MVP
	Feat 18	0.3212 ± 0.0796	0.3231 ± 0.1094	0.2446 ± 0.1478	0.4286 ± 0.1009	0.2699 ± 0.1105
	Feat 50	0.2374 ± 0.0530	0.2802 ± 0.1110	0.2179 ± 0.1701	0.3402 ± 0.0793	0.2086 ± 0.0789
L1-norm	Feat 196	0.4748 ± 0.1212	0.3154 ± 0.1460	0.2217 ± 0.1392	0.3597 ± 0.1128	0.2995 ± 0.1706
	Feat 293	0.1498 ± 0.1256	0.3104 ± 0.1690	0.2774 ± 0.1182	0.6083 ± 0.2066	0.3155 ± 0.1425
	Feat 378	0.0525 ± 0.0637	0.1898 ± 0.1380	0.2782 ± 0.1184	0.5697 ± 0.1914	0.1652 ± 0.1186
	Feat 426	0.2315 ± 0.0693	0.3213 ± 0.1580	0.3747 ± 0.1238	0.3820 ± 0.1034	0.2906 ± 0.1501
	Feat 540	0.2467 ± 0.0591	0.3490 ± 0.1531	0.3242 ± 0.0965	0.3919 ± 0.1787	0.2908 ± 0.1318
SEN	Feat 615	0.1600 ± 0.0325	0.1864 ± 0.0908	0.2342 ± 0.1238	0.3418 ± 0.1588	0.1852 ± 0.0766
	Feat 733	0.2613 ± 0.0724	0.3046 ± 0.1837	0.4197 ± 0.1099	0.3354 ± 0.0858	0.3437 ± 0.0961
	Feat 795	0.2175 ± 0.0536	0.3343 ± 0.1465	0.4140 ± 0.1088	0.3162 ± 0.0715	0.3490 ± 0.1617
	Feat 836	0.6661 ± 0.0408	0.7094 ± 0.0279	0.7490 ± 0.0495	0.6783 ± 0.0431	0.7142 ± 0.0390
	Feat 857	0.7657 ± 0.0378	0.7635 ± 0.0276	0.7927 ± 0.0521	0.7302 ± 0.0429	0.7658 ± 0.0285
PEN	Feat 928	0.7174 ± 0.0613	0.7040 ± 0.0595	0.7429 ± 0.0480	0.7964 ± 0.0401	0.7751 ± 0.0431
	Feat 1051	0.7543 ± 0.0430	0.7628 ± 0.0387	0.7815 ± 0.0417	0.7849 ± 0.0553	0.7469 ± 0.0385
	Feat 1146	0.7610 ± 0.0270	0.7363 ± 0.0368	0.7490 ± 0.0550	0.7675 ± 0.0692	0.7319 ± 0.0353

TABLE 2: Classification results obtained for automated detection of HVAs using various classifiers with SCT domain features and hold-out CV.

Classifier	Class					Performance measure				OA (%)
		TP	TN	FP	FN	Precision (%)	Sensitivity (%)	Specificity (%)	F-score (%)	
	N	60	227	0	0	97.75 ± 2.63	98.33 ± 2.35	99.38 ± 0.74	98.00 ± 1.50	
	MS	60	227	6	0	88.60 ± 3.77	96.99 ± 1.39	96.72 ± 1.30	92.56 ± 1.90	
RF	MR	49	238	1	11	96.31 ± 1.87	87.99 ± 4.62	99.13 ± 0.42	91.94 ± 3.20	95.66
	AS	59	228	1	1	92.38 ± 3.59	95.33 ± 1.82	97.90 ± 1.04	93.79 ± 1.63	
	MVP	59	228	5	1	98.22 ± 1.24	93.32 ± 4.56	99.56 ± 0.31	95.67 ± 2.79	
	N	60	231	0	0	98.69 ± 1.35	99.66 ± 0.74	99.64 ± 0.37	99.17 ± 0.82	
	MS	57	234	3	3	90.16 ± 1.65	97.66 ± 0.91	97.28 ± 0.49	93.76 ± 1.28	
KNN	MR	59	232	2	1	97.91 ± 1.46	93.99 ± 3.45	99.48 ± 0.35	95.89 ± 2.19	97.00
	AS	58	233	3	2	98.24 ± 2.11	91.99 ± 3.20	99.57 ± 0.52	94.99 ± 2.20	
	MVP	57	234	1	3	96.41 ± 2.56	97.33 ± 1.90	99.04 ± 0.71	96.85 ± 1.76	
	N	60	236	0	0	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00	
	MS	59	237	1	1	96.13 ± 2.12	98.99 ± 0.91	98.98 ± 0.56	97.54 ± 1.52	
KSRC	MR	60	236	0	0	98.66 ± 2.17	97.66 ± 1.90	99.66 ± 0.55	98.15 ± 1.89	98.66
	AS	58	238	0	2	99.00 ± 1.49	96.99 ± 2.47	99.74 ± 0.38	97.96 ± 1.29	
	MVP	59	237	3	0	96.70 ± 1.99	96.66 ± 1.17	99.15 ± 0.51	96.66 ± 1.01	
	N	60	238	0	0	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00	
	MS	59	239	1	1	98.24 ± 1.02	99.01 ± 0.10	99.00 ± 0.63	99.13 ± 0.14	
DLKSRN	MR	60	238	1	0	99.02 ± 0.23	98.76 ± 1.25	99.88 ± 0.16	99.00 ± 0.26	99.23
	AS	60	238	0	1	99.18 ± 0.64	97.95 ± 1.38	99.82 ± 0.01	98.72 ± 1.33	
	MVP	59	239	1	1	96.89 ± 1.02	97.36 ± 1.29	99.56 ± 0.11	97.51 ± 1.82	

the intraclass variations of PEN features for 36th, 57th, 128th, 251st, and 346th frequency components in all classes. The parameters such as mean and standard deviation values of features whose intraclass variations given in boxplots of Figure 3 are shown in Table 1. It is noted that each feature has distinct mean values for each of the pathological classes

Confusion matrices (Actual class columns: N, MS, MR, AS, MVP; Predicted class rows)

(a) RF classifier

	N	MS	MR	AS	MVP
N	60	0	0	0	0
MS	0	60	0	0	0
MR	0	5	49	1	5
AS	0	1	0	59	0
MVP	0	0	1	0	59

(b) KNN classifier

	N	MS	MR	AS	MVP
N	60	0	0	0	0
MS	0	57	0	2	1
MR	0	1	59	0	0
AS	0	2	0	58	0
MVP	0	0	2	1	57

(c) KSRC

	N	MS	MR	AS	MVP
N	60	0	0	0	0
MS	0	59	0	0	1
MR	0	0	60	0	0
AS	0	0	0	58	2
MVP	0	1	0	0	59

(d) DLKSRN

	N	MS	MR	AS	MVP
N	60	0	0	0	0
MS	0	59	0	0	1
MR	0	0	60	0	0
AS	0	1	0	60	0
MVP	0	0	1	0	59

FIGURE 6: Confusion matrix of classifier obtained using 1200 dimensional feature vectors with PCG signals: (a) RF classifier; (b) KNN classifier; (c) KSRC; and (d) DLKSRN.

TABLE 3: Results of the classification using the DLKSRN classifier with ten-fold CV.

HVDs	Measures (%)	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10	Average
N	Precision	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0 ± 0.00
	Sensitivity	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0 ± 0.00
	Specificity	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0 ± 0.00
	F-score	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0 ± 0.00
MS	Precision	100.0	90.90	100.0	95.23	100.0	95.23	95.23	100.0	100.0	100.0	97.65 ± 3.27
	Sensitivity	100.0	100.0	95.00	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.50 ± 1.58
	Specificity	98.75	98.75	100.0	98.73	100.0	100.0	100.0	98.75	100.0	98.73	99.37 ± 0.66
	F-score	100.0	95.23	100.0	97.56	100.0	97.56	100.0	97.43	97.56	100.0	98.53 ± 1.69
MR	Precision	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0 ± 0.00
	Sensitivity	100.0	95.00	90.00	95.00	100.0	100.0	100.0	100.0	100.0	100.0	98.00 ± 3.49
	Specificity	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0 ± 0.00
	F-score	100.0	97.43	100.0	97.43	100.0	100.0	100.0	100.0	94.73	100.0	98.95 ± 1.83
AS	Precision	100.0	100.0	95.23	100.0	100.0	95.00	100.0	100.0	100.0	100.0	99.02 ± 2.06
	Sensitivity	100.0	100.0	100.0	100.0	95.00	100.0	100.0	100.0	95.00	100.0	99.00 ± 2.10
	Specificity	100.0	100.0	100.0	98.73	100.0	100.0	100.0	97.46	100.0	100.0	98.98 ± 1.80
	F-score	94.73	100.0	97.56	100.0	100.0	100.0	100.0	100.0	100.0	97.56	98.98 ± 1.80
MVP	Precision	95.23	100.0	100.0	95.00	100.0	95.23	100.0	95.23	100.0	100.0	98.06 ± 2.49
	Sensitivity	100.0	100.0	95.00	100.0	100.0	100.0	95.00	95.00	100.0	100.0	98.50 ± 2.41
	Specificity	100.0	100.0	98.75	100.0	97.50	100.0	100.0	100.0	100.0	98.73	98.00 ± 1.54
	F-score	95.23	97.43	97.43	100.0	100.0	97.43	100.0	97.56	97.56	97.43	98.00 ± 1.54
All	OA	100.0	99.33	99.04	100.0	100.0	98.05	98.11	99.25	99.86	98.84	99.24 ± 0.74

(AS, MS, MR, MVP) and normal class. The SEN feature for more than 300 frequency components of the SCT-based time-frequency matrix has a lower mean value for the normal class as compared to the pathological classes. Similarly, more than 230 PEN features have lower mean values for the normal class, and more than 200 L1-norm features have higher mean values for the AS class. The pathological signature for MS is the presence of diastolic murmurs [73], and murmurs are observed between the systolic interval of PCG recording in MVP pathology [74]. In MS and AS pathologies, the murmurs have low-pitch sounds. Similarly, the high-pitch sounds are observed in the PCG recording during AR-based HVA [14]. The aforementioned pathological changes on the PCG recording affect the morphologies of the SCT-based time-frequency matrices. Hence, the features from the time-frequency matrices have distinct mean and standard deviation values. We have also used the analysis of variance- (ANOVA-) based test [75] to verify the statistical significance of SCT-based time-frequency features. It is observed from the ANOVA test that all 1200 features extracted from the SCT-based time-frequency representation of PCG recording have p values less than 0.001 and is significant for the detection of HVAs.

The classification results of RF, KNN, KSRC, and DLKSRN are shown in Table 2. In this work, we have considered five random trials based on the hold-out CV to evaluate the performance of each classifier. The performance metrics
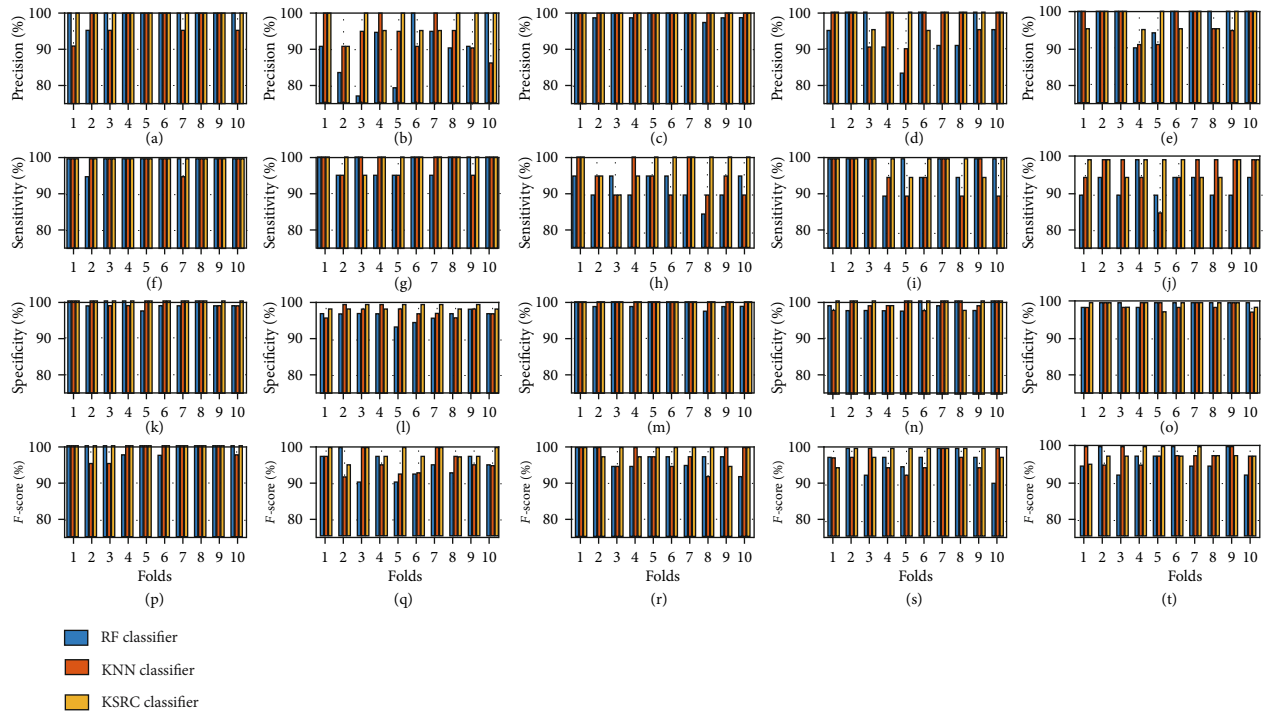
FIGURE 7: Precision values of the KSRC, RF, and KNN classifiers obtained using the 10th fold cross-validation for (a) normal class, (b) MS class, (c) MR class, (d) AS class, and (e) MVP class. Sensitivity values of the KSRC, RF, and KNN classifiers obtained using the 10th fold cross-validation for (f) normal class, (g) MS class, (h) MR class, (i) AS class, and (j) MVP class. Specificity values of the KSRC, RF, and KNN classifiers obtained using the 10th fold cross-validation for (k) normal class, (l) MS class, (m) MR class, (n) AS class, and (o) MVP class. *F*-score values of the KSRC, RF, and KNN classifiers obtained using the 10th fold cross-validation for (p) normal class, (q) MS class, (r) MR class, (s) AS class, and (t) MVP class.

TABLE 4: Overall accuracy values obtained using the DLKSRN classifier for various number of neurons in the 1st and 2nd hidden layer of validation and test sets for N vs. MS vs. MR vs. AS vs. MVP classification scheme.

Number of neurons		Overall accuracy (%)	
1st hidden layer	2nd hidden layer	Validation set	Test set
200	100	94.60	95.86
400	200	96.00	96.26
600	400	96.20	96.66
800	600	97.80	99.23
1000	800	95.80	97.33

are shown in the mean and standard deviation format in Table 2. Also, we have shown the confusion matrices of RF, KNN, KSRC, and DLKSRN classification models in Figures 6(a)–6(d), respectively. It is evident that the average number of true positives is high using DLKSRN as compared to the RF, KNN, and KSRC classifiers. The number of true positive (TP), true negative (TN), false positive (FP), and false-negative (FN) values is listed in Table 2 for normal and other pathological classes. It can be observed that DLKSRN has less number of average FN and FP values for each class. The values of the metrics such as precision, sensitivity, specificity, and *F*-score are also high using DLKSRN as compared to other classification approaches. It is also observed that the average accuracy of DLKSRN is higher than

KSRC, RF, and KNN classifiers. The DLKSRN classification results for 10-fold CV are shown in Table 3. The classification results of KSRC, RF, and KNN classifiers with 10-fold CV are depicted in Figure 7. The precision values for normal, MS, MR, AS, and MVP classes over all 10 folds are depicted in Figures 7(a)–7(e). Similarly, the sensitivity values for all classes with ten-fold CV are shown in Figures 7(f)–7(j). In Figures 7(k)–7(o), we have shown the variation in specificity values for all ten folds for all classes. Moreover, the *F*-score values are depicted in Figures 7(p)–7(t) for normal, MS, MR, AS, and MVP classes. It is observed from the results that for MS class, the sensitivity is 95% after three folds and reached 100% for other folds. Similarly, the specificity values of DLKSRN for the first, second, fourth, eighth, and tenth fold are more than 98.5%. Similarly, for the MR class, the sensitivity values of KSRC for the second, third, and fourth folds are less than the remaining folds. For the MR class, the specificity values for all folds of both DLKSRN and KSRC are 100%. Similar variations are also observed for other pathological cases using the DLKSRN classifier. It is also observed that the sensitivity, specificity, precision, and *F*-score values of DLKSRN are higher for all folds as compared to RF classifiers for each class. In some folds, the KNN and KSRC have demonstrated similar sensitivity, specificity, and *F*-score values. The overall accuracy of DLKSRN for 10-fold CV using SCT-based time-frequency features is 99.24%. This value is higher than the average accuracy values of KSRC (98.90%), KNN (97.12%), and RF (96.12%) classifiers. It

TABLE 5: Classification results obtained for automated detection of HVAs using the DLKSRN classifiers with SCT domain features and hold-out CV.

Database used	Class	Precision (%)	Sensitivity (%)	Specificity (%)	F-score (%)	OA (%)
	N	98.07 ± 2.07	100 ± 0.00	99.48 ± 0.56	99.01 ± 1.06	
	MS	93.87 ± 2.98	95.99 ± 0.90	98.38 ± 0.80	94.90 ± 1.73	
MHSMD	MR	96.42 ± 2.35	96.66 ± 2.36	99.05 ± 0.63	96.50 ± 0.69	96.79 ± 0.18
	AS	98.30 ± 1.68	94.33 ± 1.90	99.57 ± 0.43	96.25 ± 0.77	
	MVP	97.23 ± 2.29	96.33 ± 2.17	99.22 ± 0.56	96.64 ± 1.01	

TABLE 6: Summary of automated detection of HVA developed using PCG signals using the same database.

Methods used for feature extraction	Classifiers	Classes	Accuracy (%)
Morphological features extracted from PCG recording [77]	SVM	N, MS, MR	91.23
Wavelet entropies as features from PCG [86]	ANFIS	N, PS, MS	98.33
Multilevel basis selection- (MLBS-) based wavelet features extracted from PCG [76]	SVM	N, AS, MR, AR	97.56
Entropy and energy fraction-based features [78]	SVM	N, TI, PS, MI, MS	97.17
Wavelet and MFCC features obtained from PCG [44]	SVM	N, AS, MS, MR, MVP	97.90
Magnitude and phase features extracted using SST of PCG [36]	Random forest	N, AS, MS, MR	95.13
Features extracted using CT of PCG [13]	Multiclass composite classifier	HC, AS, MS, MR	98.33
DNN [79]	WaveNet	N, MS, MR, AS, MVP	98.20
Proposed work (features evaluated in SCT domain of PCG)	DLKSRN	N, MS, MR, AS, MVP	99.24

can be noted that the nonlinear features extracted using SCT of PCG recording are able to classify the HVAs accurately using DLKSRN. In this study, the parameters of the DLKSRN classifier such as the number of neurons in the 1st and 2nd hidden layers are selected based on the maximum accuracy values in the validation and test sets. The variations in the overall accuracy values with hidden neurons in the 1st and 2nd hidden layers are shown in Table 4. It can be observed from the table that the overall accuracy value of the DLKSRN classifier is high when the number of neurons in the 1st and 2nd hidden layers are 800 and 600, respectively. The overall accuracy value decreases by increasing the number of neurons in both hidden layers. Similarly, for the MHSMD database, the classification results obtained using the DLKSRN classifier are shown in Table 5. It is observed that the proposed SCT-based time-frequency domain features combined with the DLKSRN classifier have obtained an overall accuracy value of 96.79%. The sensitivity and specificity values are greater than 94% for each class using the DLKSRN classifier. Moreover, we have tested the effectiveness of our proposed approach with 15 recorded PCG signals. The DLKSRN model which has been trained using the features from the public database has been used to test the performance of the private database. The LN, SEN, and PEN features are extracted from all 15 recorded PCG signals. The trained DLKSRN model successfully predicted all 15 feature vectors of PCG recordings as normal class thereby showing the effectiveness of the proposed approach for real-time precision of HVAs.

The objective of this study is the HVA detection using nonlinear features extracted from the SCT-based time-frequency analysis of PCG recording. The proposed features are found to be discriminative with the lowest p values obtained using the statistical test. The classification results obtained using the hold-out and 10-fold CV-based PCG instance selection reveal that the proposed approach has obtained an overall accuracy of more than 99% for the detection of HVA. A comparison with the existing algorithms for automated HVA detection is shown in Table 6. Safara et al. [76] developed the automated approach using wavelet packet decomposition-based feature extraction technique and SVM classifier for the discrimination of MR, AS, and AR-based HVAs with PCG recordings. They have achieved an accuracy of 97.56%. Maglogiannis et al. [77] used the SVM classifier coupled with the morphological features (standard values of S1 and S2 peaks and other features) for the detection of MR and MS pathologies and reported an accuracy of 91.23% in classifying two HVAs. Moreover, Zheng et al. [78] employed the energy fraction and energy-based features coupled with the SVM classifier for the automated discrimination of HVAs such as tricuspid insufficiency (TI), pulmonary stenosis (PS), mitral insufficiency (MI), and AS. They have obtained an overall accuracy of 97.17% in classifying the four HVAs. The time-frequency domain magnitude and phase features extracted using the SST of PCG signal have been used in [36] for the discrimination of AS, MS, and MR classes. They have obtained an overall accuracy value of 95.12% in classifying the three classes. The combination of MFCC- and DWT-based features extracted PCG signal, along with SVM classifier, has been used for automated HVA detection with an overall accuracy of 97.9% [44]. The CT-based time-frequency features obtained from PCG and composite

classification model yielded an overall accuracy value of 98.33% [13]. Oh et al. [79] have proposed a waveNet-based DNN model for the classification of HVAs using PCG recordings and obtained an overall accuracy value of 98.20%. The proposed approach demonstrated higher classification performance as compared to the existing algorithms for automated HVA detection. The method reported in [13] has classified AS, MS, and MR pathologies using PCG. However, in the present work, we have considered MVP pathology along with AS, MS, and MR ailments for the development of an automated HVA detection system. The advantages of the proposed HVA detection approach are given as follows:

(a) The SCT has demonstrated better time-frequency localization for both normal and pathological PCG signals as compared to CT

(b) The proposed approach used the nonlinear features from different frequency components of SCT-based time-frequency representation of the PCG signal

(c) The DLKSRN based on the ELM-autoencoder and KSRC is proposed for the classification of HVAs

(d) The proposed approach is tested using the recorded PCG signals

In this work, the local features from the frequency components of the time-frequency representation of the PCG signal are evaluated. The two-dimensional convolutional autoencoder [80] can be used for the extraction of learnable features from the SCT-based time-frequency representation of the PCG signal for the classification of HVAs. The sparse residual entropy features [81] and wavelet bispectrum-based features [82] can be used for the detection of HVAs from the PCG signal. The convolutional neural network [83], convolutional attention-based network [84], and other deep learning methodologies [85] can be used for the detection of HVAs without using extracted features from PCG recordings.

4. Conclusion

A novel HVA detection approach using PCG signals is proposed. This approach used SCT to compute the time-frequency representation of PCG recording. The nonlinear features (LN, SEN, and PEN) are computed from the frequency components of time-frequency representation. The DLKSRN classifier is used to discriminate automatically into four categories of HVA classes using the extracted features. The proposed approach demonstrated an average accuracy of 99.23% and 99.24% using hold-out and 10-fold CV methods. The proposed approach is also evaluated using the recorded signal, and the result obtained shows the practicality of the proposed approach. In the future, we intend to extend this method to detect coronary artery disease and psychological stress using PCG signals. The approach can also be implemented in real-time for IoMT applications.

Data Availability

The codes and the classification results of the proposed work are available upon request to the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] B. Iung and A. Vahanian, "Epidemiology of valvular disease in the adult," *Nature Reviews Cardiology*, vol. 8, no. 3, pp. 162–172, 2011.

[2] F. Tarasoutchi, M. W. Montera, M. Grinberg et al., "Diretriz Brasileira de Valvopatias - SBC 2011/ I Diretriz Interamericana de Valvopatias - SIAC 2011," *Arquivos Brasileiros de Cardiologia*, vol. 97, no. 5, Supplement 1, pp. 1–67, 2011.

[3] R. O. Bonow, B. Carabello, A. C. de Leon et al., "ACC/AHA guidelines for the management of patients with valvular heart disease. Executive summary. A report of the American College Of Cardiology/American Heart Association Task Force On Practice Guidelines (committee on management of patients with valvular heart disease)," *The Journal of Heart Valve Disease*, vol. 7, no. 6, pp. 672–707, 1998.

[4] M. Jessup, W. T. Abraham, D. E. Casey et al., "2009 focused update: Accf/aha guidelines for the diagnosis and management of heart failure in adults: a report of the american college of cardiology foundation/american heart association task force on practice guidelines: developed in collaboration with the international society for heart and lung transplantation," *Circulation*, vol. 119, no. 14, pp. 1977–2016, 2009.

[5] R. B. Hinton and K. E. Yutzey, "Heart valve structure and function in development and disease," *Annual Review of Physiology*, vol. 73, no. 1, pp. 29–46, 2011.

[6] J. T. Butcher, C. A. Simmons, J. N. Warnock et al., "Mechanobiology of the aortic heart valve," *Journal of Heart Valve Disease*, vol. 17, no. 1, p. 62, 2008.

[7] L. Chaothawee, "Diagnostic approach to assessment of valvular heart disease using MRI—part i: a practical approach for valvular regurgitation," *Heart Asia*, vol. 4, no. 1, pp. 38–43, 2012.

[8] H. W. Goo, I.-S. Park, J. K. Ko et al., "CT of congenital heart disease: normal anatomy and typical pathologic conditions," *Radiographics*, vol. 23, suppl_1, pp. S147–S165, 2003.

[9] M. Galderisi, B. Cosyns, T. Edvardsen et al., "Standardization of adult transthoracic echocardiography reporting in agreement with recent chamber quantification, diastolic function, and heart valve disease recommendations: an expert consensus document of the european association of cardiovascular imaging," *European Heart Journal - Cardiovascular Imaging*, vol. 18, no. 12, pp. 1301–1310, 2017.

[10] K. Savino and G. Ambrosio, "Handheld ultrasound and focused cardiovascular echography: use and information," *Medicina*, vol. 55, no. 8, p. 423, 2019.

[11] J. F. Glockner, D. L. Johnston, and K. P. McGee, "Evaluation of cardiac valvular disease with mr imaging: qualitative and quantitative techniques," *Radiographics*, vol. 23, no. 1, pp. e9–e9, 2003.

[12] F. M. Cáceres-Lóriga, H. Pérez-López, J. Santos-Gracia, and K. Morlans-Hernandez, "Prosthetic heart valve thrombosis:

pathogenesis, diagnosis and management," *International Journal of Cardiology*, vol. 110, no. 1, pp. 1–6, 2006.

[13] S. K. Ghosh, R. Ponnalagu, R. Tripathy, and U. R. Acharya, "Automated detection of heart valve diseases using chirplet transform and multiclass composite classifier with PCG signals," *Computers in Biology and Medicine*, vol. 118, p. 103632, 2020.

[14] V. N. Varghees and K. Ramachandran, "A novel heart sound activity detection framework for automated heart sound analysis," *Biomedical Signal Processing and Control*, vol. 13, pp. 174–188, 2014.

[15] M. Altuve, L. Suárez, and J. Ardila, "Fundamental heart sounds analysis using improved complete ensemble EMD with adaptive noise," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 1, pp. 426–439, 2020.

[16] O. el Badlaoui, A. Benba, and A. Hammouch, "Novel PCG Analysis Method for Discriminating between Abnormal and Normal Heart Sounds," *IRBM*, vol. 41, no. 4, pp. 223–228, 2020.

[17] D. B. Springer, L. Tarassenko, and G. D. Clifford, "Logistic regression-hsmm-based heart sound segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 822–832, 2016.

[18] H. Dubey, A. Monteiro, N. Constant et al., "Fog computing in medical internet-of-things: architecture, implementation, and applications," in *Handbook of Large-Scale Distributed Computing in Smart Healthcare*, pp. 281–321, Springer, 2017.

[19] O. Faust, E. J. Ciaccio, and U. R. Acharya, "A review of atrial fibrillation detection methods as a service," *International Journal of Environmental Research and Public Health*, vol. 17, no. 9, p. 3093, 2020.

[20] A. K. Dwivedi, S. A. Imtiaz, and E. Rodriguez-Villegas, "Algorithms for automatic analysis and classification of heart sounds-asystematic review," *IEEE Access*, vol. 7, pp. 8316–8345, 2018.

[21] S. K. Ghosh, P. R. Nagarajan, and R. K. Tripathy, "Heart sound data acquisition and preprocessing techniques: a review," in *Handbook of Research on Advancements of Artificial Intelligence in Healthcare Engineering*, pp. 244–264, IGI Global, 2020.

[22] C. Ahlstrom, P. Hult, P. Rask et al., "Feature extraction for systolic heart murmur classification," *Annals of Biomedical Engineering*, vol. 34, no. 11, pp. 1666–1677, 2006.

[23] P. Wang, C. S. Lim, S. Chauhan, J. Y. A. Foo, and V. Anantharaman, "Phonocardiographic signal analysis method using a modified hidden Markov model," *Annals of Biomedical Engineering*, vol. 35, no. 3, pp. 367–374, 2007.

[24] B. M. Whitaker, P. B. Suresha, C. Liu, G. D. Clifford, and D. V. Anderson, "Combining sparse coding and time-domain features for heart sound classification," *Physiological measurement*, vol. 38, no. 8, p. 1701, 2017.

[25] M. Zabihi, A. B. Rad, S. Kiranyaz, M. Gabbouj, and A. K. Katsaggelos, "Heart sound anomaly and quality detection using ensemble of neural networks without segmentation," in *in 2016 Computing in Cardiology Conference (CinC)*, pp. 613–616, Vancouver, BC, Canada, 2016, IEEE.

[26] C. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds," in *2016 Computing in Cardiology Conference (CinC)*, pp. 621–624, Vancouver, BC, Canada, 2016, IEEE.

[27] S. R. Bhatikar, C. DeGroff, and R. L. Mahajan, "A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics," *Artificial Intelligence in Medicine*, vol. 33, no. 3, pp. 251–260, 2005.

[28] R. J. Lehner and R. M. Rangayyan, "A three-channel microcomputer system for segmentation and characterization of the phonocardiogram," *IEEE Transactions on Biomedical Engineering*, vol. BME-34, no. 6, pp. 485–489, 1987.

[29] G. Amit, N. Gavriely, and N. Intrator, "Cluster analysis and classification of heart sounds," *Biomedical Signal Processing and Control*, vol. 4, no. 1, pp. 26–36, 2009.

[30] F. Beritelli, G. Capizzi, G. L. Sciuto, C. Napoli, and F. Scaglione, "Automatic heart activity diagnosis based on gram polynomials and probabilistic neural networks," *Biomedical Engineering Letters*, vol. 8, no. 1, pp. 77–85, 2018.

[31] C. D. Papadaniil and L. J. Hadjileontiadis, "Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features," *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 4, pp. 1138–1152, 2013.

[32] R. K. Tripathy, S. K. Ghosh, P. Gajbhiye, and U. R. Acharya, "Development of automated sleep stage classification system using multivariate projection-based fixed boundary empirical wavelet transform and entropy features extracted from multi-channel EEG signals," *Entropy*, vol. 22, no. 10, p. 1141, 2020.

[33] S. Patidar, R. B. Pachori, and N. Garg, "Automatic diagnosis of septal defects based on tunable- _Q_ wavelet transform of cardiac sound signals," *Expert Systems with Applications*, vol. 42, no. 7, pp. 3315–3326, 2015.

[34] Y. Soeta and Y. Bito, "Detection of features of prosthetic cardiac valve sound by spectrogram analysis," *Applied Acoustics*, vol. 89, pp. 28–33, 2015.

[35] E. Delgado-Trejos, A. Quiceno-Manrique, J. Godino-Llorente, M. Blanco-Velasco, and G. Castellanos-Dominguez, "Digital auscultation analysis for heart murmur detection," *Annals of Biomedical Engineering*, vol. 37, no. 2, pp. 337–353, 2009.

[36] S. K. Ghosh, R. K. Tripathy, R. Ponnalagu, and R. B. Pachori, "Automated detection of heart valve disorders from the PCG signal using time-frequency magnitude and phase features," *IEEE Sensors Letters*, vol. 3, no. 12, pp. 1–4, 2019.

[37] E. Kay and A. Agarwal, "Dropconnected neural networks trained on time-frequency and inter-beat features for classifying heart sounds," *Physiological measurement*, vol. 38, no. 8, pp. 1645–1657, 2017.

[38] S. Ari, K. Hembram, and G. Saha, "Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8019–8026, 2010.

[39] G. Redlarski, D. Gradolewski, and A. Palkowski, "A system for heart sounds classification," *PloS one*, vol. 9, no. 11, p. e112673, 2014.

[40] S. Choi, "Detection of valvular heart disorders using wavelet packet decomposition and support vector machine," *Expert Systems with Applications*, vol. 35, no. 4, pp. 1679–1687, 2008.

[41] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[42] T. Nilanon, J. Yao, J. Hao, S. Purushotham, and Y. Liu, "Normal/abnormal heart sound recordings classification using convolutional neural network," in *2016 Computing in Cardiology Conference (CinC)*, pp. 585–588, Vancouver, BC, Canada, 2016, IEEE.

[43] A. D. Ricke, R. J. Povinelli, and M. T. Johnson, "Automatic segmentation of heart sound signals using hidden Markov models," in *Computers in Cardiology, 2005*, pp. 953–956, Lyon, France, 2005, IEEE.

[44] Yaseen, G. Y. Son, and S. Kwon, "Classification of heart sound signal using multiple features," *Applied Sciences*, vol. 8, no. 12, p. 2344, 2018.

[45] J. Li, L. Ke, and Q. du, "Classification of heart sounds based on the wavelet fractal and twin support vector machine," *Entropy*, vol. 21, no. 5, p. 472, 2019.

[46] C. Li and M. Liang, "A generalized synchrosqueezing transform for enhancing signal time-frequency representation," *Signal Processing*, vol. 92, no. 9, pp. 2264–2274, 2012.

[47] J. Xu, L. Durand, and P. Pibarot, "Nonlinear transient chirp signal modeling of the aortic and pulmonary components of the second heart sound," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 10, pp. 1328–1335, 2000.

[48] Jingping Xu, L. G. Durand, and P. Pibarot, "Extraction of the aortic and pulmonary components of the second heart sound using a nonlinear transient chirp signal model," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 3, pp. 277–283, 2001.

[49] S. Mann and S. Haykin, "The chirplet transform: physical considerations," *IEEE Transactions on Signal Processing*, vol. 43, no. 11, pp. 2745–2761, 1995.

[50] Z. Peng, G. Meng, F. Chu, Z. Lang, W. Zhang, and Y. Yang, "Polynomial chirplet transform with application to instantaneous frequency estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 9, pp. 3222–3229, 2011.

[51] Y. Yang, Z. Peng, G. Meng, and W. Zhang, "Spline-kernelled chirplet transform for the analysis of signals with time-varying frequency and its application," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 3, pp. 1612–1621, 2012.

[52] B. Bozkurt, I. Germanakis, and Y. Stylianou, "A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection," *Computers in Biology and Medicine*, vol. 100, pp. 132–143, 2018.

[53] D. Pei, M. Burns, R. Chandramouli, and R. Vinjamuri, "Decoding asynchronous reaching in electroencephalography using stacked autoencoders," *IEEE Access*, vol. 6, pp. 52889–52898, 2018.

[54] S. K. Ghosh, R. K. Tripathy, M. R. A. Paternina, J. J. Arrieta, A. Zamora-Mendez, and G. R. Naik, "Detection of atrial fibrillation from single lead ECG signal using multirate cosine filter bank and deep neural network," *Journal of Medical Systems*, vol. 44, no. 6, p. 114, 2020.

[55] K. Sun, J. Zhang, C. Zhang, and J. Hu, "Generalized extreme learning machine autoencoder and a new deep neural network," *Neurocomputing*, vol. 230, pp. 374–381, 2017.

[56] J. V. Frances-Villora, A. Rosado-Muñoz, J. M. Martínez-Villena, M. Bataller-Mompean, J. F. Guerrero, and M. Wegrzyn, "Hardware implementation of real-time Extreme Learning Machine in FPGA: analysis of precision, resource occupation and performance," *Computers & Electrical Engineering*, vol. 51, pp. 139–156, 2016.

[57] R. Tripathy, M. R. A. PATERNINA, and P. Pattanaik, "A new method for automated detection of diabetes from heart rate signal," *Journal of Mechanics in Medicine and Biology*, vol. 17, no. 7, p. 1740001, 2017.

[58] B. Zhang, F. Karray, Q. Li, and L. Zhang, "Sparse representation classifier for microaneurysm detection and retinal blood vessel extraction," *Information Sciences*, vol. 200, pp. 78–90, 2012.

[59] R. K. Tripathy, M. R. A. Paternina, J. G. Arrieta, A. Zamora-Méndez, and G. R. Naik, "Automated detection of congestive heart failure from electrocardiogram signal using Stockwell transform and hybrid classification scheme," *Computer Methods and Programs in Biomedicine*, vol. 173, pp. 53–65, 2019.

[60] L. Zhang, W.-D. Zhou, P.-C. Chang et al., "Kernel sparse representation-based classifier," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1684–1695, 2011.

[61] J. Yin, Z. Liu, Z. Jin, and W. Yang, "Kernel sparse representation based classification," *Neurocomputing*, vol. 77, no. 1, pp. 120–128, 2012.

[62] Z. Chen, W. Zuo, Q. Hu, and L. Lin, "Kernel sparse representation for time series classification," *Information Sciences*, vol. 292, pp. 15–26, 2015.

[63] C. Liu, D. Springer, Q. Li et al., "An open access database for the evaluation of heart sound algorithms," *Physiological Measurement*, vol. 37, no. 12, pp. 2181–2213, 2016.

[64] A. Castro, T. T. Vinhoza, S. S. Mattos, and M. T. Coimbra, "Heart sound segmentation of pediatric auscultations using wavelet analysis," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3909–3912, Osaka, Japan, 2013, IEEE.

[65] L. Sharma, R. Tripathy, and S. Dandapat, "Multiscale energy and eigenspace approach to detection and localization of myocardial infarction," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 7, pp. 1827–1837, 2015.

[66] M. el-Segaier, O. Lilja, S. Lukkarinen, L. Sörnmo, R. Sepponen, and E. Pesonen, "Computer-based detection and analysis of heart sound and murmur," *Annals of Biomedical Engineering*, vol. 33, no. 7, pp. 937–942, 2005.

[67] B. El-Asir, L. Khadra, A. Al-Abbasi, and M. Mohammed, "Time-frequency analysis of heart sounds," in *Proceedings of Digital Processing Applications (TENCON'96)*, vol. 2, pp. 553–558, Perth, WA, Australia, Australia, 1996, IEEE.

[68] R. Tripathy and S. Dandapat, "Detection of myocardial infarction from vectorcardiogram using relevance vector machine," *Signal, Image and Video Processing*, vol. 11, no. 6, pp. 1139–1146, 2017.

[69] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[70] C. Bandt and B. Pompe, "Permutation entropy: a natural complexity measure for time series," *Physical review letters*, vol. 88, no. 17, p. 174102, 2002.

[71] R. Tripathy, L. Sharma, and S. Dandapat, "Detection of shockable ventricular arrhythmia using variational mode decomposition," *Journal of medical systems*, vol. 40, no. 4, p. 79, 2016.

[72] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.

[73] C. Ahlstrom, K. Hoglund, P. Hult, J. Haggstrom, C. Kvart, and P. Ask, "Assessing aortic stenosis using sample entropy of the phonocardiographic signal in dogs," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 2107–2109, 2008.

[74] A. Gavrovska, G. Zajić, I. Reljin, and B. Reljin, "Classification of prolapsed mitral valve versus healthy heart from phonocardiograms by multifractal analysis," *Computational and Mathematical Methods in Medicine*, vol. 2013, 10 pages, 2013.

[75] A. Cuevas, M. Febrero, and R. Fraiman, "An ANOVA test for functional data," *Computational Statistics & Data Analysis*, vol. 47, no. 1, pp. 111–122, 2004.

[76] F. Safara, S. Doraisamy, A. Azman, A. Jantan, and A. R. A. Ramaiah, "Multi-level basis selection of wavelet packet decomposition tree for heart sound classification," *Computers in Biology and Medicine*, vol. 43, no. 10, pp. 1407–1414, 2013.

[77] I. Maglogiannis, E. Loukis, E. Zafropoulos, and A. Stasis, "Support vectors machine-based identification of heart valve diseases using heart sounds," *Computer Methods and Programs in Biomedicine*, vol. 95, no. 1, pp. 47–61, 2009.

[78] Y. Zheng, X. Guo, and X. Ding, "A novel hybrid energy fraction and entropy-based approach for systolic heart murmurs identification," *Expert Systems with Applications*, vol. 42, no. 5, pp. 2710–2721, 2015.

[79] S. L. Oh, V. Jahmunah, C. P. Ooi et al., "Classification of heart sound signals using a novel deep wavenet model," *Computer Methods and Programs in Biomedicine*, vol. 196, p. 105604, 2020.

[80] Y. Luo and N. Mesgarani, "Conv-TasNet: surpassing ideal time-frequency magnitude masking for speech separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 8, pp. 1256–1266, 2019.

[81] C. S. Viswabhargav, R. Tripathy, and U. R. Acharya, "Automated detection of sleep apnea using sparse residual entropy features with various dictionaries extracted from heart rate and EDR signals," *Computers in Biology and Medicine*, vol. 108, pp. 20–30, 2019.

[82] R. K. Tripathy and S. Dandapat, "Automated detection of heart ailments from 12-lead ECG using complex wavelet sub-band bi-spectrum features," *Healthcare Technology Letters*, vol. 4, no. 2, pp. 57–63, 2017.

[83] U. B. Baloglu, M. Talo, O. Yildirim, R. San Tan, and U. R. Acharya, "Classification of myocardial infarction with multi-lead ECG signals and deep CNN," *Pattern Recognition Letters*, vol. 122, pp. 23–30, 2019.

[84] S. L. Oh, E. Y. Ng, R. San Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats," *Computers in Biology and Medicine*, vol. 102, pp. 278–287, 2018.

[85] R. Tripathy and U. R. Acharya, "Use of features from RR-time series and EEG signals for automated classification of sleep stages in deep neural network framework," *Biocybernetics and Biomedical Engineering*, vol. 38, no. 4, pp. 890–902, 2018.

[86] H. Uğuz, "Adaptive neuro-fuzzy inference system for diagnosis of the heart valve diseases using wavelet transform with entropy," *Neural Computing and Applications*, vol. 21, no. 7, pp. 1617–1628, 2012.