Usage of classical auscultation techniques is qualitative and subjective. The decision making process entirely depends upon the expertise and experience of the physicians conducting the observations. In order to make the outcome of the observations quantitative and independent of expertise and experience of the observer, Automated Computerized Analysis (ACA) of heart murmurs is best suited. Efficacy of classification of heart murmurs largely depends upon the feature extraction process. Murmurs, being a non-stationery signal, features in the time domain, frequency domain, spatial domain and phase domain are necessarily to be extracted. Such features will act as input to the classifier based on machine learning. In the present work, a new technique for extraction of features of heart murmurs is proposed to enhance its suitability for classification. Features were extracted in the time domain, Frequency domain, Time-Frequency domain and Phase space domain. As the size of the feature vectors were found to be very large, feature dimension reduction method was applied to reduce the number of feature vectors by eliminating redundant or dependent features. These features were then applied as input to the classifier for murmur screening. Four types of murmurs namely musical quality, coarse quality, soft quality and blowing like quality were considered for feature extraction. The heart signal data were obtained from open sources. PASCAL Heart Sound challenge Data set 2011 [1]. Various techniques were applied at different stages. DWT was employed for denoising, feature extraction and dimensionality reduction purposes after segmentation using energy envelopogram technique in combination with Gaussian Smoothing Filter (GSF). Support Vector Machine (SVM) with linear kernel function, Linear Discriminant Analysis (LDA), K-Nearest Neighbors (K-NN) and Random Forest were employed for classification.
Neighbors (kNN) Algorithm, Probabilistic Neural Network (PNN) techniques were applied for classification of the feature vectors to draw the conclusion regarding the screening of the murmurs. The present work confirmed that it is feasible to screen heart murmurs as normal or abnormal effectively using the techniques stated. There are scopes for further research work in the same direction to discriminate the causes of abnormalities present in the heart.

**Keywords:** PCG, ACA, Murmur classification, SVM, PCA, DWT

**Abbreviations used in the text:**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACA</td>
<td>Automated Computerized Auscultation</td>
</tr>
<tr>
<td>BPF</td>
<td>Band Pass Filter</td>
</tr>
<tr>
<td>CVS</td>
<td>Cardiovascular System</td>
</tr>
<tr>
<td>CWT</td>
<td>Continuous Wavelet Transform</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform</td>
</tr>
<tr>
<td>DR</td>
<td>Dimensionality Reduction</td>
</tr>
<tr>
<td>DWT</td>
<td>Discrete Wavelet Transform</td>
</tr>
<tr>
<td>ECG</td>
<td>Electro Cardiogram</td>
</tr>
<tr>
<td>EMD</td>
<td>Empirical Mode Decomposition</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FS</td>
<td>Feature Selection</td>
</tr>
<tr>
<td>GDA</td>
<td>Generalized Discriminant Analysis</td>
</tr>
<tr>
<td>GSF</td>
<td>Gaussian Smoothing Filter</td>
</tr>
<tr>
<td>HHT</td>
<td>Hilbert-Huang Transform</td>
</tr>
<tr>
<td>HPF</td>
<td>High Pass Filter</td>
</tr>
<tr>
<td>HSS</td>
<td>Heart Sound Signal</td>
</tr>
<tr>
<td>kNN</td>
<td>k-Nearest Neighbor Algorithm</td>
</tr>
<tr>
<td>LDA</td>
<td>Linear Discriminant Analysis</td>
</tr>
<tr>
<td>LKF</td>
<td>Linear Kernel Function</td>
</tr>
<tr>
<td>LPF</td>
<td>Low Pass Filter</td>
</tr>
<tr>
<td>LS-</td>
<td>Least Square Support Vector</td>
</tr>
<tr>
<td>SVM</td>
<td>Machine</td>
</tr>
<tr>
<td>MFCC</td>
<td>Mel Frequency Cepstrum Coefficients</td>
</tr>
<tr>
<td>MHMM</td>
<td>Multiple Hidden Markov Model</td>
</tr>
<tr>
<td>NBAY</td>
<td>Naïve Bays</td>
</tr>
<tr>
<td>NMF</td>
<td>Non-negative Matrix</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PCG</td>
<td>Phonocardiogram</td>
</tr>
<tr>
<td>PKF</td>
<td>Polynomial Kernel Function</td>
</tr>
<tr>
<td>PNN</td>
<td>Probabilistic Neural Network</td>
</tr>
<tr>
<td>RF</td>
<td>Random Forest Algorithm</td>
</tr>
<tr>
<td>RMS</td>
<td>Root Mean Square</td>
</tr>
<tr>
<td>SKF</td>
<td>Sigmoid Kernel Function</td>
</tr>
<tr>
<td>STFT</td>
<td>Short-time Fourier Transform</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>WPT</td>
<td>Wavelet Packet Transform</td>
</tr>
</tbody>
</table>
Introduction

Visualization of heart sound and its analysis using various advanced techniques now become a subject matter of immense interest for the researchers in the field of computational biology and signal processing of non-stationary signals like heart sound signals. Condition of the heart vis-à-vis the cardiovascular system can be cost effectively assessed from the valuable information extracted from the acoustical sounds generated by the heart due to its mechanical operations. The need of automatic analysis of heart sound stems from the fact that classical auscultations using acoustical stethoscope and expertise of the clinicians have been ruling this arena since long back. However, due to inter-observer variability, it is now felt that a standardize analysis method be developed to assist the clinicians in detecting the abnormal functionalities of heart leading to cardiovascular disorders and finally to causalities. The condition of the heart is reflected through the sounds it produces. Such acoustical sounds are mainly classified into three categories: Heart Sounds, Heart Murmurs and Heart Noises. The present work deals with the investigation of the most suitable techniques for the analysis of heart murmurs in detecting the abnormalities and the types of disorders in the cardiovascular system. The Automated Computerized Auscultation (ACA) will be proved a cost saving solution for condition based monitoring of heart based on investigation of heart sounds and murmurs. This in turn will supersede other costlier and time consuming techniques like Electrocardiogram, Echocardiography, Angiography etc. Currently, there are some commercially available ACA software products available [2]. However, these products focus mainly on the visual perception and play back functionalities. Automated diagnosis functionalities are still limited tasks such as murmur detection, screening and diagnosis.

Heart murmurs are different from standard heart sounds like S1, S2, S3 and S4. Their root cause is rushing of blood from one chamber to another during systolic and diastolic phases [3]. They may or may not be normal. The occurrence of vibrations due to turbulent flow of the flow of blood and associated resonance created in the surrounding tissues are the root cause of heart murmurs. The frequency spectrum of murmurs lies in the range of 10 Hz to 1 KHz and is considered higher than the heart sounds. Healthy hearts also can produce murmurs under certain conditions like if the flow rate of blood through the valves is larger than the average value of the flow rate. The reason behind appearance of the heart murmurs are generally, valvular defects, leakage through the interior separation walls between the heart chambers (atrial septal defect or ventricular septal defect), heart defects by birth, physiological disorders like high body temperature, anemia etc.

As blood flows through the blood vessels, it creates a murmur known as Innocent murmurs, which are very common in neonatal [4]. Pathological murmurs are caused due to disorders in heart. They are generated when blood passes through a tapered or leaky blood vessel. The physical symptoms caused on the appearance of the pathological murmurs normally include shortness of breath, swelling of legs, fatigues, chest pressure, dizziness etc. Typical shape of heart murmurs is shown in Fig. 1 below for visual perception.
Ailing heart valves may cause, normally, two types of murmurs: *Stenotic murmur* (if the heart valves are incapable of fully opening) and *Regurgitant murmur* (if the heart valves are not capable of restricting blood flow completely in the opposite direction). Septum defect may lead to leakage of blood and is responsible for generation of heart murmurs.

Depending on the time of occurrence, shape, point of origin, radiation, intensity, pitch and quality, the murmurs are classified into seven categories as presented under [5]:

**Table 1: Categorization of heart murmurs based on their characteristics**

<table>
<thead>
<tr>
<th>Time of occurrence</th>
<th>Systolic murmur (if it occurs during the systolic phase) and diastolic murmur (if it occurs during diastolic phase).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape</td>
<td>Crescendo (if the intensity continuously increases), decrescendo (if the intensity progressively decreases), Crescendo-Decrescendo (if the intensity progressively increases and the progressively decreases).</td>
</tr>
<tr>
<td>Point of origin/location</td>
<td>Depending on the auscultation area from where the murmurs are best-heard using stethoscope. Four prominent locations namely aortic, pulmonic, tricuspid and mitral regions are the most prominent locations to consider. Moreover, the posture of the patient is also important during auscultation whether the patient is lying down or facing up.</td>
</tr>
<tr>
<td>Radiation</td>
<td>It refers to the direction of movement of the murmurs. Normally it radiates along the direction of blood flow.</td>
</tr>
<tr>
<td>Intensity</td>
<td>It characterizes the loudness of the murmurs and the murmurs graded between Grades I to Grade VI depending on the nature of the acoustical intensity.</td>
</tr>
<tr>
<td>Pitch</td>
<td>It may be low, medium or high and is decided based on whether they can be identified using the bell or diaphragm of the stethoscope.</td>
</tr>
<tr>
<td>Quality</td>
<td>Based on quality of the acoustic signal, murmurs can be grouped as blowing, harsh, rumbling or musical.</td>
</tr>
</tbody>
</table>

A brief description of the associated pathological condition of the heart with the murmurs is presented in a tabular form in Table 2 below:
Table 2: Murmur types and the probable causes [6]

<table>
<thead>
<tr>
<th>Heart Condition</th>
<th>Type</th>
<th>Probable cause</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innocent Flow Murmur</td>
<td>Normal</td>
<td>Increased cardiac output.</td>
</tr>
<tr>
<td>Innocent Vibratory Murmur</td>
<td>Normal</td>
<td>Vibration of heart tissue(s).</td>
</tr>
<tr>
<td>Aortic Stenosis (AS)</td>
<td>Abnormal</td>
<td>Inability of an aortic valve being able to fully open during systole.</td>
</tr>
<tr>
<td>Mitral Stenosis (MS)</td>
<td>Abnormal</td>
<td>Reduced mitral valve orifice.</td>
</tr>
<tr>
<td>Pulmonary Stenosis</td>
<td>Abnormal</td>
<td>Obstruction of blood flow across the pulmonary valve due to narrowed valve opening.</td>
</tr>
<tr>
<td>Ventricular Septal Defect (VSD)</td>
<td>Abnormal</td>
<td>Defect in ventricular septum causing blood flowing from left ventricle to right ventricle.</td>
</tr>
<tr>
<td>Aortic Regurgitation (AR)</td>
<td>Abnormal</td>
<td>Improper closure of aortic valve during diastole.</td>
</tr>
<tr>
<td>Mitral Regurgitation (MR)</td>
<td>Abnormal</td>
<td>Improper closure of mitral valve during systole.</td>
</tr>
<tr>
<td>Atrial Septal Defect (ASD)</td>
<td>Abnormal</td>
<td>Defect in atrial septum causing blood flowing from left atrium to right atrium.</td>
</tr>
</tbody>
</table>

Innocent murmurs are mostly found in the murmurs generated by the heart. ACA can be effective in murmur diagnosis by avoiding the wastage of costly medical resources like echocardiography, angiography and others. However, still there are scopes for extensive researches for quantization using numerical characterization and proper classification of murmurs using ACA for superior analysis for effective fixing of the heart condition. The present work thus aims to classify the patterns and quality of the murmurs using numerical categorization to conclude in a logical manner for murmur screening.

Literature Survey

Since the introduction of PCG, advancements had been made to improve the effectiveness of visualizing the heart sounds through PCG for diagnosis. It was realized earlier on throughout the PCG analysis developments that techniques such as frequency analysis and spectral analysis are greatly beneficial. Van Vollenhoven et al. (1969) [7] developed an analog based frequency analysis system to enhance the interpretation of PCGs in cases of combined heart murmur complications. Sarkady et al., (1976) [8] developed one of the earlier microcomputer systems that were capable of recording and processing the PCG signals. Spectrogram based analysis techniques were developed for PCG diagnosis by Showalter et. al., (1979) [9]. They further developed parameter extractions based on the spectral measurements and the time measurements to quantitatively describe the acoustic characteristics of the aortic ejection click and the first heart sound in the case of aortic stenosis. Ahlstrom et al. (2001) [10] reported the feature types and extraction techniques for systolic heart murmur classification only. Pretorius et al.(2010) [11] considered the use of STFT to extract features by summing the coefficients in the equally spaced frequency windows and segmented time domain bins. Kumar et al.(2010) [12] used the Lyapunov exponent estimation to characterize the degree of chaos in a dynamical system, which was then used to detect the existence of murmurs as well.
as used as part of the feature set. Zheng et al., (2015) [13] employed wavelet packet decomposition and energy based entropy to extract the most dominant features of HSS. The classification process was carried out using optimal kernel function. The proposed method exhibited high accuracy, specificity and sensitivity to the tune of 97.17%, 98.55% and 93.48% respectively. Yaseen and Kwon (2018) [14] created database of 5 categories of HSS from various sources containing one normal and four abnormal categories. MFCC and DWT were employed to extract the features and for the classification purpose SVM, DNN and kNN were utilized. The diagnosis accuracy has been reported as 97%.

Normalized average Shannon energy with time domain characteristics was employed by Ahmad et. al. (2019) [15] to segment the PCG signal to extract the features. They used iterative backward elimination technique for reduction of feature vectors.

In order to detect and screen the types of cardiovascular disorders, an MFCC feature vector with a dimension of 26 was adopted for training seven different types of SVM and kNN based classifier. Out of these classifiers, SVM classifier with Gaussian function kernel exhibited a high classification accuracy of 92.6%.

**Methodology**

In order to reflect the physical characteristics of the PCG signal, time domain features namely variance, standard deviation, entropy, peak amplitude, RMS value, crest factor, impulse factor, shape factor, energy content and clearance factor are mainly used in machine learning. Time domain features are analytically simple and easy to compute. Timing, configuration, location, quality, intensity, duration and radiation are the characteristics of the murmurs used for the sake of characterization of the murmurs during ACA.

![Figure 2: Block diagram representation of the Framework of the present work](image)
Four common types of murmur quality descriptors namely, coarse quality, soft quality, blowing quality and musical quality are in use for testing and training purposes. The different stages of proposed murmur analysis for screening have been implemented using the techniques as discussed in the paragraphs to follow.

**Preprocessing stage:**

It is used to improve the quality of the processed raw data by removing noises embedded in the signal during acquisition while preserving the useful information. Denoising of the PCG signal can be done in many ways those include use of linear filters like HPF, LPF and BPF or nonlinear filtering like Kalman Filter or by using wavelet transform technique including either CWT or DWT [16]. In the present work DWT based denoising has been employed.

**Segmentation Process:**

Segmentation is the process to identify the important segments/portions of the PCG that can be achieved by adopting many techniques based on the analysis of the energy Envelope, various Features, Time-Frequency/Wavelet analysis etc. In the present work, Gaussian Smoothing Filter and entropy based envelop analysis method have been employed for segmentation process [17].

**Feature Extraction Process:**

It helps in mining the features of the signal that may be of interest for analysis purpose. Extraction of features can be done in many ways those include FFT, DWT, Wavelet analysis, S-transform, HHT, MFCC etc. In the present work, DWT, HHT and MFCC techniques are in use for feature extraction [17].

**Feature Reduction/Dimensionality Reduction Process:**

As all the features extracted from the signal may not be of prime importance for classification and may be correlated to other features, hence feature reduction or feature selection becomes necessary to lower down the feature dimension to be applied to the classifier to reduce computational complexity. Among the techniques used for feature reduction, PCA, GA, GP, GDA, NMF and RF are prominent for the purpose. In the present work, PCA, NMF and RF have been utilized for feature reduction process [17].

**Classification Process:**

This process is used to categorize or group the information extracted from features of the signal based on given criteria to extract the decision regarding the nature of the signal for diagnostic purposes. Classifiers used by previous workers include SVM with different Kernel function including LKF, PKF, GRKF and SKF, LS-SVM and compared with BP-ANN and HMM, ANN, ANN & SVM, PNN, MHMM, PCA-DHMM, ANFIS and HMM, AIS and Fuzzy k-NN MLP, RBF and SVM classifier, a combination of DT, k-NN, Bays Net, MLP and SVM, Naive Bayes classifier. In the present work, we have used SVM, k-NN, PNN and NBAY for classification.

Making a combination of various techniques for feature extraction, dimensionality reduction and classification, total nine models have been designed in which the data set obtained from
open sources are inputted. The performances of the various combinations are tested in terms of classification accuracy of each model [17].

**Experiment Design and Results**

**Dataset**

PCG data sets from PASCAL [1] are used for murmur screening through classification. As a part of challenges namely segmentation and classification purposes, PASCAL data set was initially made available in 2011. The dataset is divided into two subsets marked as Dataset A and Dataset B. iStethoscope, a Pro iPhone App was used for collecting the Dataset A whereas DigiScope, a digital stethoscope, was engaged in a clinical setting in hospital for creating the dataset B. The files in the datasets are available in .wav format. Dataset A, the smaller one, contains 176 auscultations data as compared to 656 auscultations in Dataset B. Dataset A contains normal beats, murmurs and extra systole. Dataset B contains normal beats, murmurs, extra systole, extra heart sounds and artifacts. The data, examined by experts, are tagged with the types of abnormalities present so that the same can be used for training and testing purposes during classification. In the present work, Dataset B was utilized as input.

**Computational Environment**

In this work, a MATLAB signal processing program is used to develop the computer based monitoring system of heart based on PCG signal. This program uses the heart sounds signals which are the input and processed through sophisticated signal processing algorithms before a final diagnosis can be made.

**Preprocessing**

In the preprocessing stage of the experiment, the PCG signal has been processed for the removal of the Baseline wander, Normalization of the signal and denoising of the signal to get rid of various types of noises embedded in the signal during acquisition. Baseline Wander (BLW) is a common problem appeared during acquisition of any natural and feeble signal. It shifts the reference level in an unpredicted manner causing hindrances in proper extraction of the signal parameters. It is thus necessary to remove such wandering of the baseline of the signal. Adaptive Smoothing Filter (ASF) with a higher window length of 2.2 sec with an iteration number equal to 5 has been employed in the current work for the removal of BLW.

Normalization of a signal is a technique to change the range of the signal by increasing or decreasing the sampled values of the signal through multiplication of the signal by a predefined factor based on a mathematical function. The aim of the normalization process of a signal lies in the fact that normalization of a signal removes redundancy of amplitude data so that storage of the data occupies less space at the same time less data are to be handled for processing. Normalization can be done both in time as well as in amplitude domain. Sliding window normalization technique has been applied in the present work. As the PCG signals are corrupted by noise during acquisition process, they need to be made free from such noises. This is done through denoising process. In the present experiment,
DWT based decomposition, thresholding and Inverse DWT based synthesis are used to denoise the PCG signal. A mother wavelet is chosen for decomposition and reconstruction process. In the present work, for the purpose of denoising the PCG signals a combination of sym 20 as the MWT with decomposition level (DL) of 10 and Bayes Soft as the thresholding function are utilized [16].

**Design of Models**

In order to make a comparison of performances of various techniques applied for feature extraction, feature reduction and classification methods, nine models with various combinations of standard techniques as mentioned below in a tabular form has been designed.

**Table 3: Details of the models used**

<table>
<thead>
<tr>
<th>Model No.</th>
<th>Techniques applied for Feature extraction</th>
<th>Techniques applied for Feature reduction</th>
<th>Techniques applied for Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>DWT</td>
<td>PCA</td>
<td>SVM</td>
</tr>
<tr>
<td>II</td>
<td>MFCC</td>
<td>PCA</td>
<td>SVM</td>
</tr>
<tr>
<td>III</td>
<td>HHT</td>
<td>PCA</td>
<td>SVM</td>
</tr>
<tr>
<td>IV</td>
<td>DWT</td>
<td>RF</td>
<td>SVM</td>
</tr>
<tr>
<td>V</td>
<td>DWT</td>
<td>GDA</td>
<td>SVM</td>
</tr>
<tr>
<td>VI</td>
<td>DWT</td>
<td>NMF</td>
<td>SVM</td>
</tr>
<tr>
<td>VII</td>
<td>DWT</td>
<td>PCA</td>
<td>KNN</td>
</tr>
<tr>
<td>VIII</td>
<td>DWT</td>
<td>PCA</td>
<td>PNN</td>
</tr>
<tr>
<td>IX</td>
<td>DWT</td>
<td>PCA</td>
<td>NBAY</td>
</tr>
</tbody>
</table>

**Results**

All nine models developed have been tested for classification accuracy and the results obtained are tabulated. In all the models, to maintain the uniformity, out of 656 auscultation files available in Dataset B, 2500 beats are selected at random and used as the input to the models. Out of 2500, 300 beats are normal beats without murmurs and 2200 beats are with murmurs. 210 normal beats are used for training and 190 beats are used for testing purposes for classification. Out of 2200 abnormal beats, 1540 beats are utilized for training the classifier and 660 beats are employed for testing purpose.

**Table 4: Classification Accuracy obtained for various models**

<table>
<thead>
<tr>
<th>Model</th>
<th>Classification Techniques used</th>
<th>Classification Accuracy obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model I</td>
<td>SVM</td>
<td>97.02%</td>
</tr>
<tr>
<td>Model II</td>
<td>SVM</td>
<td>96.95%</td>
</tr>
<tr>
<td>Model III</td>
<td>SVM</td>
<td>91.28%</td>
</tr>
<tr>
<td>Model IV</td>
<td>SVM</td>
<td>93.33%</td>
</tr>
<tr>
<td>Model V</td>
<td>SVM</td>
<td>90.09%</td>
</tr>
<tr>
<td>Model VI</td>
<td>SVM</td>
<td>92.48%</td>
</tr>
<tr>
<td>Model VII</td>
<td>KNN</td>
<td>94.69%</td>
</tr>
<tr>
<td>-----------</td>
<td>-----</td>
<td>--------</td>
</tr>
<tr>
<td>Model VIII</td>
<td>PNN</td>
<td>88.59%</td>
</tr>
<tr>
<td>Model IX</td>
<td>NBAY</td>
<td>89.55%</td>
</tr>
</tbody>
</table>

**Conclusions**

Generation of the murmurs especially pathological murmurs are due to defects or disorders in heart valves and vessels. Analysis of heart murmurs refines the auscultation data and provides sufficient information relating to the pathological signs in the cardiac cycle. In the present work, the three-stage detection process including Preprocessing, Signal Processing and Classification have been implemented using standard techniques. Nine models of ACA have been designed and developed. The dataset from PASCAL, an open source for heart sound data, has been utilized as input with normal as well as abnormal murmurs. From the result, it is clear that the murmur screening problem can be solved in a better manner by using supervised classification technique such as SVM. The combination of DWT for denoising and feature extraction, PCA for feature reduction and SVM for classification provides a superior result. However, the present work deals only with the murmur screening process, i.e., to understand whether the observed murmur is innocent or pathological one. Further scopes lies with the present work for detection of the type and cause of fault in the heart or cardiovascular system.

**References**

M. Sheraz Ahmad, Junaid Mir, Muhammad Obaid Ullah, Muhammad Laiq Ur Rahman Shahid, Muhammad Adnan Syed, “An efficient heart murmur recognition and cardiovascular disorders classification system”, Australasian Physical & Engineering Sciences in Medicine, 42:733–743, 2019