Heart Sounds Classification using Loudness Features and Gaussian Mixture Model

Vishwanath Madhava Shervegar [a*]

DOI:

ABSTRACT

This paper represents a new automatic method of classifying the heart sound status using the loudness features of the heart sound. The method includes the following three main steps. First, the heart sound, which is usually found noisy, is heavily filtered by a 8^{th} order Chebyshev Type-II filter. The event synchronous method is later used to segment the filtered heart sound into the first heart sound, systole, second heart sound and diastole. In the second step, the loudness feature is represented using the mean rows of its spectrogram. The third step categorises the heart sound using the Gaussian Mixture Model approach. With a success rate of 97.77 percent, the suggested method has been evaluated on a huge database of heart sounds containing over 3000 recordings.

Keywords: Phonocardiography; event synchronous segmentation; loudness; spectrogram; GMM classifier.

1. INTRODUCTION

Heart sounds entail crucial heart function information. In conditions of heart abnormalities, such as valve dysfunctions and rapid blood flow, additional sounds are heard in regular heart sounds, which can be employed in pathology diagnosis [1-3]. Heart valve disorders are one of the main causes for heart diseases, which may lead to loss of life. Heart sounds are produced by the heart's electromechanical activity with each heartbeat [4]. In the case of normal heart sounds, heart valves create time variable low frequency transient signals. The turbulence in blood flow caused by stenosis or regurgitation through the heart valves causes pathological cardiac sounds or murmurs [5]. Automated classification of pathology in heart sounds are been reported in literature for over 50 years but accurate diagnosis remains a significant challenge. Gerbarg et al (1963) were the first to publish on the automatic classification of pathology in heart sounds using a threshold-based method (specifically to aid the identification of rheumatic heart disease in children). In general methods for heart sound classification can be grouped into four categories: (1) artificial neural network-based classification; (2) support vector machine-based classification; (3) HMM-based classification and clustering-based classification (4). For relative brevity, some of the prominent resent works in this field are listed in Table 1. Table 1 also gives details on the evaluation of the method, and training-test sets used in these approaches. Uguz (2012a) [6] used an ANN-Fuzzy approach with wavelet transform to classify three-class namely normal, pulmonary stenosis, and mitral stenosis. With a 50/50 train/test split of a dataset of 120 subjects, success rate of 100% in terms of sensitivity, 95.24% in terms of specificity, and 98.33% average accuracy is been reported for the three classes. Uguz (2012b) [7] also used time-frequency as an input to an ANN. A total of 120 heart sound recordings, split 50/50 into train/test, was used. They reported 90.48% sensitivity, 97.44% specificity and 95% accuracy for a three-class classification problem (normal, pulmonary and mitral stenosis heart valve diseases). Ari et al (2010) [8] used a least square SVM (LSSVM) method for classification of normal and abnormal heart sounds using features based on wavelet transform. The performance evaluation of the proposed method was done using 64 recordings consisting of normal and pathological cases. LSSVM was trained and tested on a dataset with 50/50 (train/test) split, the authors reported 86.72% accuracy on

[a] *E&CE Department, Mangalore Institute of Technology & Engineering, Moodbidre, DK, Karnataka, India.*
**Corresponding author: E-mail: vishwa@hotmail.co.in;*

their test dataset. Zheng [9] et al (2015) decomposed heart sounds using wavelet packets and then extracted the energy fraction and sample entropy as features for the SVM input. They tested on 40 normal and 67 pathological patients and reported a success rate 97.17% in terms of accuracy, 93.48% in terms of sensitivity and 98.55% in terms of specificity. Patidar [10] et al (2015) used the tunable-Q wavelet transform as an input to LSSVM with varying kernel functions. A dataset of 4628 cycles from 163 heart sound recordings was used (and an unknown number of patients) and a 98.8% sensitivity and 99.3% specificity is reported, but without stratifying patients (having mutually exclusive patients in testing and training sets), and therefore overfitting to their data. Maglogiannis [11] et al (2009) used Shannon energy and frequency features from four frequency bands (50–250, 100–300, 150–350, 200–400 Hz) to develop an automated diagnosis system for the identification of heart valve diseases based on an SVM classifier. Testing on 38 normal and 160 heart valve disease patients they reported an 87.5% sensitivity, 94.74% specificity and 91.43% accuracy. Gharehbaghi [12] et al (2015) used frequency band power over varying length frames during systole as input features and used a growing-time SVM (GTSVM) to classify pathological and innocent murmurs. When using a 50/50 train/test split (from 30 patients with AS, 26 with innocent murmurs and 30 normal), they reported 86.4% sensitivity and 89.3% specificity. Saracoglu (2012) [13] applied HMM in an unconventional manner, by fitting an HMM to the frequency spectrum extracted from entire heart cycles. The exact classification procedure included the training of four HMMs, and then evaluation of the posterior probability of the features for classification of the recordings. They optimized the HMM parameters and PCA-based feature selection on a training set and reported 95% sensitivity, 98.8% specificity and 97.5% accuracy on a test dataset of 60 recordings. Quiceno-Manrique [14] et al (2010) used a simple kNN classifier with features from various time-frequency representations on a subset of 16 normal and 6 pathological patients. They reported 98% accuracy for discriminating between normal and pathologic beats. However, the kNN classifier parameters were optimized on the test set, indicating a likelihood of over-training. Avendano-Valencia [14] et al (2010) also employed time-frequency features and kNN approach for classifying normal and murmur patients. In order to extract the most relevant time-frequency features, two specific approaches for dimensionality reduction are presented in their method, namely, feature extraction by linear decomposition, and tiling partition of the time-frequency plane. The experiments were carried out using 26 normal and 19 pathological recordings and they reported an average accuracy of 99.0% when using 11-fold cross-validation with grid-based dimensionality reduction. In our work, heart sounds are classified by using the loudness index features obtained using spectrogram. In order to reduce the complexity and burden of computations, maximum and minimum values of the loudness index derived from spectrogram have been used in this work as compared to usage of spectral centroid.

2. PRE-PROCESSING OF THE HEART SOUNDS

In this study we have used the dataset of heart sounds available from the Physionet database [15], The raw heart sound of this database is further pre-processed for de-noising and normalisation.

2.1 Heart sound De-noising

Heart sounds obtained from diagnostic tools are generally contaminated with noise from various sources. These sounds hinder the early detection of mild heart sounds in the phonocardiogram. So filtering of noise to remove such artefacts becomes essential. The filtering is to be done taking in to account that there is no loss of diagnostic information required for analysis of the phonocardiogram, but removing all unwanted components. The heart sound taken from the Physionet database [15] is contaminated with various types of noises. The heart sound selected is heavily filtered to remove the maximum noise from the sound. A 6^{th} order chebyshev low pass filter is used for this purpose. The filtered heart sound is then put through the second phase of analysis namely segmentation. The filtered heart sound later scaled to the required amplitude levels by normalisation. The normalisation procedure involves dividing each heart sound sample by the maximum value of the heart sound sample in each heart sound.

2.2 Segmentation of Heart Sound

After de-noising and normalisation of heart sound the phonocardiogram is segmented by making use of the occurrence of cardiac events. Some previous studies have used the electro-cardiogram (ECG) recordings to segment the heart sounds [15]. However, these signals are not available all the time along with the PCG collections data. Spectrum analysis using Daubechies four and five wavelet coefficients [16] provide a alternate approach to separate the original heart sound signals into individual cycles. The signals are down sampled by a factor of four resulting in highest power to the first and second heart sounds (S1 and S2) in each cycle. Segmentation based on cycle frequency and dynamic clustering in time-frequency domains is been considered [17]. We have addressed the procedure for segmentation of heart sound using the event synchronous method [18]. Event synchronous method is a fast implementation of segmentation using spectrogram. It reduces the computational time and complexity involved in other methods. Fig. 1 illustrates a sample abnormal heart sound taken from the physionet database with clear distinguishable segmentation boundaries. A brief step by step procedure utilised is explained here:

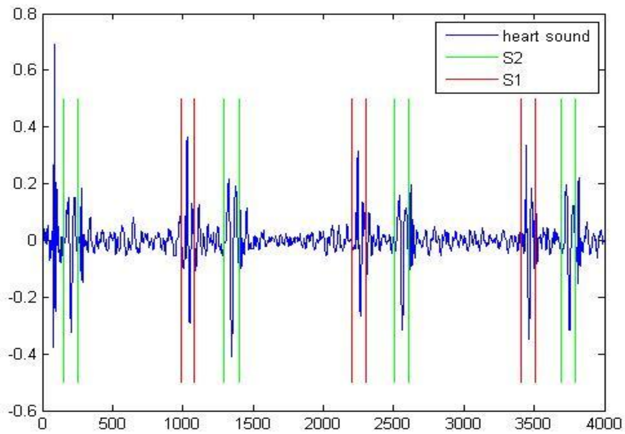

Fig. 1. Segmented heart sound (blue) and Segmentation boundaries (S1 red, S2 green)

1. Spectrogram, is obtained from the cardiac sound signal, at roughly 3ms window.

2. For a better interpretation of heart sound murmur pitch and loudness the spectrogram is converted into bark scale and then smoothened using a hanning window [19].
3. The sensation of loudness is derived from the spectrogram by summing the amplitudes of all frequency bands of the sound: $L_{dB}(t) = \frac{\sum_{k=1}^{N} E_k(t)}{N}$ (1) E_k represents the magnitude of the kth frequency band present in the spectrogram. There is a total of N such bands.
4. Event detection function is found by differentiating the loudness index function to obtain peaks that correspond to the onset and offset transients.

5. Systole and diastole are recognised by their durations with systole being shorter than the diastole.
6. A single cardiac cycle is obtained from the segmentation of all sounds.

3. FEATURE EXTRACTION

3.1 Extraction of Loudness Features

It is known from auscultation and also from literature that there exists silence period in the normal heart sounds namely systole and diastole where the intensity of the sound is minimal. In patients suffering from valve disorders such as stenosis and regurgitation certain high pitch sounds are also audible. These sounds have less loudness (grade 1-grade 4 sounds) compared to the first and second heart sounds. For grade 5 and grade 6 abnormal sounds, the loudness is much more than the first and second heart sounds. More energy is concentrated in the systoles and diastoles of abnormal sounds than the normal ones. Clearly loudness index is a measure than can be possibly used to identify and distinguish heart sounds with various indices.

To determine the features for classification we repeat the steps mentioned in the previous stage, namely 1 to 3. Maximum and minimum values of the heart sound loudness index are found and are used as the two features for the classification of these sounds. Maximum loudness indices are high for abnormal sounds while the minimum loudness indices are low for normal sounds. Fig 2 shows a single normal cardiac cycle with loudness function. Fig 3 shows a single abnormal cardiac cycle with loudness function. Sharp peaks corresponding to murmurs are visible in the abnormal sounds.

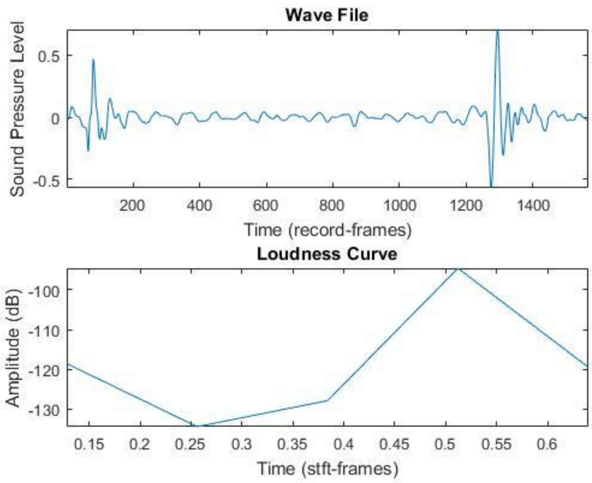

Fig. 2. Normal heart sound and the loudness curve

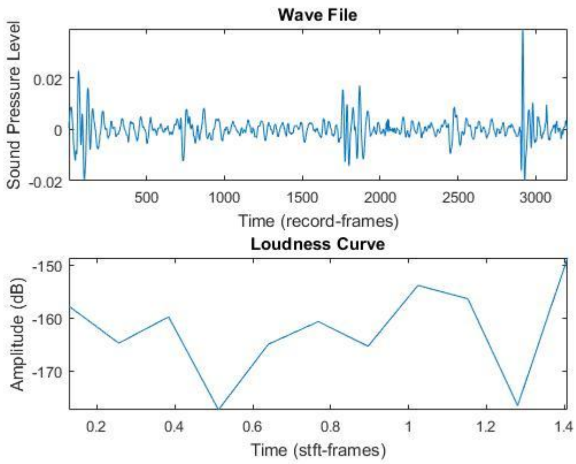

Fig. 3. Abnormal Heart sound with loudness curve

4. CLASSIFICATION OF HEART SOUNDS USING GMM MODEL

Literature survey reveals many classification methodologies used for heart sound classification as listed in Table 1. Euclidean distance (or in general, Minkowski distance) is a commonly used distance metric. The drawback of this Euclidean metric is that the largest scaled feature gets dominated. Solution to this problem could be normalization of the continuous features to a common range of variance. Another drawback of Euclidean metric is that linear correlation between features can distort the distance metric. Euclidean distance always gives hyper-spheroid clusters. Mahalanobis distance, used in GMM, is expected to overcome all the above limitations where it provides hyper-ellipsoidal clusters. The database consisting of two sound types is assumed to be generated by the two Gaussian processes, of course each Gaussian could have generated by many biochemical processes inside the heart having arbitrary stochastic distribution. Therefore, the entropy of the resultant distribution tries to become highest. Hence the resultant distribution is Gaussian which is having the maximum entropy among any other possible stochastic distributions. The probability distribution function (pdf) is given by equation 4.

$$P\ (x_{n|}\lambda) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1}(x-\bar{x})\right) \tag{4}$$

Where d is the dimensionality. ML estimators of μ and ∑ are computed by

$$\bar{x} = \frac{1}{N}(\textstyle\sum_n x_n \tag{2}$$

$$\textstyle\sum = \frac{1}{N}(\sum_n (x-\bar{x})(x-\bar{x})^T \tag{5}$$

The objective function is formed given by Eq. (4), by summing the class conditional density over all the classes for a feature in the feature vector; and again taking the product for all the features, assuming the features are linearly independent. Such likelihood based objective function for optimization is maximized by EM [20] algorithm. It is a nonlinear optimization method which optimizes the log likelihood over the entire feature space, including both observed data and hidden information embedded in the data. The EM algorithm consists of Expectation step (E-step) and Maximization step (M-step). In E-step the posterior density based on conditional density using Bayes rule [21] is computed. In the M-step the initial model is replaced with a new model which is a better one to represent the features such that the log-likelihood is more than that of the previous iteration. The iterations are repeated until the new estimate will give same model and there will not be any improvement in the model. The algorithm is briefly described here:

Input: A set of N feature vectors $E = \{x_1, x_2, \ldots x_N\}$ model structure $\Lambda = \{\mu_k, \Sigma_k, \alpha_k\}, k = 1 \ldots K,$ where μ's and ∑'s are parameters for the Gaussian models and α's are prior parameters subject to $\alpha_k \geq 0, \forall k \ and\ \Sigma_k \alpha_k = 1.$

Output: Trained model parameters Λ that maximizes the data likelihood $P(E|\Lambda) = \prod_n \sum_k \alpha_k p(x_{n|}\lambda_k)$ (6)

And a partition of data vectors given by the cluster identity vectorY = {$y_{1\ldots}$, y_N}, y_n ∈ {1..., K}.

Steps:

1. (Initialization) Initialize the model parameters Λ.
2. (E-Step) The posterior probability of model k , given a data vector x_n and current model parameters Λ , is estimated as $P(k|x_n, \Lambda) = \frac{\alpha_k p(x_{k|}\lambda_k)}{\sum_j \alpha_j p(x_{n|}\lambda_k)}$ (7)where the pdf p(e|λ) is given in (1)
3. (M-Step) The ML re-estimation of model parameters Λ is given by $\mu_k^{(new)} = \frac{\sum_n P(k|x_n,\Lambda)x_n}{\sum_n P(k|x_n,\Lambda)}$

(8) $\Sigma_k^{(new)} = \frac{\sum_n P(k|x_n,\Lambda)(x_n-\bar{x}_n)(x_n-\bar{x}_n)^T}{\sum_n P(k|x_n,\Lambda)}$ (7)$\alpha_k^{(new)} = \frac{1}{N}\sum_n P(k|x_n, \Lambda)$ (9)

4. (Stop) if P(E| Λ) converges; otherwise go back to Step 2;

For each data vector x_n, set $y_n = arg_k \max\ (\alpha_k p(x_{n|}\lambda_k))$ (10)

Table1. Comparative results of previous segmentation methods

Author	Database	Recording length	Classification method	features	Se (%)	Sp (%)	Acc (%)	Notes on database
Uguz (2012a)	40 normal, 40 pulmonary and 40 mitral stenosis	-	ANN	Wavelet	100	95.24	98.33	50-50 train-test split
Uguz (2012b)	40 normal, 40 pulmonary and 40 mitral stenosis	-	ANN	Time-Frequency	90.48	97.44	95	50-50 train-test splits
Ari et.al (2010)	64 patients (normal and pathological)	Each 8 cycles	SVM	Wavelet	-	-	86.72	50-50 train-test splits
Zheng et.al (2015)	40 normal and 67 pathological	-	SVM	Wavelet	93.48	98.55	97.17	Cross- Validation
Patidar et.al (2015)	Total 4628 heart cycles, 626 normal and 4002 pathological	-	SVM	Wavelet	98.8	99.3	98.9	80% training 20% test
Gharehbaghi et al (2015)	30 normal, 26 innocent and 30 AS	Each 10s	SVM	Frequency	86.4	89.3	-	50-50 train-test split
Sarcoghlu et.al (2012)	40 normal, 40 pulmonary and 40 mitral stenosis	-	HMM	DFT and PCA	95	98.8	97.5	50-50 train-test split
Quiceno-Manrique et al (2010)	16 normal and 6 pathological	-	kNN	Time-Frequency	-	-	98	Cross- Validation
Avendano-Valencia et al (2010)	26 normal and 19 pathological	-	kNN	Time-Frequency	99.56	98.45	99.0	Cross- Validation

Table 2. Scoring of all sounds

sounds	TP	TN	FP	FN	Se(%)	Sp(%)	Acc(%)
Normal	109	112	0	3	97.13	100	98.33
Abnormal	112	109	3	0	100	97.13	98.33
Total	221	221	3	3	98.33	98.33	98.33

5. RESULTS AND DISCUSSION

5.1 Database Used

Heart sound present in the Physionet heart sound database [15] were taken from several people around the world, collected at either a clinical or nonclinical environment, from pathological patients. The Challenge training set consists of five databases (A through E) containing a total of over 3000 heart sound recordings. The duration of heart sound varies from 5 seconds to just over 120 seconds. The heart sound recordings were obtained from different locations on the body. Four locations used for acquisition are aortic area, pulmonic area, tricuspid area and mitral area. Abnormal heart sounds were taken from patients with a confirmed cardiac diagnosis. The patients suffered from a variety of illnesses such as heart-valve defects and coronary artery disease. Heart valve defects cover mitral valve prolapse, mitral regurgitation, aortic stenosis and valvular surgery Pathological patients included both children and adults. Each subject/patient have contributed between one and six heart sound recordings. All recordings have a sampling rate of 2,000 Hz and are in .wav format.

5.2 Classification of Heart Sounds

The heart sounds selected from the physionet database are pre-processed by passing the sounds through a sixth order type 1 chebyshev filter to remove the redundant noise. The heart sounds are normalised in the range [22,23]. The normalised heart sounds are processed using the event synchronous segmentation procedure. This results in cycles of heart beat of approximately 3 s duration. Two datasets are formed consisting of abnormal and normal heart sounds apiece, using the sounds in the physionet database. In each of the datasets two groupings are done one for training set and the other for test set. 50% of the dataset sounds make up the training set while the remaining 50% sounds make up the test set. The sounds are then clustered using the GMM classifier. Fig 4 shows the log likelihood function plot for the sounds. The function is a negative function and it increases steadily and then stabilises to a constant value for a total of 101 iterations. Fig 5 shows the scatter plot of all sounds. Two features were used the maximum loudness along x axis and minimum loudness along y axis. Fig 6 shows the EM plot of all sounds. The points in yellow are the abnormal sounds with higher loudness values while the sounds in green are the normal sounds with lower loudness values. The centroids are also shown and the ellipse represents the points closer to the centroid. A total of 112 sounds of each type- abnormal and normal were used for training purpose. Similarly, 112 sounds of each type were used for testing purpose as well. Out of 112 sounds, 3 sounds in normal dataset misclassified as abnormal sounds due to the presence of noise that remained even after filtering. This is shown in Table 2. Both Normal sounds and abnormal sounds gave a similar accuracy of 98.13%. So the overall accuracy was 100%. The sensitivity of normal sounds was 97.33% while the abnormal sounds gave 100%. The specificity of normal sounds was 100% while the abnormal sounds gave 97.33%. This gave an overall sensitivity and specificity of 98.13%.

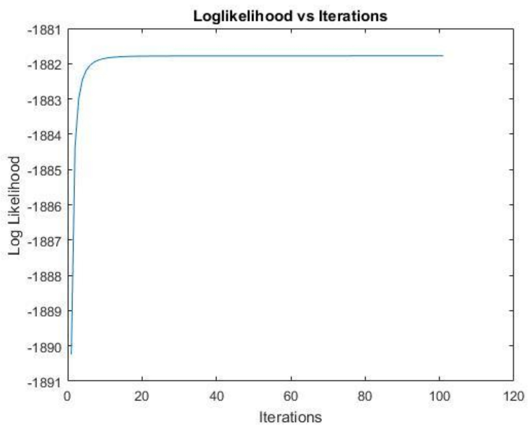

Fig. 4. Log likelihood variations versus iterations

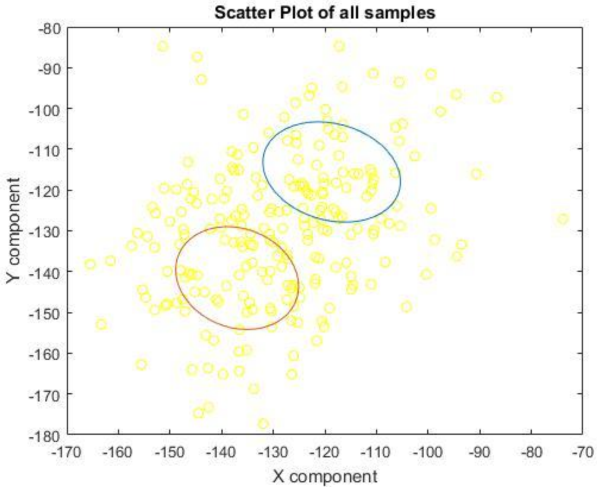

Fig. 5. Scatter plot of all observations

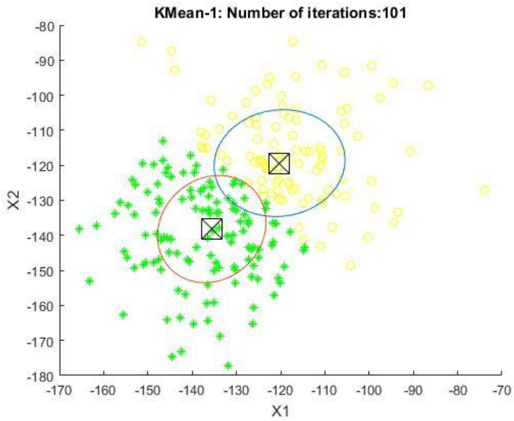

Fig. 6. EM plot of all observations

6. CONCLUSION

In this paper GMM based classifier is presented for heart sound classification after pre-processing and segmenting the heart sound. The classification scheme is based on the maximum and minimum loudness indices extracted from the spectrogram of the heart sound. GMM overcomes the limitations of Euclidean metric even though the dataset has different range of variations. The model scales down the data to zero mean and unit variance. This study highlights that classification of heart sound using GMM classifier is comparable to a state of the art classifier with high accuracy of 98.13 %. The accuracy exceeds the performance of other classifiers mentioned in literature, considering the large dataset used in this work.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Kumar D, Carvalho P, Antunes M, Paiva RP, Henriques J. Heart murmur classification with feature selection. In2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. IEEE. 2010 :4566-4569

2. Reed TR, Reed NE, Fritzson P. Heart sound analysis for symptom detection and computer-aided diagnosis. Simulation Modelling Practice and Theory. 2004;12(2):129-46.
3. Bozkurt B, Germanakis I, Stylianou Y. A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection. Computers in Biology and Medicine. 2018;100:132-43.
4. Syed Z, Leeds D, Curtis D, Nesta F, Levine RA, Guttag J. A framework for the analysis of acoustical cardiac signals, IEEE Transactions on Biomedical Engineering. 2007;54(4):651–662. DOI:10.1109/tbme.2006.889189.
5. Webster JG. The physiological measurement handbook. CRC Press, Boca Raton; 2014.
6. Uguz H. Adaptive neuro-fuzzy inference system for diagnosis of the heart valve diseases using wavelet transform with entropy Neural Comput. Appl 21. 2012a:1617-28.
7. Uguz H A. biomedical system based on artificial neural network and principal component analysis for diagnosis of the heart valve diseases J. Med. Syst. 2012b;36:61-72.
8. Zheng Y N, Guo X M and Ding X R. A novel hybrid energy fraction and entropy-based approach for systolic heart murmurs identification Expert Syst. Appl. 2015;42:2710-21.
9. Patidar S, Pachori RB, Garg N. Automatic diagnosis of septal defects based on tunable-Q wavelet transform of cardiac sound signals Expert Syst. Appl. 2015;42:3315-26.
10. Gharehbaghi A, Ekman I, Ask P, Nylander E and Janerot-Sjoberg B 2015 Assessment of aortic valve stenosis severity using intelligent phonocardiography Int. J. Cardiol. 2012;198:58-60.
11. Gharehbaghi A, Dutoir T, Sepehri A, Hult P and Ask P An automatic tool for pediatric heart sounds segmentation Computing in Cardiology (Hangzhou: IEEE) 2011 :37–40.
12. Saracoglu R. Hidden Markov model-based classification of heart valve disease with PCA for dimension reduction Eng. Appl. Artif. Intell. 25:1523-8.
13. Quiceno-Manrique AF, Godino-Llorente JI, Blanco-Velasco M, Castellanos-Dominguez G 2010 Selection of dynamic features based on time-frequency representations for heart murmur detection from phonocardiographic signals Ann. Biomed. Eng. 38:118-37.
14. Available: <u>www.physionet.org/challenge/2016/:accessed</u> date: March 1 2017.
15. Syed Z, Leeds D, Curtis D, Nesta F, Levine RA, Guttag J. A framework for the analysis of acoustical cardiac signals. IEEE Transactions on Biomedical Engineering. 2007;54(4):651–662. DOI:10.1109/tbme.2006.889189.
16. Faizan J, Venkatachalam PA, Ahmad Fadzil MH. A signal processing module for the analysis of heart sounds and heart murmurs. Journal of Physics. Conf. Ser. 2006;34(1):1098.
17. Tang H, Li T, Qiu T, Park Y. Segmentation of heart sounds based on dynamic clustering. Biomedical Signal Processing and Control. 2012;7(5):509 –516. DOI: 10.1016/j.bspc.2011.09.002.
18. Vishwanath Shervegar M, Ganesh V Bhat, Automatic segmentation of Phonocardiogram using the occurrence of the cardiac events, Informatics in Medicine Unlocked. 2017;9:6-10. Available: http://dx.doi.org/10.1016/j.imu.2017.05.002
19. Zwicker E, Fastl H. Psychoacoustics: Facts and Models, Springer Verlag, Berlin, 2nd Edition; 1999.
20. Bilmes JA. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models, Technical Report, U.C. Berkeley, April; 1998.
21. Zhong S. Probabilistic model based clustering of complex data, Ph.D. Thesis, University of Texas at Austin, August; 2003.
22. Ari S, Hembram K, Saha G. Detection of cardiac abnormality from PCG signal using LMS based least square SVM classier Expert Syst. Appl. 2010;37:8019-26.
23. Avendano-Valencia LD, Godino-Llorente JI, Blanco-Velasco M and Castellanos-Dominguez G Feature extraction from parametric time-frequency representations for heart murmur detection Ann. Biomed. Eng. 2010;38:2716-32.

Biography of author(s)

Vishwanath Madhava Shervegar
E&CE Department, Mangalore Institute of Technology & Engineering, Moodbidre, DK, Karnataka, India.

He works as an Associate Professor in Mangalore Institute of Technology and Engineering, Moodbidre, Karnataka INDIA. He holds PhD from Visvesvaraya Technological University, Belagavi, Karnataka, INDIA. He has published few papers in refereed journals such as Elsevier. He is also the reviewer in a few of the Journals from Elsevier and Springer. He is also in the Editorial Board of an Informatics in Medicine Unlocked Journal from Elsevier. Apart from Paper publication he also published a patent on Optical sensing device for Cardiac Auscultation.

DISCLAIMER
This chapter is an extended version of the article published by the same author(s) in the following journal.
Shervegar and Bhat Porto Biomed. J., 3:1, 2018.