2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Wavelet based event detection in the phonocardiogram of prolapsed mitral valve

Madhurima Patra[1], Rajeshwari BS[2], Arnab Sengupta[3], Amit Patra[2], Nirmalya Ghosh[2]

***Abstract*—Mitral valve prolapse (MVP) is one of the cardiovascular valve abnormalities that occurs due to the stretching of mitral valve leaflets, which develops in around 2 percent of the population. MVP is usually detected via auscultation and diagnosed with an echocardiogram, which is an expensive procedure. The characteristic auscultatory finding in MVP is a mid-to-late systolic click which is usually followed by a high-pitched systolic murmur. These can be easily detected on a phonocardiogram which is a graphical representation of the auscultatory signal. In this paper, we have proposed a method to automatically identify patterns in the PCG that can help in diagnosing MVP as well as monitor its progression into Mitral Regurgitation. In the proposed methodology the systolic part, which is the region of interest here, is isolated by preprocessing and thresholded Teager-Kaiser energy envelope of the signal. Scalogram images of the systole part are obtained by applying continuous wavelet transform. These scalograms are used to train the convolutional neural network (CNN). A two-layer CNN could identify the event patterns with nearly 100% accuracy on the test dataset with varying sizes (20% - 40% of the entire data). The proposed method shows potential in the quick screening of MVP patients.**

***Index Terms*—Valvular disease, Mitral valve prolapse, Systolic clicks, Scalogram, Convolutional neural network.**

I. Introduction

The human heart has four valves: Mitral and Tricuspid valve (to regulate blood flow from the atria to the ventricles) and Pulmonic and Aortic valves (to regulate blood flow from the ventricles to their respective arteries). Valvular heart diseases (VHDs) occur when any of these valves are diseased or damaged. The valves either become narrow (‘stenotic’) or flappy (‘regurgitant’) or both, which hampers their normal functioning. Depending upon the site and type of the lesion, VHDs can lead to further complications like arrhythmias, heart failure, stroke, pulmonary embolism, and coronary artery disease. VHD patients can only be treated by valve replacement surgeries, and an early and definitive diagnosis is essential to reduce mortality and morbidity [1].

Listening to the sounds of the body, usually with the aid of a stethoscope, is called auscultation. There are different areas of the chest wall from which different valvular heart sounds can be auscultated, and because each disease has a characteristic pattern of sounds, auscultation provides a first step in the clinical diagnosis of the disease. Damaged valves disrupt the normal laminar flow of blood through them and produce abnormal heart sounds called ‘murmurs’. The nature and severity of the valvular disease determine the quality and timing of the murmur [2]. However, the accuracy of traditional auscultation has been called into question multiple times [3] [4] [5]. Alternatively, a microphone can amplify the heart sounds and record as phonocardiogram (PCG) [6]. And then computers might be trained to automatically detect and categorize murmurs according to the disease. The present study makes one such attempt to identify early ejection clicks and murmurs in the PCG signals of MVP patients recorded using an electronic stethoscope.

Mitral valve prolapse (MVP) occurs due to the deposition of excessive and redundant tissue in the mitral valve apparatus. This leads to the prolapse of the mitral leaflets into the atrium during ventricular contraction. According to Harrison’s textbook of internal medicine, [1] the auscultation of MVP reveals a mid to late systolic click (which occurs 0.14 seconds or more after the first heart sound, S1) and is thought to be due to sudden tensing of the chordae tendinae or the leaflet as it reaches its maximal point during prolapse. It is usually followed by a mid to late systolic high pitched murmur, which is due to the blood regurgitating back into the atrium [1]. Different maneuvers can be used to differentiate the murmur of MVP from other valvular diseases. MVP is also the most common abnormality leading to primary Mitral Regurgitation (MR) [1]. Usually, patients with MVP are not subjected to mitral valve repair or replacement surgeries unless they are symptomatic or develop MR. Thus, early detection of an MVP evolving into an MR allows clinicians to act in time to repair the damage. Attempts have been made to detect EC in children using time growing neural network [7].

In the present study, we have made an attempt to help doctors in diagnosing MVP by identifying EC and murmurs using continuous wavelet transform. Wavelet has been in use for a long time for signal denoising, detecting transients in signal, and signal compression. Wavelets could also detect and classify the first and second cardiac sounds (S1 and S2) [7], [8], [9]. Current study uses scalograms and a machine learning approach to identify and categorize three distinct patterns of MVP heart sounds - systolic ejection clicks (EC), systolic ejection click with late systolic murmur, and systolic click with mid systolic murmur.

Convolutional neural network (CNN) architecture is

[1] Madhurima Patra is with Kalyani Govt. Engineering College West Bengal, India madhurimap811@gmail.com

[2] Rajeshwari BS, Dr. Nirmalya Ghosh and Prof. Amit Patra are with Department of Electrical Engineering, Indian Institute of Technology, Kharagpur, West Bengal-721302 rajeshwaribs@iitkgp.ac.in, {nirmalya,amit}@ee.iitkgp.ac.in

[3] Arnab Sengupta, MBBS is currently with Queen Square Institute of Neurology, University College London. arnabs0411@gmail.comm

inspired by the coordination and connectivity pattern of neurons within the human brain. For computers, images are a numerical matrix of values representing color intensity of each pixel. The CNN architecture uses images as input, specific layers to identify patterns and features and predict class labels from the extracted features. The elementary layers (conv, pooling and Rectified Linear Unit-ReLu) identify simple to complex features progressively, remove redundant spatial information, and reduce the spatial size of the generated feature maps. The last few layers (fully connected layer and softmax layer) of the network integrate all the detected features and perform classification. While training, the network moves back and forth in accordance with feedforward and backpropagation algorithm to acquire the finest possible accuracy [10] [11].

Paper organization: Section II describes the Methodology to segment PCG, extract region of interest, generate scalogram images and classify using a CNN. Results and Discussions are detailed in Sections III-B and IV respectively, followed by Conclusion in Section V .

II. METHODOLOGY

MVP can be present with a combination of systolic click and murmur complex. In this study, three distinct patterns were identified and labeled by a clinician (Arnab Sengupta, MBBS) which are - EC, EC with late systolic murmur, and EC with mid systolic murmur. These patterns can help in the diagnosis of MVP as well as monitor its progression into MR. Since the three events appear in the systole part, the region of interest is the systolic part in each cycle. The signal was preprocessed, denoised and segmented to extract the systole part of the signal. Then applying continuous wavelet transform on the selected systolic portion, scalograms were generated. These scalograms were used as input to train the CNN for event pattern detection. The entire process is depicted in Figure 1.

A. Pre-processing

PCG signals were down-sampled by a factor of 2 to reduce computational load, followed by normalization by the maximum signal value to minimize the effect of variation in amplitude. Downsampling by a factor of 2 will reduce the computational complexity and the frequency ranges of the filter banks used for scalogram image generation. The PCG signal will not have useful information in the frequency range above 1000Hz. Downsampling the signal by 2 will reduce the sampling frequency to 4kHz and the filter banks will be defined in the range of 0 to 2kHz, which helps to capture the time-frequency pattern without loss of information. Hence this downsampling reduces the computational demand of the method. The high-frequency murmur components were removed using 10^{th} order low-pass Butterworth filter with a cutoff frequency of 300 Hz.

B. Wavelet denoising

Even after filtering, any remaining murmur signal was removed by wavelet denoising. The performance of discrete wavelet denoising depends on the selection of: (1) mother wavelet, (2) number of decomposition levels, (3) detail and approximation levels for reconstruction, (4) threshold value applied on wavelet coefficients, (5) threshold function [12].

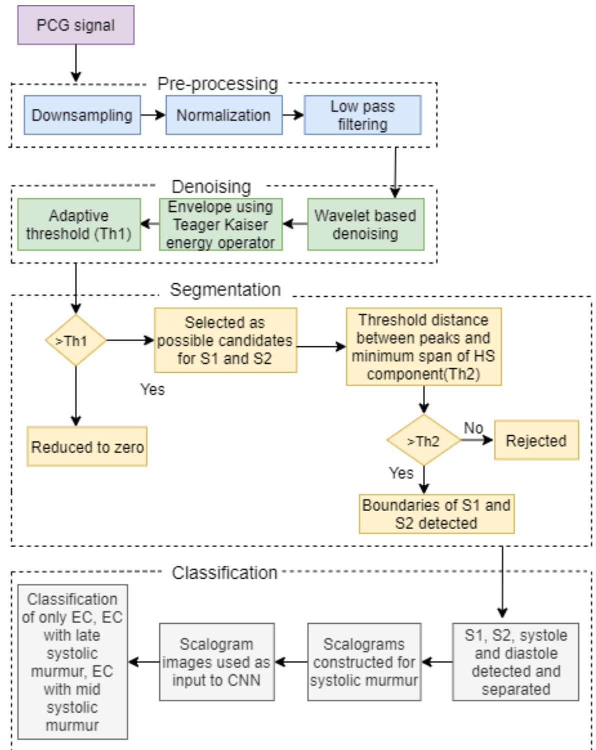

Fig. 1: Proposed methodology for MVP event detection in PCG heart sounds(HS).

For PCG signal denoising, various orthogonal wavelet families like *Coiflet*, *Daubechies*, *Symlet* have been used [12]. Based on the highest cross-correlation values between scaled wavelet and signal, in the proposed algorithm, *Sym20* has been used as the mother wavelet with decomposition up to 5 levels, resulting in five detail level coefficients and one approximation level (0-62.5Hz). The 3^{rd}, 4^{th} and 5^{th} level detail coefficients span most of the frequency range of heart sound components and were thus used for reconstruction of the signal as shown in Figure 2(c). For threshold estimation, *rigrsure* method was used which computes a low threshold value, followed by soft threshold function where coefficients lower than the threshold are set to zero, and remaining coefficients are shrunk towards the threshold value causing no discontinuity in the resulting signal [13] [14].

C. Teager-Kaiser energy operator

Envelope computation helps detecting the onset and offset of signal segments. To identify distinct patterns in MVP, onset and offset of the region of interest were identified using the envelope of the denoised signal computed with Teager-Kaiser energy operator using Equation 1.

$$\psi[p(n)] = p[n]^2 - p[n+1]p[n-1] \quad (1)$$

A distinct advantage of using this operator is that the energy waveform is obtained using the instantaneous amplitude values (*p[n]*), providing better localization properties as compared to other commonly used envelope detection methods like Shannon energy and Hilbert transform [15].

D. Boundary detection

An adaptive threshold based on mean (μ) and standard deviation (σ) of the signal reconstructed from wavelet transform (Figure 2(c)) was computed using Equation 2.

$$Th = (\mu + 2 * \sigma)/1000 \tag{2}$$

For detecting the onset and offset of the heart sound components, threshold **Th** is applied on the envelope as follows:

If x(i)>Th: The corresponding value is retained

If x(i)<Th: The corresponding value is set to zero

Signal segments with $x(i) >$ Th are candidates for S1 and S2. The remaining extra peaks were removed using the following rules: Two consecutive peaks should be separated in time $>=$ 20% of the average heart cycle duration; and the minimum time span of the peak should be 150 samples. The remaining peaks were identified as S1 and S2 based on the fact that the systole period (S1 to S2) is generally shorter than the diastole period (S2 to S1). Using this algorithm, the fundamental heart sound components- S1, S2, systolic murmur, and diastolic part were separated. An example signal at each step of the segmentation process is illustrated in Figure 2. For MVP, the systolic part contains the ejection click and murmur. The patterns in the systolic part were used for MVP progression detection.

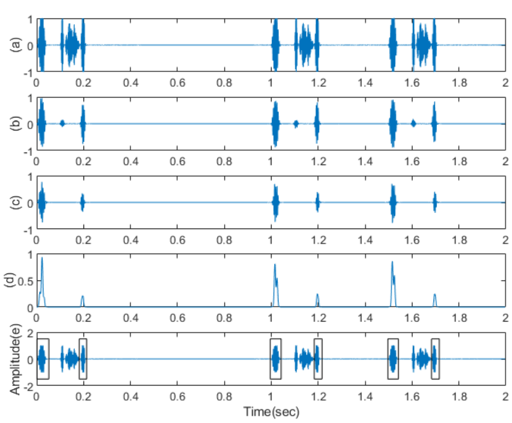

Fig. 2: (a) Original signal, (b) Low pass filtering, (c) Discrete Wavelet Transform(detail levels *d3,d4,d5*), (d) Teager-Kaiser energy operator, (e) S1, S2 identified and separated.

E. Scalogram

Continuous wavelet transform (CWT) helps in analyzing signals in the time-frequency domain. A plot of absolute values of CWT coefficients against time is the scalogram. These scalograms are preferred for better time localization of short-duration, high-frequency events, and better frequency localization of low-frequency, longer-duration events. CWT makes use of an analysis window (wavelet) which can be translated and scaled (stretched or compressed). The mathematical expression for CWT is:

$$W(p,q) = \frac{1}{\sqrt{p}} \int_{-\infty}^{\infty} x(t)\psi(\frac{(t-q)}{p}) \tag{3}$$

Equation 3 shows that a wavelet is shifted by 'q' in time and scaled by 'p' factor prior to computing correlation with the signal x(t). The CWT automatically adjusts the time and frequency resolution.

The scalogram is generated using filter banks (analytic Morse (3, 60)) with 10 'voices' per octave with center-frequency range 0 to 2000Hz. The coefficients obtained by passing the signal through filter banks generate time-frequency representation of the signal by decomposing the signal in the time-frequency plane. These coefficients in different frequency bands in log scale on y-axis and time on x-axis are scaled and written as an image of size 227x227x3 using *jet128* color scheme. These scalogram images are used as input to the shallow CNN for classification of event patterns of PCG signal in cases of prolapsed mitral valve.

F. Convolutional neural network construction

In this study, a simple 2D CNN with the architecture of Figure 3 was used for the classification of the MVP event patterns.

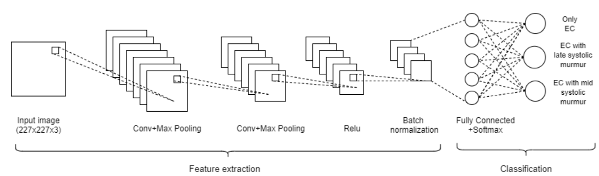

Fig. 3: CNN architecture.

III. RESULTS

A. Database

The data was obtained from an open-source database [16] and was collected from random sources, such as books (Auscultation skills CD, Heart sound made easy) and websites (48 different websites provided the data including Washington, Texas, 3M, Michigan and so on). The database was prepared by authors in [17] using Cool Edit software with PCG signal of each subject having 3 cardiac cycle and was sampled at 8kHz. The data has a total of 1000 audio files, 200 for each category of Aortic Stenosis, Mitral Regurgitation, Mitral Stenosis, Mitral Valve Prolapse (MVP), Healthy subjects (Normal). The present study uses only MVP data.

Ground-truth was generated by a clinician after visual and auditory examination of PCG signals. Few signals with wrong class labels or with less than 2 cardiac cycles were removed from the study. Systolic part from MVP subjects were identified with three classes: Class 1 (78), Class 2 (93), Class 3 (75) and were manually marked based on the patterns as depicted in Figure 4 and used as ground-truth for supervised learning and performance evaluation.

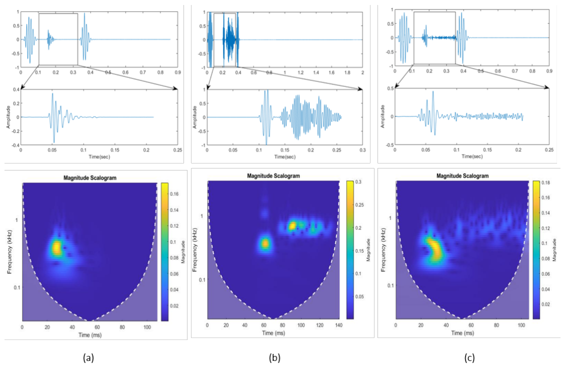

Fig. 4: Systolic part of one cardiac cycle of PCG signal and it's scalogram: (a) Only EC, (b) EC with late systolic murmur, (c) EC with mid systolic murmur.

B. Model evaluation

CNN was trained and tested for individual cardiac cycles with the data split into training, validation, and testing sets at patient level such that no subject is present in more than one dataset to remove any bias. To avoid over fitting the model is generalized with different split ratios of training, testing, and validation sets. CNN performance for test dataset is summarized in Table I.

TABLE I: Classifier performance

Parameters	**Data distribution:Training-Validation-Testing**			
	50-10-40	**50-20-30**	**50-25-25**	**60-20-20**
Accuracy	98.98%	100%	100%	100%
Sensitivity	97.30%	100%	100%	100%
Specificity	98.53%	100%	100%	100%
Precision	96.77%	100%	100%	100%
FScore	98.63 %	100%	100%	100%

IV. Discussion

The results demonstrate that the proposed method performs excellently even for a low proportion of the training set. This also highlights the generalization characteristic of the classifier. While manual ground-truth generation took 4 hours, the proposed automated method took $\simeq$3 minutes for training the CNN in an 8GB RAM workstation. Hence, the proposed method can easily be translated into clinical applications.

Diagnosing an MVP by listening to the murmur can be difficult. Maneuvers that increase the preload (like squatting, isometric exercises, etc.) decrease the intensity of murmur and cause the click-murmur complex to shift farther from S1 or even disappear; whereas maneuvers that decrease preload demonstrate the opposite. This gives an important clue towards the diagnosis of MVP. Thus, by automatically detecting the changes in the click-murmur complex with these maneuvers, this system will help the clinician in making a diagnosis. Furthermore, there is a high probability that MVP might progress into an MR at which point surgical treatment becomes necessary. This system can thus help in the follow-up of a patient with early MVP as it progresses into an MR (where the regurgitant murmur will become more prominent). It can alert the physician to changes in the murmur that might require further evaluation.

V. Conclusion

This present study uses wavelet transform and CNN to identify and categorize EC, EC with late systolic murmur, and EC with mid-systolic murmur in MVP. This will not only help physicians to diagnose this disease but also allow proper monitoring of the disease. An extension of this system for all valvular diseases will be a powerful tool in the clinical realm. It would allow physicians to accurately diagnose and follow-up these patients. Moreover, if this system was coupled to the prudent detector or packaged on a mobile platform, it can potentially be used as a replacement for echocardiogram in low resource conditions.

References

[1] E. Braunwald, A. S. Fauci, D. L. Kasper, S. L. Hauser, D. L. Longo, and J. L. Jameson, *Harrison's principles of internal medicine*. McGraw Hill, 2001.

[2] V. Kumar, A. K. Abbas, and J. C. Aster, *Robbins basic pathology e-book*. Elsevier Health Sciences, 2017.

[3] S. K. Gardezi, S. G. Myerson, J. Chambers, S. Coffey, J. d'Arcy, F. R. Hobbs, J. Holt, A. Kennedy, M. Loudon, A. Prendergast, *et al.*, "Cardiac auscultation poorly predicts the presence of valvular heart disease in asymptomatic primary care patients," *Heart*, vol. 104, no. 22, pp. 1832–1835, 2018.

[4] M. Mehta, T. Jacobson, D. Peters, E. Le, S. Chadderdon, A. J. Allen, A. B. Caughey, and S. Kaul, "Handheld ultrasound versus physical examination in patients referred for transthoracic echocardiography for a suspected cardiac condition," *JAAC: Cardiovascular Imaging*, vol. 7, no. 10, pp. 983–990, 2014.

[5] S. Mangoine and L. Nieman, "Cardiac auscultatory skills of internal medicine and family practice trainees," *Journal of the American Medical Association*, vol. 278, pp. 717–22, 1997.

[6] A. C. Guyton and J. E. Hall, "Textbook of medical physiology 11th ed," *Elsiever Saunders*, pp. 788–817, 2006.

[7] A. Gharehbaghi, T. Dutoit, P. Ask, and L. Sörnmo, "Detection of systolic ejection click using time growing neural network," *Medical engineering & physics*, vol. 36, no. 4, pp. 477–483, 2014.

[8] P. Vikhe, S. Hamde, and N. Nehe, "Wavelet transform based abnormality analysis of heart sound," in *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*. IEEE, 2009, pp. 367–371.

[9] J. Chebil and J. Al-Nabulsi, "Classification of heart sound signals using discrete wavelet analysis," *International Journal of Soft Computing*, vol. 2, no. 1, pp. 37–41, 2007.

[10] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*. IEEE, 2017, pp. 1–6.

[11] I. Arel, D. C. Rose, and T. P. Karnowski, "Deep machine learning-a new frontier in artificial intelligence research [research frontier]," *IEEE computational intelligence magazine*, vol. 5, no. 4, pp. 13–18, 2010.

[12] P. K. Jain and A. K. Tiwari, "An adaptive thresholding method for the wavelet based denoising of phonocardiogram signal," *Biomedical Signal Processing and Control*, vol. 38, pp. 388–399, 2017.

[13] D. L. Donoho, "De-noising by soft-thresholding," *IEEE transactions on information theory*, vol. 41, no. 3, pp. 613–627, 1995.

[14] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[15] P. K. Banerjee and N. B. Chakrabarti, "Noise sensitivity of teager-kaiser energy operators and their ratios," in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 2015, pp. 2265–2271.

[16] [Online]. Available: https://github.com/yaseen21khan/Classification-of-Heart-Sound-Signal-Using-Multiple-Features-/blob/master/README.md

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