A Pre-screening Technique for Coronary Artery Disease with Multi-channel Phonocardiography and Electrocardiography

Yue Rong*, Matthew Fynn, and Sven Nordholm

Abstract

In this chapter, we present a new noninvasive coronary artery disease (CAD) pre-screening technique using multi-channel phonocardiography (PCG) and electrocardiography (ECG). After a brief overview of the state-of-the-art in PCG, we present a multichannel PCG and ECG instrument. With the aid of a background-noise microphone integrated into each digital stethoscope, adaptive signal processing methods are then proposed to suppress the environmental noise to improve the integrity of the recorded heartbeat signal. Finally, we investigate important factors that can impact the performance of a neural network based CAD classifier using heart auscultation. Among them, the integrity of the heartbeat signal plays a key role.

Index Terms

Auscultation, coronary artery disease, cardiovascular disease, electrocardiography, phonocardiography.

I. Introduction

Cardiovascular disease (CVD) including coronary artery disease (CAD) is the leading cause of mortality and morbidity in the world [1], contributing 31% towards all global deaths. Early diagnosis of CAD is important to prevent further development of this disease. Standard methods for diagnosis of CAD such as coronary angiography and myocardial perfusion imaging require specialised equipment and clinical setting. Although these methods are effective in diagnosing CAD, they are highly costly and expose patients to radiation. Table I shows the performance of different diagnostic methods in Australia in 2012 [2]. The costs were retrieved from http://www9.health.gov.au/mbs in February, 2023, with item numbers 11704, 55132, 61324, 57360, 38310 for the five stress tests in order, respectively.

On the other hand, heart auscultation (the interpretation of heart sounds by a physician) is a cost-effective tool for the pre-screening of CAD. It is well established that partial obstruction of coronary arteries causes disruption of normal, laminar flow and generates flow turbulence and so modifies heart sounds. When the coronary arteries blockage is prevalent, murmur sounds frequently occur. However, auscultation skills are difficulty to acquire. Since the heart sound acquired by a stethoscope is often contaminated by various internal (e.g. breathing) and external (e.g. friction) noises, it can be hard for the human auditory system to identify abnormal heart sounds related to CAD. With the aid of computer technology and highly sensitive electronics, digital stethoscopes can be used to detect sounds that are below the human hearing threshold and frequency range. As a result, phonocardiogram (PCG) signal processing combined with computer-aided classification attract much interest over the last decade as a low-cost and noninvasive tool for prescreening of CAD.

In this chapter, we present a multi-channel phonocardiography (PCG) and electrocardiography (ECG) based CAD pre-screening technique, developed by researchers at Curtin University in collaboration with Ticking Heart, a health-tech start-up. With the aid of a background-noise microphone integrated into each digital stethoscope, adaptive signal processing methods are proposed to suppress the environmental noise to improve the integrity of the recorded heartbeat signal. We investigate important factors that can impact the performance of a neural network based CAD classifier using heart auscultation.

Y. Rong is with the School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, WA 6102, Australia (e-mail: y.rong@curtin.edu.au).

TABLE I
DIAGNOSTIC METHODS FOR CORONARY ARTERY DISEASE DETECTION

Stress Test	Cost $	Sensit. %	Specif. %	Advantages	Disadvantages
Electrocardiography	33	68	70-77	1) Assessment of exercise capacity 2) Cost effective 3) First line test in absence of contraindications	1) Lowest sensitivity of all stress tests: risk of false negative test 2) Lower diagnostic accuracy in women
Echocardiography (exercise)	240	80-85	84-86	1) Assessment of exercise capacity, cardiac structure/function 2) No radiation 3) High specificity	False negatives in single vessel/circumflex territory ischaemia (increased sensitivity with cycle ergometry)
Nuclear perfusion study	653	85-90	70-75	1) Exercise capacity can be assessed 2) High sensitivity	1) Radiation 2) False positives due to higher sensitivity/diaphragmatic attenuation
CT coronary angiogram	728	85-99	64-90	1) High negative predictive value (in low to intermediate risk subjects)	1) Radiation 2) Functional effect of stenosis not usually assessed, nor exercise capacity
Coronary angiogram	2438	≈ 100	≈ 100	Gold standard	1) Invasive 2) Radiation 3) Functional effect of stenosis not routinely assessed

The rest of this chapter is organized as follows. In Section II, a brief overview of the state-of-the-art PCG is provided. An introduction of the multichannel PCG and ECG instrument is presented in Section III. Adaptive PCG noise cancellation algorithm is proposed in Section IV to maintain the integrity of the signal-of-interest. Key factors affecting machine learning based CAD classification algorithms are discussed in Section V. Conclusions are drawn in Section VI.

A list of acronyms used in this chapter is given below.

ANN	artificial neural network
BNM	background-noise microphone
CAD	coronary artery disease
CIC	Computing in Cardiology
CNN	convolutional neural network
CSA	cardiac sonospectrographic analyzer
CVD	cardiovascular disease
ECG	electrocardiography
FIR	finite impulse response
FL	filter length
HM	heart-sensor microphone
HSMM	hidden semi-Markov model
LTI	linear time-invariant
MLP	multilayer perceptron
MSE	mean-squared error
NLMS	normalized least mean squares
NN	nearest neighbor
PCG	phonocardiography
PSD	power spectral density
SST	synchrosqueezing transform
SVM	support vector machine

II. OVERVIEW OF THE STATE-OF-THE-ART PCG

PCG signals are digital recordings of heart sound. They provide a convenient primary diagnostic tool for detecting CAD murmurs [3]. Figure 1 displays a PCG signal of a healthy subject, where the regions of systole and diastole, and major heart sounds S1 and S2 are shown.

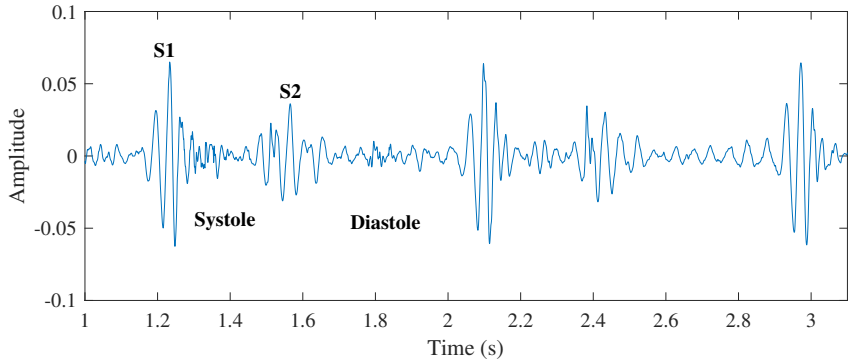

Fig. 1. Phonocardiogram of a healthy subject.

Murmur sounds from the coronary arteries blockage affect the basic characteristics of PCG signals, which provides the opportunity to detect these changes using machine learning techniques. Many algorithms have been proposed to classify normal and abnormal heart cycles, aiming at achieving high accuracy, sensitivity, and specificity from their data sets, which are defined below

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + F_p + F_n + T_n} \tag{1}$$

$$\text{Sensitivity} = \frac{T_p}{T_p + F_n} \tag{2}$$

$$\text{Specificity} = \frac{T_n}{T_n + F_p}. \tag{3}$$

Here T_p, T_n, F_p, and F_n represent the number of true positives, true negatives, false positives, and false negatives, respectively.

Most PCG-based CAD detection techniques follow a similar set of methodology, which consists of extracting features from PCG data for the machine learning algorithms to learn. Multiple studies have undertaken time-frequency analysis techniques to extract feature vectors from the heart sound signal. In [4], an artificial neural network (ANN) model was applied to classify heart sounds using wavelet based feature extraction. Based on the ANN classification model, a computer-aided diagnosis system was described in [5] for multiple pathological cases using wavelet decomposition. Support vector machine (SVM) along with wavelet packet decomposition have been applied in [6] to detect valvular heart sounds as normal or abnormal. In a 2010 study [7], a method was proposed to improve the performance of least-squares SVM to diagnose pathological sounds.

A digital electronic stethoscope named the cardiac sonospectrographic analyzer (CSA) was tested in [8] for the detection of coronary artery microbruits, where the CSA showed high sensitivity and specificity for the detection of significant early CAD in an outpatient setting. In [9], nine different types of acoustic features from five overlapping frequency bands were obtained and analysed for the identification of murmur sounds associated with CAD. The result confirms that there is a potential in PCG for the diagnosis of CAD. A dual-input neural network for CAD detection using both ECG and PCG was developed in [10].

Access to labelled heart sound recordings plays a key role in developing machine learning based classification algorithms. A list of the currently available open access data sets is provided in a recent paper [11]. Among these data sets, a large public database was created for the 2016 PhysioNet/Computing in Cardiology (CIC) challenge to classify normal and abnormal heart sound recordings [12]. In this database, there are six different data sets available, where each set is composed of heartbeat measurements recorded by different research groups at different institutions. Each recording is labelled as ‘normal’ or ‘abnormal’ according to expert diagnosis. A currently largest pediatric heart sound data set is presented in [11], which

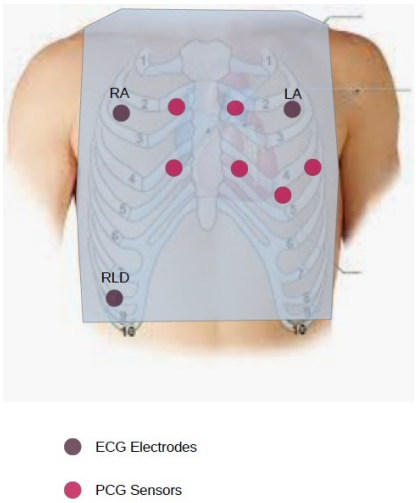

Fig. 2. A vest holding six digital stethoscopes and one three-led ECG sensor (RA, LA and RLD stand for right arm, left arm and right leg drive, respectively)

Fig. 3. Single stethoscope. Left: HM facing upward. Right: BNM facing upward.

contains $215,780$ manually annotated heart sounds. In addition to binary labels, each murmur has been manually annotated by an expert annotator according to its timing, shape, pitch, grading, and quality. The data set [11] is used in the 2022 PhysioNet challenge in detecting abnormal heart function from multi-location PCG recordings of heart sounds [13].

An open access simultaneously recorded PCG and ECG database is presented recently in [14], where the recording device includes circuitry for three-lead ECG, two digital stethoscope channels for PCG acquisition, and two auxiliary channels to capture the ambient noise. To the best of our knowledge, there is no open access database with multichannel (more than two channels) PCG and ECG recordings. In the next section, we present a multichannel PCG and ECG instrument.

III. Multichannel PCG and ECG instrument

A measurement system has been built by Ticking Heart [15], a health-tech start-up, that incorporates six digital stethoscopes and one three-lead ECG sensor onto a wearable vest that simultaneously measures heartbeat signals and applies machine learning methodologies for pre-screening CAD. The digital stethoscopes are placed in clinically advised positions, see Figure 2. In particular, four stethoscopes are located on the left side of the chest to detect sounds from the mitral, aortic, pulmonary, and tricuspid valves. The other two PCG sensors are placed on the right side of the chest to detect sounds from the ascending aorta artery. Compared with systems having only a single stethoscope or two stethoscopes [14], using multiple stethoscopes can improve the performance of classification [16]. Moreover, multichannel PCG and ECG signal processing has potential to separate the microphone signals with a stronger mapping to the cause of the signal waveform.

As shown in Figure 3, each stethoscope has two microphones, where one microphone is the heart-sensor microphone (HM) located behind the diaphragm, and the other one is the background-noise microphone

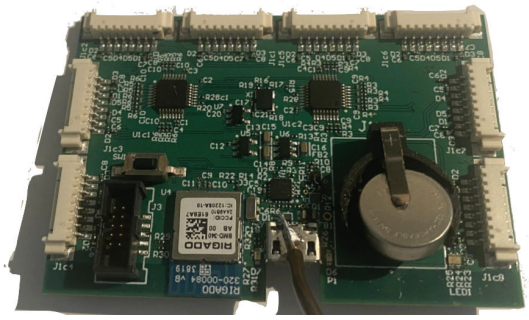

Fig. 4. Multichannel PCG and ECG data collection board.

(BNM) located at the other end of the stethoscope. The HM acquires the heart signal plus part of the background noise, while the BNM mainly picks up the background noise. Using such two-microphone configuration, the background noise can be reduced from the HM, which contributes to successful diagnosis techniques, as acquired signals from the system are cleaner with higher signal integrity. All signals from the stethoscopes and the ECG sensor are routed to a data collection board as shown in Figure 4, for further analog signal conditioning prior to digitalization. Note that Figures 3 and 4 are not to scale.

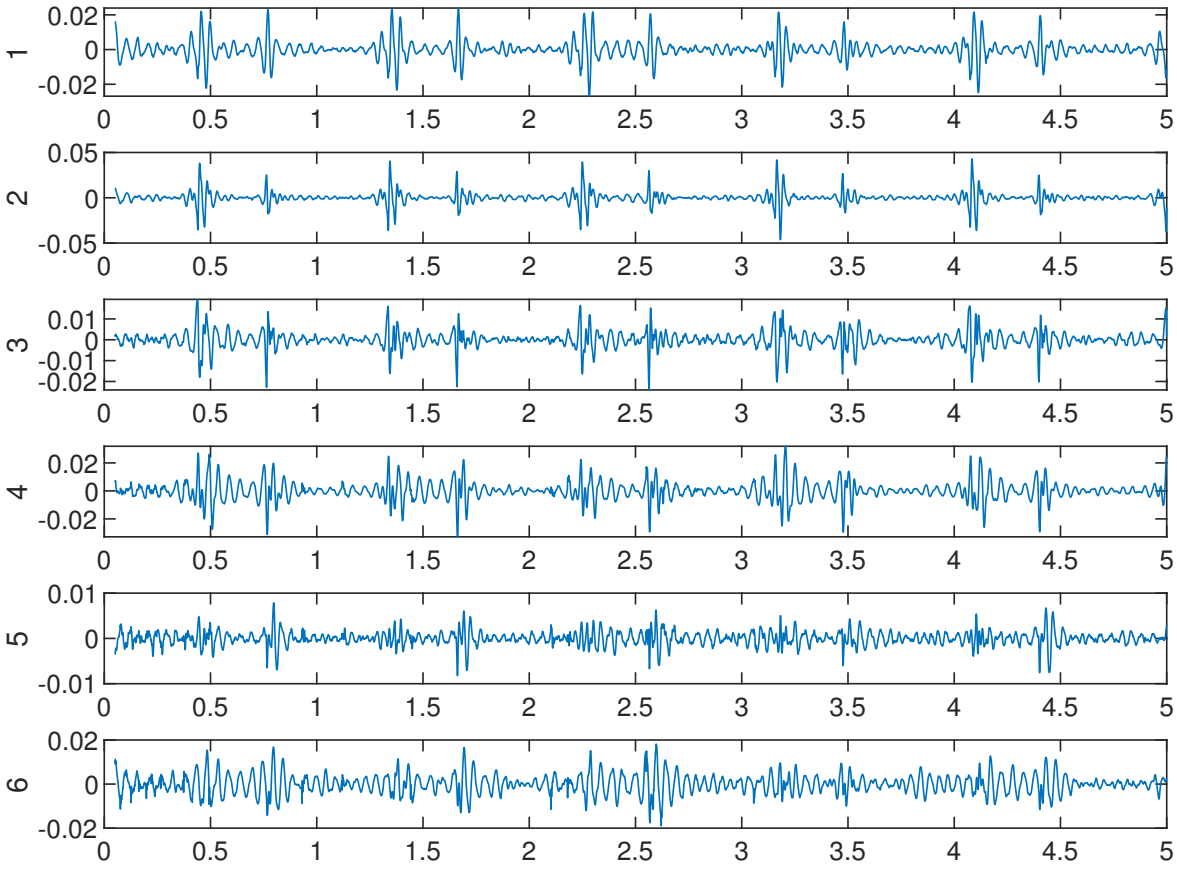

Fig. 5. Waveform of heartbeat signals acquired by the HMs in six digital stethoscopes located in Figure 6. The x-axis is time in seconds and the y-axis represents the amplitude.

An example of the PCG signals measured simultaneously from six digital stethoscopes held by TH's wearable vest is given in Figure 5, which shows the waveforms of the signals acquired by the HM of

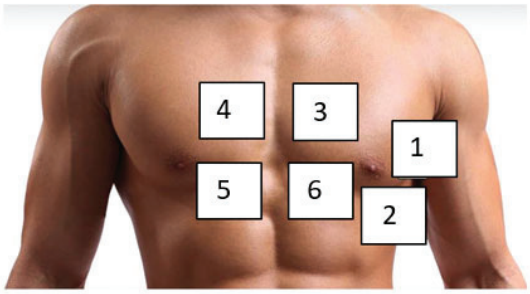

Fig. 6. Positions of the digital stethoscopes for the PCG signals shown in Figure 5.

each stethoscope sampled at 2 kHz and passed through a high-pass filter with a cutoff frequency at 20 Hz. Thus, the frequency band of signals in Figure 5 is from 20 Hz to 1 kHz. The positions of the six stethoscopes on the chest are displayed in Figure 6. Note that the six HM signals are synchronized at the signal sample level. We can observe that the signal from Sensor 2 has the highest amplitude, where the pattern of S1 and S2 can be clearly seen, while Sensor 5 yields the weakest signal as it is furthest away from the heart. An advantage of multichannel synchronized PCG signals is that they can provide more information about the heartbeat and murmur sounds, which can be utilized by machine learning algorithms to improve the accuracy of CAD classification.

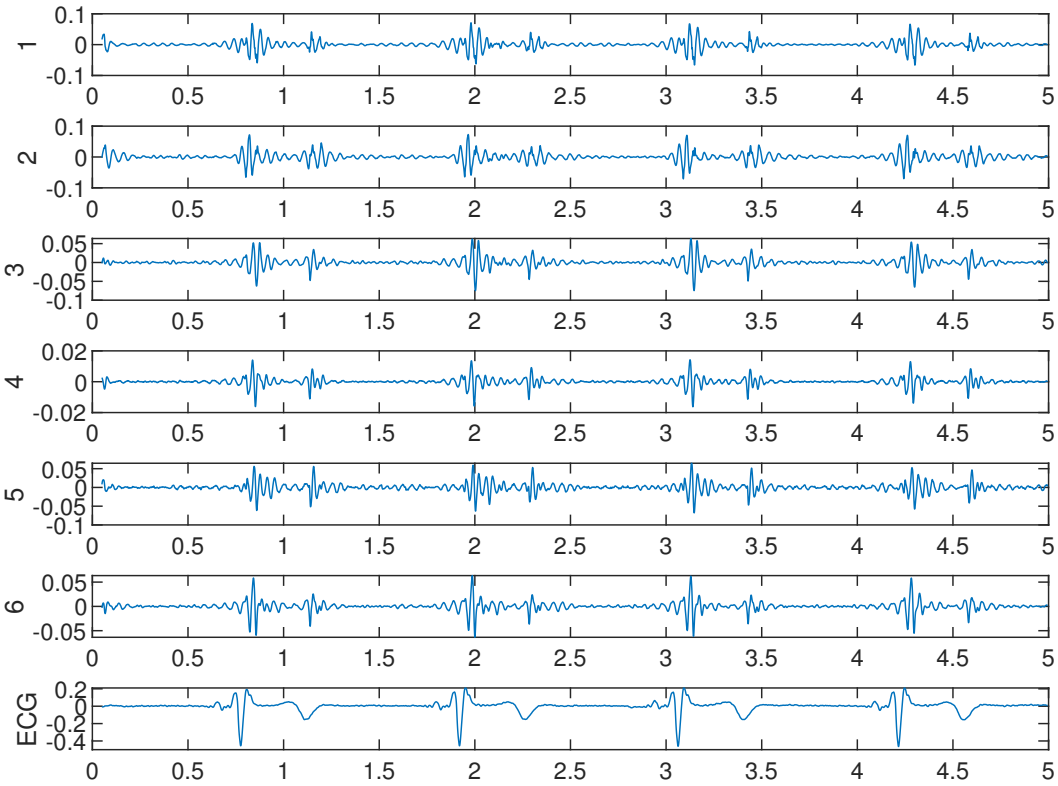

Fig. 7. Waveform of PCG signals and ECG signal recorded simultaneously. The x-axis is time in seconds and the y-axis represents the amplitude.

Waveforms of six PCG signals and one ECG signal recorded simultaneously are shown in Figure 7. Synchronization at the signal sample level is achieved through the same multichannel analogue-to-digital converter. We can observe that in one heart cycle, the first peak in ECG (i.e., the R peak) appears slightly ahead of the S1 peak in the PCG signals, while the second peak in ECG coincides with the S2 peak in the

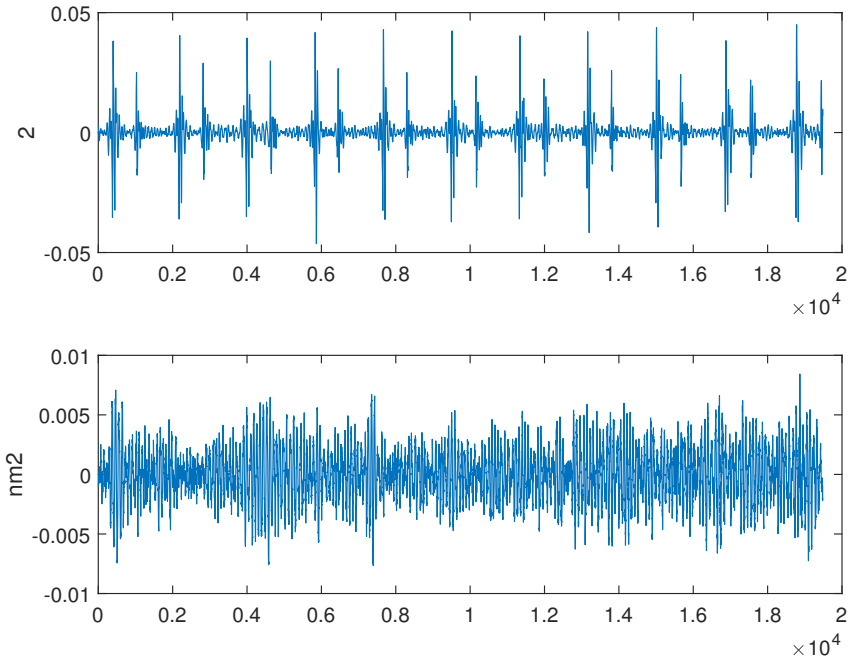

Fig. 8. Waveform of the HM (top) and BNM (bottom) signals from Sensor 2.

PCG signals. In the case of Figure 7, the ECG signal appears to be less noisy than PCG signals. Jointly processing the ECG and PCG signals would provide more information on the feature of the heartbeat, which can feed into a classifier to improve the performance of CAD classification.

Figure 8 shows the waveform of the signals acquired by the HM and BNM on Sensor 2 in Figure 5. We can see that since the signals were recorded in a quiet office environment, the BNM signal (which is mainly external noise) has a much lower amplitude than the HM signal. Moreover, the pattern of S1 and S2 is not visible in the BNM signal, making it possible to cancel the noise in the HM based on the BNM signal without compromising the signal-of-interest, as discussed in detail in the next section. A clinical study is planned in 2023 to evaluate the performance of the vest in real patients.

IV. Noise cancellation

As shown in the last section, background noise can couple into the HM and corrupt acoustic heartbeat measurements during heart auscultation. This can decrease heart signal integrity, as the noise cannot be filtered out using conventional frequency-selective filters if it lies within the frequency band of interest (e.g. 10-600 Hz). By using the BNM as a reference for noise, noise cancellation filtering techniques can be applied to attenuate unwanted background noise and restore integrity to the desired signal.

A. Noise cancellation model

A noise cancellation model of the two-microphone stethoscope system is shown in Figure 9. The desired heartbeat signal to be measured $d(n)$ is corrupted by various background noise sources $v(n)$ to produce the signal acquired by the HM as $x(n) = d(n)+v_1(n)$. Without any information on $v_1(n)$, it is not possible to remove it from $d(n)$. The BNM measures the noise sources without heartbeat signal $d(n)$, depicted as $v_2(n)$. However, the HM and BNM do not detect the noise in the same way, i.e., $v_1(n) \neq v_2(n)$. This indicates that the desired signal cannot be obtained via a direct subtraction [17]. Instead, a Wiener filter produces an estimate of $v_1(n)$, denoted as $\hat{v}_1(n)$, via the observational measurements of $v_2(n)$. This

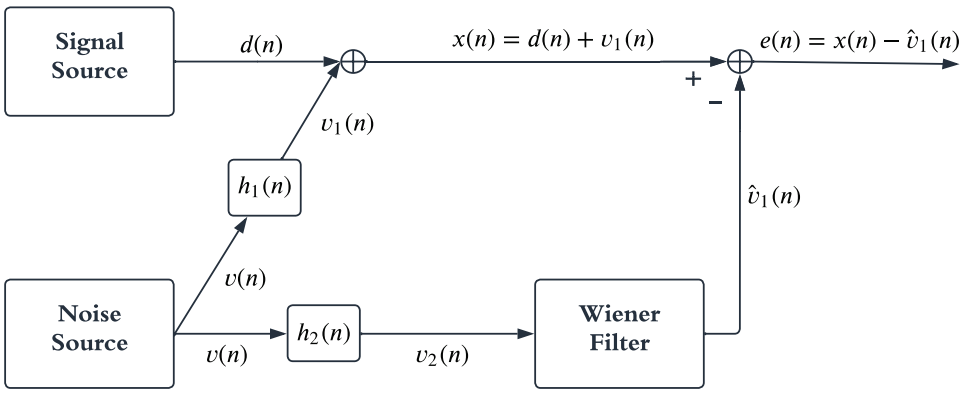

Fig. 9. Noise cancellation model of the two-microphone stethoscope.

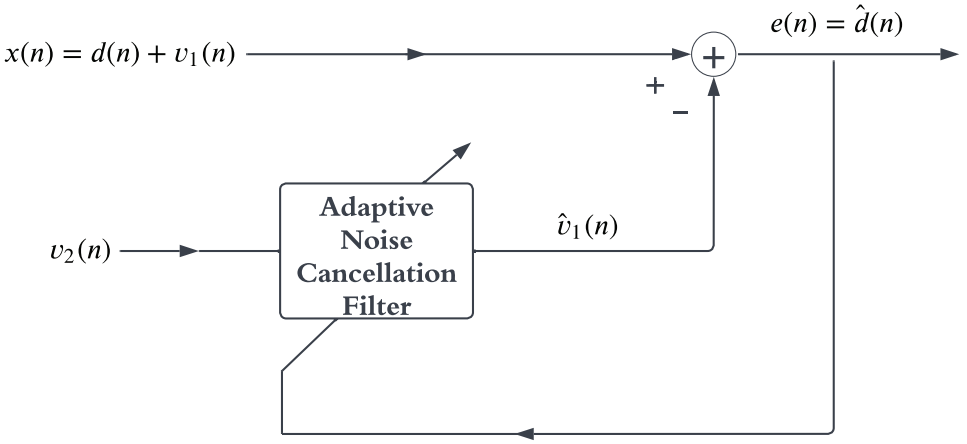

Fig. 10. Adaptive noise canceller.

estimate is subtracted from $x(n)$ to attenuate the background noise. Here, the Wiener-Hopf equation for a finite impulse response (FIR) filter is given by

$$\mathbf{R}_{v_2}\mathbf{w} = \mathbf{r}_{xv_2} \tag{4}$$

where $\mathbf{w}$ is the FIR coefficient vector of the Wiener filter, $\mathbf{R}_{v_2}$ is the autocorrelation matrix of $v_2(n)$, and $\mathbf{r}_{xv_2}$ is the cross-correlation vector of $x(n)$ and $v_2(n)$. The observed signals $x(n)$ and $v_2(n)$ are acoustic signals and their relative time delay is unknown. An extra delay is inserted in $x(n)$ to take into account non-causality [18].

In environments where the background noise is constantly changing, adaptive filters are needed to update filter coefficients $\mathbf{w}$ in (4) in real time. An adaptive filter model is shown in Figure 10, where the output of the adaptive filter aims to minimize the mean-squared error (MSE) of estimating $v_1(n)$, thus $e(n)$ is the MSE of the desired signal $d(n)$. We use a normalized least mean squares (NLMS) algorithm to update the coefficients of the adaptive filter $\mathbf{w}(i)$. For each sample $x(i)$ from the HM, the error history is given by

$$e(i) = x(i) - \mathbf{w}^T(i)\mathbf{v}_2(i) \tag{5}$$

where $\mathbf{v}_2(i) = (v_2(i), v_2(i+1), \ldots, v_2(i+N))^T$, N is the filter length (FL), and $(\cdot)^T$ denotes the matrix transpose. The filter coefficient vector is updated for each sample as

$$\mathbf{w}(i+1) = (1-\alpha)\mathbf{w}(i) + \mu\frac{\mathbf{v}_2(i)e(i)}{\|\mathbf{v}_2(i)\|^2} \tag{6}$$

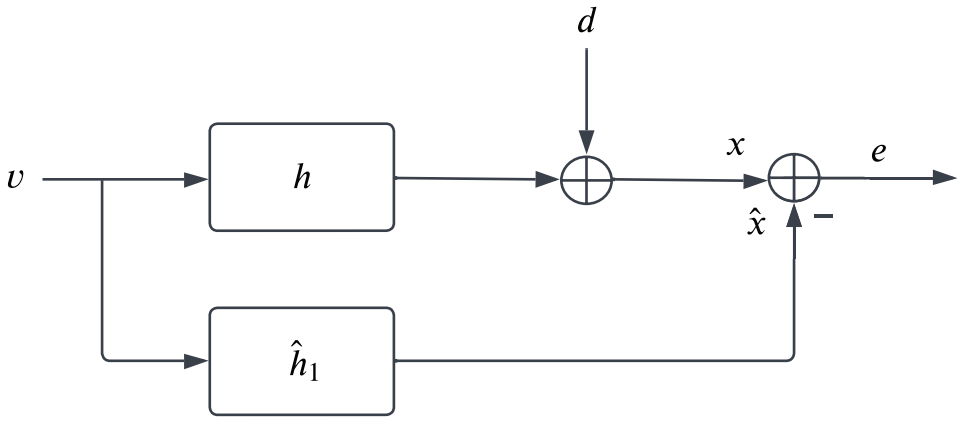

Fig. 11. Linear system model.

where α is the leakage coefficient, $\|\cdot\|$ stands for the vector Euclidean norm, and μ is the step size determining the size at which the filter coefficients are updated. The algorithm specified by (5) and (6) is a variable leaky NLMS algorithm, which has the potential to significantly outperform the standard LMS algorithm.

B. Coherence function

The performance of the NLMS based method in suppressing the background noise is closely related to the coherence function of signals from the HM and BNM. The coherence between the HM and BNM channels indicates how much noise can be attenuated at particular frequencies [19]. The coherence function of two signals $x(n)$ and $v(n)$ (also referred to as the coherence-squared function) is defined by [20]

$$\gamma_{vx}^2(f) = \frac{|S_{vx}(f)|^2}{S_{vv}(f)S_{xx}(f)} \tag{7}$$

where $S_{vv}(f)$ and $S_{xx}(f)$ are the auto spectral densities and $S_{vx}(f)$ is the cross spectral density. For all frequencies f, (7) is bounded by

$$0 \leq \gamma_{vx}^2(f) \leq 1. \tag{8}$$

If we apply (7) to a noise-free linear time-invariant (LTI) system governed by $\bar{x}(n) = v(n) * h(n)$, where $h(n)$ is the system impulse response and $*$ denotes the convolution, it can be shown [21] that $\gamma_{v\bar{x}}^2(f) = 1$. Let us rearrange Figure 9 to Figure 11, where the system of $\hat{h}_1(n)$ represents the concatenated system of $h_2(n)$ and the Wiener filter. It has been shown in [21] that to minimize the MSE of $e(n)$ in Figure 11, there is $\hat{H}_1(f) = H_1(f)$, where $H_1(f)$ and $\hat{H}_1(f)$ are the frequency-domain representation of $h_1(n)$ and $\hat{h}_1(n)$, respectively. Based on this and assuming that $d(n)$ and $v(n)$ are uncorrelated, we can calculate the coherence function in (7) as

$$\begin{aligned}
\gamma_{vx}^2(f) &= \frac{|H_1(f)|^2 R_{vv}^2(f)}{R_{vv}(f)\left(|H_1(f)|^2 R_{vv}(f) + R_{dd}(f)\right)} \\
&= \left(1 + \frac{R_{dd}(f)}{|H_1(f)|^2 R_{vv}(f)}\right)^{-1} \\
&= \left(1 + \frac{R_{dd}(f)}{R_{xx}(f) - R_{dd}(f)}\right)^{-1}
\end{aligned} \tag{9}$$

where $R_{vv}(f)$ and $R_{dd}(f)$ are the auto spectral densities of $v(n)$ and $d(n)$, respectively. From (9), we have

$$1 - \gamma_{vx}^2(f) = \frac{R_{dd}(f)}{R_{xx}(f)} \tag{10}$$

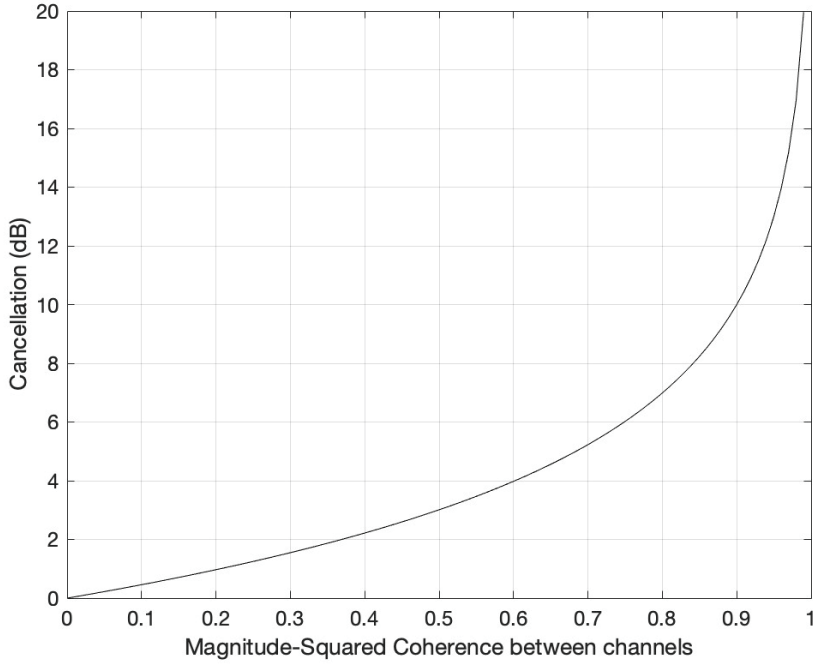

Fig. 12. Adaptive cancellation versus the squared coherence between $v_2(n)$ and $x(n)$.

which establishes the link of the LTI noise cancellation approach to the coherence function.

One of the major problems applying adaptive noise cancellation techniques in real acoustic environments is the low coherence between the noise corrupting the desired signal and the noise measured by the noise sensor. For the scenario depicted in Figure 9, a coherence value less than unity indicates that the system relating $x(n)$ and $v_2(n)$ does not fulfill the theoretical assumptions [20], thus the noise at those particular frequencies cannot be completely removed. The coherence function assumes an LTI relationship and stationary signals. The ensemble average is used so it furthermore assumes ergodicity. The amount of possible noise cancellation for this scenario as a function of frequency is given below and plotted in Figure 12

$$C(f) = \frac{1}{1 - |\gamma^2_{v_2 x}(f)|}. \tag{11}$$

C. Noise cancellation performance

Four different sources of background noise were tested on heartbeat measurements including a single 300 Hz tone, multiple tones consisting of 200 Hz, 300 Hz and 500 Hz, hospital/clinic noise, and breathing noise. The heartbeat measurements were of the second author and offer no diagnostic insight, thus no ethical approval was required for this research. A FireFace UCX was used with a MATLAB interface to allow simultaneous playback of the background noise through a Fostex 6301B speaker whilst recording from the stethoscope. FireFace UCX is a USB and FireWire audio interface manufactured by RME Audio [22]. The stethoscope was taped to the chest, making sure that the BNM was exposed.

Heartbeat measurements of 15-second duration were taken for each background noise scenario. The measurements were taken under breath-held conditions, except for the breathing noise scenario, where the speaker was turned off. The FireFace UCX collected data at 44.1 kHz sampling frequency, which was re-sampled down to 2 kHz. The adaptive noise cancellation algorithm in Section IV-A was implemented in MATLAB R2022a and run on a HP laptop.

The leakage coefficient $\alpha = 0.001$ was used in (6). The filter length N and step size μ were first tuned on the single 300 Hz tone. The combination that achieved the best noise attenuation was then used for the other background noise scenarios after comparing the performance to a conventional LMS algorithm reviewed in [23]. For each case, the coherence function between the HM and the BNM was plotted via a Welch estimator in MATLAB to indicate expected noise cancellation performance. The Welch estimator applies a 1024 length Hann window with 512 overlap calculated across 1024 samples. Power spectral density (PSD) plots and spectrograms generated through MATLAB allowed visualization of the performance.

When a single 300 Hz tone was played through the speaker, the coherence function between the HM and BNM is shown in Figure 13. It can be seen that the coherence function is unity at 300 Hz, indicating that a complete attenuation of the tone is possible using linear techniques.

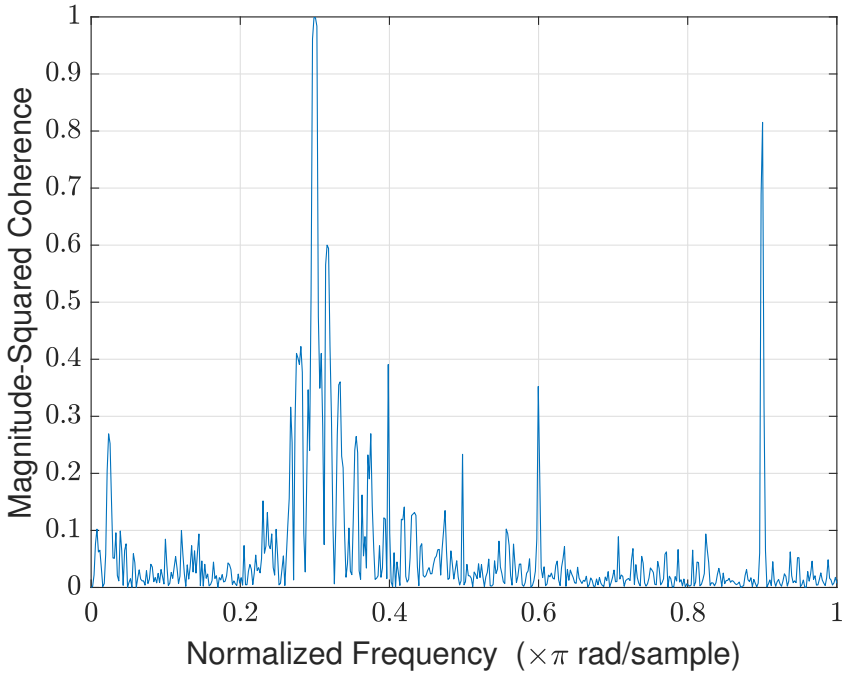

Fig. 13. 300 Hz tone background noise coherence function.

The NLMS algorithm was tested with $\mu = 0.05$ and $N = 256$, 512, and 1024. Figure 14 shows the PSD of the filtered HM signals and the unfiltered HM signal. It can be seen that for all three N, the amount of noise attenuation is 27 dB.

Next, we chose $N = 512$ for tuning the step size μ of the NLMS algorithm. Values of $\mu = 0.01$, 0.05, 0.1, and 0.5 were tested. Figure 15 shows the PSD of the filtered HM signal at various μ. It can be observed that the noise attenuation increased as μ increased, where almost complete attenuation (approx. 35 dB) was achieved at $\mu = 0.5$. This agrees with the coherence function which suggests a complete attenuation is possible. A 300 Hz tone is a periodic signal, suggesting why the measured attenuation was high. This may not be the case when other non-periodic noise sources are present.

In the following experiments, we set $N = 512$ and $\mu = 0.1$. In the second experiment, a multi-tone signal at 200 Hz, 300 Hz, and 500 Hz was played through the speaker. The coherence between the HM and BNM signals is displayed in Figure 16. It can be seen that the three tones had unity magnitude-squared coherence, as well as the first harmonic of each tone. This indicates that the first two harmonics of all three tones can be fully attenuated. In order to test the ability of the algorithm to adapt to different tones, the NLMS algorithm was tested on the scenario where each tone was played for 5 seconds while a heartbeat was recorded. It is shown in Figure 17 that a near complete noise attenuation was observed

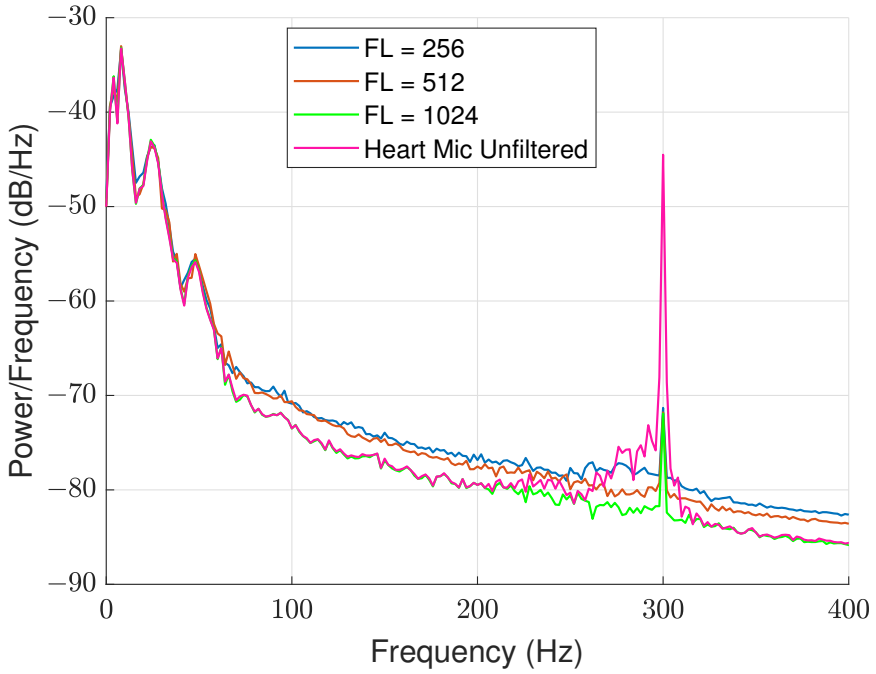

Fig. 14. PSD comparison of various filter lengths with 300 Hz tone background noise and $\mu = 0.05$.

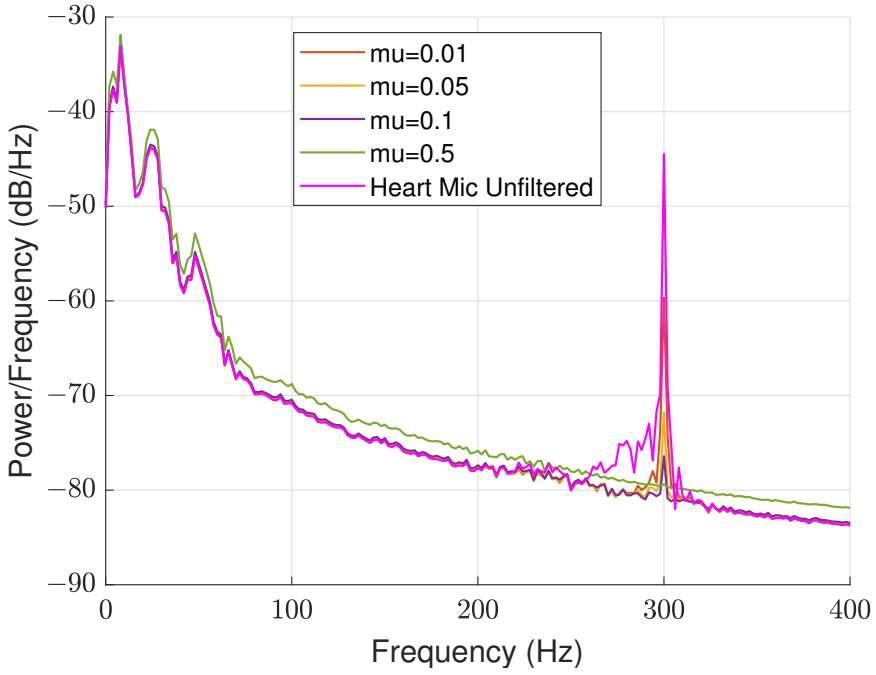

Fig. 15. PSD comparison of various μ with 300 Hz tone background noise and $N = 512$.

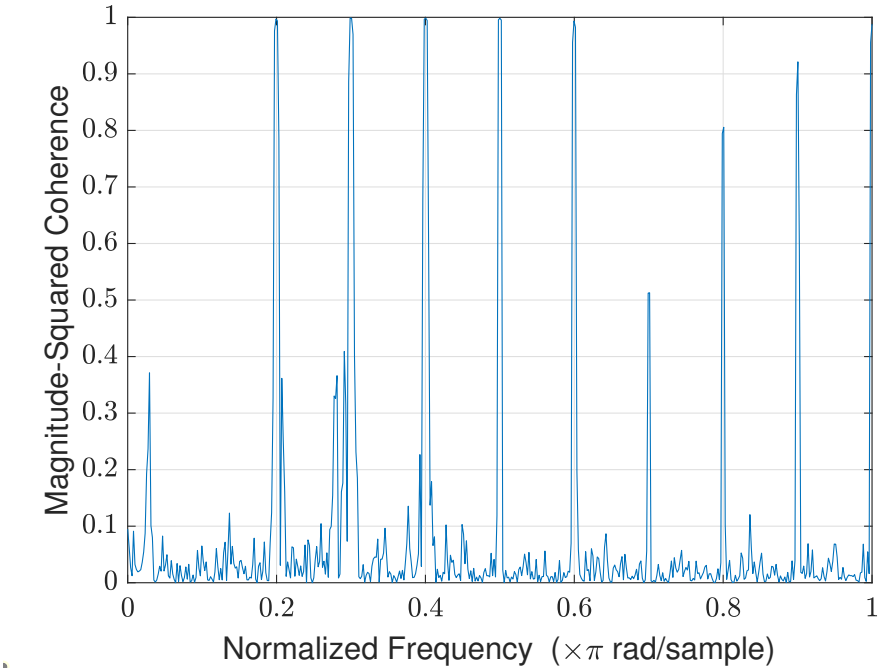

Fig. 16. 200, 300, and 500 Hz tone background noise coherence function.

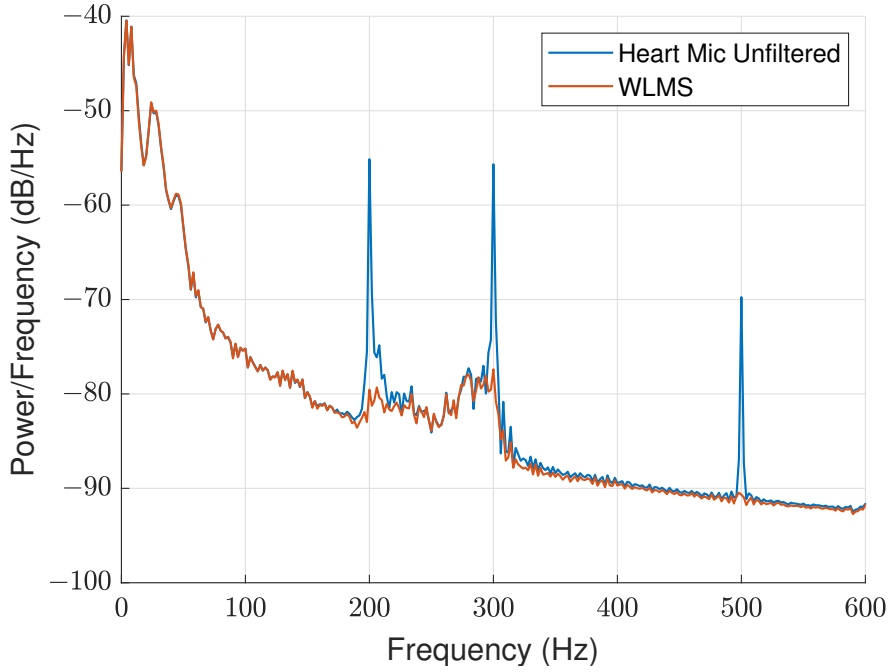

Fig. 17. PSD comparison of BNM signal, filtered and unfiltered HM signal in multi-tone background noise.

in the PSD comparison of the filtered and unfiltered HM signal. This can be visually observed in the spectrogram shown in Figure 18. Here, the BNM, HM and filtered HM signal spectrograms are displayed side-by-side, where the 200 Hz, 300 Hz, and 500 Hz tones were approximately attenuated by 24.5 dB, 21.4 dB, and 20.3 dB, respectively.

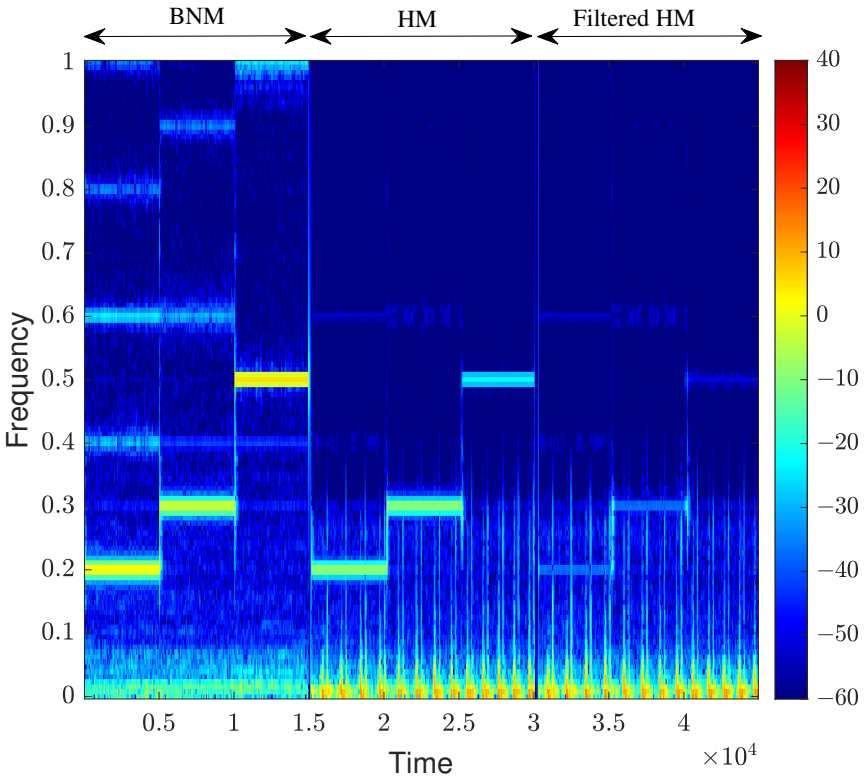

Fig. 18. Spectrogram comparison of filtered and unfiltered HM signal in multi-tone background noise.

In the third experiment, hospital/clinic background noise was played through the speaker while taking a heartbeat measurement. As this type of noise can be encountered in practical hospital and clinic environment, it is of high interest to attenuate this type of non-stationary noise within the frequency of interest. The magnitude-squared coherence between the HM and BNM signals is shown in Figure 19. It can be seen that the coherence function varies from 0.3 to 0.7 (not unity) for the frequency band between 200 Hz and 500 Hz. Thus, linear optimal filtering will not be able to achieve a complete attenuation of the background noise. According to Figure 12, around 3-5 dB noise suppression is possible at a magnitude-squared coherence value of 0.5. There is visual evidence of attenuation displayed in the spectrogram comparison in Figure 20, where we can see that the noise energy at 500 Hz, and the band between 200 and 300 Hz is slightly suppressed.

Many studies in the literature on the detection of CAD require that subjects hold their breath whilst heartbeat measurements are taken [16]. Although this is ideal for CAD detection, extending the stethoscope system to detect lung disease could be performed in unison, where the breathing noise can contain critical information. Thus, it is of great interest to separate the heartbeat sound from the breathing sound, so they can be studied and processed separately.

For the last experiment, the heartbeat signal was measured without holding breath and no extra external background noise sound was played. The magnitude-squared coherence function between the HM and BNM signals is shown in Figure 21. It can be seen that the coherence function was less than 0.3 across the whole frequency band. This suggests that only less than 1 dB of breathing noise suppression is achievable based on Figure 12. This was confirmed in Figure 22, where the HM signal before and after filtering appear identical. Non-linear methods may be explored to suppress the breathing noise and enhance

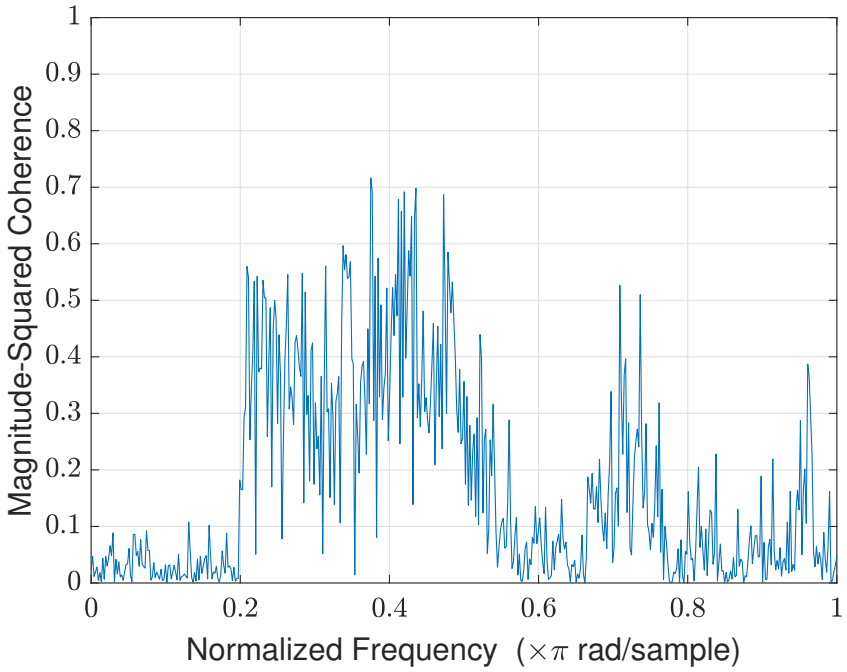

Fig. 19. Hospital/clinic background noise coherence function.

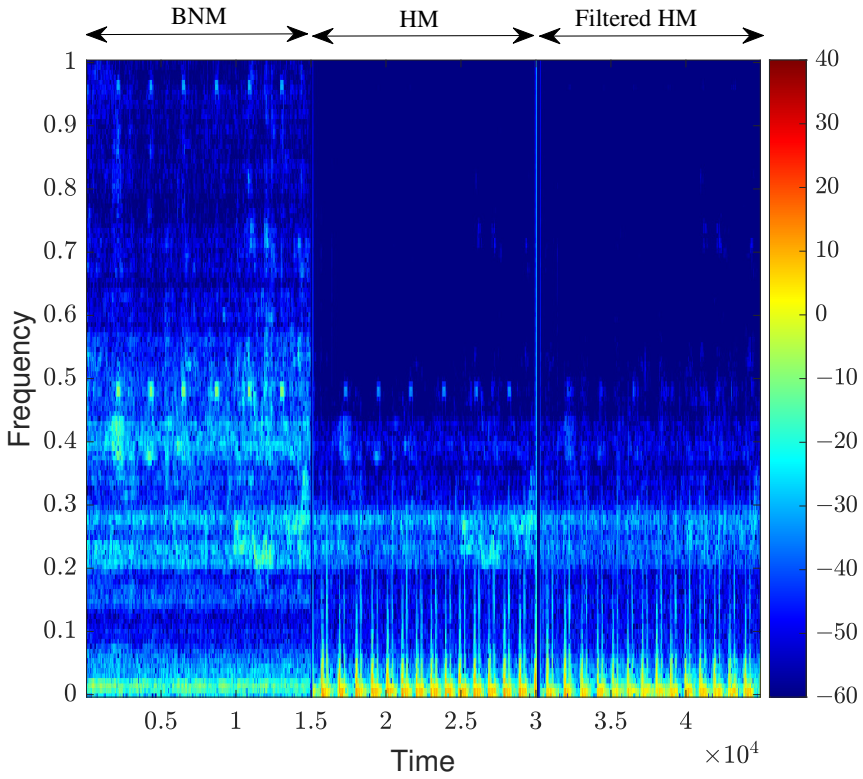

Fig. 20. Spectrogram comparison of filtered and unfiltered HM signal in hospital/clinic background noise.

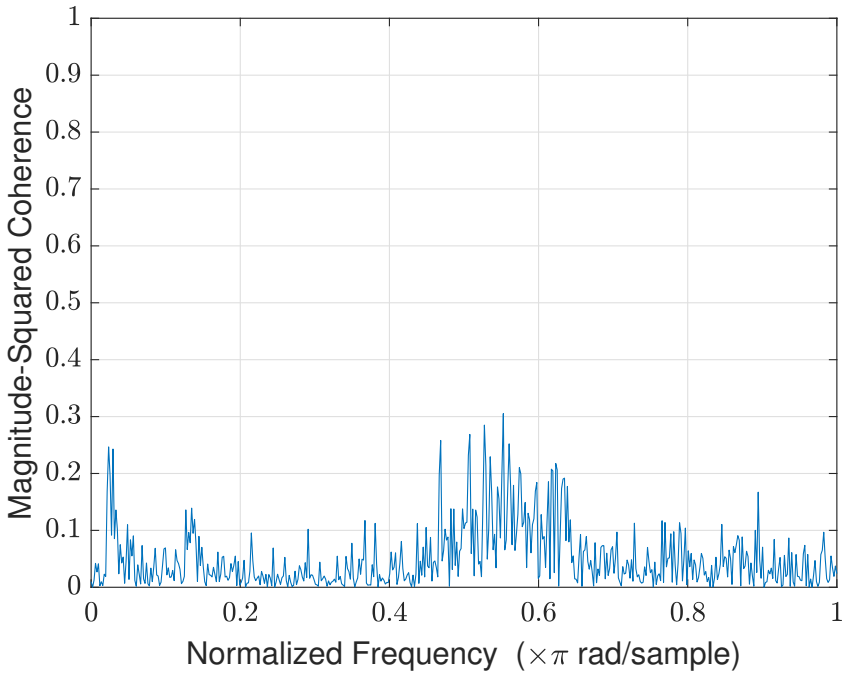

Fig. 21. Breathing noise coherence function.

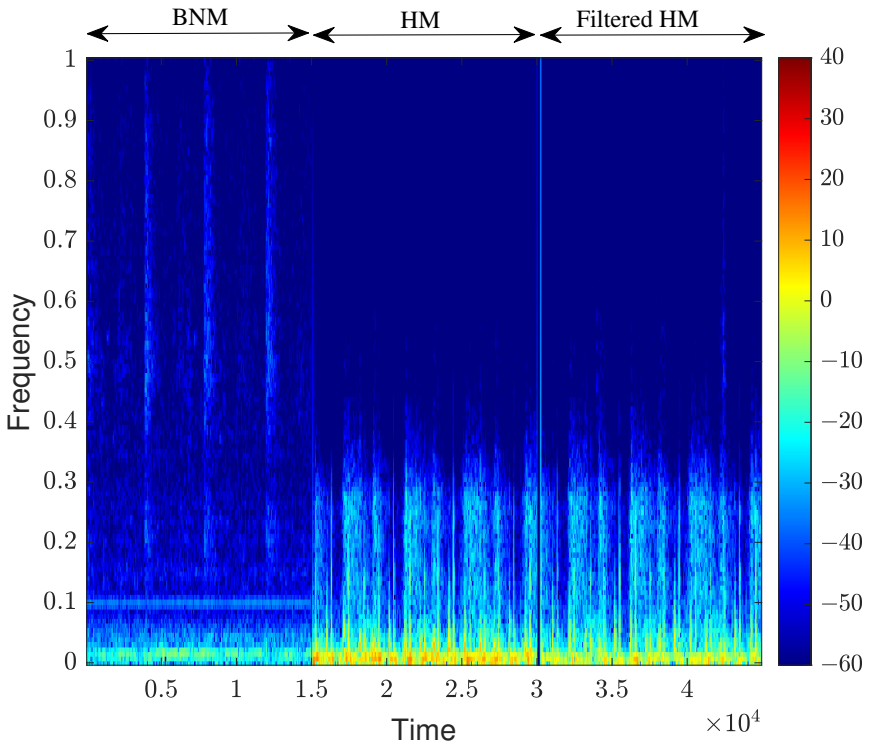

Fig. 22. Spectrogram comparison of filtered and unfiltered HM signal with breathing noise.

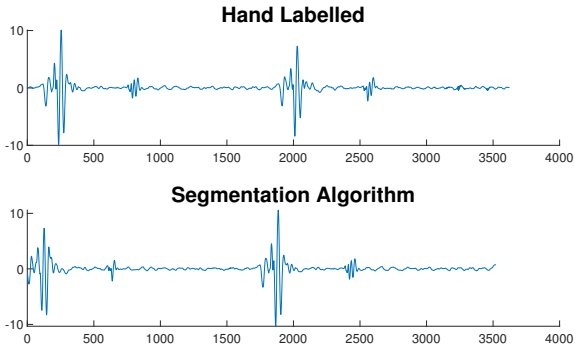

Fig. 23. Hand labelled (top) versus HSMM-regression based segmentation algorithm (bottom).

the heart signal. These other methods should be carefully designed as to not suppress murmurs that are present in CAD patients, as the diagnostic process will be negatively affected.

V. Machine learning based classification

There are many factors which can affect the results of a neural network based CAD classifier. In this section, we discuss four of these factors: segmentation of heart cycle, the integrity of heartbeat signal, data set size and the neural network structure. We demonstrate that among these factors, the integrity of heartbeat signal has great impact on the performance of a classifier.

A. Heart cycle segmentation

Prior to input into a neural network based classifier, heartbeat signals are usually segmented into heart cycles. Let us take [24] as an example. In [24], the heartbeat data input to the SVM algorithm is segmented into three epochs, each containing two full heart cycles. Thus, one epoch commences at S1 of the first cycle and finishes at the end of diastole in the second cycle.

The heartbeat signal segmentation can be implemented manually following the methods in [24]. However, hand-segmenting heartbeat data can be time-consuming and is often not practical for many applications which require automatic data acquisition and pre-processing. In this case, a logistic regression-hidden semi-Markov model (HSMM) based heart sound segmentation algorithm adapted in [25] can be applied to provide automatic heart cycle segmentation. The segmentation algorithm [25] is a probabilistic model-based approach, which identifies positions that correspond to S1, systole, S2, and diastole. Figure 23 displays a hand-segmented heartbeat and a heartbeat segmented via the HSMM-regression algorithm.

It can be seen that the heartbeat cycle starts directly at S1 when using the HSMM-regression algorithm, and there is a small delay shown in Figure 23 where the hand-segmentation is based on. Ideally, this small shift would not affect the classification results. To test the sensitivity of such shift, an adjustment factor was included in the pre-processing algorithms. This adjustment factor causes a shift in the heartbeat epochs after segmentation using the HSMM-regression algorithm, allowing a slight delay before S1. In particular, 8 different adjustment factors were tested: $0, 0.006, 0.016, \ldots, 0.066$. The adjustment factor is multiplied with the amount of samples in each epoch to obtain the shift value, which is the amount of samples the segmented epoch is delayed by. Figure 24 displays the same epoch subjecting to four different adjustment factors. Note that the higher adjustment factors provide a better representation of the hand-segmented epochs in Figure 23, where there is a slight delay before the first S1.

For each adjustment factor, 40 normal and 40 abnormal heartbeats from data set *a* in [12] were segmented into three epochs using the HSMM-regression algorithm. The k-nearest neighbor (k-NN) SVM algorithm was implemented on the data set of 80 segmented subjects. The algorithm in [24] including the whole process of performing the synchrosqueezing transform (SST) on each epoch, feature extraction

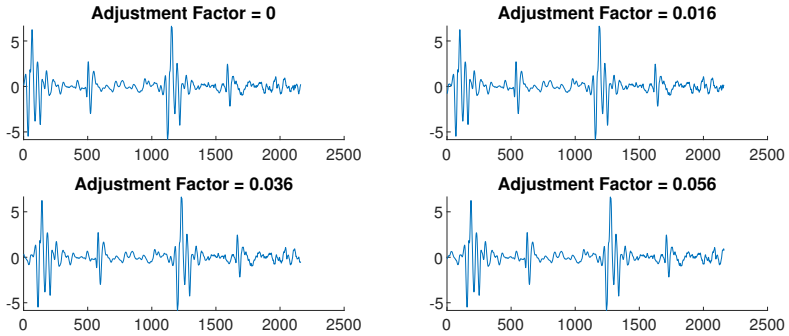

Fig. 24. Epochs with various adjustment factors.

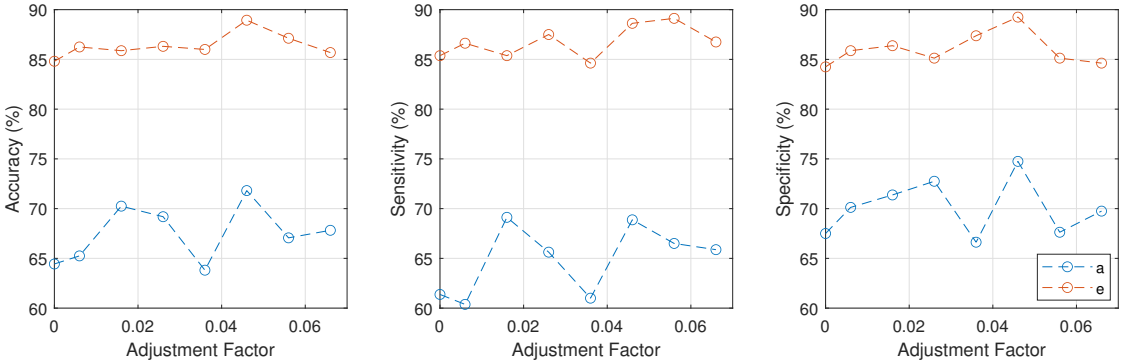

Fig. 25. Accuracy, sensitivity and specificity versus the adjustment factor for data sets *a* and *e*.

and selection, and training the SVM classifier was repeated each time. The highest accuracy, sensitivity, and specificity were recorded for each test.

The accuracy, sensitivity, and specificity plotted as a function of the adjustment factor is displayed in Figure 25. Let us focus on the curves with 'a' at this time. It can be seen that the accuracy varied from 63.8% to 71.8% during the adjustment factor variation test. This 8% difference between the minimum and maximum score occurred when the adjustment was 0.036 and 0.046, respectively. This shows that the SVM algorithm is quite sensitive with respect to pre-processing the heartbeat data. Ideally, the variation should be kept at a minimum as it is impossible to segment each subject's heartbeat in the same way. As it stands, a variation of 8% is too large. Note that when hand-segmenting data set *a*, the accuracy was 68%. Further investigations need to be conducted to determine the nature of this accuracy variation for different adjustment factors, where the behaviour was also apparent in the sensitivity and specificity plots in Figure 25. We can investigate utilizing the simultaneous ECG recording (shown in Figure 7) for a better defined segmentation. Perhaps segmentation is not necessary, and algorithms can be designed that bypass this pre-processing technique to achieve the best results.

B. Signal integrity

The same test in Section V-A was repeated using 40 normal and 40 abnormal heartbeats from data set *e* in [12]. Performing the segmentation test on different heartbeat recordings indicates how the performance varies over different environment. We can observe from Figure 25 that for all adjustment factors, the accuracy, sensitivity, and specificity for data set *e* was consistently higher than those of data set *a*. As both data sets were acquired in different environment by separate research groups, the amount of noise corrupting the signals was not the same. From the results, we conclude that data set *e* contained less noise in the measurements, due to the accuracy varying between 85% and 90% for each adjustment factor, as

opposed to 64% and 71% for data set *a*. This clearly indicates the importance of the data acquisition process. Although both sets of heartbeat signals appeared to be the same visually when filtering out noisy measurements, it is clear that critical features were corrupted in data set *a*, which negatively affected the classification process.

For data set *e*, the maximum accuracy again corresponds to an adjustment factor 0.046. However, the variation was not as large as that in data set *a*, with a minimum accuracy of 84.8% and maximum accuracy of 88.9%. The smaller variation in accuracy can also be explained by the fact that data set *e* was cleaner. Nevertheless, this 4.1% difference still has the ability to degrade the specification goal, despite having cleaner signals.

From the results in Section IV, we expect that due to its capability to suppress the background noise, the digital stethoscope with a BNM can obtain measurements with high signal integrity. Due to this reason, it is expected that the accuracy, sensitivity, and specificity of the PCG system in Section II can be higher than those of data set *e*, fulfilling the specification of exceeding 80%.

C. Data set size

The previous tests consist of 80 subjects in total. In this section, we investigate how the data set size affects the classification accuracy, where it is hypothesized that larger data sets will produce better results. Data sets *a* and *e* were combined to form larger training sets, with 160 and 240 subjects separately trained. Each set had equal numbers of normal and abnormal heartbeats, as well as equal measurements from both data sets. There is not enough clean data to extend the testing beyond 240 subjects. The adjustment factor of 0.046 was used on each data set for heart cycle segmentation, as this factor had the highest performance in the previous test according to Section V-A.

Table II displays the accuracy, sensitivity, and specificity for different data set sizes made by combining heartbeats from data sets *a* and *e*. It can be seen from Table II that when increasing the data set size, the accuracy and specificity increased by 1.2% and 3.8%, respectively. However, the sensitivity decreased by 1.4%. More tests with different data sets and sizes need to be investigated to truly understand the nature between the results and data set size. In saying this, it was encouraging to observe the accuracy and specificity increase.

TABLE II
CAD CLASSIFICATION RESULTS OBTAINED FROM DIFFERENT DATA SET SIZES

Data set size	Accuracy [%]	Sensitivity [%]	Specificity [%]
160	73.6	72.5	74.7
240	74.8	71.1	78.5

From the previous tests, the highest accuracy achieved in data sets *a* and *e* was 71.8% and 88.9%, respectively. When joining these data sets together the accuracy was not at the midpoint, instead achieving just 73.6%. Thus, the lower quality data negatively affected the entire data set. This again indicates the importance of signal integrity. Even if most of the data is of high quality, a certain number of poorer measurements arising from body movements or unexpected noise have the potential to disproportionately compromise the results.

TABLE III
CAD CLASSIFICATION ACCURACY OF TWO DIFFERENT NEURAL NETWORKS

Neural network	Data set *a*	Data set *e*
SVM	71.8%	88.9%
CNN	73.8%	95.0%

D. Neural network structure

Lastly, we study the impact of the neural network structure on the performance of a CAD classifier by comparing the SVM based classifier [24] and the convolutional neural network (CNN) driven classifier in [26], which contains two convolution layers followed by a multilayer perceptron (MLP) network. An adjustment factor 0.046 was adopted for the SVM approach, which has the highest accuracy. No adjustment was used for the CNN-based classifier, i.e., the HSMM-regression based segmentation algorithm was directly applied. The accuracy results of both approaches are shown in Table III. We can see that for both data sets *a* and *e*, the CNN based neural network achieved higher accuracy. This is mainly due to the fact that it has a larger number of layers and parameters than the SVM to adapt to the features of the heartbeat signals, rendering a higher classification accuracy. However, we note that a deeper neural network in general needs a larger amount of labelled training data. Interestingly, we also observe that the improvement in data set *e* is 6%, while only 2% for data set *a*. This suggests that by using deep neural network to cleaner data set, one can achieve a higher accuracy.

VI. Conclusions

In this chapter, we discussed a CAD pre-screening technique using PCG. A multi-channel PCG and ECG measurement device was presented. For the topic of adaptive noise cancellation in PCG signals, we showed in this chapter that the coherence function can be used to predict the performance of a NLMS-based method in suppressing the background noise. By applying existing machine learning based CAD classification algorithms to heart sound recordings in the open database of the 2016 PhysioNet/Computing in Cardiology Challenge, we demonstrated that the integrity of the heart sound signal has a significant impact on the accuracy of classification. Another discovery we showed in the chapter is that the identification of the correct starting point of heartbeat cycle affects the classification results, which raises an interesting question if segmentation of the PCG signal is absolutely necessary in machine learning based classification methods.

References

[1] G. A. Roth *et. al.*, "Global burden of cardiovascular diseases and risk factors, 1990-2019: Update from the GBD 2019 study," *Journal of the American College of Cardiology*, vol. 76, no. 25, pp. 2982-3021, 2020.

[2] A. McLellan and D. Prior, "Cardiac stress testing stress electrocardiography and stress echocardiography," *Australian Family Physician*, vol. 41, no. 3, pp. 119-122, Mar. 2012. [Online]. Available: http://www.racgp.org.au/afp/2012/march/cardiac-stress-testing/.

[3] A. K. Kumar and G. Saha, "Interpretation of heart sound signal through automated artifact-free segmentation," *Heart Research Open Journal*, vol. 2, no. 1, pp. 25-34, 2015.

[4] I. Cathers, "Neural network assisted cardiac auscultation," *Artificial Intelligence in Medicine*, vol. 7, no. 1, pp. 53-66, 1995.

[5] T. R. Reed, N. E. Reed, and P. Fritzson, "Heart sound analysis for symptom detection and computer-aided diagnosis," *Simulation Modelling Practice and Theory*, vol. 12, no. 2, pp. 129-146, 2004.

[6] 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.

[7] 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.

[8] A. N. Makaryus, J. N. Makaryus, A. Figgatt, D. Mulholland, H. Kushner, J. L. Semmlow, J. Mieres, and A. J. Taylor, "Utility of an advanced digital electronic stethoscope in the diagnosis of coronary artery disease compared with coronary computed tomographic angiography," *The American Journal of Cardiology*, vol. 111, no. 6, pp. 786-792, 2013.

[9] S. Schmidt, C. Holst-Hansen, J. Hansen, E. Toft, and J. Struijk, "Acoustic features for the identification of coronary artery disease," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 11, pp. 2611-2619, Nov. 2015.

[10] H. Li, X. Wang, C. Liu, Y. Wang, P. Li, H. Tang, L. Yao, and H. Zhang, "Dual-input neural network integrating feature extraction and deep learning for coronary artery disease detection using electrocardiogram and phonocardiogram," *IEEE Access*, vol. 7, pp. 146457-146469, Oct. 2019.

[11] J. Oliveira *et al.*, "The CirCor DigiScope dataset: From murmur detection to murmur classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 6, pp. 2524-2535, June 2022.

[12] G. D. Clifford, C. Liu, B. Moody, D. Springer, I. Silva, Q. Li, and R. G. Mark, "Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016," in *Proc. Computing in Cardiology Conference*, Vancouver, BC, Canada, Sep. 11-14, 2016.

[13] George B. Moody PhysioNet Challenge, https://moody-challenge.physionet.org/2022/.

[14] A. Kazemnejad, P. Gordany, and R. Sameni, "An open-access simultaneous electrocardiogram and phonocardiogram database," *bioRxiv preprint*. [Online]. Available: https://doi.org/10.1101/2021.05.17.444563.

[15] Ticking Heart Pty Ltd, https://www.tickingheart.com.

[16] P. Samanta, A. Pathak, K. Mandana, and G. Saha, "Classification of coronary artery diseased and normal subjects using multi-channel phonocardiogram signal," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 2, pp. 426-443, 2019.
[17] D. Della Giustina, M. Riva, F. Belloni, and M. Malcangi, "Embedding a multichannel environmental noise cancellation algorithm into an electronic stethoscope," *International Journal of Circuits, Systems and Signal Processing*, vol. 5, no. 2, pp. 184-191, 2011.
[18] B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong, and R. C. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692-1716, Dec. 1975.
[19] J. A. Zhang, N. Murata, Y. Maeno, P. N. Samarasinghe, T. D. Abhayapala, and Yuki Mitsufuji, "Coherence-based performance analysis on noise reduction in multichannel active noise control systems," *The Journal of the Acoustical Society of America*, vol. 148, no. 3, pp. 1519-1528, 2020.
[20] J. S. Bendat, A. G. Piersol, *Random Data: Analysis and Measurement Procedures; John Wiley & Sons*: Hoboken, NJ, USA, 2011.
[21] M. Fynn, S. Nordholm, and Y. Rong, "Coherence function and adaptive noise cancellation performance of an acoustic sensor system for use in detecting coronary artery disease," *MDPI Sensors*, vol. 22, no. 17, 6591, Sep. 2022.
[22] Fireface UCX, https://www.rme-audio.de/fireface-ucx.html.
[23] S. Dixit, D. Nagaria, "LMS adaptive filters for noise cancellation: A review". *Int. J. Electr. Comput. Eng.* 2017, 7, 2520.
[24] A. Pathak, P. Samanta, K. Mandana, and G. Saha, "Detection of coronary artery atherosclerotic disease using novel features from synchrosqueezing transform of phonocardiogram," *Biomedical Signal Processing and Control*, vol. 62, 102055, 2020.
[25] 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, 2015.
[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 *Proc. Computing in Cardiology Conference*, Vancouver, BC, Canada, Sep. 11-14, 2016.