2022

An audio processing pipeline for acquiring diagnostic quality heart sounds via mobile phone

Davoud Shariat Panah

Andrew Hines

Joseph McKeever

See next page for additional authors

Funder: Science Foundation Ireland

Authors
Davoud Shariat Panah, Andrew Hines, Joseph McKeever, and Susan McKeever

Computers in Biology and Medicine 145 (2022) 105415

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

An audio processing pipeline for acquiring diagnostic quality heart sounds via mobile phone

Davoud Shariat Panah [a,*], Andrew Hines [b], Joseph A. McKeever [c], Susan McKeever [a]

[a] School of Computer Science, Technological University Dublin, Dublin, Ireland
[b] School of Computer Science, University College Dublin, Dublin, Ireland
[c] UCD Academy, Beacon Hospital, Sandyford, Dublin, Ireland

ARTICLE INFO

Keywords:
Phonocardiogram
Heart sound capture
Quality assessment
Quality enhancement
Mobile phone
Subjective listening test

ABSTRACT

Recently, heart sound signals captured using mobile phones have been employed to develop data-driven heart disease detection systems. Such signals are generally captured in person by trained clinicians who can determine if the recorded heart sounds are of diagnosable quality. However, mobile phones have the potential to support heart health diagnostics, even where access to trained medical professionals is limited. To adopt mobile phones as self-diagnostic tools for the masses, we would need to have a mechanism to automatically establish that heart sounds recorded by non-expert users in uncontrolled conditions have the required quality for diagnostic purposes. This paper proposes a quality assessment and enhancement pipeline for heart sounds captured using mobile phones. The pipeline analyzes a heart sound and determines if it has the required quality for diagnostic tasks. Also, in cases where the quality of the captured signal is below the required threshold, the pipeline can improve the quality by applying quality enhancement algorithms. Using this pipeline, we can also provide feedback to users regarding the cause of low-quality signal capture and guide them towards a successful one. We conducted a survey of a group of thirteen clinicians with auscultation skills and experience. The results of this survey were used to inform and validate the proposed quality assessment and enhancement pipeline. We observed a high level of agreement between the survey results and fundamental design decisions within the proposed pipeline. Also, the results indicate that the proposed pipeline can reduce our dependency on trained clinicians for capture of diagnosable heart sounds.

1. Introduction

Cardiovascular diseases (CVDs) are currently the leading cause of death worldwide. According to the World Health Organization (WHO), over 17 million people die of CVDs each year, and it is expected that this figure will rise to 23 million by 2030 [1]. While this is a large number, the good news is that most cardiovascular diseases are manageable, provided that they are diagnosed as early as possible [2].

For more than 200 years, cardiac auscultation has been considered a cost-effective heart health screening method [3]. In this technique, a physician listens to the patient's heart sounds and analyzes the timing, duration, frequency, intensity, and quality of the sounds [4]. Listening to the heart sounds in conjunction with performing a general examination and taking a clinical history enables trained clinicians to diagnose a whole host of CVDs [5].

Recently, detecting heart disease through automatic analysis of heart sounds has been an active area of research [6–11]. Heart sound signals have been utilized to develop data-driven heart abnormality prediction systems that are able to classify heart sounds into normal and abnormal categories. In some cases, such heart abnormality detection systems have achieved acceptable accuracies. However, these systems are decision support systems, targeted for use by trained clinicians who can use digital stethoscopes to capture the heart sounds from the patients and consequently verify the validity of such signals for diagnostic purposes. This would limit the adoption of such systems in situations where trained medical professionals are not available, which is the case in underdeveloped regions of the world [12]. At the same time, according to WHO, more than 75% of the mortalities that are related to CVDs occur in low and middle-income countries [1].

In the last decade, the penetration rate of mobile technologies, and in

* Corresponding author.
E-mail addresses: d19127274@mytudublin.ie (D. Shariat Panah), andrew.hines@ucd.ie (A. Hines), joseph.mckeever@beaconhospital.ie (J.A. McKeever), susan.mckeever@tudublin.ie (S. McKeever).

https://doi.org/10.1016/j.compbiomed.2022.105415
Received 20 January 2022; Received in revised form 22 February 2022; Accepted 14 March 2022
Available online 24 March 2022

particular smartphones, has been growing rapidly even in underdeveloped parts of the world [13]. The current generation of mobile phones is generally equipped with different sensors such as microphones, cameras, and accelerometers. Also, they benefit from powerful processors that make complex on-device computations feasible. Mobile technologies have the potential to provide personalized healthcare interventions, especially in the areas of the world where access to trained medical professionals is limited. Mobile technologies have been successfully employed to support the screening of infectious diseases such as COVID-19 [14] as well as non-infectious diseases like diabetes [15] and cancer [16]. Also, recently they have been employed for the initial screening of heart disease. For example, heart sound signals captured using smartphones have been used to build data-driven models to detect heart disease [17,18]. While such studies have successfully employed mobile-based heart sound signals to detect CVDs, the proposed diagnostic systems lack any mechanism to establish that the captured signals are of diagnosable quality. In other words, such systems should be utilized only by trained medical professionals who have physical access to patients to ensure that the captured signals are valid for the diagnostic task. This would still hinder the adoption of such mobile phone-based systems as self-diagnostics or tele-medicine tools in situations where access to trained clinicians is limited.

The adoption of mobile phones to capture heart sound signals poses different data validity and consistency challenges. First, the users of such consumer devices are generally non-experts who might not have auscultation skills and consequently might not know how to capture valid heart sounds for diagnostic purposes. Also, such users would use mobile phones as a capture device in uncontrolled environments such as their homeplace where different noises and disturbances might corrupt the captured signals. Lastly, unlike digital stethoscopes that benefit from noise reduction or cancellation technologies [19–21], mobile phones might lack such capabilities and, as a result, could be more prone to ambient noise.

Addressing the challenges mentioned above would be the first step towards developing a self-diagnostics or tele-medicine system based on mobile phones for heart disease screening. In this regard, our broad goal is to enable non-expert users to capture diagnostic quality heart sounds using mobile phones. In order to solve this problem, two questions must be answered: First, how to capture the heart sounds independently from clinicians using mobile phones, and second, how to process the captured sounds to deliver a signal with the necessary fidelity for use in a data-driven classification model. This study focuses on the second question. We should note that the placement of the mobile phone at different auscultation sites and collecting multiple heart sounds is a separate important problem that is out of the scope of this study. We assume that the guidance will be provided to the users regarding the correct placement of the device. This study focuses on the quality of the heart sound capture, as opposed to addressing the challenges involved in active capturing of the heart sounds. Therefore, the precision of the placement of the capture will not be addressed in this study.

In this paper, we propose a heart sound quality assessment and enhancement (QAE) pipeline for mobile-based heart sound signals. The QAE pipeline includes mechanisms to evaluate the quality of the heart sound signals captured using mobile phones and apply signal quality enhancement algorithms adaptively. This pipeline also allows us to provide feedback to users regarding the validity of the recorded heart sounds and guide them towards an acceptable signal capture. We conducted a survey of a group of clinicians to fill in the gaps we found in the literature regarding the characteristics of a diagnosable heart sound signal and used the results of this survey to inform and validate the design decisions within the QAE pipeline.

The remainder of the paper is structured as follows: Section 2 provides an overview of the related work on quality assessment and enhancement of heart sounds. In Section 3, the QAE pipeline is presented. In Section 4, the details of the QAE pipeline validation, including the dataset generation process and survey, are presented. Results are provided in Section 5. In Section 6, the results are discussed. Conclusions and future directions are presented in Section 7.

2. Related work

2.1. Clean heart sounds

To characterize a clean heart sound recording that would be valid for diagnostic purposes, we reviewed clinical and non-clinical sources.

In terms of the minimum number of heartbeats needed in a recording, there is no consensus among non-clinical sources. For example, some studies [10,11,22] have utilized short-length heart sound segments (around 1 s) to develop data-driven models. According to Ref. [11], an average heartbeat cycle is 0.8 s long. As a result, a 1-s heart sound segment would contain at least one heartbeat cycle. Other studies have used longer segments. Three seconds [6,7,9,23], 4 s [8], 5 s [24–28], and 10 s [29] are all different values mentioned for the length of the heart sound segments in non-clinical sources. However, according to a clinical source [30], listening to any spot on the chest should take at least 5 s. This figure aligns with the length of the longer heart sound segments employed in non-clinical studies.

Regarding the impact of noise on heart sounds, non-clinical studies have emphasized the destructive impact of noise and mentioned that noise in heart sound data could cause data-driven heart abnormality detection systems to misclassify heart sounds [31–35]. Such noises can be broadly categorized as internal noise like digestive and respiratory noise, ambient noise like background speech, and noise due to movement like body movement [32]. Clinical sources have also pointed out the harmful impact of noise and interference on the accuracy of auscultation. In Ref. [36], it has been emphasized that ambient noise must be minimized at the time of auscultation. Coviello [37] has noted the destructive impact of the movement noise, and [38] pointed out that respiratory noises can also be disturbing in some cases.

2.2. Heart sound quality enhancement

This section reviews the different heart sound processing techniques that have been employed in literature to enhance the equality of the recordings and consequently decrease the possibility of misclassification of heart sounds by a classifier.

Various methods have been proposed in the literature to reduce or eliminate noise in heart sound recordings. Filtering is the simplest technique that has been frequently employed. Low-pass [39,40] and band-pass [41–43] filters with different cut-off frequencies have been employed to eliminate noise from heart sound recordings. The main disadvantage of this approach is that in some cases, the frequency range of noise and disturbances might overlap with the frequency range of heart sounds, and as a result, such filtering techniques would not be able to reduce the noise effectively [44].

Another denoising approach that researchers have widely employed is wavelet-based denoising [32,34,35,45–47]. According to Messer et al. [35], wavelet coefficients of heart sounds are much larger than those due to noise. Therefore, the coefficients that are smaller than a particular threshold are considered as noise and disregarded. In their work, Kumar et al. [46] proposed an algorithm based on Q-wavelet transform and signal second difference to remove short duration distortions (artifacts) from heart sound recordings. They evaluated their method using four different murmur heart sounds contaminated with different artifacts and achieved an average accuracy of 96.13%. Gradolewski et al. [32] have proposed an adaptive denoising algorithm for heart sound signals recorded using mobile devices in noisy environments that combines wavelet transform with a time-delay neural network. Like their previous work [45], they evaluated the performance of their algorithm on various intensities of pink and white noise. Pink noise is common in biological systems and is characterized by the noise power decreasing inversely with signal frequency, while in the case of white noise, the noise power is constant in different frequency intervals. It must be noted that color noise is only one of the noise types that could corrupt mobile phone-based heart sound recordings in noisy environments, and the aforementioned studies have not evaluated the performance of their methods using heart sounds corrupted with a variety of noise types.

In addition to the approaches mentioned above, some researchers have tried to identify the parts of the heart sound signal that might be noise-free or less corrupted by noise. Kumar et al. [48] have used the periodicity characteristic of heart sounds to detect a small part of the signal that is not noisy and used that as a reference to distinguish noise from heart sounds. Li et al. [49] proposed a method to select the sub-sequence of heart sound with the least noise based on the cyclostationary property of heartbeat events, meaning that parts of the signal less corrupted by noise have a greater periodicity. The main issue with these techniques is that they assume there is a part of the signal in a heart sound recording that is either noise-free or nearly clean. However, this assumption may not always be valid, especially in cases where strong continuous noise is present.

The studies presented rely on enhancing the heart sound signals irrespective of their noise content. Enhancement presents overall noise reductions but it also adds artifacts and corruption. This will lead to increased information loss in cases where enhancement was not required.

2.3. Heart sound quality classification

As an alternative to noise reduction and elimination, recently, some researchers have tried to assess the quality of heart sounds by developing rule-based or data-driven signal quality assessment systems [50–56]. In other words, the quality of the signals is evaluated using a quality assessment system, and signals with an acceptable level of quality are used for further processing. Springer et al. [54] defined nine signal quality features and trained a logistic regression model to classify heart sound recordings into good- and poor-quality categories. Their model was able to classify mobile phone recordings with an accuracy of 82.2%. In Ref. [55], Das et al. proposed device-agnostic features to automatically identify the quality of heart sound recordings in near-real-time. Some of these features were derived from the autocorrelation waveform and others from the signal spectrum. They used PhysioNet dataset [57] to train their quality assessment model and tested it using data collected by a low-cost smartphone-based stethoscope, achieving an accuracy of 75%. Grooby et al. [52], proposed a data-driven signal quality assessment system for neonatal heart and lung sounds. They extracted over 180 features from heart and lung sounds and trained different classification models, including SVM, KNN and Decision tree. Their model achieved an accuracy of 93% in classifying heart sounds into high- and low-quality categories. In their recent work, Tang et al. [56] annotated the samples of a large dataset with more than 7800 recordings in terms of their quality. They extracted ten different features, including but not limited to kurtosis, energy ratio and degree of periodicity and trained a binary SVM to classify signals into acceptable and unacceptable categories. Their model achieved an accuracy of 94.3% in 10-fold cross-validation.

Although some of the data-driven signal quality assessment systems mentioned above achieved acceptable accuracy, it is worth noting that training such models requires a large amount of labeled data that could be hard to obtain, especially in the case of mobile phone-based heart sounds. In addition to this, it has been shown that data-driven heart sound classification models can be biased towards the sensors employed to collect the data used to train such systems [58]. As a result, heart sound quality classification models built using datasets collected by different sensors could potentially suffer from such bias. In addition, in comparison to rule-based systems, such heart sound quality classification systems provide less information regarding the type and level of noise in the signals and may not generalize well in cases where heart sounds are contaminated with new forms of noise. Also, as mentioned earlier, heart sound quality classification systems try to categorize recordings into acceptable and unacceptable classes and discard low-quality recordings. However, capturing heart sounds with an acceptable level of quality might not be possible in noisy environments, especially when heart sounds are recorded using mobile phones. As a result, applying the quality enhancement algorithms to heart sound signals would be inevitable in such cases.

3. QAE pipeline

In this section, we present the proposed quality assessment and enhancement pipeline. Fig. 1 illustrates the QAE pipeline design. As shown in this diagram, the pipeline includes six stages of quality assessment (QA1-QA6). Also, it includes three stages of processing, out of which two are quality enhancement stages (QE1 and QE2). This design also enables us to give feedback to users regarding the cause of the signal rejection at four different decision points in the pipeline (F1–F4).

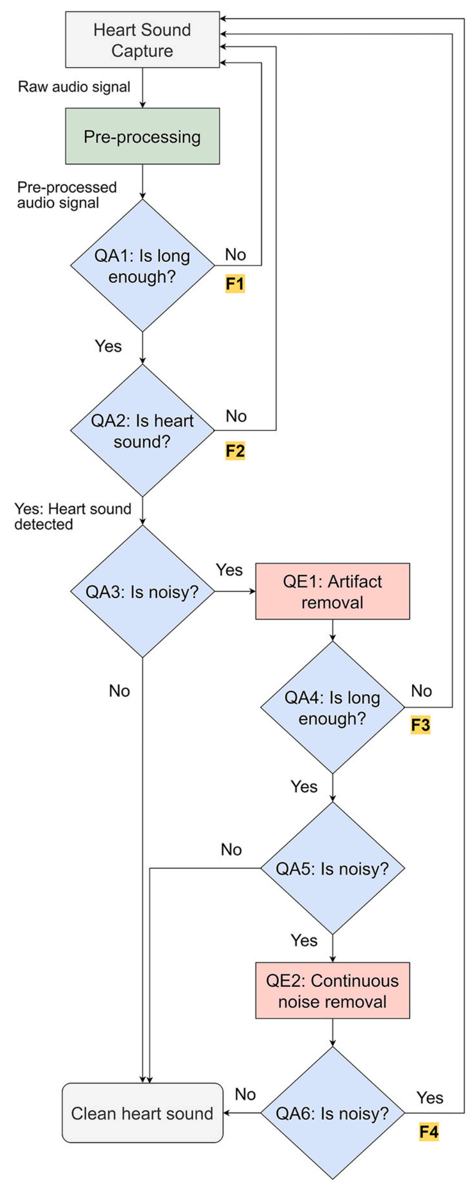

Fig. 1. Signal quality assessment and enhancement (QAE) pipeline. This pipeline includes six stages of quality assessment (QA1-QA6), two stages of quality enhancement (QE1 and QE2), and four feedback decision points (F1–F4).

As shown in Fig. 1, the first step in the pipeline is audio capture. In this step, a heart sound signal is captured using a mobile phone. Then, this audio signal is pre-processed in the second step, which includes resampling and amplitude normalization. Microphones of different mobile phones might have different digital sampling rates; thus, we down-sample the recordings to 2000 Hz to standardise the data to a level that maintains salient heart sound information and removes non-salient higher frequency data using a polyphase anti-aliasing filter, as in Ref. [54]. Also, in order to minimize the variation in amplitudes across recordings, amplitude normalization is performed, using the following equation (as in Ref. [44]):

$$S_{\text{norm}}(t) = \frac{S(t)}{\max(|S|)} \tag{1}$$

where $S(t)$ is the value of the signal at time t, and $\max(|s|)$ is the maximum of the absolute value of the signal at time t.

After the pre-processing step, the audio signal goes through multiple quality assessment and enhancement stages (if needed). Following, we explain the role and function of these QAE stages.

3.1. Signal quality assessment

The pipeline includes six stages of quality assessment (QA1-QA6):

QA1: The duration of the captured audio signal is computed. Signals longer than or equal to 8 s can pass this stage, and those shorter than 8 s are rejected. This length of heart sound recording was chosen in accordance with the values employed in previous studies, as discussed in Section 2.1.

QA2: Using three features, including the degree of periodicity [59], frequency band ratio, and energy ratio [55], we aim to determine if a captured audio signal meets a minimum quality threshold for heart sound diagnostics. This quality assessment stage rejects signals that do not contain any heart sounds early in the pipeline. The degree of periodicity shows how periodic the signal is. As stated in Ref. [59], a heart sound signal with a lower noise level has a greater degree of periodicity. As a result, we can use this feature to detect non-periodic signals. The frequency band ratio and energy ratio represent the ratio of the energy concentration in the lower frequency range (between 24 Hz and 200 Hz) and the total energy of the signal. Unlike noise with a wide frequency range, heart sounds are generally low-frequency sounds [56]. As a result, these two features can give an indication of the level of noise in the signal. We determined the thresholds for these three features empirically by analyzing multiple heart sound and noise signals from the Pascal dataset [60]. This dataset contains heart sounds and noise signals that were captured using mobile phones. Using the thresholds mentioned above, we could accurately detect recordings containing heart sounds with an accuracy over 95%. In the case the values of the features computed for the captured signal were higher than predefined thresholds, the signal will pass this stage. Otherwise, it will be rejected. A signal is classified as acceptable if one or both of the following conditions holds:

- Degree of Periodicity $\geq$ 1.6 AND Energy Ratio $\geq$ 0.4 AND Frequency Band Ratio $\geq$ 0.3
- Degree of Periodicity $\geq$ 3 AND Energy Ratio $\geq$ 0.4 AND Frequency Band Ratio $\geq$ 0.2

In all other cases, the signal is rejected. Fig. 2 illustrates the phonocardiogram of an acceptable heart sound signal and Fig. 3 shows the phonocardiogram of a rejected audio signal.

QA3: This quality assessment stage determines whether the signal is clean enough to be used in a data-driven heart abnormality detection model. It calculates the signal-to-noise ratio (SNR) and an SNR greater than or equal to 10, is considered clean. Otherwise, the signal will go through quality enhancement stages. This threshold was determined by analyzing multiple heart sounds from healthy and pathologic subjects available in the Pascal dataset. In fact, we estimated the SNR values for different heart sounds with different noise levels. Then we determined this threshold by calculating the average SNR values of the recordings with the required quality. Fig. 4 shows the phonocardiogram of a clean heart sound signal with an SNR equal to 15, and Fig. 5 illustrates a signal with an SNR equal to 5 that is passed towards quality enhancement stages. SNR is calculated using the following equation (as in Ref. [45]):

$$SNR = 10 \times \log \frac{\text{signal power}}{\text{noise power}} \tag{2}$$

Where *signal power* is the power of the heart sound signal, and *noise power* represents the power of the noise in the signal, calculated by comparing the original and noisy heart sound signals. Algorithm 1 shows the pseudocode for SNR calculation.

Algorithm 1. SNR calculation

Algorithm 1 SNR calculation

Input: noisy_signal, clean_signal
Output: SNR

```
noise ← noisy_signal - clean_signal;
SNR ← 10*Log10(sum(power(clean_signal,2))/
      sum(power(noise,2)));
return SNR;
```

QA4: like QA2, the purpose of this stage is to compute the length of the heart sound recording with the difference that the signals longer or equal to 6 s are kept, and shorter signals are rejected. **QA5 and QA6**: in these two stages, SNR is computed (similar to QA3).

3.2. Signal quality enhancement

This section provides the details of the quality enhancement stages (QE1 and QE2) of the pipeline.

QE1: In this stage of quality enhancement, an artifact removal algorithm is applied to the signal. According to Ref. [46], artifacts are short-duration transient distortions in the signal. Thus, the artifact removal algorithm aims to remove such transient noises from the signal.

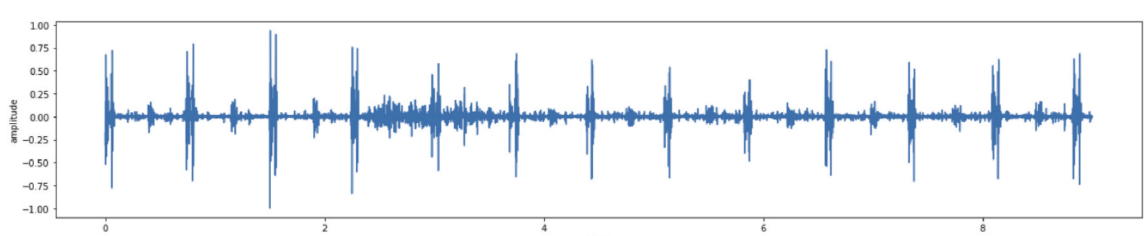

Fig. 2. Phonocardiogram of a signal that is detected as a heart sound signal in the QA2 stage of the pipeline (periodicity = 9.5, energy ratio = 0.90, frequency band ratio = 0.36).

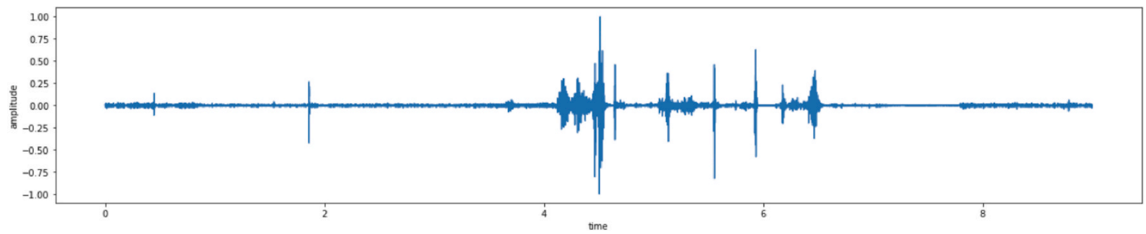

Fig. 3. Phonocardiogram of an audio signal that is classified as noise signal and rejected in the QA2 stage of the pipeline (frequency band ratio = 0.08).

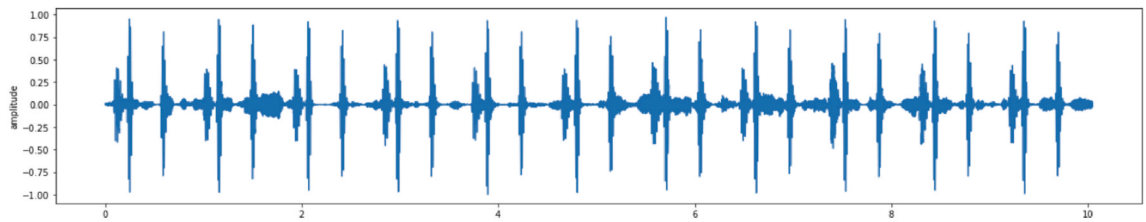

Fig. 4. Phonocardiogram of a heart sound signal that is classified as clean (SNR = 15) in the QA3 stage of the pipeline. This Signal is classified as clean.

Algorithm 2 shows the pseudocode of the artifact removal algorithm. After applying this algorithm, the length and the SNR are calculated respectively (QA4 and QA5 stages), and if they were higher than the required thresholds (as discussed in Section 3.1) signal will be considered clean. Otherwise, it will go through the second quality enhancement stage (QE2).

Algorithm 2. Artifact removal

Algorithm 2 Artifact removal

```
Input: noisy_signal
Output: denoised_signal
num_fft ← 256;
hop_len ← 64;
denoised_signal ← [];
coeffs ← abs(short_time_fourier_transform
(noisy_signal,num_fft,hop_len));
segment_len ← num_samples(noisy_signal)//
num_frequency_bands(noisy_signal);
threshold ← mean(coeffs[(num_fft/2)-13,:])+std(coeffs[num_fft/2)-13,:]);
while i<num_frequency_bands(noisy_signal) do
    if coeffs[(num_fft/2)-13,i]<threshold then
        append(denoised_signal,
        noisy_signal[i*segment_len:i*segment_len
        +segment_len]);
    else
        Continue;
    end if
end while
return denoised_signal;
```

QE2: in this stage, a continuous noise removal algorithm based on wavelet analysis is applied to the heart sound signal. Table 1 shows the parameters of the wavelet-based denoising algorithm. We followed the wavelet parameters in Ref. [45]. As for the wavelet decomposition level, we used the software library's scale levels 2–10, where 2 is a low but perceptible level of noise reduction and increases gradually to level 10 if needed. This minimizes the amount of heart sound signal data that is removed from the recording. Algorithm 3 shows the pseudocode of the wavelet-based continuous noise removal algorithm.

Algorithm 3. Continuous noise removal

Algorithm 3 Continuous noise removal

```
Input: noisy_signal, clean_signal
Output: denoised_signal
decomposition_level ← 1;
SNR ← GetSNR (noisy_signal,clean_signal);
while SNR<10 do
    decomposition_level ← decomposition_level+1;
    denoised_signal ← WaveletDenoising
    (noisy_signal,minimaxi,hard,mln,coeif5, decomposition_level);
    SNR ← GetSNR(denoised_signal,clean_signal);
    if decomposition_level == 10 AND SNR<10 then
        Break;
    end if
end while
return denoised_signal;
```

3.3. User feedback

The QAE pipeline allows us to give feedback to users regarding the quality of the captured audio signal at four different feedback points (F1–F4). Following, we provide the details of the feedback given at each of these feedback points.

F1: the captured signals that are shorter than 8 s are rejected at the QA2 stage. Therefore, at this point of the pipeline, the user can be informed that the captured signal is short, and a longer signal must be captured.

F2: noise signals that do not contain heart sounds are rejected at the QA3 stage. This could happen in cases where the user places the microphone sensor on an area of the chest that is not close enough to the heart. Consequently, the user can be informed that the microphone sensor is in a wrong place, and no heart sound was detected in the signal.

F3: heart sound signals that are shorter than 6 s are rejected at the

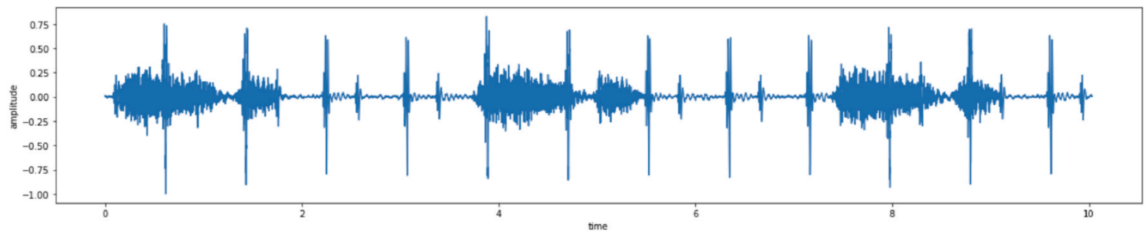

Fig. 5. Phonocardiogram of a heart sound signal that is classified as noisy (SNR = 5) in the QA3 stage of the pipeline and passed towards quality enhancement stages.

Table 1
Parameters and their corresponding values for wavelet-based denoising algorithm (adapted from Ref. [45]).

Wavelet Parameter	Value
Wavelet	Coeif5
Threshold selection	Minimaxi
Type of thresholding	Hard thresholding
Rescaling function	Multiple Level Noise estimate (MLN)
Decomposition level	From 2 to 10

QA4 stage. If we had a signal shorter than 6 s at this point of the pipeline, it would mean that at least 25% of the captured signal was removed in the artifact removal stage (QE1). This is because only signals that are at least 8 s long can pass the first stage of quality assessment (QA1). Therefore, the user can be informed that the heart sound signal is corrupted with a large amount of transient noise.

F4: heart sound signals with an SNR lower than 10 are rejected at the QA5 stage. Given that both artifact removal and continuous noise removal algorithms are applied before this stage of quality assessment, a low SNR value at this stage would mean that neither of these quality enhancement algorithms could considerably decrease the noise level of the signal. As a result, the user can be informed that a strong continuous or transient noise is present in the signal.

4. Pipeline validation

In this section, we overview the QAE pipeline validation process. In order to validate the QAE pipeline, we conducted a survey of a group of clinicians. This survey includes a subjective listening test with 20 heart sound recordings with a variety of noise types and intensities. The results of this survey is used to inform and validate the design decisions within the QAE pipeline. Following, we first explain the data generation process and provide the details of the synthetic heart sound dataset we used in our survey. Then, the details of the survey will be provided.

4.1. Dataset generation

We generated a synthetic heart sound dataset by adding different noise types with different intensities to clean heart sound recordings. This dataset was utilized in a subjective listening test that is described in Section 4.2. Table 2 summarizes the heart sound types, noise types, and SNRs that have been used to generate the dataset.

As shown in Table 2, five different clean heart sound recordings were used, including one normal, two murmurs, and two extra heart sounds. Some of these clean heart sounds were chosen from Pascal dataset and some others from available recordings on YouTube. Then eleven noise types (as summarized in Table 2) with different intensities were added to these clean heart sound recordings. The noises that were added to clean signals can be categorized into four classes: color, internal, movement, and ambient noise. Color noises were generated through simulation. Internal and ambient noises were collected from various publicly available datasets, and movement noises were captured using a mobile phone from the body surface. These noises were added to the clean heart sound signals in five different SNR levels: -5, 0, 5, 10, and 15.

Table 2
Clean heart sounds, noise types, and SNRs used to generate synthetic noisy heart sound recordings.

Heart Sound	Noise Type	SNR
Normal, Murmur (2 types), Extra Heart Sound (S3, S4)	Color Noise (white, pink, red), Internal Noise (deep breathing, fast breathing, coughing, digestive sounds), Movement Noise (sensor movement, body movement), Ambient noise (speech, music)	−5, 0, 5, 10, 15

Table 3 summarizes the heart sound type, noise type, noise category, SNR, and length of each recording available in the dataset. Using the procedure mentioned above, we generated eleven noisy heart sound recordings (Recordings 2–7 and 9–13). We fed these recordings into the QAE pipeline. Some were classified as clean, a few were rejected, and quality enhancement algorithms of the QAE pipeline were applied to seven recordings. In addition to the eleven noisy recordings that are summarized in Table 3, we also used these seven recordings in our subjective listening test (Recordings 14–20). Lastly, we added two clean recordings to the dataset, including a short-length normal heart sound and a murmur recording (Recordings 1 and 8).

4.2. Survey

In order to inform and validate the design decisions within the QAE pipeline, we designed a survey that includes a subjective listening test. We will use the results of the survey to evaluate the fundamental design decision in the pipeline. This survey was developed using the Go Listen platform [61] and is available online.[1] Before the actual survey, we conducted a pilot study with a group of non-clinicians to verify that the answers returned were suitable in terms of structure and granularity. Thirteen clinicians with auscultation skills and experience participated in the actual survey. These include one general practitioner, one cardiologist, and eleven consultants with different specialties.

Six multiple choice questions were presented in the survey and are reproduced in Appendix 1 (Table 8). The survey starts with a question regarding the profession of the respondents (Question 1). Then, the respondents were asked to listen to 20 heart sound recordings (as summarized in Table 3) and determine if each recording was clear and long enough to be used as part of a diagnostic exercise (Question 2). After the listening test, the respondents answered four questions regarding the criteria they used to form their heart sound quality judgments (Questions 3–6). One of these questions is about the minimum number of heartbeats, and the other ones were asked to determine the impact of different noise types on auscultation.

The survey follows three goals: first, by asking clinicians to rate the quality of the heart sound recordings, we try to find out what types and intensities of noise could make heart sounds undiagnosable. This will help us to determine the characteristics of a good quality heart sound that could be used for diagnostic purposes and consequently utilized in a

[1] https://golisten.ucd.ie/task/ab-test/6151d187b9536a2771028510.

Table 3
Details of the heart sound recordings available in the dataset used in subjective listening test.

Recording #	Heart Sound Type	Noise Type	Noise Category	SNR	Length (sec)
1	Normal	-	-	35	2
2	S3	Body movement	Movement	5	10
3	S3	Coughing	Internal	−5	10
4	S3	Pink	Color	0	10
5	S4	Music	Ambient	10	10
6	S4	Red	Color	5	10
7	S4	Speech	Ambient	15	10
8	Murmur 1	-	-	35	10
9	Murmur 1	Fast breathing	Internal	10	10
10	Murmur 1	Sensor movement	Movement	0	10
11	Murmur 2	Digestive sound	Internal	0	10
12	Murmur 2	White	Color	15	10
13	Normal	Deep breathing	Internal	5	10
14	S3	Body movement	Movement	14	7.6
15	S3	Coughing	Internal	3	9.8
16	S3	Pink	Color	3	8.4
17	S4	Red	Color	6.5	8
18	Murmur 1	Sensor movement	Movement	16	8
19	Murmur 2	Digestive sound	Internal	1	10
20	Normal	Deep breathing	Internal	6	9.5

data-driven classification system. Second, comparing the respondents' ratings with the outputs of the quality assessment stages of the QAE pipeline will let us find out whether our threshold of quality is accurate or not. Lastly, by asking respondents to rate both the original and processed recordings, we can find out if the quality enhancement algorithms applied to recordings change the clinicians' opinions regarding the recordings' quality. This will enable us to determine the impact of the quality enhancement algorithms of the QAE pipeline on the diagnosability of the heart sound recordings.

5. Results

5.1. Heart sound quality ratings

Fig. 6 depicts the subjective quality ratings of the 20 heart sound recordings (R1-R20). Blue and red colors show the proportions of the respondents who selected "No" and "Yes", respectively.

5.1.1. Quality ratings and noise categories

Table 4 summarizes the average quality ratings and standard deviations for four different noise type groupings. As shown in Table 4, recordings contaminated with ambient noise have a very low average quality rating. In fact, the majority of the respondents believed that those recordings do not have the required quality. Heart sounds with color noises received an average quality rating of 0.46. In the case of color noise, all respondents determined the recording with the red noise (recording 6) as a good quality heart sound, while recordings with white and pink noise (recording 4 and 12) have pretty low average quality ratings. Internal and movement noise groupings have roughly similar average quality ratings with 0.56 and 0.58, respectively.

5.1.2. Quality ratings and noise duration

We categorized the recordings into two groups in terms of the noise duration: heart sounds contaminated with short-duration noises, including movement noise, digestive sounds, and coughing, and the ones that have long-duration noises, including color noise, ambient noise, fast breathing, and deep breathing. Table 5 summarizes the average quality ratings and standard deviations for recordings with short- and long-duration noises. As shown in Table 5, the average quality rating for heart sound recordings with long-duration noises is roughly half of those with short-duration noises.

5.1.3. Quality enhancement impact

As we mentioned in Section 4.1, quality enhancement algorithms were applied to seven noisy heart sound recordings. Recordings number 14 to 20 are the outputs of denoising algorithms (as summarized in Table 3). Table 6 shows the average quality ratings and SNRs of those recordings before and after applying the denoising algorithms. All of the recordings that quality enhancement algorithms were applied to have SNR values between −5 and 5.

As shown in Table 6, in the case of four recordings (Recordings 2, 3, 4, and 10), applying the quality enhancement algorithms increased the average quality ratings. As for the other three recordings, the average quality ratings decreased (Recordings 6, 11, and 13). Also, we can see that applying denoising algorithms led to an increase in the SNR values of all recordings, although in some cases, like recordings 11 and 13, this increase is marginal.

5.2. Number of heartbeats

Fig. 7 illustrates the proportions of respondents who selected different ranges for the minimum number of heartbeats needed in a heart sound recording to be diagnosable. As we can see, most of the respondents (76%) believe that they need to listen to at least 6 to 10 heartbeats at one location on the chest before using the heart sound towards a diagnostic. This range is aligned with the figures that we used in the quality assessment stages of the QAE pipeline for the minimum duration of the heart sound recordings.

5.3. Noise impact

This section provides the details of the responses to the last three questions of the survey. Fig. 8 shows the proportions of the survey respondents who rated the internal, movement, and ambient noise categories in terms of the disruptiveness.

As shown in Fig. 8, respondents believe that all these three categories of noise are, to some extent, disruptive. Most respondents determined internal and movement noises as somewhat disruptive, and a minority selected the limited disruption option. However, compared to the internal noise category, a larger majority (85%) of the respondents found movement noises as somewhat disruptive. We can see a similar pattern for the ambient noise category, with the difference being that 31% of the respondents believe that ambient noises are very disruptive. It is worth noting that the responses given to the last three questions of the survey are aligned with the average quality ratings of heart sounds with different noises, as reported in Section 5.1.1.

5.4. Quality ratings and pipeline outputs

In this section, we draw a comparison between the average quality ratings of the recordings in the listening test and the outputs of the QAE pipeline. To do so, we categorized the recordings into acceptable and unacceptable groups based on their average quality ratings. Heart sound recordings that over 50% of the respondents determined as good quality were considered acceptable and the other recordings were placed in the unacceptable category.

As shown in Table 7, in eight out of twenty cases, the subjective quality ratings are aligned with the outputs of the QAE pipeline, and in tweleve cases, they contradict. Out of these, the QAE pipeline recognized eight recordings as noisy due to a lower than threshold SNR, while respondents determined those as good quality heart sounds. In the

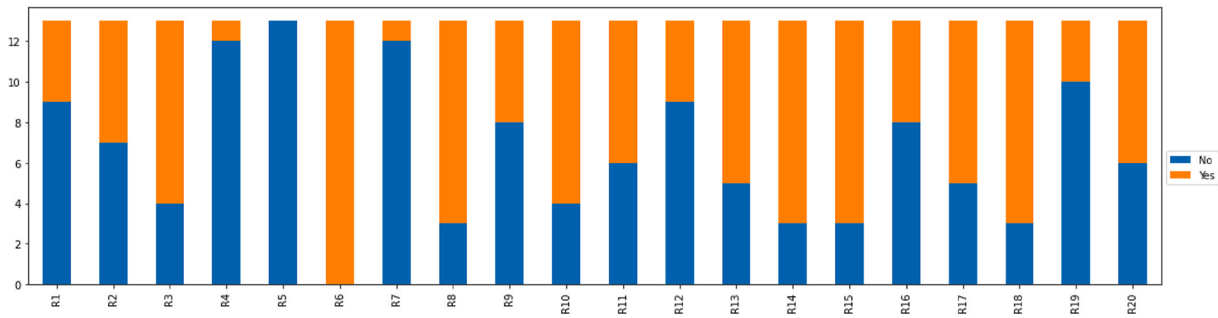

Fig. 6. The subjective heart sound quality ratings. Blue and red colors show the proportions of the respondents who chose “No” and “Yes”, respectively. 20 recordings are mapped along the horizontal axis and the vertical axis shows the number of respondents who chose any of the two options. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 4
Average quality ratings and standard deviations for each category of noise.

Recording #	Noise Category	Ratings Mean	Ratings std
4, 6, 12	Color	0.46	0.48
3, 9, 11, 13	Internal	0.56	0.13
2, 10	Movement	0.59	0.20
5, 7	Ambient	0.04	0.06

Table 5
Average quality ratings and standard deviations for recordings with long- and short-duration noise.

Recording #	Noise Duration	Ratings Mean	Ratings std
2, 3, 10, 11	Short	0.60	0.11
4, 5, 6, 7, 9, 12, 13	Long	0.35	0.36

Table 6
Average quality ratings and SNRs of seven heart sound recordings before and after applying the denoising algorithms.

Noisy Recording #	Denoised Recording #	Average Quality Rating Before QE	Average Quality Rating After QE	SNR Before QE	SNR After QE
2	14	0.46	0.77	5	14
3	15	0.69	0.77	−5	3
4	16	0.08	0.38	0	3
6	17	1.00	0.62	5	6.5
10	18	0.69	0.77	0	16
11	19	0.54	0.23	0	1
13	20	0.62	0.54	5	6

quality assessment stages of the pipeline, the recordings with an SNR below 10 are either passed towards quality enhancement stages or rejected. However, as we can see, in some cases, such recordings were determined as good-quality signals by respondents. One example of this is a recording that is contaminated by cough noise (Recording 3). While the SNR of this recording is −5, the respondents found it as a good quality heart sound.

There are also four cases where the signals were classified as clean by the QAE pipeline while survey respondents found them as low-quality heart sounds. An example of this is a recording with speech noise (Recording 7). While this signal has an SNR equal to 15, which is higher than the threshold, the respondents identified it as a low-quality recording.

The last column of Table 7 summarizes the feedbacks given when the captured signals are rejected. There are five cases where feedbacks are provided to users. Out of these, in three cases, feedbacks were accurate (Recording 1, 19, and 20), and in two cases (Recordings 15, 17), users were asked to recapture the heart sound while survey respondents found those signals as good quality heart sounds.

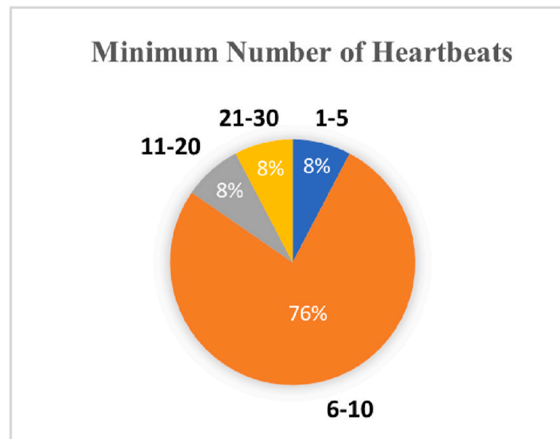

Fig. 7. The proportions of the respondents who selected four different ranges for the minimum number of heartbeats.

6. Discussion

6.1. Signal quality assessment

As we discussed in Section 3.1, the QAE pipeline includes six stages of signal quality assessment. Out of these, in two stages (QA1 and QA4), the duration of the signal is computed to determine if the signal contains a minimum number of heartbeats, and if a captured signal does not meet this quality threshold, it will be rejected. In Section 5.2, we showed that over 90% of the survey respondents stated they need to listen to at least six heartbeats at a particular auscultation site before they can use the heart sounds for diagnostics (Fig. 7). In other words, from the point of view of most respondents, a diagnosable heart sound must contain at least six heartbeats. As we discussed in Section 2.1, different durations of heart sound signals have been utilized by researchers to develop data-driven classification models. The findings from our survey confirm that a diagnosable heart sound recording is at least around 5 s long, which aligns with the studies that employed longer duration signals such as [24–28]. Also, 10 s length has been reported as sufficient by most clinicians who participated in our survey, which is in line with the approach taken in Ref. [29].

At four stages of the QAE pipeline (QA2, QA3, QA5, and QA6), the captured signal is analyzed for noise corruption. The subjective heart sound quality ratings (Section 5.1) show that clinicians determined 8 out of 20 recordings as undiagnosable. In other words, only 60% of the recordings were found clean enough for diagnostic tasks. These results

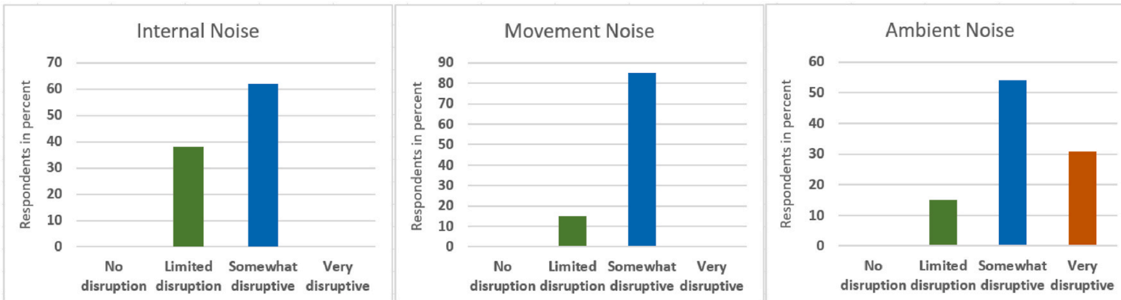

Fig. 8. The proportions of the respondents who chose each of the four options for each noise category.

Table 7
Groupings of the heart sound recordings in terms of quality, outputs of the QAE pipeline, and the provided feedbacks in the case of signal rejection.

Recording #	Acceptable Quality	QAE Pipeline Output	User Feedback
1	No	Rejected due to short length	Short-length signal. Recapture needed.
2	No	Passed to QE due to high noise	-
3	Yes	Passed to QE due to high noise	-
4	No	Passed to QE due to high noise	-
5	No	Accepted as clean	-
6	Yes	Passed to QE due to high noise	-
7	No	Accepted as clean	-
8	Yes	Accepted as clean	-
9	No	Accepted as clean	-
10	Yes	Passed to QE due to high noise	-
11	Yes	Passed to QE due to high noise	-
12	No	Accepted as clean	-
13	Yes	Passed to QE due to high noise	-
14	Yes	Accepted as clean	-
15	Yes	Rejected after QE due to high noise	A high level of noise was detected. Recapture needed.
16	No	Rejected after QE due to high noise	-
17	Yes	Rejected after QE due to high noise	A high level of noise was detected. Recapture needed.
18	Yes	Accepted as clean	-
19	No	Rejected after QE due to high noise	A high level of noise was detected. Recapture needed.
20	Yes	Rejected after QE due to high noise	A high level of noise was detected. Recapture needed.

indicate that it is necessary to estimate the quality of the captured signals in terms of noise contamination in the quality assessment stages of the pipeline and enhance or reject the signals that do not meet a specific threshold of quality.

Analysis of the heart sound quality ratings in Section 5.1 shows that the heart sounds with ambient noises received the lowest quality ratings from survey respondents compared to recordings contaminated by other noise types. The answers to questions regarding the impact of the different noise groupings on the heart sound quality in Section 5.3 also confirm that ambient noises were more disrupting compared to the other noise groupings (Fig. 8). Also, a considerably lower average quality rating for the heart sounds with long-duration noises indicates that long-duration noises have a more detrimental effect on the heart sounds' quality than short-duration ones. These findings are in agreement with the hypothesis that the duration and type of a noise corruption on the heart sound signal are indicators of signal quality. Previous studies [31–35] highlighted that data-driven heart anomaly detection systems could potentially misclassify heart sounds due to noise corruption. While minimizing ambient noise is recommended [36] and Coviello [37] highlighted the destructive impact of movement noise, internal noises (e.g. digestive and respiratory sounds) cannot be mitigated through environmental setups. The survey results are in agreement with the literature that all classes of noise can cause disruption but highlight that ambient noise was more disruptive.

Comparison of the average quality ratings of the recordings with the pipeline outputs in Section 5.4 shows that in twelve out of twenty cases, the pipeline's quality assessment disagreed with the opinions of the clinicians. In eight of those cases, the pipeline rejected signals that clinicians found of diagnosable quality and, in four of the cases, the signals were categorized as clean while the survey respondents deemed those as low-quality heart sounds. We should note that in this study, our goal is to understand the types and intensities of noise that cause clinicians to reject heart sounds due to their low quality. Ultimately, this will enable us to align the QAE pipeline with the clinician's opinions, which will lead to a lower chance of presenting undiagnosable heart sounds to the classification model.

As we discussed in Section 3.1, in three stages of the QAE pipeline (QA3, QA5, and QA6), the quality of the heart sound signal is estimated by calculating the SNR. Four recordings (Recording 5, 7, 9, and 12) were classified as clean heart sounds by the QAE pipeline while identified as undiagnosable by the respondents of the survey. These recordings were classified as clean by the pipeline because they all have higher than threshold SNRs. However, analyzing the relationship between the average quality ratings, duration, and type of noise contaminations indicates that these two noise characteristics are also indicators of the heart sound quality. These results highlight that SNR should not be relied upon in isolation to estimate heart sound quality and should be used in conjunction with other content feature analysis and signal characteristics such as type and duration of the noise. Studies such as [32,45] focused on evaluating the effectiveness of the quality enhancement algorithms on limited noise classes (pink/white noise). The findings of our study suggest that a more nuanced approach to noise is valuable where noise characteristics and SNR are considered together. Our future studies will compare the performance of data-driven heart sound classifiers for noisy and denoised versions of a signal to further establish the benefits of the pre-processing decision to denoise the signal for different noise types and durations.

6.2. *Signal quality enhancement*

In addition to the quality assessment stages, the QAE pipeline includes two stages of signal quality enhancement in which transient and continuous noises are removed from the captured signal. In Section 5.1.3 of the results, we observed the impact of the quality enhancement algorithms of the pipeline on the perceived quality of the heart sound recordings. Denoising algorithms were applied to seven recordings. Out of these, the average quality ratings went up in four cases and decreased in three cases. If we consider the cases where quality enhancement led to a decrease in average quality ratings (Recordings 6, 11, 13), we can observe that at least over 50% of respondents believed that these recordings were diagnosable before applying the denoising algorithms. These results indicate that applying quality enhancement algorithms to good quality heart sounds leads to an increased information loss and in fact, reduces the diagnosability of heart sounds.

As we discussed in Section 2.2, heart sound denoising has been recognized as a necessary pre-processing step towards building a heart abnormality detection system in the literature. A variety of denoising algorithms were applied irrespective of the noise content of the signal, which includes a range of methods, from simpler techniques like filtering to more complex ones such as wavelet-based denoising (e.g. Refs. [34,35,46]). As denoising algorithms must distinguish between noise and signal, some studies, e.g. Kumar et al. [48] try to identify noise-free parts of the heart sound to allow noise estimation. However, any denoising alters the signal and may introduce corrupting artifacts as well as restoring the heart sounds. The survey results also indicate that applying quality enhancement algorithms degrades the diagnosability of heart sounds in cases where enhancement is not required (Table 6). This finding reinforces the importance of assessing the heart sounds in terms of the noise content before applying denoising algorithms to such signals. Such a quality assessment enables us to limit the usage of denoising algorithms only to cases where quality enhancement is needed. In addition to this, it allows a decision on the type and aggressiveness of the denoising to be applied.

6.3. User feedback

While some studies (e.g. Refs. [52,54–56]) have attempted to classify the heart sound recordings in terms of the quality, their focus was a binary classification as to whether to discard low-quality recordings. However, for the mobile phone heart sound capture use case, this study sought to explore the quality as a continuum, where there is an expectation of some potential noise but an objective to deal with it through feedback to the user to assist in recapture with an acceptable quality or application of appropriate signal enhancement based on quality assessment.

In this regard, at four decision points in the QAE pipeline, feedbacks are provided to users regarding the quality of the heart sound capture. In Section 5.4, we showed that in cases where the pipeline rejected the captured signals, feedbacks were provided regarding the cause of the signal rejection. Such feedbacks can inform the users regarding the cause of an unacceptable heart sound capture. This, in turn, decreases the reliance on trained clinicians to capture and validate the heart sounds. However, it should be noted that the accuracy of such feedbacks is influenced by the accuracy of the heart sound quality assessment. In other words, to provide more accurate feedback, we need to improve the quality threshold that has been employed in the quality assessment stages of the pipeline.

7. Conclusion

The ability to capture heart sounds that can be used for diagnostic purposes independently from clinicians is essential to building a self-diagnostic or tele-medicine system for heart health screening. From a clinical point of view, a great advantage would be to establish a level of heart sound quality that enables the distinction between a normal and abnormal sounding heart which would, in turn, allow for further appropriate investigation. To address this problem, in this paper, we proposed a heart sound quality assessment and enhancement pipeline for signals captured by mobile phone devices.

In order to inform and validate the design decisions within the pipeline, a survey was conducted. We observed a high level of agreement between the survey results and fundamental design decisions in the pipeline. We showed that it is possible to automatically estimate the quality of the heart sound signals by analyzing the signal characteristics. We also showed that we could increase the diagnosability of low-quality heart sounds by applying quality enhancement algorithms. These findings indicate that the proposed pipeline can reduce our dependency on clinicians to capture valid heart sound signals. Survey results indicate that noise has a destructive impact on the diagnosability of heart sound signals. The type, intensity, and duration of the noise determine the severity of this harmful impact. As a result, the captured signals must be analyzed in terms of the quality and enhanced or rejected if not meet a minimum threshold of quality.

We observed that in a few cases, undiagnosable heart sounds were determined as clean signals by the QAE pipeline. The survey results indicate that such cases can be minimized by analyzing the characteristics of noise contamination such as intensity, duration, and type of the noise. As a result, in the future, we will explore the possibility of improving the signal quality assessment by designing a more complex quality threshold that not only includes signal-to-noise ratio but also takes into account the other important characteristics of noise contamination. Such an improvement in the pipeline will also enable us to provide more accurate and specific feedbacks regarding the cause of the signal rejection that will, in turn, reduce our dependency on trained clinicians for heart sound capture.

The survey results confirm the validity of the design decisions in the pipeline and shows the usefulness of the proposed pipeline from the point of view of clinicians. Findings from our study allow us to better understand the different ways we can improve the common approaches taken in the field regarding the assessment and enhancement of the heart sound signals in the future. In this study, we did not explore the impact of the proposed pipeline on the performance of the heart sound classification systems. In our next phase of work, we will develop a classification model and compare the performance of that model on both unprocessed signals and the ones processed by the QAE pipeline. The results of this experiment will be utilized to optimise the thresholds and algorithms used in the QAE pipeline and consequently increase the overall performance of the pipeline.

Another future direction will be to develop a mobile application prototype to investigate the heart sound capture process in real-world scenarios. Using this mobile prototype, we will investigate how can non-expert users be guided to capture valid heart sound signals using mobile phones. Such a prototype will also enable us to evaluate the QAE pipeline's capabilities in providing feedback to users and guiding them towards a successful heart sound capture.

Declaration of competing interest

This is to confirm that the authors do not have any conflict of interest relationships (commercial, financial, legal, professional) with any organizations or people that could influence our research.

Acknowledgement

This work was conducted with the financial support of the Science Foundation Ireland Centre for Research Training in Digitally-Enhanced Reality (d-real) under Grant No. 18/CRT/6224. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Appendix 1

Table 8
Survey questions and their corresponding choices

#	Question	Choices
1	Please specify your profession.	● Cardiologist ● Cardiology fellow ● General practitioner ● Medical student ● Other medical specialty ● None of the above
2	In your opinion, is this heart sound recording clear and long enough to be used as part of a diagnostic exercise? (This question is repeated 20 times – one for each recording)	● Yes ● No
3	How many heartbeats would you need to listen to at one location on the chest before you can use the sound as part of a diagnostic exercise?	● 1-5 ● 6-10 ● 11-20 ● 21-30 ● Over 30
4	In the scenario where internal sounds (such as respiratory or digestive sounds) are present when you are listening to a patient's heartbeats: To what extent do these sounds disrupt your ability to assess the heart sounds, such that you might even have to re-listen?	● No disruption ● Limited disruption ● Somewhat disruptive ● Very disruptive
5	In the scenario where sounds due to the movement (such as chest piece or body movement) are present when you are listening to a patient's heartbeats: To what extent do these sounds disrupt your ability to assess the heart sounds, such that you might even have to re-listen?	● No disruption ● Limited disruption ● Somewhat disruptive ● Very disruptive
6	In the scenario where ambient sounds (such as phone ringing or people talking) are present when you are listening to a patient's heartbeats: To what extent do these sounds disrupt your ability to assess the heart sounds, such that you might even have to re-listen?	● No disruption ● Limited disruption ● Somewhat disruptive ● Very disruptive

References

[1] WHO, URL, https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1. (Accessed 7 December 2021).
[2] Cardiovascular diseases (cvds), URL, https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds. (Accessed 7 December 2021).
[3] M.R. Montinari, S. Minelli, The first 200 years of cardiac auscultation and future perspectives, J. Multidiscip. Healthc. 12 (2019) 183–189, https://doi.org/10.2147/JMDH.S193904, pMID: 30881010 PMCID: PMC6408918.
[4] A.J. Taylor (Ed.), Learning Cardiac Auscultation: from Essentials to Expert Clinical Interpretation, Springer-Verlag, London, 2015, https://doi.org/10.1007/978-1-4471-6738-9. URL, https://www.springer.com/gp/book/9781447167372.
[5] E.M. Brown, Heart Sounds Made Easy, Elsevier Health Sciences, 2008.
[6] J. Rubin, R. Abreu, A. Ganguli, S. Nelaturi, I. Matei, K. Sricharan, Recognizing abnormal heart sounds using deep learning, arXiv:1707.04642 [cs]ArXiv: 1707.04642 (10 2017). URL, http://arxiv.org/abs/1707.04642.
[7] B. Xiao, Y. Xu, X. Bi, J. Zhang, X. Ma, Heart sounds classification using a novel 1-d convolutional neural network with extremely low parameter consumption, Neurocomputing 392 (2020) 153–159, https://doi.org/10.1016/j.neucom.2018.09.101.
[8] S. Alam, R. Banerjee, S. Bandyopadhyay, Murmur detection using parallel recurrent & convolutional neural networks, arXiv:1808.04411 [cs, eess, stat]ArXiv: 1808.04411 (8 2018). URL, http://arxiv.org/abs/1808.04411.
[9] T. Li, Y. Yin, K. Ma, S. Zhang, M. Liu, Lightweight end-to-end neural network model for automatic heart sound classification, Information 12 (2) (2021) 54, https://doi.org/10.3390/info12020054.
[10] M. Alkhodari, L. Fraiwan, Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings, Comput. Methods Progr. Biomed. 200 (2021), 105940, https://doi.org/10.1016/j.cmpb.2021.105940.
[11] Y. Chen, S. Wei, Y. Zhang, Classification of heart sounds based on the combination of the modified frequency wavelet transform and convolutional neural network, Med. Biol. Eng. Comput. 58 (9) (2020) 2039–2047, https://doi.org/10.1007/s11517-020-02218-5.
[12] Global health workforce statistics database, URL, https://www.who.int/data/gho/data/themes/topics/health-workforce. (Accessed 7 December 2021).
[13] C.S. Wood, M.R. Thomas, J. Budd, T.P. Mashamba-Thompson, K. Herbst, D. Pillay, R.W. Peeling, A.M. Johnson, R.A. McKendry, M.M. Stevens, Taking connected mobile-health diagnostics of infectious diseases to the field, Nature 566 (7745) (2019) 467–474, https://doi.org/10.1038/s41586-019-0956-2.
[14] B. Udugama, P. Kadhiresan, H.N. Kozlowski, A. Malekjahani, M. Osborne, V.Y. C. Li, H. Chen, S. Mubareka, J.B. Gubbay, W.C.W. Chan, Diagnosing covid-19: the disease and tools for detection, ACS Nano (2020), https://doi.org/10.1021/acsnano.0c02624 acsnano.0c02624PMID: 32223179 PMCID: PMC7144809.
[15] R.J. Shaw, Q. Yang, A. Barnes, D. Hatch, M.J. Crowley, A. Vorderstrasse, J. Vaughn, A. Diane, A.A. Lewinski, M. Jiang, J. Stevenson, D. Steinberg, Self-monitoring diabetes with multiple mobile health devices, J. Am. Med. Inf. Assoc. 27 (5) (2020) 667–676, https://doi.org/10.1093/jamia/ocaa007.
[16] S. Bhatt, R. Isaac, M. Finkel, J. Evans, L. Grant, B. Paul, D. Weller, Mobile technology and cancer screening: lessons from rural India, J. Global Health 8 (2) (2018).
[17] S.R. Thiyagaraja, R. Dantu, P.L. Shrestha, A. Chitnis, M.A. Thompson, P. T. Anumandla, T. Sarma, S. Dantu, A novel heart-mobile interface for detection and classification of heart sounds, Biomed. Signal Process Control 45 (2018) 313–324, https://doi.org/10.1016/j.bspc.2018.05.008.
[18] S.-H. Kang, B. Joe, Y. Yoon, G.-Y. Cho, I. Shin, J.-W. Suh, Cardiac auscultation using smartphones: pilot study, JMIR mHealth uHealth 6 (2) (2018), e49, https://doi.org/10.2196/mhealth.8946.
[19] Core digital stethoscope - electronic stethoscopes | eko, URL, https://shop.ekohealth.com/products/core-digital-stethoscope. (Accessed 16 December 2021).
[20] Littmann electronic stethoscope model 3200, URL, https://www.littmann.com/3M/en_US/littmann-stethoscopes/. (Accessed 16 December 2021).
[21] Jabes electronic stethoscope, URL, https://www.allheart.com/jabes-electronic-stethoscope/p/jsjabes3//. (Accessed 16 December 2021).
[22] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, H. Ghaemmaghami, C. Fookes, Understanding the importance of heart sound segmentation for heart anomaly detection, arXiv:2005.10480 [cs, eess, q-bio]ArXiv: 2005.10480 (9 2020). URL, http://arxiv.org/abs/2005.10480.

[23] D.M. Nogueira, C.A. Ferreira, E.F. Gomes, A.M. Jorge, Classifying heart sounds using images of motifs, mfcc and temporal features, J. Med. Syst. 43 (6) (2019) 168, https://doi.org/10.1007/s10916-019-1286-5.

[24] V. Arora, R. Leekha, R. Singh, I. Chana, Heart sound classification using machine learning and phonocardiogram, Mod. Phys. Lett. B 33 (26) (2019), 1950321, https://doi.org/10.1142/S0217984919503214.

[25] S. Aziz, M.U. Khan, M. Alhaisoni, T. Akram, M. Altaf, Phonocardiogram signal processing for automatic diagnosis of congenital heart disorders through fusion of temporal and cepstral features, Sensors 20 (13) (2020) 3790, https://doi.org/10.3390/s20133790.

[26] M.U. Khan, S. Aziz, K. Iqtidar, G.F. Zaher, S. Alghamdi, M. Gull, A two-stage classification model integrating feature fusion for coronary artery disease detection and classification, Multimed. Tool. Appl. 4 (2021), https://doi.org/10.1007/s11042-021-10805-3. URL, http://link.springer.com/10.1007/s11042-021-10805-3. (Accessed 29 June 2021).

[27] P. Langley, A. Murray, Heart sound classification from unsegmented phonocardiograms, Physiol. Meas. 38 (8) (2017) 1658–1670, https://doi.org/10.1088/1361-6579/aa724c, publisher: IOP Publishing.

[28] N. Mei, H. Wang, Y. Zhang, F. Liu, X. Jiang, S. Wei, Classification of heart sounds based on quality assessment and wavelet scattering transform, Comput. Biol. Med. 137 (2021), 104814, https://doi.org/10.1016/j.compbiomed.2021.104814.

[29] K.-K. Tseng, C. Wang, Y.-F. Huang, G.-R. Chen, K.-L. Yung, W.-H. Ip, Cross-domain transfer learning for pcg diagnosis algorithm, Biosensors 11 (4) (2021) 127, https://doi.org/10.3390/bios11040127.

[30] Cardiac exam ii: auscultation | protocol, URL, https://www.jove.com/v/10124/cardiac-exam-ii-auscultation. (Accessed 22 October 2021).

[31] H. Naseri, M. Homaeinezhad, H. Pourkhajeh, Noise/spike detection in phonocardiogram signal as a cyclic random process with non-stationary period interval, Comput. Biol. Med. 43 (9) (2013) 1205–1213, https://doi.org/10.1016/j.compbiomed.2013.05.020.

[32] D. Gradolewski, G. Magenes, S. Johansson, W. Kulesza, A wavelet transform-based neural network denoising algorithm for mobile phonocardiography, Sensors 19 (4) (2019) 957, https://doi.org/10.3390/s19040957.

[33] H. Tang, T. Li, T. Qiu, Noise and disturbance reduction for heart sounds in cycle-frequency domain based on nonlinear time scaling, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 57 (2) (2010) 325–333, https://doi.org/10.1109/TBME.2009.2028693.

[34] P.K. Jain, A.K. Tiwari, An adaptive thresholding method for the wavelet based denoising of phonocardiogram signal, Biomed. Signal Process Control 38 (2017) 388–399, https://doi.org/10.1016/j.bspc.2017.07.002.

[35] S.R. Messer, J. Agzarian, D. Abbott, Optimal wavelet denoising for phonocardiograms, Microelectron. J. 32 (12) (2001) 931–941, https://doi.org/10.1016/S0026-2692(01)00095-7.

[36] D.M. Shindler, Practical cardiac auscultation, Crit. Care Nurs. Q. 30 (2) (2007) 166–180, https://doi.org/10.1097/01.CNQ.0000264260.20994.36.

[37] J.S. Coviello, Auscultation Skills: Breath & Heart Sounds, Lippincott Williams & Wilkins, 2013.

[38] N. Ranganathan, V. Sivaciyan, F.B. Saksena, The Art and Science of Cardiac Physical Examination, JP Medical Ltd, 2015.

[39] J.P. de Vos, M.M. Blanckenberg, Automated pediatric cardiac auscultation, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 54 (2) (2007) 244–252, https://doi.org/10.1109/TBME.2006.886660, event: IEEE Transactions on Biomedical Engineering.

[40] C.N. Gupta, R. Palaniappan, S. Swaminathan, S.M. Krishnan, Neural network classification of homomorphic segmented heart sounds, Appl. Soft Comput. 7 (1) (2007) 286–297, https://doi.org/10.1016/j.asoc.2005.06.006.

[41] S. Vernekar, S. Nair, D. Vijaysenan, R. Ranjan, A novel approach for classification of normal/abnormal phonocardiogram recordings using temporal signal analysis and machine learning, in: 2016 Computing in Cardiology Conference (CinC), IEEE, 2016, pp. 1141–1144.

[42] W. Zhang, J. Han, Towards heart sound classification without segmentation using convolutional neural network, in: 2017 Computing in Cardiology (CinC), IEEE, 2017, pp. 1–4.

[43] Z. Abduh, E.A. Nehary, M. Abdel Wahed, Y.M. Kadah, Classification of heart sounds using fractional fourier transform based mel-frequency spectral coefficients and traditional classifiers, Biomed. Signal Process Control 57 (2020), 101788, https://doi.org/10.1016/j.bspc.2019.101788.

[44] P. Chen, Q. Zhang, Classification of heart sounds using discrete time-frequency energy feature based on s transform and the wavelet threshold denoising, Biomed. Signal Process Control 57 (2020), 101684, https://doi.org/10.1016/j.bspc.2019.101684.

[45] D. Gradolewski, G. Redlarski, Wavelet-based denoising method for real phonocardiography signal recorded by mobile devices in noisy environment, Comput. Biol. Med. 52 (2014) 119–129, https://doi.org/10.1016/j.compbiomed.2014.06.011.

[46] A.K. Kumar, G. Saha, Improved computerized cardiac auscultation by discarding artifact contaminated pcg signal sub-sequence, Biomed. Signal Process Control 41 (2018) 48–62, https://doi.org/10.1016/j.bspc.2017.11.001.

[47] M.N. Ali, E.-S.A. El-Dahshan, A.H. Yahia, Denoising of heart sound signals using discrete wavelet transform, Circ. Syst. Signal Process. 36 (11) (2017) 4482–4497, https://doi.org/10.1007/s00034-017-0524-7.

[48] D. Kumar, P. Carvalho, M. Antunes, R.P. Paiva, J. Henriques, Noise detection during heart sound recording using periodicity signatures, Physiol. Meas. 32 (5) (2011) 599–618, https://doi.org/10.1088/0967-3334/32/5/008.

[49] T. Li, T. Qiu, H. Tang, Optimum heart sound signal selection based on the cyclostationary property, Comput. Biol. Med. 43 (6) (2013) 607–612, https://doi.org/10.1016/j.compbiomed.2013.03.002.

[50] Q.-u.-A. Mubarak, M.U. Akram, A. Shaukat, F. Hussain, S.G. Khawaja, W.H. Butt, Analysis of pcg signals using quality assessment and homomorphic filters for localization and classification of heart sounds, Comput. Methods Progr. Biomed. 164 (2018) 143–157, https://doi.org/10.1016/j.cmpb.2018.07.006.

[51] H. Naseri, M.R. Homaeinezhad, Computerized quality assessment of phonocardiogram signal measurement-acquisition parameters, J. Med. Eng. Technol. 36 (6) (2012) 308–318, https://doi.org/10.3109/03091902.2012.684832.

[52] E. Grooby, J. He, J. Kiewsky, D. Fattahi, L. Zhou, A. King, A. Ramanathan, A. Malhotra, G.A. Dumont, F. Marzbanrad, Neonatal heart and lung sound quality assessment for robust heart and breathing rate estimation for telehealth applications, IEEE J. Biomed. Health Inf. (2020), https://doi.org/10.1109/JBHI.2020.3047602, 1–1.

[53] D.H. Eiríksdóttir, R.G. Sæderup, D. Riknagel, H. Zimmermann, M. Plocharski, J. Hansen, J.J. Struijk, S.E. Schmidt, Quality assessment of maternal and fetal cardiovascular sounds recorded from the skin near the uterine arteries during pregnancy, in: 2019 Computing in Cardiology (CinC), IEEE, 2019, p. 1.

[54] D.B. Springer, T. Brennan, N. Ntusi, H.Y. Abdelrahman, L.J. Zühlke, B.M. Mayosi, L. Tarassenko, G.D. Clifford, Automated signal quality assessment of mobile phone-recorded heart sound signals, J. Med. Eng. Technol. 40 (7–8) (2016) 342–355, https://doi.org/10.1080/03091902.2016.1213902.

[55] D. Das, R. Banerjee, A.D. Choudhury, S. Bhattacharya, P. Deshpande, A. Pal, K. M. Mandana, Novel features from autocorrelation and spectrum to classify phonocardiogram quality, in: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2017, pp. 4516–4520.

[56] H. Tang, M. Wang, Y. Hu, B. Guo, T. Li, Automated signal quality assessment for heart sound signal by novel features and evaluation in open public datasets, BioMed Res. Int. 2021 (2021) 1–15, https://doi.org/10.1155/2021/7565398.

[57] C. Liu, D. Springer, Q. Li, B. Moody, R.A. Juan, F.J. Chorro, F. Castells, J.M. Roig, I. Silva, A.E.W. Johnson, Z. Syed, S.E. Schmidt, C.D. Papadaniil, L. Hadjileontiadis, H. Naseri, A. Moukadem, A. Dieterlen, C. Brandt, H. Tang, M. Samieinasab, M. R. Samieinasab, R. Sameni, R.G. Mark, G.D. Clifford, An open access database for the evaluation of heart sound algorithms, Physiol. Meas. 37 (12) (2016) 2181–2213, https://doi.org/10.1088/0967-3334/37/12/2181.

[58] D.S. Panah, A. Hines, S. Mckeever, Exploring Composite Dataset Biases for Heart Sound Classification, AICS, 2020, pp. 145–156.

[59] T. Li, H. Tang, T. Qiu, Y. Park, Best subsequence selection of heart sound recording based on degree of sound periodicity, Electron. Lett. 47 (15) (2011) 841, https://doi.org/10.1049/el.2011.1693.

[60] P. Bentley, G. Nordehn, M. Coimbra, S. Mannor, The pascal classifying heart sounds challenge 2011 (chsc2011), URL, http://www.peterjbentley.com/heartchallenge/index.html.

[61] D. Barry, Q. Zhang, P.W. Sun, A. Hines, Go listen: an end-to-end online listening test platform, J. Open Res. Software 9 (1) (2021) publisher: Ubiquity Press.