IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 8 October 2023, accepted 15 October 2023, date of publication 18 October 2023, date of current version 31 October 2023.
Digital Object Identifier 10.1109/ACCESS.2023.3325749

RESEARCH ARTICLE

HeartWave: A Multiclass Dataset of Heart Sounds for Cardiovascular Diseases Detection

SAMI ALRABIE AND AHMED BARNAWI
Faculty of Computing and Information Technology (FCIT), King Abdulaziz University, Jeddah 21589, Saudi Arabia
Corresponding author: Sami Alrabie (sawadalrabie@stu.kau.edu.sa)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board of the Department of Research and Studies in Health Affairs in Taif City, King Abdulaziz City for Science and Technology (KACST), KSA (Registration number: HAP-02-T-067 approval and approval number: 719).

ABSTRACT The auscultation of heart sounds has proven advantageous for the early diagnosis of cardiovascular conditions. Various methods have been proposed for the automatic analysis of heart sounds to reduce subjectivity in diagnosis and alleviate physicians' workload. However, the effectiveness of these methods heavily depends on the amount and quality of the heart sound datasets used and the availability of publicly accessible datasets that include the most common and difficult classes. In this study, we introduce the HeartWave dataset, a comprehensive heart sound dataset comprising recordings from nine distinct classes of the most common heart sounds from all classes and subclasses of cardiovascular diseases, documented, with enough samples, good quality, and well labelled, with a focus on the hard and difficult cases of diagnosis. The dataset includes a total of 1353 recordings of heart sounds. Notably, this dataset includes extremely rare and difficult-to-diagnose classes. In order to establish a reliable reference standard, a team of experienced cardiologists actively participated in the entire annotation process. The length of audio recordings is substantial, allowing for the extraction of multiple heartbeats from a single recording through the use of segmentation techniques. Moreover, our dataset takes into consideration the standard cardiologist practices to enable the capture of specific heart sounds associated with corresponding clinical locations. According to our post analysis of the dataset, the average signal-to-noise ratio of our proposed dataset surpasses that of the widely known PhysioNet/CinC 2016 public dataset by about two folds, ensuring a cleaner acoustic signal. Our proposed dataset provides a valuable resource for training and evaluating machine learning models aimed at automated heart sound classification and diagnosis.

INDEX TERMS Heart sound, murmurs, cardiovascular diseases detection, valvular diseases, artificial intelligence, signal to noise ratio (SNR).

I. INTRODUCTION

Cardiovascular diseases (CVD) present a significant global health challenge, emerging as the primary cause of death across the globe. CVD encompasses a wide range of conditions, such as coronary heart diseases, valve diseases, genetic heart diseases, and more. Disturbingly, projections indicate that the burden of CVD is expected to worsen in the upcoming years. By 2030, it is estimated that CVD will be responsible for over 23 million deaths worldwide [1]. These statistics underscore the urgent need to address heart diseases promptly and implement effective strategies to reduce their prevalence and associated mortality rates.

Furthermore, it is crucial to acknowledge the detrimental impact of the absence or scarcity of primary health centers and cardiologists, especially in developing or underdeveloped countries. The lack of accessible and timely healthcare services contributes to delays in diagnosing heart diseases, leading to increased fatalities or diminished quality of life within communities [2].

Heart sounds can be a valuable tool in the early diagnosis of cardiovascular disease (CVD), enabling timely intervention

The associate editor coordinating the review of this manuscript and approving it for publication was Szidonia Lefkovits.

© 2023 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

and treatment [3]. However, accurately detecting CVDs through auscultation alone is challenging, especially due to subjective interpretations among physicians, particularly in low-resource countries [4]. It is estimated, a skilled cardiology fellow may diagnose CVDs with an accuracy of solely 56.2% [5]. Artificial intelligence (AI)-powered screening systems for disease detection have gained significant traction [6], [7], [8]. In line with this, machine learning and deep learning-based approaches for phonocardiogram (PCG) classification, in both the one-dimensional time-domain PCG signal [9], [10], [11] and in the two-dimensional time-frequency representations [12], [13], have emerged to support physicians in their decision-making process. However, developing a robust AI solution capable of accurately characterizing murmurs and anomalies in patients with different CVDs requires a large, well-annotated dataset.

The heart valves open and close in response to pressure changes across the valves. While the opening of the valves is usually silent,the closure of the valves produces vibrations that are perceived as heart sounds. When the mitral valve and tricuspid valve are closed, S1 (the first heart sound) is produced, and S2 (the second heart sound) is produced when the aortic and pulmonary valves are closed. The first heart sound S1 and the second heart sound S2 are normal heart sounds and indicate a healthy heart, and any other sounds indicate an unhealthy heart. [14] (illustrated in Figure 1). There are six notable datasets available, including the CirCor DigiScope Dataset [2], Github open-access Dataset [15], PhysioNet/CinC Challenge 2016 Dataset [16], Heart Sounds Shenzhen (HSS) Dataset [17],Michigan Heart Sound and Murmur Database [18] and the PASCAL Heart Sound Challenge Dataset [19]. However, these annotated PCG datasets have limitations. Some provide limited information on general heart sound evaluation or specific elements like murmurs, extra sounds, and severity levels (Normal, Mild, and Moderate/Severe) [17], [18], [19]. The highly clean nature of the Github dataset does not reflect real-world scenarios. Additionally, these datasets have a limited number of samples, which is insufficient for training robust deep-learning models. Most datasets, excluding the PASCAL Heart Sound Challenge Dataset, do not address the identification of low-quality heart sounds affected by environmental or physiological noise. In addition of that, none of the datasets provide samples of hard and rare PCG diagnosis which can be considered as a major drawback.

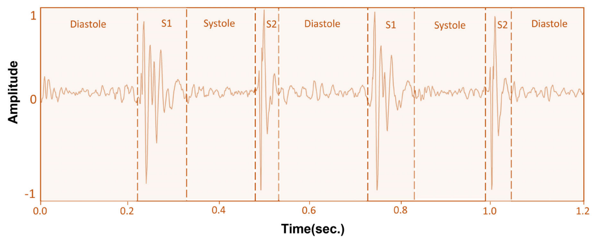


FIGURE 1. The figure depicts a standard representation of a phonocardiogram (PCG) signal, illustrating the four states of the cardiac cycle: systole, (S1), diastole, and (S2).

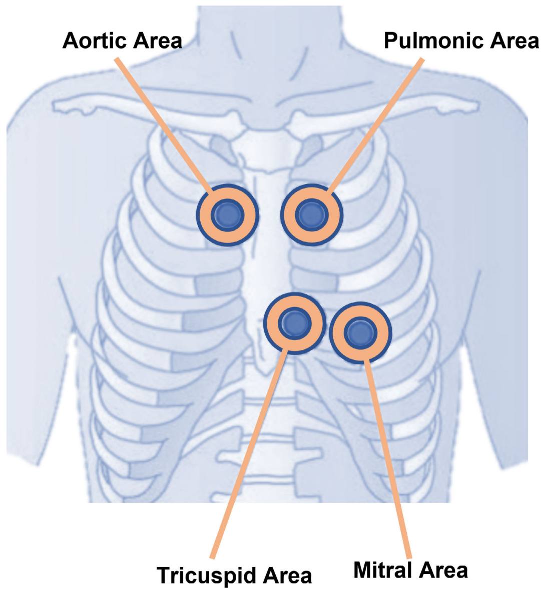


FIGURE 2. Heart auscultation collection positions.

To address these challenges,we systematic framework and system to produce a novel clinically collected dataset. The main contributions of this dataset are as follows:

- **Comprehensive Heart Sound Dataset:** The heart sound dataset presented in this study is one of the largest and most diverse available to date. It comprises 1353 records. Unlike other datasets, this dataset includes a comprehensive range of approximately 9 distinct classes of heart sounds, some of which are not found in any other existing dataset. The classes encompass various cardiac conditions such as Normal, Aortic stenosis, Aortic regurgitation, Pulmonic stenosis, Pulmonary regurgitation, Tricuspid stenosis, Tricuspid regurgitation, Mitral stenosis, Mitral regurgitation. This extensive coverage of heart sound classes enhances the dataset's potential for comprehensive analysis and diagnostic applications. It's important to note that the dataset includes normal heart sounds S1 and S2, and murmurs, which are a type of abnormal heart sound, making them part of the heart sounds. If we specify the specialization in cardiac murmurs, it may exclude the normal heart sounds (S1, S2). Therefore, we believe that defining it as 'heart sounds ' will encompass all classes, both normal and abnormal heart sounds.
- **Location Information:** The dataset incorporates essential location information where the sample is captured. This supplementary metadata not only enriches the dataset but also provides researchers with valuable contextual details, allowing them to perform source location-based investigations.
- **Inclusion of All Murmur Grades:** A significant additional contribution of our dataset is the inclusion of

murmur grades ranging from 1 to 6. This feature sets our dataset apart, as it reflects the comprehensive nature of murmurs found in real-world scenarios. By referring to echocardiography, we were able to assign murmur grades to the recorded heart sounds accurately. This comprehensive coverage of murmur grades further enhances the dataset's utility for training and evaluating machine-learning models in heart sound analysis and diagnosis.

The rest of this paper is organized as follows: In Section II, we provide background on most common heart sounds,, cardiovascular diseases, and the available public heart sound datasets in detail. Section III discusses data collection methods such as subject recruitment, participant demographics, instrumentation, and label annotation methods. The proposed heartwave dataset, as described in Section IV, and in Section V, dataset analysis and evaluation, including signal quality assessment and distribution visualization methods, are presented. Section VI draws on the discussion of this paper. Finally, Sections VII and VIII,VIII present future work, the main conclusion, and acknowledgements, respectively.

II. BACKGROUND

A. HEART SOUNDS

The normal functioning of the heart involves the opening and closing of cardiac valves in each cardiac cycle, which generates vibrations that produce the main heart sounds. To ensure optimal auscultation, the stethoscope should be placed at specific positions(illustrated in Figure 2) on the patient's chest [20].:

1) Aortic valve area: Second intercostal space, right sternal boundary.
2) Pulmonary valve area: Second intercostal space, left sternal boundary.
3) Tricuspid valve area: Left lower sternal border.
4) Mitral valve area: Fifth intercostal space, midclavicular line (apex area)

Each heart sound has several characteristics that are evaluated during auscultation, including:

1) Origin: The anatomical location on the patient's chest wall where the sound is most easily heard. Anatomical landmarks such as the midclavicular lines are used to determine the precise position.
2) Intensity: The loudness of the sound, which is subjectively assessed during auscultation. Electronic recording of a phonocardiogram (PCG) can provide a more objective measure of sound intensity by analyzing the amplitude of the sound's vibrations.
3) Duration: The length of time the sound is heard, which can be short or long. The duration of a sound influences its perception, such as whether it is perceived as a click, pop, or murmur.
4) Pitch: The frequency of the sound's oscillations, determining its high or low pitch. High-frequency sounds are best heard using the stethoscope's diaphragm, while low-frequency sounds are better perceived with the stethoscope's bell.
5) Quality: The balance of frequencies in the sound, defining its characteristics as sharp, dull, booming, cracking, blowing, loud, or melodious.
6) Timing: The correlation between the timing of the sound and a specific phase of the cardiac cycle, either systole or diastole.

The first heart sound (S1), commonly referred to as "lub," is produced by the closure of the mitral (M1) and tricuspid (T1) valves. The closure of the tricuspid valve occurs immediately after the closure of the mitral valve, resulting in these two sounds being perceived as one. The wider diameter of the atrioventricular (AV) valves generates a low-pitched sound that is best heard using the stethoscope's bell. The anatomical complexity of the AV valve mechanism contributes to the longer duration of the first heart sound compared to the second sound [21], [22].

The second heart sound (S2), often referred to as "dub," occurs at the beginning of diastole when the semilunar valves close. S2 has a higher pitch, shorter duration, and is best heard using the stethoscope's diaphragm due to the larger pressure gradients that cause valve closure. The second sound is generated by the closure of the aortic (A2) and pulmonic (P2) valves. A2 and P2 may be heard separately during inspiration. S2 is split due to the delayed closure of the pulmonic valve and the earlier closure of the aortic valve. During inspiration, the reduced intrathoracic pressure allows the distensible pulmonary circulation to receive a larger stroke volume from the right ventricle, leading to an extended right ventricular systole and a delayed P2 sound. Increased pulmonary vascular capacitance during inspiration also decreases venous return to the left atrium, resulting in a shorter left ventricular systole and an earlier A2 sound. In cases where the aortic valve closure is delayed, a phenomenon known as paradoxical splitting of S2 occurs, where P2 is heard before A2 [21], [23].

In some cases, a third heart sound (S3) may be detected in normal individuals. It is a low-pitched sound that occurs early in diastole as shown in figure 3 and is best heard near the apex of the heart. S3 is produced by vibrations caused by rapid ventricular filling. In children and adults after exercise, the presence of S3 can be considered normal [21].

A fourth heart sound (S4) may be produced by atrial systole. It is a low-pitched sound that occurs during late diastolic or presystolic phases of the heart as shown in figure 3. S4 is commonly heard in young children but is often associated with cardiovascular disease in adults. In individuals with atrial fibrillation, where atrial kick is absent, S4 may not be present. In cases of tachycardia, S4 and S3 sounds may merge, resulting in a summation gallop.

B. MOST COMMON CARDIOVASCULAR DISEASES

In order to build a well representative dataset, we have conducted an analysis with a number of physicians and

TABLE 1. Cardiac conditions.

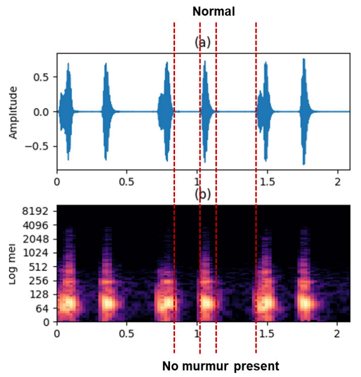
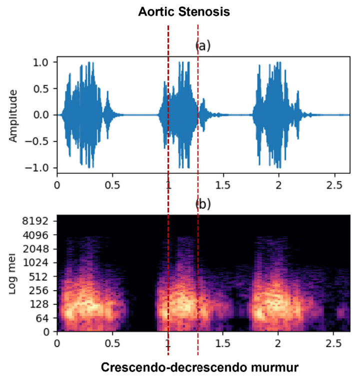
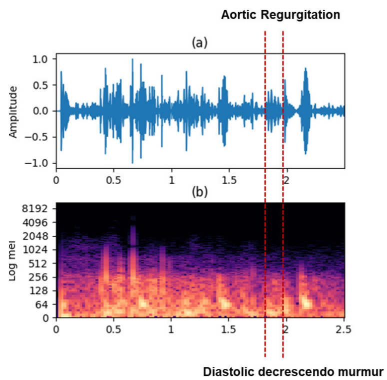
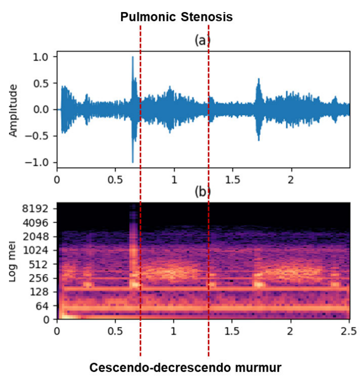
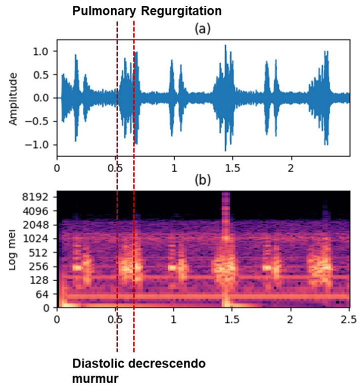
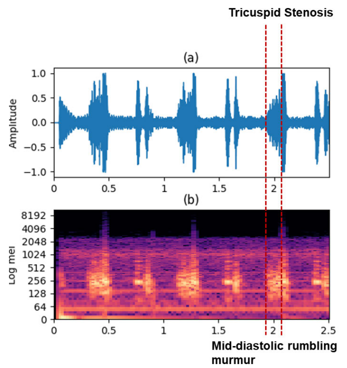
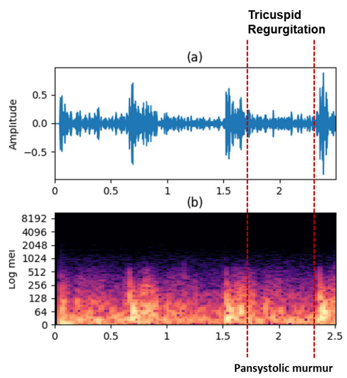
Temporal and Spectral heart sound representation of cardiac disease class	Analysis
	The normal heart sound primarily consists of two distinct components: the first heart sound (S1) and the second heart sound (S2). Temporally, S1 possesses a longer duration compared to S2 and serves as the auditory signal for the onset of the systolic phase. Upon spectrographic examination, unique frequency characteristics between S1 and S2 emerge. Specifically, S1 displays a frequency spectrum situated within the lower range of approximately 20 to 50 Hz. Contrarily, S2 is characterized by a more acute and elevated frequency profile, generally ranging between 50 and 100 Hz. This heightened frequency is a consequence of the closure of semilunar valves at the inception of the diastolic phase.
	In Aortic Stenosis (AS), a disruption is caused to the normal flow of blood from the left ventricle to the aorta during the heart's contraction phase, known as systole. Distinct characteristics in both time and frequency domains are exhibited by these heart sounds. In the temporal domain, systolic ejection murmur begins shortly after the first heart sound (S1) and extends to the second heart sound (S2). This murmur is typically crescendo-decrescendo in nature, meaning it starts softly, becomes louder, and then softens again towards the end of the systole. When examined with a spectrogram, distinct frequency patterns due to altered hemodynamics are revealed. A broader frequency range is displayed by the harsh, crescendo-decrescendo murmur, often extending from 300 to 500 Hz or even higher.
	In patients diagnosed with Aortic Regurgitation (AR), a pathological condition characterized by the retrograde flow of blood from the aorta back into the left ventricle during diastole, the heart sounds demonstrate unique attributes in both the temporal and frequency domains. In the temporal domain, an early diastolic decrescendo murmur can be detected shortly after the second heart sound (S2) and diminishes as the diastolic phase progresses. Corresponding spectrographic analysis, distinct frequency traits emerge due to the altered hemodynamics involved in AR. The decrescendo murmur typically exhibits a narrower frequency spectrum compared to normal heart sounds, often confined within a range of 200 to 300 Hz
	In Pulmonic Stenosis (PS), a pathological narrowing occurs at the pulmonic valve, leads to altered heart sounds. Typically, a systolic ejection murmur is heard, starting after the first heart sound (S1) and ending near the second heart sound (S2). This murmur follows a crescendo-decrescendo pattern, with a variable frequency spectrum, often ranging from 100 Hz to possibly over 400 Hz, depending on the severity of the stenosis. It should be noted that this specific range can vary based on the severity and characteristics of the stenosis.

TABLE 2. Cardiac conditions(continued).

Temporal and Spectral heart sound representation of cardiac disease class	Analysis
	Pulmonary Regurgitation (PR), a dysfunctional pulmonic valve allows for the retrograde flow of blood from the pulmonary artery back into the right ventricle during the heart's relaxation phase. Regarding temporal characteristics, a diastolic decrescendo murmur is detectable in cases of Pulmonary Regurgitation. This murmur emerges immediately following the second heart sound (S2) and gradually attenuates as diastole progresses. This diastolic decrescendo murmur is confined within a narrow frequency spectrum, often situated between 200 and 400 Hz. That is observed well in spectrographic analysis.
	Patients with tricuspid Stenosis (TS) experience a narrowing or obstruction in the tricuspid valve and altered hemodynamics, resulting in distinctive heart sound characteristics manifesting in both the temporal and frequency domains. A mid-diastolic monotonic rumbling murmur is evident in the temporal representation. This murmur follows the opening snap, a brief high-frequency sound, and continues until just before the first heart sound (S1). Corresponding spectral analysis reveals that this murmur shows a lower frequency spectrum, falling within approximately 50 to 150 Hz range. This lower frequency band corresponds to the reduced flow velocity and turbulence generated by the stenotic tricuspid valve, which obstructs the regular inflow of blood from the right atrium to the right ventricle during diastole.
	Tricuspid Regurgitation (TR) causes blood to flow backward from the right ventricle into the right atrium during the heart's contraction. Unlike many other heart conditions, TR's murmur remains constant throughout systole, starting with the first heart sound (S1). When analyzed spectrographically, this pansystolic murmur falls within a frequency range typically ranging from approximately 100 Hz and extending beyond 300 Hz.

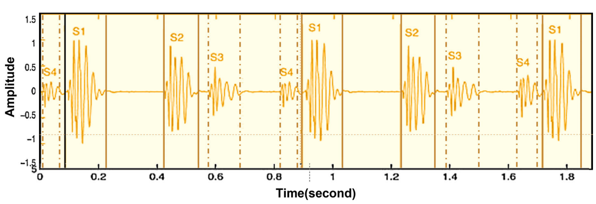
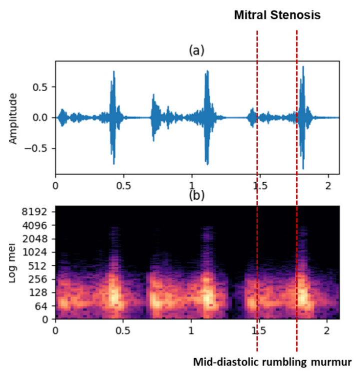
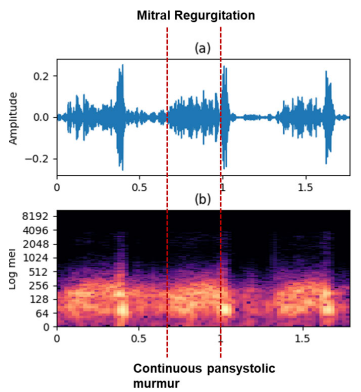


FIGURE 3. **The figure depicts a standard representation of a phonocardiogram (PCG) signal, illustrating the heart sounds: (S1), (S2), (S3), and (S4).**

consultants in several hospitals in Saudi Arabia and Egypt. We have arranged for meetings and questionsinars aimed at understanding how the healthcare provider precieve some condtions are more common, difficult or rare than others, we ended up with agreeing with them that there are at least nine PCG diagnosis (including the normal condition) that we could start with Table Table 1,2,3 displays a detailed analysis of the heartbeat sound of each class. We have provided a sample representation of each class using both time and time-frequency representation in order to get insights how it might be sometimes challenging to realize the features of each representation. It is observed that some cases are more difficult than others and some consideration need to be taken for successful classification. table 4 shows the

TABLE 3. Cardiac conditions(continued).

Temporal and Spectral heart sound representation of cardiac disease class	Analysis
	Patients with tricuspid Stenosis (TS) experience a narrowing or obstruction in the tricuspid valve and altered hemodynamics, resulting in distinctive heart sound characteristics manifesting in both the temporal and frequency domains. A mid-diastolic monotonic rumbling murmur is evident in the temporal representation. This murmur follows the opening snap, a brief high-frequency sound, and continues until just before the first heart sound (S1). Corresponding spectral analysis reveals that this murmur has a lower frequency spectrum, falling within the 50 to 150 Hz range. This lower frequency band corresponds to the reduced flow velocity and turbulence generated by the stenotic tricuspid valve, which obstructs the regular inflow of blood from the right atrium to the right ventricle during diastole.
	In Mitral Regurgitation (MR), the mitral valve's disfunction causes blood backflow into the left atrium during systole.Temporally, a continuous pansystolic murmur is observed, starting almost simultaneously with the first heart sound (S1) and ending when the second heart sound (S2) occurs. Unlike murmurs associated with other valve issues, this one maintains a consistent intensity throughout systole.Spectrographic analysis also reveals specific frequency traits, reflecting the altered hemodynamics in MR. It is observed that the energy spans a broader frequency range compared to normal heart sounds, typically ranging from about 60 Hz up to 600 Hz or even higher.

summary of the comparison of each class along with sample availability in terms of how common the diagnosis is and diagnosis difficulty according to the specialists opinions and pre-processing of data samples.

C. AVAILABLE HEART SOUND DATASETS

Currently, there exist several sound datasets, and six of them are widely utilized for algorithm development and evaluation. In this section, we briefly discussed the descriptions of these datasets. Table 5 show a summary of the currently available open-access datasets.

1) CIRCOR DIGISCOPE DATASET

The CirCor DigiScope dataset stands as the largest pediatric heart sound collection to date. It comprises a total of 5282 recordings obtained from the four primary auscultation locations of 1568 patients. These recordings provide a wealth of data, totaling over 312 hours of heart sound signals. The patients' ages range from 0.1 to 356.1 months, with an average of 73.4 months and a standard deviation of 50.3 months. The duration of the recordings varies between 4.8 and 80.4 seconds, with an average duration of 22.9 seconds and a standard deviation of 7.4 seconds. Through a semi-supervised annotation scheme, experts performed detailed annotations, covering the timing, shape, pitch, grading, quality, and location of each murmur. This meticulous approach yielded a total of 215780 manually annotated heart sounds [2].

2) HEART SOUNDS SHENZHEN (HSS) DATASET

The Heart Sounds Shenzhen (HSS) dataset is a comprehensive collection of Phonocardiogram (PCG) signals, consisting of 845 recordings obtained from 170 individuals. These individuals represent a diverse range of patients with various heart conditions, including coronary heart disease, fibrillation, valvular heart disease, and congenital heart disease. The dataset provides valuable insights into the acoustic characteristics of different heart diseases.The PCG recordings in the HSS dataset are sampled at a rate of 4 kHz, ensuring high-fidelity representation of the heart sounds. Each recording, which is approximately 30s, is labeled with one of three class labels: Normal, Mild, and Moderate/Severe, which indicate the severity of heart disease. It is important to note that the dataset does not explicitly specify which valves or diseases are classified as severe in the Moderate/Severe class.The audio recordings in the dataset were captured using an Electronic Stethoscope, specifically the Eko CORE model from the United States. The recordings are stored in the widely-used .wav format, ensuring compatibility and ease of use for researchers and practitioners [17].

TABLE 4. Characteristics of heart sounds in diagnostic classes.

Diagnostic Heart Sound	Best heard area	Intensity (Loudness)	Duration (Length)	Pitch (Frequency)	Quality (Characteristics)	Timing (Cardiac Cycle Phase)	Murmur configuration	Availability - Diagnosis
Normal	Mitral Area or Aortic Area	Normal 'lub-dub' sound(S1 is louder than S2)	S1 is longer than S2	S1 is lower in pitch than S2.	Rhythmic	No murmur	No murmur configuration	Common - Easy
Aortic Stenosis	Aortic Area	Loud	Short	High	Harsh	Systole	Decrescendo or mid-systolic crescendo-decrescendo	Common -Easy
Aortic Regurgitation	Aortic Area	Variable	Variable	High	Blowing	Diastole	Mid-diastolic crescendo-decrescendo or decrescendo	Common - Difficult
Pulmonic Stenosis	Pulmonic Area	Loud	Short	High	Harsh	Systole	Crescendo-decrescendo	Rare - Difficult
Pulmonary Regurgitation	Pulmonic Area	Variable	Variable	High	Blowing	Diastole	May have an early diastolic decrescendo murmur in severe cases	Common - Difficult
Tricuspid Stenosis	Tricuspid Area	Variable	Variable	High	Blowing	Diastole	Plateau-shaped or decrescendo-crescendo	Rare - Difficult
Tricuspid Regurgitation	Tricuspid Area	Variable	Variable	High	Blowing	Systole	May be associated with a systolic thrill in severe cases	Rare - Difficult
Mitral Stenosis	Mitral Area	Variable	Variable	Low	Rumbling	Diastole	Plateau-shaped or decrescendo-crescendo	Common - Easy
Mitral Regurgitation	Mitral Area	Variable	Variable	High	Blowing	Systole	Crescendo-decrescendo (mild, moderate) and holo-systolic (severe)	Common - Easy

3) GITHUB OPEN-ACCESS DATASET

The Github open-access dataset comprises a collection of 1000 PCG recordings in the .wav format. These recordings have been classified into five distinct categories: Normal (N), Aortic stenosis (AS), Mitral regurgitation (MR), Mitral stenosis (MS), and Mitral valve prolapse (MVP). The data was gathered from diverse sources such as books and websites. All recordings were sampled at a rate of 8 kHz. Each category contains 200 recordings, which vary in duration. On average, the recordings are approximately 3 seconds long, with the shortest signal length observed in the dataset being 1.125 seconds [15].

4) PHYSIONET/CINC CHALLENGE 2016 DATASET

The PhysioNet/CinC 2016 Challenge's heart sound dataset is a comprehensive compilation of nine distinct databases. This comprehensive collection consisted of a total of 2435 heart sound records obtained from 1297 patients. The duration of the phonocardiogram (PCG) signals in the dataset varies between 8 and 312.5 seconds. In order to maintain consistency across different devices with varying sampling rates, all PCG signals have been downsampled to a uniform rate of 2000 Hz. The heart sounds were captured from four auscultation positions: aortic, pulmonary, tricuspid, and mitral. The subjects included in the dataset encompass a wide range of conditions, including heart valve diseases and coronary artery diseases, and consist of both healthy (normal) and diseased (abnormal) individuals.However, it is important to note that the dataset exhibits a significant imbalance, with a greater number of normal records compared to abnormal ones. The recordings were made in various settings, including both clinical and non-clinical environments, and were categorized as normal, abnormal, or uncertain [16].

5) PASCAL HEART SOUND CHALLENGE DATASET

The PASCAL Heart Sound Challenge dataset is divided into two subsets: Dataset A and Dataset B. The dataset consists of sound clips collected from four different regions of the body: aortic (AR), pulmonary (PR), tricuspid (TR), and mitral (MR).Dataset A was created through crowd-sourcing using the iStethoscope Pro iPhone app. It contains 176 sound files in .wav format, categorized into five classes. The first category includes 31 files of normal heart sounds, serving as the training set for healthy conditions. The second category consists of 34 files classified as training murmurs, indicating potential heart disorders. The third category contains 19 files labeled as training extra heart sounds, which may be indicative of certain heart conditions. The fourth category, named Artifact, includes 40 files with various sounds unrelated to heart conditions. Lastly, the fifth category comprises 52 unlabeled files used for testing.Dataset B was compiled during a clinical trial conducted in hospitals using the DigiScope digital stethoscope. It comprises 656 sound files in .wav format, divided into four classes. The first category contains 320 files representing normal heart sounds for training. The second category consists of 95 files classified as training murmurs, indicating abnormal heart conditions. The third category includes 46 files representing extrasystole sounds, which signify irregular heart rhythms. The fourth category comprises 195 unlabeled files for testing [19].

III. COLLECTION METHODOLOGY

In this section, we will provide a detailed explanation of the collection methodology, beginning with the subject recruitment method, participant demographics, and the tools used.

A. SUBJECT RECRUITMENT

The dataset was gathered from three prominent hospitals that offer specialized cardiovascular healthcare services: the National Heart Institute in Cairo, Egypt; King Abdulaziz Specialist Hospital-Taif, KSA; and King Faisal Medical Complex-Taif, KSA. Data collection took place from September 9, 2022, to January 30, 2023. Consequently, the sample population represents the adult population with cardiovascular and valvular diseases effectively. The Institutional Review Board of the Department of Research and Studies in Health Affairs in Taif City, King Abdulaziz City for Science and Technology (KACST), KSA (Registration number: HAP-02-T-067 approval and approval number: 719), granted ethical approval for the study. All participants agreed to participate in this data collection study with informed consent.

B. PARTICIPANT DEMOGRAPHIC

Participants in this study come from a wide range of backgrounds. Individuals from diverse adults age groups, ethnic backgrounds, and geographical areas are included in the dataset. The sample population includes both males and females, assuring gender representation. Because the study's focus is on cardiovascular and valvular medical conditions the participants are mostly individuals who have been diagnosed with such conditions. The recruitment strategy guaranteed that the sample population accurately reflected the adult population suffering from these health problems.

FIGURE 4. Developed In-house stethoscope.

C. INSTRUMENT

We had to build our instruments to collect the data professionally, as the commercial equipment is expensive and does not help us collect efficiently and quickly, as the time of the clinical centers is critical. In this subsection, we explore the composed digital stethoscope and the developed data collection app.

1) IN HOUSE DIGITAL STETHOSCOPE

Stethoscopes are essential tools used by cardiologists and physicians to assess heart conditions by listening to physiological sounds. While newer electronic stethoscopes have the capability to collect various physiological sounds digitally, integrating them with mobile applications for recording heart sounds can be expensive due to the need for a license to access the electronic stethoscope's Application Programming Interface (API), which can be quite costly.

To address this issue, we have developed in-house digital stethoscope(are shown in Figure 4. The development process involves the following steps:

1) We utilized the 3M Littman Classic III Stethoscope as the base. The rubber tube of the stethoscope was cut from the diaphragm side.
2) An Electret Microphone Condenser was embedded inside the rubber tube. Specifically, we employed the uxcell Electret Microphone Condenser Pickup with a size of 6mm x 3.5mm.
3) The microphone was connected to the suitable amplifier, and its cable was then linked to the iRig HD-2 audio interface.
4) Finally, the audio interface was connected to an Apple iPad, enabling the digital stethoscope to interface with the device.

TABLE 5. A summary of the currently available open access datasets.

Dataset	Class	Recordings numbers	Total	Recordings lengths (s)	Sampling frequency	Acquisition Device	Limitation
CirCor DigiScope Dataset [2]	Timing, shape, pitch, grading, quality, and location of each murmur	-	5282	5-168 s	4 KHz	Littmann 3200 stethoscope	Lacks adult population representation
PhysioNet Dataset [24]	Normal Abnormal	2575 655	3240	5-120 s	2 KHz	Digital stethoscope	Lacks others heart diseases
Pascal - A Dataset [19]	Normal Murmur Extrasystole Artifact	45 48 27 56	167	1-30 s	44.1 KHz	iStethoscope Pro iPhone app	Small numbers of recordings Lacks others heart diseases
Pascal - B Dataset [19]	Normal Murmur Extrasystole	167 69 39	279	1-30 s	44 KHz	Digital stethoscope	Lacks others heart diseases
Github open-access Dataset [25]	Normal Aortic stenosis (AS) Mitral valve prolapse (MVP) Mitral stenosis (MS) Mitral regurgitation (MR)	200 200 200 200 200	1000	Roughly 3 s	8 KHz	Collected from different sources	Lacks others heart diseases Small numbers of recordings
Heart Sounds Shenzhen [17]	Normal Mild Moderate/Severe	- - -	845	30 s on average	4 KHz	Electronic Stethoscope	Lacks others heart diseases

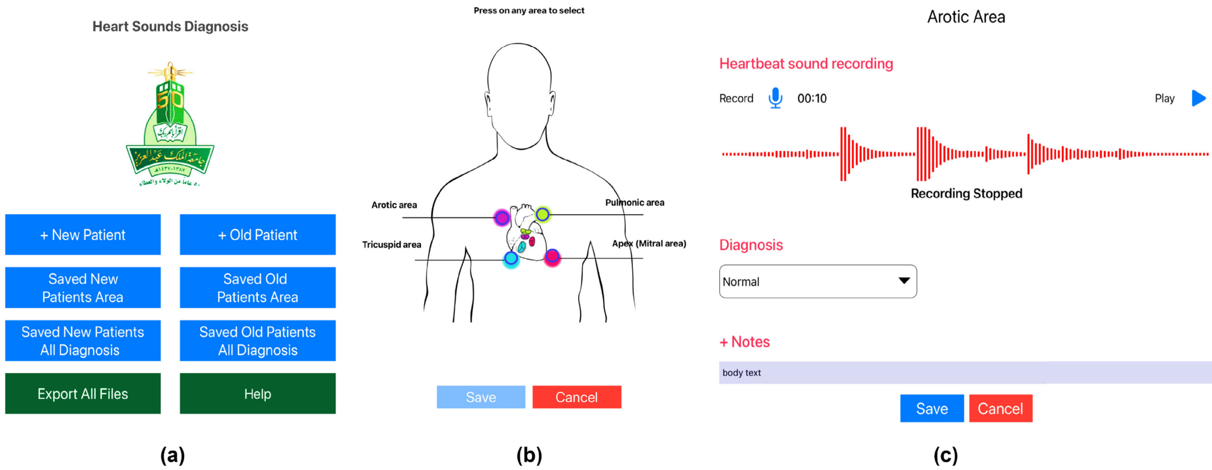


FIGURE 5. Data Collection App (a) Main Page UI,(b) Location setection UI and (b) Data collection and annotation UI.

By following these steps, we have developed a digital stethoscope that offers a cost-effective alternative for recording heart sounds, eliminating the need for costly licenses associated with integrating electronic stethoscopes with mobile applications.

2) DATA COLLECTION APP

We have developed a mobile application specifically designed for collecting and labeling heart sounds from patients in clinics. The application features a user-friendly interface that enables physicians to seamlessly gather and label data. The main page of the application is illustrated in Figure 5(a).

To utilize the application for data collection and labeling, the following steps are followed:

1) Select the appropriate icon based on the patient's status: "old patient" if their case has been confirmed by an echocardiogram report or "new patient" if they are visiting the clinic for the first time. For this study, only patients with an echocardiogram report were included(are shown in Figure 5(a)).
2) Place the stethoscope on the relevant area of the patient's body and choose the corresponding area within the application. The physician will select the appropriate area based on the information provided in the echocardiogram report(are shown in Figure 5(b)).
3) Start recording the heart sounds by pressing the recording icon. The application will capture the audio during this time.
4) Once the recording is complete, choose the specific diagnosis from the dropdown list based on the findings in the echocardiogram report(are shown in Figure 5(c)).
5) Save the heart sound as a WAV format file along with its corresponding label by pressing the save icon. The file will be stored in separate folders based on the collection area and specific diagnosis(are shown in Figure 5(c)).

By following these steps, the application enables physicians to conveniently collect and label heart sounds from patients, ensuring organized storage of data for further analysis and research purposes.

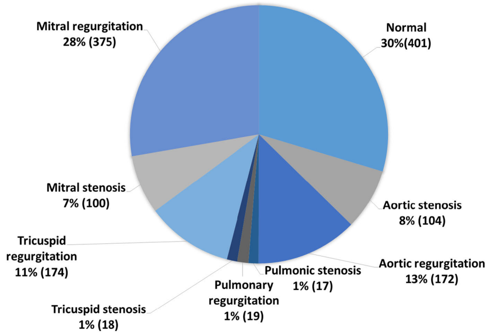


FIGURE 6. The percentage and samples number of each class in the proposed dataset.

TABLE 6. A comprehensive summary of the HeartWave dataset.

Diagnosis	Recordings	SNR	Duration (Sec.)
Normal	401	0.00318	18.32
Aortic regurgitation	172	0.00298	25.06
Aortic stenosis	104	0.00346	25.69
Pulmonic stenosis	17	0.00304	20.81
Pulmonary regurgitation	19	0.00403	19.89
Tricuspid stenosis	18	0.00318	19.95
Tricuspid regurgitation	147	0.00310	23.52
Mitral stenosis	100	0.00347	25.27
Mitral regurgitation	375	0.00274	24.61
Overall	1353	0.00324	22.57

D. LABEL ANNOTATIONS METHODOLOGY

To ensure accurate labels for effective prediction of cardiovascular disease classification problems, our study employed a robust annotation approach by utilizing echocardiogram reports of the patients. Each patient's report contained comprehensive details about their cardiac structure, function, and any identified abnormalities. Our team of cardiologists diligently examined each patient's echocardiogram report. They managed and supervised the whole labeling and annotation process. They carefully reviewed the descriptions of pathological conditions mentioned in the reports and listened to the corresponding heart sounds for each patient. Based on their analysis and expertise, they assigned labels to the heart sounds, indicating the presence or absence of specific diseases or pathological conditions.This process ensured accurate and validated label annotations for the heart sound based on the findings in each patient's echocardiogram report. We made sure not to include patients with multiple valve diseases. In other words, if a patient had two or three diseases, we did not collect any samples from them. We aimed to capture only the sound of a single disease along with its corresponding label to maintain the quality of our dataset.

IV. DATASET DESCRIPTION

The HeartWave dataset, developed collaboratively by King Abdul-Aziz University and three hospitals, stands as one of the largest and most comprehensive collections of heart sounds. It comprises 1353 records. The dataset offers label annotations at record levels. Additionally, the annotations indicate the specific chest area from which each recording was obtained as shown in Figure 7. In terms of patient distribution, the dataset consists of 401 recordings from healthy individuals and 952 recordings from patients with various diseases. Among the diseased patients, mitral regurgitation is the most prominently represented condition, followed by aortic regurgitation. The mitral and aortic valves are affected by rheumatic fever, so the number of samples of the two diseases is the highest.It is important to note that certain classes, such as pulmonic stenosis and pulmonic regurgitation, have relatively fewer samples compared to other disease classes. Furthermore, an important feature of the HeartWave dataset is the inclusion of murmur grades ranging from 1 to 6. These grades accurately reflect the varying severity and characteristics of murmurs found in real-world scenarios. The assignment of murmur grades was done by referencing echocardiography. On average, the record duration is 21.57 seconds, and all sound records are stored in wave (.wav) format. An overview of the dataset has been presented in Table 6 and Figure 6

V. DATASET ANALYSIS AND EVALUATION

A. SIGNAL QUALITY ASSESSMENT METHODOLOGY

Ensuring the reliability and credibility of a dataset, referred to as a gold standard (GS) reference, is of utmost importance when assessing and comparing different classification algorithms [26]. Since the diagnosis process relies on the subjective judgment and clinical expertise of physicians, it is crucial to employ a robust methodology for evaluating the quality of the dataset [27], [28]. In this particular study, we utilized a quality assessment approach that centers around the concept of signal-to-noise ratio (SNR) in order to assign signal quality labels to each individual record. The signal-to-noise ratio (SNR) is defined as the ratio of the mean or expected value (μ) of the signal to the standard deviation (σ) of the noise [29], [30]. By incorporating this definition, we gain an additional perspective for evaluating the quality of a heart sound record. It takes into account the relationship between the mean and the variability of the noise, offering valuable insights into the overall signal quality. The mathematical representation of SNR is given by the equation:

$$\mathrm{SNR} = \frac{\mu}{\sigma} \tag{1}$$

where μ represents the mean or expected value of the heart sound signal, and σ represents the standard deviation of the noise present in the signal. This equation provides a measure of the SNR specific to heart sound signals, allowing for the evaluation of the signal quality by quantifying the ratio of the signal's strength to the level of noise interference.This definition provides a quantitative measure that aids in assessing the signal quality and facilitates accurate analysis and interpretation of this dataset.

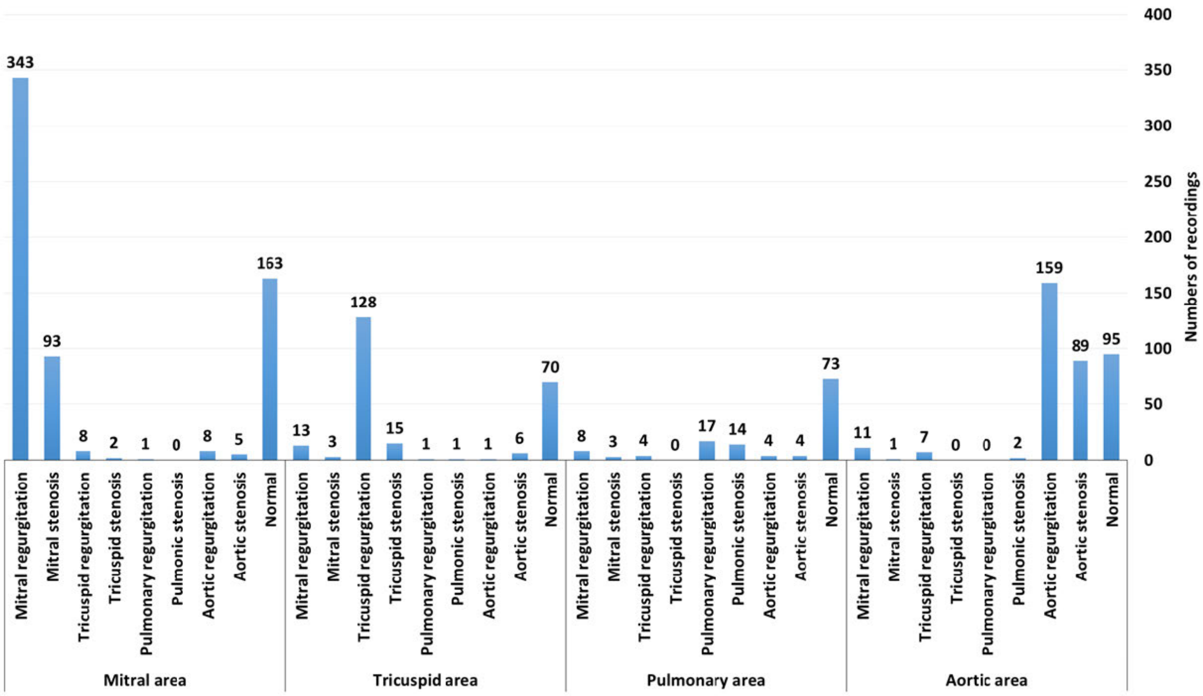


FIGURE 7. The distribution of samples from different disease classes with respect to the acquisition area.

B. DISTRIBUTION VISUALIZATION METHODOLOGY

Kernel density estimation (KDE) is a non-parametric statistical technique used to estimate the probability density function (PDF) of a random variable [31]. It is commonly used in data analysis and data visualization to understand the underlying distribution of a dataset. The KDE equation for estimating the probability density function (PDF) at a point x is given by:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{2}$$

where, $\hat{f}(x)$ represents the estimated PDF at point x,n is the number of data points in the dataset,x_i represents the i-th data point,h is the bandwidth parameter that determines the width of the kernel,$K(\cdot)$ is the kernel function, typically a symmetric and smooth function.

Commonly used kernel functions include the Gaussian (normal) distribution:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \tag{3}$$

where u is the argument of the kernel function.

By summing up the contributions of the kernel functions for each data point, scaled by the bandwidth h and the number of data points n, we obtain the KDE estimate $\hat{f}(x)$ of the PDF at point x.

In this study, we also used KDE to analysis and visualize the underlying SNR and duration based distribution our proposed dataset.

VI. DISCUSSION

Our study introduces a novel dataset featuring nine unique classes of heart sounds, including rare and challenging-to-diagnose conditions such as pulmonary stenosis, tricuspid stenosis, and tricuspid regurgitation. The dataset comprises 401 recordings from healthy individuals and 952 from patients with various heart-related diseases. Notably, mitral regurgitation is most prevalent, with 375 recordings, as shown in Figure 6.

The dataset offers in-depth insights into the tempo-spectral characteristics of the murmurs related to these specific conditions, detailed in Table 1,2,3. Figure 8 focuses on the duration of the Phonocardiogram (PCG) samples, revealing that the majority exceed 5 seconds in length. Aortic stenosis and mitral stenosis recordings have the longest average durations, clocking in at 25.69 seconds and 25.27 seconds, respectively. In contrast, Pulmonic stenosis recordings have a shorter average duration of 20.81 seconds. The data also shows variations in recording duration density, as depicted in Figure 9. Most classes exhibit a peak density of samples around 22 seconds, with aortic regurgitation as an outlier at 23 seconds. A secondary density peak appears around 30 seconds for most disease classes and at 11 seconds for the normal class.

These variations in duration have implications for the study of heart sounds. Each PCG recording contains multiple heartbeats and four distinct heart sound states (S1, Systole, S2, and Diastole), making them well-suited for heart sound segmentation methods like those proposed by Springer et al. [32]. This extensive dataset will aid in

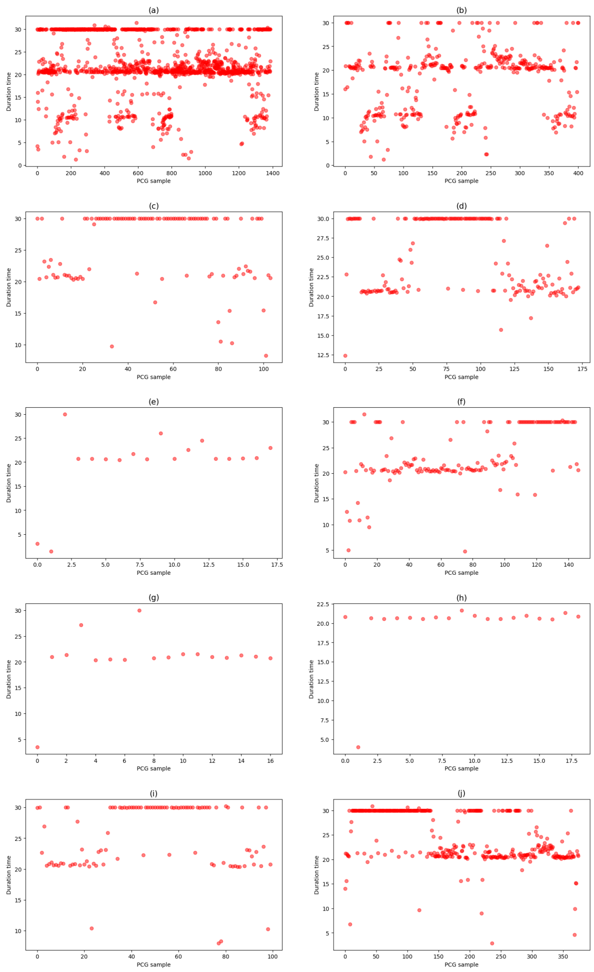


FIGURE 8. **Heart sound recordings distribution in function of duration (a) Overall, (b) Normal,(c) Aortic stenosis, (d) Aortic regurgitation, (e) Tricuspid stenosis, (f)Tricuspid regurgitation, (g) Pulmonic stenosis, (h) Pulmonary regurgitation, (i) Mitral stenosis, (j) Mitral regurgitation.**

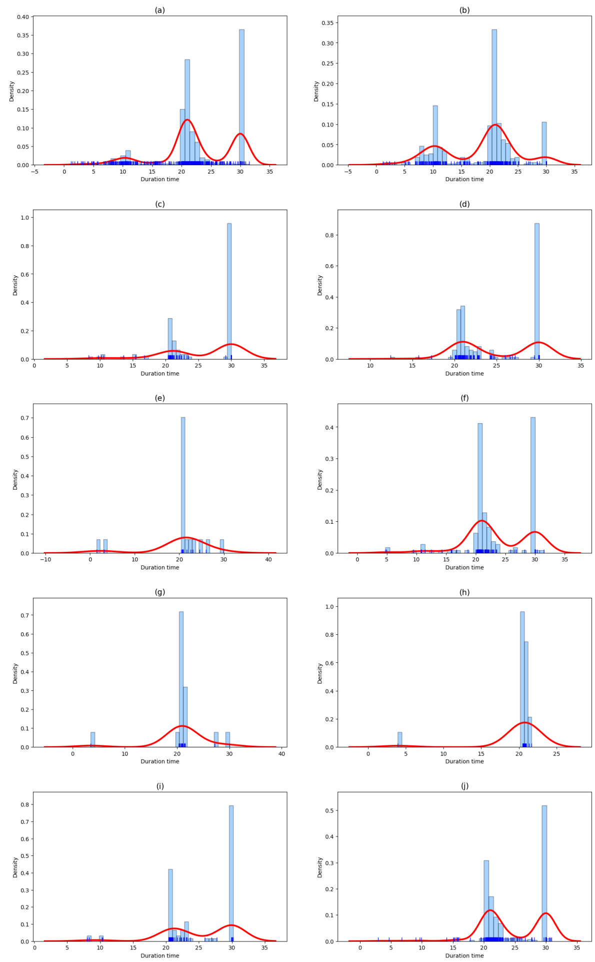


FIGURE 9. **The highest density of PCG samples for each class is observed at the following durations: (a) Overall:around 22s, (b) Normal: around 22s, (c) Aortic stenosis: around 30s, (d) Aortic regurgitation: around 23s, (e) Tricuspid stenosis: around 22s, (f) Tricuspid regurgitation: around 22s, (g) Pulmonic stenosis: around 22s, (h) Pulmonary regurgitation: around 22s, (i) Mitral stenosis: around 22s, (j) Mitral regurgitation: around 22s.**

optimizing signal duration for classification tasks, providing a valuable resource for future research.

In addition to temporal characteristics, our dataset given the importance of the area of auscultation in diagnosing cardiovascular diseases (CVDs). To this end, our dataset includes recordings collected from various clinically relevant chest locations, as depicted in Figure 2. The distribution of collected samples across these regions is showcased in Figure 7. This demonstrates a diverse representation from most classes in each location, making our dataset particularly valuable for crafting classification algorithms that are both robust and widely applicable. Notably, the mitral area has a heightened representation within our collection, while samples from the pulmonary area are somewhat sparse in comparison.

Such an imbalance in the dataset is, however, expected due to the inherent prevalence rates of certain diseases over others. For instance, mitral regurgitation samples outnumber those of other conditions. This can be attributed to the fact that rheumatic fever primarily targets the mitral valve. Subsequently, the aortic valve may be affected by the same fever, while the pulmonic and tricuspid valves generally remain unscathed. Thus, the dataset's structure mirrors the real-world incidence and prevalence of these specific conditions, offering an authentic resource for research.

A significant feature of our dataset is implementing a Signal-to-Noise-Ratio (SNR) based noise analysis technique to ensure quality control. This methodology not only facilitates the assessment of data quality but also enables the identification of noisy signals for further analysis. In order to validate the quality of phonocardiogram (PCG) samples in our dataset, we conducted a comprehensive SNR analysis, visualized in Figure11. This figure reveals that the Kernel Density Estimation (KDE) values of SNR are mostly concentrated above zero, suggesting that the majority of the PCG samples are of high quality with acceptable noise levels. This assertion is further supported by Figure 10, which displays each PCG sample's SNR in a scatterplot format for all classes.

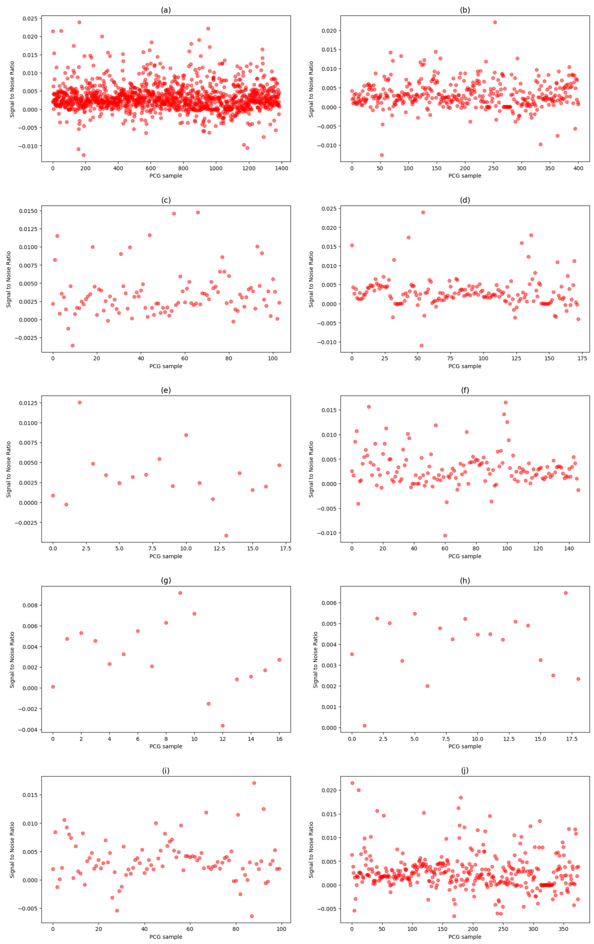


FIGURE 10. **Heart sound recordings distribution in function of the signal-to-noise ratio of (a) Overall, (b) Normal,(c) Aortic stenosis, (d) Aortic regurgitation, (e) Tricuspid stenosis, (f)Tricuspid regurgitation, (g) Pulmonic stenosis, (h) Pulmonary regurgitation, (i) Mitral stenosis, (j) Mitral regurgitation.**

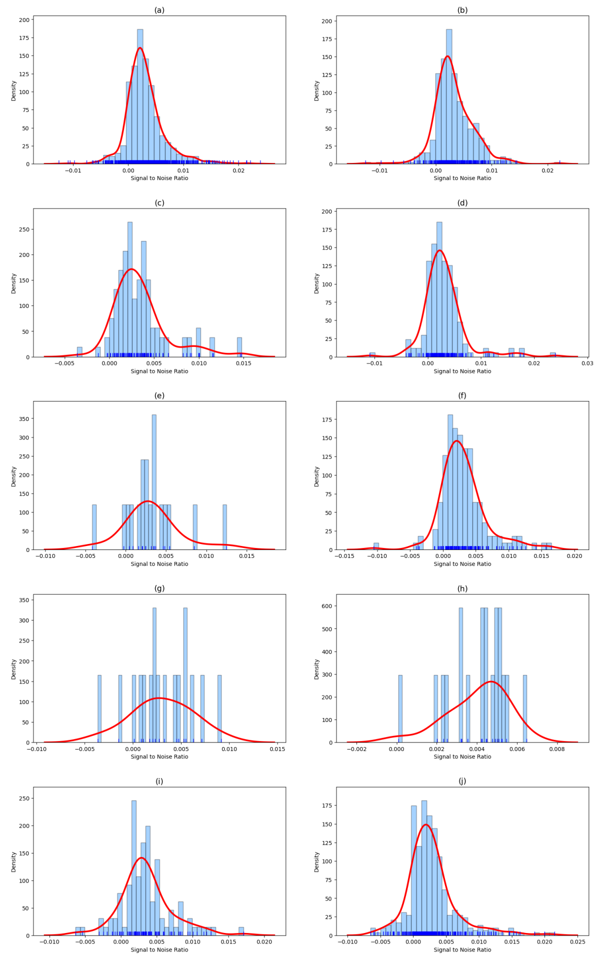


FIGURE 11. **Signal-to-noise ratio in function of density related to (a) Overall, (b) Normal,(c) Aortic stenosis, (d) Aortic regurgitation, (e) Tricuspid stenosis, (f) Tricuspid regurgitation, (g) Pulmonic stenosis, (h) Pulmonary regurgitation, (i) Mitral stenosis, (j) Mitral regurgitation.**

The trend demonstrates that the number of PCG samples with negative SNR values is minimal and statistically negligible. Moreover, the KDE curves in Figure 11 indicate that the SNR values for most classes follow a near-normal distribution, except for tricuspid regurgitation, which shows a slight skew to the right. This uniformity in SNR values across classes reaffirms the dataset's suitability for predictive modeling and classification tasks. This standardized approach for evaluating noise levels enhances the dataset's reliability, making it a valuable resource for researchers aiming to develop robust deep-learning models for detecting various cardiovascular diseases (CVDs).

To deepen our understanding of the dataset's quality, we conducted a comprehensive comparison of the average signal-to-noise ratio (SNR) between our proposed dataset and the well-known PhysioNet/CinC 2016 dataset. The individual SNR values for samples from both datasets are illustrated in Figure 12.

In our dataset, the majority of SNR values are above zero, contrasting with the PhysioNet/CinC 2016 dataset, where

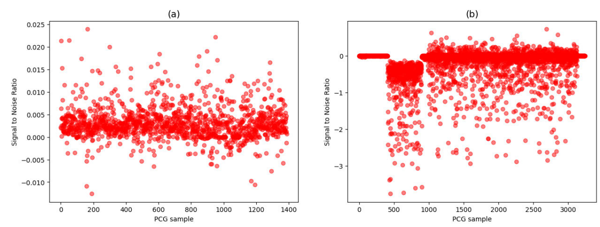


FIGURE 12. **Heart sound recordings distribution in function of the overall signal-to-noise ratio of (a) Proposed dataset, (b) PhysioNet/CinC 2016 dataset.**

the majority of values fall below zero. Despite this general trend, it's important to acknowledge that both datasets contain outliers, which are likely attributable to variations in different classes of heart sounds.

When we examine the averages, the SNR value for our proposed dataset is 0.0031, significantly higher than the -0.24 average of the PhysioNet/CinC 2016 dataset. Even when narrowing the scope to just the 'normal' class, the average SNR values for our dataset (0.0032) far exceed those of the PhysioNet/CinC 2016 dataset (-0.28).

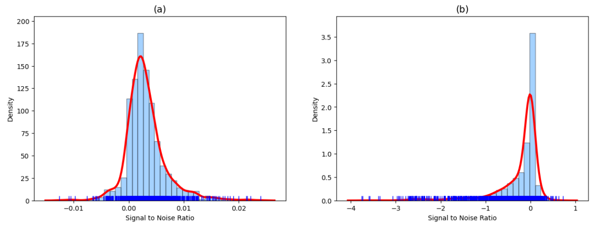


FIGURE 13. Signal-to-noise ratio in function of density related to (a) Overall of the proposed dataset, (b) Overall of the PhysioNet/CinC 2016 dataset.

For a more nuanced analysis, we also examined the abnormal classes in both datasets. The PhysioNet/CinC 2016 dataset groups all types of murmurs into one abnormal class. To create a more comparable metric, we calculated the average SNR across all pathological classes in our dataset. The result was an average SNR of 0.0033, notably higher than the average SNR of -0.063 in the PhysioNet/CinC 2016 dataset's abnormal class.

The KDE distributions of the SNR values of all samples together, depicted in Figure 13, further underscore these findings. Our dataset shows a bell-shaped distribution, suggesting a normal distribution, while the PhysioNet/CinC 2016 dataset reveals a positively skewed distribution, indicative of the presence of outliers which could affect further statistical analyses and modeling.

The superior average SNR in both normal and pathological classes of our dataset highlights its higher signal quality and increased ability to distinguish between various pathological conditions.

To ensure optimal accuracy and reliability in our heart sound research, we developed a custom-made software tool specifically tailored for data collection, sound editing, annotation, and quality assurance evaluation. This streamlined software significantly simplifies the management and evaluation of the dataset.

The involvement of expert cardiologists in the annotation process was crucial. Their expertise ensured the accuracy and consistency of the annotations, providing a strong foundation for our dataset's value. Our dataset stands out not merely because of its volume but also due to its comprehensive annotation and noise analysis approach. This dual focus offers future researchers a robust platform for crafting deep learning models adept at pinpointing a wide range of cardiopulmonary diseases.

The integration of expert annotations with noise analysis augments the dataset's usability and effectiveness, propelling further advancements in heart sound analysis. By offering a blend of expert knowledge and technical analysis, our dataset is poised to make significant contributions to the field of heart sound research.

VII. FUTURE WORKS

Despite our valuable contributions, it's important to acknowledge that the HeartWave dataset may encompass noise interference in real-world settings during collection may impact the precision of cardiovascular sound analysis.

In the future, we aim to explore the potential of incorporating machine learning and deep learning techniques into the analysis of our dataset and do benchmarking studies, which could significantly enhance the accuracy and efficiency of cardiovascular disease classification algorithms. Moreover, we will use heart-sound PCG signal segmentation techniques to augment the dataset and conduct further analysis and classification.

VIII. CONCLUSION

Our HeartWave dataset is an invaluable asset for researchers investigating cardiovascular sound analysis in adults.As this dataset has a substantial number of records diverse group of participants, it provides a comprehensive and extensive dataset for the development and evaluation of automated algorithms for the classification of cardiovascular diseases. Our dataset is comprehensive and encompasses a wide range of cardiovascular diseases (CVDs). It includes not only the most common CVDs but also those that are particularly challenging to diagnose. Additionally, our dataset covers rare and difficult to diagnosis conditions such as tricuspid stenosis,pulmonic stenos and tricuspid regurgitation. The meticulous use of a SNR-based quality assurance methodology in constructing the golden standard reference ensures the reliability of the database. The duration of the recordings in the proposed dataset allows for augmentation using various heart sound segmentation methods.. Our firm belief is that this dataset will promote the development of medical devices that are equipped with algorithms for denoising, profiling, or segmenting heart sounds, ultimately leading to better safety and early detection of heart diseases.

ACKNOWLEDGMENT

The authors would like to thank all the individuals who helped them to complete this study, especially the cardiologists, consultants, and physicians, for their time and effort. In particular, they would like to thank the following individuals for their major contributions to data collection and annotation: Ahmed Kamel Abdelghany Hassan, an Assistant Professor of cardiology at the National Heart Institute, Cairo, Egypt; Mahmoud Mazen, a Professor of cardiothoracic surgery at the National Heart Institute, Cairo; Magdy Abdulhameed Abdulhady, an MD specializing in cardiology at the Cardiology Department, King Abdulaziz Specialist Hospital, Taif, Saudi Arabia; Ibrahim Abdelmonaem Abdelhaleem, a Doctor at the National Heart Institute, Cairo, and the Faculty of Human Medicine, Zagazig University, Zagazig, Egypt; and Talal Khalid Alafif, a Doctor at the King Fahad Hospital Jeddah, Saudi Arabia.

REFERENCES

[1] R. Lozano, M. Naghavi, K. Foreman, S. Lim, K. Shibuya, and V. Aboyans, "Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the global burden of disease study 2010," *Lancet*, vol. 380, no. 9859, pp. 2095–2128, 2012.

[2] J. Oliveira, F. Renna, P. D. Costa, M. Nogueira, C. Oliveira, C. Ferreira, A. Jorge, S. Mattos, T. Hatem, T. Tavares, A. Elola, A. B. Rad, R. Sameni, G. D. Clifford, and M. T. Coimbra, "The CirCor DigiScope dataset: From murmur detection to murmur classification," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 6, pp. 2524–2535, Jun. 2022.

[3] A. Bourouhou, A. Jilbab, C. Nacir, and A. Hammouch, "Heart sounds classification for a medical diagnostic assistance," *Int. J. Online Biomed. Eng.*, vol. 15, no. 11, pp. 88–103, 2019.

[4] U. Alam, O. Asghar, S. Khan, S. Hayat, and R. Malik, "Cardiac auscultation: An essential clinical skill in decline," *Brit. J. Cardiol.*, vol. 17, no. 1, p. 8, 2010.

[5] S. Mangione, L. Z. Nieman, E. Gracely, and D. Kaye, "The teaching and practice of cardiac auscultation during internal medicine and cardiology training: A nationwide survey," *Ann. Internal Med.*, vol. 119, no. 1, pp. 47–54, 1993.

[6] N. Huda, S. Khan, R. Abid, S. B. Shuvo, M. M. Labib, and T. Hasan, "A low-cost, low-energy wearable ECG system with cloud-based arrhythmia detection," in *Proc. IEEE Region Symp. (TENSYMP)*, Jun. 2020, pp. 1840–1843.

[7] S. B. Shuvo, S. N. Ali, S. I. Swapnil, T. Hasan, and M. I. H. Bhuiyan, "A lightweight CNN model for detecting respiratory diseases from lung auscultation sounds using EMD-CWT-based hybrid scalogram," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 7, pp. 2595–2603, Jul. 2021.

[8] S. S. Alam, S. B. Shuvo, S. N. Ali, F. Ahmed, A. Chakma, and Y. M. Jang, "Benchmarking deep learning frameworks for automated diagnosis of ocular toxoplasmosis: A comprehensive approach to classification and segmentation," 2023, *arXiv:2305.10975*.

[9] M. B. Er, "Heart sounds classification using convolutional neural network with 1D-local binary pattern and 1D-local ternary pattern features," *Appl. Acoust.*, vol. 180, Sep. 2021, Art. no. 108152.

[10] P. T. Krishnan, P. Balasubramanian, and S. Umapathy, "Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network," *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 505–515, 2020.

[11] S. B. Shuvo, S. N. Ali, S. I. Swapnil, M. S. Al-Rakhami, and A. Gumaei, "CardioXNet: A novel lightweight deep learning framework for cardiovascular disease classification using heart sound recordings," *IEEE Access*, vol. 9, pp. 36955–36967, 2021.

[12] M. Deng, T. Meng, J. Cao, S. Wang, J. Zhang, and H. Fan, "Heart sound classification based on improved MFCC features and convolutional recurrent neural networks," *Neural Netw.*, vol. 130, pp. 22–32, Oct. 2020.

[13] S. B. Shuvo, S. S. Alam, S. U. Ayman, A. Chakma, P. D. Barua, and U. R. Acharya, "NRC-Net: Automated noise robust cardio net for detecting valvular cardiac diseases using optimum transformation method with heart sound signals," 2023, *arXiv:2305.00141*.

[14] L. P. Tilley, *Manual of Canine and Feline Cardiology*. Amsterdam, The Netherlands: Elsevier, 2008.

[15] G.-Y. Son and S. Kwon, "Classification of heart sound signal using multiple features," *Appl. Sci.*, vol. 8, no. 12, p. 2344, Nov. 2018.

[16] 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. Comput. Cardiology Conf. (CinC)*, Sep. 2016, pp. 609–612.

[17] F. Dong, K. Qian, Z. Ren, A. Baird, X. Li, Z. Dai, B. Dong, F. Metze, Y. Yamamoto, and B. W. Schuller, "Machine listening for heart status monitoring: Introducing and benchmarking HSS—The heart sounds Shenzhen corpus," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 7, pp. 2082–2092, Jul. 2020.

[18] University of Michigan Health System. (2015). *Michigan Heart Sound and Murmur Database (MHSDB)*. Accessed: Apr. 2022. [Online]. Available: http://www.med.umich.edu/lrc/psb/heartsounds/index.htm

[19] P. Bentley, G. Nordehn, M. Coimbra, and S. Mannor. (2011). *The PASCAL Classifying Heart Sounds Challenge 2011*. [Online]. Available: http://www.peterjbentley.com/heartchallenge/index.html

[20] S. E. Koetting, "The physiological origins of heart sounds and murmurs: The unique interactive guide to cardiac diagnosis," *J. Investigative Surg.*, vol. 10, no. 6, pp. 397–399, Jan. 1997.

[21] A. Rijnberk and H. De Vries, *Medical History and Physical Examination in Companion Animals*. Berlin, Germany: Springer, 1995.

[22] L. Fetters and J. Tilson, *Evidence Based Physical Therapy*. Philadelphia, PA, USA: FA Davis, 2018.

[23] A. J. Higgins and J. R. Snyder, *The Equine Manual E-Book*. Amsterdam, The Netherlands: Elsevier, 2013.

[24] C. Liu, D. Springer, Q. Li, B. Moody, R. A. Juan, F. J. Chorro, F. Castells, J. M. Roig, I. Silva, and A. E. Johnson, "An open access database for the evaluation of heart sound algorithms," *Physiological Meas.*, vol. 37, no. 12, pp. 2181–2213, Dec. 2016.

[25] *Classification of Heart Sound Signal Using Multiple Features*. Accessed: May 17, 2023. [Online]. Available: https://github.com/yaseen21khan/Classification-of-Heart-Sound-Signal-Using-Multiple-Features

[26] S. Picard, C. Chapdelaine, C. Cappi, L. Gardes, E. Jenn, B. Lefevre, and T. Soumarmon, "Ensuring dataset quality for machine learning certification," in *Proc. IEEE Int. Symp. Softw. Rel. Eng. Workshops (ISSREW)*, Oct. 2020, pp. 275–282.

[27] E. Grooby, J. He, J. Kiewsky, D. Fattahi, L. Zhou, A. King, A. Ramanathan, A. Malhotra, G. A. Dumont, and F. Marzbanrad, "Neonatal heart and lung sound quality assessment for robust heart and breathing rate estimation for telehealth applications," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 12, pp. 4255–4266, Dec. 2021.

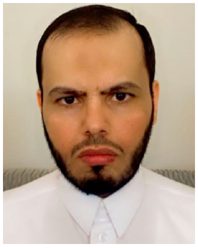
[28] S. N. Ali, S. B. Shuvo, M. I. S. Al-Manzo, M. Hasan, and T. Hasan, "An end-to-end deep learning framework for real-time denoising of heart sounds for cardiac disease detection in unseen noise," *IEEE Access*, vol. 11, pp. 87887–87901, 2023.

[29] D. I. Hoult and R. E. Richards, "The signal-to-noise ratio of the nuclear magnetic resonance experiment," *J. Magn. Reson.*, vol. 24, no. 1, pp. 71–85, Oct. 1976.

[30] S. Kiranyaz, M. Zabihi, A. B. Rad, T. Ince, R. Hamila, and M. Gabbouj, "Real-time phonocardiogram anomaly detection by adaptive 1D convolutional neural networks," *Neurocomputing*, vol. 411, pp. 291–301, Oct. 2020.

[31] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962.

[32] D. B. Springer, L. Tarassenko, and G. D. Clifford, "Logistic regression-HSMM-based heart sound segmentation," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 4, pp. 822–832, Apr. 2016.



SAMI ALRABIE received the M.Sc. degree in computer science from King Abdulaziz University, Saudi Arabia, where he is currently pursuing the Ph.D. degree with the Faculty of Computing and Information Technology. He is a part-time Lecturer with Arab Open University. His research interests include artificial intelligence, AI in biomedical applications, and biomedical signal processing.

AHMED BARNAWI received the M.Sc. degree from the University of Manchester Institute of Science and Technology (UMIST), U.K., in 2001, and the Ph.D. degree from the University of Bradford, U.K., in 2005. He is currently a Professor with the Faculty of Computing and Information Technology, King Abdulaziz University. He is also the Managing Director of the KAU Cloud Computing and Big Data Research Group. He has published more than 120 papers in peer-reviewed journals and conferences. His research interests include wireless communications, artificial intelligence, cloud computing, and robotics.