SPRSound: Open-Source SJTU Paediatric Respiratory Sound Database

Qing Zhang[1,2§], Jing Zhang[2,3,4§], Jiajun Yuan[4,5,6§], Huajie Huang[1], Yuhang Zhang[1], Baoqin Zhang[7], Gaomei Lv[8], Shuzhu Lin[9], Na Wang[8], Xin Liu[10], Mingyu Tang[3], Yahua Wang[3], Hui Ma[3], Lu Liu[3], Shuhua Yuan[3], Hongyuan Zhou[11], Jian Zhao[1,2], Yongfu Li[1,2*], Yong Yin[2,3,4*], Liebin Zhao[4*], Guoxing Wang[1,2], Yong Lian[1,2]

***Abstract*—It has proved that the auscultation of respiratory sound has advantage in early respiratory diagnosis. Various methods have been raised to perform automatic respiratory sound analysis to reduce subjective diagnosis and physicians' workload. However, these methods highly rely on the quality of respiratory sound database. In this work, we have developed the first open-access paediatric respiratory sound database, SPRSound. The database consists of 2,683 records and 9,089 respiratory sound events from 292 participants. Accurate label is important to achieve a good prediction for adventitious respiratory sound classification problem. A custom-made sound label annotation software (SoundAnn) has been developed to perform sound editing, sound annotation, and quality assurance evaluation. A team of 11 experienced paediatric physicians is involved in the entire process to establish golden standard reference for the dataset. To verify the robustness and accuracy of the classification model, we have investigated the effects of different feature extraction methods and machine learning classifiers on the classification performance of our dataset. As such, we have achieved a *score* of 75.22%, 61.57%, 56.71%, and 37.84% for the four different classification challenges at the event level and record level.**

***Index Terms*—Paediatric Respiratory Sound Database, Respiratory Sound Classification, Quality Assurance**

Manuscript received XXXX XX, XXXX; accepted XXXX XX, XXXX. Date of publication XXXX XX, XXXX; date of current version XXXX XX, XXXX. This work is supported in part by the National Key Research and Development Program of China under Grant No. 2019YFB2204500, in part by the Science, Technology and Innovation Action Plan of Shanghai Municipality, China under Grant No. 22015821400, in part by the Science and Technology Innovation-Biomedical Supporting Program of Shanghai Science and Technology Committee under Grant No. 19441904400 and in part by the Program for artificial intelligence innovation and development of Shanghai Municipal Commission of Economy and Informatization under Grant No. 2020-RGZN-02048 and in part by the Translational Medicine Cross Research Fund of Shanghai Jiao Tong University (No. YG2021QN109 and YG2022QN095). This paper was recommended by Associate Editor XXX.

[§]Qing Zhang, Jing Zhang and Jiajun Yuan contributed equally to this work.
*The corresponding authors are Yongfu Li (e-mail: yongfu.li@sjtu.edu.cn), Yong Yin (e-mail: yinyong@scmc.com.cn) and Liebin Zhao (e-mail: zhaoliebin@126.com).
The authors[1] are with the Department of Micro-Nano Electronics and MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University, China.
The authors[2] are with Pediatric AI Clinical Application and Research Center, Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, China.
The authors[3] are with Department of Respiratory Medicine, Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, China.
The authors[4] are with the Shanghai Engineering Research Center of Intelligence Pediatrics (SERCIP), China.
The author[5] is with School of Computer Engineering and Science, Shanghai University, China.
The author[6] is with Sanya Maternity and Child Care Hospital, China.
The author[7] is with Department of Paediatrics, Taicang Affiliated Hospital of Soochow University, The First People's Hospital of Taicang, China.
The authors[8] are with Linyi City People Hospital, China.
The author[9] is with Department of Paediatrics, Fengcheng Hospital, China.
The authors[10] are with Department of Pediatrics, Fifth Affiliated Hospital of Harbin Medical University, China. The authors[11] is with Shanghai Tuoxiao Intelligent Technology Co., Ltd, China.

I. Introduction

Respiratory disease, one of the top global causes of death, has seriously threatened our life and social economic [1]. To prevent deteriorating of health condition and reduce death rate, early diagnosis for timely treatment are of great importance [2]. Lung auscultation plays an important role in identifying adventitious respiratory sounds and the related respiratory diseases [2]. It is performed using a stethoscope by experienced physicians [3]. However, the inter-listener variability among physicians and the lack of quantitative measurements make lung auscultation a subjective process [4]–[6]. Therefore, it is necessary to develop automatic solutions to address the present challenge.

The state-of-the-art solutions [7]–[38] have been reported to detect adventitious respiratory sounds automatically. However, the robustness and accuracy of these solutions highly rely on the quality of the respiratory sound database [39]. Therefore, it is essential to build high-quality and standardized respiratory sound databases for both algorithm development and evaluation.

Presently, there are three main open-access respiratory sound database, namely the International Conference on Biomedical and Health Informatics (ICBHI) 2017 database [40], King Abdullah University Hospital (KAUH) database [41], and the HF_Lung_V1 database [42]. However, the participants in these databases are mainly adults with a wide age range. Since the characteristics of children's respiratory sounds are slightly different from adults [43], [44], it is important to consider this population group when we want to develop a holistic digital auscultation system. Besides, the present respiratory sound databases do not support the identification of low-quality respiratory sounds as the noisy sounds from the environment or within the human body can inevitably affect the signal quality [45]. Furthermore, in the present respiratory sound databases, the label annotation of each respiratory sound is either done by one physician or lack of proper quality assurance methodology.

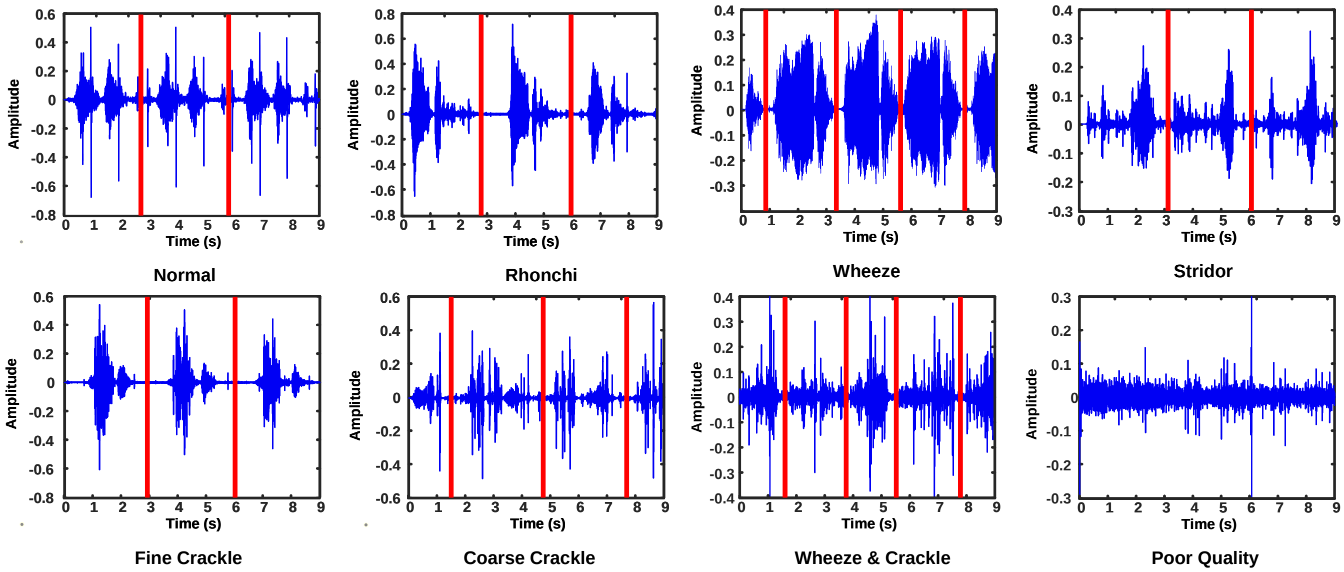


Fig. 1. Visualization of respiratory sounds waveform, including *Normal*, *Rhonchi*, *Wheeze*, *Stridor*, *Coarse Crackle*, and *Fine Crackle*. *Poor Quality* refers to the respiratory sound records with poor signal quality (e.g., noisy). Red lines refer to the respiratory sound event boundaries.

TABLE I
THE CHARACTERISTICS OF THE RESPIRATORY SOUNDS.

Sound Type	Continuity	Pitch	Acoustics	Duration	Frequency Range	Disorder
Normal	-	-	-	-	[50 Hz, 2.5 kHz]	-
Rhonchi	Continuous	Low	Musical	$\geqslant$ 80 ms	$\leqslant$ 200 Hz	Bronchitis
Wheeze	Continuous	High	Musical	$\geqslant$ 80 ms	$\geqslant$ 400 Hz	Asthma, Foreign body
Stridor	Continuous	High	Musical	$\geqslant$ 250 ms	$\geqslant$ 500 Hz	Foreign body
Coarse Crackle	Discontinuous	Low	Non-musical	∼15 ms	∼350 Hz	Chronic Bronchitis
Fine Crackle	Discontinuous	High	Non-musical	∼5 ms	∼650 Hz	Pneumonia

To overcome these aforementioned challenges, we have developed a paediatric respiratory sound database, SPRSound. The main contributions of this paper are presented as follows:

1) **Open-Access Database:** The first open-source paediatric respiratory sound database (SPRSound) consists of 2,683 records and 9,089 respiratory sound events from 292 paediatric participants, dedicated to respiratory sound analysis of paediatric population. *Poor Quality* is firstly incorporated in the respiratory sound database to encourage researchers to identify low-quality sounds and perform accurate analysis. Specifically, there are 1,785 *Normal*, 233 continuous adventitious sound (*CAS*), 347 discontinuous adventitious sound (*DAS*), 131 *CAS & DAS*, and 187 *Poor Quality* records. The number of *Normal*, *Rhonchi*, *Wheeze*, *Stridor*, *Coarse Crackle*, *Fine Crackle*, and *Wheeze & Crackle* events are 6,887, 53, 865, 17, 66, 1,167, and 34.
2) **Label Annotation Methodology:** A custom-made sound label annotation software (SoundAnn) has been developed to perform sound editing, sound annotation, and quality assurance evaluation. A team of 11 experienced paediatric physicians is involved in the entire process to establish golden standard reference for the dataset. This enables us to reduce the inter-listener variability, improve the quality assurance process, and ensure the reliability of the database.
3) **Benchmark Results:** A series of experiments based on the database has been conducted for respiratory sound classification at event level (Task 1) and record level (Task 2), which demonstrates the effectiveness and reliability of our database. The experimental results demonstrate that the mel-frequency cepstral coefficient (MFCC) method with naive bayes (NB) and support vector machine (SVM) achieve the *score* of 75.22%, 61.57%, 56.71%, and 37.84% for Task 1-1, 1-2, 2-1, and 2-2, respectively.

The rest of this paper is organized as follows. Section II describes the types of respiratory sounds and the present open-access respiratory sound database. Section III provides the materials and methods of our SPRSound Database, including the subject recruitment, participant demographic, recording instrument, label annotation methodology, quality assurance methodology, and the description of our SPRSound. Section IV details the classification tasks and evaluation methodology, which includes the evaluation metrics, classification framework and its results. Section VI discusses the strengths and limitations of our database, and makes comparison between our database and others.

Fig. 2. An example of *CAS* record with 1 *Wheeze* and 2 *Normal* respiratory sound events.

II. Preliminaries

A. Type of Respiratory Sounds

Respiratory sound provides important information in the diagnosis of pulmonary disorder through acoustic analysis [2]. Table I reports the characteristics of different types of respiratory sound, including continuity, pitch, acoustics, duration, frequency range, and the associated diseases. Specifically, the respiratory sounds are divided into normal and adventitious sounds. Adventitious sounds can be further classified as continuous adventitious sounds (*CAS*, including *Rhonchi*, *Wheeze*, and *Stridor*) and discontinuous adventitious sounds ((*DAS*, including *Coarse Crackle* and *Fine Crackle*) sounds based on the duration [46].

Rhonchi is an important symptom of bronchitis, which is a low-pitched and snoring-like sound with a frequency range and duration of less than 200 Hz and more than 80 ms, respectively [47], [48]. *Wheeze* is a high-pitched and whistling-like sound, which is utilized for diagnosing asthma and foreign body [47], [48]. Its frequency range and duration time are more than 400 Hz and longer than 80 ms, respectively. *Stridor* is a high-pitched and musical sound with a frequency exceeding 500 Hz and duration exceeding 250 ms [47]. *Coarse Crackle* and *Fine Crackle* are explosive and non-musical sounds with high and low pitches, respectively [49]–[51]. The frequency range and duration of *Coarse Crackle* are around 350 Hz and 15 ms, respectively, which is associated with chronic bronchitis [52]. *Fine Crackle* is synonymous with pneumonia with a frequency range and duration of around 650 Hz and 5 ms, respectively [52]. Fig. 1 illustrates the waveform of respiratory sounds of different types, with red lines denoting the boundaries of each respiratory sound event.

Prior works have analyzed respiratory sounds at various levels, including event level [31]–[33] and record level [8], [20]. At the event level, respiratory sounds are analyzed after manually annotating the specific location of the respiratory sounds while respiratory sounds analysis is performed on the whole records at record level. Since each record consists of multiple respiratory sound events, we simplify the types of respiratory sounds at record level as *Normal*, *CAS*, *DAS*, and *CAS & DAS* according to the presence/absence of continuous and discontinuous respiratory sound events. As illustrated in Fig. 2, a record is labeled as *CAS* at the record level, which consists of 2 *Normal* and 1 *Wheeze* respiratory sound events.

B. Open-Access Respiratory Sound Database

Presently, there are three open-access respiratory sound databases (ICBHI 2017 [40], KAUH [41], and HF_Lung_V1 databases [42]), which are important for both algorithm development and evaluation. However, these databases are mainly based on adults population with a wide age range.

1) ICBHI Database: The International Conference on Biomedical and Health Informatics (ICBHI) 2017 database [40] consists of 920 respiratory sound records from 126 participants, aged between 0.25 and 93 years old. ICBHI 2017 database provides respiratory sound annotation at event level, including 3,646 *Normal*, 1,864 *Crackle*, 886 *Wheeze*, and 506 *Crackle & Wheeze*.

2) King Abdullah University Hospital (KAUH) Database: King Abdullah University Hospital (KAUH) database [41] consists of 112 respiratory sound records from 112 participants, aged between 21 and 90 years old. The database provides respiratory sound annotation at record level, including 35 *Normal*, 23 *Crepitations*, 41 *Wheeze*, 8 *Crackle*, 1 *Bronchial*, 2 *Wheeze & Crackle*, and 2 *Bronchial & Crackle*.

3) HF_Lung_V1 Database: The HF_Lung_V1 database [42] consists of 9,765 respiratory sound records from 279 participants older than 20 years old. The database provides respiratory sound annotation at event level, including 15,606 *Crackle*, 8,457 *Wheeze*, 4,740 *Rhonchi*, and 686 *Stridor*.

III. Open-Source SJTU Paediatric Respiratory Sound (SPRSound) Database

A. Subject Recruitment

The study was conducted at the paediatric respiratory department of Shanghai Children's Medical Center (SCMC), Shanghai Jiao Tong University affiliated hospital, China from May 2021 to May 2022. SCMC is one of the national children's medical centers with the rank of "Grade 3, Class A", the highest level of hospital classification in China, which provides the most comprehensive paediatric specialties and healthcare services throughout China. Thus, the sample population is a good representative of the paediatric population within China. All participants voluntarily participate in the study have informed consent from their parents or guardians. Ethical approval was obtained from the Institutional Review Board of SCMC (approval No. SCMCIRB-K2019056-1).

B. Participant Demographic

Table II summarises the detailed demographics of the database. Fig. 3 illustrates the distribution of gender, age, and health condition of participants. A total of 292 participants with 140 female (F) and 152 male (M) were involved. Except for 20 participants without respiratory diseases as the control group, the primary diagnoses of these participants were asthma, bronchitis, pneumonia (non-severe), pneumonia (severe), and other respiratory diseases with the number of 33, 19, 177, 20, and 23 respectively. The age of participants ranges from 0.2 years to 16.2 years with a mean age of 5.5 years.

C. Instrument

The respiratory sounds were collected with a digital stethoscope (Yunting model II Stethoscope, Yunting II). The

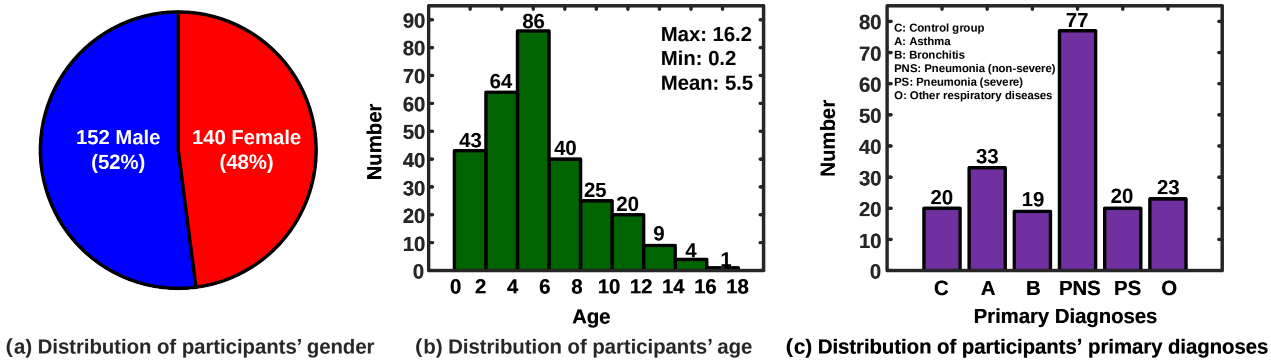


Fig. 3. (a) Distribution of participants' gender in our database. (b) Distribution of participants' age. (c) Distribution of participants' primary diagnoses, including Control group (C), Asthma (A), Bronchitis (B), Pneumonia (non-severe) (PNS), Pneumonia (severe) (PS), and Other respiratory diseases (O).

TABLE II
DEMOGRAPHICS OF THE PARTICIPANTS.

Primary Diagnoses	Participants #	Gender	Age
Asthma	33 (11.3%)	19 M, 14 F	6.7±2.9
Bronchitis	19 (6.5%)	10 M, 9 F	4.1±2.3
Pneumonia (non-severe)	177 (60.6%)	87 M, 90 F	5.0±3.1
Pneumonia (severe)	20 (6.8%)	8 M, 12 F	5.0±5.0
Other respiratory diseases	23 (7.9%)	13 M, 8 F	6.6±4.5
Control group	20 (6.8%)	13 M, 7 F	6.5±3.0
Total	292	152 M, 140 F	5.5±3.4

sampling frequency and the quantization resolution of the digital stethoscope are 8 kHz and 16 bit, respectively. The collection process was performed by experienced physicians, well trained in specimen collection and clinical measurements. In our preliminary study, we observe that respiratory sounds of pediatric participants are weaker than those of adults. Besides, the respiratory sounds are highly affected by the heart sound when acquired at the front chest. Therefore, respiratory sounds were acquired at four back locations, including left posterior, left lateral, right posterior, and right lateral (Fig. 4). The collection duration for each location lasted over 9 seconds to ensure at least two respiratory cycles (a respiratory cycle includes a breathing in (inhalation) followed by a breathing out (exhalation)) [53]. During the collection process, the infants are either resting on the arms of their parents or in sitting, supine or prone positions. Participants were required to keep quiet and cooperate with the physicians. Respiratory sounds recorded by the digital stethoscope were saved in waveform record file format (.wav) and uploaded to the cloud server for subsequent analysis and annotation.

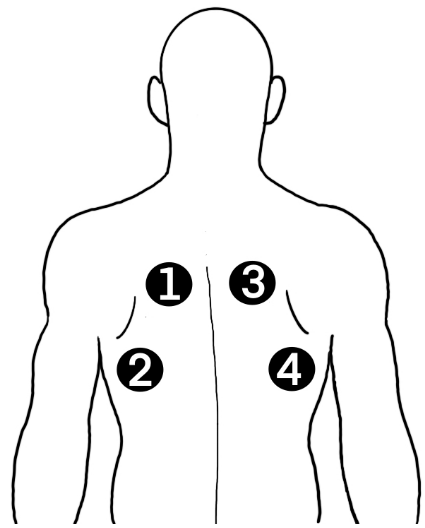


Fig. 4. Chest locations (including (1) left posterior, (2) left lateral, (3) right posterior, and (4) right lateral) used to collect respiratory sounds.

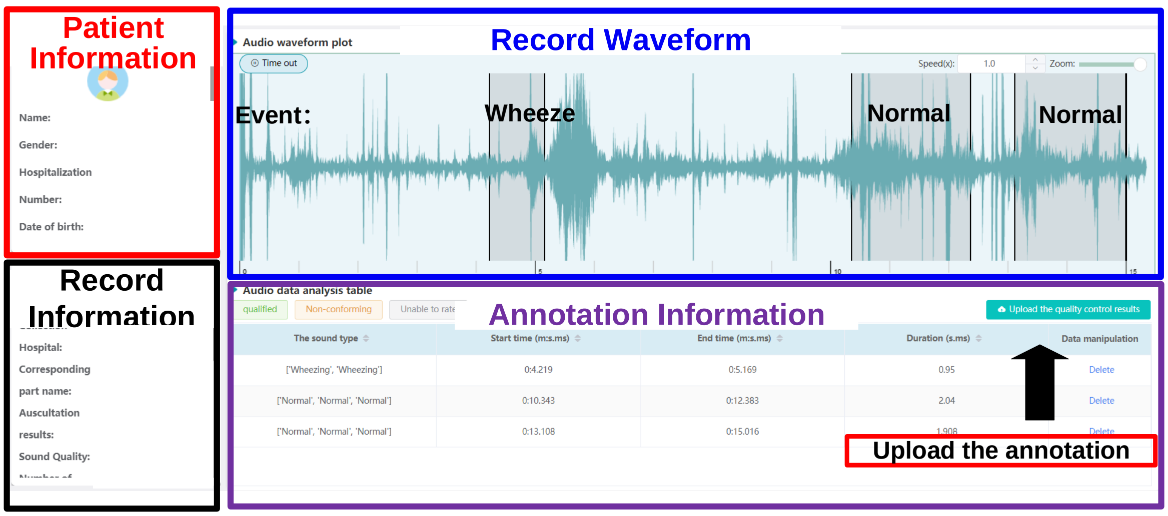
D. Label Annotations Methodology

Accurate label is important to achieve a good prediction for adventitious respiratory sound classification problem [54]. A custom-made sound label annotation software (SoundAnn) has been developed to perform sound editing, sound label annotations, and quality assurance evaluation. Each respiratory sound record is playback online and the corresponding waveform is illustrated in the annotation user interface (UI) of the SoundAnn software with sound loudness and time on the vertical and horizontal axes, respectively. Respiratory sound events are identified and labelled from the records with the start and end time position by physicians. Note that respiratory sound records with poor signal quality, which is hard for physicians to decide their perception of quality acoustically, were labelled as *Poor Quality* at the record level. As illustrated in Fig. 5, a record ('40877908_4.0_0_p2_3440') was loaded and annotated as *CAS* at the record level, including 2 *Normal* and 1 *Wheeze* respiratory sound events.

E. Quality Assurance Methodology

A well trusted dataset as a gold standard (GS) reference is of great importance for benchmarking different classification algorithms [55]. Since the diagnosis is subjective and highly relies on the physician's clinical experience, an excellent quality assurance methodology would be required to build a reliable dataset [45]. As illustrated in Fig. 6, our quality assurance methodology involves the signal quality assessment on each respiratory sound record by a team of 11 experienced

(a) The record loading UI of the annotation software



(b) The annotation UI of the annotation software

Fig. 5. Respiratory sound annotation software (a) The record loading UI (b) The annotation UI.

paediatric physicians. For poor/low signal quality record, it is labeled as *Poor Quality* records [56]. To the best of our knowledge, this is the first open source respiratory sound database incorporating with *Poor Quality* label, encouraging researchers to identify low-quality sounds and perform accurate analysis. For record with better signal quality, it is segmented into multiple respiratory sound events and annotated by 3 physicians independently. The labels are termed as GS reference when there is 100% consistent with the label annotations. When there inconsistent in the label annotations, the records would be forwarded to 2 authoritative physicians for further adjudication.

F. Database Description

SPRSound database is the first open-access paediatric respiratory sound database, jointly developed by Shanghai Jiao Tong University and its affiliated hospitals. The database contains 2,683 records and 9,089 respiratory sound events from 292 participants with a total duration of 8.2 hours. The label annotations at event and record levels are provided in the database. At the event level, the number of *Normal*, *Rhonchi*, *Wheeze*, *Stridor*, *Coarse Crackle*, *Fine Crackle*, and *Wheeze & Crackle* are 6,887, 53, 865, 17, 66, 1,167, and 34, respectively. At the record level, the number of *Normal*, *CAS*, *DAS*, *CAS & DAS*, and *Poor Quality* records are 1,785, 233, 347, 131, and 187, respectively. The mean duration of respiratory sound events and records are 1.3s and 11s, respectively.

SPRSound database provides free access for all registered users and it can be downloaded publicly [1]. Respiratory sound records are saved in wave (.wav) format with naming rules as follows: The name of each record contains 5 pieces of information, separated with underscores, namely the number, age, gender, the record location, and the record

[1]https://github.com/SJTU-YONGFU-RESEARCH-GRP/Lung-Sound-Database

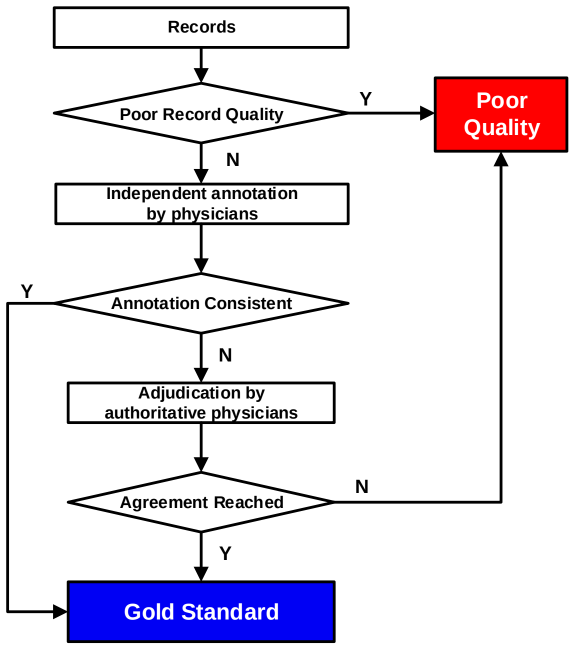


Fig. 6. The quality assurance evaluation process to establish gold standard annotations.

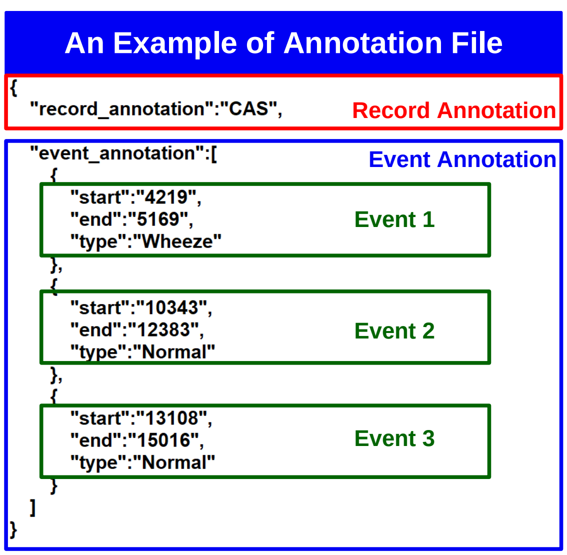


Fig. 7. An example of the label annotation file ('40877908_4.0_0_p2_3440') in JSON format.

number of the participants. The label annotations of each record is saved in JavaScript Object Notation (JSON) format with the same filenames, including the label annotations at the record level and event level (the beginning, the end, and the corresponding type of the respiratory sound event). Fig. 7 illustrates an example of the label annotation file ('40877908_4.0_0_p2_3440') in the SPRSound database.

TABLE III
THE DETAILED STATISTICS OF THE TRAINING AND TESTING SETS OF OUR DATABASE.

Level	Type	Training	Testing-1 (intra-patient)	Testing-2 (inter-patient)	Total
Record	Normal	1,303	241	241	1,785
	CAS	126	42	65	233
	DAS	248	75	24	347
	CAS & DAS	95	19	17	131
	Poor Quality	177	2	8	187
	Total	1,949	379	355	2,683
Event	Normal	5,159	688	1040	6,887
	Rhonchi	39	14	0	53
	Wheeze	452	108	305	865
	Stridor	15	2	0	17
	Coarse Crackle	49	14	3	66
	Fine Crackle	912	175	80	1,167
	Wheeze & Crackle	30	3	1	34
	Total	6,656	1,004	1,429	9,089

IV. EVALUATION METHODS

A. Training and Testing Sets

To verify the robustness and accuracy of the classification model, the variability within patients (intra-patient) and between patients (inter-patient) should be considered [34], [57], [58]. In this work, we have split the SPRSound database into training and testing sets, where the testing sets consist of two subsets for both intra-patient (testing-1) and inter-patient (testing-2) evaluation. The training set consists of 1,949 records and 6,656 respiratory sound events from 251 participants. The testing-1 dataset is derived from the same participants with training dataset, containing a total of 379 records and 1,004 respiratory sound events for intra-patient evaluation. The testing-2 dataset contains a total of 355 records and 1,429 respiratory sound events from another 41 participants for inter-patient evaluation, which can be further expand with the development of the database. The statistical details are reported in Table III.

B. Main Tasks

The development of respiratory sound classification models at event and recording levels are vital essential for the real-time monitoring [1]. Thus, given the SPRSound database, two levels of classification tasks are proposed in this work.

Task 1 (Respiratory Sound Classification at Event Level). Task 1-1 is a binary-class classification, aiming at classifying the respiratory sound events as *Normal* and *Adventitious*. Task 1-2 is a multi-class classification, aiming at classifying the respiratory sound events as *Normal*, *Rhonchi*, *Wheeze*, *Stridor*, *Corase Crackle*, *Fine Crackle*, and *Wheeze & Crackle*.

Task 2 (Respiratory Sound Classification at Record Level). Task 2-1 is a ternary class classification, aiming at classifying the respiratory sound records as *Normal*, *Adventitious*, and *Poor Quality* records. Task 2-2 is a multi-class classification, aiming at classifying the respiratory sound records as *Normal*, *CAS*, *DAS*, *CAD & DAS*, or *Poor Quality*.

C. Evaluation Metrics

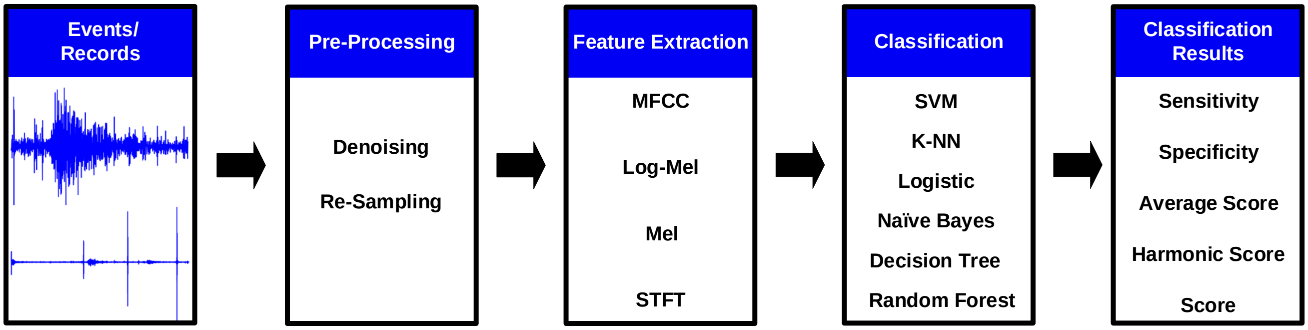


Fig. 8. The flow of respiratory sound classification framework for event level and record level.

Sensitivity (*SE*) and *specificity* (*SP*) are commonly used as the measures of diagnostic accuracy in the medical field [59]. ICBHI scientific challenge has employed *average score* (*AS*) and *harmonic score* (*HS*) as the evaluation metrics based on *SE* and *SP* [40]. In this work, we introduce *score* to make a comprehensive model evaluation for each task. The detailed definitions are as follows.

***Definition** 1 (Sensitivity (SE)): SE* denotes the ratio between the number of correctly recognized adventitious sounds (events/records) and the total number of adventitious sounds (events/records).

$$SE = \frac{\#\ of\ correctly\ recognized\ adventitious\ sounds}{\#\ of\ total\ adventitious\ sounds}, \tag{1}$$

***Definition** 2 (Specificity (SP)): SP* denotes the ratio between the number of correctly recognized *Normal* sounds (events/records) and the total number of *Normal* sounds (events/records).

$$SP = \frac{\#\ of\ correctly\ recognized\ normal\ sounds}{\#\ of\ total\ normal\ soundss}, \tag{2}$$

***Definition** 3 (Average Score (AS)): AS* denotes the arithmetic mean of *SE* and *SP*.

$$AS = \frac{SE + SP}{2}, \tag{3}$$

***Definition** 4 (Harmonic Score (HS)): HS* denotes the harmonic mean of *SE* and *SP*.

$$HS = \frac{2 * SE * SP}{SE + SP}, \tag{4}$$

***Definition** 5 (Score):* Overall, the score for each task is the arithmetic mean of *AS* and *HS*.

$$Score = \frac{AS + HS}{2}, \tag{5}$$

D. Classification Framework for Database Quality Evaluation

In this work, a respiratory sound classification framework for both event (Task 1) and record (Task 2) level is provided with three steps, including (i) pre-processing, (ii) feature extraction, and (iii) classification. The whole respiratory sound records/events with different lengths are taken as the inputs to perform pre-processing and feature extraction. The feature vectors after feature extraction have the same size and are fed into the classifier for classification. Notice that deep learning (DL) models highly rely on large amount of data while machine learning (ML) models have demonstrated remarkable success on small dataset with efficient feature extraction methods [56], [60], [61]. Therefore, we decide to employ ML models to minimize the imbalance problem in dataset and avoid overfitting problems without applying data augmentation techniques. Different feature extraction methods along with ML models are discussed based on our database as follows.

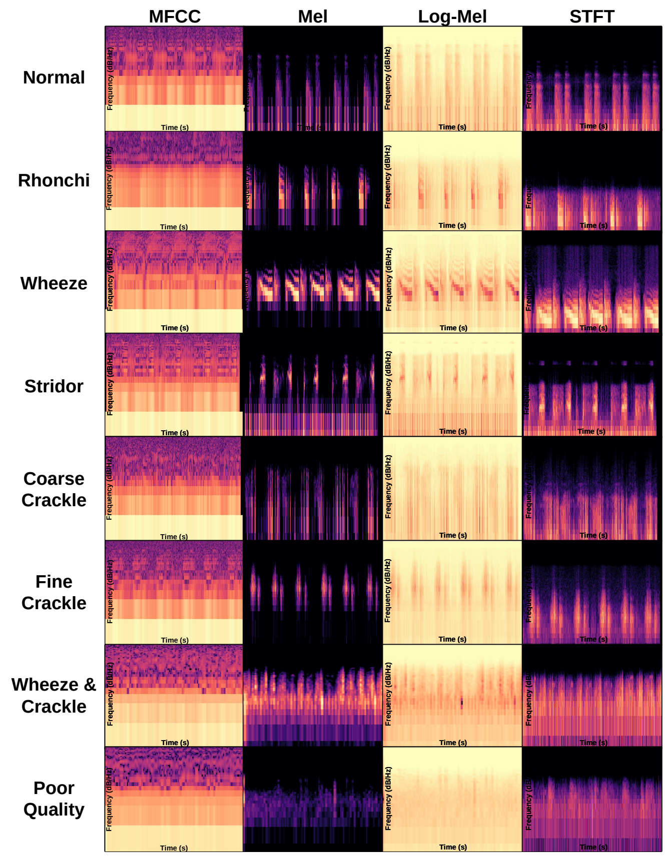


Fig. 9. The spectrogram representation of different features (including MFCC, Mel, Log-Mel, and STFT) for different types of respiratory sounds.

Pre-processing Methods with signal de-noising and re-sampling are necessary to enhance the classification performance [1]. Note that the background noises are

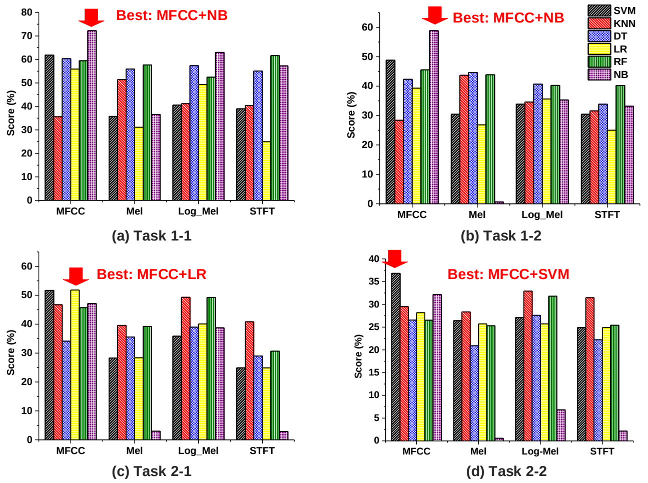


Fig. 10. Intra-patient comparison with different feature extraction methods and different classifiers based on the testing-1 set.

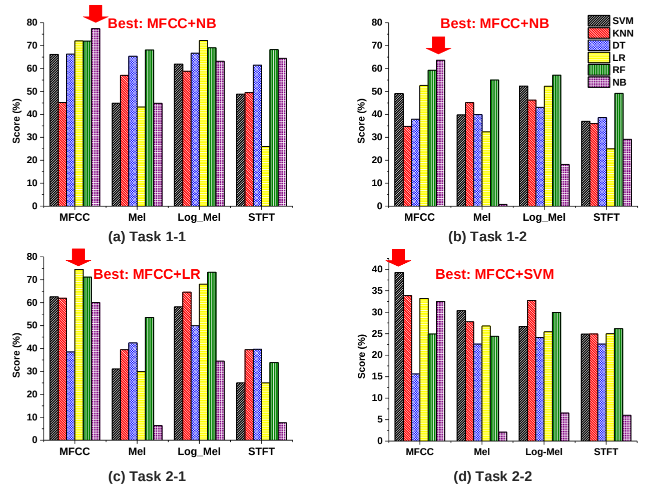


Fig. 11. Inter-patient comparison with different feature extraction methods and different classifiers based on the testing-2 set.

unavoidably recorded during the auscultation process. As described in Section II-A, the frequency of respiratory sounds ranges from 50 Hz to 2.5 kHz. A 5-th Butterworth band-pass filter with cut-off frequencies of 50 Hz and 2.5 kHz is designed for suppressing the background noises to obtain relatively pure respiratory sound signals. Respiratory sound signals are re-sampled to 8 kHz following Computerized Respiratory Sound Analysis (CORSA) guidelines [39].

Feature Extraction Methods plays a key role in the sound analysis frameworks, reducing redundant information and improving the classification accuracy [62]. Various feature extraction methods have been explored in the state-of-the-art works, such as mel-frequency cepstral coefficient (MFCC) [8], [15], [28], [30], [63]–[68], log-mel spectrogram (Log-Mel) [62], [69], [70], mel spectrogram [34], [71]–[73] (Mel), and short-time fourier transform (STFT) spectrogram [31], [74]–[76]. Fig. 9 illustrates the spectrogram representation of different features for various types of respiratory sounds. In this work, we have implemented these commonly used feature extraction methods for comparison.

Classification Models based on the feature extraction and selection have been experimental in the prior works [10], [15], [16], [20], [21], [24], [26]. Machine learning models are commonly used to overcome the over-fitting problem when data set is small [60] and have demonstrated significant success in the respiratory sound analysis tasks, such as support vector machine (SVM) [8]–[14], k-nearest neighbor method (K-NN) [15]–[19], logistic regression (LR) [20]–[22], naive bayes (NB) [16], [19], decision tree (DT) [23], [24], and random forest (RF) [21], [25]. We have also explore the effect of the aforementioned classifiers to obtain basic classification results for comparison.

E. Evaluation Results

We have implemented and evaluated the classification framework for each task with machine learning models using the scikit-learn library [77] and investigated the effects of different feature extraction methods and classification models

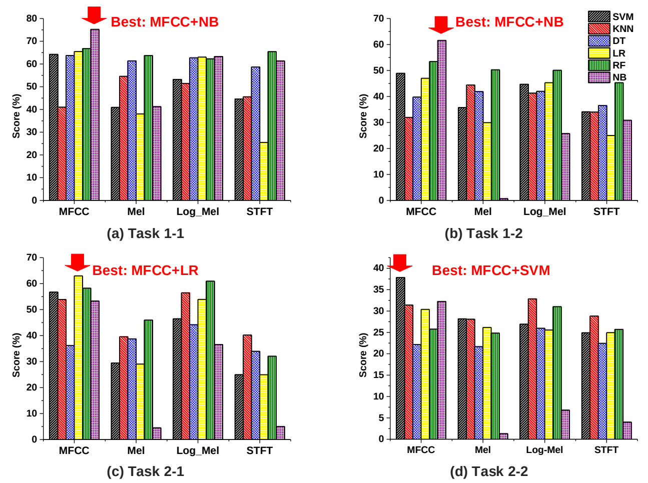


Fig. 12. Score comparison with different feature extraction methods and different classifiers for each classification task on the combined testing set (combination of testing-1 and testing-2 sets).

on classification performance. A total of 24 combinations with different feature extraction methods (including MFCC, Log-Mel, Mel, and STFT) and classifiers (including SVM, KNN, DT, LR, RF, and NB) have been implemented for each task. Specifically, the window length and the number of coefficients in MFCC are the length of input signals and 128, respectively. Hanning window with frame length and hop size of 2048 and 512 samples is adopted in Mel and Log-Mel. Hanning window with frame length and hop size of 80 and 40 samples is adopted in STFT. The kernel function and the parameter C in SVM are set as polynomial kernel and 2, respectively. The number of neighbors is set as 5 in KNN. L2 regularization is used as the penalty term in LR. Gaussian naive bayes classifier is adopted in NB. The classification criterion and splitter in DT are set as “gini” and “best”, respectively. The number of trees and the classification criterion in RF are set as 100 and “gini”. To make a fair comparison, we apply the One-vs-the-Rest (OvR) strategy for all classifiers. Min-max normalization method are employed

TABLE IV
SUMMARY OF THE CLASSIFICATION RESULTS FOR EACH TASK.

Task	Testing-1 (intra-patient)					Testing-2 (inter-patient)					Combined Testing				
	SE (%)	SP (%)	AS (%)	HS (%)	Score (%)	SE (%)	SP (%)	AS (%)	HS (%)	Score (%)	SE (%)	SP (%)	AS (%)	HS (%)	Score (%)
Task 1-1	57.59	90.55	74.07	70.41	72.24	75.83	79.04	77.44	77.40	77.42	67.66	83.62	75.64	74.80	75.22
Task 1-2	38.61	89.10	63.85	53.87	58.86	51.93	77.88	64.90	62.31	63.61	45.95	82.35	64.15	58.99	61.57
Task 2-1	38.24	69.71	53.97	49.38	51.68	58.49	66.81	62.65	62.37	62.51	47.11	68.26	57.68	55.74	56.71
Task 2-2	18.38	70.54	44.46	29.16	36.81	23.58	64.32	43.95	34.51	39.23	20.66	67.43	44.04	31.63	37.84

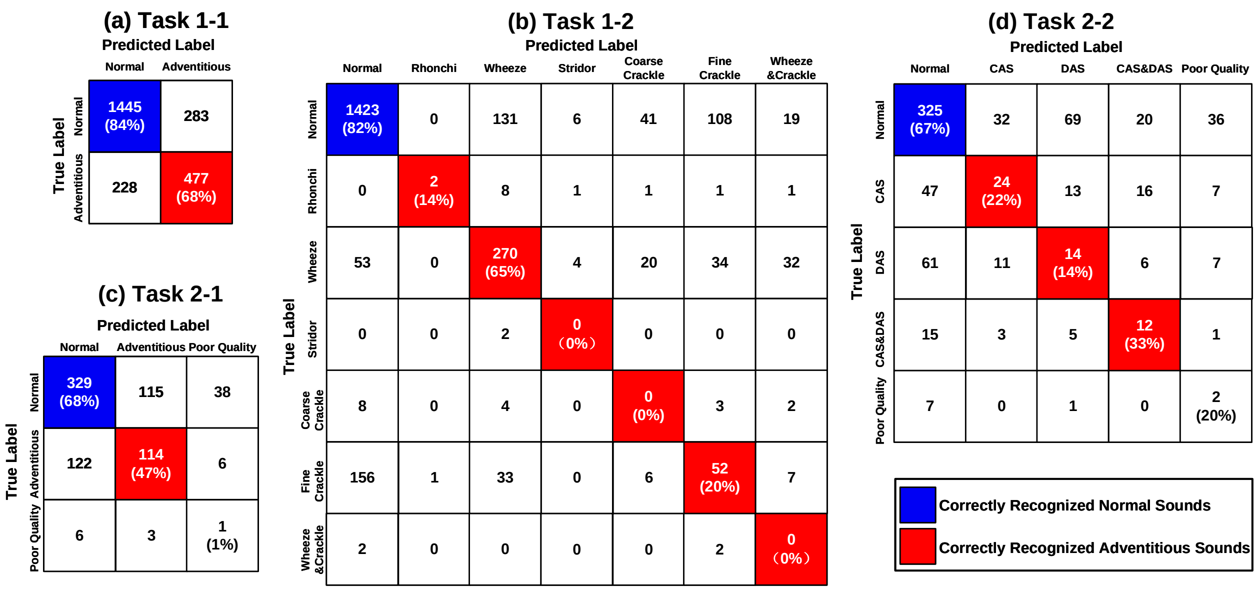


Fig. 13. Confusion matrix of Task 1-1 (a), Task 1-2 (b), Task 2-1 (c), and Task 2-2 (d).

after feature extraction to eliminate systematic bias with different auscultation devices.

Figs. 10, 11, and 12 illustrate the intra-patient, inter-patient, and the overall classification comparison with different combinations on the testing-1 and testing-2 sets, respectively. We observe that the MFCC method achieves the best classification performance for all tasks, which sufficiently models the response of the human auditory system and provides features with good discrimination and low correlation [78]. For Task 1-1 and Task 1-2, the MFCC method with the NB classifier achieves the highest *score*, suggesting that low correlation MFCC features yield good classification performance with NB classifier [79]. The MFCC method with LR classifier and SVM classifier achieve a relatively high classification performance for Task 2-1 and Task 2-2.

Following the experimental comparison results, the MFCC method along with NB and SVM models are employed for Task 1 and Task 2, respectively. As reported in Table IV, the *score* of Task 1-1, Task 1-2, Task 2-1, and Task 2-2 are 75.22%, 61.57%, 56.71%, and 37.84%, respectively. The results show that the inter-patient *score* is slightly higher than the intra-patient *score* due to the difference in proportions of adventitious sounds. Specifically, the proportions of *Wheeze* in the adventitious sounds are 34% and 78% in intra-patient and inter-patient testing sets, respectively. Furthermore, the classification model achieves relatively high accuracy in recognizing *Wheeze*. Therefore, the *SE* of the inter-patient testing set is higher than that of the intra-patient testing set, resulting in slightly higher *score*. Overall, the intra-patient and inter-patient variations of our database are small, indicating the effectiveness of our database for algorithm development and evaluation. Fig. 13 illustrated the confusion matrix for each task, where the x-axis and y-axis refer to the predicted and true labels, respectively. The diagonal values refer to the number of correctly recognized respiratory sounds (events/records). We observe that the best classification performance on Task 1 and Task 2 occurs at *Normal*, which is the majority class. While the worst performance occurs in the minority classes for each task.

V. DISCUSSIONS

A. Discussion of Open Source Database

In this study, we have developed the first open-access respiratory sound database (SPRSound) dedicated to infants, children, and adolescents (the details are shown in Section III). Fig. 14 illustrates the age distribution comparison with the prior open-access databases described in Section II-B. Note that the participants in HF_Lung_V1 database [42] are aged older than 20 years while the detailed demographic information is lost. Besides, *Poor Quality* respiratory records have been first incorporated in the respiratory sound database, facilitating researchers to develop innovative algorithms for

TABLE V
INTER-PATIENT CLASSIFICATION RESULTS BASED SPRSOUND DATABASE, ICBHI 2017 DATABASE, AND THE DATA FUSION DATABASE.

Task	Database	SE (%)	SP (%)	AS (%)	HS (%)	Score (%)
Task 1-1	Ours, 2-classes	75.83	79.04	77.44	77.40	77.42
	ICBHI 2017, 2-classes	63.98	43.89	53.93	52.06	53.00
	Data Fusion, 2-classes	59.71	63.46	61.58	61.53	61.55
Task 1-2	Ours, 7-classes	51.93	77.88	64.90	62.31	63.61
	ICBHI 2017, 4-classes	20.39	70.87	45.63	31.67	38.65
	Data Fusion, 6-classes	23.95	74.65	49.30	36.26	42.78
Task 2-1	Ours, 3-classes	58.49	66.81	62.65	62.37	62.51
	ICBHI 2017, 2-classes	77.95	28.35	53.15	41.57	47.36
	Data Fusion, 3-classes	71.67	66.85	69.26	69.17	69.22
Task 2-2	Ours, 5-classes	23.58	64.32	43.95	34.51	39.23
	ICBHI 2017, 4-classes	31.89	29.92	30.90	30.87	30.89
	Data Fusion, 5-classes	25.00	72.01	48.51	37.11	42.81

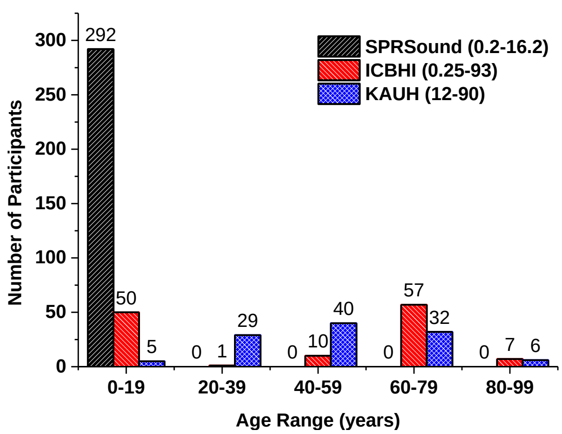


Fig. 14. Age comparison between our database, ICBHI 2017 [40], and KAUH database [41].

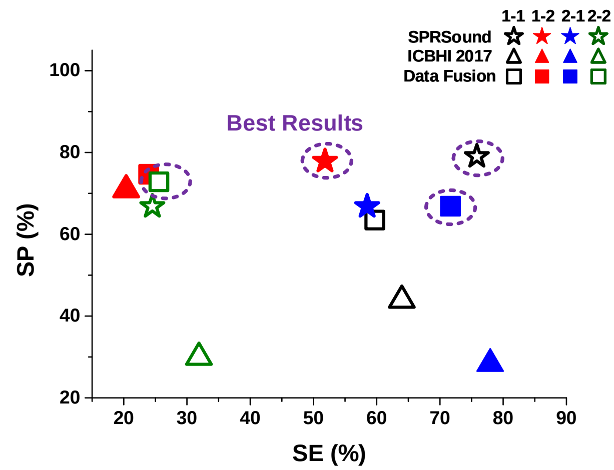


Fig. 15. The illustration of the inter-patient classification results with SPRSound, ICBHI 2017 [40], and the data fusion database.

low-quality sounds identification. We have also introduced a quality assurance methodology to build golden standard (GS) reference with multiple experienced physicians, which reduces the subjectivity of the diagnosis and generates a well-trusted database. In addition, our database aims to encourage researchers and engineers to perform accurate respiratory sound analysis and develop wearable devices (e.g., [80]–[85]) for long-term and mobility monitoring. The availability of our database with well-characterized clinical records from participants and GS reference has significant value in medical education and biomedical field.

There are some limitations which need to be overcome in the near future. Currently, the participants of the database are from China, there is a need to obtain participants from various world to build a more robust database. Besides, physicians involved in the study are from China, which may contribute bias to the annotation process as they have the same clinical training. Database can be more enriching with more physicians from other countries participated in the annotation process. Therefore, the number of participants, the sources of participants, and the number of physicians should be expanded with the deepening and development of the research.

B. Discussion of the Classification Results Based on SPRSound Database

In this work, we focus on respiratory sound classification at event level (Task 1-1 and Task 1-2) and record level (Task 2-1 and Task 2-2) based on our database. The experimental results of our classification framework have been presented in Section IV, with the *score* of 75.22%, 61.57%, 56.71%, and 37.84% on Task 1-1, Task 1-2, Task 2-1, and Task 2-2, respectively. These classification results can be further improved as there are few factors that affect the classification performance. Firstly, there is a class imbalance problem in our database due to the low probability of obtaining adventitious respiratory sounds, causing the classifiers ignoring the minority classes (e.g., *Stridor*). As reported in Table III, the number of *Normal* and *Stridor* respiratory sound events are 6,887 and 17 respectively, which is 405× different. Data-level and algorithm-level methods should be considered to address the class imbalanced problem and improve the classification performance [86]. Besides, we employ a 5th band-pass filter for de-noising, more adaptive pre-processing methods [87]–[90] to filter out influential noise are yet to investigate. In addition, more superior feature extraction methods and classifiers for better classification performance can be developed in future studies.

C. Discussion of the Classification Results Based on Data Fusion Database

We have conducted experiments with our classification framework based on the data fusion database, integrated with our SPRSound database and other open-access databases. Notice that ICBHI 2017 database [40] contains 920 records and 6,902 respiratory sounds collected using different stethoscopes (e.g., 3M Littmann Classic II SE and 3M Litmmann 3200). Besides, ICBHI 2017 database is conducted in Europe, which expand the diversity of the data fusion database. However, the KAUH database [41] only provides annotations at the record level. The HF Lung V1 database [42] contains a large amount of adventitious respiratory sound events (29,489) annotated by a single physician, which can contribute to a huge bias toward the data fusion database. Therefore, only ICBHI 2017 database [40] is assembled in the data fusion database.

Table V reports the classification performance of the data fusion database, which is build by merging the training and testing sets of SPRSound database and ICBHI 2017 database, respectively. To accommodate the types of respiratory sounds, we combine the *Corase Crackle* and *Fine Crackle* together as *Crackles* in the data fusion database. Notice that the official testing set of ICBHI database is for inter-patient comparison, Table V also reports the inter-patient classification performance based on the SPRSound database and ICBHI 2017 database for comparison. To make a fair comparison, the classification results are obtained after conducting experiments with different feature extraction methods and classifiers. Finally, the MFCC method along with LR and SVM classifiers are employed for Task1 and Task 2 in ICBHI 2017 database. The MFCC method with SVM classifier is adopted in the data fusion database for both Task 1 and Task 2. We observe that our database outperforms ICBHI 2017 database and the data fusion database at the event level (Task 1-1 and Task 1-2). The reason is that our database with quality assurance methodology enables us to build golden standard (GS) reference and ensure the reliability of our database. Compared to ICBHI 2017 database, the data fusion database has improved the classification performance at both the event level and record level, indicating the integrability and quality of our database.

VI. Conclusion

We have built an open-access respiratory sound database in paediatric population, containing 2,683 records and 9,089 respiratory sound events from 292 participants. A quality assurance methodology to build golden standard reference has been adopted to ensure the reliability of our database. The *Poor Quality* respiratory records have been provided in our database to facilitate the researchers to identify the respiratory sound with poor signal quality. We have investigated the effects of different feature extraction methods and classifiers and provided classification results for each respiratory sound classification task. We hope that the open-access respiratory sound database could inspire the development of medical device equipped with automatic respiratory sound classification algorithm to improve the safety of children with respiratory disease.

References

[1] A. Marques, A. Oliveira, and C. Jácome, “Computerized adventitious respiratory sounds as outcome measures for respiratory therapy: a systematic review,” *Respiratory Care*, vol. 59, no. 5, pp. 765–776, 2014.

[2] R. X. A. Pramono, S. Bowyer, and E. Rodriguez-Villegas, “Automatic adventitious respiratory sound analysis: A systematic review,” *PloS One*, vol. 12, no. 5, p. e0177926, 2017.

[3] N. Sengupta, M. Sahidullah, and G. Saha, “Lung sound classification using cepstral-based statistical features,” *Computers in Biology and Medicine*, vol. 75, pp. 118–129, 2016.

[4] A. Gurung, C. G. Scrafford *et al.*, “Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: A systematic review and meta-analysis,” *Respiratory Medicine*, vol. 105, no. 9, pp. 1396–1403, 2011.

[5] L. C. Puder, S. Wilitzki *et al.*, “Computerized wheeze detection in young infants: comparison of signals from tracheal and chest wall sensors,” *Physiological mMasurement*, vol. 37, no. 12, p. 2170, 2016.

[6] J. Zhang, H.-S. Wang *et al.*, “Real-world verification of artificial intelligence algorithm-assisted auscultation of breath sounds in children,” *Frontiers in Pediatrics*, vol. 9, p. 627337, 2021.

[7] A. Kandaswamy, C. S. Kumar *et al.*, “Neural classification of lung sounds using wavelet coefficients,” *Computers in Biology and Medicine*, vol. 34, no. 6, pp. 523–537, 2004.

[8] R. Palaniappan and K. Sundaraj, “Respiratory sound classification using cepstral features and support vector machine,” in *Proc. IEEE Recent Advances in Intelligent Computational Systems*, 2013, pp. 132–136.

[9] M. Lozano, J. A. Fiz, and R. Jané, “Automatic Differentiation of Normal and Continuous Adventitious Respiratory Sounds Using Ensemble Empirical Mode Decomposition and Instantaneous Frequency,” *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 2, pp. 486–497, 2016.

[10] P. Bokov, B. Mahut *et al.*, “Wheezing recognition algorithm using recordings of respiratory sounds at the mouth in a pediatric population,” *Computers in Biology and Medicine*, vol. 70, pp. 40–50, 2016.

[11] S. İçer and Şerife Gengeç, “Classification and analysis of non-stationary characteristics of crackle and rhonchus lung adventitious sounds,” *Digital Signal Processing*, vol. 28, pp. 18–27, 2014.

[12] F. Jin, F. Sattar, and D. Goh, “New approaches for spectro-temporal feature extraction with applications to respiratory sound classification,” *Neurocomputing*, vol. 123, pp. 362–371, 2014.

[13] R. Palaniappan and K. Sundaraj, “Respiratory sound classification using cepstral features and support vector machine,” in *Proc. IEEE Recent Advances in Intelligent Computational Systems*, 2013, pp. 132–136.

[14] D. Emmanouilidou, K. Patil *et al.*, “A multiresolution analysis for detection of abnormal lung sounds,” in *Proc. IEEE Engineering in Medicine and Biology Society*, 2012, pp. 3139–3142.

[15] C.-H. Chen, W.-T. Huang *et al.*, “Using k-nearest neighbor classification to diagnose abnormal lung sounds,” *Sensors*, vol. 15, no. 6, pp. 13 132–13 158, 2015.

[16] S. Ulukaya, I. Sen, and Y. P. Kahya, “Feature extraction using time-frequency analysis for monophonic-polyphonic wheeze discrimination,” in *Proc. IEEE Engineering in Medicine and Biology Society*, 2015, pp. 5412–5415.

[17] F. Jin, S. Krishnan, and F. Sattar, “Adventitious Sounds Identification and Extraction Using Temporal–Spectral Dominance-Based Features,” *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 11, pp. 3078–3087, 2011.

[18] S. Alsmadi and Y. P. Kahya, “Design of a DSP-based instrument for real-time classification of pulmonary sounds,” *Computers in Biology and Medicine*, vol. 38, no. 1, pp. 53–61, 2008.

[19] R. Naves, B. H. Barbosa, and D. D. Ferreira, “Classification of lung sounds using higher-order statistics: A divide-and-conquer approach,” *Computer Methods and Programs in Biomedicine*, vol. 129, pp. 12–20, 2016.

[20] L. Mendes, P. Carvalho *et al.*, “Robust features for detection of crackles: an exploratory study,” in *Proc. IEEE Engineering in Medicine and Biology Society*, 2014, pp. 1473–1476.

[21] L. Mendes, I. Vogiatzis *et al.*, “Detection of wheezes using their signature in the spectrogram space and musical features,” in *Proc. IEEE Engineering in Medicine and Biology Society*, 2015, pp. 5581–5584.

[22] L. Mendes, I. M. Vogiatzis *et al.*, "Detection of crackle events using a multi-feature approach," in *Proc. IEEE Engineering in Medicine and Biology Society*, 2016, pp. 3679–3683.

[23] D. Oletic, B. Arsenali, and V. Bilas, "Low-Power Wearable Respiratory Sound Sensing," *Sensors*, vol. 14, no. 4, pp. 6535–6566, 2014.

[24] G. Chambres, P. Hanna, and M. Desainte-Catherine, "Automatic detection of patient with respiratory diseases using lung sound analysis," in *Proc. International Conference on Content-Based Multimedia Indexing*, 2018, pp. 1–6.

[25] H. Mukherjee, H. Salam, and K. Santosh, "Lung health analysis: adventitious respiratory sound classification using filterbank energies," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 35, no. 14, p. 2157008, 2021.

[26] M. A. Tocchetto, A. S. Bazanella *et al.*, "An embedded classifier of lung sounds based on the wavelet packet transform and ANN," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 2975–2980, 2014.

[27] D. Chamberlain, R. Kodgule *et al.*, "Application of semi-supervised deep learning to lung sound analysis," in *Proc. International Conference of the IEEE Engineering in Medicine and Biology Society*, 2016, pp. 804–807.

[28] B.-S. Lin and B.-S. Lin, "Automatic wheezing detection using speech recognition technique," *Journal of Medical and Biological Engineering*, vol. 36, no. 4, pp. 545–554, 2016.

[29] Y. Shi, G. Wang *et al.*, "Classification of sputum sounds using artificial neural network and wavelet transform," *International Journal of Biological Sciences*, vol. 14, no. 8, p. 938, 2018.

[30] M. Markandeya, U. R. Abeyratne *et al.*, "Severity Analysis of Upper Airway Obstructions: Oesophageal Pressure Versus Snoring Sounds," in *Proc. IEEE Biomedical Circuits and Systems Conference*, 2019, pp. 1–4.

[31] Y. Ma, X. Xu *et al.*, "LungBRN: A smart digital stethoscope for detecting respiratory disease using bi-resnet deep learning algorithm," in *Proc. IEEE Biomedical Circuits and Systems Conference*, 2019, pp. 1–4.

[32] Ma, Yi, X. Xu *et al.*, "Live Demo: LungSys-Automatic Digital Stethoscope System For Adventitious Respiratory Sound Detection," in *Proc. IEEE Biomedical Circuits and Systems Conference*, 2019, pp. 1–1.

[33] Y. Ma, X. Xu, and Y. Li, "LungRN+ NL: An Improved Adventitious Lung Sound Classification Using Non-Local Block ResNet Neural Network with Mixup Data Augmentation," in *Interspeech*, 2020, pp. 2902–2906.

[34] J. Acharya and A. Basu, "Deep Neural Network for Respiratory Sound Classification in Wearable Devices Enabled by Patient Specific Model Tuning," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 3, pp. 535–544, 2020.

[35] N. Baghel, V. Nangia, and M. K. Dutta, "ALSD-Net: Automatic lung sounds diagnosis network from pulmonary signals," *Neural Computing and Applications*, vol. 33, no. 24, pp. 17 103–17 118, 2021.

[36] H. Zhu, J. Lai *et al.*, "Automatic pulmonary auscultation grading diagnosis of Coronavirus Disease 2019 in China with artificial intelligence algorithms: A cohort study," *Computer Methods and Programs in Biomedicine*, vol. 213, p. 106500, 2022.

[37] H. Zhu, J. Lai *et al.*, "Automatic pulmonary auscultation grading diagnosis of Coronavirus Disease 2019 in China with artificial intelligence algorithms: A cohort study," *Computer Methods and Programs in Biomedicine*, vol. 213, p. 106500, 2022.

[38] B. Liu, Z. Wen *et al.*, "Energy-Efficient Intelligent Pulmonary Auscultation for Post COVID-19 Era Wearable Monitoring Enabled by Two-Stage Hybrid Neural Network," in *Proc. IEEE International Symposium on Computer-Based Medical Systems*, 2022, pp. 1–5.

[39] A. Sovijarvi, J. Vanderschoot, and J. Earis, "Standardization of computerized respiratory sound analysis," *European Respiratory Review*, vol. 10, no. 77, pp. 585–585, 2000.

[40] B. M. Rocha, D. Filos *et al.*, "An open access database for the evaluation of respiratory sound classification algorithms," *Physiological Measurement*, vol. 40, no. 3, p. 035001, 2019.

[41] M. Fraiwan, L. Fraiwan *et al.*, "A dataset of lung sounds recorded from the chest wall using an electronic stethoscope," *Data in Brief*, vol. 35, p. 106913, 2021.

[42] F.-S. Hsu, S.-R. Huang *et al.*, "Benchmarking of eight recurrent neural network variants for breath phase and adventitious sound detection on a self-developed open-access lung sound database-HF_Lung_V1," *PLoS One*, vol. 16, no. 7, p. 0254134, 2021.

[43] J. F. Kanga and S. S. Kraman, "Comparison of the lung sound frequency spectra of infants and adults," *Pediatric Pulmonology*, vol. 2, no. 5, pp. 292–295, 1986.

[44] H. Pasterkamp, R. E. Powell, and I. Sanchez, "Lung sound spectra at standardized air flow in normal infants, children, and adults." *American Journal of Respiratory and Critical Care Medicine*, vol. 154, no. 2, pp. 424–430, 1996.

[45] E. Grooby, J. He *et al.*, "Neonatal Heart and Lung Sound Quality Assessment for Robust Heart and Breathing Rate Estimation for Telehealth Applications," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 12, pp. 4255–4266, 2021.

[46] T. Grzywalski, M. Piecuch *et al.*, "Practical implementation of artificial intelligence algorithms in pulmonary auscultation examination," *European Journal of Pediatrics*, vol. 178, no. 6, pp. 883–890, 2019.

[47] A. Bohadana, G. Izbicki, and S. S. Kraman, "Fundamentals of Lung Auscultation," *New England Journal of Medicine*, vol. 370, no. 8, pp. 744–751, 2014.

[48] H. Pasterkamp, P. L. Brand *et al.*, "Towards the standardisation of lung sound nomenclature," *European Respiratory Journal*, vol. 47, no. 3, pp. 724–732, 2016.

[49] M. Sarkar, I. Madabhavi *et al.*, "Auscultation of the respiratory system," *Annals of Thoracic Medicine*, vol. 10, no. 3, p. 158, 2015.

[50] R. Palaniappan, K. Sundaraj, and N. U. Ahamed, "Machine learning in lung sound analysis: A systematic review," *Biocybernetics and Biomedical Engineering*, vol. 33, no. 3, pp. 129–135, 2013.

[51] B. M. Rocha, D. Pessoa *et al.*, "Automatic Classification of Adventitious Respiratory Sounds: A (Un)Solved Problem?" *Sensors*, vol. 21, no. 1, 2021.

[52] M. Munakata, H. Ukita *et al.*, "Spectral and waveform characteristics of fine and coarse crackles," *Thorax*, vol. 46, no. 9, pp. 651–657, 1991.

[53] Emmanouilidou, Dimitra and McCollum, Eric D and Park, Daniel E and Elhilali, Mounya, "Computerized lung sound screening for pediatric auscultation in noisy field environments," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 7, pp. 1564–1574, 2017.

[54] K. K. Guntupalli, P. M. Alapat *et al.*, "Validation of Automatic Wheeze Detection in Patients with Obstructed Airways and in Healthy Subjects," *Journal of Asthma*, vol. 45, no. 10, pp. 903–907, 2008.

[55] S. Picard, C. Chapdelaine *et al.*, "Ensuring dataset quality for machine learning certification," in *Proc. International Symposium on Software Reliability Engineering Workshops*, 2020, pp. 275–282.

[56] P. P. Shinde and S. Shah, "A Review of Machine Learning and Deep Learning Applications," in *Proc. International Conference on Computing Communication Control and Automation*, 2018, pp. 1–6.

[57] Sapey, E and Bayley, D and Ahmad, A and Newbold, P and Snell, N and Stockley, R A, "Inter-relationships between inflammatory markers in patients with stable COPD with bronchitis: intra-patient and inter-patient variability," *Thorax*, vol. 63, no. 6, pp. 493–499, 2008.

[58] J. F.-S. Lin, V. Joukov, and D. Kulić, "Classification-based segmentation for rehabilitation exercise monitoring," *Journal of Rehabilitation and Assistive Technologies Engineering*, vol. 5, p. 2055668318761523, 2018.

[59] A.-M. Šimundić, "Measures of diagnostic accuracy: basic definitions," *EJIFCC*, vol. 19, no. 4, p. 203–211, 2009.

[60] K. Pasupa and W. Sunhem, "A comparison between shallow and deep architecture classifiers on small dataset," in *Proc. International Conference on Information Technology and Electrical Engineering*, 2016, pp. 1–6.

[61] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 3, pp. 685–695, 2021.

[62] L. Pham, H. Phan *et al.*, "CNN-MoE Based Framework for Classification of Respiratory Anomalies and Lung Disease Detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 8, pp. 2938–2947, 2021.

[63] N. Jakovljević and T. Lončar-Turukalo, "Hidden markov model based respiratory sound classification," in *Proc. International Conference on Biomedical and Health Informatics*, 2017, pp. 39–43.

[64] J. Liu, Y. Yin *et al.*, "Bowel Sound Detection Based on MFCC Feature and LSTM Neural Network," in *Proc. IEEE Biomedical Circuits and Systems Conference*, 2018, pp. 1–4.

[65] X. H. Kok, S. Anas Imtiaz, and E. Rodriguez-Villegas, "A Novel Method for Automatic Identification of Respiratory Disease from Acoustic Recordings," in *Proc. International Conference of the IEEE Engineering in Medicine and Biology Society*, 2019, pp. 2589–2592.

[66] D. Perna and A. Tagarelli, "Deep auscultation: Predicting respiratory anomalies and diseases via recurrent neural networks," in *Proc. IEEE International Symposium on Computer-Based Medical Systems*, 2019, pp. 50–55.

[67] S. Ntalampiras, "Collaborative framework for automatic classification of respiratory sounds," *IET Signal Processing*, vol. 14, no. 4, pp. 223–228, 2020.

[68] K. Kochetov and A. Filchenkov, "Generative Adversarial Networks for Respiratory Sound Augmentation," in *Proc. International Conference on Control, Robotics and Intelligent System*, 2020, pp. 106–111.

[69] B. McFee, C. Raffel *et al.*, "librosa: Audio and music signal analysis in python," in *Proc. Python in Science Conference*, 2015, pp. 18–25.

[70] W. Song, J. Han, and H. Song, "Contrastive Embeddind Learning Method for Respiratory Sound Classification," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2021, pp. 1275–1279.

[71] J. Li, C. Wang *et al.*, "Explainable CNN With Fuzzy Tree Regularization for Respiratory Sound Analysis," *IEEE Transactions on Fuzzy Systems*, vol. 30, no. 6, pp. 1516–1528, 2022.

[72] T. Nguyen and F. Pernkopf, "Lung Sound Classification Using Snapshot Ensemble of Convolutional Neural Networks," in *Proc. IEEE Engineering in Medicine and Biology Society*, 2020, pp. 760–763.

[73] M. T. García-Ordás, J. A. Benítez-Andrades *et al.*, "Detecting Respiratory Pathologies Using Convolutional Neural Networks and Variational Autoencoders for Unbalancing Data," *Sensors*, vol. 20, no. 4, 2020.

[74] J. Amoh and K. Odame, "Deep Neural Networks for Identifying Cough Sounds," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 5, pp. 1003–1011, 2016.

[75] J. Amoh and K. Odame, "DeepCough: A deep convolutional neural network in a wearable cough detection system," in *Proc. IEEE Biomedical Circuits and Systems Conference*, 2015, pp. 1–4.

[76] G. Serbes, S. Ulukaya, and Y. P. Kahya, "An Automated Lung Sound Preprocessing and Classification System Based OnSpectral Analysis Methods," in *Proc. International Conference on Biomedical and Health Informatics*, 2017, pp. 45–49.

[77] F. Pedregosa, G. Varoquaux *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[78] V. Tiwari, "MFCC and its applications in speaker recognition," *International Journal on Emerging Technologies*, vol. 1, no. 1, pp. 19–22, 2010.

[79] I. Rish *et al.*, "An empirical study of the naive Bayes classifier," in *Proc. IJCAI 2001 workshop on Empirical Methods in Artificial Intelligence*, vol. 3, no. 22, 2001, pp. 41–46.

[80] Z. Li and X. Wu, "Wheeze detection using fractional Hilbert transform in the time domain," in *Proc. IEEE Biomedical Circuits and Systems Conference*, 2012, pp. 316–319.

[81] A. Rao and K. Odame, "Estimating the short-time bandwidth of wheeze sounds," in *Proc. IEEE Biomedical Circuits and Systems Conference*, 2015, pp. 1–4.

[82] J. Xu, M. Konijnenburg *et al.*, "A 36 μW 1.1 mm 2 Reconfigurable Analog Front-end for Cardiovascular and Respiratory Signals Recording," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 4, pp. 774–783, 2018.

[83] X. Tang, Q. Hu, and W. Tang, "A Real-time QRS Detection System with PR/RT Interval and ST Segment Measurements for Wearable ECG Sensors using Parallel Delta Modulators," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 4, pp. 751–761, 2018.

[84] S.-Y. Lee, P.-W. Huang *et al.*, "Electrocardiogram and Phonocardiogram Monitoring System for Cardiac Auscultation," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 6, pp. 1471–1482, 2019.

[85] K. Zhao, H. Jiang *et al.*, "Long-Term Bowel Sound Monitoring and Segmentation by Wearable Devices and Convolutional Neural Networks," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 5, pp. 985–996, 2020.

[86] A. Ali, S. M. Shamsuddin, and A. L. Ralescu, "Classification with class imbalance problem," *International Journal of Advances in Soft Computing and its Applications*, vol. 5, no. 3, 2013.

[87] C.-Y. Wang, Y.-H. Yang *et al.*, "A wireless pulse oximetry system with active noise cancellation of motion artifacts," in *Proc. IEEE Biomedical Circuits and Systems Conference*, 2012, pp. 148–151.

[88] M. Maslik, T. S. Lande, and T. G. Constandinou, "A Clockless Method of Flicker Noise Suppression in Continuous-Time Acquisition of Biosignals," in *Proc. IEEE Biomedical Circuits and Systems Conference*, 2018, pp. 1–4.

[89] A. M. Mekala and S. Chandrasekaran, "Heart sound interference cancellation from lung sound using dynamic neighbourhood learning-particle swarm optimiser based optimal recursive least square algorithm," *International Journal of Biomedical Engineering and Technology*, vol. 34, no. 2, pp. 133–151, 2020.

[90] A. Manzoor, Q. Pan *et al.*, "Analysis and Detection of Lung Sounds Anomalies Based on NMA-RNN," in *Proc. IEEE International Conference on Bioinformatics and Biomedicine*, 2020, pp. 2498–2504.