Physiological Measurement

An open access database for the evaluation of respiratory sound classification algorithms

To cite this article before publication: Bruno M Rocha *et al* 2019 *Physiol. Meas.* in press https://doi.org/10.1088/1361-6579/ab03ea

An Open Access Database for the Evaluation of Respiratory Sound Classification Algorithms

Bruno M. Rocha[1], Dimitris Filos[2], Luís Mendes[3,1], Gorkem Serbes[4], Sezer Ulukaya[5,6], Yasemin P. Kahya[6], Nikša Jakovljević[7], Tatjana L. Turukalo[7], Ioannis M. Vogiatzis[2], Eleni Perantoni[2], Evangelos Kaimakamis[2], Pantelis Natsiavas[2], Ana Oliveira[8], Cristina Jácome[8], Alda Marques[8], Nicos Maglaveras[2,9], Rui Pedro Paiva[1], Ioanna Chouvarda[2], Paulo de Carvalho[1]

Abstract

Background and Objectives: During the last decades, there has been a significant interest in the automatic analysis of respiratory sounds. However, currently there are no publicly available large databases with which new algorithms can be evaluated and compared. Further developments in the field are dependent on the creation of such databases. *Methods*: This paper describes a public respiratory sound database, compiled for an international competition, the first scientific challenge of the IFMBE's International Conference on Biomedical and Health Informatics. The database includes 920 recordings acquired from 126 participants and two sets of annotations. One set contains 6898 annotated respiratory cycles, some including crackles, wheezes, or a combination of both, and some with no adventitious respiratory sounds. In the other set, precise locations of 10775 events of crackles and wheezes were annotated. *Results*: The best system that participated in the challenge achieved an average score of 52.5% with the respiratory cycle annotations and an average score of 91.2% with the event annotations. *Conclusion*: The creation and public release of this

[1]Centre for Informatics and Systems (CISUC), Department of Informatics Engineering, University of Coimbra, Portugal
[2]Lab of Computing, Medical Informatics, and Biomedical Imaging Technologies, Aristotle University of Thessaloniki, Greece
[3]AIBILI - Association for Innovation and Biomedical Research on Light, Coimbra, Portugal
[4]Department of Biomedical Engineering, Yildiz Technical University, 34220, Istanbul, Turkey
[5]Department of Electrical and Electronics Engineering, Trakya University, 22030, Edirne, Turkey
[6]Department of Electrical and Electronics Engineering, Boğaziçi University, 34342, Istanbul, Turkey
[7]Department of Power, Electronics, and Communication Engineering, Faculty of Technical Sciences, University of Novi Sad, Serbia
[8]Respiratory Research and Rehabilitation Laboratory (Lab3R), School of Health Sciences (ESSUA) and Institute for Biomedicine (iBiMED), University of Aveiro, Portugal
[9]Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, USA

database will be useful to the research community and can bring attention to the respiratory sound classification problem.

---

1. Introduction

Respiratory conditions are the third leading cause of death worldwide [1] and generate a significant burden for public health systems [2]. Because early diagnosis and routine monitoring of patients with respiratory conditions are important for timely interventions, considerable research endeavours have been focused on these issues in the last couple of decades [3].

Respiratory conditions and their progress are typically assessed through spirometry and lung auscultation, the former measuring the volume of air mobilised in respiration and the latter assessing airflow through the tracheabronchial tree, via the sounds produced. Spirometry is the most useful and commonly available test of lung function. This simple lung function measurement, which has well established normal values, is effective and well validated for the diagnosis and monitoring of upper and lower airway abnormalities [4]. However, one major limitation of spirometry is its dependence on patient's motivation and cooperation [3], with high potential for error if effort is suboptimal. Additionally, traditional spirometers are typically only used in some clinical settings due to their high cost, challenges of patient guiding, and required calibration [5].

Concurrently, the existence of respiratory conditions may be assessed through the auscultation of respiratory sounds. The stethoscope is the main tool for lung auscultation in clinical practice. Auscultation is typically performed on the anterior and posterior chest [6]. The expert clinician is trained to listen to and recognise the pathologic findings, such as the presence of adventitious sounds (e.g., crackles, wheezes). Although auscultation devices have evolved from analog to digital, thus enabling storing, analysis, and visualisation in computer systems, digital auscultation is not yet a mature and fully computational procedure. Conventional auscultation has some drawbacks that limit its use in research due to: 1) impossibility of providing continuous monitoring; 2) the need of having an expert to detect presence/absence and clinical meaning of normal/abnormal sounds [7]; 3) its inherent inter-listener variability [8]; 4) human audition and memory limitations [9]. These drawbacks hinder the effectiveness of conventional auscultation as a way of monitoring and managing respiratory conditions. Automated respiratory sound analysis could potentially overcome these limitations.

This work aimed to leverage digital auscultation in terms of data and algorithms availability for diagnostic purposes. The stimuli to accomplish this goal was to a large extent the need for inclusion of respiration sound sensors in wearable technology for chronic obstructive pulmonary disease (COPD) monitoring, raised by the FP7 project WELCOME (Wearable Sensing and Smart Cloud Computing for Integrated Care to COPD Patients with Comorbidities) [10]. Such an ambition for recording respiratory sounds in daily life via wearable sensors would also require their automated analysis in a cloud computing environment, and the automatic detection of adventitious respiratory sounds for inclusion of such findings in a decision support system for clinician support. The effort presented in this paper, the generation of a respiratory sounds database and its initial testing via a scientific challenge, was considered a necessary investment towards introducing digital auscultation as a connected health technology, for its wide adoption from hospital to primary care and home care settings.

2. Background

Respiratory sounds are relevant indicators of respiratory health and respiratory conditions, as they are directly related to movement of air, changes within the lung tissue, and position of secretions within the tracheobronchial tree [11].

Typically, respiratory sounds are classified as normal or adventitious. Adventitious respiratory sounds are superimposed on normal respiratory sounds and can be discontinuous (crackles) or continuous (wheezes) [12].

Crackles are discontinuous, explosive, and non-musical adventitious respiratory sounds that occur frequently in cardiorespiratory conditions [13]. They are usually classified as fine and coarse crackles based on their duration, loudness, pitch, timing in the respiratory cycle (i.e., inspiration or expiration), and relationship to coughing and changing body position [14]. Appearance of crackles may be an early sign of respiratory disease [15]. Number of crackles per breath is associated with the severity of the disease in patients with interstitial lung conditions [16]. Moreover, the waveform and timing of crackles may have clinical significance in differential diagnosis of cardiorespiratory conditions [13]. When present, crackling sounds in patients with lung fibrosis are typically fine, repetitive, and end inspiratory, whereas those associated with chronic airways obstruction (e.g., COPD, emphysema or bronchiectasis) are coarse, less repeatable, and occur early in inspiration [17].

Wheezes are musical respiratory sounds that usually last more than 100 ms [13]. They are a common clinical sign in patients with obstructive airway conditions, such as asthma and COPD [15]. Marini et al. [18] have demonstrated that there is an association between the degree of bronchial obstruction and the presence and characteristics of wheezes. The strongest association has been obtained when the degree of bronchial obstruction is compared to the proportion of the respiratory cycle occupied by wheezing [19].

In the last decades, many researchers have developed methods for the automatic detection or classification of adventitious respiratory sounds [20–26]. Most systems comprise two steps: 1) relevant features are extracted from the signal; 2) extracted features are used to detect or classify adventitious respiratory sound events (i.e., crackles and wheezes). In developing a detection or classification algorithm, especially if machine learning techniques are used, it is important to take note of how the data are used to train, test, and validate the algorithm.

The most common features employed in the literature were identified in a recent systematic review [7]. These include Mel-frequency cepstral coefficients (MFCCs) [27], spectral features [28], entropy [29], and wavelet coefficients [30].

Table I: List of Publicly Available Respiratory Sound Databases

Auscultation Skills: Breath and Heart Sounds, 4th edition [39]
East Tennessee State University repository [40]
Fundamentals of Lung and Heart Sounds [41]
Heart and Lung Sounds Reference Library [42]
Littmann repository [43]
Lung Sounds: An Introduction to the Interpretation of the Auscultatory Finding [44]
R.A.L.E. repository [37]
Secrets Heart & Lung Sounds Workshops [45]
SoundCloud repository [46]
The Chest: Its Signs and Sounds [47]
Understanding Heart Sounds and Murmurs [48]
Understanding Lung Sounds, the 2nd edition [49]
Understanding Lung Sounds, the 3rd edition [38]

Machine learning algorithms proposed in the literature comprise empirical rule-based methods [31], support vector machines [32], artificial neural networks [33], Gaussian mixture models [34], k-nearest neighbours [35], and logistic regression models [36].

The small number of patients involved in most studies that proposed automatic classification of respiratory sounds have hindered their generalisability. Although it is possible to achieve very good classification results with small samples by customising the algorithm to fit the data, as the number of patients increases to dozens or hundreds, the features learned from small datasets typically fail to generalise [33].

A core problem in the field is the lack of publicly available large databases, which can serve to develop algorithms and compare results. Although most works used in-house data collections, there were 13 publicly available databases among the data sources of the 77 articles covered by the systematic review of Pramono et al. [7]. These databases are from 4 online repositories and 9 audio CD companion books, shown in Table I. The most referenced databases are the R.A.L.E. repository [37] and the audio CD from Understanding Lung Sounds 3rd edition [38]. As these repositories and CDs were designed for teaching, they generally include a small number of examples of each type of respiratory sound. Most of these sounds are clear and do not include environmental noise, commonly present in clinical practice, thus are not suitable for generating realistic classification models.

To overcome these difficulties, we compiled a database of respiratory sounds [50] and organised a scientific challenge at the International Conference on Biomedical and Health Informatics (ICBHI) 2017. Details of the database will be further discussed in the next section, but it is currently available to the research community (http://bhichallenge.med.auth.gr/). In the website, it is possible to download the audio files, both sets of annotations, and clinical/demographic information about the participants.

Table II: Demographic Information of Database (NA: Not Available)

Number of recordings	920
Sampling frequency (number of recordings)	4 kHz (90); 10 kHz (6); 44.1 kHz (824)
Bits per sample	16
Average recording duration	21.5 s
Number of participants	126: 77 adults, 49 children
Sex	79 male, 46 female (NA: 1)
Age (mean ± standard deviation)	43.0 ± 32.2 $years$ (NA: 1)
Age of adult participants	67.6 ± 11.6 $years$ (NA: 1)
Age of child participants	4.8 ± 4.6 $years$
BMI of adult participants	27.2 ± 5.4 kg/m^2 (NA: 2)
Weight of child participants	21.4 ± 17.2 kg (NA: 5)
Height of child participants	104.7 ± 30.8 cm (NA: 7)

3. Data Collection

The ICBHI Scientific Challenge database contains audio samples, collected independently by two research teams in two different countries, over several years. Ethical approval was obtained from the Ethics Committees of the appropriate institutions. The database contains 920 annotated audio samples from 126 participants and it is currently the largest publicly available database. Hence, it has the potential to be used as a benchmark in the field. The database is described in Table II (as the relationship between body-mass index (BMI) and obesity is not linear in children, we opted to instead publish height and weight values for child participants, i.e., younger than 19 years-old, as recommended by clinical investigation guidelines [51].

3.1. Respiratory Research and Rehabilitation Laboratory of the School of Health Sciences, University of Aveiro

Audio samples recorded by the research team of the Respiratory Research and Rehabilitation Laboratory (Lab3R) of the School of Health Sciences, University of Aveiro (ESSUA) constitute most of the database. Samples were recorded in Aveiro, Portugal, at ESSUA and Hospital Infante D. Pedro, in Porto, Portugal, at Hospital Santa Maria and Lusíadas, and at the Faculty of Health Sciences, University of Southampton, England. Sounds from five studies conducted by this research team were included in the database. All recordings followed the computerised respiratory sounds analysis guidelines for short-term acquisitions [52]. Sounds were recorded from the trachea and six chest locations: left and right anterior, posterior, and lateral. Sounds were collected in clinical and non-clinical (home) settings. The acquisition of respiratory sounds was performed on participants of all ages. Participants included patients with lower respiratory tract infections, upper respiratory tract infections, pneumonia, COPD, asthma, bronchiolitis, bronchiectasis, and cystic fibrosis.

In three studies, the sounds were collected sequentially with a digital stethoscope (Welch Allyn Meditron Master Elite Plus Stethoscope Model 5079-400). In the other studies, the sounds were collected simultaneously using either seven stethoscopes (3M Littmann Classic II SE) with a microphone in the main tube or seven air-coupled electret microphones (C 417 PP, AKG Acoustics) located into capsules made of teflon. Respiratory sounds were acquired using the Computerised Lung Auscultation – Sound System (CLASS) [53].

3.2. Aristotle University of Thessaloniki

Respiratory sounds were acquired by the research team of the Aristotle University of Thessaloniki (AUTH) at the Papanikolaou General Hospital, Thessaloniki and at the General Hospital of Imathia (Health Unit of Naousa), Greece. Sounds were collected sequentially from six chest locations, as shown in Fig. 1, with a digital stethoscope (WelchAllyn Meditron Master Elite Plus Stethoscope Model 5079-400 or 3M Litmmann 3200). During the recordings, the participants were seated and were asked to produce events of cough, speech, laughter, and throat clearing. The acquisition of respiratory sounds was performed on adult and elderly patients. All patients had COPD with comorbidities (e.g., heart failure, diabetes, hypertension). These recordings were acquired as part of the European project WELCOME [10].

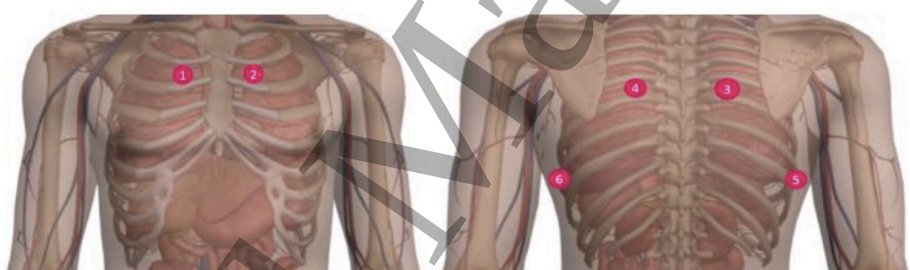

Figure 1: Chest locations for the recording of respiratory sounds.

4. Data Annotation

4.1. Lab3R

The most common method to evaluate the robustness of algorithms to detect adventitious respiratory sounds is the annotation of sounds by respiratory health professionals [54]. Two respiratory physiotherapists and one medical doctor, with experience in visual-auditory crackles/wheezes recognition, independently annotated the sound files in terms of presence/absence of adventitious sounds and identification of breathing phases. However, as annotation is a difficult and slow process, all the sound files in the Lab3R database were annotated by only one expert. The Respiratory Sound Annotation Software was used to annotate the sound files [55]. Fig. 2 reproduces a sample of the annotation process.

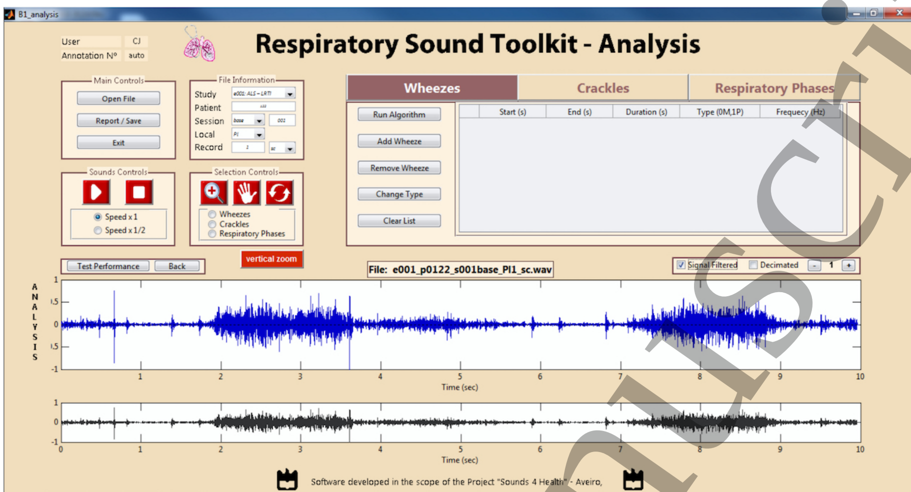

Figure 2: A sample of the respiratory sound annotation process (Lab3R).

4.2. AUTH

Three experienced physicians, two specialised pulmonologists and one cardiologist, annotated the sound files using Audacity 2.0.6 [56], a free, open source, cross-platform software for recording and editing sounds. The following sounds were discriminated in the annotation process: normal (respiratory sound), fine crackles, coarse crackles, wheezing, speech, cough, artefact. Fig. 3 reproduces a sample of the annotation process.

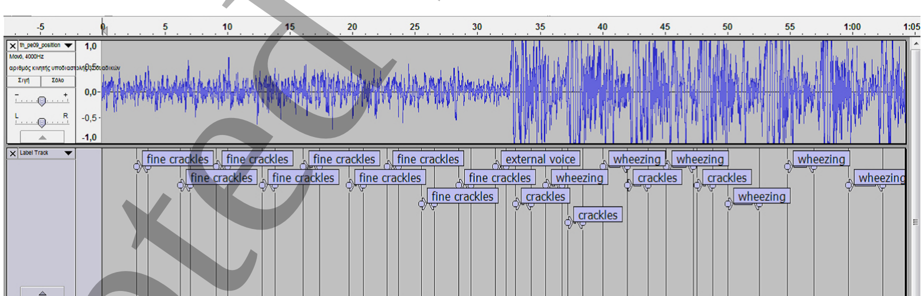

Figure 3: A sample of the respiratory sound annotation process (AUTH).

5. ICBHI Challenge

In the unofficial phase of the challenge, 18 systems were submitted from 5 international research teams. In the official phase, 3 of those teams uploaded a total of 11 entries. The 2 best teams presented their algorithms at ICBHI 2017. Below is an overview of their systems.

5.1. Serbes, Ulukaya, and Kahya (SUK) team

In this database, respiratory sounds have different sampling rates. Therefore, as a preprocessing step, each lung sound segment is resampled to 4000 Hz. Afterwards, a 12th order Butterworth band-pass filter having 120 and 1800 Hz cut-off frequencies is employed on raw lung sound signals to eliminate noise components that are due to coughing, intestinal sounds, stethoscope motion, speech, and heart sounds. Respiratory sounds have low and high frequency components and these are overlapped in both time and frequency domains. Therefore, a resonance based signal separation method, which uses non-dynamic Tunable Q-Factor Wavelet Transform [57], is employed to decompose respiratory sounds into low and high resonance, and residual channels. At the end of resonance based decomposition, it can be observed that crackle components are separated into low resonance, wheeze components are separated into high resonance, and noise components are separated into residual channels.

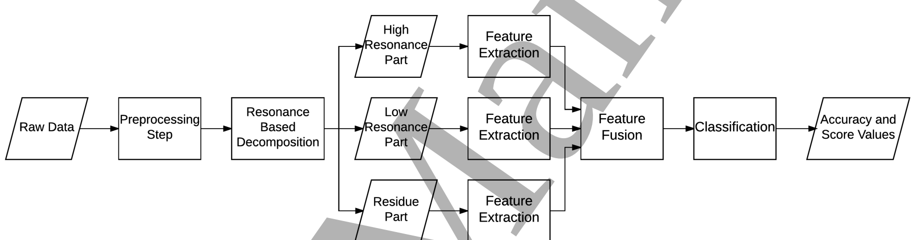

Figure 4: Step by step flowchart of SUK preprocessing and classification system.

When the duration of crackle and wheeze signals is considered, it is known that crackles are short (typically less than 20 ms) signals while wheezes are long (more than 100 ms) signals [12]. Therefore, in order to extract robust features, Short Time Fourier Transform (STFT) is applied to low/high resonance and residual components, and time-frequency distributions of separated channels are obtained. By doing that, the individual time-frequency behaviour of each class type (crackle, wheeze or normal classes) is obtained without the undesirable overlaps in the frequency domain. However, each crackle, wheeze, and normal signal segment has different durations and, to stabilize the learning algorithm, each output of STFT is integrated over time, resulting in the power distribution of separated signal components (high/low resonance and residual) over frequency. Moreover, as a fast adaptive time-scale representation based feature extraction method, Tunable Q-Factor Wavelet Transform [58] is also applied to low/high resonance and residual components resulting in wavelet coefficients. To decrease the number of features and address the curse of dimensionality, statistical (mean, standard deviation, kurtosis, minimum, maximum and skewness) and spectral (linear energy, Teager-Kaiser energy, Shannon entropy) features are derived from wavelet coefficients. At the end of the feature extraction step, to increase the discriminative power of the learning algorithm, feature level fusion is applied to STFT and wavelet based features.

In the learning step of the proposed study, to reduce the complexity of the learning model, Principal Component Analysis (PCA) is applied to the feature set, preserving 90% of the variance. Prior to feeding the classification algorithm, extracted features are normalized to [-1, +1] range. Finally, both STFT based spectral features and wavelet transform plus STFT based features are fed into a Support Vector Machine classifier with a grid search parameter optimization. A step by step flowchart of the proposed system is depicted in Fig. 4. A more detailed explanation of this system can be found in Serbes et al. [59].

5.2. Jakovljević and Lončar-Turukalo (JL) team

The audio recordings in this database varied in sampling frequency, recording locations, number of samples per class, and levels of different types of real noise. To overcome this diversity in the database, all recordings were resampled to 4000 Hz, preserving the relevant frequency range [60-2000 Hz] for the identification of wheezes and crackles [14]. The team investigated two schemes for removal of low frequencies and heart beat sounds: low order bandpass filter (denoted T1) and high order finite impulse response filter with $f_c = 100Hz$, and constant group delay $\tau_g = 1024 samples$ obtained by Hamming window function.

The stationary noise in audio files has been suppressed using spectral subtraction (SS) [60]. It is performed on the signal segmented into 30 ms long frames shifted by 15 ms using Hann window function. For a frame captured at a time instant t, discrete Fourier transform (DFT), $X(k,t)$, is obtained at each frequency bin k. The noise magnitude spectrum $|D(k)|$ is estimated as the mean value of $|X(k,t)|$ over 1% of the frames with minimum energy in the audio signal, excluding invalid (zero energy) frames. In the first approach, denoted SS1 (1), the denoised magnitude spectrum $|X_d(k,t)|$ is obtained by subtracting the magnitude spectra of stationary noise, setting the negative magnitude values to 1% of $|X(k,t)|$:

$$X_d(k,t) = \begin{cases} |X(k,t)| - |D(k)| & |X(k,t)| > |D(k)| \\ 0.01|X(k,t)| & \text{else} \end{cases} \tag{1}$$

The second approach, denoted SS2 (2), additionally reduces the level of musical noise introduced by magnitude spectrum subtraction. As that breath sound should be dominant in the signal, for each k the estimated noise level $|D(k)|$ has been iteratively reduced by 10%, until in at least 60% of frames $|X(k,t)| > |D(k)|$ is fulfilled. The denoised magnitude spectrum is obtained by:

$$X_d(k,t) = \begin{cases} |X(k,t)| - |D(k)| & |X(k,t)| > |D(k)| \\ |X(k,t)|^2 & \text{else} \end{cases} \tag{2}$$

To accommodate quadrature scaling, $|X(k,t)|$ has to be range normalised. To suppress sudden drops of magnitude, $|X(k,t)|$ is monitored in 5 successive frames. If $|X(k,t)| < |D(k)|$ in at least 3 of 5 adjacent frames, the frequency bin k is marked as noise. An entire frame is considered as corrupted by noise and set to zero ($|X_d(k,t)| = 0$, for each k) if more than 70% of the bins are marked as noise.

The synthesis step merges the obtained denoised magnitude spectrum with the original phase spectrum $argX(k,t)$ and the reconstructed signal is the sum of overlapping segments obtained by inverse DFT of $X_d(k,t)$.

The estimation of MFCCs on denoised signals is done every 10 ms using 30 ms long windows. 16 equal-width overlapped channels in mel-frequency domain divides the frequency range of interest [50, 2000 Hz]. The coefficient C_0, representing energy in the selected frequency band, is discarded as it significantly correlates with heart beat sound. The cepstral coefficients are z-normalised per record to remove variations caused by the remaining noise. We introduce first time derivatives of MFCCs to track their dynamics and to decorrelate successive feature vectors. The cardinality of the feature vectors, which include both static and dynamic MFCCs, is $d = 30$.

A respiration cycle for each sound class at each location is represented as a left-to-right HMM with S states (Fig. (5)), because the acoustic content of the same sound class varies depending on the recording location. A full HMM description contains: initial state probabilities ($\mathbf{\Pi}$), state transition matrix ($\mathbf{A}$), and emitting probability density function for each state (b_s). A state emitting probability density function (pdf), for a given d-dimensional observation $\mathbf{o}$, $b_s(\mathbf{o})$, is defined as the weighted mixture of M Gaussians characterised by their mean and covariance matrix. As commonly done, the same number of mixture components is used for each state.

The HMM parameters are estimated maximising the likelihood that the models will generate the training sequence using expectation maximisation algorithm (Baum-Welch estimation) [61]. The initial parameters were obtained by the time equidistant partition of the observation sequence between states, as the sample mean $\boldsymbol{\mu}_s$ and the covariance matrix $\mathbf{\Sigma}_s$. In case of several mixture components per state, means ($\boldsymbol{\mu}_i$) were obtained by random sampling from normal distribution $\mathcal{N}(\boldsymbol{\mu}_s, \mathbf{\Sigma}_s)$, and covariance matrices ($\mathbf{\Sigma}_i$) as the corresponding sample covariance matrix ($\mathbf{\Sigma}_i = \mathbf{\Sigma}_s$). The transition probabilities (a_{ij}) were initialised to 0.5, except for a_{SS}, initialised to 1. With these settings, the algorithm converged in 6-12 iterations.

During the test phase, each HMM (λ_c), $c = 1 \ldots, 28$ is aligned with an unknown observation sequence and the classification decision is based on the maximum likelihood criterion.

Additionally, an ensemble of classifiers trained over 10 different folds was evaluated. All classifiers had the same model complexity (28 models with 5 states and 1 Gaussian per state) and were trained with a single learning method. The majority voting was used as a decision scheme. It should be noted that computational complexity of this approach is 10 times greater. A more detailed explanation of this system can be found in Jakovljević and Lončar-Turukalo [62].

5.3. Baseline System

To account for the different sampling rates, we first resample each audio recording to 4000 Hz. Then, using the MIR Toolbox [63], we extract 13 MFCCs in 10 ms frames with 5 ms overlap (due to the low stationarity of the signal) and

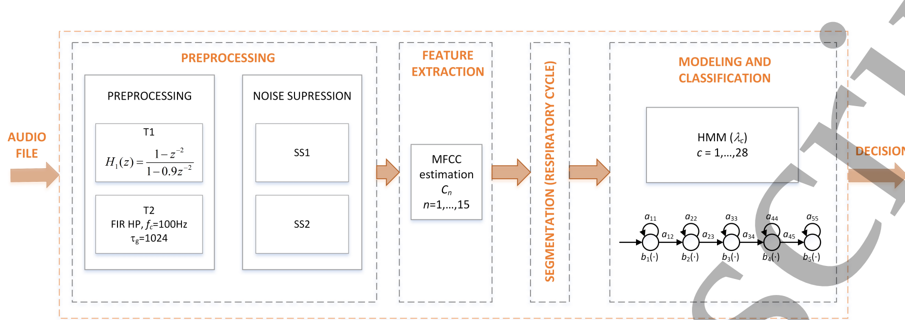

Figure 5: Step by step flowchart of JL preprocessing and classification system.

Table III: Determination Rules

		Entry's output			
		Crackles	Wheezes	Crackles+Wheezes	Normal
Reference label	Crackles (C)	C_c	C_w	C_b	C_n
	Wheezes (W)	W_c	W_w	W_b	W_n
	Crackles+Wheezes (B)	B_c	B_w	B_b	B_n
	Normal (N)	N_c	N_w	N_b	N_n

compute their mean for each respiratory cycle (event, in the case of Experiment 2, as discussed below). Finally, we employ a pruned decision tree to classify each respiratory cycle/event.

5.4. Evaluation Metrics

For each entry, we computed two performance measures: average score (AS) and harmonic score (HS). AS (5) is the average of sensitivity (SE) (3) and specificity (SP) (4), while HS (6) is the harmonic mean of SE and SP. Table III shows the determination rules employed to calculate SE and SP.

$$SE = (C_c + W_w + B_b)/(C + W + B) \tag{3}$$

$$SP = N_n/N \tag{4}$$

$$AS = (SE + SP)/2 \tag{5}$$

$$HS = (2 * SE * SP)/(SE + SP) \tag{6}$$

6. Experiment 1: ICBHI Scientific Challenge (Annotation of Respiratory Cycles)

6.1. Data Preparation for the Challenge

The challenge was structured in two phases: unofficial and official. During each phase, data from the two aforementioned databases were divided into training

(60%) and testing (40%) sets. The ground-truth annotations comprised four classes of respiratory cycles: containing crackles, containing wheezes, containing both, or normal (i.e., not exhibiting crackles or wheezes). A respiratory cycle is the sequence of events during which a human being inhales (inspiration) and exhales (expiration) a given volume of air through the respiratory system [64]. Fig. 6 shows an example of an annotated sound recording.

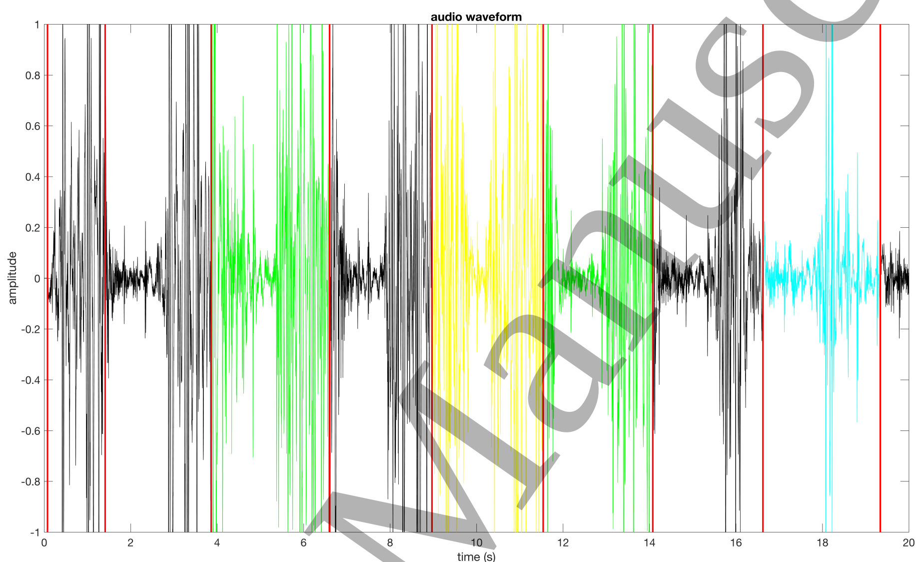

Figure 6: A segment including eight respiratory cycles: the third and sixth contain crackles (green), the eighth contains wheezes (blue), the fifth contain both crackles and wheezes (yellow), and the others are normal (black). Respiratory cycle boundaries are represented by vertical lines (red).

During the official phase of the challenge, 2063 respiratory cycles from 539 recordings derived from 79 participants were included in the training set, while 1579 respiration cycles from 381 recordings derived from 49 patients were included in the testing set. Additional details about the distribution of the adventitious respiratory sounds between training and testing partitions can be found in Table IV. As depicted, effort was paid to represent in analogy the different types of findings in training and testing data sets. Table V presents the demographic information distributed by training and testing sets.

6.2. Aggregated Results

This section presents the best results achieved by the two finalists of the challenge, Serbes, Ulukaya, and Kahya (SUK) team and Jakovljević and Lončar-Turukalo (JL) team. During the challenge, the size of the algorithms each team could send was restricted. After the challenge, each finalist sent a new algorithm to be evaluated, unrestricted in size. All the submitted algorithms used the same sound recordings for training. Fig. 7 shows the scores for the new (JL new and

Table IV: Summary of the Training and Testing Sets Used in Experiment 1

	Training Set			Testing Set		
	Lab3R	AUTH	All	Lab3R	AUTH	All
participants	72	7	79	38	11	49
recordings	507	32	539	317	64	381
crackles	1104	111	1215	588	61	649
wheezes	459	42	501	273	112	385
crackles+wheezes	335	28	363	106	37	143
normal	1740	323	2063	1216	363	1579

Table V: Demographic Information of Training and Testing Sets (NA: Not Available)

	Training Set	Testing Set
Number of participants	51 adults, 28 children	28 adults, 21 children
Sex	47 male, 32 female	34 male, 14 female (NA: 1)
Age	45.9 ± 31.6 *years*	39.6 ± 33.2 *years* (NA: 1)
Age of adults	67.8 ± 12.6 *years*	67.9 ± 9.4 *years* (NA: 1)
Age of children	6.0 ± 5.2 *years*	3.2 ± 3.2 *years*
BMI of adults	27.0 ± 6.0 kg/m^2 (NA: 1)	27.4 ± 3.9 kg/m^2 (NA: 1)
Weight of children	25.8 ± 21.5 kg (NA: 5)	16.5 ± 8.7 kg
Height of children	112.7 ± 34.4 cm (NA: 6)	95.8 ± 24.1 cm (NA: 1)
Respiratory cycle duration	2.73 ± 1.21 s	2.65 ± 1.11 s
Wheeze duration	0.63 ± 0.84 s	0.56 ± 0.76 s
Crackle duration	0.05 ± 0.19 s	0.05 ± 0.15 s

SUK new) and the best old (JL old and SUK old) systems submitted by each finalist, as well as the scores attained by the baseline model.

First, we note that the scores are much lower than those reported in the literature. As will be shown later, the length of the annotated cycles might be relevant to explain these results. Then, when considering Harmonic Score (HS), we see that both teams outperform the baseline. Regarding Average Score (AS), only SUK's systems surpass the baseline. This difference is due to great discrepancies between the systems' specificities and sensitivities.

Fig. 8 shows specificity (SP) and sensitivity (SE) values for the same systems. We see that the baseline SP is much higher than SE, meaning the baseline system classified most cycles as 'normal'. Only one system outperforms the baseline in both measures: SUK new.

Fig. 9 shows SE values for each class except 'normal'. Baseline SE is only shown for 'crackles', as it is 0 for both 'wheezes' and 'crackles + wheezes'. This means that the baseline system only considered the two most represented classes in the database, 'normal' and 'crackles', ignoring the other two. JL new was the system that achieved higher SE in each class, at the cost of reduced SP.

6.3. Results per Participant

In this section, we analyse the results for each participant. However, we have to acknowledge some limitations of this analysis. First, the number of cycles per

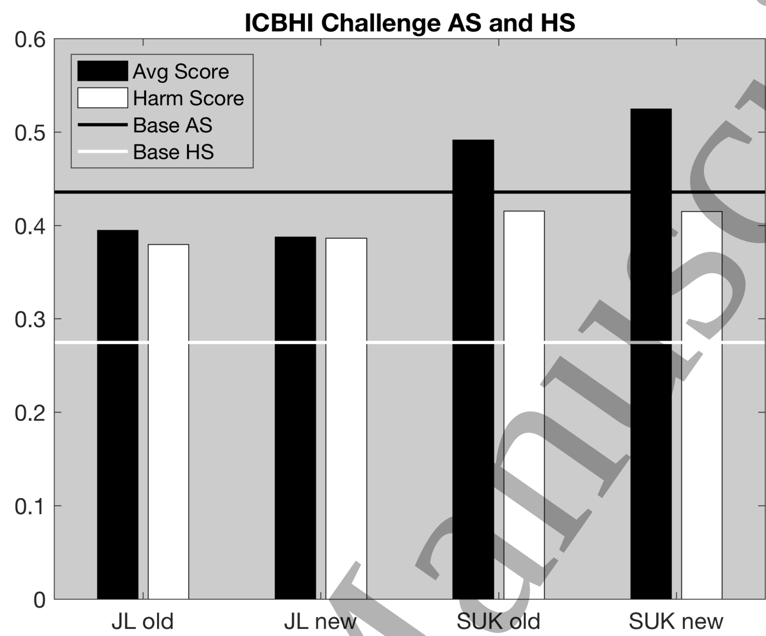

Figure 7: Average Score and Harmonic Score for the respiratory cycle annotations.

Table VI: Results per Participant (Experiment 1)

	AS		HS		SP		SE	
	M	SD	M	SD	M	SD	M	SD
JL old	40%	13%	22%	20%	50%	34%	31%	28%
JL new	39%	12%	23%	19%	38%	28%	41%	29%
SUK old	47%	12%	30%	20%	71%	27%	29%	25%
SUK new	47%	11%	24%	22%	78%	21%	20%	21%
Baseline	43%	8%	15%	16%	75%	23%	12%	16%

participant is highly variable. The testing set contains 49 participants with an average of 56 cycles, standard deviation of 50, maximum of 208, and minimum of 7. Furthermore, the respiratory cycles of 16 participants are all normal, i.e., do not contain adventitious respiratory sounds. Therefore, the AS, HS, and SE values do not take into account these participants. Besides, even when the number of recorded cycles for a particular participant is high, the number of cycles containing adventitious sounds could be as low as 1, hence skewing the results. Table VI presents the mean (M) and standard deviation (SD) of the evaluation metrics for each system.

As in the aggregated results, SUK's systems outperform the baseline and JL's

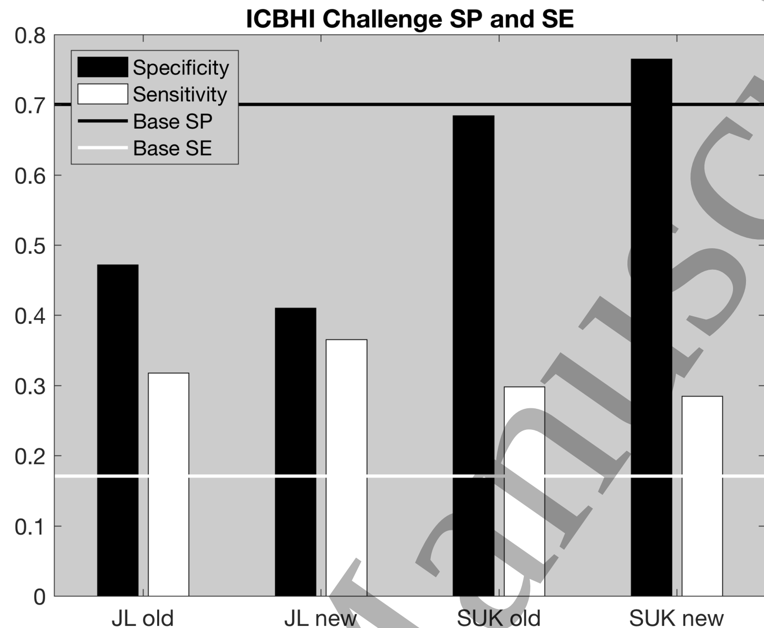

Figure 8: Specificity and sensitivity values for the respiratory cycle annotations.

systems in AS and HS. The most interesting aspect of these results to mention is the high standard deviation, as it provides evidence of the low reliability of these systems in this dataset.

7. Experiment 2: Post-Challenge (Annotation of Adventitious Respiratory Sound Events)

7.1. Data Preparation

In this experiment, even though the sound recordings were the same, the ground-truth comprised annotations of individual events of wheezes and crackles. The average duration of these events can be seen in V. For this reason, the number of classes was reduced to three: crackles, wheezes, and normal. Fig. 10 shows the annotated crackles and wheezes superimposed on the same sound recording of Fig. 6. Fig. 11 depicts the spectrogram of the same recording. The spectrogram is focused on the low frequencies (100 to 500 Hz). Although adventitious respiratory sounds are typically associated with particular signatures in the spectrogram, it is not a trivial task to annotate these events by solely inspecting the spectrogram. As the annotations correspond to events, not respiratory cycles, the events for the 'normal' class were created using a custom script. Approximately half of them are 50 ms and the other half are 150 ms events. Their positions in the audio

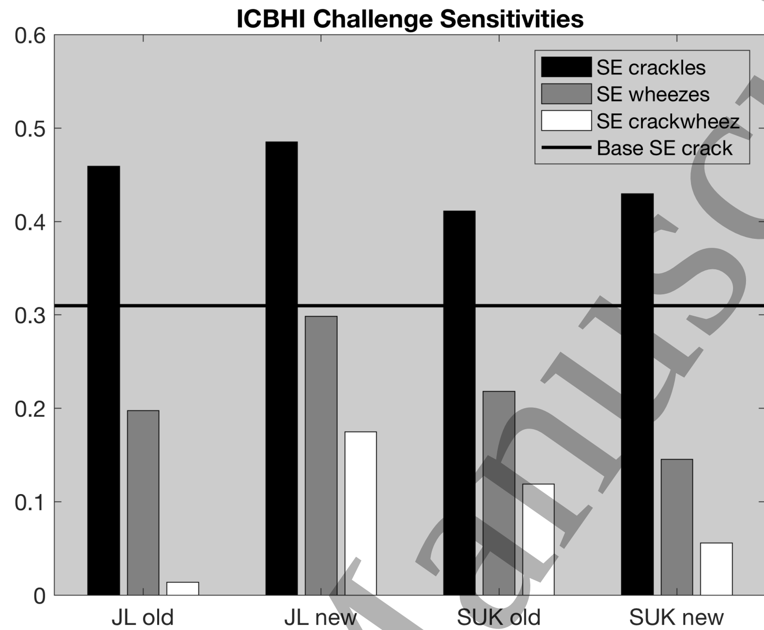

Figure 9: Sensitivity values for each class in the respiratory cycle annotations.

Table VII: Summary of the Training and Testing Sets Used in Experiment 2

	Training Set			Testing Set		
	Lab3R	AUTH	All	Lab3R	AUTH	All
participants	72	7	79	38	11	49
recordings	507	32	539	317	64	381
crackles	5808	188	5996	2760	121	2881
wheezes	1103	70	1173	518	207	725
normal	2348	665	3013	1597	415	2012

files are randomised in a unique fashion for each file, with every file containing at least one event per 2 s. Table VII provides details about the number of events in each class.

7.2. Aggregated Results

This section presents the results achieved by each finalist with the event annotations, as well as the results obtained by the baseline model. The submissions (JL new and SUK new) were adapted to the reduction in the number of classes from 4 to 3. Because some crackle events contained less than 5 frames and correlation between the frames was high, the number of HMM states in JL new had to be

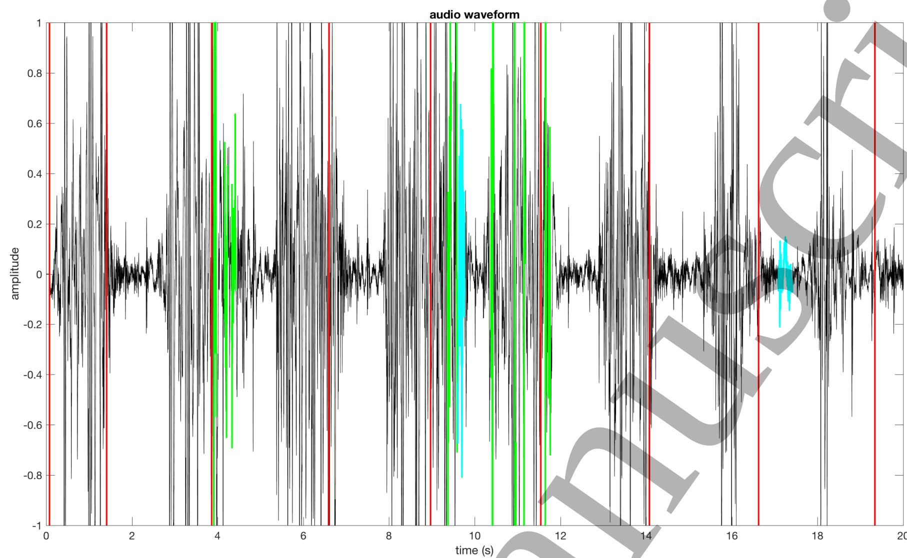

Figure 10: A segment containing 20 crackles (green) and 2 wheezes (blue) superimposed on the audio waveform (black). Respiratory cycle boundaries are represented by vertical lines (red).

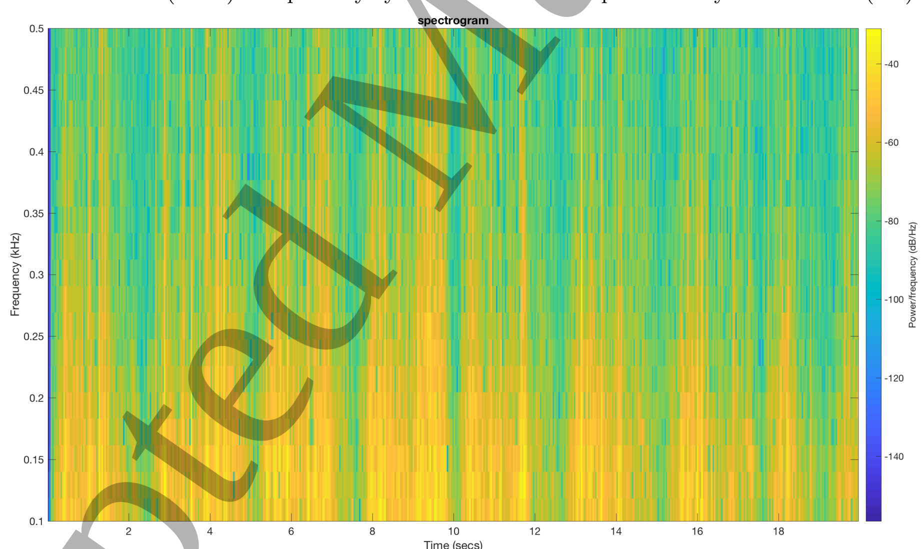

Figure 11: Spectrogram of the segment depicted in Fig. 6 and 10.

reduced from 5 to 1 in this experiment. Fig. 12 shows the scores for the event annotations.

First, we note that only SUK's system obtains higher scores than the baseline. SUK's scores are in line with the results reported in the literature. We can also see that the differences between HS and AS are not significant for any system.

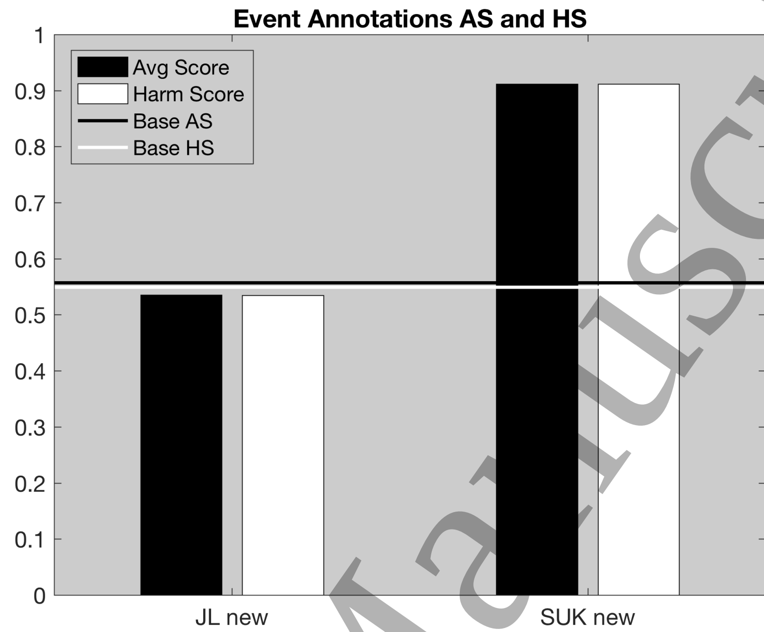

Figure 12: Average Score and Harmonic Score for the event annotations.

Fig. 13, which displays sensitivity and specificity values, confirms the reason for similar HS and AS: the disparity between SE and SP is quite smaller than the one seen with the cycle annotations.

Fig. 14 shows sensitivity values for each class except ‘normal’. SUK’s system attains almost perfect SE for ‘crackles’, while the baseline’s does not reach 80% and JL’s is below 60%. For ‘wheezes’, SUK’s SE is almost 80%, while JL’s is above 50% and the baseline’s is less than 10%.

7.3. Results per Participant

In this section, we analyse the results for each participant. However, we have to acknowledge that the recordings of 16 participants do not present adventitious respiratory sounds. Thus, the AS, HS, and SE values do not take into account these participants. Besides, although the variability in the number of events per participant has been reduced in this experiment, the number of adventitious respiratory sounds each participant presents is still highly variable, hence skewing the results. Table VIII presents the mean (M) and standard deviation (SD) of the evaluation metrics for each system.

Although the standard deviations are still large, SUK’s system seems clearly better suited for this problem than the other systems.

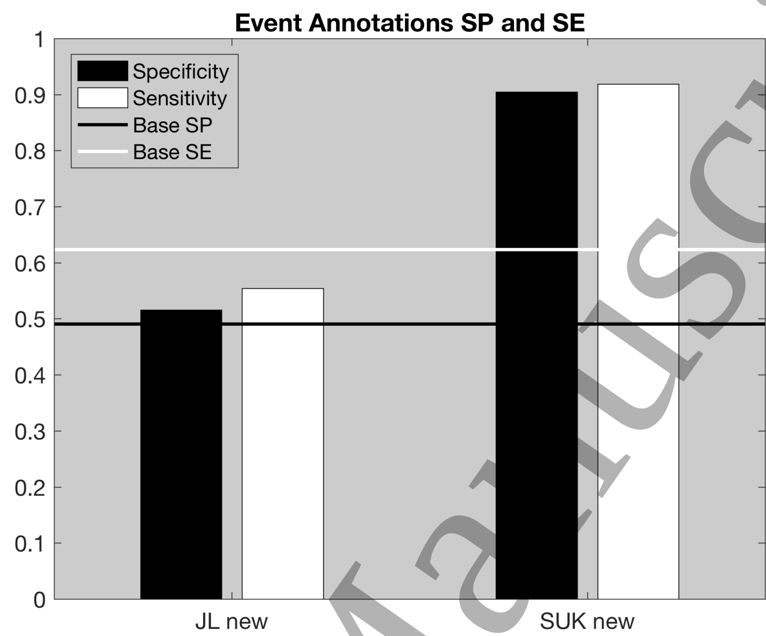

Figure 13: Specificity and sensitivity values for the event annotations.

Table VIII: Results per Participant (Experiment 2)

	AS		HS		SP		SE	
	M	SD	M	SD	M	SD	M	SD
JL new	55%	15%	49%	19%	51%	18%	61%	22%
SUK new	82%	21%	80%	25%	86%	22%	82%	22%
Baseline	46%	17%	36%	23%	55%	22%	45%	29%

8. Discussion

In this work, we have introduced a new database for the development of algorithms dedicated to the automatic classification of adventitious respiratory sounds. Additionally, we presented two experiments. In the first experiment, the ICBHI challenge, participating teams had to develop systems that classify respiratory cycles as one of four possible classes: containing crackles, containing wheezes, containing both crackles and wheezes, or normal (devoid of adventitious respiratory sounds). In the second experiment, participating teams had to develop systems that classify events as crackles, wheezes, or normal.

Among the factors that influence the performance of the different systems, the precision of the annotations appears to be the most important. As can be seen in the comparison between Fig. 6 and Fig. 10, adventitious respiratory

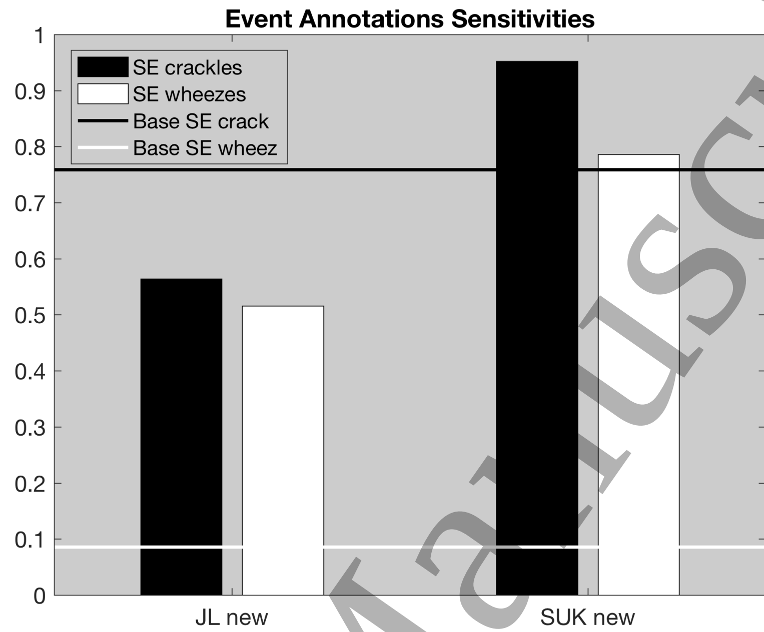

Figure 14: Sensitivity values for each class in the event annotations.

sound events are contained within respiratory cycles and are usually much shorter than them. Therefore, systems trained on precisely-annotated events have higher probability of correctly classifying crackles and wheezes than systems trained on respiratory cycles, as the adventitious respiratory sound events cover only part of a cycle.

However, only SUK's system took full advantage of the more precise annotations. We speculate that it might be due to the domain knowledge that SUK's system incorporated, i.e., carefully engineered features taking into account the peculiarities of adventitious respiratory sounds' signals. Both JL's system and the baseline could not model correctly the desired signals, showing that MFCCs may not be the most appropriate features for the discrimination of different types of respiratory sounds.

Some limitations of the database may have influenced the performance of the evaluated systems, such as the lack of healthy adult participants, the absence of gold standard annotations (i.e., annotations from multiple annotators), and the shortage of confounding noise sources.

It has been shown that healthy adults can exhibit adventitious respiratory sounds [65] and it would be useful to compare the characteristics and frequency of these sounds in healthy people and patients.

All the files in this database were annotated by a single health professional. While this has been common practice in the literature because annotation is a very time-consuming task, Dinis et al. [55] demonstrated the importance of obtaining annotations from multiple annotators and creating agreement metrics robust enough to extract reference annotations. A recent study found moderate to good agreement between observers when classifying crackles (Fleiss' kappa = 0.62) and wheezes (Fleiss' kappa = 0.59) [66].

Finally, although most files in this database contain some confounding noises, such as handling noise, cough, and speech, a robust database should contain other noise sources that would only be captured if in-the-field recordings were available. Zhang et al. [67] have analysed how an eating detection algorithm with high performance in experiments in the lab fails to generalise in the field. It would be desirable that a similar study be conducted with respiratory sound classification algorithms.

9. Conclusion

Respiratory sound classification is a complex task, as we expect to have demonstrated with the establishment of this database and the related scientific challenge. The public release of the respiratory sound database can serve many users. Researchers that want to develop algorithms for respiratory sound analysis may benefit from the access to real clinical signals and both sets of annotations, along with the presented experiments, approaches, and results. Likewise, this database can be useful for biomedical engineering education, by providing clinical signals to professors and students. Furthermore, we believe that new wearable systems and home-based measurements would enrich the field with the creation of big databases and propel digital auscultation to the era of 'big data'. That would allow for recent advances in the analysis of large databases to be applied in this field, as well as letting researchers study the impact of different factors (e.g., gender, body size, recording place, subjects' position and respiratory flow) on respiratory sound characteristics. We hope this respiratory sound database will inspire researchers interested in respiratory sound analysis to continue their venture in pushing forward this field.

Conflict of Interest

The authors of this article declare that they have no conflict of interest.

Acknowledgments

The scientific challenge was sponsored by the Working Group on Health Informatics and eHealth of the IFMBE. We acknowledge their financial and organisational support. This work was partially supported by: Fundação para a Ciência e Tecnologia (FCT) Ph.D. scholarship SFRH/BD/135686/2018; EU FP7 programme WELCOME under grant agreement 611223; Fundo Europeu de Desenvolvimento Regional (FEDER) through Programa Operacional Competitividade e Internacionalização (COMPETE) and FCT under the project UID/BIM/04501/2013 and POCI-01-0145-FEDER-007628 – iBiMED; the Ministry of Education, Science and Technological Development of the Republic of Serbia under grant agreements TR32035 and TR32040; Bogazici University Research Fund under grant number 16A02D2 and the Ph.D. scholarship (2211) from Turkish Scientific and Technological Research Council (TUBITAK).

References

[1] World Health Organization. The top 10 causes of death [online] (2017).

[2] G. J. Gibson, R. Loddenkemper, B. Lundbäck, Y. Sibille, Respiratory health and disease in Europe: the new European Lung White Book., The European respiratory journal 42 (2013) 559–63. doi:10.1183/09031936.00105513. URL http://erj.ersjournals.com/content/42/3/559

[3] A. Marques, A. Oliveira, C. Jácome, Computerized adventitious respiratory sounds as outcome measures for respiratory therapy: a systematic review., Respiratory care 59 (5) (2014) 765–76. doi:10.4187/respcare.02765. URL http://rc.rcjournal.com/content/59/5/765.short

[4] D. Hayes, S. S. Kraman, The Physiologic Basis of Spirometry, Respiratory care i (2009) 1717–1726. URL http://rc.rcjournal.com/content/54/12/1717.short

[5] E. C. Larson, E. Saba, S. Kaiser, M. Goel, S. N. Patel, Pulmonary Monitoring Using Smartphones, Springer, 2017, pp. 239–264. doi:10.1007/978-3-319-51394-2_13. URL http://dx.doi.org/10.1007/978-3-319-51394-2_13

[6] S. Fleming, A. Pluddemann, J. Wolstenholme, C. Price, C. Heneghan, M. Thompson, Diagnostic Technology : Automated lung sound analysis for asthma, Tech. Rep. July, Primary Care Diagnostic Horizon Scanning Centre Oxford (2011). URL https://www.oxford.dec.nihr.ac.uk/files/reports-and-resources/horizon-scanning-report0016-automated-lung-sound-analysis.pdf

[7] R. X. A. Pramono, S. Bowyer, E. Rodriguez-Villegas, Automatic adventitious respiratory sound analysis: A systematic review, PLOS ONE 12 (5) (2017) e0177926. doi:10.1371/journal.pone.0177926. URL https://doi.org/10.1371/journal.pone.0177926

[8] A. Gurung, C. G. Scrafford, J. M. Tielsch, O. S. Levine, W. Checkley, Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: A systematic review and meta-analysis, Respiratory Medicine 105 (9) (2011) 1396–1403. arXiv:NIHMS150003, doi:10.1016/j.rmed.2011.05.007. URL http://dx.doi.org/10.1016/j.rmed.2011.05.007

[9] S. Reichert, R. Gass, C. Brandt, E. Andrès, Analysis of respiratory sounds: state of the art, Clinical medicine. Circulatory, respiratory and pulmonary medicine 2 (2008) 45–58.
URL https://doi.org/10.4137/CCRPM.S530

[10] I. Chouvarda, V. Kilintzis, H. Karis, E. Kaimakamis, E. Perantoni, N. Maglaveras, L. Mendes, C. Lúcio, C. Teixeira, J. Henriques, P. Carvalho, R. Paiva, S. D'Arcy, N. Philip, O. Chetelat, J. Wacker, M. Rapin, C. Meier, J. Porchet, I. Frerichs, A. Raptopoulos, Combining pervasive technologies and cloud computing for COPD and comorbidities management, in: 2014 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH), 2014.
URL http://ieeexplore.ieee.org/xpls/abs{_}all.jsp?arnumber=7015984

[11] J. Earis, B. Cheetham, Current methods used for computerized respiratory sound analysis, European Respiratory Review.

[12] A. Sovijarvi, F. Dalmasso, J. Vanderschoot, L. P. Malmberg, G. Righini, Definition of terms for applications of respiratory sounds, Eur Respiratory Review 10 (77) (2000) 597–610.

[13] P. Piirila, A. Sovijarvi, Crackles: recording, analysis and clinical significance, Eur. Respir. J. 8 (12) (1995) 2139–2148.
URL http://erj.ersjournals.com/content/8/12/2139.short

[14] M. Sarkar, Madabhavi, Niranjan, Dogra, Auscultation of the respiratory system, Annals of Thoracic Medicine 10 (3) (2015) 158–168. doi:10.4103/1817-1737.160831.
URL http://dx.doi.org/10.4103/1817-1737.160831

[15] A. Sovijarvi, L. P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, J. E. Earis, Characteristics of breath sounds and adventitious respiratory sounds, Eur Respir Review.

[16] G. R. Epler, C. B. Carrington, E. A. Gaensler, Crackles (Rales) in the Interstitial Pulmonary Diseases, Chest 73 (3) (1978) 333–339. doi:10.1378/chest.73.3.333.
URL https://doi.org/10.1378/chest.73.3.333

[17] A. Nath, L. Capel, Inspiratory crackles—early and late, Thorax.
URL http://dx.doi.org/10.1136/thx.29.2.223

[18] J. Marini, D. Pierson, L. Hudson, The Significance of Wheezing in Chronic Airflow Obstruction, American Review of Respiratory Disease.
URL http://www.atsjournals.org/doi/abs/10.1164/arrd.1979.120.5.1069

[19] R. P. Baughman, R. G. Loudon, Quantitation of wheezing in acute asthma, Chest 86 (5) (1984) 718–722. doi:10.1378/chest.86.5.718.
URL https://doi.org/10.1378/chest.86.5.718

[20] R. B. Urquhart, J. McGhee, J. E. S. Macleod, S. W. Banham, F. Moran, The diagnostic value of pulmonary sounds: A preliminary study by computer-aided analysis, Computers in Biology and Medicine 11 (3) (1981) 129–139. doi:10.1016/0010-4825(81)90002-0.
URL https://doi.org/10.1016/0010-4825(81)90002-0

[21] R. L. Murphy, E. A. Del Bono, F. Davidson, Validation of an automatic crackle (rale) counter., The American Review of Respiratory Disease 140 (4) (1989) 1017–20. doi:10.1164/ajrccm/140.4.1017.
URL http://www.atsjournals.org/doi/10.1164/ajrccm/140.4.1017

[22] B. Sankur, Y. P. Kahya, E. Guler, T. Engin, Comparison of AR-based algorithms for respiratory sounds classification, Computers in Biology and Medicine 24 (1) (1994) 67–76. doi:10.1016/0010-4825(94)90038-8.
URL https://doi.org/10.1016/0010-4825(94)90038-8

[23] K. E. Forkheim, D. Scuse, H. Pasterkamp, A Comparison of Neural Network Models for Wheeze Detection, in: WESCANEX 95. Communications, Power, and Computing. Conference Proceedings., IEEE, no. 95, 1995, pp. 214–219.
URL https://ieeexplore.ieee.org/abstract/document/493973/

[24] M. Du, F. Chan, F. Lam, J. Sun, Crackle Detection and Classification Based on Matched Wavelet Analysis, The 19th IEEE Engineering in Medicine and Biology Society Conference Proceedings.
URL https://ieeexplore.ieee.org/document/757031/

[25] S. A. Taplidou, L. J. Hadjileontiadis, Wheeze detection based on time-frequency analysis of breath sounds, Computers in Biology and Medicine 37 (8) (2007) 1073–1083. doi:10.1016/j.compbiomed.2006.09.007.
URL https://doi.org/10.1016/j.compbiomed.2006.09.007

[26] M. Bahoura, Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes, Computers in biology and medicine 39 (9) (2009) 824–843.
URL https://doi.org/10.1016/j.compbiomed.2009.06.011

[27] N. Nakamura, M. Yamashita, S. Matsunaga, Detection of patients considering observation frequency of continuous and discontinuous adventitious sounds in lung sounds, Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2016-Octob (2016) 3457–3460. doi:10.1109/EMBC.2016.7591472.
URL https://ieeexplore.ieee.org/document/7591472/

[28] P. Bokov, B. Mahut, P. Flaud, C. Delclaux, Wheezing recognition algorithm using recordings of respiratory sounds at the mouth in a pediatric population,

Computers in Biology and Medicine 70. doi:10.1016/j.compbiomed.2016.01.002.
URL https://doi.org/10.1016/j.compbiomed.2016.01.002

[29] X. Liu, W. Ser, J. Zhang, D. Y. T. Goh, Detection of adventitious lung sounds using entropy features and a 2-D threshold setting, in: 10th International Conference on Information, Communications and Signal Processing (ICICS), IEEE, 2015, pp. 1–5. doi:10.1109/ICICS.2015.7459851.
URL https://ieeexplore.ieee.org/document/7459851/

[30] S. Ulukaya, G. Serbes, Y. P. Kahya, Overcomplete discrete wavelet transform based respiratory sound discrimination with feature and decision level fusion, Biomedical Signal Processing and Control 38 (September) (2017) 322–336. doi:10.1016/j.bspc.2017.06.018.
URL http://dx.doi.org/10.1016/j.bspc.2017.06.018

[31] C. Pinho, A. Oliveira, C. Jácome, J. Rodrigues, A. Marques, Automatic Crackle Detection Algorithm Based on Fractal Dimension and Box Filtering, Procedia Computer Science 64 (October) (2015) 705–712. doi:10.1016/j.procs.2015.08.592.
URL https://doi.org/10.1016/j.procs.2015.08.592

[32] M. Lozano, J. A. Fiz, 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 20 (2) (2016) 486–497. doi:10.1109/JBHI.2015.2396636.
URL https://ieeexplore.ieee.org/document/7021891/

[33] D. Chamberlain, R. Kodgule, D. Ganelin, V. Miglani, R. R. Fletcher, Application of Semi-Supervised Deep Learning to Lung Sound Analysis Application of Semi-Supervised Deep Learning to Lung Sound Analysis, in: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), no. August, IEEE, 2016, pp. 804–807. doi:10.1109/EMBC.2016.7590823.
URL http://ieeexplore.ieee.org/document/7590823/

[34] B. S. Lin, B. S. Lin, Automatic Wheezing Detection Using Speech Recognition Technique, Journal of Medical and Biological Engineering 36 (4) (2016) 545–554. doi:10.1007/s40846-016-0161-9.
URL http://dx.doi.org/10.1007/s40846-016-0161-9

[35] R. Naves, B. H. G. Barbosa, D. D. Ferreira, Classification of lung sounds using higher-order statistics: A divide-and-conquer approach, Computer Methods and Programs in Biomedicine 129. doi:10.1016/j.cmpb.2016.02.013.
URL https://doi.org/10.1016/j.cmpb.2016.02.013

[36] L. Mendes, I. M. Vogiatzis, E. Perantoni, E. Kaimakamis, I. Chouvarda, N. Maglaveras, J. Henriques, P. Carvalho, R. P. Paiva, Detection of crackle events using a multi-feature approach, in: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2016, pp. 3679–3683. doi:10.1109/EMBC.2016.7591526.
URL http://ieeexplore.ieee.org/document/7591526/

[37] D. Owens, R.A.L.E. Lung Sounds 3.0, Computers, Informatics, Nursing 5 (3) (2002) 9–10.
URL http://journals.lww.com/cinjournal/Citation/2002/09010/R_A_L_E__Lung_Sounds_3_0.4.aspx

[38] S. Lehrer, Understanding lung sounds, the 3rd edition, W.B. Saunders, 2002.
URL https://books.google.pt/books?id=LURrAAAACAAJ{&}source=gbs{_}book{_}other{_}versions

[39] W. Lippincott, Auscultation Skills: Breath and Heart Sounds, Wolters Kluwer, 2009.

[40] East Tennessee State University. Pulmonary Breath Sounds [online] (2002).

[41] R. L. Wilkins, J. E. Hodgkin, B. Lopez, Fundamentals of lung and heart sounds, Mosby, 2004.

[42] D. Wrigley, Heart and Lung Sounds Reference Library, PESI Healthcare, 2011.

[43] M. Littmann. Basic Lung Sounds Tutorial [online] (2018).

[44] S. S. Kraman, Lung sounds: an introduction to the interpretation of the auscultatory finding, American College of Chest Physicians, Northbrook, IL (1993) 19–21.

[45] S. Mangione, Secrets Heart & Lung Sounds Workshop: Audio CD, Hanley & Belfus.

[46] SoundCloud. Lung Sounds [online] (2018).

[47] G. Druger, The chest: its signs and sounds, Humetrics Corporation, 1973.

[48] A. G. Tilkian, M. B. Conover, Understanding heart sounds and murmurs with an introduction to lung sounds, WB Saunders Company, 2001.

[49] S. Lehrer, Understanding Lung Sounds, the 2nd edition, Saunders, Philadelphia, PA.

[50] B. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, R. Paiva, I. Chouvarda, P. Carvalho, Maglaveras, A Respiratory Sound Database for the

Development of Automated Classification, in: Precision Medicine Powered by pHealth and Connected Health, 2018, pp. 33–37. doi:10.1007/978-981-10-7419-6_6.
URL http://dx.doi.org/10.1007/978-981-10-7419-6_6

[51] International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use, ICH HARMONISED TRIPARTITE GUIDELINE: CLINICAL INVESTIGATION OF MEDICINAL PRODUCTS IN THE PEDIATRIC POPULATION E11, in: Int. Conf. Harmon. Tech. Requir. Regist. Pharm. Human Use, 2000.
URL http://www.ema.europa.eu/docs/en{_}GB/document{_}library/Scientific{_}guideline/2009/09/WC500002926.pdf

[52] M. Rossi, A. Sovijarvi, P. Piirila, L. Vannuccini, F. Dalmasso, J. Vanderschoot, Environmental and subject conditions and breathing manoeuvres for respiratory sound recordings, European Respiratory Review.

[53] A. Machado, A. Oliveira, C. Jácome, M. Pereira, J. Moreira, J. Rodrigues, J. Aparício, L. M. T. Jesus, A. Marques, Usability of Computerized Lung Auscultation–Sound Software (CLASS) for learning pulmonary auscultation, Medical & Biological Engineering & Computingdoi:10.1007/s11517-017-1697-8.
URL http://link.springer.com/10.1007/s11517-017-1697-8

[54] K. K. Guntupalli, P. M. Alapat, V. D. Bandi, I. Kushnir, Validation of automatic wheeze detection in patients with obstructed airways and in healthy subjects, Journal of Asthma 45 (10) (2008) 903–907. doi:10.1080/02770900802386008.
URL https://doi.org/10.1080/02770900802386008

[55] J. Dinis, G. Campos, J. Rodrigues, A. Marques, Respiratory Sound Annotation Software, Proceedings of the International Conference on Health Informatics (2012) 183–188doi:10.5220/0003756301830188.
URL http://doi.org/10.5220/0003756301830188

[56] AudacityTeam, Audacity(R): Free Audio Editor and Recorder (2014).

[57] I. W. Selesnick, Sparse signal representations using the tunable Q-factor wavelet transform, Proceedings of SPIE: Wavelets and Sparsity XIV 8131 (2011) 81381U–81381U–15. doi:10.1117/12.894280.
URL https://doi.org/10.1117/12.894280

[58] I. W. Selesnick, Wavelet transform with tunable Q-factor, IEEE Transactions on Signal Processing 59 (8) (2011) 3560–3575. doi:10.1109/TSP.2011.2143711.
URL https://ieeexplore.ieee.org/document/5752263/

[59] G. Serbes, S. Ulukaya, Y. P. Kahya, An Automated Lung Sound Preprocessing and Classification System Based on Spectral Analysis Methods, in:

Precision Medicine Powered by pHealth and Connected Health, Vol. 66, 2018, pp. 45–49. doi:10.1007/978-981-10-7419-6_8.
URL http://dx.doi.org/10.1007/978-981-10-7419-6_8

[60] M. Berouti, R. Schwartz, J. Makhoul, Enhancement of speech corrupted by acoustic noise, ICASSP '79. IEEE International Conference on Acoustics, Speech, and Signal Processing 4 (May 1979) (1979) 208–211. doi:10.1109/ICASSP.1979.1170788.
URL http://ieeexplore.ieee.org/document/1170788/

[61] M. Gales, S. Young, The Application of Hidden Markov Models in Speech Recognition, Found. Trends Signal Process. 1 (3) (2008) 195–304. arXiv:arXiv:1011.1669v3, doi:http://dx.doi.org/10.1561/2000000004.
URL http://dx.doi.org/10.1561/2000000004

[62] N. Jakovljević, T. Lončar-Turukalo, Hidden Markov Model Based Respiratory Sound Classification, in: Precision Medicine Powered by pHealth and Connected Health, Springer, Singapore, 2018, pp. 39–43. doi:10.1007/978-981-10-7419-6_7.
URL http://dx.doi.org/10.1007/978-981-10-7419-6_7

[63] O. Lartillot, P. Toiviainen, A matlab toolbox for musical feature extraction from audio, International Conference on Digital Audio (2007) 1–8.
URL http://dafx.labri.fr/main/papers/p237.pdf

[64] W. G. Hale, V. A. Saunders, J. P. Margham, W. G. Hale, Dictionary of biology, rev. and updated. Edition, HarperCollins, Glasgow, 2003.
URL https://www.worldcat.org/title/collins-dictionary-of-biology/oclc/1017279313?referer=di{&}ht=edition

[65] A. Oliveira, A. Marques, Respiratory sounds in healthy people: A systematic review, Respiratory Medicine 108 (4) (2014) 550–570. doi:10.1016/j.rmed.2014.01.004.
URL http://dx.doi.org/10.1016/j.rmed.2014.01.004

[66] H. Melbye, L. Garcia-Marcos, P. Brand, M. Everard, K. Priftis, H. Pasterkamp, Wheezes, crackles and rhonchi: simplifying description of lung sounds increases the agreement on their classification: a study of 12 physicians' classification of lung sounds from video recordings, BMJ Open Respir. Res. 3 (1) (2016) e000136. doi:10.1136/bmjresp-2016-000136.
URL http://bmjopenrespres.bmj.com/lookup/doi/10.1136/bmjresp-2016-000136

[67] S. Zhang, R. Alharbi, M. Nicholson, N. Alshurafa, When Generalized Eating Detection Machine Learning Models Fail in the Field, in: Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, 2017, pp. 613–622. doi:10.1145/3123024.3124409.
URL https://doi.org/10.1145/3123024.3124409