Enabling effective breathing sound analysis for automated diagnosis of lung diseases

Wassila Lalouani [a], Mohamed Younis [b], Roland N. Emokpae Jr. [c],
Lloyd E. Emokpae [c],[*]

[a] *Department of Computer and Information Science, Towson University, USA*
[b] *CSEE Dept., Univ. of Maryland, Baltimore County, Baltimore, MD, USA*
[c] *LASARRUS Clinic and Research Center Inc., Baltimore, MD, USA*

ARTICLE INFO

ABSTRACT

With the emergence of the COVID-19 pandemic, early diagnosis of lung diseases has attracted growing attention. Generally, monitoring the breathing sound is the traditional means for assessing the status of a patient's respiratory health through auscultation; for that a stethoscope is one of the clinical tools used by physicians for diagnosis of lung disease and anomalies. On the other hand, recent technological advances have made telehealth systems a practical and effective option for health status assessment and remote patient monitoring. The interest in telehealth solutions have further grown with the COVID-19 pandemic. These telehealth systems aim to provide increased safety and help to cope with the massive growth in healthcare demand. Particularly, employing acoustic sensors to collect breathing sound would enable real-time assessment and instantaneous detection of anomalies. However, existing work focuses on autonomous determination of respiratory rate which is not suitable for anomaly detection due to inability to deal with noisy data recording. This paper presents a novel approach for effective breathing sound analysis. We promote a new segmentation mechanism of the captured acoustic signals to identify breathing cycles in recorded sound signals. A scoring scheme is applied to qualify the segment based on the targeted respiratory illness by the overall breathing sound analysis. We demonstrate the effectiveness of our approach via experiments using published COPD datasets.

1. Introduction

Respiratory diseases are worrisome for medical providers and constitute a major cause of hospitalization. These illnesses can be fatal, especially for elders and people with weak immune systems. The COVID-19 pandemic has made respiratory illnesses a big concern, where we have witnessed the unfortunate loss of lives at a historic rate just due to the pandemic. According to the World Health Organization (WHO) global impact import (Forum of International Respiratory Societies, 2017), respiratory diseases are the leading causes of death and disability in the world. Chronic obtrusive pulmonary disease (COPD) is the most prevalent, with more than 65 million people suffering from COPD each year and 3 million deaths, making it the third leading cause of fatality worldwide. Some other respiratory diseases of concern include Asthma, Pneumonia, acute respiratory tract infections, tuberculosis, lung cancer,

emphysema, Bronchiolitis, Alpha-1 and cystic fibrosis. Traditionally, some of these respiratory diseases can be diagnosed using stethoscopes for assessing breathing regularity or coughs (Turner & Bothamley, 2014). Such diagnostics opt to determine the frequency of breathing and/or cough events over a given time interval to ascertain the presence of abnormal sounds such as wheezing, crackles, rhonchi, and others. Given the use of conventional stethoscopes, the diagnostics are often performed during in-person physician visits.

With the rise in healthcare cost and stress on facilities caused by high demand, especially during the COVID-19 pandemic, telehealth systems have become attractive options where patients can be monitored remotely without the need for visiting clinics and hospitals. A telehealth system consists of a set of sensors that are attached to the patient's body; these sensors share their data, either raw or preprocessed, with remote healthcare providers to assess the patient conditions. The major advancements in wireless technologies and microelectronics have made these networked solutions quite effective and inspired automation in the diagnostic process, most notably for lung diseases (Gurung et al., 2011; Rao et al., 2019). This workflow is illustrated in Fig. 1, which depicts the use of a telehealth system for remote auscultation and monitoring of respiratory symptoms. Several algorithms have been proposed for automated processing of breathing sound recordings. Examples include LifeShirt (Grossman, 2004), Vitalojak (Smith & Woodcock, 2008), and Pulmotrack (Vizel et al., 2010), and many others (Hall et al., 2020), where either digital signal processing (DSP) or machine learning techniques are employed to detect adventitious patterns (anomaly) in the collected acoustic measurements. In the context of respiratory illness, such anomaly detection is conducted at the level of breathing cycle.

We note that in a telehealth system data is streamed; thus, it is necessary to analyze the collected data and apply anomaly detection in real-time (Ahmad et al., 2017). Automatic detection of breathing cycles in a noisy recording is challenging, not to mention the variability of the breathing pattern among individuals based on health conditions, activities, body composition, age, race, gender, etc. In essence, the effectiveness of a respiratory telehealth system is dependent on the recording conditions including the patient's proximity to the microphone, sensor sensitivity and breathing intensity. The recorded sound waveforms are also affected by the ambient noise in case of contactless and body-attached sensors, respectively (Massaroni et al., 2019). Existing algorithms do not deal with the noise in real-time and rely on simple signal processing primitives (Cohen-McFarlane et al., 2019). In this paper we address the aforementioned issues and promote a novel automated data-driven segmentation (AUDAS) mechanism of recorded chest sound to detect breathing cycles in real-time. AUDAS strives to optimally partition the streamed data based on contextual information, specifically the respiratory illness that is being tracked. In other words, the quality of a segment is gauged in terms of the exhibited anomalous pattern. In this paper we use the International Conference on Biomedical Health and Informatics (ICBHI) database (Moussavi, 2006) for illustrating the utility of our segmentation mechanism and assessing its performance. The results show the outstanding performance of AUDAS. The main contribution of this paper is summarized as follow:

- The development of a context aware acoustic signal segmentation mechanism for medical diagnostics. Our mechanism achieves accurate disease classification with real-time symptom detections.
- The design of a lightweight classifier for the detection of breathing cycles and development of a scoring method to increase the efficiency of the segmentation using a learning approach.
- Demonstrating the utility of our segmentation mechanism and assessing its performance using popular datasets.

The paper is organized as follows. The next section discusses related work in the literature. Section 3 enumerates the challenges and provides an overview of the proposed solution. Section 4 describes AUDAS in detail. The validation results are reported in Section 5 and the paper is concluded in Section 6.

2. Related work

The use of DSP techniques has been popular for conducting automated respiratory sound analysis (Moussavi, 2006). However, with the recent major advances in machine learning algorithms and the availability of cloud resources, data-driven methodologies have become the most attractive option, especially with their ability to personalize the analysis and conduct it in real-time (Kim et al., 2021a; Srivastava et al., 2021). Anomaly detection based on streamed data has been attracting increased attention (Ahmad et al., 2017); yet the focus has been on matching certain signal patterns, e.g., temporal patterns. In our case, the diagnostic process is not only based on breathing sounds (features) but also on the breathing rate; both of which vary, and hence signal segmentation would be necessary. Quite a few approaches, e.g. (Mukherjee et al., 2021),- (Ma et al., 2020), have been proposed to classify the lung sound and

Table 1
A comparative summary of the capabilities of AUDAS to those of published approaches.

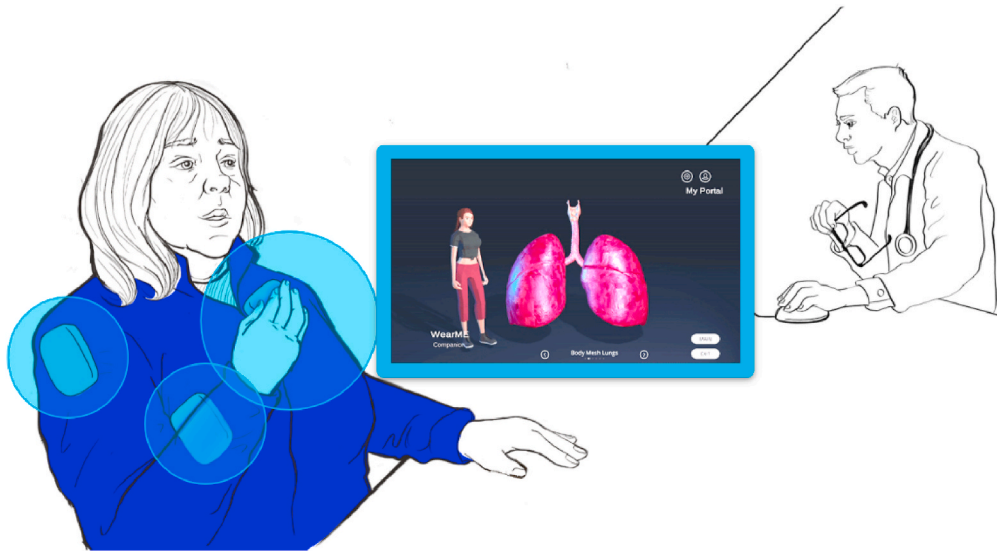| Features and Usage	Automated Segmentation	Lung Sound Classification	Breathing Cycle	Lung Function Assessment
AUDAS	✓	✓	✓	✓
Ahmad et al. (Ahmad et al., 2017)		✓		
Mukherjee et al. (Mukherjee et al., 2021)	✓	✓		
Rao et al. (Rao et al., 2018)	✓	✓		
Demir et al. (Demir et al., 2020)	✓	✓		
Fraiwan et al. (Fraiwan et al., 2021)				✓
Hsiao et al. (Hsiao et al., 2020)	✓		✓	
Jacome et al. (Jácome Cristina et al., 2019)	✓	✓	✓	

Fig. 1. Illustration of a telehealth system used for remote patient monitoring of respiratory illness symptoms.

detect anomalies. AUDAS complements these approaches. Basically, in order to apply the lung sound classifier, the breathing cycle needs to be determined. Moreover, the breathing rate plays a major role in the diagnostics, and a lung sound classifier alone would not suffice (Nicolò et al., 2020).

Segmenting a breathing sound signal is complicated by the background noise (Emokpae et al., 2021). Some work has focused on tackling the noise effect at the sensor level by factoring in other modality (Gupta et al., 2021). However, acoustics continue to be predominantly the most acceptable means for diagnosing respiratory illness and automating the analysis of collected sound signals is highly desired for the feasibility of the corresponding telehealth systems. The incorporation of breathing sound analysis is becoming more popular in telehealth systems both for monitoring illness symptoms and fitness during exercises (Gu et al., 2017)- (Oletic & Bilas, 2016). However, the breathing cycle is assumed to be fixed in size despite variation among individuals. Such an assumption makes data segmentation quite simple, yet it is not practical and can lead to delayed response to alarming development, e.g., in the case of COPD.

Partitioning a signal into segments of different sizes based on the contextual features is quite challenging. Some prior studies considered such segmentation for modality other than acoustics. For example, Haddad, and Najafizadeh (Haddad & Najafizadeh, 2019) factored in the spatial distribution of active cortical neurons in the temporal partitioning of the EEG signal into segments. Obviously, the contextual features for sound signals is different from EGG signals and other sensor modality. Work on segmentation of sound signals either considers simple statistical measures, e.g., covariance, (Paul and Shoukat Choudhury, 2015), or pursue a greedy approach with unbounded complexity with no notion of global optimality (Bosse et al., 2018). Some have mainly focused on collecting accurate training data to develop a segmentation classifier. For example, Kong et at (Kong et al., 2019). use time and frequency masking to detect specific sound patterns to overcome weakly labeled data. Meanwhile, Martín-Morató et al. (Martín-Morató et al., 2021) and Fraiwan et al. (Fraiwan et al., 2021) assume fixed size segments and pursue crowdsourcing to determine the boundary of specific sound patterns.

There have also been some published approaches that investigate dynamic window sizes for segmenting breathing cycles. Hsiao et al. (Hsiao et al., 2020) promote dynamic segmentation to detect inspiratory or expiratory sounds using the encoded spectrogram on an attention-based decoder. Although their approach achieves high segmentation accuracy, it involves a complicated architecture and does not suit real-time respiratory monitoring. Jacome et al. (Jácome Cristina et al., 2019) propose a model for breathing phase detection based on Faster R–CNN object detection system. However, these dynamic segmentation approaches are limited to detecting respiratory periods and cannot be applied to assess lung function status (Emokpae et al., 2022). For patients with COPD, being able to assess changes in lung function will enable early treatment intervention. Traditionally, these changes are typically captured by calculating both the forced expiratory volume in 1-s (FEV1) and the forced vital capacity (FVC) through gold-standard spirometry. Whereby the ratio of FEV1 to FVC can be used to diagnose if a patient has COPD (Pellegrino et al., 2005). In our recent publication (Emokpae et al., 2022), we show that we can leverage features of breathing cycles with normal and deep breaths to estimate lung function status. The proposed AUDAS approach enables automated segmentation of breathing features which can be used to compute FEV1 and FVC, for effective COPD diagnosis. Unlike prior work, AUDAS factors in the contextual features in the segmentation process to detect abnormal sounds and symptoms in order to optimize the selection of the segmentation option that better reflects the conditions of the monitored individual. Please refer to Table 1 for a comparative summary of AUDAS to published approaches.

3. Design goals and approach overview

We are considering a telehealth system that consists of wearable sensors and is connected to remote facilities or caregivers through

wireless communication links. The sensor data is processed to assess the wellbeing of the monitored individual. Such processing is performed by a computationally capable server or using cloud resources. The paper further focuses on wearable acoustic sensors that are attached to the chest and provide sound recording to diagnose respiratory illness conditions. The balance of this section highlights the technical challenges associated with the automation of chest sound analysis, and how AUDAS tackles these challenges.

3.1. Requirements and challenges

Generally, telehealth data has streaming nature and is characterized by irregularity and correlation over time. Particularly these characteristics are quite influential for systems that utilize acoustic signals to assess lung conditions, where the breathing cycle and sound vary from one person to another and even for the same person based on activities and illness. Appropriate segmentation of the collected data is necessary to adapt to changes. These characteristics make the traditional anomaly detection algorithms unsuitable. Specifically, we are confronted with the following challenges:

(1) The wearable sensors are geared for timely and continual assessment of patient's conditions and generate streamed data at a high sampling rate; hence, it is important to employ data segmentation that enables performing diagnostics in real-time.
(2) The collected data is inherently correlated, and distinct data streams provide different features (varying dimensionality); inappropriate data segmentation may prevent the learning mechanism from capturing important features and degrade the diagnosis accuracy.
(3) Like traditional anomaly detection, the telehealth real-time data can be very noisy. The source of the noise includes sounds from other body organs, e.g., heart, or due to external noise from the environment. This is illustrated in Fig. 2, which shows a combination of breathing sounds with S1/S2 heart sounds collected by our wearable sensor used in breathing analysis (Emokpae et al., 2021). Ensuring accurate anomaly detection in presence of noise is necessary for accurate diagnosis.

We argue that signal processing techniques for segmentation constitute a more deterministic way to distinguish the breathing cycles, yet they do not factor in the variability between the patients and the impact of significant noises within the data streaming. Data-driven approaches are more appropriate to deal with the aforementioned challenges, and cope with such variability.

3.2. Approach overview

Overall, the design goals of an anomaly detection mechanism for telehealth systems are: (i) enabling continual incorporation of streaming data transfer under constrained devices and network resources, (ii) providing timely reports to users (patients and providers) to avoid missing any alarming patient's conditions, and (iii) striving to accurately reflect the status of the participant. To achieve these goals, AUDAS pursues extraction of contextual information to segment the collected data to meaningful records that provide the most evidence about the user status. The irregularity of the user's activities and sensing capabilities constitute an important challenge. To illustrate, acoustic measurements of breathing could be irregular depending on the type of activities and the biomedical state of the participant. On the other hand, the sensors may be relocated according to the user movement, which creates variability even when there are no changes in the user's health conditions. Extracting the most relevant set of samples is a very challenging problem and is subject to the data semantic.

AUDAS pursues an adaptive data segmentation process that factors in the relevance to the underlying application. The collected data will be considered in batches; the batch size depends on the sampling rate, buffer size and communication link/protocol for the wearable device, e.g., acoustic sensor in case of lung function monitoring. For each batch, AUDAS strives to find the best segmentation.
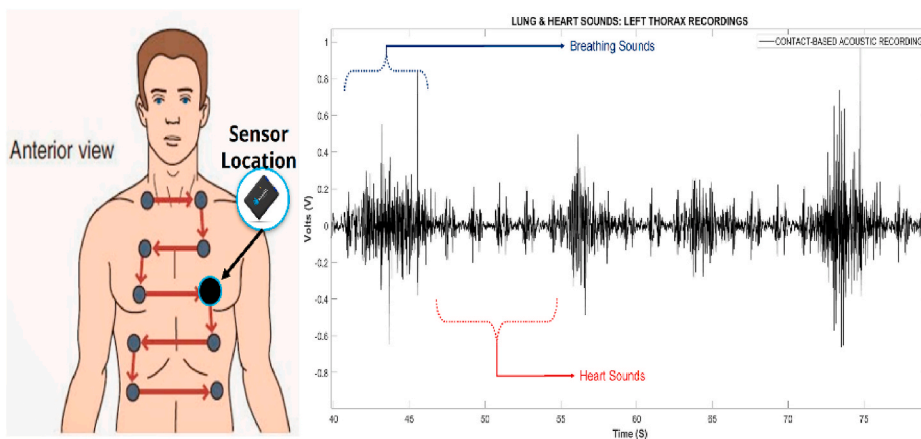

Fig. 2. Recording of lung and heart sounds taken from left thorax with our wearable sensors. It shows noise due to regular heart sounds that affect the analysis of breathing information.

The quality of a certain combination of segments is based on the utility of the individual segments with respect to the diagnostic objective. To assess the quality of a segment, AUDAS employs a machine learning (ML) model based on segment patterns that are being detected from the labeled data. In the case of respiratory diseases, the ML model considers distinct spectral features within annotated breathing cycles, coughs, wheezes, etc. The ML model will be initially trained offline and could further be adjusted over time. The data segment is used as an input (test data) to the ML and the classification confidence is used to assign a score for the segment. AUDAS employs a varying time window size to determine possible segments; each is scored through the trained ML model. Finally, all segment options within a batch are considered to determine the best combination that maximizes the average segment score.

4. Detailed AUDAS design

As pointed out earlier, existing schemes for automated respiratory disease diagnostics usually analyze the breathing sound to infer the rate and detect adventitious lung sounds, e.g., wheezing, crackles, etc. However, these schemes assume that the signal is already segmented accurately, and their objective is just to classify the sounds. Furthermore, they are sensitive to the background noise and quality of the recorded breathing sounds. In addition, the microphone position on the chest could vary due to motion and physical activities. AUDAS strives to overcome these issues and achieves a robust assessment of the patient's chest conditions. As explained in the balance of this section, AUDAS slices the sound recording while factoring in the overall diagnostics goal of the telehealth system. The idea is to qualify the various segmentation options based on the importance of exhibited features to the goal of distinguishing the various sound patterns. A classifier is then applied to the individual segments to detect the symptoms of underlying disease in order to provide real-time diagnostics.

4.1. Segmenting sound recording

As pointed out in Section 2, traditional segmentation algorithms do not consider relevant contextual information in the process, and hence their performance is degraded in the presence of noise and variability in the sensor measurements. AUDAS overcomes such a shortcoming and applies a dynamic data stream segmentation strategy that factors in the correlation and the contextual significance of the collected data samples. Let σ denote the output of a classifier that is built offline in order to detect normal and abnormal sound patterns in the data as specified by the application, e.g., detecting wheezes. Such a classifier is used by AUDAS to distinguish relevant parts of the data and assess the quality of a segment. AUDAS selects a slicing window W of data points, reflecting samples of the acoustic signal. We regard such a set of data points as an m-dimensional feature vector $\overline{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,m})$. Unlike existing segmentation schemes, the size of W in AUDAS is not fixed and varies based on the relevance of the features.

Varying the window size comes at a cost of computational complexity. Assuming that the measurements are disseminated from the
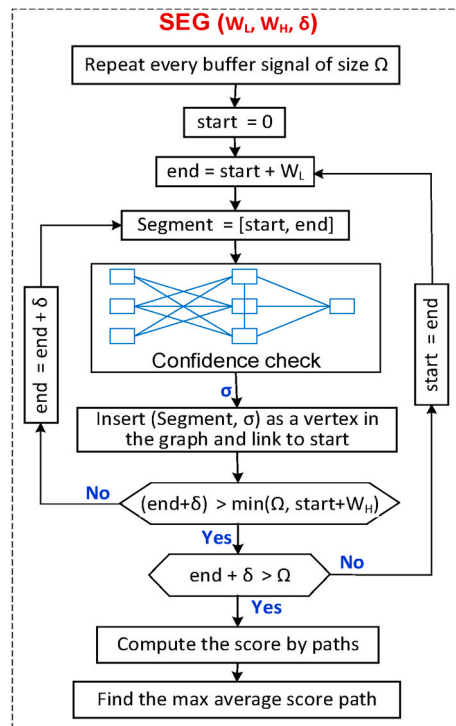

Fig. 3. Flowchart description of the proposed AUDAS algorithm for segmenting chest sound recordings.

wearable system to the server in batches of size Ω data points. Dynamic segmentation in AUDAS is a combinatorial problem that can be mapped to the problem of partitioning a set into k subsets, which has a runtime complexity of $O(2^{\Omega})$. Such a partitioning problem is known to be NP-Hard. To mitigate the complexity, we consider the cross-correlation among different segments and formulate the problem as dynamic programming to determine the most rewarding non-overlapping partitions, where rewarding here implies relevance to the sound pattern detection process and the overall objective of accurate anomaly detection. Given the objective function *MaxPart* which returns the maximum reward decomposition $D = (r_1, \ldots, r_n)$, the segmentation optimization is defined by the following recursive function:

$$MaxPart(r_i, r_j) = \max\{rwd(r_i, r_j) + MaxPart(r_i + 1, r_i), rwd(r_i + 1, r_j)\} \tag{1}$$

where r_i, and r_j are the current sample, and the sample at the beginning of the last considered segment, respectively, $MaxPart(r_i, r_j)$ is the maximum reward decomposition from sample r_i to W, and rwd is the confidence score of the ML classifier, which will be explained in the next subsection. Dynamic programming can reduce the runtime complexity to $O(n^2)$ where n is the maximum number of possible samples; yet such complexity continues to be high for large batches.

To limit the complexity, AUDAS exploits the semantics of a window size with respect to the application. For example, in the context of respiratory diseases, the breathing cycle reflects the window that is considered, where the breathing sound and frequency are used by the classifier to detect anomalies. Hence, AUDAS employs a lower and upper bounds, denoted by W_L, and W_H, for the window size. It has been noted in (Nuckowska et al., 2019) that an inspiration and expiration cycle can take as much as 5 s and as small as 2 s. In addition, AUDAS considers the sampling frequency of the acoustic signal in order to avoid processing irrelevant segments, where a step size δ is used to align the window size boundaries with the data points. The segmentation heuristic goes as follows. A data point corresponding to W_L is considered as a segment and fed to the classifier to obtain the corresponding σ. The window size is iteratively extended by δ and the associated σ is calculated, accordingly. Such a gradual increase of window size, i.e., growing it by δ, seizes when reaching W_H. Fig. 3 provides a summary of steps for AUDAS. As captured by the inner loop in the figure, AUDAS considers all feasible window sizes in the range $[W_L, W_H]$.

While Eq. (1) considers every data point as the possible start of a window, by considering W_L to be the least window size, AUDAS reduces the number of combinations by a factor of W_L. Consequently, the runtime complexity of the segmentation process becomes $O\left(\left(\frac{n}{W_L}\right)^2\right)$. The outer loop in Fig. 3 reflects the start of a new segment, assuming minimum spacing, i.e., at $i.W_L$, where $i = 1, 2, \ldots$, $\lceil(\Omega.\delta)/W_L\rceil$. Overall, the number of possible segments, ν, that AUDAS considers are $\frac{W_H - W_L}{\delta}\left\lfloor\frac{\Omega.\delta}{W_L}\right\rfloor + \frac{\Omega.\delta - \left\lfloor\frac{\Omega.\delta}{W_L}\right\rfloor W_L}{\delta}$, where $\frac{W_H - W_L}{\delta}$ reflects the number of iterations in the inner loop and the term $(\Omega.\delta - \lfloor(\Omega.\delta)/W_L\rfloor W\_L)/\delta$ captures the case the block size is not a multiple of W_L.

4.2. Best segmentation option

As explained in the previous subsection, AUDAS explores the various segmentation options within a batch and associates a score for each of the considered segments. Basically, segments are ranked based on their utility for the application. To capture the feature relevance, AUDAS employs a ML model that is to be trained offline. As shown in Fig. 3, the model will be consulted for each possible segment. The model is to generate a score, σ, that reflects the feature relevance. In AUDAS, σ is taken to be the probabilistic confidence of the output of the ML classifier. This fits scenarios where the classifier maps the segment to one a discrete set of classes with probabilistic ranges for each segment. The score in this case reflects how closely the features of the segment match those of the picked class, and hence a higher confidence implies better segment score. To illustrate, we can apply Support vector machines (SVM), and use the probabilistic score for each class, or alternatively capture the result of a softmax function at the output of a neural network to associate a confidence level of the classification. In other words, if the telehealth system is monitoring the heart or lung of a patient, the relevance of the features provided by a segment for assessing the conditions of such an organ is the ranking criterion. As shown in Fig. 3, the model will be consulted for each possible segment. The question that remains is how to pick the best set of segments. Two clear requirements can be pointed out: (i) the picked segments should not overlap, and (ii) the set of segments should cover the entire data batch.

While quite a few selection criteria could be considered, AUDAS favors maximizing the average segment score as an objective of the segment set selection optimization. The rationale is that the entire data batch can be covered by sets of different cardinality. Maximizing the average score implies the incorporation of the fewest and most indicative segments. To illustrate, let us compare two sets X and Y with scores {0.75, 0.70, 0.80} and {1., 0.40, 0.30, 0.70}, respectively. While Y yields the best total score, it includes segments with low scores. The segments in set X, on the other hand, have consistently high scores. The average score for X and Y is 0.75 and 0.6, respectively, which makes X preferable. We note that the variance will not be a better metric than average score since it may yield a set of segments with consistently low scores.

The segment selection optimization is also a combinatorial problem; yet the two aforementioned requirements, specifically non-overlap and coverage, enable pursuing an efficient solution. AUDAS models the segments as vertices in a graph and captures adjacency (consecutiveness in time) among segments using edges. The weight of an outbound edge of a vertex reflects the score (σ) of the corresponding segment. By using such a graph model, the selection optimization becomes finding the best (least average cost) path from the beginning of the batch to its end. Numerous algorithms can be employed to solve such a shortest path in a graph problem with runtime complexity of $\Theta(E + \nu \log\nu)$, where ν is the number of segments and E is the number of edges in the graph (Sniedovich, 2006). Fig. 4 explains the segment selection process through an example where two segmentation options are considered. The example shown

in Fig. 4 is mainly for illustration purposes, in practice the number of segments created with AUDAS will be larger. In the figure, a link exists between each pair of consecutive segments, e.g., ξ_1 and ξ_2. The presence of multiple outgoing links implies multiple segment sequences, and hence multiple segmentation options for the data batch. Furthermore, the links from segment ξ'_4 to ξ_5 and to ξ'_5 in the graph reflect adjacency as ξ'_4 ends at 30, where both ξ'_5 and ξ_5 start. Clearly the number of possible segmentations depends on W_H, W_L and σ. AUDAS strives to provide a general solution to determine the most rewarding segmentation for accurate diagnosis of respiratory diseases.

5. Validation experiments

To validate the effectiveness of AUDAS, we consider a case study of detecting acute symptoms that are indicative of possible COPD exacerbations. In this section, we provide the details of the study and present the obtained results.

5.1. COPD detection classifier

Data Collection: We used a dataset of lung diseases collected from real patients. The dataset, namely, ICBHI (Rocha et al., 2019), contains labeled respiratory diseases including: COPD, Bronchiectasis, Asthma, upper and lower respiratory tract infection, Pneumonia, Bronchiolitis. The dataset also includes respiratory sounds, with categorized acute symptoms such as Crackle, Wheeze, Both (Crackle & Wheeze), and Normal. The data is for 128 patients (64 with COPD) and 920 audio recordings, for a total of 5.5 h with sampling frequency that vary between 4 and 44.1 kHz. The data has been segmented according to the breathing cycle. In practice, such segmentation is not provided in the telehealth system and typically involves further investigation by a specialty provider, e.g., pulmonologist or respiratory therapist. Although technique of signal processing can be used for that purpose, the variability of the signal prevents such techniques from converging to the most useful segmentation for anomaly detection. For training, we carefully selected 9 subjects (patients) to assess the quality of the overall segmentation. The selection criterion is based on the balance between COPD patient and normal subjects while considering all their respective annotated segments. The remaining subjects are further separated into two sets; one is used for the breathing/non-breathing classification and the other is used for COPD diagnosis. Within each category we have used cross validation to provide the final results for the classifiers. Such a split of the dataset mitigates the effect of any overlap between the training and test. To assess the effectiveness of AUDAS in handling noisy data, we augment the training dataset with incoherent segments that correspond to incorrect cycles of crackles and wheezes.

Feature Extraction: It has been shown that spectral features enable the distinction among cough sounds (Pramono et al., 2016). We leverage such a finding to distinguish among indicative and inconclusive breathing sound segments. The following is the list of features used in the experiment: the energy peak of the signal envelope, the zero-crossing rate, the phase power ratio, the spectral centroid, skewness, roll-off, spread factor, bandwidth, kurtosis, spectral flatness. In addition, we have used the mean and standard deviation of the Mel-frequency Cepstral Coefficients (MFCC) and the Crest Factor. The overall dataset has been normalized. An SVM classifier is employed to score the segment options. The classifier is based on an RBF kernel where $\gamma = 2$ and $C = 1$. The scores are computed like indicated in eq. (1).

Disease Classifier: For this, we construct a mel-spectrogram for COPD. As the measurements are of contact-based sensors, the energy of the signal is concentrated in different frequency bands. Thus, we applied signal transformations to keep 75 percent of the energy for each spectrogram. To detect COPD symptoms in each of the AUDAS-generated segments, we employ convolutional neural network (CNN) architecture. The CNN contains four convolutional layers with kernel size 3×3. The convolutional layers contain respectively 64, 32,28, 64 neurons. We then introduce a nonlinear layer using Relu for activation, while the last layer applies a Softmax function with two outputs. We use max-pooling layers in between the convolution layers. The classifier architecture is depicted in Fig. 5.
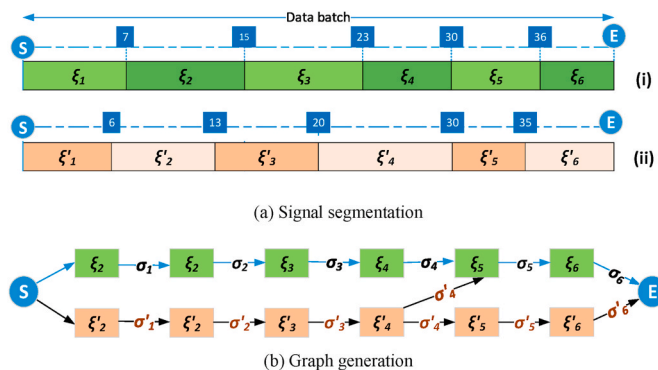

(a) Signal segmentation

(b) Graph generation

Fig. 4. Illustrating the modeling of the various segmentation options as a graph to enable the selection of the best set of segments.
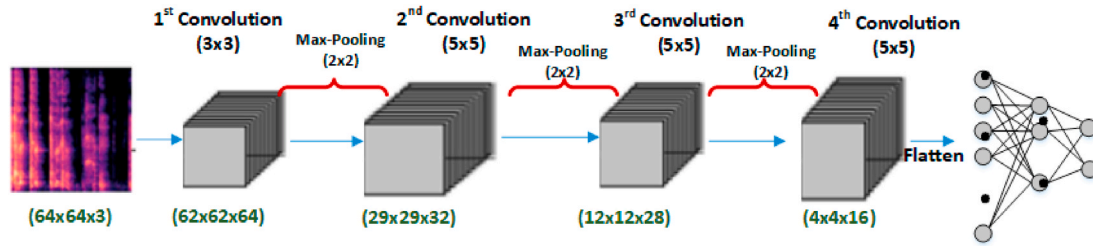
Fig. 5. Illustrating the design of the employed CNN classifier for COPD diagnosis.

5.2. Metrics and results

First, we validate the performance of the scoring of individual segments. Fig. 6 shows the performance of the underlying SVM classifier for detecting valid segment and discriminating crackles, wheezes and normal sounds from invalid ones. The performance is reported in terms of accuracy, precision, recall and F1 score. The plot reflects scores for actual breathing segments in the dataset and the incoherent ones. The classifier performance for incoherent segments is marked as "NonBreath" in the figure. The results in Fig. 6 demonstrate the effectiveness of AUDAS in detecting breathing cycles, with performance exceeding 90% for all metrics. We have also assessed the utility of the segments generated by AUDAS in respiratory disease diagnosis, specifically COPD. Fig. 7 reports the performance of the CNN architecture employed to detect COPD. As seen in the figure, the CNN-based COPD classifier achieves outstanding results in terms of accuracy, precision, recall and F1-score. We later compare AUDAS with fixed segmentation.

Next, we evaluate the effectiveness of AUDAS' dynamic adjustment of the window size. The performance is assessed in terms of the accuracy of the generated segments relative to the ground truth (annotated segments in the dataset). The performance is also compared to segmentation using fixed window settings with various window sizes. Basically, we use the annotated segments as an input to the CNN classifier and note the COPD detection accuracy. We then do the same for segments generated by AUDAS and the fixed window (static) approach. The latter simply reflects running our CNN model against fixed size segments. Fig. 8 reports the COPD classification accuracy relative to that achieved using the annotated segments, where a value of one implies that the real-time segmentation is perfectly matching reality. We note that for this figure, we have resampled the sound signals so that the segmentation is based only on 4 KHz sampling rate. Such a step is motivated by the fact that the dataset is based on mixed sampling rates. The largest window size in the experiment is 5 s, which corresponds to the largest inspiration and expiration cycle as is pointed out in (Nuckowska et al., 2019). Hence, the largest window size in figure (5 sec.) corresponds to 20,000 samples. Fig. 8 show that AUDAS significantly outperforms the version of the COPD classifier that uses fixed window sizes. The results show diminished relative accuracy for increased window sizes. This is attributed to the fact that actual segment sizes vary and could be much smaller than the fixed window; consequently, the segmentation accuracy degrades. Yet with respiratory illness the breathing cycle length is irregular and the dynamic approach of AUDAS is deemed more appropriate as confirmed by the results. For normal conditions AUDAS is found to produce relatively stable segment sizes.

As explained in Section 4, AUDAS employs dynamic programming in order to provide the most insightful approximation while considering the least execution time. To assess the divergence from the exact segmentation and the effectiveness of our objective function over many segments, we tracked the accumulated error trends, measured using mean absolute error (MAE), for the static (fixed size) segmentation and that of AUDIS. Based on the results in Fig. 9, AUDAS achieves the best MAE due the irregularity of the segmentation and inability to static segmentation to capture such variability.

Fig. 10 demonstrates the performance of AUDAS in terms of anomaly detection accuracy while varying the signal to noise ratio (SNR). To study the impact of the noise, we have considered recordings with −50 to 15 dB of SNR range. We partitioned the data according to the SNR for each segment, and then measured the accuracy on the COPD detection for each segment. The figure shows the observed accuracy for the different SNR ranges. Clearly, AUDAS achieves a major gain in accuracy relative to fixed-size segmentation.
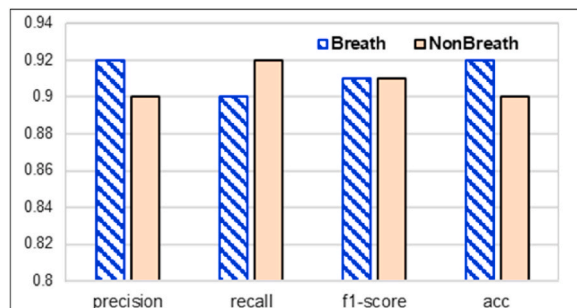

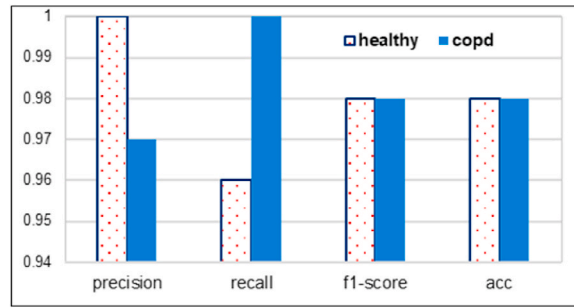**Fig. 6.** Assessment metric for the scoring function of AUDAS.

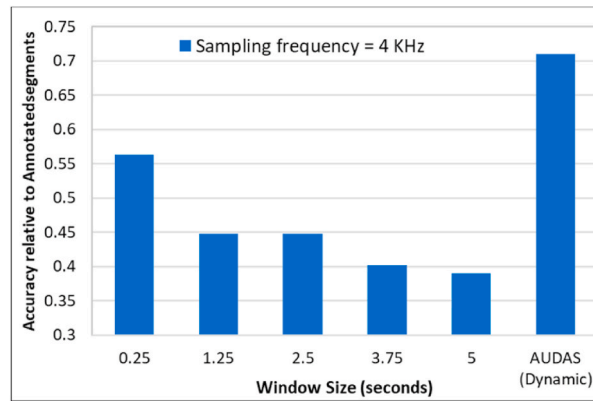**Fig. 7.** COPD Detection based on segments generated by AUDAS.

Fig. 8. The effect of window size on the segmentation accuracy.
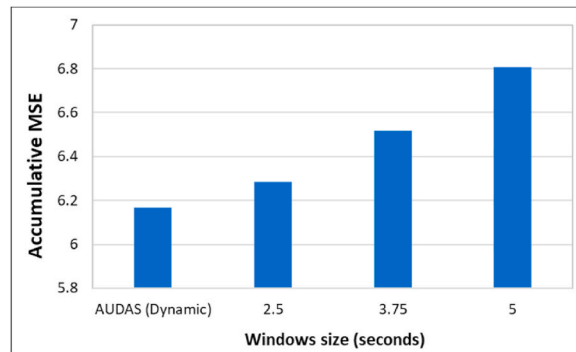

Fig. 9. AUDAS′ dynamic setting of the window size minimizes the segmentation error.

Such a gain scales when the SNR increases, given the positive impact of the signal quality (features) on the diagnostics. As expected, a high SNR enables AUDAS to match the accuracy to the annotated data.

We again note that AUDAS enables not only the detection of diseases but also the extraction of acute symptoms. The latter can be used to assess severity and track progression. Published approaches for detecting lung diseases, e.g., (Fraiwan et al., 2021), do not have such capability as strong as AUDAS. We have conducted additional experiments to highlight such an advantage. Fig. 11 studies the performance of AUDAS in terms of accurate detection of abnormal chest sounds detection. The results in the figure assess the detectability of abnormal sound patterns, e.g., wheezes, crackles, etc., for AUDAS in comparison with the approach of Fraiwan et al. (Fraiwan et al., 2021). The underlying classifier used in such an approach is based on VGG16 (a CNN that is 16 layers deep) and assumes fixed window sizes. The results of AUDAS in that figure are based on feeding the generated (dynamic) segments to the classifier of (Kim et al., 2021b). We partitioned the data according to AUDAS segmentation algorithm and compared the detection of symptoms using (Kim et al., 2021b). As indicated by the figure the dynamic segmentation has improved the detectability of abnormal chest sounds. As illustrated in Fig. 11, AUDAS achieves a major gain in accuracy relative to fixed-size segmentation. The effectiveness
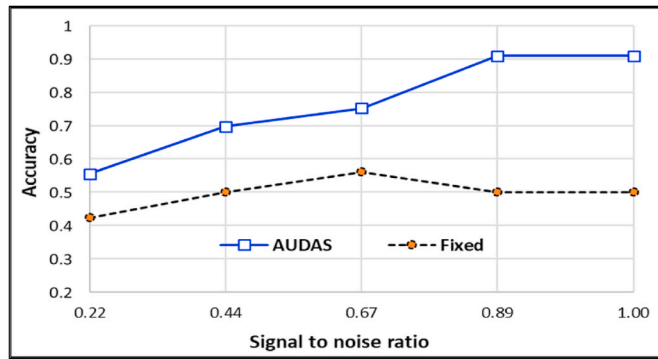
Fig. 10. The effect of SNR on COPD detection accuracy for AUDAS as compared to annotated (ground truth) and fixed segmentation sizes.
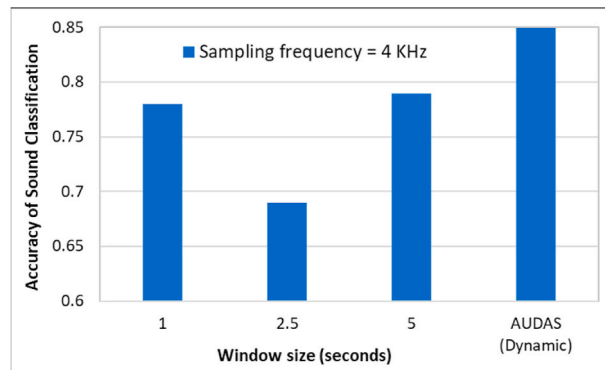

Fig. 11. The effect of the segmentation on the accurate detection of abnormal chest sound.

of AUDAS in detecting lung diseases and abnormal symptoms (sounds) is achieved due to its context aware segmentation.

To assess the suitability of AUDAS for real-time processing of streamed sound data, we have used an Arduino platform that is clocked at 16 MHz. The detection of normal/abnormal using AUDAS is estimated to be 0.3 s. We note that more capable embedded platforms, e.g., Arduino DUE, could achieve more than 5 times reduction in execution time. Prior work, e.g., (Nuckowska et al., 2019), has pointed out that an inspiration and expiration cycle takes as much as 5 s and as little as 2 s. Hence, AUDAS can robustly meet the timing constraint, i.e., complete processing one cycle before the end of the next cycle. In other words, AUDAS can be applied in real-time. The whole algorithm including the CNN-based disease diagnosis took about 26.84s using Arduino; we note, however, that the CNN classifier is to be executed on the cloud and hence its execution can be expedited massively. To compare the computational overhead with competing dynamic segmentation schemes, we have implemented the approach of (Fraiwan et al., 2021), which also employs a CNN model. We have assessed the computational complexity in terms of the number of parameters, where our CNN model for COPD detection is found to involve 34,542 parameters compared to 191,042 parameters for the baseline approach; in other words, the complexity of AUDAS is approximating 5 times less.

6. Conclusions and future work

The COVID-19 pandemic has brought respiratory illness to the spotlight and highlighted the importance of telehealth systems. Analyzing the lung sound and monitoring the breathing rate are the conventional means for detecting and tracking symptoms of respiratory diseases. Automating such a process in a telehealth system requires attaching acoustic sensors to the patient's chest and recognizing the breathing cycles within the lung sound signal. However, determining the breathing cycles in real-time is complicated by the fact that they vary over time depending on the patient's conditions and activities. This paper has presented, AUDAS, a novel context-aware segmentation mechanism. AUDAS applies a scoring scheme to qualify the segment based on the targeted respiratory illness. To overcome the variability of the breathing rate, AUDAS pursues a dynamic window size to detect segment boundary and applies a dynamic programming method to determine the best segmentation of a lung sound recording. We have demonstrated the effectiveness of AUDAS via experiments using a published COPD dataset. In the future, we plan to study the applicability of AUDAS to other sensor modalities and investigate the possibility of using disease-based classification to fine tune the segment boundary recognition module within AUDAS. In addition, we envisage to adapt our segmentation algorithm to personalized diagnostics, e.g., personalize the window size.

Credit author statement

Wassila Lalouani: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - Original Draft. **Mohamed Younis:** Conceptualization, Methodology, Formal analysis, Writing - Review & Editing, Supervision, Funding acquisition. **Lloyd E. Emokpae:** Conceptualization, Methodology, Formal analysis, Writing - Review & Editing, Supervision, Funding acquisition. **Roland N. Emokpae:** Data Curation, Validation, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

References

Ahmad, S., Lavin, A., Purdy, S., & Agha, Z. (2017). Unsupervised real-time anomaly detection for streaming data. *Neurocomputing, 262*, 134–147.

Bosse, S., Koerdt, M., & Schmidt, D. (2018). Robust and adaptive signal segmentation for structural monitoring using autonomous agents. *Proceedings, 2*(No. 3), 105.

Cohen-McFarlane, M., Goubran, F., & Knoefel, R. (2019). Comparison of silence removal methods for the identification of audio cough events. In *Proceeding of Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1263–1268). Jul.

Demir, F., Ismael, A. M., & Sengur, A. (2020). Classification of lung sounds with CNN model using parallel pooling structure. *IEEE Access, 8*, 105376–105383.

Emokpae, L. E., Emokpae, R. N., Jr., Bowry, E., Saif, J. B., Mahmud, M., Lalouani, W., … Joyner, R. L., Jr. (2021). A wearable multi-modal acoustic system for breathing analysis. Special Issue on COVID-19 Pandemic Acoustic Effects *Journal of Acoustical Society of America, 151*, 1033–1038. https://doi.org/10.1121/10.0009487.

Emokpae, L. E., et al. (2022). A wearable multi-modal acoustic system for breathing analysis. *Journal of the Acoustical Society of America, 151*(No. 2), 1033–1038.

Forum of International Respiratory Societies. (2017). *The global impact of respiratory disease.* Sheffield: European Respiratory Society, 2nd Edition, ISBN: 9781849840873.

Fraiwan, M., et al. (2021). Recognition of pulmonary diseases from lung sounds using convolutional neural networks and long short-term memory. *Journal of Ambient Intelligence and Humanized Computing.* https://doi.org/10.1007/s12652-021-03184-y. Apr.

Grossman, P. (2004). The LifeShirt: A multi-function ambulatory system monitoring health, disease, and medical intervention in the real world. *Studies in Health Technology and Informatics, 108*, 133–141.

Gu, F., Niu, J., Das, S. K., He, Z., & Jin, X. (2017). Detecting breathing frequency and maintaining a proper running rhythm. *Pervasive and Mobile Computing, 42*, 498–512.

Gupta, P., Wen, H., Di Francesco, L., & Ayazi, F. (2021). Detection of pathological mechano-acoustic signatures using precision accelerometer contact microphones in patients with pulmonary disorders. *Scientific Reports, 11*, 13427.

Gurung, A., Scrafford, C. G., Tielsch, J. M., Levine, O. S., & Checkley, W. (2011). Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: A systematic review and meta-analysis. *Respiratory Medicine, 105*(No. 9), 1396–1403.

Haddad, A. E., & Najafizadeh, L. (2019). Source-informed segmentation: A data-driven approach for the temporal segmentation of EEG. *IEEE Transactions on Biomedical Engineering, 66*(No. 5), 1429–1446. May.

Hall, J. I., Lozano, M., Estrada-Petrocelli, L., Birring, S., & Turner, R. (2020). The present and future of cough counting tools. *Journal of Thoracic Disease, 12*(No. 9), 5207–5223.

Hsiao, C.-H., et al. (2020). Breathing sound segmentation and detection using transfer learning techniques on an attention-based encoder-decoder architecture. In *2020 42nd annual international conference of the IEEE engineering in medicine & biology society (EMBC)* (pp. 754–759).

Jácome Cristina, et al. (2019). Convolutional neural network for breathing phase detection in lung sounds. *Sensors (Basel, Switzerland), 19*(No. 8), 1798. Apr.

Kim, Y., Hyon, Y., Jung, S. S., et al. (2021a). Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning. *Scientific Reports, 11*, #17186.

Kim, Y., Hyon, Y., Jung, S. S., et al. (2021b). Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning. *Scientific Reports, 11*, 17186.

Kong, Q., Xu, Y., Sobieraj, I., Wang, W., & Plumbley, M. D. (2019). Sound event detection and time–frequency segmentation from weakly labelled data. *IEEE/ACM Trans. Audio, Speech and Lang. Proc., 27*(4), 777–787. April.

Martín-Morató, I., Harju, M., & Mesaros, A. (2021). Crowdsourcing strong labels for sound event detection. In *Proc. of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)* (pp. 246–250).

Massaroni, C., Nicolò, A., Lo Presti, D., Sacchetti, M., Silvestri, S., & Schena, E. (2019). Contact-based methods for measuring respiratory rate. *Sensors (Basel), 19*(No. 4), 908. Feb.

Ma, Y., Xu, X., & Li, Y. (2020). LungRN+NL: An improved adventitious lung sound classification using non-local block ResNet neural network with mixup data augmentation. In *Proc. of Interspeech.*

Moussavi, Z. (2006). Fundamentals of respiratory sounds and analysis. *Synthesis Lectures on Biomedical Eng., 1*(No. 1), 1–68.

Mukherjee, H., et al. (2021). Automatic lung health screening using respiratory sounds. *Journal of Medical Systems, 45*(No. 19).

Nicolò, A., Massaroni, C., Schena, E., & Sacchetti, M. (2020). The importance of respiratory rate monitoring: From healthcare to sport and exercise. *Sensors (Basel), 20*(No. 21), 6396.

Nuckowska, M. K., Gruszecki, M., Kot, J., et al. (2019). Impact of slow breathing on the blood pressure and subarachnoid space width oscillations in humans. *Scientific Reports, 9*(No. 1), 6232.

Oletic, D., & Bilas, V. (2016). Energy-efficient respiratory sounds sensing for personal mobile asthma monitoring. *IEEE Sensors Journal, 16*, 8295–8303.

Paul, R., & Shoukat Choudhury, M. A. A. (2015). Novel data segmentation methods for data driven process analyses. *Computer Aided Chemical Engineering, 37*, 1727–1732.

Pellegrino, R., et al. (2005). Interpretative strategies for lung function tests. *European Respiratory Journal, 26*(No. 5), 948–968.

Pramono, R. X., Imtiaz, S. A., & Rodriguez-Villegas, E. (2016). A cough-based algorithm for automatic diagnosis of pertussis. *PLoS One, 11*(9), e0162128. https://doi.org/10.1371/journal.pone.0162128. Sep 1 PMID: 27583523; PMCID: PMC5008773.

Rao, A., Huynh, E., Royston, T. J., Kornblith, A., & Roy, S. (2018). Acoustic methods for pulmonary diagnosis. *IEEE Reviews in Biomedical Engineering, 12*, 221–239.

Rao, A., Huynh, E., Royston, T. J., Kornblith, A., & Roy, S. (2019). Acoustic methods for pulmonary diagnosis. *IEEE Rev Biomed Eng., 12*, 221–239.

Rocha, B. M., et al. (2019). An open access database for the evaluation of respiratory sound classification algorithms. *Physiological Measurement, 40*(No.3), 35001. Mar.

Smith, J., & Woodcock, A. (2008). New developments in the objective assessment of cough. *Lung, 186*, 48–54.

Sniedovich, M. (2006). Dijkstra's algorithm revisited: The dynamic programming connexion. *Control and Cybernetics, 35*, 599–620.

Srivastava, A., Jain, S., Miranda, R., Patil, S., Pandya, S., & Kotecha, K. (2021). Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. *PeerJ. Computer science, 7*, e369.

Turner, R. D., & Bothamley, G. H. (2014). How to count coughs? Counting by ear, the effect of visual data and the evaluation of an automated cough monitor. *Respiratory Medicine, 108*(No. 12), 1808–1815.

Vizel, E., et al. (2010). Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions. *Cough, 6*(#3). May.