BRAIN. Broad Research in Artificial Intelligence and Neuroscience

e-ISSN: 2067-3957 | p-ISSN: 2068-0473

Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science Evaluation Center, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)

2024, Volume 15, Issue 3, pages: 188-196.

Submitted: July 2nd, 2024| *Accepted for publication:* September 20th, 2024

Lung Sounds Anomaly Detection with Respiratory Cycle Segmentation

Constantin Constantinescu [1]
Remus Brad [2]
Adrian Bărglăzan [3]

[1] Lucian Blaga University of Sibiu, Romania
ctin.constantinescu@ulbsibiu.ro
https://orcid.org/0000-0002-9743-9794
[2] Lucian Blaga University of Sibiu, Romania
remus.brad@ulbsibiu.ro
https://orcid.org/0000-0001-8100-1379
[3] Lucian Blaga University of Sibiu, Romania
adrian.barglazan@ulbsibiu.ro
https://orcid.org/0009-0009-4420-3639

Abstract: *Employing machine learning algorithms in the medical field has proven successful for some time now. Mostly computer vision techniques have been applied to medical images, while medical sound data has been somewhat overlooked. By using electronic stethoscopes, it is now possible to process both heartbeats and lung sounds. While some products are available for detecting anomalies in heartbeats, addressing lung-related anomalies presents a more intricate challenge. Applying a deep learning approach is hindered by insufficient data. Although some datasets do exist, the size and diversity of the data are too small for comprehensive analysis. This paper introduces a novel technique for detecting anomalies in lung sounds: first by combining two datasets, second by automatically segmenting each sound into respiratory cycles, and third by employing GFCCs as sound features.*

Keywords: *anomaly detection; respiratory sounds; deep learning*

How to cite: Constantinescu, C., Brad, R., & Bărglăzan, A. (2024). Lung sounds anomaly detection with respiratory cycle segmentation. *BRAIN: Broad Research in Artificial Intelligence and Neuroscience*, 15(3), 188-196.

https://doi.org/10.70594/brain/15.3/14

1. Introduction

Sound anomaly detection has received significant attention recently, particularly in industrial applications for predictive maintenance of equipment and automatic quality assurance of products, showing promising results. In the medical field, computer vision has predominantly been utilized for tasks such as tumor and mass detection in medical images like CT scans and ultrasound images, while other areas where artificial intelligence could be beneficial have been overlooked. When considering sound within the medical activity, the stethoscope comes to mind. Owned and used by every medical doctor to listen to heartbeats or respiratory sounds, the stethoscope has not been left out by the progress of modern technology, giving rise to electronic versions.

Some modern devices make it possible for physicians to record the sounds they hear, opening the door to automated sound processing. Some electronic stethoscopes already integrate heartbeat anomaly detection. However, in the case of respiratory sounds, the problem is more complex due to the various events occurring during respiration, typically at low frequencies. Therefore, automatic anomaly detection in respiratory sounds becomes a challenging problem, with deep learning having the best chance of offering promising results. However, deep learning necessitates a substantial amount of data for effective training, which is currently lacking. Although some datasets are available, they contain insufficient data for comprehensive model training.

Therefore, this paper proposes a solution where two datasets, containing different data represented in distinct ways, are combined. The objective is to unify the data into a standardized format, addressing the challenge of insufficient training data for deep learning in the context of automatic respiratory sound anomaly detection.

2. Related work

With the subject being a task in a competition, some work has been done during the ICBHI Challenge event. The paper (Siddhartha, Tom, Kwatra, & Jain, 2021) introduces RespiraNET, a neural network designed for anomaly detection in respiratory sounds. To overcome the challenge of limited data, the authors suggest various methods of data augmentation, along with a transfer learning technique. Their experiments yielded promising results in two-class classification, achieving an ICBHI score of 77%, calculated as the mean of sensitivity and specificity.

Another paper (Dar, Srivastava, & Mishra, Lung anomaly detection from respiratory sound database (sound signals), 2023) discusses the use of two datasets in the development of a lung sound anomaly detection algorithm. Upon closer examination of the datasets, they appear to be remarkably similar, if not identical, despite being downloaded from different sources. Disregarding this aspect, the authors introduce a complex method that involves extracting more than ten features and training a 9-layer model. Their algorithm's performance, measured in terms of True Positive Rate (TPR), True Negative Rate (TNR), and Testing Accuracy, yields the best results with corresponding rates of 0.963, 0.932, and 0.948. The same authors also published another technique (Dar, Srivastava, & Lone, Spectral features and optimal Hierarchical attention networks for pulmonary abnormality detection from the respiratory sound signals, 2022) using Bark Frequency Cepstral Coefficients and Hierarchical Attention Networks using the same two datasets.

Most researchers typically treat this task as a classification problem with multiple classes, assuming sufficient data availability for each class. An anomaly detection approach is taken by the authors in this paper (Cozzatti, Simonetta, & Ntalampiras, 2022). They extracted MFCCs from the ICBHI Database (Rocha, et al., 2019) and utilized an autoencoder for classification. Notably, they adopted a ”weakly-supervised” method, training the model only on normal recordings. In this case, the best True Positive Rate (TPR), True Negative Rate (TNR), and accuracy (ACC) achieved were 0.58, 0.61, and 0.60, respectively.

A comprehensive analysis was conducted by the authors of this paper (Pham, Phan, Palaniappan, Mertins, & McLoughlin, 2021). To determine the most effective feature, various spectrograms such as Log-Mel, Gamma, or CQT were extracted from the sound. These spectrograms were then divided into fixed-sized patches and input into a Convolutional-Dense Neural Network (C-DNN) based model. The optimal results were observed with the Log-Mel and Gamma spectrograms, with a slight advantage in favor of Gamma in the two-category classification of respiratory cycles. With the newly discovered information, the authors propose a robust framework. When applied to a respiratory cycle anomaly detection task, this framework achieves a Sensitivity, Specificity, and ICBHI Score of 0.90, 0.78, and 0.84.

The paper (Sengupta, Sahidullah, & Saha, 2016) demonstrates how cepstral-based statistical features perform better than wavelet-based features, while (Manzoor, et al., 2020) focuses on getting better results using a variation of a recurrent neural network. Also focusing on improving the learning model, (Fernando, Sridharan, Denman, Ghaemmaghami, & Fookes, 2022) use a temporal convolution network for the job. (Senthilnathan, Deshpande, & Rai, 2020) have a slightly different approach, trying to detect anomalies in breathing sounds directly recorded with a smartphone, not a professional electronic stethoscope. Also, tackling the problem of insufficient data for training, (Le, Bang, Le, & Choo, 2023) propose a lightweight model that employs a technique called feature fool exploitation to detect the anomalies.

3. Datasets used in the study

3.1. ICBHI 2017 Challenge Database

The ICBHI dataset was created for the scientific challenge organized at the International Conference on Biomedical Health Informatics 2017 (ICBHI 2017 Challenge, n.d.). It is a respiratory sound database containing audio samples collected by two research teams, one from Portugal and the other from Greece. The sounds were recorded using either electronic stethoscopes from various manufacturers or a microphone array. Recordings were obtained from 126 patients and had a total duration of 5.5 hours. Each recording varies in duration, ranging from 10 to 90 seconds. Respiratory experts labeled the data, establishing respiratory cycles and identifying the presence of events such as crackles and wheezes. This information is stored in a separate annotation file for each recording. The recording setup details, including information about the device, stethoscope position, and patient number, are encoded in the file names.

Additionally, a separate file provides the diagnosis for each patient. The dataset encompasses both healthy patients and those with conditions such as Asthma, Pneumonia, Bronchiectasis, Chronic Obstructive Pulmonary Disease, Lower Respiratory Tract Infection, or Upper Respiratory Tract Infection. While demographic information for each patient is available in the dataset, it is not utilized in this paper. (Rocha, et al., 2019)

3.2. A dataset of lung sounds

Without having a particular name, recordings from 112 patients are included in this dataset, containing only one recording per subject (Fraiwan, Fraiwan, Khassawneh, & Ibnian, 2021). The patients are of various ages, with an almost equal number of men and women, and detailed demographic information is provided. Each recording varies in length, ranging in this case from 5 to 30 seconds.

For every recording, three types of filters are available. In this instance, we exclusively processed recordings with diaphragm mode filtration, as it is the most commonly used by doctors in their daily activities. The data is labeled, providing information about the presence of respiratory events, such as crackles and wheezes, as well as the final diagnosis.

Among the patients, 35 are healthy subjects, while 77 have been diagnosed with conditions including asthma, heart failure, pneumonia, bronchitis, pleural effusion, lung fibrosis, or COPD. In contrast to the ICBHI dataset, the respiratory cycles are not determined in this case. Both the annotation data and the setup details for each patient are encoded in the file names and also structured in a separate sheet. A single type of recording device was used specifically the Littmann 3200 Electronic Stethoscope. (Fraiwan, Fraiwan, Khassawneh, & Ibnian, 2021)

4. Preprocessing

In this paper, our primary focus is to automatically determine whether a person is healthy or sick. For this purpose, we needed to extract from the annotations only the information relevant to us. We processed the annotation files from each dataset, specifically extracting the diagnosis information. Patients labeled as *Healthy* or *Normal* were assigned a label of 0, while those diagnosed with any disease were assigned a label of 1.

Having recordings from two datasets and various recording devices, the sound data was recorded with different sample rates, ranging from 4kHz to 44.1 kHz. Consequently, we standardized each recording by resampling it to a rate of 4kHz. Additionally, each sound was divided into frames, with each frame containing 1024 samples with an overlap of 512 samples.

Sound segmentation plays a very important role in the success of our proposed method. On one hand, even with the combination of two datasets, the available data for training a neural network remains limited. By breaking down the recordings into smaller segments, we generate numerous smaller samples for training, as opposed to having fewer larger samples. This helps in the training of the neural network. On the other hand, each recording varies in length, ranging from 5 seconds to 90 seconds. Handling the recordings as they are is impractical, if not impossible, so segmenting them into smaller units provides uniformly sized samples for training.

The easiest way to segment each recording involves arbitrarily setting a segment length, calculating the duration of each recording, and dividing it by the chosen length, leaving out the rest of the division. However, with this method, there is a risk of splitting recordings in the middle of inspiration or expiration.

A more effective approach is to segment each recording into respiratory cycles. Although one of the datasets already includes this information, we believe that automatically segmenting respiratory recordings into cycles is essential. This is particularly important as our research aims to develop a generic algorithm capable of detecting anomalies in any respiratory audio data. Moreover, manually determined cycles have different lengths, which will make it difficult for us to process them later on. Nevertheless, we used this information to validate our automatic segmentation. Collaborating with a physician and using the available annotations, we established that the average duration of a complete respiratory cycle is 2.5 seconds.

The cycles are not readily visible or easily determinable in the classic waveform representation of the sound (Figure 1(a)) but can be distinguished in the power spectrogram, as shown in Figure 1(b).

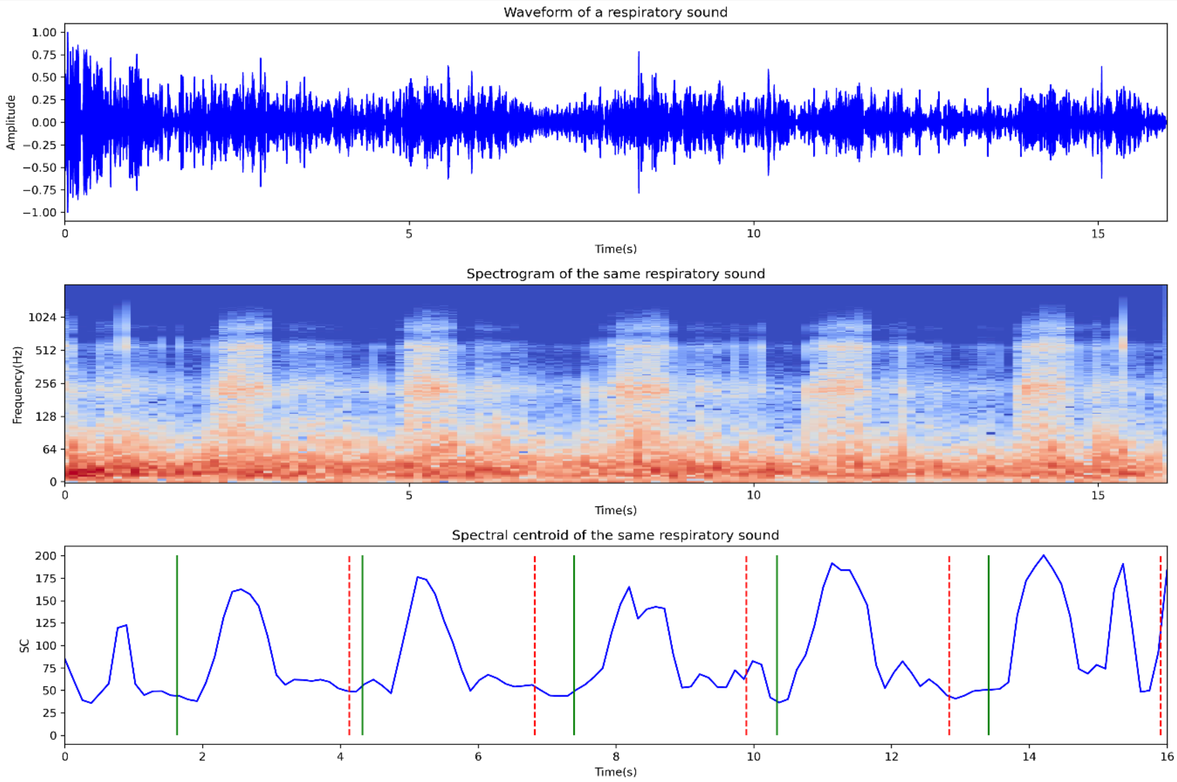

Figure 1. Respiratory sound representation with segmentation

Therefore, we employed the spectral centroid, a frequency domain feature in sound that indicates the so-called "center of gravity" of the magnitude spectrum. In other words, it shows the frequency band with the most energy. We calculate the spectral centroid for each frame *t*, doing essentially a weighted mean of the frequencies, where the weights are the magnitude of a certain frequency bin (Constantinescu & Brad, 2023), as expressed in the following equation (1):

$$SC_t = \frac{\sum_{n=1}^{N} m_t(n) \cdot n}{\sum_{n=1}^{N} m_t(n)} \quad (1)$$

The maximum values of the spectral centroid indicate the moments with the most energy in sound, which, in our case, correspond to the respiratory cycles. Upon observation, we noted that the peaks are not situated in the middle of the respiratory cycles but rather approximately 0.8 seconds after the start of inspiration and 1.7 seconds before the end of expiration. Using these established values, we extract the 2.5-second respiratory cycle around each spectral centroid peak. In Figure 1(c), a visual representation of the calculated spectral centroid, the continuous green vertical line indicates the start of the respiratory cycle, while the dashed red vertical line marks the end. If a spectral centroid peak is too close to the start or end of the recording, we exclude it, as it may result in an incomplete cycle that could disrupt our algorithm later. Additionally, since every patient is different, the respiratory rate is different, resulting in different values for the length of the respiratory cycles. We maintain our established 2.5-second cycle length. This may lead to some cycles overlapping slightly, a situation that will not lead to any issues for the future usage of a machine learning algorithm.

5. Experimental results

5.1. Feature extraction

After preprocessing the recordings and annotations, we ended up with a set of data vectors, each representing a sound with a fixed length of 2.5 seconds. Each vector contains one respiratory cycle and is labeled as 1 if it originates from a sick person, or 0 if it is from a healthy individual.

Despite opting for a deep learning approach in this experiment, which theoretically doesn't need feature extraction, the current state of the art suggests that, in the case of sound, better results are achieved by extracting Mel Frequency Cepstral Coefficients (MFCCs) from the sound and training the model with them (Huang, et al., 2023). Taking a step further, we are using Gammatone Frequency Cepstral Coefficients (GFCCs), a set of coefficients that have demonstrated excellent results in speech recognition (Liu, 2018). GFCCs are particularly useful for speech and music processing, as they can simulate the characteristics of the auditory system, being particularly effective in capturing the fine spectral details of the speech sounds that are important for speech processing. They have been used in various audio and speech processing tasks such as speech recognition, speech synthesis, and speaker identification. These coefficients have the potential to achieve similar success in our case due to their effectiveness in capturing the fine spectral details of the sounds.

GFCCs are a type of sound representation similar to MFCCs. While MFCCs are obtained by applying a Mel filter bank to convert the linear frequency scale of the signal to the Mel-scale, GFCCs employ Gammatone filter banks to be applied to the sound. This choice aims to better replicate how the human auditory system perceives sound. A Gammatone filter is designed to mimic the spectral shape of the human auditory system, simulating the behavior of the basilar membrane in the inner ear. This membrane is responsible for analyzing the frequency content of sounds.

We implemented the GFCCs feature extraction relying on the Spafe implementation (Malek, et al., 2023) (Malek, Spafe: Simplified python audio features extraction, 2023) and following these steps (Jeevan, Dhingra, Hanmandlu, & Panigrahi, 2017):

1. The power spectrum of the signal is calculated.
2. The power spectrum is multiplied by the Gammatone filter bank.
3. A non-linear rectification is performed on the absolute value of the signal using a root cubic operation.
4. Finally, the Discrete cosine transform (DCT) is applied.

We calculate the GFCCs maintaining the sample rate and the framing parameters from the preprocessing step, with a number of 20 ceps, resulting in a coefficient matrix for each respiratory cycle. All resulting matrices have the same dimensions and are prepared to be used for training the model.

5.2. Experiment description

As previously mentioned, we merged the two available datasets. Next, we want to divide the combined dataset into three primary subsets: the training set, the testing set, and the prediction set.

Right from the beginning, we randomly selected 4 healthy patients and 4 patients with a respiratory condition from both of the original datasets to constitute the prediction set. These individuals are excluded from the dataset throughout the entire process and are only used at the end to make predictions about them. The remaining dataset is further divided into training and testing, with a ratio of 70/30.

Although the biggest challenge in general anomaly detection lies in achieving satisfactory results through an unsupervised approach, as anomalous data is challenging to record (for instance, in the case of industrial machines where anomalies in functionality occur rarely and at unpredictable times), we believe a supervised approach is more fitting and realistic for this specific task. This is because most patients seeking medical assistance are indeed ill. Consequently, the recording of anomalous lung sounds does not represent a significant problem and it would be unfortunate not to use them for training purposes.

Having prepared the data for training, we designed a deep-learning model for our neural network. As this paper is mainly focused on the preprocessing steps and the features that we extract from the sound, we went for a simple deep-learning architecture. The model is sequential and features one Flatten input layer, that converts the multidimensional input data to a one-dimensional array. The next three hidden layers are Dense layers with 512, 256 respectively 64 units, activated by the RELU function. Finally, the output layer is also a dense one, activated by the sigmoid function, since we are dealing with a binary classification problem. The model underwent optimization using the Adam optimizer, as for calculating the loss we used the binary-cross entropy function. We trained the network for one hundred epochs.

5.3. Results

With the model trained, it is time to use it for predictions on new, previously unseen data, using the 8 patients that we initially excluded from the dataset. After subjecting their recordings to the same preprocessing steps, we obtained a total of 219 respiratory cycles. Our previously trained model correctly predicted the class for 212 of them and incorrectly predicted 7. If we circle back from respiratory cycles to patients and establish a patient's class based on the majority vote of its respiratory cycles, we achieve a 100% correct prediction rate for these 8 patients.

To facilitate comparison with other proposed methods addressing this problem, we used the same four metrics that commonly appear in other papers:

- TPR – True positive rate (Sensitivity)
- TNR – True negative rate (Specificity)
- Accuracy
- ICBHI Score calculated as follows (2):

$$ICBHIS = \frac{TPR+TNR}{2} \quad (2)$$

Table 1 presents the results both per cycle and per patient using the aforementioned metrics.

Table 1: Prediction results per cycle and per patient, compared to other methods

Method	TPR (Sensitivity)	TNR (Specificity)	ACC	ICBHIS
RespireNet[1]	0.81	0.73	-	0.77
FrWCSO[2]	0.96	**0.93**	0.94	**0.94**
VAE[4]	0.58	0.61	0.60	0.59
CNN-MoE[6]	0.90	0.78	-	0.84
Ours(per cycle)	**1.00**	0.68	**0.96**	0.84
Ours(per patient)	**1.00**	**1.00**	**1.00**	**1.00**

6. Limits and Discussions

The relatively lower True Negative Rate (TNR) can be attributed to the dataset's imbalance, which has two primary causes. Firstly, grouping all diseases into one abnormal class led to a larger number of sick patients than healthy ones. Secondly, recordings from healthy patients are typically shorter than those from sick patients. This discrepancy may arise because physicians tend to verify the events they hear by placing the stethoscope in multiple locations on the patient's body or by listening for a longer time, resulting in more and longer recordings. Consequently, this leads to a higher number of respiratory cycles for the specific patient. To address this issue, data augmentation, for artificially generating more normal samples, could prove to be a potential solution. In comparison to other methods, our approach demonstrates superior outcomes. While the FRWCSO method closely resembles our results, the ambiguity of the dataset casts doubts on the relevance of their findings.

7. Conclusions

The primary objective of this research is to enhance respiratory sound anomaly detection. Our first contribution involves merging two distinct datasets and improving the deep learning model's performance by training it on a bigger dataset. Unlike most experiments that use a single dataset, our approach reflects a more realistic scenario. By combining datasets, we offer a more generic solution, aligning with our ultimate goal of deploying a lung sound anomaly detection system for use in the daily practices of healthcare professionals. The second major contribution lies in the automatic segmentation of recordings into respiratory cycles. Beyond helping in data preparation for training, this step also contributes to the development of a generic solution. While manual labeling might be more accurate, it is impractical for daily use. Lastly, utilizing Gammatone Frequency Cepstral Coefficients (GFCCs), a proven sound feature in speech processing, as a representation for our lung sound recordings, has a positive impact on the final results. The integration of these three improvements, combined with insights from previous research, has driven our algorithm to the promising results presented with relatively low complexity, marking a significant step toward a generic respiratory sound anomaly detection system.

Acknowledgments

This work was carried out during a research mobility at the Grupo de Tecnologia Medica y Audiovisual, an applied research group dedicated to the application of Information and Communication Technologies (ICTs) in biomedical research, education, and practice. The group is affiliated with the University Institute for Biomedical and Health Research (IUBS) at the University of Las Palmas de Gran Canaria (ULPGC).

We extend special thanks to Dr. Clara Peña, a specialist in family and community medicine, for providing valuable consultation on the medical aspects of this research.

References

Constantinescu, C., & Brad, R. (2023). An overview on sound features in time and frequency domain. International Journal of Advanced Statistics and IT&C for Economics and Life Sciences, 45-58. https://doi.org/10.2478/ijasitels-2023-0006

Cozzatti, M., Simonetta, F., & Ntalampiras, S. (2022). Variational autoencoders for anomaly detection in respiratory sounds. Artificial Neural Networks and Machine Learning – ICANN 2022, 333-345. https://doi.org/10.48550/arXiv.2208.03326

Dar, J. A., Srivastava, K. K., & Lone, S. A. (2022). Spectral features and optimal hierarchical attention networks for pulmonary abnormality detection from the respiratory sound signals. Biomedical Signal Processing and Control, 78, Article 103905. https://doi.org/10.1016/j.bspc.2022.103905

Dar, J. A., Srivastava, K. K., & Mishra, A. (2023). Lung anomaly detection from respiratory sound database (sound signals). Computers in Biology and Medicine, 164, Article 107311. https://doi.org/10.1016/j.compbiomed.2023.107311

Fernando, T., Sridharan, S., Denman, S., Ghaemmaghami, H., & Fookes, C. (2022). Robust and interpretable temporal convolution network for event detection in lung sound recordings. IEEE Journal of Biomedical and Health Informatics, 26(7), 2898-2908. https://doi.org/10.1109/JBHI.2022.3144314

Fraiwan, M., Fraiwan, L., Khassawneh, B., & Ibnian, A. (2021). A dataset of lung sounds recorded from the chest wall using an electronic stethoscope. Data in Brief, Article 106913. https://doi.org/10.1016/j.dib.2021.106913

Huang, D.-M., Huang, J., Qiao, K., Zhong, N.-S., Lu, H.-Z., & Wang, W.-J. (2023). Deep learning-based lung sound analysis for intelligent stethoscope. Military Medical Research, 10(1), Article 44.

ICBHI 2017 Challenge. (n.d.). Retrieved June 25, 2024, from https://bhichallenge.med.auth.gr/ICBHI_2017_Challenge

Jeevan, M., Dhingra, A., Hanmandlu, M., & Panigrahi, B. (2017). Robust speaker verification using GFCC based i-vectors. In Proceedings of the International Conference on Signal, Networks, Computing, and Systems (Vol. 395, pp. 85-91). Springer India. https://doi.org/10.1007/978-81-322-3592-7_9

Le, K.-N. T., Bang, S. Y., Le, D.-T., & Choo, H. (2023). Feature fool exploitation for lightweight anomaly detection in respiratory sound. In Future Data and Security Engineering. Big Data, Security and Privacy, Smart City and Industry 4.0 Applications (pp. 556-563). Springer. https://doi.org/10.1007/978-981-99-8296-7_40

Liu, G. K. (2018). Evaluating Gammatone Frequency Cepstral Coefficients with neural networks for emotion recognition from speech. arXiv. https://doi.org/10.48550/arXiv.1806.09010

Malek, A. (2023). Spafe: Simplified Python audio features extraction. Journal of Open Source Software, 8(82), Article 4739. https://doi.org/10.21105/joss.04739

Malek, A., Titeux, H., Borzì, S., Nielsen, C. H., Stöter, F.-R., Bredin, H., & Moerman, K. M. (2023). SuperKogito/spafe: v0.3.2. Zenodo. https://doi.org/10.5281/zenodo.7686438

Manzoor, A., Pan, Q., Khan, H. J., Siddeeq, S., Bhatti, H., Muhammd, A., & Wedagu, M. A. (2020). Analysis and detection of lung sounds anomalies based on NMA-RNN. In 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 2498-2504). IEEE. https://doi.org/10.1109/BIBM49941.2020.9313197

Pham, L., Phan, H., Palaniappan, R., Mertins, A., & McLoughlin, I. (2021). CNN-MoE based framework for classification of respiratory anomalies and lung disease detection. IEEE Journal of Biomedical and Health Informatics, 25(8), 2938-2947. https://doi.org/10.1109/JBHI.2021.3064237

Rocha, B., Filos, D., Mendes, L., Serbes, G., Ulukaya, S., Kahya, Y., De Carvalho, P. (2019). An open access database for the evaluation of respiratory sound classification algorithms. Physiological Measurement, 40(2), Article 106913. https://doi.org/10.1088/1361-6579/ab03ea

Sengupta, N., Sahidullah, M., & Saha, G. (2016). Lung sound classification using cepstral-based statistical features. Computers in Biology and Medicine, 75, 118-129. https://doi.org/10.1016/j.compbiomed.2016.05.013

Senthilnathan, H., Deshpande, P., & Rai, B. (2020). Breath sounds as a biomarker for screening infectious lung diseases. Engineering Proceedings, Article 8200. https://doi.org/10.3390/ecsa-7-08200

Siddhartha, G., Tom, F., Kwatra, N., & Jain, M. (2021). RespireNet: A deep neural network for accurately detecting abnormal lung sounds in limited data setting. In 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 527-530). IEEE. https://doi.org/10.1109/EMBC46164.2021.9630091