Challenges in Lung and Respiratory Sound Processing: Quantity and Quality of Available Data Akciğer ve Solunum Sesi İşlemedeki Zorluklar: Mevcut Verilerin Niceliği ve Niteliği

Damla Gürkan Kuntalp
Department of Electrical and Electronics Engineering, Dokuz Eylül University, Izmir, Türkiye
ORCIDs: 0000-0003-0617-7918
E-mails: damla.kuntalp@deu.edu.tr

***Abstract*—Respiratory diseases, both acute and chronic, are widespread due to exposure to harmful substances in the environment, workplace, and through personal behaviors. Furthermore, the COVID-19 pandemic has led to both short-term and long-term lung damage in survivors. Therefore, accurate identification of chronic respiratory diseases, in particular, is vital for effective management and treatment. Auscultation, the practice of listening to respiratory sounds, plays a crucial role in diagnosing respiratory diseases. By accurately interpreting these sounds, complemented by other clinical findings, specialists can make reliable diagnoses with minimal errors. However, the effectiveness of auscultation is heavily influenced by the doctor's experience and environmental noise. To address these limitations, automatic classification of respiratory sounds recorded with a digital stethoscope using expert software has emerged as a popular research area. This approach eliminates the reliance on subjective interpretation by specialists. Unfortunately, as with many biomedical signals, researchers face significant challenges. The most pressing issue is the need for high-quality, accurately labeled, and extensive lung and respiratory sound datasets. Additionally, removing noise that distorts these sound signals is another major obstacle. This brief review aims to delve into these two primary challenges and provide examples of potential solutions from relevant literature.**

Keywords*—lung sound; respiratory sound; data augmentation; noise removal

***Özetçe*—Çevresel faktörler, işyeri koşulları ve kişisel alışkanlıklar nedeniyle hem akut hem de kronik solunum hastalıkları sıklıkla görülmektedir. COVID-19 pandemisi ise, uzun vadeli sağlık sorunlarına yol açarak solunum sağlığı üzerinde ek bir yük oluşturmuştur. Bu bağlamda, özellikle kronik solunum hastalıklarının doğru teşhisi ve etkili yönetimi büyük önem taşımaktadır. Solunum seslerinin dinlenmesi (auskultasyon), geleneksel olarak solunum hastalıklarının teşhisinde kullanılan bir yöntemdir. Ancak, bu yöntemin etkinliği, hekimin deneyimine ve çevresel faktörlere bağlı olarak değişkenlik gösterebilmektedir. Bu nedenle, dijital stetoskoplarla kaydedilen solunum seslerinin bilgisayar ortamında otomatik olarak analiz edilmesi, daha objektif ve güvenilir bir teşhis yöntemi olarak öne çıkmaktadır. Ancak, bu alandaki çalışmaların ilerlemesinin önünde bazı önemli zorluklar bulunmaktadır. Bunlardan ilki, doğru etiketlenmiş ve kapsamlı solunum sesi veri setlerinin yetersizliği, ikincisi ise solunum seslerindeki gürültüyü etkili bir şekilde giderme ihtiyacıdır. Bu çalışmada, bu iki temel zorluğa değinilerek, literatürdeki olası çözüm önerileri incelenecektir.**

Anahtar Kelimeler*—solunum sesi; akciğer sesi; veri artırma; gürültü giderme

I. Introduction

Automatic lung sound classification is an important area of research due to the increasing prevail of respiratory diseases. Major risk factors for lung related morbidity include smoking, indoor and outdoor air pollution, occupational exposures, and poverty [1]. In addition, many COVID-19 pandemic survivors suffer from short or long-term damages to the lungs [2], [3]. Understanding the global and regional prevalence, morbidity, and mortality of chronic respiratory diseases is crucial for advancing prevention, screening, treatment, and research initiatives [4]. Correct identification of chronic respiratory diseases requires adequate access to and utilization of diagnostic instruments such as spirometry and chest imaging, as well as efficient and practical case-finding approaches [5]. Furthermore, efforts should be directed toward promoting early and accurate diagnosis and treatment of chronic respiratory diseases to improve long-term clinical outcomes and reduce premature mortality [6].

Accurate diagnosis and classification of lung diseases are crucial for proper management and treatment. Following the COVID-19 pandemic, naturally there has been an increase in the studies focused on the development of diverse diagnostic and classification tools [7]. Anamnesis (complaints), physical examination, and, crucially, auscultation are the primary diagnostics tools that the specialists employ in assessing the condition of the patient [8]. Every respiratory check-up includes an audio auscultation by which the medical specialist listens to sounds from the patient body with different tools (stethoscope, sonography). This shows how important sound analysis is for lung disease detection since accurately interpreting respiratory sounds, with complemented by other findings, allows specialists to make a reliable diagnosis with a minimal margin for error. However, the usefulness of auscultation completely depends on the doctor's experience, time, and external noise factors. Despite this, World Health Organization (WHO) statistics [9] reveal that 45% of WHO Member States report having less than 1 physician per 1000 population, which is the WHO recommended ratio. To overcome these disadvantages, the automatic classification of respiratory sounds recorded with a digital stethoscope through an expert software has been proposed, removing the dependence on subjective interpretation of the specialists [10].

In classification problems, it is a well-known fact that, independent from the classifier and the signal type, both quantity and quality of data have dominant influence on the success. Unfortunately, when dealing with biomedical signals, scarcity and low quality of data are the common problems posing a challenge which impacts the biomedical research and application. Respiratory and lung sounds are not exceptions in terms of these problems. Publicly available databases for research are very few and they usually offer unbalanced data which negatively affects research progress in sound classification research [11]. In addition, data quality is generally low due to many reasons such as electrical interference, muscle movements, environmental sounds, etc. Environmental noise, a common issue in audio recordings, significantly hinders the performance of deep learning systems [12]. There exist other artifacts due to recording device errors, patient movements, or incorrect placement of sensors which are undesired effects on signals that need to be removed for accurate analysis.

In order to deal with data quantity problem, increasing data number by creating new synthetic data (i.e. data augmentation) is a popular way. Contrary to the popular belief that large datasets are always essential for optimal deep learning performance [13], data augmentation has consistently demonstrated its effectiveness in improving training model performance, even with small datasets [14]. The necessity of data augmentation cannot be overstated, as previous research studies on neural network models for sound/audio classification have shown its effectiveness in addressing overfitting and reducing sensitivity to background noise and information redundancy [15]–[17]. For overcoming data quality problems, there are many methods used. These methods mostly use denoising techniques.

In summary, in the field of respiratory and lung sound processing, the lack of sufficient data, weakly labelled data, unbalanced and noisy datasets affect the overall performance of classifiers. Two primary challenges persist: the quantity and the quality of the available data. This brief literature review aims to identify recent research advancements addressing these challenges within the context of lung and respiratory sound data. Examples from the initial search findings using various keyword combinations from the recent years are selected and summarized.

In the following sections, we delve into the underlying factors contributing to these challenges. Section 2 focuses on the difficulties of obtaining sufficient amount of data, addressing issues related to data scarcity and discuss data augmentation techniques as a potential solution. Section 3 deals with the problem of data quality and presents examples of studies aimed at enhancing data quality. Section 4 briefly talks about some future research areas. Finally, Section 5 provides a general discussion.

II. Avalibility of Data for Research Purposes

A. Data Gathering

In signal classification problems, it is a well-established fact that the quantity and quality of data significantly impact success. Unfortunately, biomedical signals often suffer from scarcity and low quality, posing a significant challenge for research and application. In case of rare diseases or conditions, obtaining sufficient amount of data is particularly difficult. However, this problem is not exclusive to rare cases. Even for common diseases or health conditions, acquiring adequate data remains a major challenge in biomedical signals. One primary reason for these difficulties lies in the sensitive nature of personal health data. Strict regulations, ethical guidelines, and the need for extensive permissions make data collection and sharing arduous [18]–[21]. Additionally, performance evaluation in classification systems relies on accurately labeled data. A significant limitation arises from the requirement for expert knowledge to collect accurate data and annotate it correctly. This involves expert(s) examining the data and using their expertise to annotate, leading to lengthy and costly data collection processes. While there are limited examples of automated bio signal retrieval and annotation tools for non-technical users, aimed at streamlining ground truth collection for biomedical applications, this remains a substantial challenge [22]. Creation of large databases that is required for data hungry deep learning based classifiers is especially difficult due to the differences in biomedical signal recording devices and protocols that makes data comparison and assembly difficult [23]–[30].

Large databases are often generated from different centers, i.e., universities, hospitals, or research institutions that may have different policies regarding data sharing, hindering data combination, and large-scale analysis [27]. This necessitates additional specifications for recording equipment and protocol compatibility. Quality assurance during data acquisition is also crucial and must adhere to structured guidelines. Unfortunately, there's a lack of universal guidelines, and these often need to be tailored to specific databases. In conclusion, the limited availability and challenges in obtaining biomedical signal data are due to a combination of factors. These challenges can hinder biomedical research progress and delay the development of new diagnostic helps. To address these issues, it is essential to establish international standards for data sharing, develop reliable data privacy mechanisms, and optimize data collection processes [31], [32].

In addition, usually the primary purpose of recording biomedical data is typically not database creation but patient monitoring, diagnosis, or research. Therefore, naturally the priority of health care professionals is not accumulating or annotating data immediately. Luckily, as long as continuous storage of data is carried out, immediate visualization or annotation of analyzed data is usually not required [31].

Lung and respiratory sound data are not free from the problems explained above. There are not many publicly available databases of lung sounds exist. The most widely used one for research and development is The International Conference on Biomedical and Health Informatics (ICBHI) 2017 ICBHI Respiratory Sound Database, a public database [33]. Database comprises 5.5 hours of recordings obtained from seven different chest locations, namely trachea, left and right anterior, posterior, and lateral. Recordings encompass a total of 6898 breathing cycles, which were labeled by respiratory specialists as containing crackles, wheezes, a combination of both, or no abnormal respiratory sounds. The sample frequencies of the sounds in the datasets range from 4 kHz to 44.1 kHz, and the recordings vary in duration from 10 $sec.$ to 90 $sec.$ Each recording is composed of a certain number of breathing cycles with corresponding annotations of the beginning and the end, and the presence/absence of crackles and/or wheezes. The cycle duration ranges from 0.2 $sec.$ to 16 $sec.$ and the average cycle duration is 2.7 $sec.$ The database includes 6898 different respiratory cycles with 3642 normal cycles, 1864 crackles, 886 wheezes, and 506 cycles containing of both crackles and wheezes. The database was compiled from 126 individual participants over several years by two separate study teams located in two countries. A total of 920 audio samples were recorded using heterogeneous types of equipment, namely Meditron, LittC2SE, Litt3200 stethoscopes, and AKGC417L microphone. The respiratory cycles were also categorized into eight distinct conditions by experts: Upper Respiratory Tract Infection (URTI), Chronic Obstructive Pulmonary Disease (COPD), Bronchiectasis, Pneumonia, Bronchiolitis, Asthma, Lower Respiratory Tract Infection (LRTI), and Healthy. We provided a detailed information about the ICBHI database above to be able to underline some facts. Despite currently being the largest database and serving as a benchmark for most of the research in this field, it is not immune to the aforementioned common shortcomings. ICBHI database suffers from the problems of very scarce data for some diseases offering an unbalanced data set which creates serious difficulty in classification studies that include these diseases of limited data. Distribution of the data among disease classes also does not reflect the corresponding prevalence of the diseases in society. Usage of different types of equipment (one microphone and three different digital stethoscopes), sampling with different frequencies, existence of very noisy, even unusable data, recordings of auscultation from different locations of chest from different subjects, and recordings of different durations are the aforementioned problems of creating a high quality biomedical signal database.

In addition to ICBHI database, there are other publicly available databases. A relatively new database, CoronaHack-Respiratory-Sound-Dataset [34], include sound files of Corona-affected subjects and subjects who does not have corona to classify the respiratory sounds of healthy vs Corona-affected patients. Multiple categories of respiratory sound files include breathing-deep, breathing-shallow, cough-heavy, cough-shallow, counting-fast, counting-normal, vowel-a, vowel-e, and vowel-o. User Health ailments and Corona Symptoms are also provided. It includes 1397 user's data from different countries. It provides data only for a specialized area of research. Respiratory Disease Detection [35] database includes 160 subject's audio files. 6$sec.$ of the respiration cycle of a patient is recorded. Database includes 7 diseases and healthy classes. One microphone and three stethoscope is used to gathered data. A Dataset of Lung Sounds [36] database includes respiratory sounds from 120 subjects (35 healthy and 77 unhealthy). Normal, asthma, COPD, BRON, heart failure, lung fibrosis, and pleural effusion are the labels provided for the data. For each patient disease diagnosis and the lung sound type is given in the annotation file. Data base also have provides labeled data for normal (35), crepitation (23), wheeze (41), crackle (8), bronchial (1), wheeze and crackle (2), and bronchial and crackle (2) sounds. Asthma Detection Dataset Version 2 [37] is a specialized collection of audio samples designed to facilitate research in diagnosing asthma and other lung conditions. This dataset is self-created consisting of sound files segmented into $1.5 - 5$ $sec.$ to ensure consistency and manageability for analysis. The dataset is a good source for developing and testing ML models aimed at detecting asthma through the analysis of lung sounds. Database includes approximately 170 samples from the Respiratory Sound Database and 212 samples from ICBHI database, and the remaining data is original to the database. In total, database includes 4 disease class, asthma (288 samples), bronchial (104 samples), COPD (401 samples), pneumonia (255 samples), and healthy (133 samples) class. Another publicly available lung sound database is Clinical Audio Database [38]. This database offers free lung, breath, heart, and ambient sounds in MP3 format. Data is presented in MP3 format for easy accessibility and offers a broader range of audio samples, including heart and breath sounds. Unfortunately, amount of lung sound data is very limited (10 recordings of 10 $sec.$ to 1.4 $sec.$) for studies of research purposes, but is suitable for educational purposes and simulations. Pulmonary Sound Dataset [39] includes 532 audio samples of crepitation, normal, rhonchi and wheezing. Each file is 10 s long. Disease labels are not provided, therefore database serves to a limited research area. There are also several publicly available lung sound databases, often intended for educational purposes that contain a limited number of examples for each type of respiratory sound. These datasets, typically found on repositories or CDs, often feature clear sounds without environmental noise, making them less suitable for training realistic classification models. Some examples are online available at

- http://www.rale.ca
- https://www.easyauscultation.com/lung-sounds
- https://www.thinklabs.com/sound-library
- https://github.com/soundcloud

B. Data Scarcity-Data Augmentation

Creating a large, well-annotated sound recording dataset is a time-consuming and resource-intensive process, hindering the development of efficient classification systems. Unfortunately, many application domains including lung sound studies do not have access to big data; moreover, they suffer from a class imbalance amid available databases. Especially in the field of deep learning, model performance often improves with the quantity of available training data. Data augmentation techniques effectively generate synthetic datasets (images, sounds, text, etc.) and have consistently demonstrated their effectiveness in enhancing the performance of training models for small datasets, even in the context of sound classification tasks [14]. Given the over-parameterized nature of neural network models, data augmentation is essential for mitigating overfitting and reducing sensitivity to background noise and information redundancy [40]. Previous research studies on sound/audio classification tasks have confirmed the effectiveness of data augmentation in improving model performance [15], [16]. Data augmentation also provides a promising solution to the legal and practical challenges surrounding clinical data. By expanding and improving the quality of training datasets, data augmentation techniques enable the development of more robust deep learning models. Additionally, they address the issue of class imbalance prevalent in many clinical databases [41].

Data augmentation involves generating synthetic data from existing real data while preserving the original class labels. In the context of lung and respiratory sound classification, real data refers to sounds directly recorded from patients, while synthetic data is artificially created using computer-based methods. Effective data augmentation requires domain-specific customization to ensure that applied transformations accurately reflect realistic variations and preserve the critical features that differentiate various classes. In essence, data augmentation should be tailored to the specific classification task. The selection of appropriate augmentation techniques depends on the characteristics of the lung sound dataset and the desired outcomes. For example, to improve the model's robustness to noise, adding noise or frequency masking could be beneficial. Conversely, to increase the dataset's diversity, time stretching or pitch shifting might prove effective [41]. The choice of data augmentation schemes can significantly impact the generalization ability of synthetic samples. Inaccurate choices may lead to poor classifier performance. In the medical domain, insufficient data, particularly for images and sound, remains a major challenge. In addition, existing databases often suffer from class imbalance, as exemplified by the ICBHI dataset [42]. So, data augmentation is a necessity in most cases.

Data augmentation methods can be divided into four main groups: time domain augmentation, frequency domain augmentation, time-frequency domain augmentation, and other more advanced methods. Augmentation in time domain includes time stretching which increases/reduces the sampling rate of an audio signal without affecting its pitch, i.e., modifying the tempo of the audio, pitch shifting i.e., changing the pitch of the audio, noise addition to simulate real world conditions, and extracting random segments of the audio [41].

Nguyen et al. [43] employ time stretching to increase or decrease the sampling rate of an audio signal without altering its pitch. This technique is used to double the number of segments for both the wheeze and crackle classes. Additionally, random sampling rate adjustments, uniformly distributed within $\pm 10\%$ of the original sampling rate, are implemented. On the doubled training set, additional data augmentation methods, such as volume adjusting, noise addition, pitch adjusting, and speed adjusting, are randomly applied based on predefined probabilities. Data augmentation in both time domain and time-frequency domain is used to account for the class imbalance of the ICBHI and the multi-channel lung sound dataset created by the authors. They report that the proposed systems mostly outperform all state-of-the-art lung sound classification systems for the adventitious lung sounds and respiratory diseases of both datasets. Gairola et al. [44] propose a streamlined CNN-based model, RespireNet, in conjunction with data augmentation techniques like device-specific fine-tuning, concatenation-based augmentation, blank region clipping, and smart padding. These methods effectively utilize the small ICBHI dataset, comprising only 6898 breathing cycles, which is insufficient for training a robust deep network model. Their approach achieves a 2.2% improvement over state-of-the-art results for $4-$class classification.

In frequency domain, shifting the frequency spectrum of the audio, randomly masking certain frequency bands, and adding harmonics to the audio signal are methods to create synthetic data. In time-frequency domain data augmentation approaches, adding Gaussian noise to the time-frequency representation of the audio, randomly masking regions of the spectrogram, and adding harmonics to the audio signal, spectrogram flipping by reversing the rows or columns of pixels vertically or horizontally, respectively are some common methods [45].

Arı et al. [46] propose a data augmentation method specifically tailored for respiratory sound classification. Their approach involves input transformation and migration techniques applied to the data from ICBHI database. Experimental results demonstrate that the proposed data augmentation methods can enhance separation performance compared to baseline methods. Notably, these methods are easily adaptable to existing automated auscultation systems.

Nguyen et al. [43] also use vocal tract length perturbation (VTLP) for data augmentation. VTLP is representative of a group of data augmentation schemes that generate new samples through perturbing or distorting the recording spectra of the existing training samples. VTLP is applied to enlarge the dataset for all classes for both the original training set and the time stretched data. VTLP selects a random wrap factor for each recording and maps the frequency of the signal bandwidth to a new frequency. They choose wrap factor from a uniform distribution between 0.9 and 1.1 and set the maximum signal bandwidth to between $3.2\ kHz$ and $3.8\ kHz$. VTLP is applied directly to the mel filter bank rather than distorting each spectrogram. Additionally, they double the log-mel features by adding the flipped log-mel features (in frequency axis).

Ma et al. [47] employed two basic forms data augmentation for the training data: audio stretching (speeding up or down) as well as Vocal Tract Length perturbation to address the imbalance problem and to improve the robustness of the model. Proposed model has been compared with the state-of-the-art works using the official ICBHI 2017 challenge dataset. A performance score of 52.26%, which is improved by $2.1-12.7\%$ compared to the state-of-the-art models is achieved.

In addition to traditional data augmentation methods, innovative transformation techniques have been proposed to generate synthetic training samples, including random erasing, scaling, masking (frequency and time), standardization, and trimming [48].

Other augmentation methods include data mixing, i.e., combining different lung sound recordings to create new samples; or more advanced methods like using Generative Adversarial Networks (GANs) to generate new, realistic lung sound samples, and Cycle-Consistent Adversarial Networks (CycleGAN)-based augmentation using CycleGANs to translate lung sounds from one domain to another (e.g., healthy to diseased), and Variational Autoencoders (VAE) [49]–[51].

The limited availability of respiratory signals has hindered the accuracy of computer-aided diagnosis. Traditional data augmentation approaches can distort the inherent characteristics of the time-frequency representation (TFR) of the signal. Jayalakshmy et al. [52] propose using conditional GANs to augment the dataset. The signals are analyzed using wavelet-based TFR, specifically scalograms. To demonstrate the impact of data augmentation, classification was performed with three pretrained classifiers. The results indicate a significant improvement in accuracy, from 81.37% to 98.75%, achieved using a ResNet-50 model trained on the augmented dataset.

Generative Adversarial Networks (GANs) have demonstrated remarkable success in synthesizing realistic images. However, their application in audio generation is less prevalent due to the scarcity of available datasets for developing accurate models. Yella et al. [53] propose WaveGAN, a variant of GANs, as a solution for raw audio synthesis in a supervised setting for classification tasks. Their method showcases one approach for augmenting speech datasets using GANs. By deploying WaveGAN on existing open-source datasets, they generate synthetic, larger datasets to develop an accurate sound-based diagnosis tool.

Saldanha et al. [54] aim to synthesize respiratory sounds of various categories using different variants of VAE, including MLP-VAE, CVAE, and Conditional VAE. They compare the impact of augmenting the imbalanced dataset on the performance of various lung sound classification models. The quality of the synthesized respiratory sounds was evaluated using metrics such as Frechet Audio Distance, Cross-Correlation, and Mel Cepstral Distortion. Significant improvements in classification performance metrics were observed for certain minority classes when the imbalanced dataset was augmented, while marginal improvements were noted for other classes. The results demonstrate that deep learning-based lung sound classification models offer promising solutions compared to traditional methods and can achieve substantial performance gains when trained on augmented, imbalanced datasets.

Soni et al. [55] used augmentation-based contrastive learning methodology as the base line for comparison to supervised learning using limited labeled data. For generating contrastive views, they use augmentations to hide spectrogram information (time and frequency dimensions) from the model and applied splitting, time masking, frequency masking, spectrogram masking, and spectrogram masking and splitting. The study paves the path for medical applications of contrastive learning that leverage clinical information.

III. Data Quality-Noise

Data quality is another problematic area when it comes to biomedical signals. Several reasons, mostly inevitable, deteriorate the quality of signals. Electrical interference, muscle movements, environmental sounds, and other factors can introduce noise into biomedical signals, lowering the quality of signals to lead difficulty in making analysis. In addition, there exist other artifacts due to recording device errors, patient movements, or incorrect placement of sensors which are undesired effects on signals that need to be removed for accurate analysis.

Lung sound data suffers from bunch of distorting effects unique to biomedical signals. Therefore, preprocessing techniques as the first stage to enhance the quality of the auscultation recordings is a necessity in most of the cases. Successful lung sound extraction from lung auscultations recordings requires precise filtering. The choice of preprocessing techniques depends on the specific characteristics of the lung sound data, the desired features, and the machine learning algorithm used. Experimentation is often necessary to find the optimal combination of techniques for a given task.

Noise in lung sounds can be broadly divided into two main groups: interference of heart sound and ambient noise. The sounds should be cleaned from these interferences by data cleaning procedures. However, data cleaning must maintain a high peak-signal-to-noise ratio without removing desirable signal information [56]. To get a clean auscultation signal, other noise factors such as periodic noise, DC offsets, baseline errors, harmonic noise, etc. must also be addressed. These types of noises are relatively easier to handle.

A. Heart Sound Removal

One of the primary challenges in lung sound analysis is the unavoidable interference of heart sounds during recording. The heart sound (HS) is the major noise, which complicates the lung sound signal processing and can interfere with the clinical interpretation of lung sounds, particularly in the low-frequency range at low flow rates. It's therefore desirable to eliminate the influence of HS on lung sound recordings. The heart and lung sound signals overlap in the time and frequency domains, therefore removing HS interference from respiratory sound recordings is a challenging task. In addition, lung and heart sounds are weaker than environment noises, and have frequency bands which overlap significantly with noise frequencies. Conventional denoising methods may not be practical due to the noisy nature of the lung sound as well as its spectral overlap with different noise sources [57].

Most common techniques in related area can be grouped as wavelet transform-based methods such as multiresolution decomposition, time frequency filtering, machine learning-based methods, adaptive filtering (wiener filtering, spectral subtraction, non-stationary noise reduction), and hybrid approaches (Combines multiple methods to achieve better results.). Below we give some examples of hearth sound removal studies.

Pouyani et al. [58] proposed an adaptive technique based on Discrete Wavelet Transform and Artificial Neural Network (DWT-ANN). This new method mixes the multi-resolution property of DWT with ANN as a nonlinear. Authors reports that proposed method significantly enhanced compared to employment of only DWT method. Hossain et al. [59]

also employed a wavelet transform-based adaptive denoising technique to mitigate heart sounds from lung sounds. However, wavelet-based filtering significantly reduces the average power of lung sounds across the entire frequency range, leading to noticeable changes in the original signal's spectrum, which poses an important problem on feasibility of the method. Mondal et al. [60] proposed a method based on Empirical Mode Decomposition (EMD) technique for reducing the undesired heart sound interference from the desired lung sound signals. In the study, the mixed signal is split into several components containing heart sound, environmental noise etc. The results of the analysis of the synthetic and real-time recorded mixed signal show that the proposed method has ability to remove efficiently the HS interference, without any degradation of the quality of the reconstructed LS signal. Singh et al. [61] examined six denoising techniques namely, Wavelet, Savitzky Golay Moving average filter, FIR, Median filter and Butterworth filter for heart sound removal and evaluated the results in terms of signal to noise ratio. They reported the superiority of wavelet denoising technique over others.

Yamuna et al. [62] employed Adaptive Variational Mode Decomposition (AVMD) technique to remove heart sound contaminants from lung sounds. The proposed AVMD method initially breakdown the noisy lung sound signal into a collective of bandlimited modes called Variational Mode Functions (VMF). Then, based on the frequency spectrum, the HS is filtered out from the LS. The real time lung sound data is collected from 95 participants and the performance of VMD technique is evaluated using the statistical metrics measures. These experimental results are found to be superior and outperform all other recently proposed techniques. Sangeetha and Periyasamy [63] proposed an Enhanced Variational Mode Decomposition (E-VMD) technique to remove HS interference from LSs effectively. The E-VMD method automatically determines the mode number for signal decomposition based on the characteristics of variational mode functions (VMFs) such as normalized permutation entropy, kurtosis index, extreme frequency domain, and energy loss coefficient. This method improves denoising accuracy and computational efficacy, making it a useful tool for improving the analysis of LS signals and assisting in medical diagnostics. In comparison to other denoising methods such as EMD, ensemble empirical mode decomposition (EEMD), complementary ensemble empirical mode decomposition (CEEMD), singular spectrum analysis (SSA), and VMD, the new E-VMD method demonstrates superior denoising outcome. There are various applications of singular spectrum analysis (SSA) in biomedical signal denoising, one of them is removing heart sounds from lung sounds. A crucial preprocessing step in many heart sound cancellation methods involves localizing the primary heart sound components. This paper employs singular spectrum analysis (SSA), a robust time series analysis technique, to achieve this goal. Despite the frequency overlap between heart and lung sound components, two distinct trends can be observed in the eigenvalue spectra. This enables the identification of a subspace that contains more information about the underlying heart sound. To evaluate the performance of the proposed method, both artificially mixed and real respiratory signals were used. Selecting an appropriate window length for SSA results in high-quality decomposition and low computational cost. The proposed method was compared to well-established methods that utilize the wavelet transform and signal entropy for heart sound component detection. Results demonstrate that the proposed method outperforms the wavelet-based method in terms of false detection and correlation with the underlying heart sounds. While the performance of the proposed method is slightly superior to the entropy-based method, its execution time is significantly lower [64].

Pourazad et al. [65] proposed method for HS removal, based on time-frequency filtering, demonstrated promising results in preserving lung sound characteristics. By using multiresolution decomposition of wavelet transform coefficients to localize HS segments, these segments were removed from the original lung sound record, and missing data was estimated through $2D$ interpolation in the time-frequency (TF) domain. The signal was then reconstructed in the time domain. The use of TF-filtering introduced no noticeable clicks or artifacts in the reconstructed signal. A common method to minimize the effect of heart sounds is to filter the sound with linear high-pass filters which, however, also eliminates the overlapping spectrum of breath sounds. Iyer et al. [66] used adaptive filtering to reduce heart sounds without significantly affecting breath sounds. The technique is found to reduce the heart sounds by $50 - 80$ percent. Singh et al. [57] proposed a $2-$level multi-ensemble filtering model with 43 filters that analyze 15 other parameters to denoise and extract lung sound from lung auscultations to eliminate these drawbacks. Combining LMS, NLMS, and RLS is used for denoising process. For heart sound identification, an ensemble of Savitzky-Golay, FIR equiripple, Butterworth, Chebyshev, Elliptic, and wavelet filters is used. This selective combination of filters improves PSNR by 20% compared to sole filter performance, while signal entropy, crest factor, root mean squared error, kurtosis, etc. also improves for different scenarios.

Molaie et al. [67] considered the chaotic behavior of respiratory sound to study the stretching and folding features extracted from the curves selected out of the trajectory. This research group developed the method for heart sound suppression using static discrete wavelet transform with an autoregressive and moving average (ARMA) model.

Groopy et al. [68] presented novel artificial intelligence-based Non-negative Matrix Factorisation (NMF) and Non-negative Matrix Co-Factorisation (NMCF) methods for neonatal chest sound separation. To assess these methods and compare them with existing single-channel separation methods, an artificial mixture dataset was generated comprising heart, lung, and noise sounds. Overall, both methods outperform the existing method especially in artificial data case. Maximum of 1.12 dB signal quality improvement is achieved for the real-world dataset.

Lozana et al. [69] investigate empirical mode decomposition (EMD) and the S-method steepest gradient-based reconstruction algorithm for cancellation of heart sound noise from lung sound. EMD is to decompose nonstationary and nonlinear signals into intrinsic mode function components (IMFs). The main advantage of this method is that instantaneous frequency

(IF) estimated by EMD provides information about the frequency content of the sound signals at each time instant. The use of selected IMFs for feature extraction has the advantage of low computational cost and less time consumption. The method was evaluated using four statistical evaluation metrics: the root mean square error (RMSE), the signal to noise ratio (SNR), the normalized root mean square error (NRMSE), and the percentage of correlation coefficient (percent CC). The maximum SNR of 39.65 dB, the lowest RMSE of 0.00051, the lowest NRMSE value of 0.00022, and the highest percent CC of 98.45 was obtained. Hence the algorithm successfully eliminates the heart sound noise from lung sounds for accurate detection of lung disorder.

Al-Naggar [70] introduces a novel LS filtering method capable of simultaneously separating heart sounds (HS) and noise interference (NI). This method leverages the least mean squares (LMS) algorithm in conjunction with adaptive noise cancellation (ANC). In the second step, the reference input of LMS-ANC is modulated to facilitate the combination of HS and NI signals. The resulting signal is then subtracted from the primary signal (original lung sound recording-LS). The effectiveness of the method is evaluated using power spectral density (PSD). The results demonstrate a clear visual difference in PSD between normal and abnormal LS recordings.

B. Ambient Noise Removal

Typically, the lung sound recording is done in a clinical environment where different sources of ambient noises may be present. Therefore, lung sound signal is usually corrupted with different forms of contaminations which include background noises, power line/Radio Frequency (RF) interferences, environmental noises, and recording artifacts. For a successful automated lung sound based classification, the recorded lung sound signal should be free from any noise which would hinder signal analysis and diagnosis. As a result, numerous studies have focused on ambient noise reduction and lung sound enhancement [71].

A noise suppression method was proposed by Baharanchi et al. [72] for enhancing the respiratory sound signals corrupted by AWGN based on singular spectrum analysis (SSA) combined with discrete cosine transform (DCT). Results are significantly superior to the wavelet. Chang [73] compared the performance of the adaptive filter based on the least mean square (LMS), the dual-channel spectral subtraction, and the independent component analysis (ICA) in ambient noise reduction from pulmonary sounds. The breath sound signals were artificially contaminated with babble noise and ambulance vehicle noise. His comparative study showed that the dual-channel spectral subtraction method was more efficient in removing the ambient noise. Li et al. [74] used adaptive noise cancellation (ANC) based on LMS for environmental noise reduction and preprocessing lung sounds to classify them into two classes: normal and abnormal. Lu et al. [75] applied a real-time LMS adaptive filter for reducing the ambulance siren noise in remote auscultation of the lung sounds.

Syahputra et al. [76] performed a wavelet analysis for noise reduction in respiratory sounds. Haider et al. [77] applied Savitzky-Golay filter to denoise breath sounds corrupted by different levels of Gaussian noise. Meng [78] demonstrated lung sound denoising by serially integrating an FIR band-pass filter, a modified wavelet filter, and an adaptive filter. Shi et al. [79] proposed a method for characteristic extraction and recognition of lung sounds. Wavelet denoising is employed to reduce noise in the collected lung sounds. Wavelet decomposition is then used to extract characteristic coefficients. Through wavelet transform, the analysis of lung sounds across different frequency bands and time positions is enabled, along with noise reduction. The denoised lung sounds are confirmed to be more easily distinguishable by professional doctors.

Singh et al. [80] evaluated six denoising techniques for denoising pulmonary sounds, viz. wavelet, Savitzky Golay, moving average filter, FIR, median filter, and Butterworth filter. They reported the results based on the signal-to-noise ratio (SNR) and showed that the wavelet method performed better. Pouyani et al. [58] used a combination of wavelet transform and artificial neural network (DWT-ANN) for denoising lung sounds. Their results showed that the DWT-ANN performed better than the wavelet transform method.

Fava et al. [81] highlight the effectiveness of KNN and LogitBoost classifiers in data cleaning and enhancing auscultation quality for DNNs. KNN demonstrated a unique ability to identify local similarities and reject outliers, resulting in a clean dataset and improved learning efficiency. Emmanouilidou et al. [82], [83] demonstrated respiratory sound denoising in infants with an average age of 12.2 months using the adaptive spectral subtraction method. They investigated several types of noise, such as environmental sounds, patient-specific noises such as crying, and mechanical noise.

IV. Future Aspects

Future studies seems to focus on several key areas including automatic data labeling, explainability, and federative learning. The lack of sufficient labeled data has emerged as a significant barrier to the advancement of sound classification. This can be attributed to several factors including data privacy concerns, time-consuming nature of data collection, and high dependency on expert knowledge for effective annotation. To address the limitations of labeled data, contrastive learning methods, along with other unsupervised learning methods, leverage unlabeled data to learn meaningful representations [84], [85].

Explainability is another important research area when it comes to biomedical signal processing. Facilitating clinical interpretation, improving model development, ensuring regulatory compliance, and addressing ethical considerations in biomedical signal classification studies are important. Explainable models help establish trust between healthcare providers and patients. When users understand how a model arrives at its conclusions, they are more likely to trust and adopt it in clinical settings. Explainability enhances transparency, allowing users to understand the underlying logic and reasoning behind the model's predictions. This can be particularly important in critical applications where transparency is essential. Explainable models provide insights into the factors that influence the model's decisions. This can help healthcare providers understand the underlying medical reasoning and make informed clinical interpretations. Explainable models can help identify errors or limitations in the model's logic. This information can be used to improve the model's performance and reliability. By understanding the model's reasoning, developers can more effectively debug and troubleshoot issues. Finally, in many regulatory environments, explainability is a requirement for deploying AI models in healthcare. Explainable models can help ensure compliance with relevant regulations and standards [86]–[90].

The success of deep learning is largely attributed to the availability of data. Data samples are frequently collected on edge devices, such as smartphones, vehicles, and sensors, and may not be shared due to privacy concerns. With the emergence of the Internet of Medical Things (IoMT), massive volumes of healthcare sensor data (HSD) are being transmitted over the Internet, which presents various security challenges. Healthcare data is highly sensitive and essential for patient care. Automatic classification of HSD offers significant value in safeguarding patient privacy. Edge computing-based federated learning has introduced novel opportunities and challenges in this context. Federated learning provides a collaborative framework where multiple clients work together to solve machine learning problems under the guidance of a central aggregator. This decentralized approach ensures data privacy by maintaining training data locally on each device. Federated learning adheres to two key principles: local computing and model transmission. This setup mitigates some of the systematic privacy risks and costs associated with traditional centralized machine learning methods. The original data of each client remains private and is not shared or transferred. Instead, devices utilize their local data for training, upload their trained models to the server for aggregation, and receive updated models from the server to achieve the learning goal [91]–[94].

V. Discussion

The field of respiratory sound classification has witnessed significant advancements, driven by the increasing prevalence of respiratory diseases and the limitations of traditional diagnostic methods. This review has highlighted two primary challenges: the scarcity and quality of available lung sound datasets.

The limited availability of large, high-quality lung sound datasets poses a major obstacle in training robust classification models. While there are publicly available datasets, many of them are relatively small and may not capture the full spectrum of respiratory diseases or variations in patient demographics. Solutions to these challenges are presented by data augmentation methods. Generating synthetic data from existing real data can significantly expand the size and diversity of datasets, improving model performance. In addition to traditional techniques like time stretching, pitch shifting, noise addition, and spectrum correction, advanced new techniques based on deep learning such as GAN will be utilized more to augment lung sound data. Collaborative Data Sharing can also ease the problem. Encouraging collaboration among researchers and institutions to share and pool lung sound data can help address the scarcity issue. Establishing data sharing platforms and standardized data formats can facilitate this process. Finally, leveraging crowdsourcing platforms to collect lung sound data from a wider population can contribute to building larger datasets.

Noise contamination and the variability in recording conditions can significantly degrade the quality of lung sound data. These factors can hinder the accuracy of classification models. Applying noise reduction algorithms to remove or minimize the impact of environmental and inherit noise such as heart sound can improve data quality. Establishing standardized recording protocols can help ensure consistent data quality across different studies. This includes guidelines for recording equipment, patient positioning, and environmental conditions. Implementing quality control measures during data collection, annotation, and preprocessing can help identify and address data quality issues.

Addressing the challenges of data scarcity and quality is crucial for advancing respiratory sound classification. The exceptional performance of deep learning methods in pattern recognition tasks has significantly influenced modern sound classification. Despite advancements in deep learning, these methods still face challenges due to insufficient data availability in audio/sound-related tasks. The scarcity of sound data negatively impacts the performance of deep learning methods, particularly CNNs [95]. By employing data augmentation techniques, fostering data sharing, and implementing noise reduction strategies, researchers can develop more accurate and robust classification models. Future research should focus on developing innovative approaches to overcome these challenges and improve the diagnostic capabilities of lung sound analysis.

Future studies should also address the significant challenges posed by large data volumes, data storage, and management in health research. Biomedical signals often generate massive amounts of data, demanding robust infrastructure and software for storage, management, and analysis. While the popularity of big data is recent, the underlying challenges have persisted for a long time and have been actively pursued in health research. Big data in healthcare focuses on datasets that are too extensive, complex, or rapidly generated for traditional healthcare providers to process and interpret using existing tools. This is especially critical in the context of a growing population with an aging demographic and the evolving paradigm of shifting towards prevention, early intervention, rapid and reliable diagnostic tools, and optimal management of health conditions [96], [97].

VI. Conclusion

Respiratory sound classification has witnessed significant advancements, driven by the growing prevalence of respiratory diseases and the limitations of traditional diagnostic methods.Two primary challenges persist in this field: the scarcity and quality of available lung sound datasets. Addressing these challenges is crucial for developing accurate and robust classification models. This brief literature review aims to identify recent research advancements addressing these challenges within the context of lung and respiratory sound data. Examples from the initial search findings using various keyword combinations from the recent years are selected and summarized. By addressing these challenges and exploring innovative solutions, researchers can further advance the field of respiratory sound classification and develop more accurate and reliable diagnostic tools.

REFERENCES

[1] Soriano JB (+ 231 colleagues), et al. Prevalence and attributable health burden of chronic respiratory diseases, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet Respiratory Medicine 2020; 8(6): 585-596.

[2] Narin A, Pamuk Z. Effect of different batch size parameters on predicting of COVID19 cases. Journal of Intelligent Systems with Applications 2020; 3(2): 69-72.

[3] Narin A, Isler Y. Detection of new coronavirus disease from chest x-ray images using pre-trained convolutional neural networks. Journal of the Faculty of Engineering and Architecture of Gazi University 2021; 36(4): 2095-2107.

[4] Labaki WW, Han MK. Chronic respiratory diseases: a global view. The Lancet Respiratory Medicine 2020; 8(6): 531-533.

[5] Labaki WW, Han MK. Improving detection of early chronic obstructive pulmonary disease. Annals of the American Thoracic Society 2018; 15(Supplement 4): S243-S248.

[6] Martinez CH, Mannino DM, Jaimes FA, Curtis JL, Han MK, Hansel NN, Diaz AA. Undiagnosed obstructive lung disease in the United States. Associated factors and long-term mortality. Annals of the American Thoracic Society 2015; 12(12): 1788-1795.

[7] Boehm A, Pizzini A, Sonnweber T, Loeffler-Ragg J, Lamina C, Weiss G, Tancevski I. Assessing global COPD awareness with Google Trends. European Respiratory Journal 2019; 53: 1900351.

[8] Talley NJ, O'COnnor S. Clinical Examination: A Systematic Guide to Physical Diagnosis. Elsevier Health Sciences, 6th edition, 2009.

[9] Wolrd Health Organization. Density of physicians, 2017.

[10] Troncoso A, Ortega JA, Seepold R, Madrid NM. Non-invasive devices for respiratory sound monitoring. Procedia computer science 2021; 192: 3040-3048.

[11] Sun Y, Wong AKC, Kamel MS. Classification of imbalanced data: A review. International Journal of Pattern Recognition and Artificial Intelligence 2009; 23(4): 687-719.

[12] Dong X, Yin B, Cong Y, Du Z, Huang X. Environment sound event classification with a two-stream convolutional neural network. IEEE Access 2020; 8: 125714-125721.

[13] Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel MA, Al-Amidie M, Farhan L. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of Big Data 2021; 8: 53.

[14] Zhao YX, Li Y, Wu N. Data augmentation and its application in distributed acoustic sensing data denoising. Geophysical Journal International 2022; 228(1): 119-133.

[15] Abeber J. A review of deep learning based methods for acoustic scene classification. Applied Sciences 2020; 10(6): app10062020.

[16] Bahmei B, Birmingham E, Arzanpour S. CNN-RNN and data augmentation using deep convolutional generative adversarial network for environmental sound classification. IEEE Signal Processing Letters 2022; 29: 682-686.

[17] Feurer M, Hutter F. Hyperparameter optimization. Book chapter in: Automated Machine Learning: Methods, Systems, Challenges, 2019, pp. 3-33.

[18] Ravizza A, De Maria C, Di Pietro L, Sternini F, Audenino AL, Bignardi C. Comprehensive review on current and future regulatory requirements on wearable sensors in preclinical and clinical testing. Frontiers in Bioengineering and Biotechnology 2019; 7: 313.

[19] Tettey F, Parupelli SK, Desai S. A review of biomedical devices: classification, regulatory guidelines, human factors, software as a medical device, and cybersecurity. Biomedical Materials & Devices 2024; 2(1): 316-341.

[20] Isler Y, Olcuoglu LT, Yeniad M. Data security and privacy issues of implantable medical devices. Natural and Engineering Sciences 2018; 3(3), 12-22.

[21] Richards M. The collection, linking and use of data in biomedical research and health care: ethical issues. Nuffield Council of Bioethics, 2015.

[22] Lourenco A, da Silva HP, Carreiras C, Alves AP, Fred ALN. A web-based platform for biosignal visualization and annotation. Multimedia Tools and Applications 2014; 70: 433-460.

[23] Huang Y, Gottardo R. Comparability and reproducibility of biomedical data. Briefings in Bioinformatics 2013; 14(4): 391-401.

[24] Wang Y, Liu L, Wang C. Trends in using deep learning algorithms in biomedical prediction systems. Frontiers in Neuroscience 2023; 17: 1256351.

[25] Kalaivani K, Kshirsagarr PR, Sirisha Devi J, Bandela SR, Colak I, Nageswara Rao, J, Rajaram A. Prediction of biomedical signals using deep learning techniques. Journal of Intelligent & Fuzzy Systems 2023; 44(6): 9769-9782.

[26] Taye MM. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. Computers 2023; 12(5): 91.

[27] Degirmenci M, Yuce YK, Isler Y. Motor imaginary task classification using statistically significant time domain and frequency domain EEG features. Journal of Intelligent Systems with Applications 2022; 5(1): 49-54.

[28] Sayilgan E, Yuce YK, Isler Y. Investigating the effect of flickering frequency pair and mother wavelet selection in steady-state visually-evoked potentials on two-command brain-computer interfaces. IRBM 2022; 43(6): 594-603.

[29] Uzun R, Isler Y, Toksan M. Use of support vector machines to predict the success of wart treatment methods. 2018 Innovations in Intelligent Systems and Applications Conference (ASYU), 2018.

[30] Sayilgan E, Yuce YK, Isler Y. Evaluating steady-state visually evoked potentials-based brain-computer interface system using wavelet features and various machine learning methods. Book chapter in: Brain-Computer Interface, InTech Publications, 2021.

[31] Penzel T, B. Kemp B, Klosch G, Schlogl A, Hasan J, Varri A, Korhonen I. Acquisition of biomedical signals databases. IEEE Engineering in Medicine and Biology Magazine 2001; 20(3): 25-32.

[32] Pramono RXA, Bowyer S, Rodriguez-Villegas E. Automatic adventitious respiratory sound analysis: A systematic review. PloS One 2017; 12(5): e0177926.

[33] https://paperswithcode.com/sota/audio-classification-on-icbhi-respiratory

[34] https://www.kaggle.com/datasets/praveengovi/coronahack-respiratory-sound-dataset

[35] https://www.kaggle.com/datasets/shivam316/respiratory-disease-dataset-processed-audio-files

[36] https://www.kaggle.com/datasets/arashnic/lung-dataset

[37] https://www.kaggle.com/datasets/mohammedtawfikmusaed/asthma-detection-dataset-version-2

[38] https://www.thesimtech.org/av-stimuli

[39] https://data.mendeley.com/datasets/fr7zvy8j5s/1

[40] Horwath JP, Zakharov DN, Megret R, Stach EA. Understanding important features of deep learning models for segmentation of high-resolution transmission electron microscopy images. Computational Materials 2020; 6(1): 108.

[41] Abayomi-Alli OO, Damasevicius R, Qazi A, Adedoyin-Olowe M, Misra S. Data augmentation and deep learning methods in sound classification: A systematic review. Electronics 2022; 11(22): 3795.

[42] Rocha BM, Filos D, Mendes L, Serbes G, Ulukaya S, Kahya YP, Jakovljevic N, Turukalo TL, Vogiatzis IM, Perantoni E, Kaimakamis E, Natsiavas P, Oliveira A, Jacome C, Marques A, Maglaveras N, Paiva RP, Chouvarda I, de Carvalho P. An open access database for the evaluation of respiratory sound classification algorithms. Physiological Measurement 2019; 40(3): 035001.

[43] Nguyen T, Pernkopf F. Lung sound classification using co-tuning and stochastic normalization. IEEE Transactions on Biomedical Engineering 2022; 69(9): 2872-2882.

[44] Gairola S, Tom F, Kwatra N, Jain M. RespireNet: A deep neural network for accurately detecting abnormal lung sounds in limited data setting. 2021 IEEE 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 527-530.

[45] Geron, Hands-On Machine Learning with Scikit-Learn Keras and Tensorflow: Concepts Tools and Techniques to Build Intelligent Systems, O' Reilly Media, 2019.

[46] Berna A, Alcin OF, Sengur A. A lung sound classification system based on data augmenting using ELM-Wavelet-AE. Turkish Journal of Science and Technology 2022; 17(1): 79-88.

[47] Ma Y, Xu X, Li Y. LungRN+NL: An improved adventitious lung sound classification using non-local block resnet neural network with mixup data augmentation. Interspeech, 2020, pp. 2902-2906.

[48] Araujo T, Aresta G, Mendonça L, Penas S, Maia C, Carneiro A, Mendonca AM, Campliho A. Data augmentation for improving proliferative diabetic retinopathy detection in eye fundus images. IEEE Access 2020; 8: 182462-182474.

[49] Qian Y, Hu H, Tan T. Data augmentation using generative adversarial networks for robust speech recognition. Speech Communication 2019; 114: 1-9.

[50] Ramesh V, Vatanparvar K, Nemati E, Nathan V, Rahman MM, Kuang J. CoughGAN: Generating synthetic coughs that improve respiratory disease classification. 2020 IEEE 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020, pp. 5692-5688.

[51] Mertes S, Baird A, Schiller D, Schuller BW, Andre E. An evolutionary-based generative approach for audio data augmentation. 2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP), 2020.

[52] Jayalakshmy S, Priya L, Sudha GF. Synthesis of respiratory signals using conditional generative adversarial networks from scalogram representation. Book chapter in: Generative adversarial networks for image-to-image translation. Academic Press, 2021, pp. 161-183.

[53] Nishant Y, Rajan B. Data augmentation using GAN for sound based COVID 19 diagnosis. 2021 IEEE 11th international conference on intelligent data acquisition and advanced computing systems: technology and applications (IDAACS), 2021, pp. 606-609.

[54] Saldanha J, Chakraborty S, Patil S, Kotecha K, Kumar S, Nayyar A. Data augmentation using variational autoencoders for improvement of respiratory disease classification. Plos One 2022; 17(8): e0266467.

[55] Soni PN, Shi S, Sriram PR, Ng AY, Rajpurkar P. Contrastive learning of heart and lung sounds for label-efficient diagnosis. Patterns 2022; 3(1): 100400.

[56] Singh D, Singh BK, Behera AK. Design of multi-ensemble hybrid filtering approach for removal of heart sound from lung auscultations. Journal of Information and Optimization Sciences 2024; 45(3): 785-793.

[57] Wang H, Wang LY. Multi-sensor adaptive heart and lung sound extraction. 2003 IEEE Sensors, 2003, pp. 1096-1099.

[58] Pouyani MF, Vali M, Ghasemi MA. Lung sound signal denoising using discrete wavelet transform and artificial neural network. Biomedical Signal Processing and Control 2022; 72: 103329.

[59] Hossain I, Moussavi Z. An overview of heart-noise reduction of lung sound using wavelet transform based filter. Proceedings of the IEEE 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE Cat. No. 03CH37439, 2003.

[60] Mondal A, Bhattacharya PS, Saha G. Reduction of heart sound interference from lung sound signals using empirical mode decomposition technique. Journal of Medical Engineering & Technology 2011; 35(6-7): 344-353.

[61] Singh D, Singh BK, Behera AK. Comparative analysis of Lung sound denoising technique. 2020 IEEE First International Conference on Power, Control and Computing Technologies (ICPC2T), 2020, pp. 406-410.

[62] Yamuna KS, Thirunavukkarasu S, Manjunatha B, Karthikeyan B. Elimination of heart sound from respiratory sound using adaptive variational mode decomposition for pulmonary diseases diagnosis. Journal of Intelligent & Fuzzy Systems 2024; 46(2): 3649-3657.

[63] Sangeetha B, Periyasamy R. Heart sound noise separation from lung sound based on enhanced variational mode decomposition for diagnosing pulmonary diseases. Biomedical Engineering: Applications, Basis and Communications 2024; 36(01): 2350035.

[64] Ghaderi F, Mohseni HR, S. Sanei S, Localizing heart sounds in respiratory signals using singular spectrum analysis. IEEE Transactions on Biomedical Engineering 2011; 58(12): 3360-3367.

[65] Pourazad MT, Moussavi Z, Thomas G. Heart sound cancellation from lung sound recordings using time-frequency filtering. Medical and Biological Engineering and Computing 2006; 44: 216-225.

[66] Iyer VK, Ramamoorthy PA, Fan H, Ploysongsang Y. Reduction of heart sounds from Lung sounds by adaptive filtering. IEEE Transactions on Biomedical Engineering 1986; 33(12): 1141-1148.

[67] Molaie M, Jafari S, Moradi MH, Sprott JC, Golpayegani SMRH. A chaotic viewpoint on noise reduction from respiratory sounds. Biomedical Signal Processing and Control 2014; 10: 245-249.

[68] Grooby E, Sitaula C, Fattahi D, Sameni R, Tan K, Zhou L, King A, Ramanathan A, Malhotra A, Dumont G, Marzbanrad F. Noisy neonatal chest sound separation for high-quality heart and lung sounds. IEEE Journal of Biomedical and Health Informatics 2022; 27(6): 2635-2646.

[69] Manuel L, Fiz JA, Jane R. Estimation of instantaneous frequency from empirical mode decomposition on respiratory sounds analysis. 2013 IEEE 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013.

[70] Al-Naggar, NQ. A new method of lung sounds filtering using modulated least mean square-Adaptive noise cancellation. Journal of Biomedical Science and Engineering 2013; 6(9): 869-876.

[71] Chang GC, Cheng YP. Investigation of noise effect on lung sound recognition. In 2008 IEEE International Conference on Machine Learning and Cybernetics, pp. 1298-1301, 2008.

[72] Baharanchi SA, Vali M, Modaresi M. Noise reduction of lung sounds based on singular spectrum analysis combined with discrete cosine transform. Applied Acoustics 2022; 199: 109005.

[73] Chang GC. A comparative analysis of various respiratory sound denoising methods. In 2016 IEEE International Conference on Machine Learning and Cybernetics (ICMLC) 2016, Jul 10, 2: 514-518.

[74] Li L, Xu W, Hong Q, Tong F, Wu J. Adaptive noise cancellation and classification of lung sounds under practical environment. In 2016 IEEE 10th International Conference on Anti-counterfeiting, Security, and Identification (ASID), 2016 Sep 23, pp. 39-42.

[75] Lu BY, Hsueh ML, Wu HD. Reducing the ambulance siren noise for distant auscultation of the lung sound. Acoustics Australia 2017; 45(2): 381-387.

[76] Syahputra MF, Situmeang SI, Rahmat RF, Budiarto R. Noise reduction in breath sound files using wavelet transform based filter. In IOP Conference Series: Materials Science and Engineering 2017; 190: 012040.

[77] Haider NS, Periyasamy R, Joshi D, Singh BK. Savitzky-Golay filter for denoising lung sound. Brazilian Archives of Biology and Technology 2018; 61: e18180203.

[78] Meng F, Wang Y, Shi Y, Zhao H. A kind of integrated serial algorithms for noise reduction and characteristics expanding in respiratory sound. International Journal of Biological Sciences 2019; 15(9): 1921-1932.

[79] Shi Y, Li Y, Cai M, Zhang XD. A lung sound category recognition method based on wavelet decomposition and BP neural network. International Journal of Biological Sciences 2019; 15(1): 195-207.

[80] Singh D, Singh BK, Behera AK. Comparative analysis of lung sound denoising technique. In 2020 First International Conference on Power, Control and Computing Technologies (ICPC2T). IEEE, 2020.

[81] Fava A, Dianat B, Bertacchini A, Manfredi A, Sebastiani M, Modena M, Pancaldi F. Pre-processing techniques to enhance the classification of lung sounds based on deep learning. Biomedical Signal Processing and Control 2024; 92: 106009.

[82] Emmanouilidou D, McCollum ED, Park DE, Elhilali M. Computerized lung sound screening for pediatric auscultation in noisy field environments. IEEE Transactions on Biomedical Engineering 2017; 65(7): 1564-1574.

[83] Emmanouilidou D, McCollum ED, Park DE, Elhilali M. Adaptive noise suppression of pediatric lung auscultations with real applications to noisy clinical settings in developing countries. IEEE Transactions on Biomedical Engineering 2015; 62(9): 2279-2288.

[84] Dinis J, Campos G, Rogrigues J, Marques A. Respiratory sound annotation software. In Proceedings of the International Conference on Health Informatics (HEALTHINF-2012), pp. 183-188, 2012.

[85] Fu-Shun H, Huang CJ, Kuo CY, Huang SR, Cheng YR, Wang JH, Wu YL, Tzeng TL, Lai F. Development of a respiratory sound labeling software for training a deep learning-based respiratory sound analysis model. International Forum on Medical Imaging in Asia 2021. Proceedings Volume 11792. SPIE, 2021.

[86] Wanasinghe T, Bandara S, Madusanka S, Meedeniya D, Bandara M, de la Torre Diaz I. Lung sound classification with multi-feature integration utilizing lightweight CNN model. IEEE Access 2024; 12: 21262-21276.

[87] Wanasinghe T, Bandara S, Madusanka S, Meedeniya D, Bandara M, de la Torre Díez I. Lung sound classification for respiratory disease identification using deep learning: a survey. International Journal of Online & Biomedical Engineering 2024; 20(9): 115-129.

[88] Topaloglu I, Barua PD, Yildiz AM, Keles T, Dogan S, Baygin M, Gul HF, Tuncer T, Tan RS, Acharya UR. Explainable attention ResNet18-based model for asthma detection using stethoscope lung sounds. Engineering Applications of Artificial Intelligence 2023; 126: 106887.

[89] Lo Giudice M, Mammone N, Ieracitano C, Aguglia U, Mandic D, Morabito FC. Explainable deep learning classification of respiratory sound for telemedicine applications. International Conference on Applied Intelligence and Informatics. Cham: Springer Nature Switzerland, 2022.

[90] Mercaldo F, Brunese L, Cesarelli M, Martinelli F, Santone A. Respiratory disease detection through spectogram analysis with explainable deep learning. 2023 IEEE 8th International Conference on Smart and Sustainable Technologies (SpliTech), 2023.

[91] Li L, Fan Y, Tse M, Lin KY. A review of applications in federated learning. Computers & Industrial Engineering 2020; 149: 106854.

[92] Yang Q, Liu Y, Cheng Y. Kang Y, Chen T, Yu H. Federated Learning. Springer Link, 2019.

[93] Le S, Wu J. A scalable and transferable federated learning system for classifying healthcare sensor data. IEEE Journal of Biomedical and Health Informatics 2022; 27(2): 866-877.

[94] Gafni T, Shlezinger N, Cohen K, Eldar YC, H. V. Poor HV. Federated learning: A signal processing perspective. IEEE Signal Processing Magazine 2022; 39(3): 14-41.

[95] Ykhlef H, Ykhlef F, Chiboub S. Experimental design and analysis of sound event detection systems: case studies. 2019 IEEE 6th International Conference on Image and Signal Processing and their Applications (ISPA), 2019.

[96] Andreu-Perez J, Poon CCY, Merrifield RD, Wong STC, Yang GZ. Big data for health. IEEE Journal of Biomedical and Health Informatics 2015; 19(4): 1193-1208.

[97] Kocheturov A, Pardalos PM, Karakitsiou A. Massive datasets and machine learning for computational biomedicine: trends and challenges. Annals of Operations Research 2019; 276: 5-34.