



Resilient embedded system for classification respiratory diseases in a real time



Ahlam Fadhil Mahmood^{*}, Ahmed Maamoon Alkababji, Amar Daood

Dept. of Computer Engineering, College of Engineering, University of Mosul, Mosul, Iraq

ARTICLE INFO

Keyword:

Lung sound
Respiratory diseases
Random forest
Co-design
FPGA

ABSTRACT

Listening to lung sounds using a stethoscope is still one of the most important methods to diagnose respiratory diseases. These sounds are complex and challenging to diagnose, as even trained people may misclassify them. Accurate interpretation of these sounds requires excellent experience from the treating physician. For diagnosing respiratory diseases, sounds were analyzed, and various features were extracted for the proposed hierarchical design consisting of four layers. A random forest classifier was utilized for three layers and deep learning for the last layer. An FPGA implementation of the proposed respiratory processor is validated experimentally on soft and hard resources of the Virtex-5 ML506 FPGA board. Designing the system by field programmable gate array in a hierarchical manner that allows classification without completing all four stages. The resilient four-layer system achieved the highest average accuracy of 100 %, 99.83, 99.62, 99.88, and 99.87 for COPD, Healthy, URTI, Bronchiectasis, and Pneumonia diseases while saving both power consumption 63.8 % and 54.7 % of testing time.

1. Introduction

Knowing and analyzing the characteristics of respiratory sound (RS) is an essential part of knowing the condition of patients and the extent of the injury and provides insight into the nature of the lungs. Although the stethoscope has been an invaluable diagnostic tool since its invention, it requires much expertise from clinicians, as even trainees and residents sometimes misdiagnose. With artificial intelligence techniques development, having a portable system for classifying respiratory sounds is extremely important.

In the Tenth International Lung Sound Association (ILSA) [1], classification criteria for respiratory sounds were defined into two categories: (normal and abnormal). These sounds are categorized based on various factors that help detect different disease categories. They are broadly categorized as continuous and intermittent sounds, and this distinction is due to the difference in the duration of their occurrence during breathing [2,3]. Diseases related to the respiratory system can be described by the sounds generated during breathing, which include the following sounds.

a- Fine crackle

They occur in the small airways, which are high-pitched and soft. This type occurs more frequently during breathing when inhaled. The wave recedes quickly with a duration (about 5 msec) and a frequency \geq of 650 Hz [1]. It is caused by interstitial pulmonary fibrosis, pneumonia, or congestive heart failure.

b- Coarse crackle

This type of crackle occurs in the large airways, is high and low in intensity, and lasts longer than the previous type. Rapidly damped wave deflection frequency \geq of 350 Hz, with a longer duration (about 15 ms). It occurs primarily during inhalation but can also occur during exhalation [2]. It is the same as soft crackles but usually occurs in more advanced diseases.

c- Wheeze

Wheezing is a sharp sound from the bronchi and is a regular, continuous sinusoid high-pitched with a frequency from 100 to 5000 Hz and durations of about 80–100 ms. A narrowing airway typically causes

* Corresponding author.

E-mail address: ahlam.mahmood@uomosul.edu.iq (A.F. Mahmood).

Table 1

Characteristics of adventitious sounds types.

Diseases	Pneumonia, Lung fibrosis	Chronic bronchitis, COPD, bronchiectasis	COPD and Asthma	COPD and Bronchitis,	foreign body, Epiglottitis, croup, laryngeal edema	Pneumonia and hypersensitivity pneumonia	lung tumor, Inflammation of the lung membrane	Whooping cough
Frequency (Hz)	≤ of 650	≤ of 350	100–5000	50–200	500	200–300	350	high
Duration (ms)	± 5	± 15	80–100	80–100	250	200	15	250
Continuity Types	Yes Fine Crackle	No Coarse Crackle	Yes Wheeze	Yes Rhonchus	Yes Stridor	Yes Squawk	No Pleural rub	Yes Gasp

them and may lead to a restriction of airflow, which occurs in the expiratory more than in the inspiratory. The most important diseases associated with wheezing are COPD or Asthma, which may be produced by a foreign body hindering the airway, such as a tumor [3–5].

d- Rhonchus

Rhonchus is a continuous sinusoid wave from the bronchi. They can be heard during the expiratory period or through both phases of respiration. It has a low-pitched and harmonious unusual sound with a frequency between (50–200) Hz and (80–100) ms in duration. Among the diseases that generate these sounds are bronchitis and pneumonia [1,2].

e- Stridor

Squeaking sounds are generated in the larynx or trachea by air permeating the narrowed airway. In some cases, it appears during the inhalation or exhalation phase and can also be heard in both phases. Stridor has continuous and sinusoid high-pitched of more than 500 Hz frequency and duration of greater than 250 msec. This sound is usually harsher and louder than the wheezing sound. Epiglottitis, croup, and laryngeal edema are related diseases with stridor that correlate to airway obstruction [3].

f- Squawk

Squawk is a discontinuous adventitious sound with short durations and can be hearable in the inspiration phase. They have a low pitch with frequency ranging from 200 to 300 Hz and a continuation of 200 ms duration time. They appear at the inspiratory phase when the patient suffers from common or hypersensitive pneumonia [1,3].

g- Pleural friction rub

It is a non-rhythmic and discontinuous category, which occurs during breathing when the pleural aggravating faces rub against each other, and the resulting “pleural friction” sound is low pitch. Typically, it is less than 350 Hz and only seems for a 15 ms duration period. Occurs during inspiratory and expiratory phases as a symptom of one of the following diseases: pleurisy, pericarditis, or pleural tumor [1–3].

h- Gasp

It is continuous during the inhalation, and the gasping is loud for more than 250 msec. The related breathing disease to this sound is whooping cough [3]. In Table 1, the characteristics of incidental sounds are summarized.

Great efforts have been made in many studies to classify respiratory diseases in different methods, while they lack the design of a hardware system that works in real-time and is capable of classifying a group of diseases with its symptoms automatically and with low power consumption in preparation for burying it with many devices for automatic

diagnosis. The rest of the paper contains the relevant previous works that were reviewed in Section 2, Section 3 presents materials and methods, including data processing and feature extraction, Section 4 presents the proposed system, Section 5 presents simulation results and discussions, and the sixth is devoted to conclusions and recommendations for future work.

2. Literature review

Many papers associated with automatic adventitious sound classification and analysis have been published in different directions. Ref. [6] reviews forty-nine articles that discussed the data set, sensor types, and extracted features, with the analysis techniques and performance measures used. Recently, many researchers have achieved the task of classifying respiratory sounds using numerous deep learning models, such as Convolutional Neural Networks (CNN) [7–9], Long Short-Term Memory (LSTM) [10], a hybrid CNN-LSTM Network with Focal Loss Function [4], Networks Recursive Neural Networks (RNN) [11], and residual networks (ResNet) [12,13]. These papers had different directions and achieved various goals, one being to diagnose sounds programmatically and another to build the diagnostic equipment, as shown in Table 2.

Many other related deep-learning works have been presented [20]. The authors in ref. [21] classified and reviewed developments in artificial intelligence models and strategies for collecting information through the Internet of Things for various pulmonary diseases. The cough sounds were classified into thirteenth categories (A:L) and tabulated the performance of XGBOOST, SVM, Random Forest models, and logistic regression for different diseases. To solve the problem of instability and the long time it takes to adjust the hyperparameters, the researchers [22] introduced self-tuning optimization by Particle Swarm Optimisation (PSO). Convolutional Neural Network (PSTCNN) allows the model to adjust hyperparameters automatically. In the [23–25], researchers presented a tool for diagnosing allergies and brain tumors in MRI images by analyzing the patient’s condition based on symptoms, with the help of Pythagorean Fuzzy Hypersoft and inverse mapping software.

3. Materials and methods

Diagnosing respiratory diseases requires excellent expertise, and doctors use various devices to carefully attend to patients’ breathing. The specialists must be trained to listen to many patients for an extended period, taking into account the variety of medical stethoscopes used. Designing and implementing an automatic diagnostic system that helps everyone overcome many problems. A simulation of how information is processed in the brain, artificial intelligence works as a means of absorbing, understanding, processing, and classifying information with the ability to predict what will happen in it as a result of creating cognitive ability about it. Technological advances have made deep learning the broadest field to solve many problems, but the lack of a large amount of data prevents its use and the need for training in a reasonable time. With limited data availability, random forests are an

Table 2

Summary of the classification of the respiratory sound in the reviewed articles.

Y[Ref]	Procedure	Result	Tool/dataset	advantages	disadvantages
2021 [1]	Artificial noise was added simultaneously with various deep convolutional networks for classifying seven abnormal respiratory sounds by Fourier analysis. VGG: (B1, B3, V1, V2, D1), AlexNet, ResNet, InceptionNet, and LeNet models are used to distinguish between (Coarse and fine crackle, Pleural rub, Rhonchi, Squawk Stridor, Wheeze).	AlexNet model achieved the highest classification Acc. ¹ of 1.00 % while the rest of the models were 0.95 %, and some were less than 0.95 %.	Python: 3.7 with scipy library/ From [RALE lung sounds 3.2], Thinklabs (digital stethoscope), and Easy Auscultation]. Wheeze:12, Stridor:10, Rhonchi:9, Fine Crackles:11, Squawk:8, Coarse Crackles:11, and Pleural Rub:9.	Enhance the spectrum of faint sounds to make them more potent by adding artificial noise of a subtle spectral nature. Using horizontal flip: for data augmentation by randomly flipping the spectrograms from left to right.	Deep learning usually requires a lot of data. There is still little data available to meet this restriction.
2021 [2]	First, the sounds were converted to a Mel spectrogram, and then VGG16 (Transfer learning) was used to find features and classify them by CNN. 16-layer VGG16 trained on fixed-size images and input processing with convolution layers using 3×3 approximate field kernels. Used input 256×256 size with weights pre-trained on ImageNet by freezing all five convolution blocks without a fully connected layer.	Acc. results were less than 80 % in all four votes. The rhonchi's voice was the hardest to distinguish. CNN has a better performance compared with SVM ² , and VGG16 is the best classifier for SVM and CNN. Additionally, CNN had less computation time compared with SVM.	Python library librosa/ A respiratory sound database comprised 2840 audio records containing 1222 normal breathing sounds, 297 crackling sounds, 298 wheezing, and 101 rhonchi sounds recorded in the Chungnam National University Hospital.	Making a collection of 1,918 respiratory sounds from adult lung disease patients. The performance was compared between CNN and SVM classifiers to investigate the dependency of the feature extractor.	Abnormal sounds and filter noises remain a challenge, but recent innovations in algorithms and actual sound recording respiratory systems can make significant progress.
3 2022 [4]	Proposed a new hybrid neural implement on a Focal Loss (FL) function to deal with the distortion of the training data. Firstly, the features were extracted from a short-time Fourier transform (STFT) spectrogram by a convolutional neural net (CNN) as input to a Long Short-Term Memory network (LSTM) that saves time dependencies between data and classifies four types of lung sounds, including normal, wheezing, crackle and wheezing with crackle.	Used three data splitting strategies: Sen. ³ = 47.37 %, Spe. ⁴ = 82.46 %, score 64.92 % and Acc. ¹ = 73.7 % for 60/40 splitting, Sen. = 52.78 %, Spe. = 84.26 %, score = 68.52 % and Acc. = 76.39 % for using interpatient 10-fold cross validation, and Sen. = 60.29 % and Acc. = 74.57 % using Leave-One-Out Cross Validation (LOOCV).	Python 3.7.7 using DL tool Keras. Tensorflow 2.1 was utilized as the backend of the Keras library. Intel i5 Core –9600 K 3.70 GHz CPU, 16 GB RAM-memory, and an 8-GB NVIDIA GeForce RTX 2070 GPU/ICBHI 2017 Database [14].	Evaluate performance for different (numbers of convolutional layers, dropout rates, and learning rates). Using simplified structures reduces training time. Thus, the total network accounts have become much less.	The proposal categorizes only three types and cannot detect other lung sound classes. Although the ICBHI 2017 respiratory voice database is one of the largest open-access databases, they are limited in number, so the results of the current study cannot be safely generalized. Low classification result (Acc., Sen., and Spe.).
2021 [8]	Authors propose scalable devices by deep convolutional neural networks with multiple end-to-end inputs. The proposal takes audio frames as input and is processed with demographic information using 8 processing engines and was implemented in three different cases on Artix-7 FPGA and CPU+GPU platforms. *Two study cases: cough and dyspnea detection in the first study and	The proposed Artix-7 FPGA implementation consumes 245mW with efficiency energy of 4.9 GOPS/W, which is 1.5× higher than the FPGA and GPU platforms in other works. Detection accuracy for cough=80%, dyspnea= 87.3%, and respiratory disease=83%	Python and Verilog HDL/ Case1: cough and dyspnea SDKaggle2018 [15]-41classes. Case2: Upper Respiratory Tract Infection(URTI), Lower Respiratory Tract Infection (LRTI), COPD Asthma, Pneumonia, Bronchiectasis, and Bronchiolitis public RS database [16].	Tabulate implementation results on CPU+GPU, and FPGA with 3-(CPU+GPU) platforms [Nvidia TX2 GPU+CPU, Arm A57 CPU, and Denver CPU]. Optimize memory requirements and power consumption with bit-width quantization. Diagnosis of stable cases and COPD with higher accuracy, an increase of 5% when demographic information is considered.	Great interest in architecture, platforms, and testing on two datasets with varying many factors instead of classifying multiple RS with high accuracy.

(continued on next page)

Table 2 (continued)

Y[Ref]	Procedure	Result	Tool/dataset	advantages	disadvantages
2014 [17]	respiratory sound with demographic information detection in the second. The authors proposed a portable FPGA wheezing detection system. Using a 2D bilateral filter and edge detection for spectrogram processing, polynomial thresholds image segmentation, and morphological operation with image labeling to extract wheezing features.	Experimental showed that the wheeze recognition result was (0.912) with a clock frequency of 51.97MHz. 2sec. breath can be analyzed for wheezing in 0.07956sec.	Matlab and Xilinx System Generator/breath sounds were recorded at the National Taiwan University Hospital [18].	Added dual port RAM ⁵ to increase speed and maximize hardware resource usage with pipeline technology to achieve the highest possible clock rate and data throughput. Tabulate differences between the hardware and software results.	Small testing samples consisted of 12 healthy people and 13 asthmatic patients. Fixed-point hardware operation. The limited depth of the LUT ⁶ table is the main factor that causes the discrepancy between the hardware and software results.
2022 [19]	The work proposes the development of woven capacitive sensors with a face mask equipped with reading electronics to monitor and record breathing signals at home and send notifications to clinicians. It is made of silver-coated polyamide conductor moisture-sensitive polyester filament, and interlaced filament electrodes act as a capacitor.	The operating time was estimated to be more than 5 days, but in practice, the operating time is accurately determined by the actual test because the mathematical equation does not include the state of the battery and assumes that its initial efficiency is 100%. The work does not contain a percentage of the accuracy of the diagnosis because it focuses on the physical construction of the work.	AutoCAD 2021 for designed capacitive facemask sensor, IDE for microcontroller linked, and Android TM operating system/ Only 10 volunteers, 4 females, with ages ranging between 8 and 30 and 6 males 5–31 years.	The developed sensor is a lightweight feature. The design added 1.1 g only to the face mask and reading electronics that allow the exchange of breathing data with a mobile phone or personal computer. The proposed work contains an electronic analysis with equations and figures of the embroidered sensor in the protective face mask, electronics reading, and implementation of data acquisition done with an internal 10-bit ADC of the ATmega328P microcontroller.	The proposal provides a respiratory rate and differentiates between shallow and deep breathing. The realization was focused on home-medical trying (point-of-care concept), rather than on flexibility, mobility, and liberty of movement. Did not address the use of the proposed while walking, as the noise generated was needed to add sensors for environmental monitoring and acceleration. Separate final readout unit with 43.68 Grams weight. On-site testing of the system for 10 volunteers only.

Note: ¹Acc. =accuracy, ²SVM=support Vector Machine, ³Sen.=Sensitivity, ⁴Specificity, ⁵ Random Access memory, ⁶ Look Up Table, ⁷Analog to digital converter.

Table 3

Analysis Information of the ICBHI-2017 Dataset.

Diseases	Healthy	URTI	LRTI	Asthma	COPD	Bronchiectasis	Pneumonia
Participants	26 13F,13M	14 8F,6M	2 0F,2M	1 1F,0M	64 15F,48M,1NA	13 7F,6M	6 2F,4M
Records	35	24	2	1	792	29	37
Age/sex CW(kg),CH(cm) & ABMI(kg/m ²) (kg/m ²)							
0	(7F,10 M) 8.24,68–25,125	(7F,6M) 8.12,74–32,135	(0F,2M) 7.14,64	0	0	(2F,4M) 9.5,70–17,3,90	(0F,1M) 16.7,103
1	(6F,3M) 62,170–70,183	(1F,0M) NA, NA	0	0	0	(1F,0M) 17.35	0
2	0	0	0	0	0	(1F,0M) 25.5	0
3	0	0	0	0	0	0	0
4	0	0	0	0	(0F,1M) 20.1	0	0
5	0	0	0	0	(3F,3M) 22.1–30.1	(2F,1M) 28.81–25.35	0
6	0	0	0	0	(3F,23 M) 16.5–30.1	(0F,1M) 28.4	(0F,1M) 33
7	0	0	0	(1F,0M) 33	(7F,17 M) 18.6–35.14	(0F,1M) 28.52	(2F,1M) 21–36
8	0	0	0	0	(2F,3M) 17.1–36.76	0	(0F,1M) 26
9	0	0	0	0	(0F,1M) 29.03	0	0
Equipment							
mc_AKGC417L	0	0	0	0	646	0	0
sc_LittC2SE	0	0	0	0	0	0	1
mc_LittC2SE	0	0	0	1	50	0	34
sc_Litt3200	0	0	0	0	60	0	
sc_Meditron	35	24	2	0	36	29	2
Bits per sample	16						

Note:

NA: Not Available.

CW,CH:Child Weight (kg), Child Height (cm), ABMI: Adult BMI(kg/m²), F: Female, M: Male.

0 for (0–9) year, 1 for (10–19), 2 for (20–29), 3 for (30–39), 4 for (40–49), 5 for (50–59), 6 for (60–69), 7 for (70–79), 8 for (80–89), 9 for (90–99).

AKGC417L: AKG C417L Microphone.

LittC2SE: 3 M Littmann Classic II SE Stethoscope.

Litt3200: 3 M Litmann 3200 Electronic Stethoscope.

Meditron: WelchAllyn Meditron Master Elite Electronic Stethoscope.

sc: sequential/single channel, mc: simultaneous/multichannel.

ideal solution for best training with a strong and good model.

3.1. Respiratory dataset

The ICBHI-2017 database comprises audio samples collected in two countries independently by two teams [26,27] from 126 participants with 920 records in 5.5 h (6898 cycles). Sounds were recorded from different stethoscopes and classified by experts into healthy and six different diseases (Bronchiectasis, Pneumonia, COPD, lower respiratory tract infection (LRTI), upper respiratory tract infection (URTI), and asthma). Non-healthy cases are appended to crackle and wheeze or both notation 1864 crackles cycles, 886 have wheezes, and 506 cycles have both. Table 3 summarizes the statistical analysis of the dataset.

3.2. Data pre-processing

Starting with pre-processing, each ICBHI 2017 record contains a four-column notated text file for annotations. The first column marks the beginning of the breathing cycle, the second column is at the end of the cycle, whereas the third and fourth columns are for the presence or absence of crackles and wheezes. So, the first step is to extract the wave indicated by the start and end time of the recording using WavePad audio editor and spread it into seven folders named by the disease diagnosis. Most folders contain three sub-folders one for crackles presence, a second for wheezes presence, and the last for the cycles of folders named by the of having both. Classification is done into the mentioned respiratory diseases (RD). Starting to clean audio waves from noise and extract various features for higher classification accuracy. Some are performed on the whole signal and others from their cycles due to the great diversity in the recording periods. Achieving fixed-length segments requires the use of sample padding for solutions without abrupt changes in the signal.

3.3. Features extraction

In the proposed work, different categories of features will be extracted for better classification, mainly as they are in two domains in local symptomatic lung sounds (ALS) and respiratory disease (RD).

3.3.1. Temporal features

Initially, many features can be directly evaluated from the audio respiratory signal. There are many important temporal features, including:-

- a. Root Mean Square Energy (RMSE): Is used to determine the average over time [28].

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_n x(n)} \quad (1)$$

where $x(n)$, is the respiratory signal, and N is the length.

- b. Temporal centroid (TC): It is the sum of the products of the energy envelope of signal and signal divided by the total energy envelope.
- c. Log attack time: the time needed to reach the greatest signal amplitude from the smallest time threshold [29].
- d. Zero Crossing Rate (ZCR): The rate of sign change in the audio signal in the frame [30]. In other words, it is the number of times the signal changes from negative to positive and vice versa divided by the frame length. Equations 2 and 3 determine the zero crossing rate.

$$\text{ZCR}(i) = \frac{1}{2N} \sum_{n=1}^N |\text{sgn}[x_i(n)] - \text{sgn}[x_i(n-1)]| . \quad (2)$$

$$\text{sgn}[x_i(n)] = \begin{cases} 0, x_i < 0 \\ 1, x_i \geq 0 \end{cases} \quad (3)$$

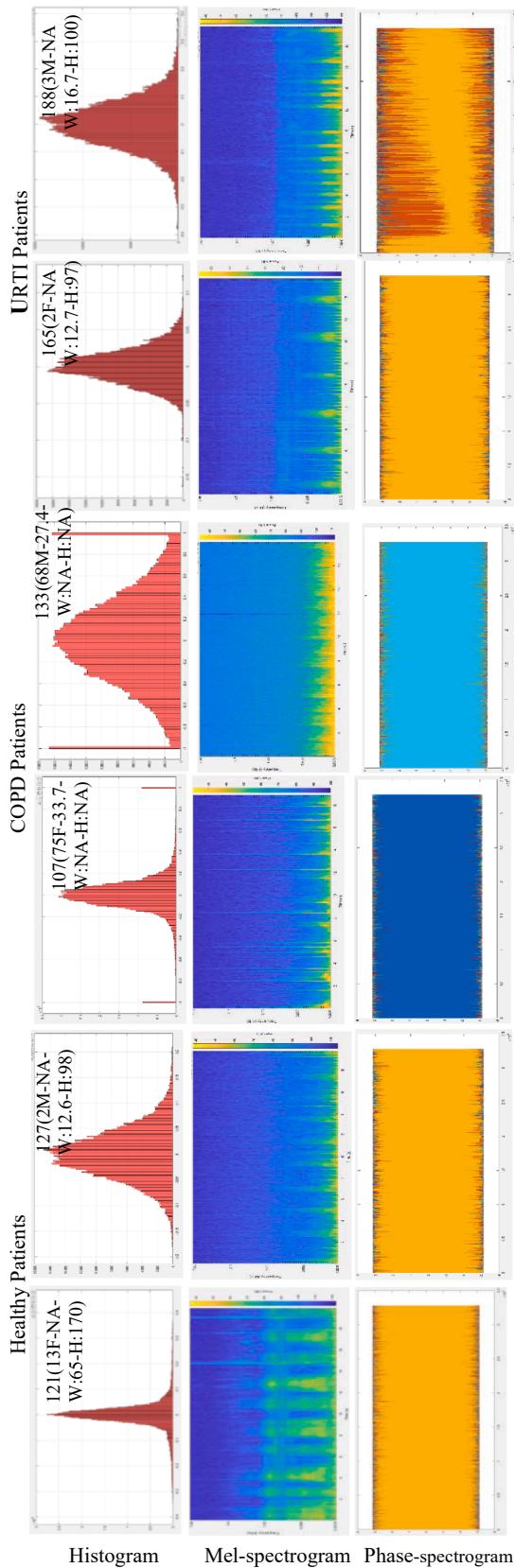


Fig. 1. Layer-4 Images features (Histogram Mel-spectrogram Phase-spectrogram) for three respiratory diseases, two patients for each case having various clinical features.

e. Amplitude Distribution (AD): These features include (mean, variance, histogram form, and energy to distinguish between different sound types.

3.3.2. Spectral features

a. Spectral centroid(SC): It is the center of gravity of the Scale Invariant Fourier Transform (SIFT) spectrum. It offers the characteristic spectral profile of the speech signal [31].

$$SC_t = \frac{\sum_{n=1}^N M_t(n) \times n}{\sum_{n=1}^N n} \quad (4)$$

Where $M_t(n)$ is the SIFT magnitude over time frame t and n frequency bins.

b. Mel-Frequency Cepstral Coefficients (MFCC): Affords the short-term power spectrum of the respiratory sound. Pre-emphasis regularizes the raw signal to extract the MFCC coefficient and minimize disturbances and noise. Further, framing the filter signal into 50 ms frames with 50 % overlapping (25 ms). Then, windowing a signal with $\alpha = 0.48$ and N number of samples/frame length(N) of 40 ms collects the closest frequency sections together, as is given in Equation (5).

$$W(n) = (1 - \alpha) - \alpha \times \cos\left(\frac{2\pi n}{N-1}\right); 0 \leq n \leq N-1 \quad (5)$$

Discrete Fourier Transform (DFT) is the next step to convert the respiratory signal $x(t)$ from the time domain to the frequency domain $X(k)$. The power spectrum of the DFT is given in equation (6).

$$x_k = \frac{1}{N} \left| \sum_{n=0}^{N-1} x(n) \times W(n) \times e^{-j2\pi nk/N} \right|^2; 0 \leq n \leq N-1 \quad (6)$$

After that, the signal is passed on a M number of triangular Mel Frequency filter banks. ($T_m(k)$) to give the respiratory hearing perceptual information as is given in Equation (7). The conversion between linear and Mel frequency is given in Equation (8).

$$ET_m = \sum_{k=0}^{k=1} T_m(k) \times x_k; m = 1 \dots M \quad (7)$$

$$Mel = 2595 \left(1 + \frac{f}{700} \right), \quad f = 70(10^{\frac{Mel}{25905}} - 1) \quad (8)$$

Finally, the Discrete Cosine Transform (DCT) of log filter bank energy delivers L number of cepstral coefficients as given in Equation (9).

$$MFCC_i = \sum_{m=1}^M \log_{10}(ET_m \times \cos(m + .5) \frac{\pi}{m}); \text{ for } i = 1, \dots, L \quad (9)$$

Thirty-nine MFCC features are extracted in the proposed system: one for the energy of the respiratory signal, twelve MFCC coefficients, and twenty-six for first and second-order MFCC feature derivatives.

The proposed system collects different feature sets and clinical features (Age, Sex, Weight) from the ICBHI2017 dataset, as listed in Table 3. And amplitude signal feature (mean, variance, energy, RMSE and temporal centroid with histogram features (maximum and minimum with their dynamic range and standard deviation), spectral centroid. Addition with frame features zero crossing rate and Thirty-nine MFCC features. And transformation to image features (Histogram distribution, Mel spectrogram and Phase features) for deep learning training model as illustrated in Fig. 1.

4. Proposed system

With the significant development of machine learning, it has become the ideal solution for most problems in different fields. One of the most

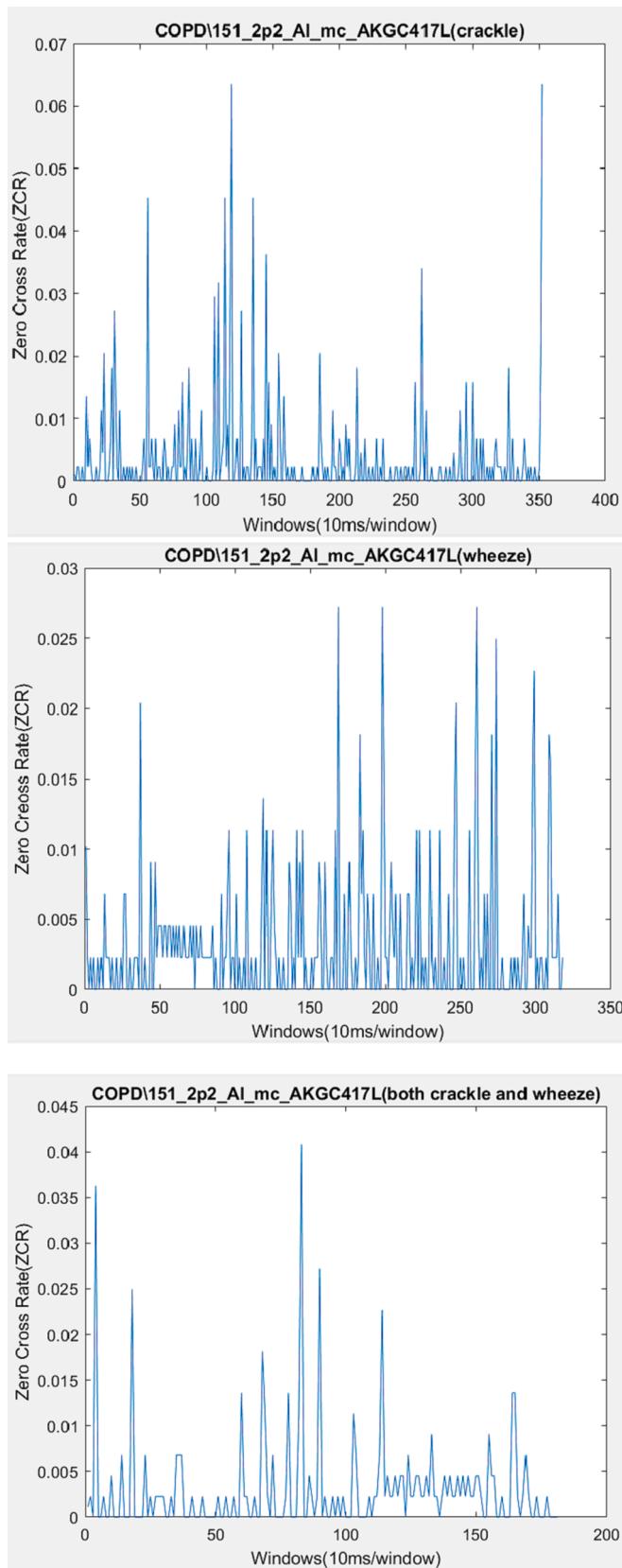


Fig. 2. Performance of ZCR feature for sub-periods in which crackling or wheezing sounds, or both, appear in patients with COPD.

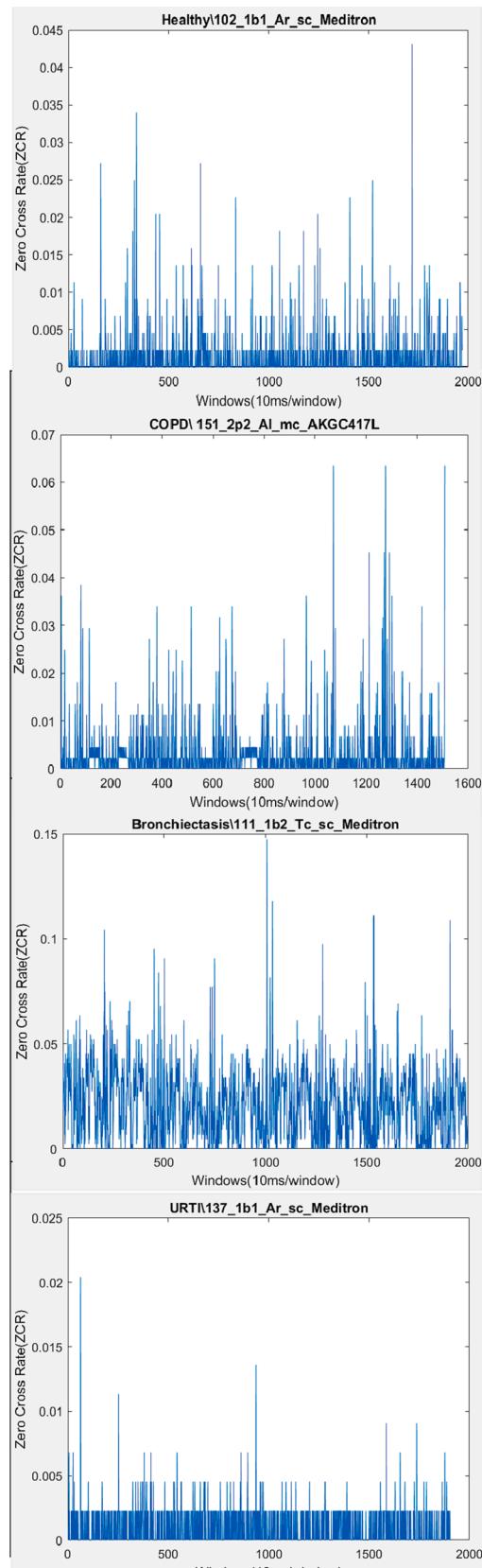


Fig. 3. Performance of the ZCR to the whole record of respiratory diseases (Healthy, COPD, Bronchiectasis, URTI).

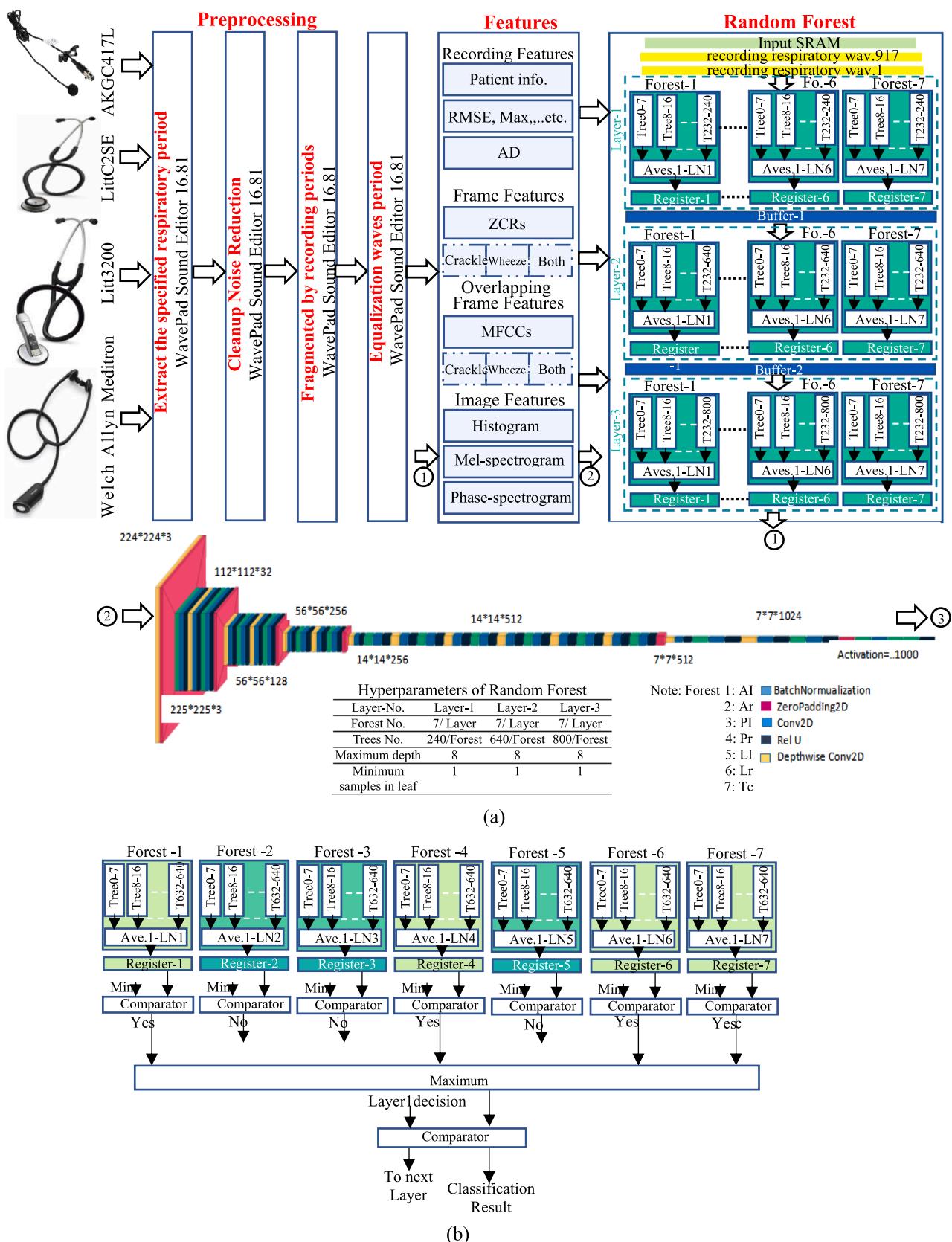


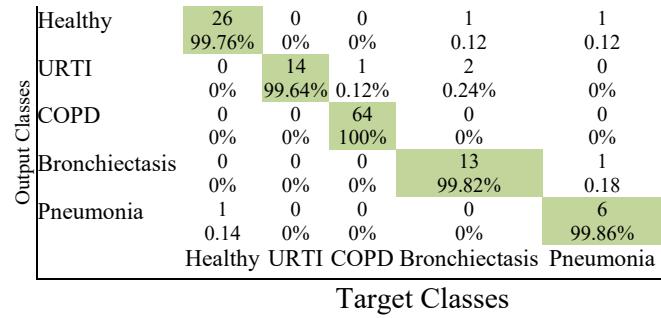
Fig. 4. (a) Architecture of the proposed respiratory processor system with the Hyperparameters of the three Random Forest Layers (b) The internal arrangement of the seven forests in each layer.

Table 4
Comparison of lightweight CNN models.

Model	Params(M)	M-adds(M)	FLOPs(M)
MobileNetV2	3.95	756	425
SqueezeNet	1.82	690	383
shuffleNetV1	4.38	1098	675

Table 5
Performance parameter of the proposed system.

Cases	Accuracy	Sensitivity	Specificity	Precision	F1-score
Healthy	99.83	98.57	99.78	98.87	98.77
URTI	99.62	98.46	99.79	98.75	98.66
COPD	100	100	100	100	100
Bronchiectasis	99.88	98.47	100	100	99.37
Pneumonia	99.87	100	99.87	98.82	99.45



Target Classes

Fig. 5. Overall Confusion matrix of the proposed system.

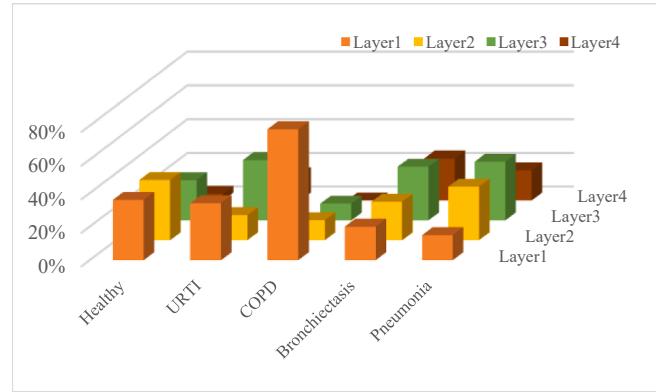


Fig. 6. Distribution of classification rates for the five diseases into the four layers.

prevalent machine learning models is random forest (RF), used for regression and classification problems. It produces a forest with various decision trees and norms estimates over many separate trees. RF classifier utilizes a bagging and random subsets process in constructing each tree to generate an unlinked forest of trees. In addition, many deep learning models can contribute to classifying respiratory diseases.

To increase the accuracy of classification, three levels were adopted. Firstly, features are extracted from the whole respiratory record for each patient. Three attribute types are used [clinical patient information (edge, sex, Adult BMI(kg/m²) and weight, height for children), Root Mean Square Energy, Temporal centroid, Log attack time, maximum, minimum, and variance] as well as the last features sets is constructed on amplitude distribution from histogram. Second, frame features, the respiratory signal is divided into equal, non-overlapping frames whose duration is set at 10 ms. Depending on the extent of the cracking period.

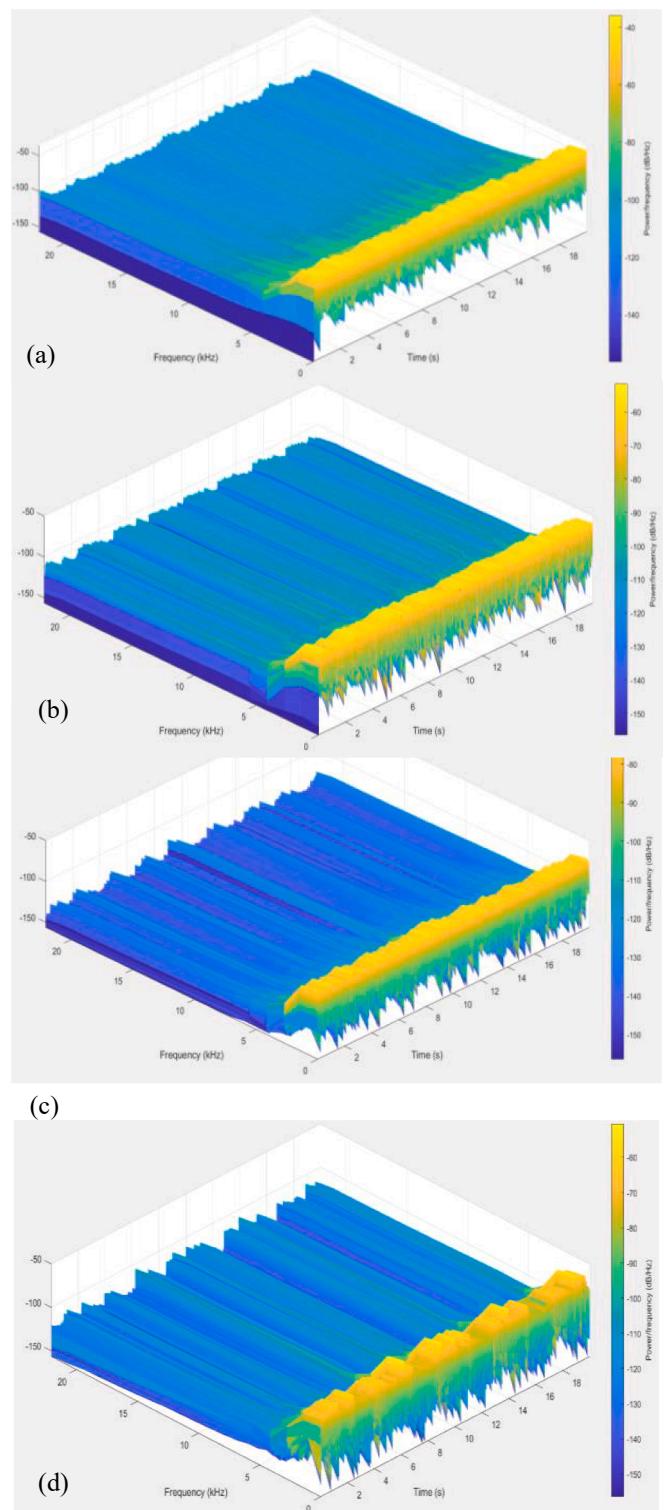


Fig. 7. 3D feature representation of four respiratory subcases: (a) Normal, (b) Wheezes, (c) Crackles, and (d) both.

For this feature type, three groups of Zero Cross Rate(ZCR) parameters are extracted [rate, count, indices] for a window length equal to 10 ms, as shown in Figs. 2 and 3. Third ZCR for respiratory sub-cycles, wheezes, crackl-Es, and both crackles and wheezes using the same window length.

The third feature type is done on the frames overlapping in extracting MFCC. In this proposed system, the frame length of 25 ms was adopted, with an overlap of 50 %. Thirty-nine features were extracted: twelve for

Table 6

Resource utilization of the (Layer1-3) on MicroBlaze soft processor and Layer-4 hardware core.

Device	MicroBlaze Processor core	Hardware Core
Max. Frequency	100MHz	
DSP48 Unit	3 %	79 %
BRAM	12 %	82.17 %
LUT	34.4 %	43.6 %
FF.	58.8 %	42.04 %

Table 7

CPU-GPU and FPGA system comparison.

System	CPU-GPU	Soft-Hard FPGA
Platform	Core i7-1165G7-NVIDIA GeForce MX350	XC5VLX50T-1FFG665C
Latency (ms)	33	17
Throughput (GOPs)	325.8	815.6
Power (W)	41.5	12.5
Efficiency (GOPs/W)	7.85	65.25

MFCC coefficients, twenty-six for first and second-order derivatives, and one for energy.

The random forest is one of the most essential supervised machine learning methods, as it can be used for classification and regression based on the ensemble learning concept. Its importance lies in its ability to process parallelism and deal with high-dimensional problems with little computational time compared to other methods. Hierarchical random forests have been proposed, and they are composed of three levels of random forests as structured for feature extraction. The first layer is dedicated to the characteristics extracted from the total respiratory recording of patients. The second layer is for properties of the frame's level, and the third is for MFCC features of overlapping frames. Seven forests in each layer, the data stops traversing all forests if the first four forests get a high diagnosis, even if they are not adjacent. Each forest has been divided into multiple subsets operated in parallel, with eight trees for each process element, as demonstrated in Fig. 4. The proposal of deep hierarchical random forests (DHRF) made it possible to complete the classification for many cases in the first layer and others in the second. Others in the third layer, and for a few cases, in the deep training model as a final stage.

Deep learning is adopted for image classification in the fourth layer [32,33]. To accomplish the deep learning stage at the lowest cost and with high accuracy, lightweight convolutional neural networks such as the ShuffleNet series, SqueezeNet series, and MobileNe series methods can be used. Table 4 demonstrates the complexity of three lightweight models in terms of the number of parameters (Params), Multiply-add operation (M-adds), and floating point of operations (FLOPs).

From the table, it is clear that the first two models are close in terms of performance in terms of consumption and that the third model has the highest consumption. Despite the low consumption of the second model, the first has higher accuracy, as will become clear later. The proposed MobileNet architecture uses six Zeropadding2D, fourteen ReLU layers, sixteen Conv and twenty-eight BatchNormalization, and fourteen DepthwidthConv2D and classifier. According to this model, the image size has been changed to match the size of the input layer 224*224*3. Training and testing of the four proposed layers were performed on a laptop with an Intel Core™ 11th Gen i7-1165G7@2.80 GHz CPU, 16 GB RAM, and NVIDIA GeForce MX 350.

their reconfigurability, performance efficiency [34], and low cost. Newer versions of FPGAs contain soft processors besides hardware resources that are considered limited, especially for larger designs [35]. Xilinx vendor supplied its Virtex-5 with a MicroBlaze soft processor [36]. It is a 32-bit reduced instruction set (RISC) computer with a Harvard architecture that was provided as the intellectual property (IP) core

in the Embedded Development Kit (EDK) [37,38]. The MicroBlaze softcore processor has a fixed feature set that includes 32-bit general-purpose registers, 32-bit instruction words with three operands, and two addressing modes. Default 32-bit address bus, extensible to 64-bit and single-issue pipeline [39–41]. It can be configured with The AMBA AXI4 for peripheral and cache interfaces, 32-bit PLB V4.6 interface, and simple synchronization for efficient RAM transfers through the LMP protocol. FSL delivers a fast, non-arbitrated streaming communication. XCL offers a fast, slave-side arbitration stream interface between external memory controllers and caches, a Microprocessor Debug Module (MDM) interface, and a trace interface for presentation analysis [39].

The ML506 board has an AC97 audio codec for audio processing. AD1981 Analog Devices audio codec supports stereo 16-bit audio with up to 48-kHz sampling [42]. Xilinx Virtex-5 board has separate audio jacks for the Microphone, Line In, Line Out, and headphones. Distribution of the works between the hardware components of the FPGA and the soft processor can help with the problem of limited resources, especially for large designs, with high flexibility for any improvement, especially with the soft processor task.

5. Results and discussion

The ICBHI dataset was classified into eight respiratory diseases as mentioned in Table 3, i.e., (1) case of asthma, (2) case of lower respiratory tract infection (LRTI), (23) URTI, (793) for COPD, (37) for Pneumonia, (29) bronchiectasis, and (35) healthy cases. There is only one sample for asthma and two for (LRTI). Therefore, they were excluded from the classification process. COPD cases are much larger than other classes, so the set is severely imbalanced. To overcome this problem, augmentation techniques were utilized, such as noise addition, pitching, and time shifting have been performed. Then, this augmentation data was randomly divided into (80, 20) independent training and testing splits. Optimum values for tuning parameters were determined using cross-validation. Through partition of the training data into four sets, stated as folds, and pending each fold, one at a time. The three remaining folds were trained and used to predict the cases in the blocked fold. The results of the proposed system are illustrated in Table 5. The overall confusion matrix is plotted in Fig. 5.

Table 5 shows the system performance when using MobilenetV2 at the last layer. When replaced with SqueezeNet, although the power consumption is slightly lower, the accuracy decreases by 0.46, 0.75, 0.34, 0.56, and 0.59 for health, urinary tract infection, COPD, bronchiectasis, and pneumonia, respectively. In contrast, the accuracy increased slightly by 0.08, 0.07, 0, 0.03, and 0.023 with the shuffleNetV1 model. On the other hand, the power consumption increased in a way that did not equal the increase in accuracy achieved, so the mobile model was adopted.

A novel hierarchy approach made it possible to finalize the classification in different strata and determine which is more important for each case. For example, in chronic obstructive pulmonary disease, the characteristics of the first layer were the most important, as shown in Fig. 6. The presence of clinical features, age, and the patient's sex prevented the error for three cases where smokers are more likely to be infected. The second and third layers had the largest share in achieving health cases and features of URTI disease in the first and third layers. The second and third for pneumonia are demonstrated in Fig. 6.

Additionally, the proposed system has four subclasses (Normal, Wheezes, Crackles, and both Wheezes and Crackles). The dataset includes 6898 respiratory cycles, 3642 regular cycles, 886 wheezes, 1864 crackles, and 506 cycles containing both wheezes and crackles. The distributions of the four sub-cases are demonstrated in Fig. 7.

Different stethoscopes were used to record the lung audios, i.e., AKGC417L, LittC2SE, Litt3200, and Meditron. For the ICBHI data, most of the percentage of registrations were for AKGC417L(63 %), LittC2SE (7 %), Litt3200(9 %), and 21 % for Meditron. General, the AKG C417L

Table 8

Comparison table of the proposed system with related research.

Ref.	Database	Methods	Classes	Performance Accuracy%	System	Platform	flexible
[43]	131 participants using mobile devices	GBT ¹	Wheezes/Non	94.6	software	NA	Not
[26]	ICBHI 2017	LPCC-MLP ²	Healthy/Non	99.22	software	NA	Not
[44]	HF_Lung_V2	CNN-BiGRU ³	I, E, C and D ⁴	93.1, 92.4, 88.4 and 87.8	Python 3.7	CPU: Intel(R) Xeon(R) Gold 6154 GPU: NVIDIA Titan V100	Not
[45]	ICBHI 2017	Co-tuning and stochastic normalization with different depths of ResNet	2-classes healthy/unhealthy 3-classes Healthy, Chronic, and Non-Chronic Diseases 4-classes Normal, Crackles, Wheezes, and both	93.77 ± 1.41 92.72 ± 1.30 64.74 ± 0.05 %	software	NA	Not
[46]	103 patients + 111 patients from ICBHI 2017	CNN + BDLSTM ⁵	Normal asthma, pneumonia, BRON, COPD, and HF ⁶	98.8, 95.6, 98.8, 100, 99, and 100	MATLAB R2020a	CPU: Intel (i7-9700) GPU: NVIDIA GeForce GTX 1070	Not
[4]	ICBHI 2017	hybrid 2D CNN-LSTM	Normal, Crackles, Wheezes, and both	76.39	Python 3.7.7	Intel Core i5-9600 K NVIDIA GeForce RTX 2070	Not
[47]	ICBHI 2017	CNN6	Normal/abnormal	75.95 ± 2.31	software	GPU: Nvidia 3090	Not
[48]	ICBHI 2017	1D wavelet smoothing, displacement artifact removal, and z-score normalization	COPD, Healthy, URTI, Bronchiectasis, Pneumoine	99, 80, 67, 100, 33	MATLAB R2019	Intel Core i7 GPU: RAM.A 4 GB	Not
[49]	ICBHI 2017	Blnet Network	Wheezes/Non	72.27	PyTorch in python	Intel i9-10900 k NIVIDA RTX2080	Not
Proposed System	ICBHI 2017	Random Forest(layer1-3) + MobileNet	Five classes (Healthy, URTI, COPD, Bronchiectasis, and Pneumoia) with four subclasses (Normal, Crackles, Wheezes, and both).	(99.76, 99.64, 100, 99.82, 99.86) & 99.5, 98.78, 97.5, and 96.9 for four subclasses	Python 3.7.7, assembly language for the soft processor VHDL for the hardware of MobileNet	Core i7-1165G7- NVIDIA GeForce MX350 & XC5VLX50T -1FFG665C	YesMicroblaze random forest program (Layer1-3) can be easily modified.

Note:

¹GBT: Gradient-Boosted Tree, ²LPCC-MLP: LineaMOr Predictive Cepstral Coefficient with Multilayer Perceptron (MLP)-based classifier, ³CNN-BiGRU: bidirectional gated recurrent unit, ⁴I, E, C, and D: Inhalation, Exhalation, continuous adventitious sound and discontinuous adventitious sound, ⁵CNN + BDLSTM: bidirectional long short-term memory, ⁶HF: heart failure.

recording microphone delivered better results due to the mic's sensitivity, not having filters and the fact that its training and testing records are larger than other types. Moreover, (70 %) of the data set was recorded by multichannel equipment and yielded better results than a single channel.

The proposed system was initially trained on a CPU and GPU pair. Librosa was used for feature extraction, sci-kit-learn for the random forest, CUDA for executing the GPU program, and TensorFlow for the last deep learning layer. Virtex-5 ML506 FPGA was adopted as a hardware platform with its Microblaze soft processor. Table 6 shows the resource utilization of the proposed system on the Virtex-5 ML506 board. The performance comparison of CPU-GPU and Soft-hard FPGA system is demonstrated in Table 7.

Table 8 describes the performance of the proposed system compared to recent studies. The summary table covers the most recent studies for lung sounds classifications.

6. Conclusion

The design and implementation of an innovative diagnostic tool for respiratory diseases that operates in real-time was proposed. The first respiratory processor was built with four layers and different features for each layer. Hierarchical design allows it to obtain high accuracy without completing the four layers to save energy and time consumption, as the transition to the next layer can be stopped the moment an accurate classification is obtained.

The dataset recorded by different devices was adopted, making the processor more comprehensive. Variations in the number of recordings for different diseases resulted in lower accuracy for some diseases despite the data argumentation and cross-validation, which can be overcome once more data is available.

Gradual distribution of features across the four layers of the whole record, frame, and overlapping frames led to reducing processor consumption while increasing classification speed. That was helped by excluding the completion of entry into the seven forests assigned to the data from different locations when high values were obtained for the

first four forests and designing the processing elements to work in parallel in each forest. Which reduced power consumption by 63.8 % and time by 54.7 % compared to completing all layers for all data, allowing it to be used with mobile devices.

Adding patients' clinical data to the system features helped save five cases from misclassification. Also, using frames with 50 % overlapping in extracting the Mel-Frequency Cepstral Coefficients in third layer features helped save 24 % of this layer cases from misclassification. Adding the classification of the four sub-periods (crackle, whistling, other than both) contributed significantly to achieving high accuracy in classifying the five diseases. Three models were tested in the last deep learning layer (SqueezeNet, MobileNetV2, and shuffleNetV1). The first model was lighter weight and had more minor accuracies. ShuffleNetV1's third model is heavier, and accuracies approximated that of the second model. So, MobileNetV2 was adopted in the hardware system.

The Resilient Respiratory Processor (RRP) is very flexible, as the soft processor unit can programmatically change the first three layers.. Embedding the proposed processor in mobile devices is very important for classifying all respiratory diseases whenever the necessary data are available, in addition to using Virtex-7 instead of Virtex-5 due to their higher performance. The system will be a valuable tool for medical professionals and can be used easily, especially since it is integrated with the AC97 audio codec in the ML506 board.

For future work, determining appropriate drug doses can be added after diagnosing cases, especially with patients' clinical data. In addition, the availability of extensive data for many respiratory diseases during the training stage helps to generalize the proposed respiratory therapist.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgments

With appreciation for the support provided by the University of Mosul, College of Engineering, Department of Computer Engineering.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bspc.2023.105876>.

References

- [1] R. Zulfiqar, F. Majeed, R. Irfan, H.T. Rauf, E. Benkhelifa, A.N. Belkacem, Abnormal Respiratory Sounds Classification Using Deep CNN Through Artificial Noise Addition, *Front. Med.* 8 (2021), 714811, <https://doi.org/10.3389/fmed.2021.714811>.
- [2] Y. Kim, Y. Hyon, S.S. Jung, S. Lee, G. Yoo, C. Chung, T. Ha, Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning, *Sci. Rep.* 11 (2021) 17186, <https://doi.org/10.1038/s41598-021-96724-7>.
- [3] R. X. Pramono, S. Bowyer, E. Rodriguez-Villegas, "Automatic adventitious respiratory sound analysis: A systematic review", *PLoS ONE* 12(5): e0177926, 43Pages, (2017). 10.1371/journal.pone.0177926.
- [4] G. Petmezas, G. Cheimariotis, L. Stefanopoulos, B. Rocha, R.P. Paiva, A. Katsaggelos, N. Maglaveras, Automated Lung Sound Classification Using a Hybrid CNN-LSTM Network and Focal Loss Function, *Sensors* 22 (2022) 1232, <https://doi.org/10.3390/s22031232>.
- [5] P. S. Faustino, "Crackle and wheeze detection in lung sound signals using convolutional neural networks", MSc. Thesis, Engenharia de Redes e Sistemas Informáticos Departamento de Ciência de Computadores, (2019). <https://repositorio-aberto.up.pt/>.
- [6] A. Rizal, R. Hidayat and H. A. Nugroho, "Signal Domain in Respiratory Sound Analysis: Methods, Application and Future Development", *Journal of Computer Sciences*, 11 (10): 1005.1016, 12Pages, (2015). 10.3844/jcssp.2015.1005.1016.
- [7] D. F. Bastos, "Differential Diagnosis of Respiratory Pathologies: a Data-Driven Approach based on Respiratory Sounds", MSc. Thesis, University of Coimbra, Biomedical Engineering, (2021). eg.uc.pt.
- [8] A.N. Mazumder, H. Ren, H. Rashid, M. Hosseini, V. Chandraseddy, H. Homayoun, T. Mohsenin, Automatic Detection of Respiratory Symptoms Using a Low Power Multi-Input CNN Processor, *IEEE Des. Test* 39 (3) (2021) 82–90, <https://doi.org/10.1109/MDAT.2021.307931.8>.
- [9] A. A. Saraiva, D. B. S. Santos, A. A. Francisco, J. V. M. Sousa, N. M. Fonseca Ferreira, S. Soares and A. Valente, "Classification of Respiratory Sounds with Convolutional Neural Network", In Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2020), 138-144, (2020), 10.5220/000896510138 0144.
- [10] L. Li, A. Ayigili, Q. Luan, B. Yang, Y. Subinuer, H. Gong, A. Zulipikaer, J. Xu, X. Zhong, J. Ren and X. Zou, "Prediction and Diagnosis of Respiratory Disease by Combining Convolutional Neural Network and Bi-directional Long Short-Term Memory Methods", *Front. Public Health* 10:881234, 9Pages, (2022), 10.3389/fpubh.2022.881234.
- [11] C. Wall, L. Zhang, Y. Yu, A. Kumar and R. Gao, "A Deep Ensemble Neural Network with Attention Mechanisms for Lung Abnormality Classification Using Audio Inputs", *Sensors* 22, 5566, 25Pages, (2022), 10.3390/s22155566.
- [12] S. Bharatia, P. Podder, M.R.H. Mondala, V.B.S. Prasath, CO-ResNet: Optimized ResNet model for COVID-19 diagnosis from X-ray images, *Int. J. Hybrid Intell. Syst.* 17 (2021) 71–85, <https://doi.org/10.3233/HIS-210008>.
- [13] R. Li, C. Xiao, Y. Huang, H. Hassan and B. Huang, "Deep Learning Applications in Computed Tomography Images for Pulmonary Nodule Detection and Diagnosis: A Review", *Diagnostics*, 12, 298, 21Pages, (2022), 10.3390/diagnostics 1202 0298.
- [14] https://bhichallenge.med.auth.gr/ICBHI_2017_Challenge.
- [15] E. Fonseca, M. Plakal, F. Font, D. P. W. Ellis, X. Favory, J. Pons, X. Serra, "General-purpose Tagging of Freesound Audio with AudioSet Labels: Task Description, Dataset, and Baseline", arXiv: 1807.09902v3 [cs.SD], (2018), 10.48550/arXiv.1807.09902.
- [16] B. M. Rocha, D. Filos, L. Mendes, G. Serbes, S. Ulukaya, Y. P. Kahya, N. Jakovljević, T. L. Turukalo, I. M. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, N. Maglaveras, R. Pedro Paiva, I. Chouvarda, P. d. Carvalho, "An open access database for the evaluation of respiratory sound classification algorithms", *Physiological measurement*, vol. 40, no. 3, Pp.035001, (2019). 10.1088/1361-6579/ab03ea.
- [17] B. Lin, T. Yen, An FPGA-Based Rapid Wheezing Detection System, *Int. J. Environ. Res. Public Health* 11 (2014) 1573–1593, <https://doi.org/10.3390/ijerph110201573>.
- [18] B. Lin, B.H. Wu, F. Chong, S. Chen, "Biomedical Engineering Applications, Basis & Communications", Vol. 18 No. 3 June (2006), 10.4015/S1016237206000221.
- [19] M. Simić, A. K. Stavrakis, A. Sinha, V. Premcevski, B. Markoski, and G. M. Stojanović, "Portable Respiration Monitoring System with an Embroidered Capacitive Facemask Sensor", *Biosensors*, 12, 339, (2022). 10.3390/bios12050339.
- [20] T. Ahmed Khaleel and J. Abdullah kareem, "A Review of Issues and Challenges to Address the Problem of Implementing Green Computing for Sustainability", *Al-Rafidain Engineering Journal*, Vol. 28, Issue 1, P. 300-311, March 2023. 10.33899/rengi.2022.135485.1197.
- [21] A. Kumar, K. Abhishek, C. Chakraborty, and N. Kryvinska, "Deep Learning and Internet of Things Based Lung Ailment Recognition Through Coughing Spectrograms", Special Section on Intelligent Big Data Analytics for Internet of Things, Services and People, *IEEE Access*, Vol. 9, Pp.95938-95948, (2021). 10.1109/ACCESS.2021.3094132.
- [22] W. Wang, Y. Pei, S. Wang, J. manuel Gorz, and Y. Dong Zhang, "PSTCNN: Explainable COVID-19 diagnosis using PSO-guided self-tuning CNN", *Biocell*, 47 (2), P373-384, (2023). 10.32604/biocell.2021.0xxx.
- [23] M. Saeed, M. Ahsan, M. H. Saeed, A. U. Rahman, M. A. Mohammed, J. Nedoma, and R. Martinek, "An algebraic modeling for tuberculosis disease prognosis and proposed potential treatment methods using fuzzy hypersoft mappings", *Biomedical Signal Processing and Control*, Vol. 80, Part 1, 104267, (2023). 10.1016/j.bspc.2022.104267.
- [24] M. Saeed, M. Ahsan, M. H. Saeed, A. Mehmood, H. A. Khalifa, and I. Mekawy, "The Prognosis of allergy-Based Diseases Using Pythagorean Fuzzy Hypersoft Mapping Structures and Recommending Medication", *IEEE Access*, vol. 10, pp. 5681-5696, (2022). 0.1109/ACCESS.2022.3141092Vol. 10, 2022.
- [25] A.U. Rahman, M. Saeed, M.H. Saeed, D.A. Zebari, M. Albahar, K.H. Abdulkareem, A.S. Al-Waisy, M.A. Mohammed, A Framework for Susceptibility Analysis of Brain Tumours Based on Uncertain Analytical Cum Algorithmic Modeling, *Bioengineering* 10 (2023) 147, <https://doi.org/10.3390/bioengineering10020147>.
- [26] B. M. Rocha, D. Filos, L. Mendes, G. Serbes, S. Ulukaya, Y. P. Kahya, N. Jakovljević, T. L. Turukalo, I. M. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques, N. Maglaveras, R. P. Paiva1, I. Chouvarda, P. Carvalho, "An Open Access Database for the Evaluation of Respiratory Sound Classification Algorithms", *Physiol. Meas.* 40(3): 035, 001, (2019), 10.1088/1361-6579/ab03 ea.
- [27] H. Mukherjee, P. Srerama, A. Dhar, S. Obaidullah, K. Roy, M. Mahmud, K. C. Santosh, Automatic Lung Health Screening Using Respiratory Sounds, *J. Med. Syst.* 45 (2021) 19, <https://doi.org/10.1007/s10916-020-01681-9>.

- [28] G. Sharma, K. Umapathy and S. Krishnan, "Trends in audio signal feature extraction methods", Applied Acoustics, 158, 107020, 21pages, (2020), <https://doi.org/10.1016/j.apacoust.2019.107020>.
- [29] H.A. Modran, T. Chamunorwa, D. Ursutiu, C. Samoilă, H. Hedesiu, Using Deep Learning to Recognize Therapeutic Effects of Music Based on Emotions, Sensors 23 (2023) 986, <https://doi.org/10.3390/s23020986>.
- [30] A. Ksibi, N.A. Hakami, N. Alturki, M.M. Asiri, M. Zakariah, M. Ayadi, Voice Pathology Detection Using a Two-Level Classifier Based on Combined CNN-RNN Architecture, Sustainability 15 (2023) 3204, <https://doi.org/10.3390/su15043204>.
- [31] K. Bhangale and M. Kothandaraman, "Speech Emotion Recognition Based on Multiple Acoustic Features and Deep Convolutional Neural Network", Electronics, 12, 839, (2023). 10.3390/ electronics12040839.
- [32] F. Hong, D. W. Liang Tay and A. Ang, "Intelligent Pick-and-Place System Using MobileNet", Electronics, 12, 621, (2023). 10.3390/ electronics12030621.
- [33] K. Sabancı, Benchmarking of CNN Models and MobileNet-BiLSTM Approach to Classification of Tomato Seed Cultivars, Sustainability 15 (2023) 4443, <https://doi.org/10.3390/su15054443>.
- [34] R.D. Sala, G. Scotti, A Novel FPGA Implementation of the NAND-PUF with Minimal Resource Usage and High Reliability, Cryptography 7 (2023) 18, <https://doi.org/10.3390/cryptography70120018>.
- [35] V.J. Vazhoth, "Embedded processors on FPGA: Hard-core vs Soft-core", Master of Science in Electrical Engineering, Padnos College of Engineering and Computing, Grand Valley State University, 2017.
- [36] Y. Gao, Designing an IEEE Floating-Point Unit with Configurable Compliance Support and Precision for FPGA-Based Sof-Processors. Master of Applied Science, in School of Engineering Science Faculty of Applied Sciences, Simon Fraser University, 2022.
- [37] Xilinx, "Vivado Design Suite User Guide Programming and Debugging", UG908 (v2022.1), (2022). www.xilinx.com.
- [38] S. Hema Chitra, K. Sarvesh, M. Preetha, "Comparison of Different Configurations of MicroBlaze Soft IP Core", International Journal of Applied Engineering Research ISSN 0973-4562 Vol. 14, No. 8, Pp.1965-1969, (2019). <http://www.ripublication.com>.
- [39] Xilinx, "MicroBlaze Processor Reference Guide", UG984 (v2020.2) November 18, (2020). www.xilinx.com.
- [40] Xilinx, "Vivado Design Suite Tutorial Embedded Processor Hardware Design", UG940 (v2021.1) June 16, (2021). www.xilinx.com.
- [41] I. Zagan, V.G. Găitan, Soft-core processor integration based on different instruction set architectures and field programmable gate array custom datapath implementation, PeerJ Comput. Sci. 9 (2023) e1300.
- [42] Xilinx, ML505/ML506/ML507 Evaluation Platform User Guide, UG347 (v3.1.1) October 7, (2009). www.xilinx.com.
- [43] S. Chatterjee, M. Rahman, T. Ahmed, N. Saleheen, E. Nemat, V. Nathan, K. Vatanparvar, J. Kuang, Assessing Severity of Pulmonary Obstruction from Respiration Phase-Based Wheeze Sensing Using Mobile Sensors, ACM Trans. Math. Software (2020), <https://doi.org/10.1145/3313831.3376444>.
- [44] F. Hsu, S. Huang, C. Huang, Y. Cheng, C.h Chen, J. Hsiao, C. Chen and F. Lai, "A Progressively Expanded Database for Automated Lung Sound Analysis: An Update", Appl. Sci., 12, 7623, (2022). 10.3390/app12157623.
- [45] T. Nguyen, F. Pernkopf, Lung Sound Classification Using Co-Tuning and Stochastic Normalization, IEEE Trans. Biomed. Eng. 69 (9) (2022) 2872–2882. <https://ieeexplore.ieee.org>.
- [46] M. Fraiwan, L. Fraiwan, M. Alkhodari, O. Hassannin, Recognition of pulmonary diseases from lung sounds using convolutional neural networks and long short-term memory, J. Ambient Intell. Hum. Comput. 13 (2022) 4759–4771, https://doi.org/10.1007/s12_652_021_03184_y.
- [47] I. Moummad and N. Farrugia, "Pretraining Respiratory Sound Representations using Metadata and Contrastive Learning", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October 22-25, (2023), New Paltz, NY. <https://imt-atlantique.hal.science/hal-04165413>.
- [48] A.A. Dawud, G. Haile, Deep Learning-Based Pneumonia Classification Based on Respiratory Sounds, J. Art. Intell. Cloud Computing 2 (2) (2023) 1–6, [https://doi.org/10.47363/JAICC/2023\(2\)115](https://doi.org/10.47363/JAICC/2023(2)115).
- [49] R. Yang, K. Lv, Y. Huang, M. Sun, J. Li, J. Yang, Respiratory Sound Classification by Applying Deep Neural Network with a Blocking Variable, Appl. Sci. 13 (2023) 6956, <https://doi.org/10.3390/app1312656>.