International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Transfer Learning Approaches for Lung Sound Detection

Jemisha A Patel[1], Rikita Gohil[2], Dwiti Pandya[3], Pratik Kahar[4]

[1,2,3,4]Assistant Professor, Vadodara Institute of Engineering, Vadodara

Abstract

lung disease is now the biggest cause of death in the whole world. For the most part, lung disease is found in its last stages after it has already progressed to a serious state. One on either hand, early detection of lung disease may aid in therapy. In today's society, technological advancements are critical to the delivery of healthcare services. To record the sounds emitted by the patients' lungs, we use electronic stethoscopes. The sounds made by the lungs give useful information for lung diagnosis. Study in the medical profession is now concentrating on the significance of diagnosing lung disease using lung acoustics, which is a current research topic. Transfer learning is critical for success in the medical system. In this study, we provide a range of transfer learning algorithms for lung sound classification, including ALEXNET, MOBILENET, VGGNET, and RESNET, among others. For the new data set, the WAVGAN model will be used to produce a new data set, which will be used to perform recognition of respiratory system noises with in Transfer learning model. With their high accuracy in classifying lung sounds, our transfer learning models might one day be used to diagnose lung disorders. Transfer learning strategies and their advantages and disadvantages will be discussed in this article. There are other dangers to consider. As a means of distinguishing between the four distinct lung sounds, in addition, it recommends future directions for research into lung sound identification.

Keywords AlexNet; Vggnet; Resnet; Random Forest; Naive Bayes; Support Vector Machine; Convolutional neural network; Wavegan; and Artificial neural network.

1. Introduction

The third greatest cause of mortality in the world is lung disease. As shown in the Health Ministry (WHO), upwards of 3 million people worldwide die every year from respiratory infections [1]. A total of 235 million individuals has asthma and a total of 200 million have chronic obstructive lung disease (COPD), with 65 million having feelings of extreme COPD. More than 8.7 million individuals are infected with tuberculosis (TB) every year, and 1–6 percent of adults experience respiratory difficulties when sleeping [2]. In this scenario, lung sounds are critical in identifying respiratory diseases and infections [10]. Lung sounds are acoustic signals produced by breathing. Physicians have commonly used an auscultatory approach to analyses lung sounds linked with various respiratory ailments. Asthma, pneumonia, and bronchiectasis are among respiratory disorders that may be diagnosed with the auscultatory approach. A manual process, however, it consumes a lot of effort and has the potential for greater or lower accuracy because of how complex the sound patterns are. There is a significant risk of missing data, which might result in underdiagnoses or inaccurate results [3].

The use of peak flow meter values as well as a questionnaire-based technique, in conjunction with lung sounds, is determined by examining lung problems. The identification of wheezes and crackles from lung sounds is accomplished by the application of property supervised learning on spectrogram-based features [1]. An alternative framework, on the other hand, may pick properties that may not be actually open to interpretation and may not have a predetermined error tolerance. A collection of features that is easy to understand, on either contrary, is more beneficial to the health sector as a whole. The identification of lung sound abnormalities has been accomplished by the use of a broad range of auditory stimuli, including Mel - frequency cepstral functionalities, feature descriptors, and time-frequency analysis approaches, among others. Different classification methods, such as neural networks and hidden Markov models, were also used. The bulk of these research either incorporate the concomitant recording of an extra flow indicator in order to detect the respiration stages through the use of classifying of the bronchial phases in order to assess the breathing phases.
After that, we'll go further into each of the issues mentioned below. Section II will address some of the most important advances in voice recognition technology. Section III contains a thorough examination of the many approaches used in the development of this framework. In Section IV, describe the result and analysis. Finally, towards the completion of this study, a list of recommendations for more research is offered.

2. Proposed System

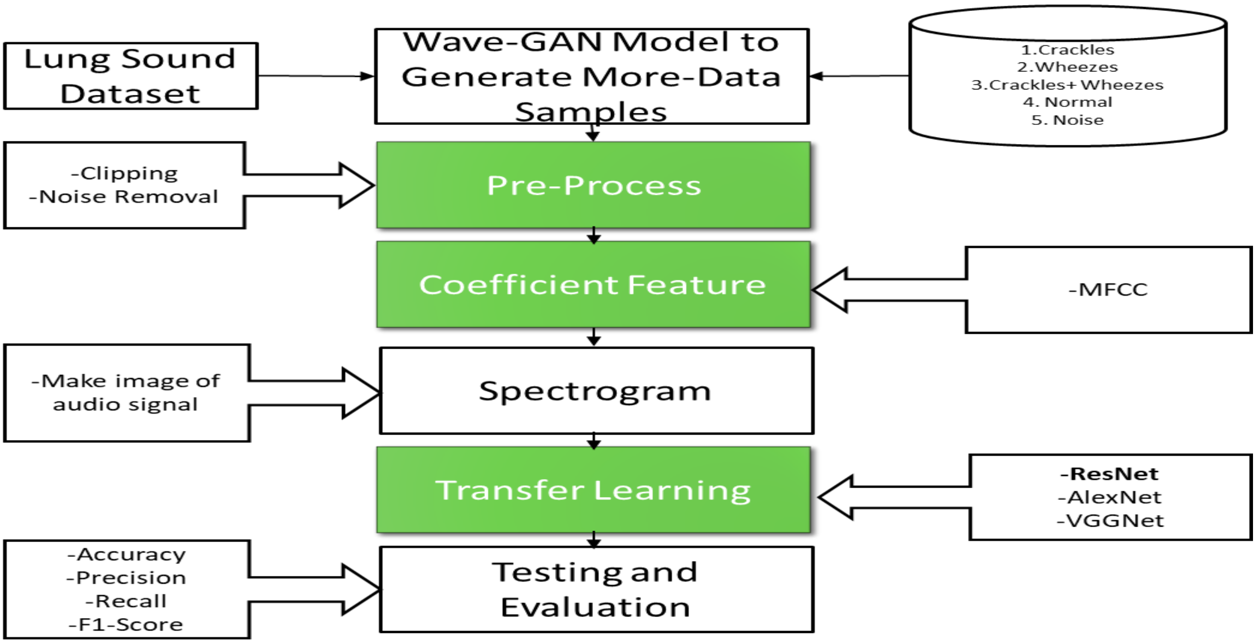

Fig. 1. Proposed System

A. Data set

In all, 920 tagged recorded interviews were submitted to the ICBHI 2017 data by 126 participants, resulting in a total of 920 entries. Analog signals were captured using a variety of headsets to capture the information. In the course of recording, the sample rate is varied between 4000Hz and 44100Hz, and the capturing period is varied between 10 seconds and 90 seconds. The duration of each recording is controlled by a specified range of breathing cycles, as well as short comments at the beginning and conclusion for every cycle, and not whether crackles and/or wheezes can be heard throughout the duration of the record. We were capable of separating voice tapes into lung cycles with the use of database labels and tags. The cycle time may vary around 0.2 to 16 seconds, including an estimated 7 seconds, depending on the cycle. 6898 respiration cycles are predicted to be saved in the dataset, with 3642 regular cycles, 1864 crackling sounds, 886 wheezes, and 506 periods which include crackles and wheezes getting recorded.

B. Data generation

In this scenario, the WaveGAN paradigm is applied for the goal of information extraction from the data. Without exception, WaveGAN is generally used for unaided raw waveform audio synthesis, with the exception of spectrograms that are similar to images. The WaveGAN model is based on the DCGAN architecture and is designed to work with it. It is possible to continually upgrade constrained morphological procedures into a more complete picture by using the transposed convolution approach, which is implemented by the DCGAN generator. WaveGAN makes changes to this transposed convolution approach in order to broaden its receptive spectrum and increase its sensitivity. an increased one-dimensional screening with a range of 25 instead of a measure of 5x5, and raising discoveries by a quarter point rather than two points on every level the discriminator, on the other hand, is enhanced by including length-25 filters into a single dimension and extending the pace from 2 to 4 times, respectively. WaveGAN now contains the same number of parameters as the previous DCGAN, computes the same number of overall outcomes, and meets all the requirements that are multidimensional when compared to the prior DCGAN.

C. Feature Extraction

MFCC: Mel Frequency (MFCC) [1, 7, 10, 11] is a harmonic spectrum that may be found in a variety of musical genres and instruments. For the purposes of this research, the computer aided design Coefficients were used to define the characteristics of sound samples. A large number of MFCCs are used by speech recognition systems in order to discriminate between various forms of speech. Because it allows for the evaluation of the relatively brief transmission lines of time - frequency indicators, the simple power sub band of wavelet-based signals was already widely used in previous survey on the proof of identity of due to defective cardiovascular look and sound in time domain data, particularly in the diagnosis of malfunctioning respiratory tones in time domain data. In order to distance between various tangential noises that arise in a single capturing at different points in time and for varying lengths of time, it is necessary to distinguish between their frequencies and structures, as these characteristics are important in identifying and categorizing the sounds in question. Thus, the MFCC is beneficial for capturing the change in signal amplitude of a recorded as a function of the signal's temporal development. An exponential scale of pitches whereby the distances between frequencies are seen as being similar through the grownup auditory cortex is referred to as the Mel scale, and this is a phenomenon known as equalization. The continuity formula is used to determine the Mel scale's harmonics. The MFCC creates its double depth image (time and frequency) which is collapsed for a one array before it is transferred to the host computer for use in other operations.

Spectrogram: Using several frequencies inside a certain waveform design, which is acknowledged as the spectrum representation, it is possible to see the intensity of a signal or the sound it produces over a lengthy period of time, which is mentioned as even the spectrum representation. In waveform analysis, this is known as a spectral representation. Also shown in the graph is how energy levels fluctuate over time as a consequence of the passage of time. By using an edge modelling method, we can convert these organ tones into spectrogram images that can then be employed as part of our hypothesized insideoutside modelling strategy. Their invention was made possible through the use of the Longinus Color matching Map, which should be a useful tool for creating power spectral pictures by using hues ranging from blue to green and yellow. Viridis Coloring Map is a useful tool for creating power spectrum pictures by using hues ranging from blue to green and yellow.

CNN [1,2,3,6]: The classification of images is performed by something like a neural net classification approach, which is described below. In this particular instance, it is termed to with a "Convolutional Neural Network" (CNN). When opposed to traditional neural networks, which connect each incoming symbol to a distinct range of constraints, CNN criteria are maintained across all input features, resulting in a more efficient algorithm. This gives the network with the opportunity to learn more about the distinctive qualities of the local region, which is helpful to all parties involved. As a result, CNN gains knowledge of the relevant qualities, and no further recommended approach is required. On top of data, CNN-based architectures extract tough attributes by using resulting solution that have been learnt from the data. These convolutional kernels are then utilized to build a deep architecture on place the emphasis. CNN is also a computing platform that is quite productive. We may be able to accomplish parameter sharing and other more efficient computing by using convolutional approaches.

D. Transfer learning Approaches:-

VGG [1,10]: In visual geometry recognition, the VGG (Visual Geometry Group) framework is a complex DCNN structure featuring numerous convolution layers that is used in conjunction with several convolution layers. A significant factor is the frequency of convolution operations performed; for example, VGG-16 has 16 layers, but VGG-19 has 19 layers. This design, known as VGG, serves as the basis for an expanded item identification model, which is constructed atop the VGG architecture. The Objects tend neural network, which has been built, is an illustration of a neural network.

Alexnet [1,5]: A large perceptron (rnn) may be able to achieve high excellent on a highly difficult dataset by using solely supervised learning methodologies, according to the findings of the AlexNet study. In the year after the debut of AlexNet, a competition was launched that has continued to this day. The Convolutional Neural Network is used to categories all contributions to the ImageNet database. CNN is a pioneer in biomedical research, ushering in a new age with AlexNet, which was created in collaboration with the National Institutes of Health and launched in 2004. Because a variety of deep learning are readily available, the mounting of AlexNet is rather basic.

Resnet [2,10]: It is a construction piece that was destroyed but which still contains a bridged connector (formerly known as a legacy connection) that permits data to flow through it without being altered in any way despite the destruction. The data signal x is converted into an output signal F by the activation curve layer. It is composed of two types: layer 1 and layer 2, which are interconnected (x). The transfer seems to be pretty comparable to those of a connection that has been skipped in this instance. The residual unit in this specific design demonstrates how it control signal x varies from those of the thanks to advances F, which is a result of the construction process itself (x). In accordance with the findings of this study, if the infrastructure has also fruitfully recreated the linear mapping that is assembled on a given spot, the improvements may be effective to minimize muscular endurance in the unavailable slabs on varying scales to essentially zero, but also guarantee that the output passes across the disconnect with next to no damage.

WAVGAN and Resnet: a fresh data set for the model using the WaveGAN architecture, then apply the noise removal approach and the clipping method for data preparation. The MFCC was utilized for efficient feature selection. The MFCC provided an audio signal, which was then used to construct an audio spectrogram. Then, for categorization, use Resnet.

3. Results And Analysis

1) Training loss:

Specifically, whenever it comes to fitness data, the trained loss measures how well the algorithm fits the data. However, when it comes to validation data, the validation error measures how often the model fits the new data.

2) Training Accuracy:

While accuracy rate is often defined in terms of productivity achieved when a model is applied to an original dataset, checking quality is typically expressed in terms of effectiveness gained when a method is built to test the data set in question. It may be good to compare these in order to detect overtraining or undertraining.

3) Testing loss:

A harm is the result of making an inaccurate predicting decision. In other words, a loss is a numerical value that indicates how accurate the model's forecast was on some instance of time. Alternatively, if the model's forecast is flawless, the risk is zero; otherwise, the loss is higher.

4) Testing Accuracy:

Precision is a measure that is used to assess which machine learning model is the most accurate in identifying relationships and trends between measurement scales based on the evidence, or training, data that was provided.

```
COPD              64
Healthy           26
URTI              14
Bronchiectasis     7
Pneumonia          6
Bronchiolitis      6
LRTI               2
Asthma             1
Name: Diagnosis, dtype: int64
```

Fig. 2. Load Data

	Patient number	Recording index	Chest location	Acquisition mode	Recording equipment
0	210	1b1	Al	sc	Meditron
0	124	1b1	Ar	sc	Litt3200
0	170	1b4	Lr	mc	AKGC417L
0	180	1b4	Al	mc	AKGC417L
0	110	1p1	Lr	sc	Meditron

Fig. 3. Fig. 3. Load Patient data

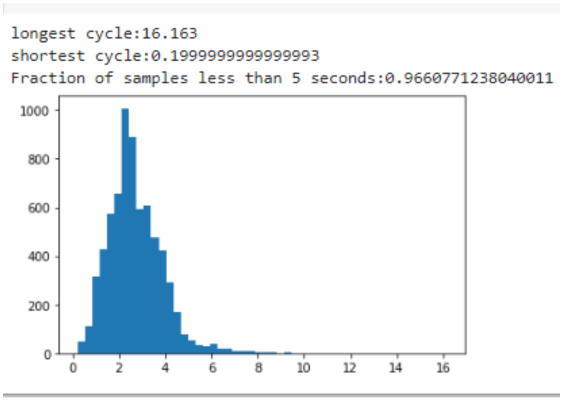

Fig.4. MFCC using CNN

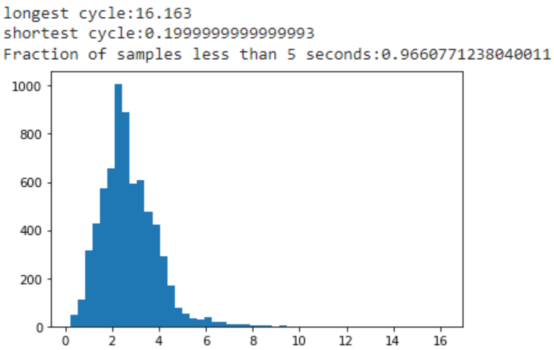

Fig.5 MFCC using Resnet

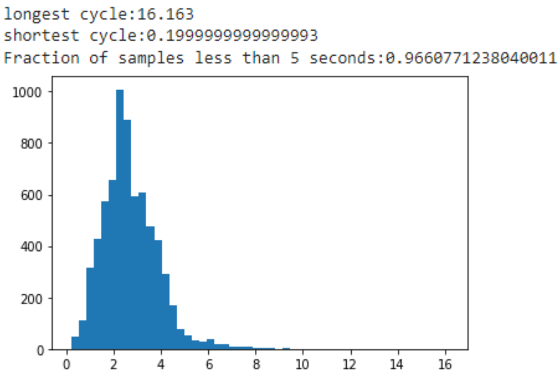

Fig. 6 MFCC using VGGnet

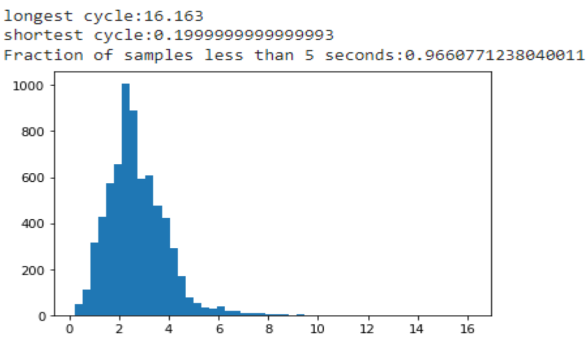

Fig.7 MFCC using WaveGAN and ResNet

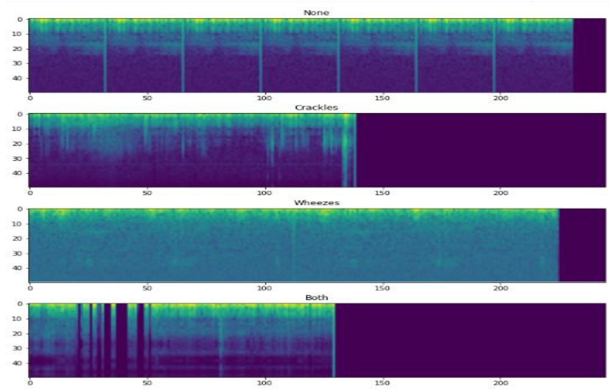

Fig. 8 Spectogram CNN

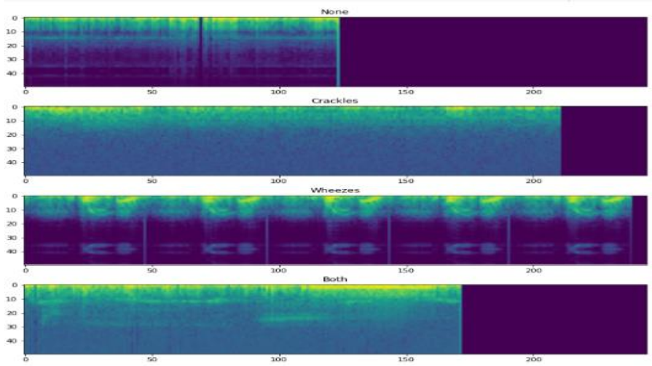

Fig. 9 Spectogram RESNET

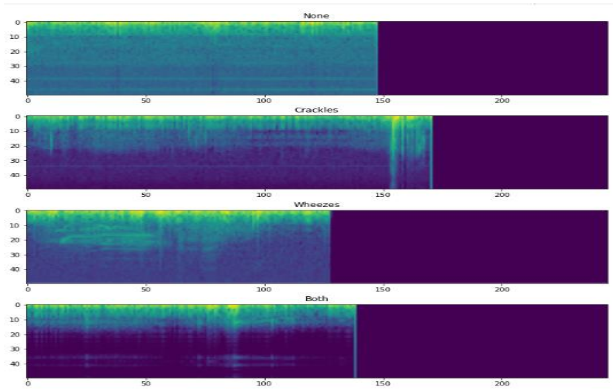

Fig. 10 Spectogram VGGNET

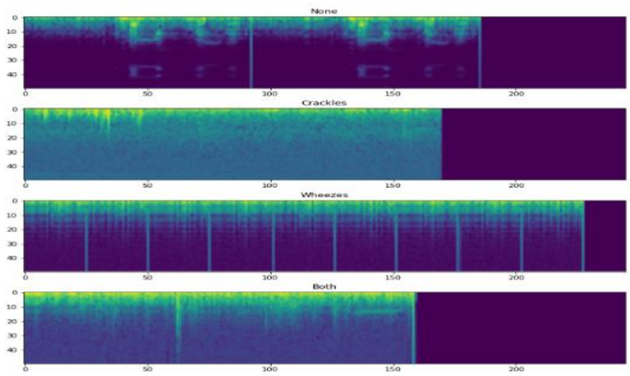

Fig. 11 Spectogram WaveGAN and RESNET

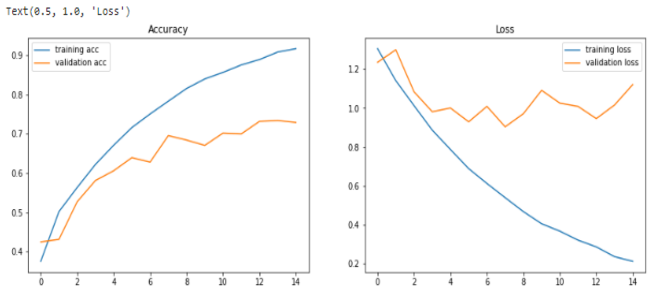

Fig.12 Acuuracy and Loss Graph CNN

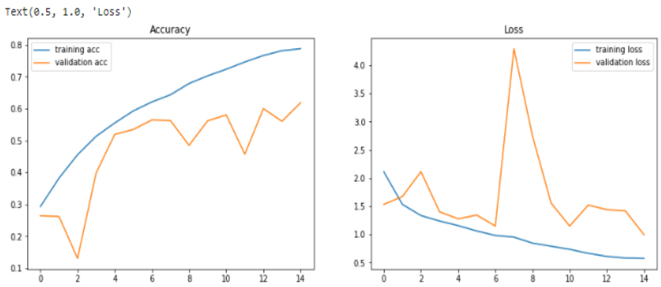

Fig. 13 Accuracy and Loss Graph RESNET

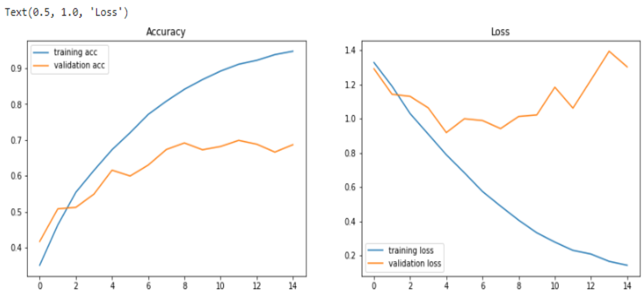

Fig.14 Accuracy and Loss Graph VGGNET

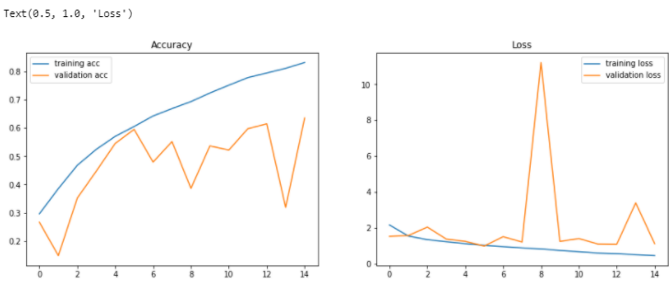

Fig.15 Accuracy and Loss WaveGAN and ResNet

TABLE I. ANALYSIS GRAPH

MODEL	PRECISION	RECALL	F1-SCORE	ACCURACY
VGGNET	0.61	0.63	0.61	0.69
RESNET	0.77	0.77	0.77	0.77
CNN	0.66	0.64	0.65	0.74
WAVEGAN + RESNET	0.83	0.82	0.83	0.82

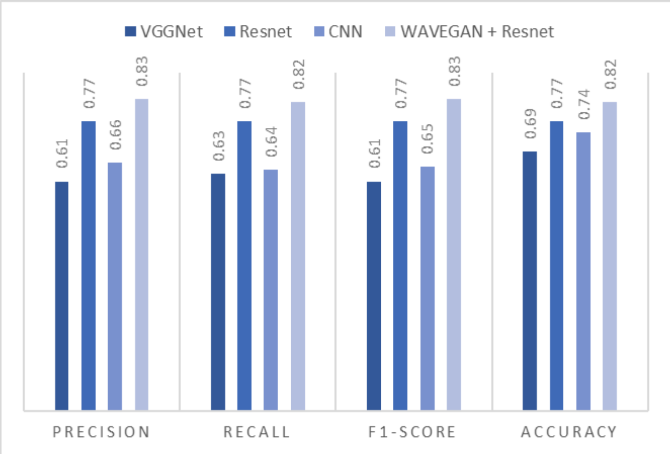

Fig.16 Analysis Graph

4. Conclusion

The major parts of the research are to determine organize lung sounds characteristic. However, it is often difficult to discern the appealing combination of the highlight, particularly if the details are too big. Because data acquired from nature is often erroneous, we can't utilize standard procedures to discover instances or estimate prices. Because meteorological data is often erroneous, we can't utilize traditional approaches to determine the composition or create numerical models. While other deep learning algorithms will result in the collection of highlights of lung sounds. The model CNN, VGG 16 devotes a significant amount of time and effort to prepare and deliver limited accuracy. RESNET50 with Wave-GAN, on the other hand, achieves 82% accuracy in each of the four types.

5. References

1. Valentyn Vaityshyn, Hanna Porieva, Anastasiia Makarenkova, “Pre-trained Convolutional Neural Networks for the Lung Sounds Classification”, 39th International Conference on Electronics and Nanotechnology (ELNANO), 2019
2. Fatih Demir, Aras Masood Ismael, Abdulkadir Sengur, “Classification of Lung Sounds with CNN model using Parallel Polling Structure”, IEEE Access Volume 8, 2020
3. Joel Than Chia Ming, Norliza Mohd Noor, Omar Mohd Rijal, Rosminah M. Kassim, Ashari Yunus, “Lung Disese Classification using Different Deep Learning Architectures and Principal Component Analysis”, 2nd International Conference on BioSignal Analysis, Processing and Systems, 2018
4. Zeenat Tariq, Sayed Khushal Shah, Yugyung Lee, “Lung Disease Classification using Deep Convolutional Neural Network”, International conference on Bioinformatics and Biomedicine, 2019
5. Anuradha D. Gunasingle, Achala C. Aponso, Harsha Thirimanna, “Early prediction of Lung Disease”, 5th International Conference for Convergence in Technology, 2019
6. Md. Ariful Islam, Irin Bandyopadhyaya, Parthasarathi Bhattacharyya and Goutam Saha, “Classification of Normal, Asthma and COPD Subjects using Multichannel Lung Sound Signals”, International Conference on Communication and Signal Processing, 2018
7. Fatih Demir, Abdulkadir Sengu, Varun Bajaj, “Convolutional neural networks based efficient approach for classification of lung diseases” Health Information Science and Systems, 2019
8. Syed Zohaib Hassan Naqvi, Dr. Mohammad Ahmed Choudary, Aleena Zaman Khan, Maheen Shakeel, “Intelligent System for Classification of Pulmonary Diseases from Lung Sound”,13th

International conference on mathematics, Actuarial science, Computer science and Statistics, 2019Gorkem Serbes, Sezer Ulukaya, Yasemin P. Kahya, “An Automated Lung sound Preprocessing and Classification System Based OnSpectral Analysis”, International Conference on Biomedical and Health Informatics, 2017

9. Ramizraja Shethwala, Shital Pathar, Tejal Patel, Prasann Barot, “Transfer Learning Aided Classification of Lung Sounds-Wheezes and Crackles” Fifth International Conference on Computing Methodologies and Communication (ICCMC 2021),2021
10. D. Jayaraj, S. Sathiamoorthy, “Random Forest based Classification Model for Lung Cancer Prediction on Computer Tomography Images”, Second International Conference on Smart and Inventive Technology (ICSSIT 2019), 2019
11. Adnan Hassal Falah, Jondri, “Lung sounds classification using stacked Autoencoder and Support Vector Machine”, 7th International Conference on Information and Communication Technology (ICoICT), 2019
12. Shreyasi Dutta, Anirban Dutta Choudary, Parijat Deshpande, Sakyajit Bhattacharta, Arpan Pal, “Automated Lung Sound Analysis for Detecting Pulmonary Abnormalities”, 39th Annual International Conference of the IEEE Engineering in Medicine & Biology society, 2017
13. H. Kamble, “Frequency Response Analysis of Respiratory Sounds and Comparative Study for Windowing Techniques” pp. 210–215, 2016.