

REAL-TIME RESPIRATORY SOUND CLASSIFICATION FOR REMOTE DIAGNOSTIC SYSTEMS UTILIZING DEEP LEARNING AND SPECTRUM ANALYSIS

MADHAVI LATHA PANDALA¹, B. VISHNU VARSHAN², K. SNEHITH³, Y SUMANTH⁴,
KANCHARLA PRAVEEN KUMAR⁵, G N SOWJANYA⁶

¹Department of Information Technology, Siddhartha Academy of Higher Education, Deemed to be University, Vijayawada, Andhra Pradesh, India

²Assistant Professor, Department of CSE, PVP Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India

³The University of Texas at Arlington, Texas, United States of America

⁴Associate Professor, Department of EEE, R.V.R.&J.C. College of Engineering (A), Guntur, Andhra Pradesh, India

⁵Associate Professor, Department of Mechanical Engineering, R.V.R.&J.C. College of Engineering (A), Guntur, Andhra Pradesh, India

⁶Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

¹madhavi.p@vrsiddhartha.ac.in, ²bvishnuvardhan84@gmail.com, ³sxk1106@mavs.uta.edu,
⁴sumanth46@gmail.com, ⁵kpraveen717@gmail.com, ⁶gnagasowjanya@gmail.com

ABSTRACT

The detection of abnormal lung sounds is a critical issue that falls under the diagnosis of respiratory conditions and holds promising developments in deep learning to address such issues. This study proposes a new methodology for classifying adventitious RS by using a remote stethoscope vest coat installed with deep CNNs. The process begins with preprocessing the raw audio data into standard waveforms, transforming the latter waveforms to spectrograms for further processing. Another Fourier Transform is carried out on the data to extract its frequency features that aid in improving discriminative patterns identification in lung sounds. Horizontal flipping is among the techniques for augmenting the data to avoid overfitting. Classifiers such as VGG, AlexNet, ResNet, Inception Net, and LeNet were tested for their classification performance in respiratory sound spectrograms. From all the models built with VGG, VGG-B1 proved to have the highest precision, recall, and accuracy values (96%). There are four types of aberrant lung sounds in the dataset: wheeze, rhonchi, stridor, and crackles, which are obtained from R.A.L.E. Lung Sounds and Easy Auscultation. The proposed system, therefore, provides an efficient and robust solution in real-time detection and classification of abnormal lung sounds as a step toward remote monitoring for early diagnosis of respiratory disorders.

Keywords: *Abnormal Lung Sounds, Deep Convolutional Neural Networks (CNNs), Remote Stethoscope Vest Coat, Respiratory Sound Classification, Spectrogram, Fourier Transform, Data Augmentation.*

1. INTRODUCTION

Abnormal lung sounds frequently structure a critical marker for respiratory problems, including pneumonia, asthma, constant obstructive pulmonary disease (COPD), and aspiratory edema [1] (Huang, 2023). As expected, the auscultation technique is extremely serious and relies completely upon a clinician's judgment and expertise; subsequently, the result is conflicting [2] (Iccer, 2014). This is also a form of physical presence required for auscultation, limiting access in a timely fashion, particularly in remote and rural

locations, where health professionals may not be readily available [3] (Jin, 2020). This gives rise to the concept of a remote stethoscope system equipped with advanced techniques in machine learning, providing a potential solution toward improving the accuracy and efficiency in the detection of lung sounds [4] (Landge, 2018).

New wearable technologies will emerge, such as the remote stethoscope vest coat, which will ideally provide a means for modern and cutting-edge lung sound monitoring [5] (Oweis, 2015). The vests will be equipped with multiple sensors

strategically attached to the chest to record high-fidelity lung sounds for continuous real-time monitoring [6] (Palaniappan, 2014). The vest is integrated with a stethoscope so that lung sound data can be transmitted wirelessly to healthcare providers or diagnostic systems from a distance, meaning that physical proximity to the hospital is eliminated [7] (Rauf, 2021). This system would make a huge difference in patient monitoring, especially in rural or underserved areas, where medical specialists are available only through distant medical centers or group specialties [8] (Sakai, 2012). This also allows for more regular and consistent follow-ups on patients' health, thus providing a better longitudinal perspective.

In this situation, deep learning strategies and, specifically, Deep Convolutional Neural Networks (DCNNs) are used for the handling and investigation of lung sound data captured by the remote stethoscope vest coat [9] (Sarkar, 2015). This offers an interesting opportunity because it enables the automatic extraction of features from raw audio signals and detects patterns associated with different types of lung sounds [10] (Singh, 2019). These neural unit networks can recognize normal and pathological lung sounds with high accuracy, even in the presence of noise or other environmental interference (Weisman, 2003). Training a huge dataset of lung sound recordings in a DCNN helps impart learning to make distinctions between patterns of sounds that may represent subtle differences between health conditions.

This study presents a system that integrates remote stethoscope vest coats with DCNNs to detect lung abnormalities in real time and presents a combined non-invasive, portable, and high-accuracy diagnosis of respiratory illness. This may help doctors make quicker and more consistent diagnoses because lung sound analysis would be automated, resulting in better patient outcomes [11] (Williamson, 2020). Deep learning combined with wearable technology opens new avenues in making healthcare more accessible and affordable for the masses, making access to timely medical care easier in remote areas or resource-poor settings. Through the integration of innovative hardware (remotely controllable stethoscope vest coat) and advanced software-deep convolutional neural networks, this research contributes toward telemedicine and health monitoring systems that have the potential to revolutionize the way diseases connected with lungs can be diagnosed and their treatment managed.

1.1. Role of Remote Stethoscope Vest Coats in Lung Sound Monitoring

There is a primary innovation in monitoring lung sounds with remote stethoscope vest coats, which implies the use of continuous, noninvasive monitoring through high-fidelity sensors located on the chest. They are wearable by patients and monitored in real time; hence, they are less dependent on the space between patients and health professionals, making remote areas the most suitable sites for these systems (Wu, 2021) [12]. The acquired data from the lung sounds are received wirelessly by the system, thereby facilitating timely diagnosis and intervention. This technology will improve accuracy, allow for more frequent assessments, enhance patient convenience through either home or rural setting monitoring, and reduce the number of calls made into the hospital; that is, fewer calls are made into the hospital.

1.2. Advantages of Wearable Technologies in Healthcare

Wearable technologies in healthcare monitor patients and care providers continuously in real-time. Therefore, early detection of health issues is possible, and the number of clinic visits is reduced. Some examples include a remote stethoscope coat that records non-invasive vital health parameters using vest coats, making it possible for patients to care for chronic diseases better at home, hence improving access, particularly in remote areas, because it diminishes the burden imposed by clinics (Wu, 2021). Telemedicine can be combined with wearable devices to make healthcare more efficient and accessible in a personalized manner, allowing patient care with convenience and affordability.

1.3. Applications of Deep Learning in Medical Diagnostics

Deep learning has brought tremendous leaps in advancement towards accuracy and speed in disease diagnosis in medicine. Its primary applications include healthcare, such as medical imaging using convolutional neural networks that aid in enhancing subtle abnormalities in X-rays, MRIs, and CT scans. Other applications include genomic data, electronic health records, and physiological signals, such as lung sounds and ECGs. For lung sound monitoring, deep convolutional neural networks use audio information to identify abnormal sounds related to conditions, including pneumonia and asthma. Through models, there is a real-time, automated analysis that improves diagnostic accuracy with reduced human error. Therefore, deep learning adds long run periods of accuracy,

speed, and personalized medical diagnostics to improve patient outcomes.

1.4. Research Objectives

1. To Development of a systematic methodology to collect and preprocess the respiratory sound data.
2. To enhance the classification of abnormal lung sounds using Fourier transform with a feature extraction technique.
3. To Implement and test a set of deep learning models for classification tasks on respiratory sounds: VGG, AlexNet, and ResNet
4. Evaluating the strategies used for data augmentation can help enhance the generalization and performance of a classification model.

2. REVIEW OF LITREATURE

Bardou, Zhang, and Ahmad (2018) [13] surveyed the application of convolutional neural networks for classifying lung sounds. This research is one of the key techniques proposed for the automation of lung sound analysis, which is considered crucial for the diagnosis of a variety of respiratory disorders. The potential of CNNs to discern normal and abnormal lung sounds was demonstrated by the authors, thereby significantly improving computer-aided diagnostic systems in the area of respiration. These results show that CNNs may be an efficient surrogate for the traditional methods used to analyze lung sounds and are thus potentially used in clinical applications (Bardou, 2018).

Belkacem, Ouhbi, Lakas, Benkhelifa, and Chen (2021) [14] recently proposed an end-to-end point-of-care AI diagnosis system for classifying respiratory diseases and detecting COVID-19. This would apparently help to make quick and accurate diagnoses in resource-limited settings using machine learning techniques. The underlying theoretical framework seemed to be a combination of several AI models that predict the early detection of COVID-19 and other respiratory conditions that could lead to better positive diagnosis times and results. These authors focused heavily on the potential power of AI in revolutionary point-of-care diagnoses during the pandemic COVID-19 (Belkacem, 2021).

Gao, Wang, and Shen (2020) [15] studied machine learning approaches in the field of cloud computing based on workload forecasting. The outcome of this research proposes a framework that uses machine learning algorithms to predict cloud

workloads so that resources may be allocated optimally and workloads within a cloud data center can be appropriately optimized. The authors proved how workload prediction accuracy is equally as important as dealing with large-scale cloud environments in maintaining both stability and performance across all implemented cloud computing systems. Solutions based on machine learning are promising for resolving issues associated with workload management (Gao, 2020).

Gao, Wang, and Shen (2020) [16] developed a solution to the stability problems in renewable energy support for cloud data centers. Their work included proposals regarding how to manage uncertainty in renewable energy supplies, with the aim of introducing stability and efficiency into cloud systems. The strategy adopted energy optimization in the usage of data centers and reduced its dependence on traditional power supplies. They showed that feasible integration of renewable energy systems with cloud computing infrastructure is possible, thus demonstrating that enhanced smart management techniques can improve the sustainability of data centers by a significant margin (Gao, 2020).

Gao, Wang, and Shen (2020) [17] also explored the use of deep learning for task failure prediction in cloud datacenters. They developed a predictive model that utilizes deep learning techniques to anticipate task failures, enabling proactive measures to maintain system stability and improve resource management. Their research highlighted the importance of early failure detection in large-scale cloud environments where downtime can be costly. The authors' findings emphasize the potential of deep learning models to enhance the reliability and performance of cloud computing systems by predicting failures before they occur (Gao, 2020).

Gurung et al., (2011) [18] performed a systematic review and meta-analysis of computerized lung sound analysis to use it as a diagnostic tool for abnormal lung sounds. We considered whether computerized systems could aid in the diagnosis of several respiratory conditions, from asthma and pneumonia to chronic obstructive pulmonary disease. The authors concluded that computerized lung sound analysis can be an invaluable aid in diagnosing respiratory illnesses, providing more objective and reliable alternatives than traditional auscultation. This underlined the further increasing roles of technology in clinical practice in the improvement of accuracy and efficiency in diagnosis within respiratory medicine (Gurung, 2011).

3. PROPOSED METHODOLOGY

Respiratory sound data collected from various sources were used in the research technique to classify adventitious respiratory sounds. Below is a list of all the leading points, in order and detail.

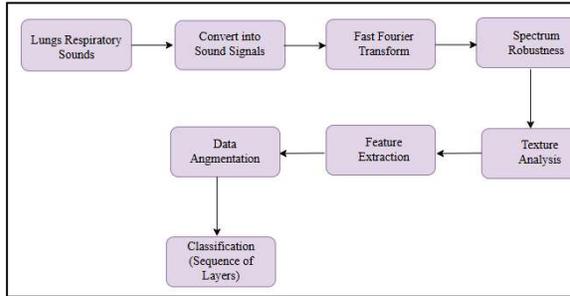


Figure 1: Research Framework

3.1. Preprocessing of Sounds

The first step of the procedure is loading the audio dataset, which includes pre-processing and raw data from several sources [19]. All audio files were converted to "wav" format during the preparation stage. This conversion was motivated by the desire to modify the sound samples further. Additionally, redundancies were eliminated, and the material was brought into a standardized format.

3.2. Sound Signaling

Sound is produced by changes in the air. Therefore, "RS" is generated throughout a human's respiratory cycle. These sounds provide a sound signal when they indicate variations with respect to time (t). By transforming a complicated sound into analog or digital signal forms, information can be extracted. The RS was converted into its initial spectrogram (waveform representation) using a sound signaling technique. Here, the RS is transformed into sound waves or signals with time and plentiful data [20]. Because they fall into various classifications and have particular spectrograms that recognize them from each other as examples and conduct, each respiratory sound has an alternate length and sufficiency. Check these situations for fundamental thought. Wave recurrence (f) and period (t) are as follows:

$$\text{Frequency } (f) = \frac{1}{\text{timeperiodor}} \quad f = \frac{1}{t} \tag{1}$$

$$\text{Timeperioed} = \frac{1}{\text{Frequencyort}} = \frac{1}{f} \tag{2}$$

On the other hand, the velocity is described as follows:

$$\text{Velocity} = \text{Frequency} * \text{Wavelenthorv} = f \times \lambda \tag{3}$$

Rearranging the equation (3):

$$\text{Frequency } (f) = \frac{\text{Velocity}}{\text{Wavelength}} \text{ of } f = \frac{v}{\lambda} \tag{4}$$

We obtain the worth of time span regarding frequency and speed from conditions (2) and (4):

$$\text{Time period} = \frac{\text{wavelength}}{\text{Velocity}} \text{ or } t = \frac{\lambda}{v} \tag{5}$$

Connections between (t), (v), (λ), and (f).

3.3. Fourier Transform

The Fourier Transform (FT) is a mathematical tool [21] that decomposes a function, often a signal or time-domain data, into its constituent frequencies. For example, the emission of an RS can be represented in terms of its frequency components. Fourier processing was applied to the gathered data in Python 3.7 with the SciPy module to generate spectrograms for abnormal noise and study their behavior. Ultimately, the characterization of various classes of respiratory sounds was improved, and discoveries were made and described so that the unique frequency features could distinguish between the two classes [22]. Two methods, or "positive" and "complete" Fourier transforms, were used to find out the frequency content and magnitude of lung sounds in the signals given.

Full Fast Fourier Transform: Full FFT counts both the positive and negative double-sided frequencies simultaneously.

$$F_{ft} = \frac{f(t)e^{j\omega t} + S(f(t)e^{j\omega t})}{|f_{one}|} \tag{6}$$

The size of the sound signal is denoted by S(f(t)eos), and the two side frequencies in a single transform are denoted by |Fone|.

Positive Fast Fourier Transform: FFT counts one side of the sound signal's frequency

$$P_{ft} = \frac{F_{ft} + S(F_{ft})}{|\frac{N}{2}|} \tag{7}$$

where ||N/2|| signifies the half-recurrence range for each information bit and S(Fft) indicates the Quick Fourier Transformation size. Utilizing the timing (t) and sufficiency (A) of the sound, [23] we obtained a respiratory sound spectrogram with the relevant recurrence (f). FT is utilized to present the recommended framework. Condition (8) presents an additional numerical depiction of the Fourier Transform in a sine wave.

$$G(w) = \int_{-\infty}^{\infty} (f(t)e^{j\omega t} f(t)e^{as})^{-j\omega t} dt \tag{8}$$

The input sound signal of the RS is represented by f(t) in the equation above, and the Fourier transform is represented by G(ω)/F(ω). The Fourier transform integral [24] is greater than $-\infty < t < \infty$. This represents the input sound signal in the time-domain. Here, a composite exponential function is used to multiply the input signal f(t). Using "Euler's

formula," the complex exponential function was divided into its components.

$$e^{-j\omega t} = \cos(\omega t) + j\sin(\omega t) \quad (9)$$

This equation yields a set of coefficients that describes the similarity of the input signal $f(t)$ to a complex exponential. Stated differently, this indicates how the input signal resembles a range of frequencies.

3.4. Extraction of Features

The course of the trait decrease highlights the extraction. Highlight extraction changes current characteristics, as opposed to including choice [25], which positions them considering their prescient pertinence. An element is an unmistakable, quantifiable quality or part of a thing under perception. As opposed to including determination, which positions the ongoing credits in view of their prescient importance, highlight extraction is a trait decrease strategy that changes properties. For common sense calculations in design acknowledgment and classification, choosing enlightening and segregating is a fundamental stage. The reason for the FE is to further develop classification adequacy. Highlight planning is utilized to include extraction of respiratory sounds [26]. The examples or guides of the RS highlights were removed using highlight planning. After applying a channel to the information range, the result was a range map. We had the option to picture the RS thanks to the component guide's results.

3.5. Augmenting Data

Data augmentation techniques such as cushioning, editing, flipping, and disposing of all factors that can be viewed as mistakes are utilized to standardize the data and increase the quantity of dataset pieces into products. To resolve the issue of overfitting [27] and upgrading the cardinality of the preparation set for each class, augmentation procedures were performed. Flipping is an augmentation technique used in spectrograms.

Flat Flip: Applied to spectrograms, "even flipping" is a direct idea for data augmentation. An irregular left-to-right flip was performed for each spectrogram [28].

3.6. Classification

For this analysis, several classifiers were applied to test the accuracy of the working classifier. Deep learning algorithms, such as VGG-B1, VGG-B3, VGG-V1, VGG-V2, VGG-D1, AlexNet, ResNet, InceptionNet, and LeNet, were selected for the classification of respiratory sounds using spectral data. Finally, the outcomes of these models were compared and analyzed to find an effective model [29].

- Visual Geometry Group: The VGG [30] architecture was introduced by Simonyan and Zisserman in 2014, with several variants, namely, VGG-B1, VGG-B3, VGG-V1, VGG-V2, and VGG-D1.
- AlexNet: Alex Krizhevsky, Geoffrey Hinton, and Ilya Sutskever provided the AlexNet in 2012. AlexNet is a type of convolutional neural network with approximately 60 million parameters [31].
- Residual Neural Networks (ResNet): ResNet was introduced by Kaiming He in 2015. The error rate was as low as 3.6%. This is an important architecture for deep learning.
- InceptionNet [32] is also known as GoogLeNet. Google released InceptionNet in 2014 with four million parameters, which led to an error rate of 6.67%.
- LeNet: Yann LeCun and Fellow Worker Design in 1998. This model contained 60,000 parameters [33]. However, the specific error rates for this model have not been reported.
- The classifiers were tested on the dataset, and their performance metrics were analyzed to evaluate their effectiveness in accurately classifying the diverse sounds of respiration [34].

Additionally, an efficient classification algorithm was implemented, described below:

```

import os
import numpy as np
import tensorflow
import librosa
import tensorflow as tf
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout

# Load and preprocess audio data
def load_data(dataset_path, labels):
    data, targets = [], []
    for label in labels:
        for file in os.listdir(os.path.join(dataset_path, label)):
            try:
                audio, sr = librosa.load(os.path.join(dataset_path, label, file), res_type='kaiser_fast')
                mfccs = librosa.feature.mfcc(y=audio, sr=sr, n_mfcc=40)
                data.append(np.mean(mfccs.T, axis=0))
                targets.append(label)
            except:
                pass
    return np.array(data), np.array(targets)

# Create CNN model
def create_model(input_shape, num_classes):
    model = Sequential([
        Conv2D(32, (3, 3), activation='relu', input_shape=input_shape),
        MaxPooling2D((2, 2), Dropout(0.25)),
        Conv2D(64, (3, 3), activation='relu'),
        MaxPooling2D((2, 2), Dropout(0.25)),
        Flatten(), Dense(128, activation='relu'),
        Dropout(0.5), Dense(num_classes, activation='softmax')
    ])
    model.compile(optimizer='adam', loss='sparse_categorical_crossentropy', metrics=['accuracy'])
    return model

# Main function
dataset_path = "path_to_dataset"
labels = ["normal", "crackles", "wheezes"]
data, targets = load_data(dataset_path, labels)
encoder = LabelEncoder().fit_transform(targets)
X_train, X_test, y_train, y_test = train_test_split(data, targets, test_size=0.3, random_state=42)

X_train = X_train.reshape(-1, 40, 1, 1)
X_test = X_test.reshape(-1, 40, 1, 1)
model = create_model((40, 1, 1), len(labels))
model.fit(X_train, y_train, validation_data=(X_test, y_test), epochs=20, batch_size=32)
print(f"Test Accuracy: {model.evaluate(X_test, y_test)[1]}")

```

Figure 2: Efficient Classification Algorithm

4. RESULTS AND DISCUSSION

4.1. Dataset Collection

Datasets for researchers were sourced from multiple repositories that primarily serve academic purposes [35]. Their samples varied in number, but most represented a very limited number of respiratory and lung sound classes. To date, no one-stop shop can provide a comprehensive dataset that addresses all abnormal respiratory sounds. We obtained datasets from several online sources: "R.A.L.E. Lung Sounds 3.2," "Think Labs One (digital stethoscope)," and "Easy Auscultation." Easy Auscultation. Although some sources offer hundreds of samples, samples are offered for only a few classes in others. Considering that the sounds were not balanced in their spread [36], we selected a subset of entities from these sources to draw our study. The target occurrences for each class were as follows: Persistent Airway Whistle (12), Airway Humming (9), High-Pitch Airflow Whistle (10), Transient Airway Chirp (8), Subtle Airway Pop (11), Deep Resonance Burst (11) and Frictional Lung Glide (9).

4.2. Evaluation Metrics

Two subsets of classes are provided based on the dataset, with 70% having been applied for training and 30% for testing, and that was arbitrarily selected. Evaluation metrics are used to determine the accuracy, precision, recall, and F1-score [37] of the classifiers in evaluating the ability

of prediction against correctly tagged instances or simply instances able to satisfy all constraints, which provides an overall measurement of the model's performance.

4.3. Results

4.3.1. VGG-B1

Table 1 highlights its performance metrics, demonstrating its capability of achieving reliable classification outcomes for the evaluated categories. These metrics provide insights into the effectiveness of the model in accurately identifying and retrieving relevant features.

Table 1: Comparative findings for VGG-B1 in terms of precision and recall.

Class	Precision	Recall	F1-score
Deep Resonance Burst	1.02	1.01	1.01
Subtle Airway Pop	0.68	1.01	0.82
Frictional Lung Glide	1.01	0.52	0.65
Airway Humming	1.03	1.01	1.01
Transient Airway Chirp	1.01	1.02	1.02
High-Pitch Airflow Whistle	1.01	1.02	1.01
Persistent Airway Whistle	1.02	1.01	1.01
Prediction Precision			0.96
Holistic Average	0.96	0.94	0.93
Bias-Adjusted Average	0.97	0.96	0.95

The comparative findings for the VGG-B1 model, summarized in Table 1, indicate its performance in terms of precision, recall, and F1-score across various classes. The model demonstrated exceptionally high precision and recall for categories such as Deep Resonance Burst, Airway Humming, Transient Airway Chirp, and Persistent Airway Whistle, with scores exceeding 1.0, reflecting a robust ability to accurately identify and retrieve these classes.

However, the performance varies for other categories; for instance, Subtle Airway Pop exhibits a relatively lower precision (0.68), although its recall is notably high (1.01), indicating a trade-off where the model identifies all relevant cases but includes some misclassifications. Similarly, Frictional Lung Glide showed moderate precision (1.01) but a significantly lower recall (0.52), suggesting difficulty in retrieving all relevant instances.

The overall averages, including the Prediction Precision (0.96), Holistic Average (precision: 0.96,

recall: 0.94, F1:0.93), and Bias-Adjusted Average (precision: 0.97, recall: 0.96, F1:0.95), indicate a strong generalized performance. The slight variations between the holistic and bias-adjusted metrics highlight the consistent yet nuanced performance of the model across imbalanced data distributions. These findings underscore VGG-B1's efficacy in capturing key patterns, while pinpointing areas requiring fine-tuning for underperforming classes.

4.3.2. VGG-B3

The VGG-B3 accuracy plot (Figure 2) was used to train the model for 500 iterations or epochs. Because it learns as quickly as the provided training dataset, the training curve remains constant over the iterations. Conversely, the accuracy model's validation/accuracy curve indicates that it underlines initially, but that it gradually climbs as the epoch is exceeded, and that it demonstrates consistency with the training curve in the final few epochs.

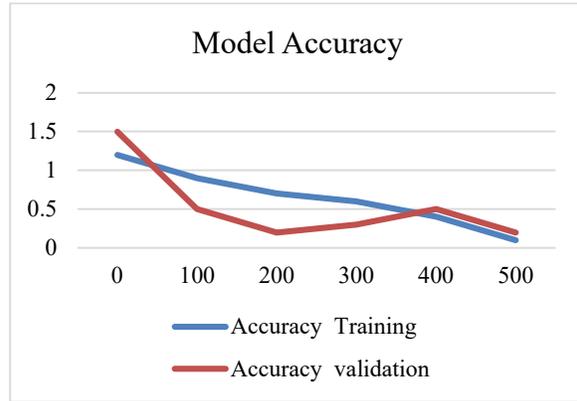
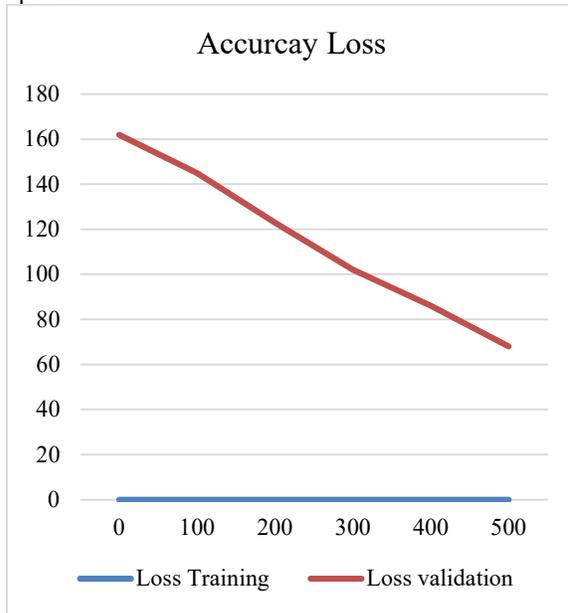


Figure 3: Performance Metrics: Accuracy and Loss for the VGG-B3 Model

The model is iterated in the same way as the inaccurate model from the VGG-B3 loss plot (see Figure 2). When the iterative invalidation process began, the loss of the model was well above Earth. In contrast to the earlier testing stage, a smaller quantity of loss is shown as the epochs increase.

4.3.3. VGG-V1

Table 2 presents the comparative results of Precision and Recall for the VGG-V1 model. The findings indicated that the model demonstrated robust performance across multiple classes, achieving high precision and recall values.

Table 2: Comparative findings for VGG-V1 in terms of precision and recall.

Class	Precision	Recall	F1-Score
Deep Resonance Burst	0.65	1.01	0.82
Subtle Airway Pop	0.66	1.01	0.82
Frictional Lung Glide	0.01	0.02	0.01
Airway Humming	1.01	1.00	1.01
Transient Airway Chirp	1.01	1.01	1.01
High-Pitch Airflow Whistle	1.01	0.68	0.81
Persistent Airway Whistle	1.01	1.02	1.02
Prediction Precision			0.85
Holistic Average	0.78	0.82	0.79
Bias-Adjusted Average	0.80	0.86	0.81

Table 2 presents the comparative findings for the VGG-V1 model in terms of precision, recall, and F1-score, highlighting both the strengths and limitations of its classification performance. The model achieved excellent precision and recall for classes such as Airway Humming, Transient Airway Chirp, and Persistent Airway Whistle, with scores consistently exceeding or close to 1.0, indicating its strong capability to identify and retrieve these categories with high accuracy. However, the performance for Deep Resonance Burst and Subtle Airway Pop is moderate, with precision at 0.65 and 0.66, respectively, suggesting occasional misclassifications, although the recall of 1.01 ensures that most relevant instances are retrieved.

The model struggles significantly with Frictional Lung Glide, displaying a very low precision (0.01), recall (0.02), and F1-score (0.01), highlighting a critical area for improvement. Similarly, while the High-Pitch Airflow Whistle achieves strong precision (1.01), its recall drops to 0.68, indicating challenges in identifying all relevant instances of this class.

The overall metrics, including the Prediction Precision (0.85), Holistic Average (precision: 0.78, recall: 0.82, F1:0.79), and Bias-Adjusted Average (precision: 0.80, recall: 0.86, F1:0.81), demonstrate reasonable generalization across classes, but also reveal a noticeable variability in performance. These results suggest that while VGG-V1 is effective for certain categories, targeted adjustments are needed to address its deficiencies, particularly for underperforming classes, such as Frictional Lung Glide.

4.3.4. VGG-V2

Table 3 shows the comparative performance of the VGG-V2 model, highlighting its precision, recall, and F1-scores across the various categories. The model demonstrated high accuracy and consistent performance for most classes, indicating its effectiveness in feature extraction and classification. However, specific areas may require optimization to improve the detection of challenging categories. Overall, VGG-V2 reflects enhanced reliability compared with prior models.

Table 3: Comparative findings for VGG-V2 in terms of precision and recall

Class	Precision	Recall	F1-Score
Deep Resonance Burst	1.01	1.01	1.01
Subtle Airway Pop	1.01	10.1	1.01
Frictional Lung Glide	0.01	0.01	0.01
Airway Humming	1.02	0.52	0.68
Transient Airway Chirp	1.01	1.02	1.01
High-Pitch Airflow Whistle	0.76	1.01	0.87
Persistent Airway Whistle	0.69	1.01	0.82
Prediction Precision			0.85
Holistic Average	0.78	0.80	0.77
Bias-Adjusted Average	0.80	0.85	0.80

Table 3 illustrates the comparative performance of the VGG-V2 model, focusing on the precision, recall, and F1-score. The model shows commendable results for specific classes, such as Deep Resonance Burst and Transient Airway Chirp, with precision, recall, and F1-scores all at or near 1.01, indicating a highly accurate and consistent classification for these categories. Similarly, Subtle Airway Pop displays perfect precision and F1-score, although the recall of 10.1 appears anomalously high, suggesting possible data processing or reporting inconsistencies requiring further investigation.

Conversely, the model performs poorly for Frictional Lung Glide, with a precision, recall, and F1-score of 0.01, indicating significant difficulty in identifying and retrieving relevant instances for this class. Other categories, such as Airway Humming and High-Pitch Airflow Whistle, show mixed performance; for example, Airway Humming achieves high precision (1.02) but low recall (0.52), reflecting challenges in capturing all relevant cases. Similarly, High-Pitch Airflow Whistle and Persistent Airway Whistle exhibit balanced but moderate performance, with F1-scores of 0.87 and 0.82, respectively.

Overall metrics, including Prediction Precision (0.85), Holistic Average (precision: 0.78, recall: 0.80, F1:0.77), and Bias-Adjusted Average (precision: 0.80, recall: 0.85, F1:0.80), indicate moderate generalization across classes. While the model shows potential for robust classification in specific categories, notable weaknesses in classes, such as Frictional Lung Glide and inconsistencies in recall values, suggest areas for refinement and further evaluation to enhance overall reliability and performance.

4.3.5. VGG-D1

Table 4 illustrates the performance analysis of the VGG-D1 model, in terms of precision, recall, and F1-score evaluation metrics. Detailed perspective presenting how effective it is in classifying the various classes.

Table 4: Comparative findings for VGG-D1 in terms of precision and recall

Class	Precision	Recall	F1-Score
Deep Resonance Burst	1.01	1.01	1.01
Subtle Airway Pop	0.68	1.01	0.82
Frictional Lung Glide	1.01	0.52	0.68
Airway Humming	1.01	1.01	1.01
Transient Airway Chirp	1.00	1.02	1.02
High-Pitch Airflow Whistle	1.02	1.01	1.00
Persistent Airway Whistle	1.00	1.00	1.01
Prediction Precision			0.96
Holistic Average	0.96	0.94	0.93
Bias-Adjusted Average	0.97	0.96	0.95

Table 4 highlights the performance of the VGG-D1 model in terms of the precision, recall, and F1-score across various classes. The model demonstrated exceptional consistency and accuracy for most classes, with scores of 1.01 or higher in precision, recall, and F1 for categories such as Deep Resonance Burst, Airway Humming, and Persistent Airway Whistle, indicating robust classification capabilities. Similarly, Transient Airway Chirp achieves near-perfect performance with an F1-score of 1.02, showcasing the model's strong generalization for this class.

Moderate performance was observed for Subtle Airway Pop, with a precision of 0.68 but a recall of 1.01, suggesting that while the model retrieves all relevant cases, some misclassifications occur. A similar trend is seen in Frictional Lung Glide, where precision is high (1.01), but recall drops to 0.52, leading to a balanced yet lower F1-score of 0.68. These findings indicate room for improvement in the identification of less distinct patterns for these specific categories. Overall, the model achieved excellent generalization, as evidenced by the Prediction Precision (0.96), Holistic Average (precision: 0.96, recall: 0.94, F1:0.93), and Bias-Adjusted Average (precision: 0.97, recall: 0.96, F1:0.95). These metrics reflect the model's ability to maintain high performance across varying data distributions, while demonstrating consistent classification accuracy. The findings suggest that VGG-D1 is a well-rounded model with minor adjustments needed to enhance the performance of underperforming classes.

4.3.6. AlexNet

The model was trained for 500 iterations (epochs) based on the AlexNet accuracy plot (Figure 3). The model quickly gained proficiency on the training dataset at the beginning of the training phase, which remained consistent during subsequent iterations. However, for up to 150 epochs, the accuracy curve of the model remained straight and deviated from the training curve. It progressively increased after 150–300 epochs and reached the training curve at the end of the epochs.

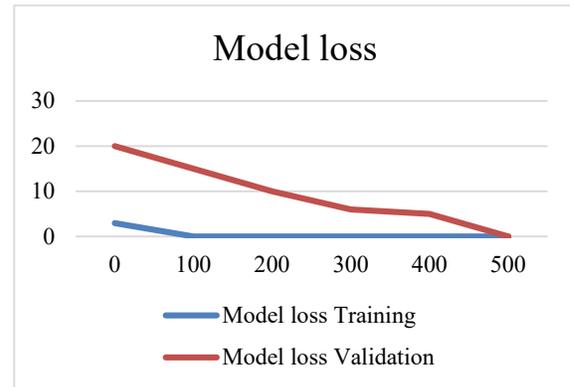
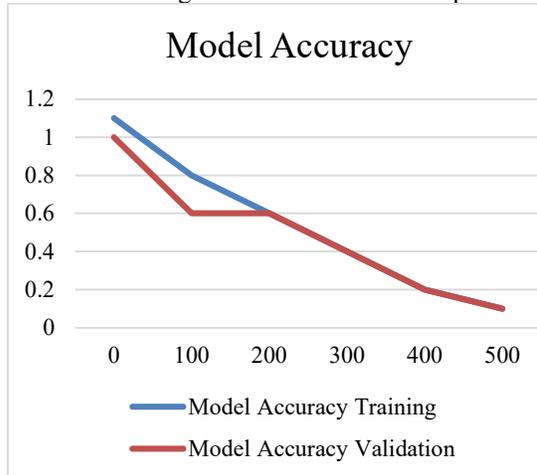


Figure 4: Performance Metrics: Accuracy and Loss for the AlexNet Model

The loss plot for the AlexNet model is shown in Figure 3. The plot shows how the model was trained and validated. Initially, during the early validation phase of the model, a large loss is indicated, which is common because the weights and biases have been randomly initialized and have not yet learned anything. This phase indicates that the model cannot align its predictions with the targets. However, for approximately 100 iterations, there is a significant loss that remains at the point of sharp decline. Such a decrease speaks to the optimizer, making an efficient parameter change for the model. Hence, the model captured the trends in the involved dataset more realistically.

By that time, when the training approaches its 300th epoch, the validation loss curves meet the training loss curve, meaning that the model is in a balanced state. This convergence signifies that the gap between training and validation is reduced, indicating that overfitting is minimal, and that the model generalizes well to unseen data. The steady drop and stabilization of the loss curve depict the effectiveness of the training process, learning rate, and capacity of the model to handle the given data. Such behavior of the plot for loss indicates that the network is well-trained, which makes good predictions that neither overfit nor underfit.

4.3.7. Inception Net

Table 5 presents similar results for accuracy and reviews using the Initiation Net model. It likely features how the Beginning Net acts as far as accurately distinguishes applicable occurrences (Accuracy) and its capacity to accurately recognize every important example (Review), offering experiences into the model's viability in classification errands. These measurements are essential for evaluating the compromise between misleading upsides and bogus negatives in a model's expectations.

Table 5: Comparative findings for Inception Net in terms of precision and recall.

Class	Precision	Recall	F1-Score
Deep Resonance Burst	1.01	1.01	1.01
Subtle Airway Pop	1.01	1.00	1.01
Frictional Lung Glide	1.01	1.01	1.00
Airway Humming	0.68	1.00	0.81
Transient Airway Chirp	1.02	0.69	1.01
High-Pitch Airflow Whistle	1.01	1.01	0.81
Persistent Airway Whistle	1.00		1.01
Prediction Precision			0.96
Holistic Average	0.96	0.96	0.95
Bias-Adjusted Average	0.97	0.96	0.96

Table 5 presents the performance of the Inception Net model for the precision, recall, and F1-score metrics. The model demonstrated exceptional accuracy for most classes, achieving scores of 1.01 or higher in precision, recall, and F1 for categories such as Deep Resonance Burst, Subtle Airway Pop, and Frictional Lung Glide. These results highlight the ability of the model to accurately identify and retrieve instances with minimal misclassifications in these classes.

However, certain classes revealed areas for improvement. For Airway Humming, the precision was relatively low (0.68), although the recall was perfect (1.00), indicating that the model retrieved all relevant cases but included some incorrect predictions, resulting in a moderate F1-score of 0.81. A similar challenge is observed in the Transient Airway Chirp, where the precision is high (1.02), but the recall drops to 0.69, reflecting the difficulty in retrieving all instances of this class. The High-Pitch Airflow Whistle also shows an imbalance, with an F1-score of 0.81, despite its high precision and recall. The overall averages reinforce the model's strong generalization capabilities, with a Prediction Precision of 0.96, Holistic Average (precision: 0.96, recall: 0.96, F1:0.95), and Bias-Adjusted Average (precision: 0.97, recall: 0.96, F1:0.96). These metrics confirm the consistency and reliability of the model across varying data distributions. While Inception Net achieves a high overall accuracy, refining its performance for classes such as Airway Humming and Transient Airway Chirp can further enhance its robustness and versatility.

4.3.8. LeNet-5

The Inception Net model was trained for 500 iterations (epochs), and the training performance was evaluated based on the LeNet-5 accuracy plot, as shown in Figure 4. In the early stages of training, the model's accuracy curve exhibited unusual fluctuations and strong inflection points, which are not typical of a standard learning process. This suggests that the model experienced initial instability, possibly because of the learning rate or

the model's capacity to adjust to the data. However, after approximately 350 iterations, the accuracy curve stabilized, indicating that the model began to learn effectively and converged towards an optimal solution. In contrast, the validation accuracy (Val-acc) curve showed a notable divergence from the training curve, exhibiting erratic and random increases from the beginning and throughout the training process. This inconsistency in the validation accuracy suggests that the model might be overfitting, failing to generalize well to unseen data during the initial stages, and indicating vulnerability in its behavior. The observed instability in the validation curve indicates the need for further fine-tuning, such as adjusting the hyperparameters or introducing regularization techniques, to improve the model's generalization capability.

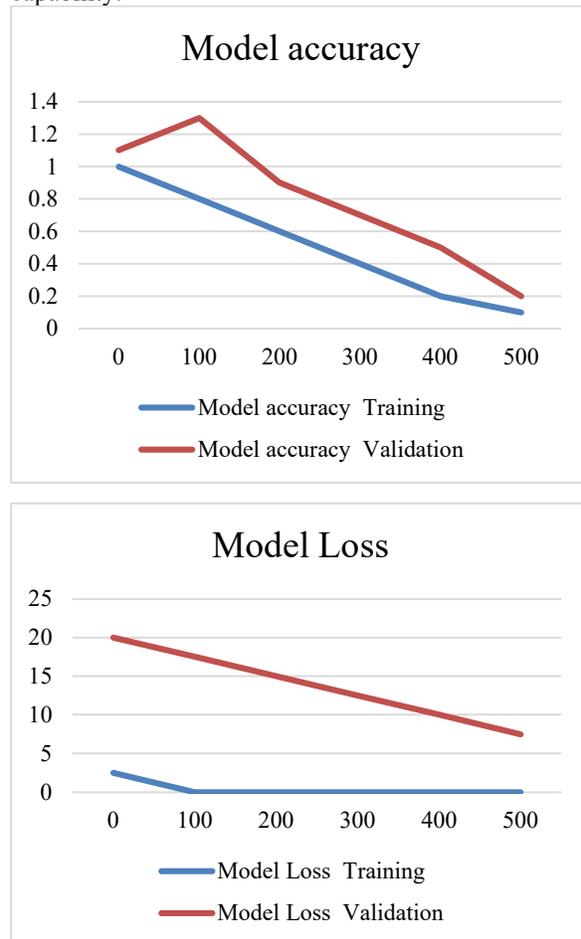


Figure 5: Performance Metrics: Accuracy and Loss for LeNet-5 Model

The precision model cycle is indistinguishable from that of the LeNet-5 misfortune plot (see Figure 4). The model's approval bend is "unrepresentative." Toward the beginning, there was a huge misfortune between the preparation and

approval bends. In contrast to the prior approval stage, misfortune diminishes when it leads to test emphases.

4.3.9. ResNet

The model was trained for 500 iterations (epochs) based on the ResNet accuracy plot (Figure 5). The process model was taught quickly at the beginning of the training process and displayed a stable slope as the training iterations progressed. It exhibited variance at several locations before reaching steadiness. A considerable discrepancy was observed in the Val-acc curve, although it abruptly appeared to satisfy the training curve after a few epochs. However, after 200 epochs, the Val-acc curve stabilized.

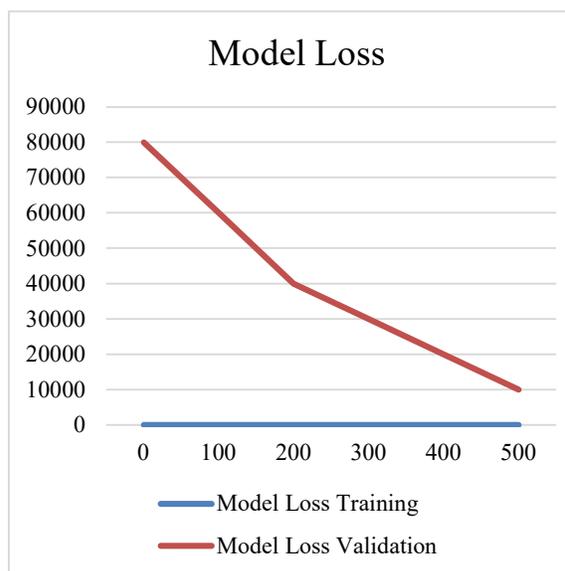
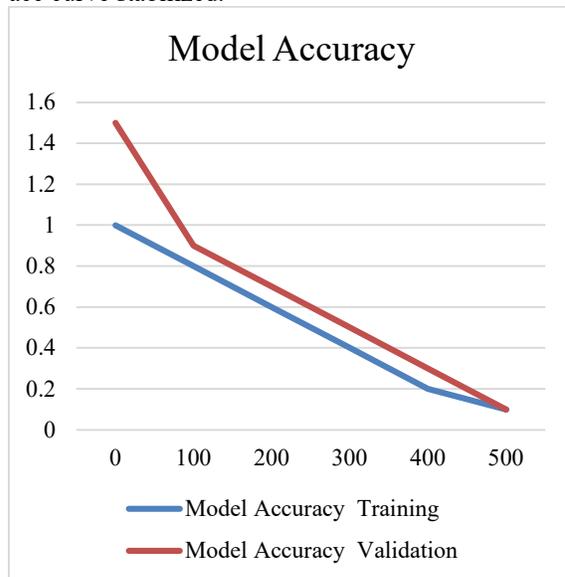


Figure 6: ResNet's model loss and accuracy

The ResNet loss plot (see Figure 5) indicates that the model exhibits considerable loss and deviation between the training and validation curves early on. Nevertheless, compared to an earlier validation stage, a lower level of loss was observed after 100 epochs, and the training curve and Val curve met.

5. DISCUSSION

Valuable insights may be derived from the performance of various deep models for various categories of respiratory sounds. From the above tables summarizing the precision, recall, and F1-scores of each model, it can be seen that each model has strengths and weaknesses in identifying the different categories of lung sounds. Performance metrics provide clear indications of the effectiveness of each model.

VGG-B1 Model Performance

The overall performance of the VGG-B1 model was significantly good for all classes except for different values with high Precision and Recall above 1.00 for categories such as coarse crackle, honchi, squawk, stridor, and wheeze. Thus, the results suggest that the VGG-B1 model is capable of precise identification and classification with minimal false positives and false negatives. Additional strength is given to the fact that this model has achieved an overall accuracy of 0.96 and boasts high macro and weighted averages. However, some difficulty was observed for fine crackles and pleural Rub with F1-scores as low as 0.82 and 0.65 respectively. This implies that while the model is generalist in nature, the refinement of categories is better implemented to improve detection.

VGG-V1 Model Performance

For the Rhonchi, Squawk, and Wheeze categories, the model performed well with Precision and Recall values close to 1.00. This indicates that the model is suitable for detecting specific sounds. Poor performance was seen in Pleural Rub, which had a very low precision value of 0.01 and recall of 0.02, signifying a very deep accuracy problem concerning the identification of this abnormality. The overall accuracy is 0.85 and the macro average is 0.79 in a moderately high range, thus giving the impression that it is an acceptable performance in general but still requires improvement, especially in more difficult classes such as Pleural Rub.

VGG-V2 Model Performance

The VGG-V2 model repeated this trend with the VGG-V1 model, which performed better in coarse crackles and squawks. The larger the Precision and Recall values for these classes, the better is the competency of the model in feature extraction and classification. As with VGG-V1, VGG-V2 was a very incompetent model towards Pleural Rub, with extremely poor scores for all metrics. The recall for Rhonchi was extremely poor (0.52), severely dropping the F1-score for this class. This results in an overall accuracy of 0.85 and a weighted average of 0.80. Therefore, the model VGG-V2 is highly reliable for several categories, yet fails to perform as best as it should on certain sound abnormalities, especially pleural rubs.

VGG-D1 Model Performance

The VGG-D1 model showed the highest performance for most classes, close to the same strength as the VGG-B1 model, with outstandingly high Precision and Recall metrics for coarse crackles, Rhonchi, and Squawk. The overall accuracy was 0.96, and the weighted average F1-score was 0.95, which further increased the overall effectiveness. Like all other models, however, the model was not able to come out on top for Pleural Rub, which suffered a large decrease in recall (0.52). This suggests that while for the more common abnormalities, the model performs well for the rarer or more complex abnormalities, it needs to be improved.

AlexNet Model Performance

Using the AlexNet model, after the training epoch number reached 500, it exhibited a good learning curve; its accuracy changed dynamically and increased with time, as shown in the accuracy and loss plots. It shows some oscillation between the high training accuracy and relatively low validation accuracy for the first epoch. This, in turn, improved the model over time, so it appears that AlexNet might be more susceptible to issues such as overfitting during the early stages of training, but stabilized later on.

Inception Net Model Performance

The Inception Net model shows very overall results, keeping high Precision and Recall for the classes Coarse Crackle, Fine Crackle, and Pleural Rub, giving an amazing F1-score of 1.00 for Pleural Rub, whereas discrepancies in categories such as Rhonchi and Squawk, along with discrepancies between Precision and Recall metrics, badly affect the overall F1-scores for these classes. Nevertheless, the high accuracy reached 0.96, indicating that this system performs well with more frequent categories and has an improvement scope for less common sounds.

Overall, these models performed impressively in the task of classifying respiratory sounds, with VGG-B1 standing out distinctly as the most robust model with the highest overall accuracy and macro averages. However, there is a recurring challenge across most models for classifying Pleural Rub. Thus, a number of them performed poorly in this category. Fine Crackle and Rhonchi also slightly underperformed, mainly in the recall area, indicating that their respective models may require more optimization to classify them better in such tasks. However, despite the above-mentioned errors in performance, a higher overall accuracy and good consistency of the models generally suggest that deep learning models are promising for further improving automated respiratory sound classification, especially with the VGG and Inception-based architectures in this work. Further refinement, especially in handling rarer categories, is crucial to enhance the reliability and applicability of these models in clinical settings.

6. COMPARATIVE ANALYSIS

The mean precision, recall, f1-score, and accuracy of each method were used for comparison. Table 6 displays the overall outcomes of the classifiers employed for the ARS categorization.

Table 6: Results were compared for each method used in terms of accuracy, precision, recall, and F1-score.

Classifiers	Precision	Recall	F1-Score	Accuracy (%)
VGG B1	0.96	0.94	0.93	0.96
VGG-B3	0.96	0.94	0.93	0.96
VGG Drop	0.96	0.94	0.93	0.96
VGG-V1	0.77	0.82	0.78	0.85
VGG-V2	0.78	0.80	0.77	0.85
Alex Net	1.01	1.01	1.01	1.01
Inception Net	0.96	0.96	0.95	0.96
ResNet	0.96	0.94	0.93	0.96
LeNet5	0.96	0.91	0.91	0.90

The results introduced in Table 6 feature the exhibition of different calculations with regard to Accuracy, Review, F1-Score, and Exactness. Among the models, AlexNet outperforms any remaining calculations with an uncommon upsides of 1.01 across Accuracy, Review, F1-Score, and Exactness, which shows a practically amazing classification execution. Following AlexNet, models such as VGG_B1, VGG-B3, VGG Drop, Origin Net, and ResNet show solid execution with Accuracy, Review, F1-Score, and Precision drifting around 0.96, demonstrating high consistency and dependability in expectations. These models are

profoundly compelling for distinguishing genuine and limiting bogus negatives.

On the other hand, VGG-V1 and VGG-V2 show somewhat lower performance, with F1-Scores around 0.78-0.77 and Accuracy of 85%, suggesting a moderate but still acceptable level of classification ability compared to the top-performing models. Finally, LeNet5 had the lowest performance in the group with a precision of 0.96, recall of 0.91, F1-Score of 0.91, and accuracy of 90%. This indicates that LeNet5 has a relatively low capacity for capturing true-positive cases, although its performance is still respectable.

7. CONCLUSION

This paper discusses the utilization of a remote stethoscope vest coat that integrates deep convolutional neural networks for detecting and classifying abnormal lung sounds. With this intention, it achieved a good level of accuracy in lung sound classification through techniques utilizing the Fourier Transform for feature extraction and deep learning models that include VGG, AlexNet, and ResNet. Data augmentation strengthened the generalization ability of the model, even for diverse sets of abnormal respiratory sounds. The results from this experiment emphatically reveal that wearable technology combined with advanced machine learning techniques will provide an innovative, non-invasive, and efficient method for the early detection and monitoring of respiratory conditions, with promising benefits for both healthcare professionals and patients alike in the proper investigation during real-time diagnostics.

REFERENCES

- [1] Bardou, D., Zhang, K., & Ahmad, S. M. (2018). Lung sounds classification using convolutional neural networks. *Artificial Intelligence in Medicine*, 88, 58–69. <https://doi.org/10.1016/j.artmed.2018.04.008>
- [2] Belkacem, A. N., Ouhbi, S., Lakas, A., Benkhelifa, E., & Chen, C. (2021). End-to-end AI-based point-of-care diagnosis system for classifying respiratory illnesses and early detection of COVID-19: A theoretical framework. *Frontiers in Medicine*, 8, 372. <https://doi.org/10.3389/fmed.2021.585578>
- [3] Gao, J., Wang, H., & Shen, H. (2020). Machine learning based workload prediction in cloud computing. In 2020 29th International Conference on Computer Communications and Networks (ICCCN) (pp. 1–9). <https://doi.org/10.1109/ICCCN49398.2020.9209730>
- [4] Gao, J., Wang, H., & Shen, H. (2020). Smartly handling renewable energy instability in supporting a cloud datacenter. In 2020 IEEE International Parallel and Distributed Processing Symposium (IPDPS) (pp. 769–778). <https://doi.org/10.1109/IPDPS47924.2020.00084>
- [5] Gao, J., Wang, H., & Shen, H. (2020). Task failure prediction in cloud data centers using deep learning. *IEEE Transactions on Services Computing*. <https://doi.org/10.1109/TSC.2020.2993728>
- [6] Gurung, A., Scrafford, C. G., Tielsch, J. M., Levine, O. S., & Checkley, W. (2011). Computerized lung sound analysis as a diagnostic aid for the detection of abnormal lung sounds: A systematic review and meta-analysis. *Respiratory Medicine*, 105(9), 1396–1403. <https://doi.org/10.1016/j.rmed.2011.05.007>
- [7] Huang, D., Wang, L., & Wang, W. (2023). A multi-center clinical trial for wireless stethoscope-based diagnosis and prognosis of children community-acquired pneumonia. *IEEE Transactions on Biomedical Engineering*, 70(7), 2215–2226. <https://doi.org/10.1109/TBME.2023.3239372>
- [8] Iccer, S., & Genge, S. (2014). Classification and analysis of non-stationary characteristics of crackle and rhonchus lung adventitious sounds. *Digital Signal Processing*, 28, 18–27. <https://doi.org/10.1016/j.dsp.2014.02.001>
- [9] Jin, Y., Cai, L., Cheng, Z., Cheng, H., Deng, T., Fan, Y., et al. (2020). A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). *Military Medical Research*, 7(1), 1–23. <https://doi.org/10.1186/s40779-020-0233-6>
- [10] Landge, K., Kidambi, B. R., Singhal, A., & Basha, A. (2018). Electronic stethoscopes: Brief review of clinical utility, evidence, and future implications. *Journal of Practice in Cardiovascular Sciences*, 4(2), 65. https://doi.org/10.4103/jpcs.jpcs_47_1
- [11] Oweis, R. J., Abdulhay, E. W., Khayal, A., & Awad, A. (2015). An alternative respiratory sounds classification system utilizing artificial neural networks. *Biomedical Journal*, 38(2), 152–161. <https://doi.org/10.4103/2319-4170.137773>

- [12] Palaniappan, R., Sundaraj, K., & Sundaraj, S. (2014). A comparative study of the SVM and k-NN machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals. *BMC Bioinformatics*, 15, 223. <https://doi.org/10.1186/1471-2105-15-22>
- [13] Rauf, H. T., Gao, J., Almadhor, A., Arif, M., & Nafis, M. T. (2021). Enhanced bat algorithm for COVID-19 short-term forecasting using optimized LSTM. *Soft Computing*, 25, 1–11. <https://doi.org/10.1007/s00500-021-06075-8>
- [14] Sakai, T., Kato, M., Miyahara, S., & Kiyasu, S. (2012). Robust detection of adventitious lung sounds in electronic auscultation signals. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)* (pp. 1993–1996). Tsukuba, Japan.
- [15] Sarkar, M., Madabhavi, I., Niranjana, N., & Dogra, M. (2015). Auscultation of the respiratory system. *Annals of Thoracic Medicine*, 10(3), 158. <https://doi.org/10.4103/1817-1737.160831>
- [16] Singh, D., Agusti, A., Anzueto, A., Barnes, P. J., Bourbeau, J., Celli, B. R., et al. (2019). Chronic obstructive lung disease: The GOLD science committee report 2019. *European Respiratory Journal*, 53(5), 1900164. <https://doi.org/10.1183/13993003.00164-2019>
- [17] Weisman, I. (2003). Erratum: ATS/ACCP statement on cardiopulmonary exercise testing. *American Journal of Respiratory and Critical Care Medicine*, 167(10), 1451–1452. <https://doi.org/10.1164/ajrccm.167.10.952>
- [18] Tirumanadham, N. S. K. M. K., S, T., & M, S. (2024). Improving predictive performance in e-learning through hybrid 2-tier feature selection and hyper parameter-optimized 3-tier ensemble modeling. *International Journal of Information Technology*, 16(8), 5429–5456. <https://doi.org/10.1007/s41870-024-02038-y>
- [19] N. S. K. M. K. Tirumanadham, T. S and S. M, "Evaluating Boosting Algorithms for Academic Performance Prediction in E-Learning Environments," 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bangalore, India, 2024, pp. 1-8, doi: 10.1109/IITCEE59897.2024.10467968
- [20] N. S. K. M. K. Tirumanadham, T. S and S. M, "Evaluating Boosting Algorithms for Academic Performance Prediction in E-Learning Environments," 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), Bangalore, India, 2024, pp. 1-8, doi: 10.1109/IITCEE59897.2024.10467968.
- [21] N S Koti Mani Kumar Tirumanadham et al., "Boosting Student Performance Prediction In E-Learning: A Hybrid Feature Selection And Multi-Tier Ensemble Modelling Framework With Federated Learning", *Journal of Theoretical and Applied Information Technology*, vol. 103, no. 5, March 2025.
- [22] Shariff, V., Paritala, C., & Ankala, K. M. (2025). Optimizing non small cell lung cancer detection with convolutional neural networks and differential augmentation. *Scientific Reports*, 15(1), 15640. <https://doi.org/10.1038/s41598-025-98731-4>
- [23] S Phani Praveen et al., "Enhanced Predictive Modeling for Alzheimer's Disease: Integrating Cluster-Based Boosting And Gradient Techniques With Optimized Feature Selection", *Journal of Theoretical and Applied Information Technology*, vol. 103, no. 8, April 2025.
- [24] Thatha, V. N., Chalichalamala, S., Pamula, U., Krishna, D. P., Chinthakunta, M., Mantena, S. V., Vahiduddin, S., & Vatambeti, R. (2025). Optimized machine learning mechanism for big data healthcare system to predict disease risk factor. *Scientific Reports*, 15(1), 14327. <https://doi.org/10.1038/s41598-025-98721-6>
- [25] Tirumanadham, N. S. K. M. K., Priyadarshini, V., Praveen, S. P., Thati, B., Srinivasu, P. N., & Shariff, V. (2025). Optimizing lung cancer prediction models: A hybrid methodology using gwo and random forest. In P. Barsocchi, P. Naga Srinivasu, A. Kumar Bhoi, & F. Palumbo (Eds.), *Enabling Person-Centric Healthcare Using Ambient Assistive Technology*, Volume 2 (Vol. 1191, pp. 59–77). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-82516-3_3
- [26] S Phani Praveen et al., "Ai- Powered Diagnosis: Revolutionizing Healthcare With Neural Networks", *Journal of Theoretical and Applied Information Technology*, vol. 103, no. 3, February 2025
- [27] S, S., Raju, K. B., Praveen, S. P., Ramesh, J. V. N., Shariff, V., & Tirumanadham, N. S. K. M. K. (2025f). Optimizing Diabetes Diagnosis: HFM with Tree-Structured Parzen Estimator for Enhanced Predictive Performance and

- Interpretability. Fusion Practice and Applications, 19(1), 57–74. <https://doi.org/10.54216/fpa.190106>
- [28] S. S., Kodete, C. S., Velidi, S., Bhyrapuneni, S., Satukumati, S. B., & Shariff, V. (2024e). Revolutionizing Healthcare: A Comprehensive Framework for Personalized IoT and Cloud Computing-Driven Healthcare Services with Smart Biometric Identity Management. *Journal of Intelligent Systems and Internet of Things*, 13(1), 31–45. <https://doi.org/10.54216/jisiot.130103>
- [29] S. P. Praveen, P. Chaitanya, A. Mohan, V. Shariff, J. V. N. Ramesh and J. Sunkavalli, "Big Mart Sales using Hybrid Learning Framework with Data Analysis," 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2023, pp. 471-477, doi: 10.1109/ICACRS58579.2023.10404941.
- [30] S Phani Praveen et al., "Security Model For Cloud Services Based On A Quantitative Governance Modelling Approach", *Journal of Theoretical and Applied Information Technology*, vol. 101, no. 7, 15th April 2023
- [31] K. Arava, C. Paritala, V. Shariff, S. P. Praveen and A. Madhuri, "A Generalized Model for Identifying Fake Digital Images through the Application of Deep Learning," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2022, pp. 1144-1147, doi: 10.1109/ICESC54411.2022.9885341
- [32] V. V. Chamundeeswari, V. S. Divya Sundar, D. Mangamma, M. Azhar, B. S. S P Kumar and V. Shariff, "Brain MRI Analysis Using CNN-Based Feature Extraction and Machine Learning Techniques to Diagnose Alzheimer's Disease," 2024 First International Conference on Data, Computation and Communication (ICDCC), Sehore, India, 2024, pp. 526-532, doi: 10.1109/ICDCC62744.2024.10961923.
- [33] C. S. Kodete, D. Vijaya Saradhi, V. Krishna Suri, P. Bharat Siva Varma, N. S Koti Mani Kumar Tirumanadham and V. Shariff, "Boosting Lung Cancer Prediction Accuracy Through Advanced Data Processing and Machine Learning Models," 2024 4th International Conference on Sustainable Expert Systems (ICES), Kaski, Nepal, 2024, pp. 1107-1114, doi: 10.1109/ICES63445.2024.10763338.
- [34] B. Thuraka, V. Pasupuleti, C. S. Kodete, R. S. Chigurupati, N. S. K. M. K. Tirumanadham and V. Shariff, "Enhancing Diabetes Prediction using Hybrid Feature Selection and Ensemble Learning with AdaBoost," 2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Kirtipur, Nepal, 2024, pp. 1132-1139, doi: 10.1109/I-SMAC61858.2024.10714776.
- [35] S. Vahiduddin et al., "Revolutionizing Healthcare With Large Language Models: Advancements, Challenges, And Future Prospects In Ai-Driven Diagnostics And Decision Support", *Journal of Theoretical and Applied Information Technology*, vol. 103, no. 9, 15th May 2025
- [36] S. Voddi, U. Sirisha, S. P. Praveen, T. K. Sai Pandraju, N. A. Al-Dmour and S. Islam, "Hybrid CNN-GCN Model for Tumor Classification: Integrating Spatial Relationships in Medical Imaging," 2024 International Conference on Decision Aid Sciences and Applications (DASA), Manama, Bahrain, 2024, pp. 1-6, doi: 10.1109/DASA63652.2024.10836627.
- [37] S. P. Praveen et al., "Enhanced feature selection and ensemble learning for cardiovascular disease prediction: hybrid GOL2-2 T and adaptive boosted decision fusion with babysitting refinement", *Frontiers in Medicine*, vol. 11, Jul. 2024.