Pulmonary Disease Classification Using Respiratory Sounds

Allen Manoj (190501015)
Janani K (190501045)
Ritunjay M (190501099)

CS18811 – PROJECT WORK [Third Review]

Project Connect, SVCEAA

Batch No : 05
Name of the Internal Guide : Ms. G. R. Khanaghavalle (Asst. Professor)
Name of the External Guide : Mr. Dinesh Koka (Co-Founder and CEO)
Name of the Industry : Onward Assist (Healthcare)
Date of Review : 16/05/2023
Domain : Artificial Intelligence

TABLE CONTENTS

PROBLEM STATEMENT

- India has 18% of the global population and an increasing burden of chronic respiratory diseases. A Pulmonary disease may be caused by infections, smoking tobacco or other forms of air pollution.
- According to the world health organization report in 2017 more than 235 million people are suffering from asthma worldwide. In addition, chronic obstructive pulmonary disease (COPD) is expected to be the third leading cause of death by 2030.
- Our project aims to diagnose the pulmonary diseases with the provided respiratory sounds using latest technologies such as deep learning and digital stethoscope to achieve high accuracy in a safe, cost-effective and non-invasive way.

CO-PO-PSO MAPPING

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2
CO1	3	3		2		3		3	3		2	2		3
CO2			3	3	2	3		3	3	2			1	3
CO3	2	3	3		3	3	2		2	2		3	2	
CO4	3	3	2	2	3			1	3	2	1	2	1	1
CO5			3		3	3		3	3	2	2	1	3	3

SUBJECT MAPPING

S.No	Subject Code	Subject Name
1	CS18601	Artificial Intelligence
2	CS18604	Machine Learning
3	CS18002	Advanced User Interface Technologies

ABSTRACT

- Pulmonary disease is a type of disease that affects the lungs and the respiratory system and may be caused by infections, smoking tobacco, or other forms of air pollution.
- Pulmonary auscultation is one of the oldest techniques used in the diagnosis of the respiratory system. Through a stethoscope, the sound of air moving inside and outside the lungs during breathing can be auscultated through chest walls allowing a physiotherapist to identify any pulmonary diseases.
- With the immense development in technology, an automated system for the Classification of Pulmonary Disease can be developed. The automated system will use Deep Learning techniques to do the classification.
- A Deep Neural Network will be developed for the classification of respiratory sounds to identify the disease in the lungs. This will also aid in the efficient identification of diseases in a safe, non-invasive, environment-friendly, and sustainable way, which will improve the lives of the patients.

ISSUES AND CHALLENGES

- Generally used methods of diagnosis involve Chest X-Ray Scans which are studied by a medical professional. X-Rays are not considered as cost-effective and is not safe for disposal in a large scale.

- Using the traditional manual stethoscope, many diseases could be misdiagnosed or go undetected due to the inability of hearing their corresponding respiratory sounds. The diagnosis is usually affected by the quality of the tool, physician experience, etc.

- ML models are used which cannot be used for larger datasets like ICBHI.

- Binary and Ternary-class classification has only been implemented.

LITERATURE SURVEY

Lung Sound Classification Using Co-Tuning and Stochastic Normalization

Authors: Truc Nguyen, Franz Pernkopf
Year of Publication: 2022
Published in: IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING

Learning:
In this paper, pre-trained ResNet models are used as backbone architectures for the classification of adventitious lung sounds and respiratory diseases.

Transferred knowledge of pre-trained models from different ResNet architectures is exploited by vanilla fine-tuning, co-tuning, stochastic normalization, and the combination of co-tuning and stochastic normalization techniques.

Furthermore, spectrum correction and flipping data augmentation are introduced to improve the robustness of our system.

LITERATURE SURVEY

Multi-channel Lung Sound Classification with Convolutional Recurrent Neural Networks

Authors: Elmar Messner, Melanie Fediuk, Paul Swatek, Stefan Scheidl, Freyja-Maria Smolle-Jüttner, Horst Olschewski, Franz Pernkopf

Year of Publication: 2021

Published in: ELSEVIER

Learning:

In this paper, an approach for multi-channel lung sound classification is presented.

In particular, a frame-wise classification framework is proposed to process full breathing cycles of multi-channel lung sound recordings with a convolutional recurrent neural network.

From the lung sound recordings, spectrogram features are extracted and compared to different deep neural network architectures for binary classification, i.e. healthy vs. pathological.

Convolutional Neural Networks based efficient approach for classification of lung diseases

Author: Fatih Demir, Abdulkadir Sengur, Varun Bajaj

Year of Publication: 2020

Published in: Health Information Science and Systems

Learning:

In this paper, the lung sound signals were initially converted to spectrogram images by using the time-frequency method. Two deep learning-based approaches were used for lung sound classification.

In the first approach, a pre-trained deep convolutional neural networks (CNN) model was used for feature extraction and a support vector machine (SVM) classifier was used in the classification of the lung sounds.

In the second approach, the pre-trained deep CNN model was fine-tuned (transfer learning) via spectrogram images for lung sound classification.

CNN-MoE based framework for classification of respiratory anomalies and lung disease detection

Authors: L. Pham, H. Phan, R. Palaniappan, A. Mertins and I. McLoughlin

Year of Publication: 2021

Published in: IEEE Journal

Learning:

This paper aims to classify anomalies in respiratory cycles and detect diseases, from respiratory sound recordings.

The framework begins with front-end feature extraction that transforms input sound into a spectrogram representation. Then, a back-end deep learning network is used to classify the spectrogram features into categories of respiratory anomaly cycles or diseases. Finally, a Teacher-Student scheme is applied to achieve a trade-off between model performance and model complexity which holds promise for building real-time applications.

LITERATURE SURVEY

Deep Neural Network for Respiratory Sound Classification in Wearable Devices Enabled by Patient Specific Model Tuning

Author: J. Acharya and A. Basu

Year of Publication: 2020

Published in: IEEE TRANSACTIONS

Learning:

In this work a deep CNN-RNN model is proposed that classifies respiratory sounds based on Mel-spectrograms.

A patient specific model tuning strategy is also implemented that first screens respiratory patients and then builds patient specific classification models using limited patient data for reliable anomaly detection.

Firstly, the proposed model is able to achieve state of the art score on the ICBHI'17 dataset.

Secondly, deep learning models are shown to successfully learn domain specific knowledge when pre-trained with breathing data and produce significantly superior performance compared to generalized models.

EXISTING WORK

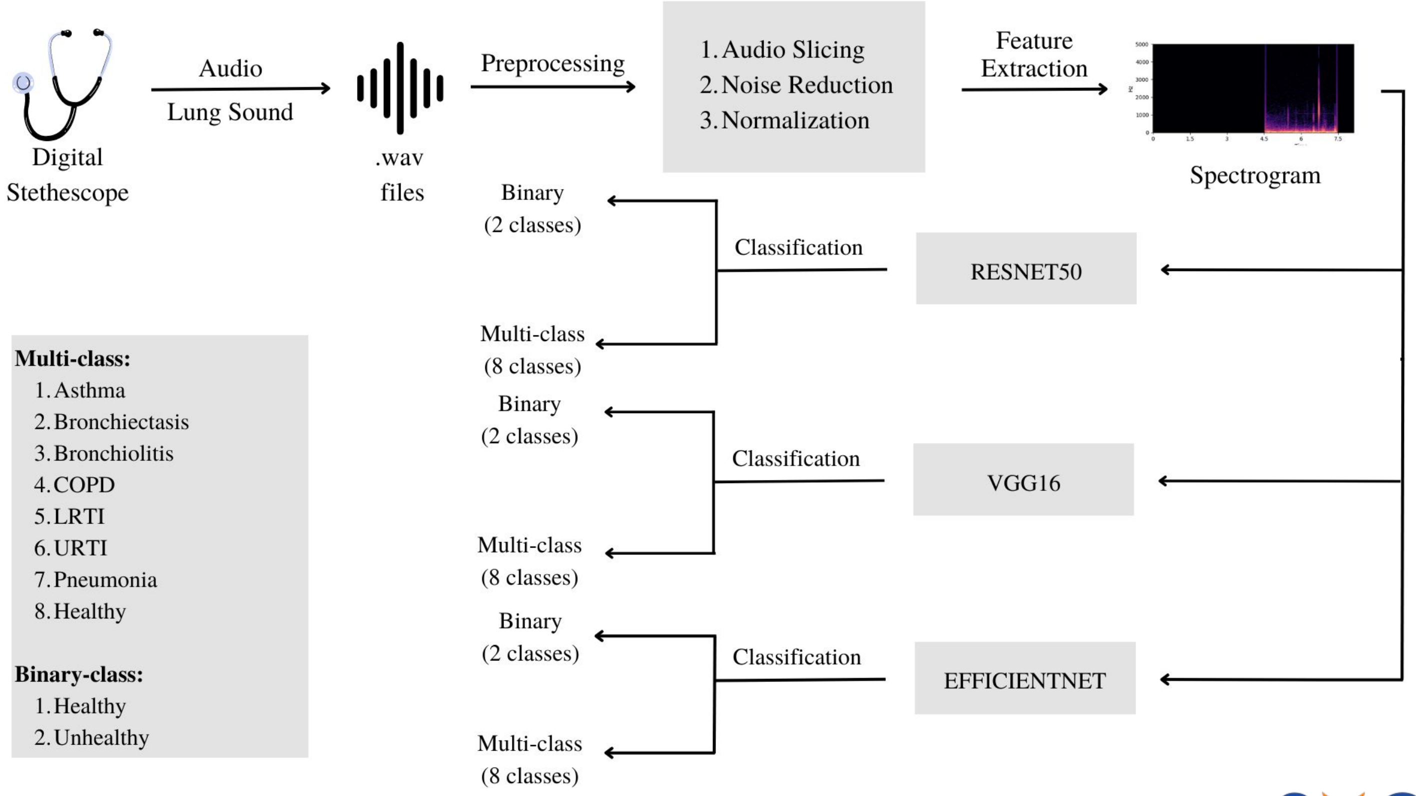
- Almost all proposed lung sound classification systems based on ICBHI 2017 dataset used conventional classifiers. The robust systems used hidden Markov models and Gaussian mixture models for MFCC features or support vector machines (SVMs) for STFT and wavelet features.

- Deep learning systems use CNNs, recurrent neural networks (RNNs) and hybrid architectures.

- Recently, CNN-based systems from diverse architectures i.e. VGGNets, ResNets or their variations have been more and more introduced.

PROPOSED WORK - ARCHITECTURE DIAGRAM

Architecture diagram of Pulmonary Disease Classification Using Respiratory Sounds



PROPOSED WORK

- Digital stethoscope is used to get the lung sounds as .wav files, then preprocessing is done on the audio files by Audio Slicing, Noise Reduction and Normalisation.
- Feature Extraction is implemented using STFT with Spectrogram wavelets to convert the audio .wav files to Spectrogram images for further process.
- Pre-trained models like VGG16, RESNET50 and EFFICIENTNETB0 have been used for classification of 2-class and Multi-class classification.
- The attributes of Binary class are, Healthy and Unhealthy.
- The attributes of Multi-class are Asthma, Chronic Obstructive Pulmonary Disease (COPD), Pneumonia, URTI, LRTI, Bronchiectasis, and Bronchiolitis.

DATASET DESCRIPTION

Dataset of respiratory cycles - Int. Conf. on Biomedical Health Informatics - **ICBHI 2017**

The database consists of a total of **5.5 hours of recordings** containing **6898 respiratory cycles**, of which 1864 contain crackles, 886 contain wheezes, and 506 contain both crackles and wheezes, in 920 annotated audio samples from **126 subjects**.

From the dataset, sound is classified as:

1. **Wheezes** : continuous high pitched sounds, identified for Chronic diseases
2. **Crackles** : discontinuous sounds, identified for Non Chronic diseases
3. Normal : Subtle sounds, no abnormalities.

The dataset predominantly consists of **Stereo**: 2 channel audio files, and few **Mono**: single channel audio files.

Recording equipment used for the dataset

1. AKG C417L Microphone (AKGC417L),
2. 3M Littmann Classic II SE Stethoscope (LittC2SE),
3. 3M Litmmann 3200 Electronic Stethoscope (Litt3200),
4. WelchAllyn Meditron Master Elite Electronic Stethoscope (Meditron)

MODULES & DESCRIPTION

MODULE 1: Pre-processing

MODULE 2: Feature extraction

MODULE 3: Model training

MODULE 4: Validation

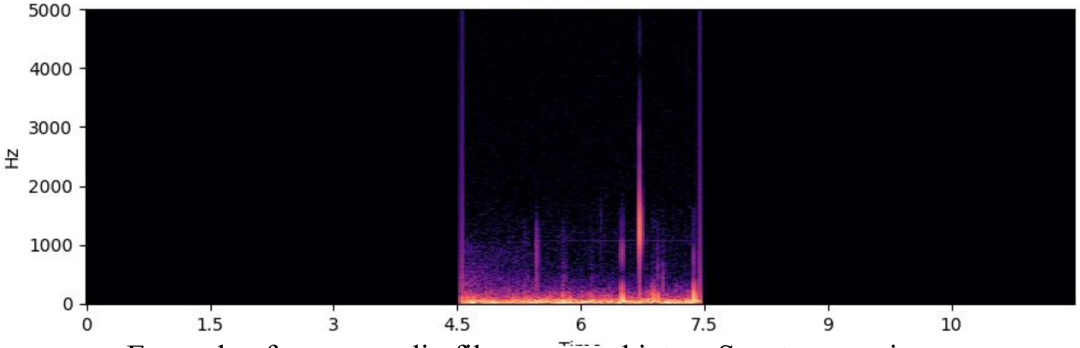
MODULES & DESCRIPTION

MODULE 1: Data Preprocessing

- **Audio Slicing**: The audio file is divided manually with a specified duration from the .txt file provided along with the .wav files, into multiple smaller segments to extract features from different parts of the signal.
- **Noise Reduction**: Any unwanted noise or artifacts are removed from the audio signal using the LMS and NLMS filter for noise reduction. The lung sounds and distinguished from the heart sounds using the frequency range.
- Heart sounds have a lower frequency that varies from 5-600 Hz, whereas lung sounds comparatively have a higher frequency that ranges above 1000Hz.

MODULE 2: Feature extraction

- Extract audio feature Mel Spectrogram.
- A mel spectrogram logarithmically renders frequencies above a certain threshold (the corner frequency). For example, in the linearly scaled spectrogram, the vertical space between 1,000 and 2,000Hz is half of the vertical space between 2,000Hz and 4,000Hz
- This feature is used to generate a spectrogram image of the audio signal.

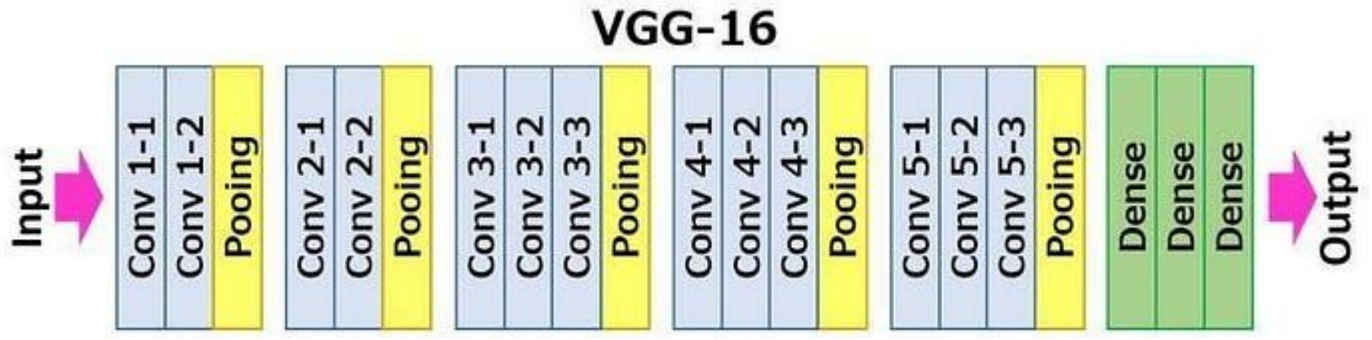


Example of a .wav audio file converted into a Spectrogram image

MODULES & DESCRIPTION

MODULE 3: Model training

- VGG16 (Visual Geometry Group):
  - Most unique thing about VGG16 is that instead of having a large number of hyper-parameters, it follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. In the end it has 2 FC(fully connected layers) followed by a softmax for output. The 16 in VGG16 refers to it has 16 layers that have weights.

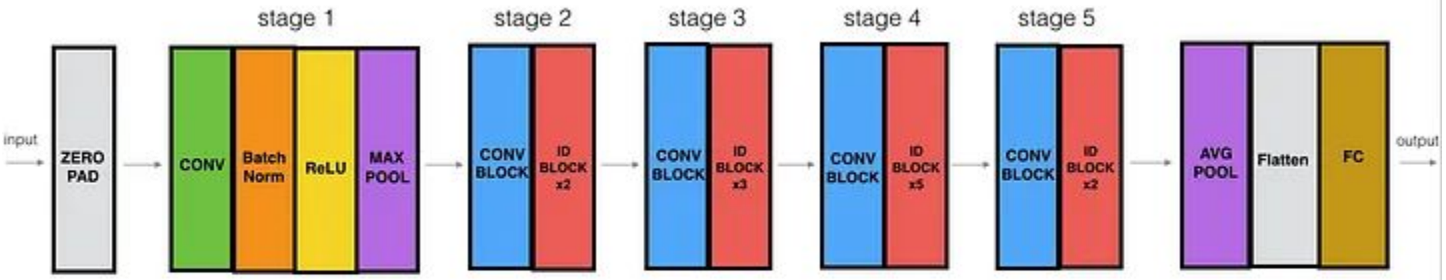


Architecture of VGG-16

MODULES & DESCRIPTION

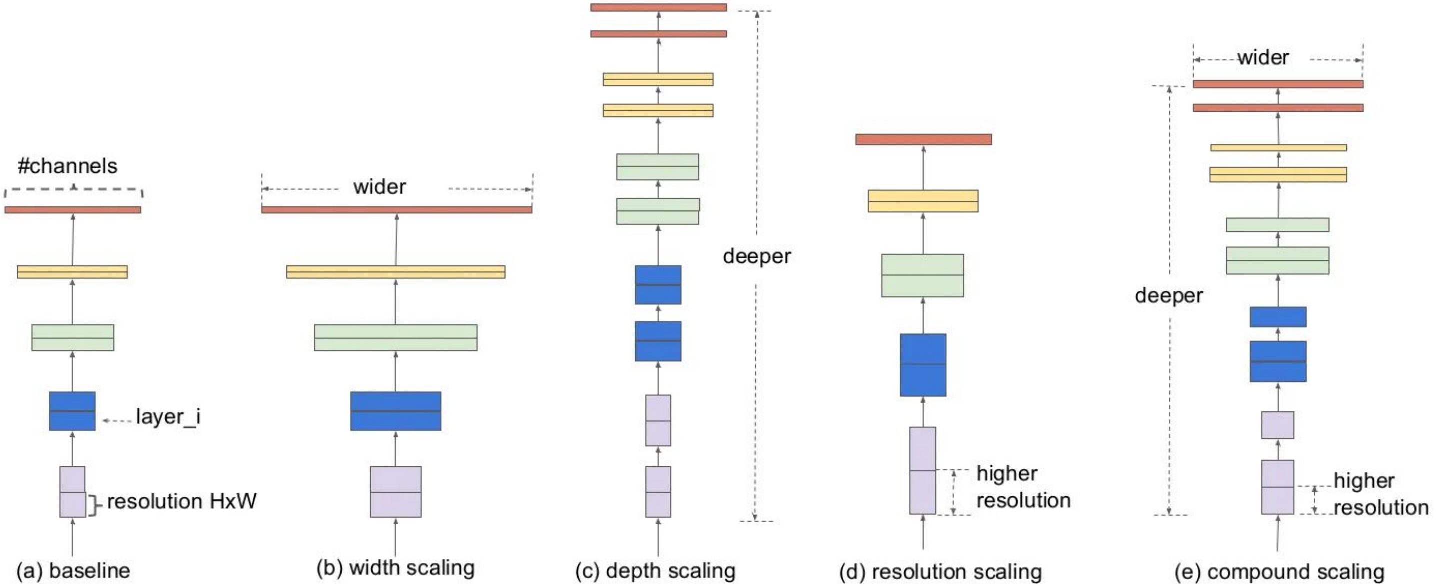
MODULE 3: Model training

- ResNet50 (Residual Networks):
  - ResNet-50 is a 50-layer convolutional neural network (48 convolutional layers, one MaxPool layer, and one average pool layer). Residual neural networks are a type of artificial neural network (ANN) that forms networks by stacking residual blocks.
  - ResNet uses skip connection to add the **output from an earlier layer to a later layer**. This helps it mitigate the vanishing gradient problem.



Architecture of ResNet50

MODULES & DESCRIPTION



Model Scaling:

(a) is a baseline network example;

(b)-(d) are conventional scaling that only increases one dimension of network width, depth, or resolution.

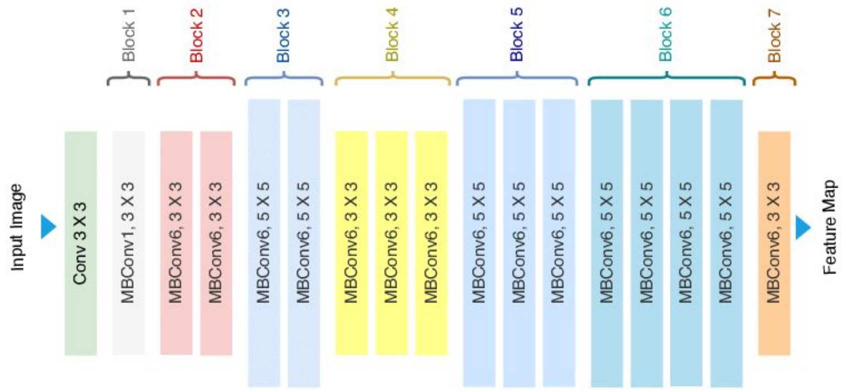
(e) Proposed compound scaling method that uniformly scales all three dimensions with a fixed ratio.

MODULE 3: Model training

- EfficientNet-B0:
  - Compound scaling does not change the operations used inside a layer of the network, instead it expands the network's width, depth and resolution. Hence, it is critical to have a good baseline network.
  - EfficientNet starts with a baseline network (N) and try to expand the length of the network (L), the width of the network (C) and the resolution (W,H) **without changing the baseline architecture**. Thus, the optimization problem can be defined as : **finding the best coefficients** for width (w),depth (d) and resolution(r) that maximizes the accuracy of the network under the constraints of the available resources.

$$\max_{d,w,r} \quad Accuracy\big(\mathcal{N}(d,w,r)\big)$$

$$\text{Memory}(\mathcal{N}) \leq \text{target\_memory}$$
$$\text{FLOPS}(\mathcal{N}) \leq \text{target\_flops}$$



Architecture of EfficientNet-B0

It is necessary to balance all the dimensions of the network (width, depth and resolution) by uniformly scaling each one of them using a **constant ratio**, this method is called **compound scaling**.

MODULES & DESCRIPTION

MODULE 4: Validation

K- fold cross validation:

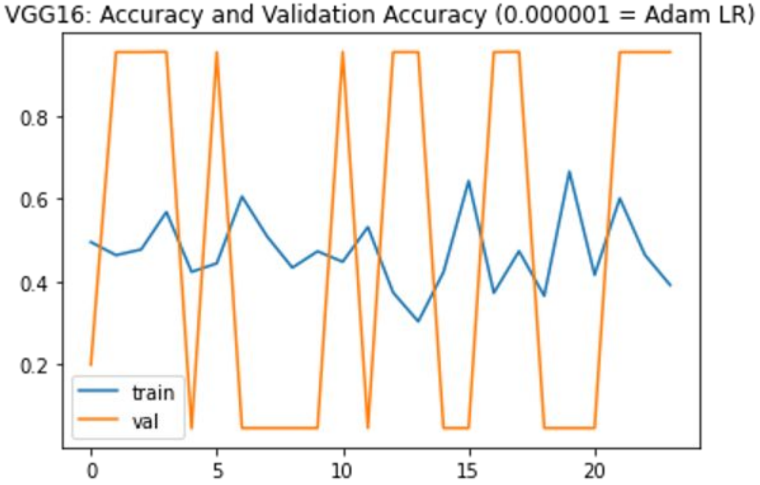
- Eight fold validation was performed for improving the accuracy and fine tuning the model (K=8).
- From our research study, we found that EfficientNet model has produced the highest accuracy and hence Eight- fold validation was performed on it.
- K-fold validation is the process of splitting the dataset for training and testing automatically for each fold.
- It splits the data into k-subsets and the holdout method is repeated k-times where each of the k subsets are used as test set and other k-1 subsets are used for the training purpose.
- This method of validation is proven to give exceptional results with respect to the accuracy.

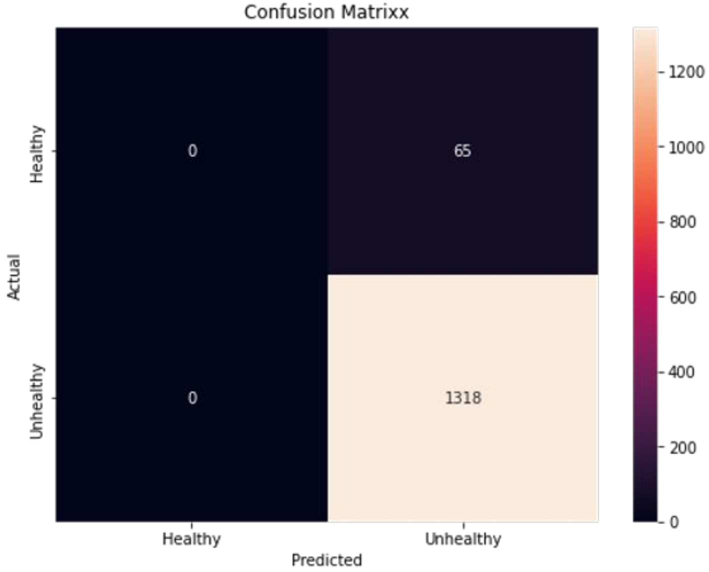
IMPLEMENTATION SCREENSHOTS

MODEL	CLASSIFICATION	ACCURACY
VGG16	BINARY	95%
VGG16	MULTI-CLASS (8 CLASS)	72%
ResNet50	BINARY	80%
ResNet50	MULTI-CLASS (8 CLASS)	67%
EfficientNet-B0	BINARY	94.07%
EfficientNet-B0	MULTI-CLASS (8 CLASS)	75%
EfficientNet-B0	MULTI-CLASS (8 CLASS) WITH 8-FOLD CROSS VALIDATION	80.33%

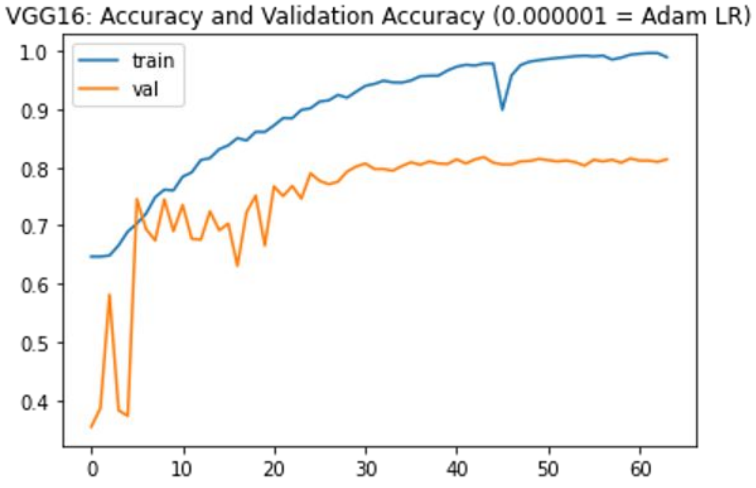
IMPLEMENTATION SCREENSHOTS

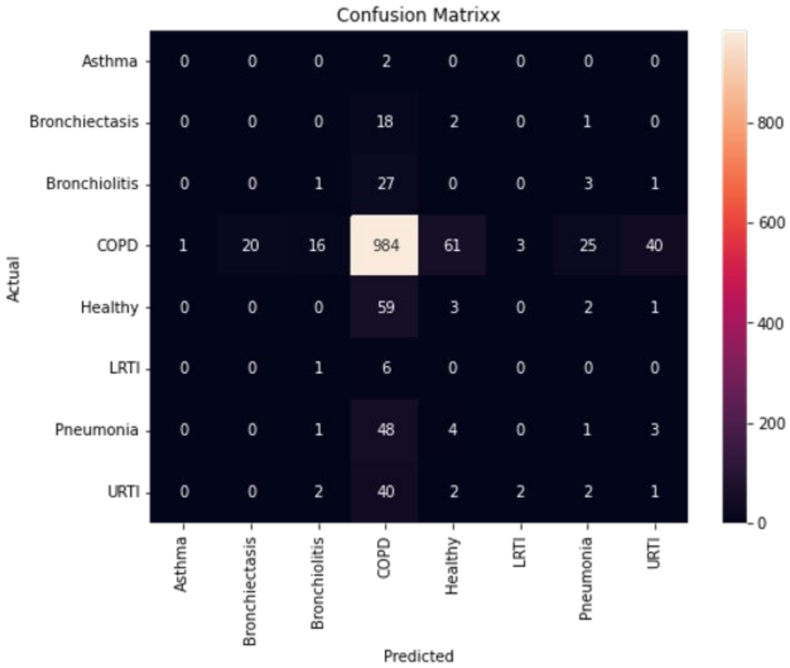
VGG16 BINARY CLASSIFICATION



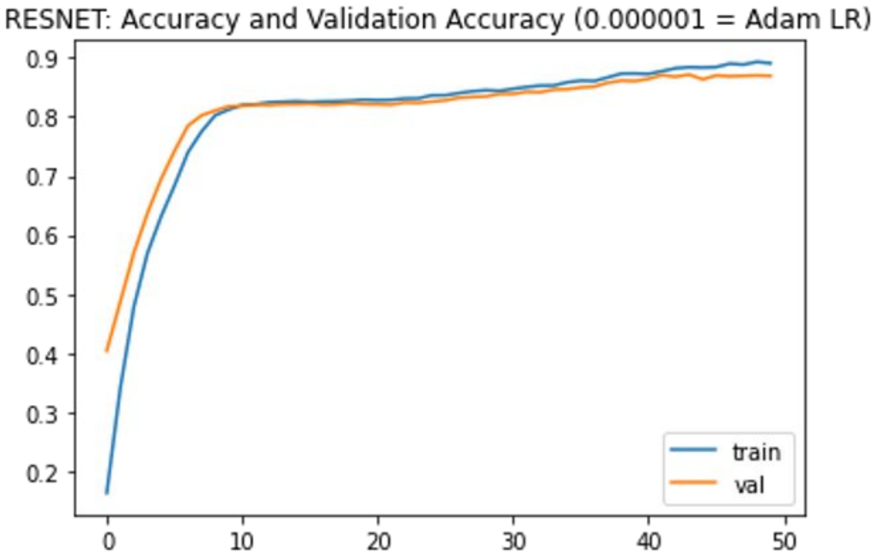


VGG16 MULTICLASS CLASSIFICATION



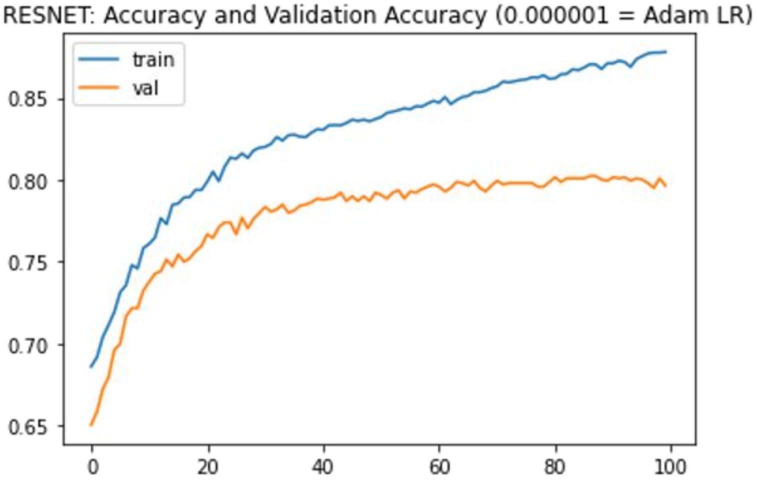


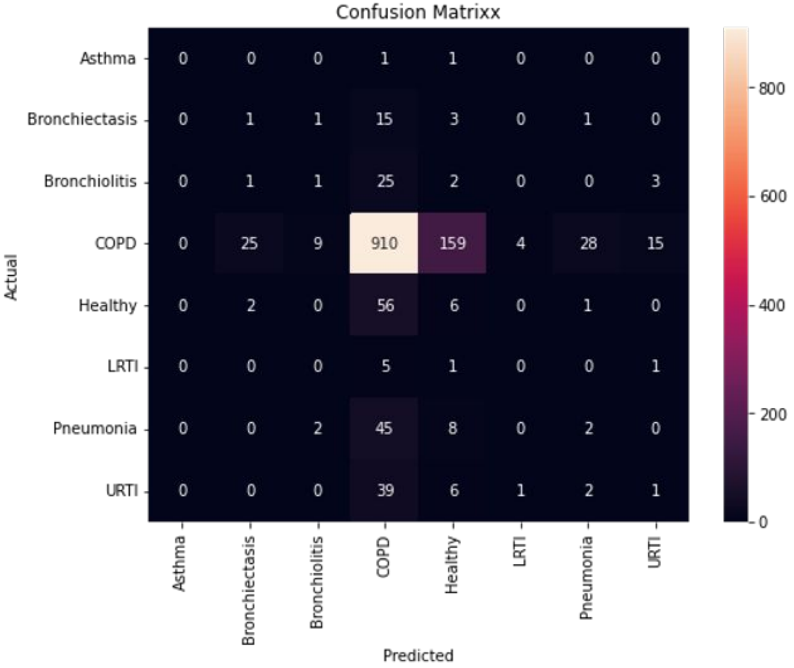
RESNET50 BINARY CLASSIFICATION



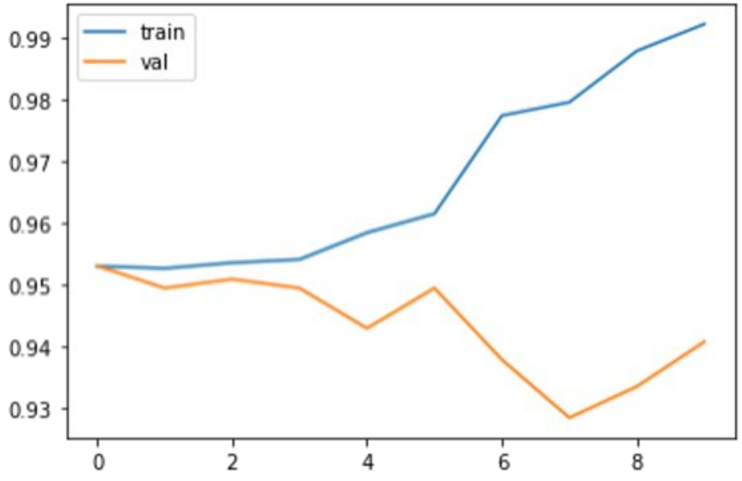
IMPLEMENTATION SCREENSHOTS

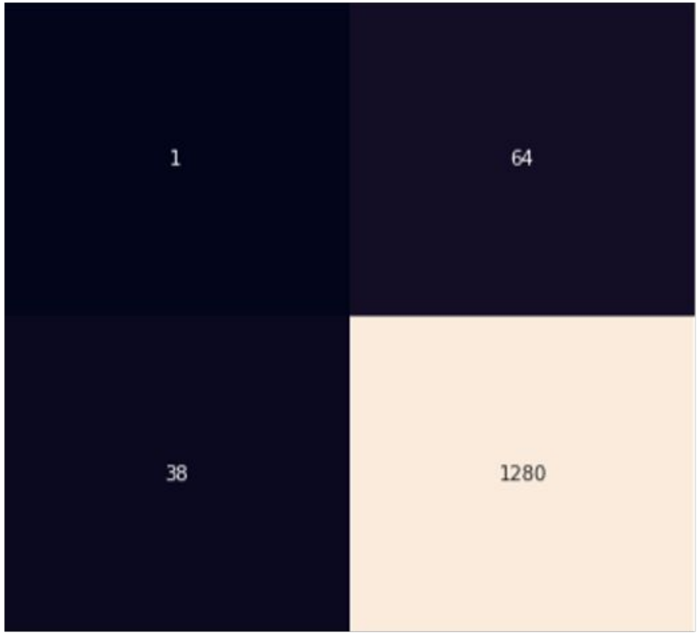
RESNET50 MULTICLASS CLASSIFICATION



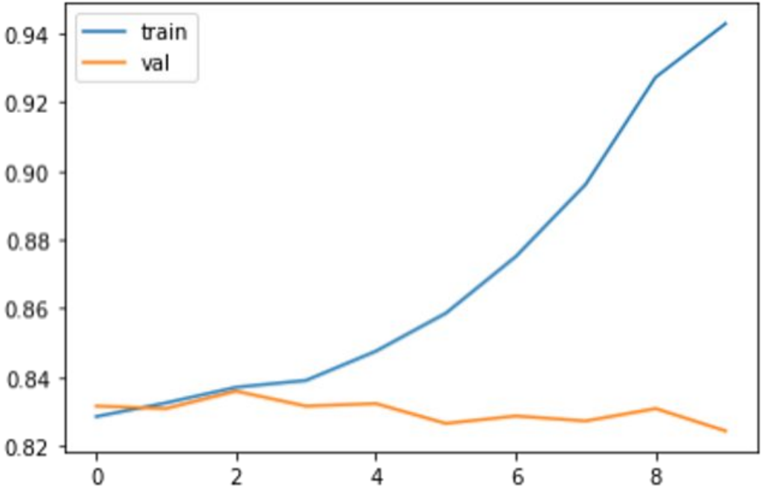


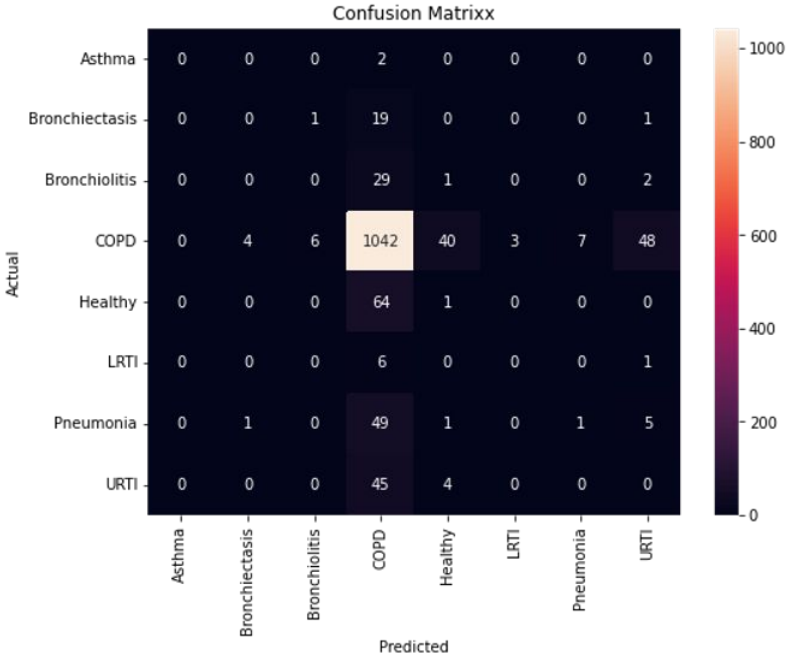
EFFICIENTNET-B0 BINARY CLASSIFICATION



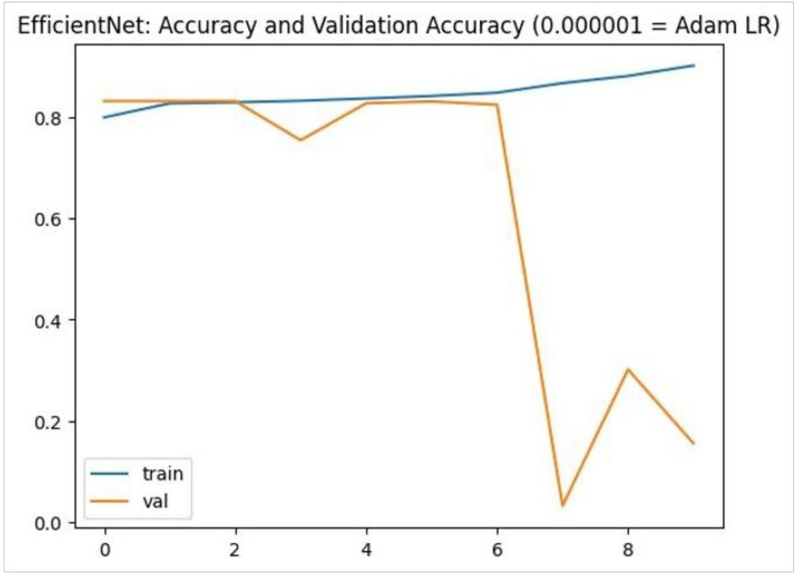


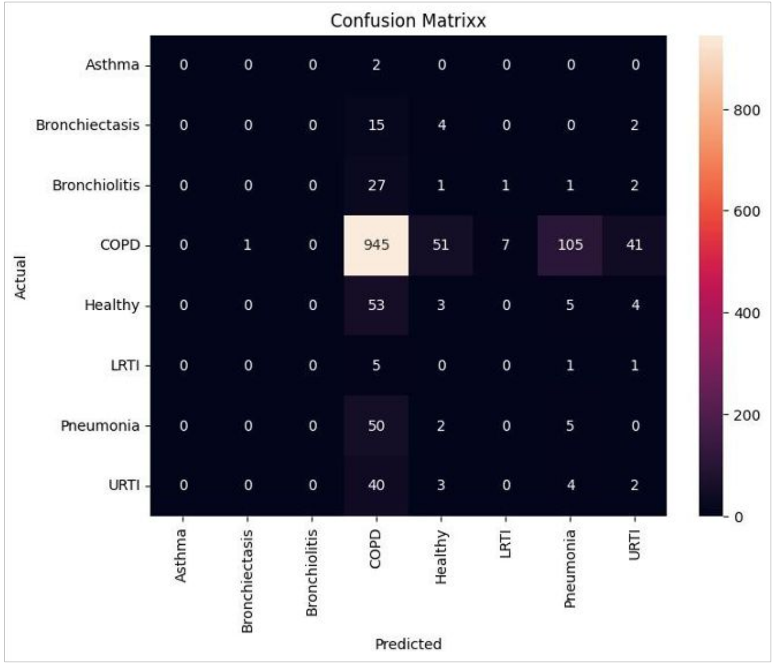
EFFICIENTNET-B0 MULTICLASS CLASSIFICATION





EFFICIENTNET-B0 MULTICLASS CLASSIFICATION WITH K-FOLD CROSS VALIDATION





SYSTEM REQUIREMENTS AND TOOLS

Hardware Requirement:

• **Random Access Memory**	:	8GB Minimum
• **Processors**	:	Intel Core i3 Processor or above

Software Requirement:

• **Operating Systems**	:	Microsoft Windows 7 or later, MacOS Montenary
• **Python Version**	:	3.8.3
• **Development Tools**	:	Anaconda Framework
• **Compatible Tools**	:	Microsoft Visual Studio Code, Spyder
• **Python Libraries**	:	TensorFlow, Keras, Librosa, Pandas, NumPy

REFERENCES

1. T. Nguyen and F. Pernkopf, "Lung Sound Classification Using Co-Tuning and Stochastic Normalization," in IEEE Transactions on Biomedical Engineering, Vol. 69, no. 9, pp. 2872-2882, Sept. 2022, doi: 10.1109/TBME.2022.3156293.

2. Demir, F., Sengur, A. & Bajaj, V. Convolutional neural networks based efficient approach for classification of lung diseases. *Health Inf Sci Syst* 8, 4. doi: 10.1007/s13755-019-0091-3 (2020)

3. Elmar Messner, Melanie Fediuk, Paul Swatek, Stefan Scheidl, Freyja-Maria Smolle-Jüttner, Horst Olschewski, Franz Pernkopf, "Multi-channel lung sound classification with convolutional recurrent neural networks, Computers in Biology and Medicine", Vol. 122, 2020, 103831, ISSN 0010-4825, doi: 10.1016/j.compbiomed.2020.103831.

4. J. Acharya and A. Basu, "Deep Neural Network for Respiratory Sound Classification in Wearable Devices Enabled by Patient Specific Model Tuning," in IEEE Transactions on Biomedical Circuits and Systems, Vol. 14, no. 3, pp. 535-544, June 2020, doi: 10.1109/TBCAS.2020.2981172.

5. L. Pham, H. Phan, R. Palaniappan, A. Mertins and I. McLoughlin, "CNN-MoE Based Framework for Classification of Respiratory Anomalies and Lung Disease Detection," in IEEE Journal of Biomedical and Health Informatics, Vol. 25, no. 8, pp. 2938-2947, Aug. 2021, doi: 10.1109/JBHI.2021.3064237.

THANK YOU