International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.037 ISSN: 2581-4621

BEYOND BREATHE - ANALYZING COUGH SOUNDS TO PREDICT LUNG INFECTIONS

Dharmendra Kumar Roy[1], Arigela Sathwika[2], Chinta Sai Deepthi[3], Dubbaka Poojitha[4], Kadari Pravalika[5]

[1]Associate Professor, Hyderabad Institute of Technology and Management (HITAM), Telangana, India.
[2]B.Tech Student of HITAM, Department of CSE, [3]B.Tech Student of HITAM, Department of CSE,
[4]B.Tech Student of HITAM, CSE Department of CSE,
[5]B.Tech Student of HITAM, Department of CSE.

Abstract— The project aims to address the critical need for accurate prediction of lung diseases to facilitate early diagnosis and personalized treatment strategies. A part from critical health diseases such as cancer and diabetes, the impact of respiratory diseases is also gradually on the rise and is becoming life-threatening for society. The early diagnosis and immediate treatment are crucial in respiratory diseases, and hence the audio of the respiratory sounds. we apply Convolutional Neural Network based deep learning methodologies to assist medical experts by providing a detailed and rigorous analysis of the medical respiratory audio data for Chronic Obstructive Pulmonary detection(COPD), Upper Respiratory Tract Infection(URTI),Pneumonia, Bronchiolitis. The output is in the form MFCC and ROC Curve(Receiver Operating Characteristic (ROC).
Keywords—Rectified Linear Unit, Deep Learning, Chronic Obstructive Pulmonary Disease, Convolution Neural Network, Receiver Operating Characteristic Curve, Upper Respiratory Tract Infection, Mel Frequency Cepstral Coefficient.

INTRODUCTION

Respiratory infections are a pervasive global health concern, contributing significantly to mortality rates across diverse age demographics. Current diagnostic approaches for identifying lung infections are plagued by issues of time inefficiency, high costs, and dependence on medical professionals for confirmation of infection presence before initiating treatment. This project seeks to revolutionize respiratory health diagnostics by introducing an innovative tool that harnesses the power of deep learning algorithms to analyze cough sounds, enabling early detection of lung infections. Respiratory lung infections are illnesses that affect the lungs and can range from mild to severe. Common examples include pneumonia, bronchitis, and tuberculosis. These infections are typically caused by bacteria, viruses, or fungi. Symptoms often include coughing, fever, chest pain, and difficulty breathing.

In the realm of deep learning techniques applied to cough sound classification, several key methods and frameworks play crucial roles. Convolutional Neural Networks (CNNs) are foundational, known for their ability to extract hierarchical features from input data. In this context, CNNs process spectrograms of cough sounds, capturing frequency and temporal patterns that distinguish between different types of coughs. Max pooling, a technique used 4 within CNNs, down samples feature maps to reduce computational complexity while preserving essential information. Activation functions like ReLU (Rectified Linear Unit) are used to introduce non-linearity, enhancing the network's capacity to model complex relationships in data. Frameworks such as TensorFlow provide powerful tools for constructing and training deep learning models, facilitating the implementation and optimization of CNN architectures. Scikit-learn (SKlearn), on the other hand, offers a comprehensive suite of machine learning tools for data preprocessing, model selection, and evaluation. Mel-frequency

DOI:10.46647/ijetms.2024.v08i03.037 ISSN: 2581-4621

cepstral coefficients (MFCC) are commonly used to represent the audio features of cough sounds, capturing the frequency distribution over time.
The objective of a Convolutional Neural Network (CNN) algorithm in deep learning for sound prediction is to achieve precise and robust identification, classification, and interpretation of audio signals. This involves transforming raw audio data into spectrograms, which are two dimensional representations that display the frequency spectrum of the audio signal over time. The CNN is then trained to detect and learn intricate patterns and features from these spectrograms through a series of convolutional layers. The ultimate goal of using CNNs for sound prediction is to achieve high levels of accuracy and generalization. This means the model should perform well not only on the training data but also on new, unseen audio samples. Applications of this technology include automatic speech recognition (ASR), music genre classification, environmental sound detection, and even complex tasks such as speaker identification and sentiment analysis in speech. By leveraging the powerful pattern recognition capabilities of CNNs, deep learning systems can deliver sophisticated and reliable sound prediction and classification solutions.

LITERATURE SURVEY

Based on the provided sources, here is a summary of the key points from each study:

The paper by Hosny et al. (2018) focuses on using deep learning to classify and predict patterns of interstitial lung diseases (ILDs) from CT scans. They employed a Convolutional Neural Network (CNN) to automatically categorize different ILD patterns, such as idiopathic pulmonary fibrosis (IPF), from a large dataset of annotated CT images.[1]
The paper "Predicting Pulmonary Function from Three-Dimensional Chest Radiographs in Cystic Fibrosis" by Castillo et al. (2019) explores the use of three-dimensional chest radiographs to predict pulmonary function in patients with cystic fibrosis. The study investigates how structural changes in the lungs observed from these radiographs can correlate with pulmonary function tests, aiming to improve monitoring and management of cystic fibrosis patients.[2]
The paper "Predicting Lung Cancer Incidence from CT Image Segmentation, Genetic Information, and Demographic Data" by Lindgren et al. (2020) focuses on integrating CT image segmentation, genetic information, and demographic data to predict lung cancer incidence. The study employs machine learning techniques to analyze these diverse datasets, aiming to enhance early detection and risk assessment for lung cancer. The research underscores the potential of combining medical imaging, genetic markers, and patient demographics to improve personalized cancer screening and prevention strategies.[3]
The study by Washko et al. (2021) predicts survival in idiopathic pulmonary fibrosis (IPF) patients using longitudinal clinical data. They employ advanced statistical methods to analyze patient records over time, aiming to identify prognostic factors that influence disease progression and survival outcomes in IPF.[4]
The paper by Sweeney et al. (2019) focuses on predicting mortality in patients with Acute Respiratory Distress Syndrome (ARDS) using clinical and biological biomarkers. The study integrates data from both clinical assessments and biological measurements to develop predictive models for ARDS mortality. The research highlights the potential of combining multiple biomarkers to improve prognostication and clinical management of ARDS, aiming to enhance patient outcomes through early identification of high-risk individuals.[5]
The study by Du et al. (2020) proposes a method to detect Chronic Obstructive Pulmonary Disease (COPD) using snapshots of the 3D lung airway tree. They employ a deep Convolutional Neural Network (CNN) for classification, achieving an accuracy of 88.2% with colorful snapshots, 88.6% with gray snapshots, and 86.4% with binary snapshots. The research demonstrates the potential of using 3D imaging and deep learning techniques to aid in the diagnosis and management of COPD, leveraging different visual representations for effective disease detection.[6]

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.037 ISSN: 2581-4621

EXISTING SYSTEM

The conventional methods of COPD diagnosis were detection which needed images to be fed as input to detect the diseases. In case of any respiratory distress situation like a heart attack or asthma attack, reaching the hospital and making the initial diagnosis via chest scan or x-ray is time-consuming, expensive, and life threatening. Also, the automated AI systems for image-based detection require training the model on vast numbers of high-quality HD images of x-ray, which is challenging to get each time. Instead, there is a requirement of a more straightforward and less resource-intensive system, which can help medical healthcare providers quickly make the initial diagnosis. The system that we have proposed is a system that detects COPD based 8 on respiratory sounds. The sound produced by the body's internal organs is very different in case of a heart attack, asthma, COPD, etc. Automated detecting such sounds to classify if a person is susceptible to COPD is a too time-saving, self-alarming method for both the patient and the doctor. The doctors can use the system for confirmed detection of COPD. In contrast, this system's future scope involves integration with smart devices and mics to record people's sounds routinely and thus predict the possibility of a case of COPD. Lung prediction systems have revolutionized the field of healthcare by leveraging cutting-edge technology to assist in the early detection and accurate diagnosis of respiratory conditions. By utilizing advanced imaging techniques such as CT scans and X-rays, these systems are able to capture detailed internal lung structures, providing healthcare professionals with a comprehensive view of the patient's respiratory health. The process of image preprocessing plays a crucial role in enhancing the quality of the data collected and removing any noise that may distort the images. This ensures that the resulting analysis is as accurate as possible. Feature extraction techniques then come into play, identifying key characteristics within the images such as texture and density, which are instrumental in detecting abnormalities that may indicate the presence of conditions like pneumonia or lung cancer. Machine learning algorithms, particularly Convolutional Neural Networks (CNNs), are then employed to analyze the images and classify them based on the identified features. These algorithms are trained on labeled datasets to learn and improve their predictive capabilities, allowing them to accurately classify images and predict lung conditions with a high degree of accuracy.

PROPOSED SYSTEM

The outlined flowchart offers a structured methodology for analyzing respiratory audio recordings, starting with the aggregation of a diverse dataset encompassing samples from various respiratory conditions and healthy controls. This comprehensive dataset ensures the robustness and generalizability of the subsequent analysis. Through preprocessing, the raw audio data is transformed into a format conducive to feature extraction, enhancing the model's ability to discern relevant patterns. 9 The subsequent processing of extracted features aims to standardize the data, mitigating potential biases and ensuring the model's robustness to variations within the dataset. Leveraging the capabilities of a Convolutional Neural Network (CNN), a powerful deep learning architecture adept at learning spatial hierarchies of features, facilitates the effective analysis of the preprocessed features, enabling the model to capture intricate relationships inherent in respiratory audio data. Post-training, the model's performance is rigorously evaluated using established metrics like accuracy, precision, recall, and F1-score, providing quantitative insights into its classification efficacy across different respiratory conditions. Furthermore, visual aids such as ROC curves enrich the interpretation of results by illustrating the trade-off between true positive and false positive rates at various classification thresholds, thus aiding in the selection of an optimal operating point. By encompassing data preprocessing, deep learning modeling, and thorough performance assessment, this approach not only enables accurate diagnosis and classification of respiratory conditions but also lays the foundation for future advancements in automated respiratory health monitoring and diagnosis through audio analysis.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.037 ISSN: 2581-4621

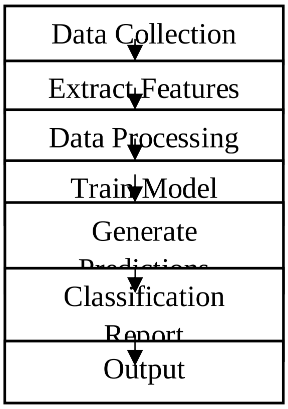

Figure.1 Proposed System

Software Requirements like Google Colab offers a Python environment, alleviating the need for explicit installation. However, specific Python libraries such as NumPy, Pandas, TensorFlow, and Scikit-learn are essential for data manipulation, machine learning modeling, and evaluation. Operating on a Linux-based environment, Google Colab supports various distributions, ensuring compatibility with Python and its libraries.

Hardware Requirements where Google Colab leverages virtual machines with varying specifications. For optimal performance, a processor with multiple cores like Intel Xeon or AMD Ryzen enhances computational efficiency during model training. Adequate RAM, preferably 16GB or more, ensures smooth processing of large datasets and complex models. A minimum of 50GB hard disk space is advisable to accommodate datasets, code files, and model checkpoints. Additionally, a standard keyboard suffices for code input and interaction within the Colab environment. Processor - Intel Xeon or AMD Ryzen RAM - 16GB Hard Disk - 50GB Keyboard - 110 Keys Enhanced

IMPLEMENTATION

The proposed system involves the development of a Convolutional Neural Network (CNN) to classify respiratory sounds into different disease categories using MFCC (Mel-Frequency Cepstral Coefficients) features extracted from audio files. The dataset used is the Respiratory Sound Database, containing recordings of respiratory sounds from patients with various respiratory conditions.

Data Preparation is the first stage. The dataset was loaded from the Kaggle input directory, and audio files were identified using the os and os.path libraries. Patient IDs were extracted from filenames to link them with diagnosis labels provided in a CSV file. The audio files were processed to extract MFCC features, which are effective in representing the spectral properties of the sounds.

Feature Extraction is then done where the MFCC features were extracted from the audio files using the librosa library. Each audio file was padded to ensure uniform feature length, necessary for CNN input.

Data Labeling and Preprocessing plays a major role where the labels corresponding to each audio file were extracted and one-hot encoded. Some classes (Asthma, LRTI) were removed to balance the dataset.

Model Architecture is the CNN model that was designed with multiple convolutional and pooling layers, followed by global average pooling and a dense output layer with softmax activation. Dropout layers were included to prevent overfitting.Model Training and Evaluation is the next process. The model was trained on the dataset with an 80-20 train-test split, using Model Checkpoint to save the best model based on validation accuracy. The model achieved good performance in classifying respiratory sounds. ROC curves and confusion matrices were plotted to visualize performance across different classes. The implementation of the CNN model for classifying respiratory sounds demonstrates its potential in assisting medical diagnosis. The project highlighted the effectiveness of

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.037 ISSN: 2581-4621

MFCC features in representing audio data and the use of CNNs in audio classification tasks. Future work can explore data augmentation techniques and model optimizations to further improve accuracy.

RESULT AND ANALYSIS

This project aimed to classify respiratory sounds into disease categories using a CNN and MFCC features from the Respiratory Sound Database. Preprocessing involved linking patient IDs with diagnosis labels and extracting MFCC features from audio files using librosa, ensuring uniform length through padding. The CNN model included multiple convolutional and pooling layers with dropout for regularization, followed by a dense output layer with softmax activation. Training on an 80-20 train-test split, with Model Checkpoint to save the best model, the expected results were good classification performance, as indicated by high accuracy, detailed in ROC curves, and confusion matrices. The project demonstrated the potential of CNNs and MFCC features in effectively classifying respiratory conditions, contributing to advancements in automated medical diagnostics.

	precision	recall	f1-score	support
Bronchiolitis	1.00	0.33	0.50	3
Bronchiectasis	0.25	0.50	0.33	2
COPD	0.89	0.99	0.93	158
Healthy	0.00	0.00	0.00	7
Pneumonia	0.00	0.00	0.00	7
URTI	0.00	0.00	0.00	5
Accuracy			0.87	182
Macro Avg	0.36	0.30	0.29	182
Weighted Avg	0.79	0.87	082	182

Figure.2 Classification Report

The ROC curve, or Receiver Operating Characteristic curve, is a crucial tool for evaluating the performance of a classification model, especially in medical diagnostics. It plots the true positive rate (sensitivity) against the false positive rate (1-specificity) at various threshold settings. The importance of the ROC curve lies in its ability to illustrate the trade-offs between sensitivity and specificity, providing a comprehensive view of the model's diagnostic ability across different thresholds. A key feature of the ROC curve is the Area Under the Curve (AUC). A higher AUC value, closer to 1, indicates better model performance, meaning the model is more effective at distinguishing between the positive and negative classes. This is particularly important in medical diagnostics where the cost of false positives (misdiagnosing a healthy patient) and false negatives (failing to diagnose a sick patient) can be significant.

By examining the ROC curve, we can choose an optimal threshold that balances sensitivity and specificity according to the clinical context, improving decision-making and patient outcomes. In summary, the ROC curve is a valuable metric for understanding and optimizing model performance in critical applications like healthcare. For the ROC curve results, we expect high Area Under the Curve (AUC) values for each respiratory disease category. A high AUC, ideally close to 1.0, indicates that the model is effective at distinguishing between different conditions. This means the model can accurately identify whether a patient has a specific disease with minimal errors. High AUC values show that the model makes few false positives (incorrectly predicting a disease when there isn't one) and few false negatives (missing a disease when it is present). Essentially, good ROC curves mean the model can reliably classify respiratory sounds, enhancing its potential use in assisting doctors with accurate diagnoses. This contributes to better automated medical diagnostics for respiratory conditions, improving patient care and treatment outcomes.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.037 ISSN: 2581-4621

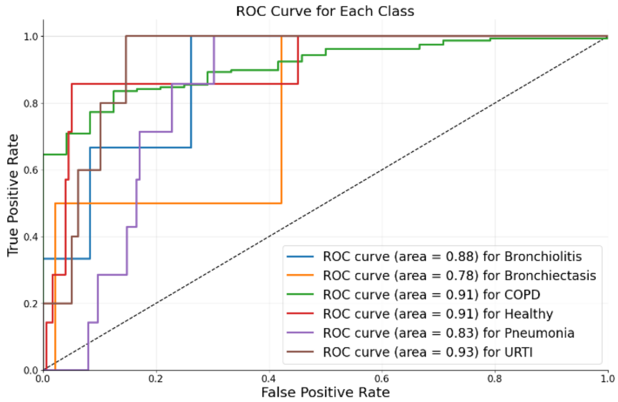

Figure.3 ROC Curve

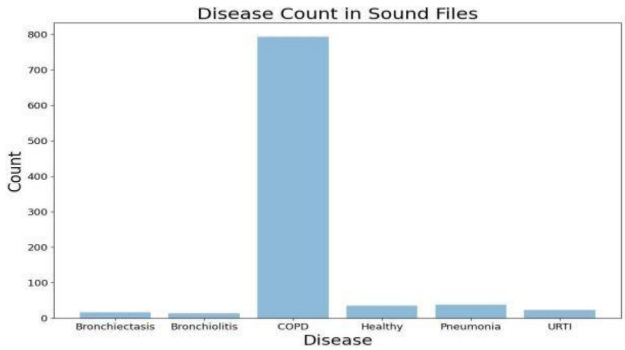

Figure.4 Graph Predicting most affected disease

The accuracy curve shows the model's performance over each training epoch, indicating how well it is learning to classify the respiratory sounds. During training, we expect to see the accuracy of both the training and validation datasets improve over time. The training accuracy should increase steadily, reflecting the model's ability to learn from the training data. The validation accuracy, which measures how well the model performs on unseen data, should also rise, ideally following a similar trend to the training accuracy but typically at a slightly lower rate. If the validation accuracy plateaus or decreases while the training accuracy continues to rise, it could indicate overfitting, where the model is too tailored to the training data and doesn't generalize well to new data. A well-performing model will show both high training and validation accuracy, demonstrating its capability to accurately classify new respiratory sound samples.

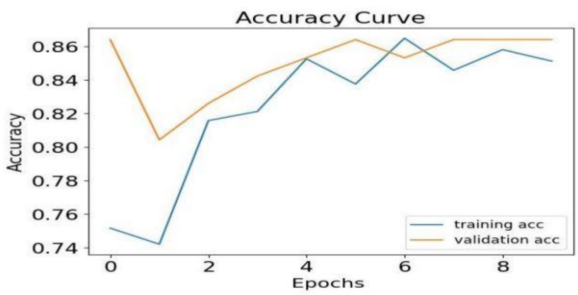

Figure.5 Accuracy Curve

Conclusion

In conclusion, the incorporation of deep learning methodologies into the analysis of cough audio data heralds a transformative era in the diagnosis of lung infections. This project, featuring its utilization of Mel-Frequency Cepstral Coefficients (MFCC) and a Convolutional Neural Network (CNN)-based

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 3 Volume No.8May - June – 2024
DOI:10.46647/ijetms.2024.v08i03.037 ISSN: 2581-4621

prediction model, presents a pioneering solution to the limitations of traditional diagnostic approaches. By streamlining the diagnostic process, reducing time and cost burdens, enhancing accessibility, and offering a non-invasive alternative, this innovation has the potential to revolutionize respiratory health diagnostics. The project's MFCC-based feature extraction and CNN-based prediction model not only offer heightened accuracy and efficiency but also pave the way for a more nuanced understanding of respiratory conditions. By harnessing the power of artificial intelligence, it opens avenues for early detection and personalized treatment strategies, thereby improving patient outcomes and alleviating the strain on healthcare systems. Looking ahead, continued research, collaboration, and refinement hold the key to realizing the full potential of this groundbreaking technology. With ongoing advancements, the vision of a more efficient, cost-effective, and accessible respiratory health diagnostic tool is within reach, promising significant benefits for individuals and healthcare systems worldwide. Through collective efforts, the journey towards a future where respiratory conditions are diagnosed swiftly, accurately, and comprehensively is well underway, offering hope for a healthier and more resilient global community.

References

[1] "Deep learning-based classification and prediction of interstitial lung disease patterns on CT" (Hosny et al, 2018)

[2] "Predicting Pulmonary Function from Three-Dimensional Chest Radiographs in Cystic Fibrosis" (Castillo et al., 2019)

[3] "Predicting Lung Cancer Incidence from CT Image Segmentation, Genetic Information, and Demographic Data" (Lindgren et al., 2020)

[4] "Predicting Survival of Idiopathic Pulmonary Fibrosis Patients using Longitudinal Clinical Data" (Washko et al., 2021)

[5] "Prediction of Acute Respiratory Distress Syndrome (ARDS) Mortality using Clinical and Biological Biomarkers" (Sweeney et al., 2019)

[6] 2020 Du et al. (2020) To detect COPD using snapshots of 3D lung airway tree and with the help of deep CNN Classification by colorful snapshot.