**Lung Disease Detection Using Ensemble Technique**

Contents

1. Introduction (250 words) .................................................................................................................................. 1

2. Problem Statement (500 words) ....................................................................................................................... 1

3. Aims and Objectives (500 words) .................................................................................................................... 1

4. Legal, Social, Ethical and Professional Considerations (250 words). .............................................................. 2

5. Background (1500 words) ................................................................................................................................ 2

6. References ........................................................................................................................................................ 2

1. **Introduction**

Chronic Obstructive Pulmonary Disease (COPD) is a progressive lung disease that significantly impacts patients’ quality of life and healthcare systems worldwide. It is characterized by persistent respiratory symptoms, airflow limitation, and progressive lung function decline. COPD includes conditions such as emphysema and chronic bronchitis, both of which lead to breathing difficulties, increased hospitalization rates, and higher mortality risks. According to the World Health Organization (WHO), COPD is a leading cause of morbidity and mortality, affecting millions of people globally, with many cases remaining undiagnosed until the disease reaches an advanced stage.

This project focuses on COPD detection using machine learning and ensemble methods. The motivation behind this research stems from the urgent need for early diagnosis and efficient risk assessment to improve patient outcomes. Traditional diagnostic methods rely on clinical assessments, spirometry tests, and imaging, which can be expensive, time-consuming, and inaccessible in resource-limited settings. Machine learning models offer a data-driven approach to identify high-risk individuals, predict disease severity, and assist in clinical decision-making.

The study is particularly relevant to the healthcare industry and medical research as early detection of COPD can reduce hospital admissions, optimize treatment plans, and improve disease management. By leveraging patient data, this project aims to develop predictive models that enhance diagnostic accuracy and provide actionable insights for healthcare professionals. With an increasing focus on personalized medicine and AI-driven diagnostics, this study contributes to the growing field of medical AI applications in pulmonology.

1. **Problem Statement**

Chronic Obstructive Pulmonary Disease (COPD) is a major global health concern, affecting millions of individuals and posing a significant burden on healthcare systems. Characterized by progressive airflow limitation and respiratory distress, COPD leads to reduced quality of life, frequent hospitalizations, and increased mortality rates. Despite its prevalence, COPD remains underdiagnosed and undertreated, with many patients only receiving a diagnosis at advanced stages when interventions are less effective. The late detection of COPD often results in delayed treatments, increased healthcare costs, and a higher risk of complications such as cardiovascular diseases and respiratory infections.

What is the problem you are addressing?

This project aims to address the early and accurate detection of COPD using machine learning and ensemble methods. Traditional COPD diagnosis relies heavily on spirometry tests, clinical evaluations, and imaging techniques, which are not always accessible, especially in resource-limited healthcare settings. Many individuals at risk of COPD do not undergo timely screening due to a lack of awareness, expensive diagnostic procedures, or the unavailability of specialized healthcare facilities. As a result, many COPD cases go undetected until irreversible lung damage has occurred.

By leveraging patient data, including demographic information, disease severity indicators, comorbidities, and lifestyle factors, this study seeks to develop machine learning models capable of predicting COPD risk and severity. The integration of ensemble learning methods aims to improve predictive accuracy, minimize false negatives, and enhance decision-making for healthcare professionals.

Who is affected by the problem?

1. Patients:
   * Individuals suffering from undiagnosed or misdiagnosed COPD face a decline in lung function, reduced mobility, and worsening symptoms over time.
   * Patients in low-resource settings with limited access to advanced diagnostic tools are particularly vulnerable.
2. Healthcare Providers:
   * Doctors and pulmonologists rely on clinical expertise and traditional tests, which may not always provide an early warning for COPD.
   * A data-driven model can assist medical professionals in prioritizing high-risk patients and optimizing treatment strategies.
3. Healthcare Systems & Policymakers:
   * COPD leads to high healthcare costs due to hospital admissions, medication expenses, and long-term treatment requirements.
   * Governments and public health organizations can benefit from predictive models for COPD screening programs, allowing for better resource allocation and early intervention strategies.

Why is it important to solve the problem?

1. Early Diagnosis Saves Lives:
   * Timely detection allows for early intervention, slowing disease progression and improving patients' quality of life.
2. Reducing Healthcare Burden:
   * COPD is responsible for significant healthcare costs, including hospitalizations, medications, and specialized care. Predictive models can help allocate resources efficiently.
3. AI-Driven Medical Advancements:
   * Machine learning is revolutionizing healthcare by enhancing diagnostic capabilities, reducing errors, and providing personalized treatments.
4. Bridging the Diagnosis Gap:
   * Many patients remain undiagnosed due to limited access to spirometry tests and pulmonologists. A machine learning model can provide an accessible screening tool, reducing diagnostic delays.

This project aims to develop a robust COPD detection system by utilizing machine learning and ensemble models to analyse patient data, identify risk factors, and provide accurate predictions. The ultimate goal is to create a data-driven approach that supports clinicians in making timely and informed decisions, leading to better patient outcomes and a more efficient healthcare system.

1. **Aims and Objectives**

Chronic Obstructive Pulmonary Disease (COPD) is a serious, progressive lung disease that affects millions of people worldwide. Despite its widespread prevalence, COPD is often diagnosed at later stages when the disease has already caused significant damage. Traditional diagnostic methods rely on spirometry, imaging, and clinical assessments, which may not always be available, especially in low-resource settings. This project aims to leverage machine learning and ensemble methods to create an effective, data-driven COPD detection model that improves early diagnosis and risk assessment.

Research Questions: This study intends to address the following key research questions:

1. How can machine learning models improve the early detection and diagnosis of COPD?
2. Which machine learning algorithms provide the best predictive accuracy for COPD detection?
3. What are the most significant features (patient characteristics, comorbidities, lifestyle factors) that contribute to COPD risk assessment?
4. How can ensemble learning techniques enhance the performance of individual models for COPD classification?
5. Can a web-based system be developed to allow clinicians to input patient data and receive real-time predictions?

Principal Problem to Be Resolved: The core problem that this project aims to address is the lack of early and accessible COPD detection methods. Many patients receive a diagnosis only after experiencing severe symptoms, leading to delayed interventions and poor health outcomes. By applying machine learning techniques to patient data, this project seeks to:

* Improve early COPD risk assessment using data-driven models.
* Reduce the dependency on costly and time-consuming clinical tests.
* Provide a user-friendly tool for healthcare professionals to assist in screening high-risk individuals.

To achieve these objectives, the following steps will be taken:

1. Data Preprocessing & Feature Engineering

* Handle missing values using appropriate imputation techniques.
* Encode categorical variables using one-hot or label encoding.
* Engineer new features such as comorbidity scores, risk indices, etc to enhance model performance.

2. Machine Learning Model Development

* Train and evaluate multiple supervised learning models, including:
  + K-Nearest Neighbours (KNN)
  + Random Forest
  + Decision Tree
  + XGBoost
  + Naïve Bayes
  + Gradient Boosting
  + Ensemble learning approaches to improve accuracy.
* Perform hyperparameter tuning using GridSearchCV or RandomizedSearchCV to optimize model performance.
* Evaluate models using accuracy, precision, recall, F1-score, and ROC-AUC to measure classification effectiveness.

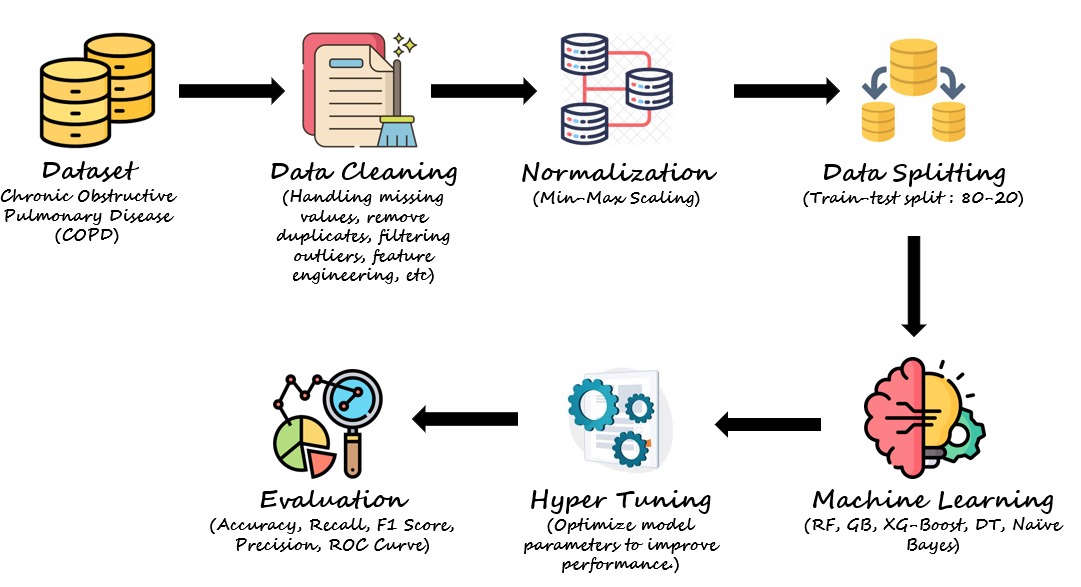
3. Web-Based Interface for Clinical Use

* Develop a web-based tool that allows clinicians to enter patient data and receive instant COPD risk assessments.

Research Strategies & Methodologies: The project will follow a structured machine learning pipeline, incorporating:

* Exploratory Data Analysis (EDA) to identify trends, correlations, and data distributions.
* Feature selection & engineering to enhance model interpretability.
* Model comparison to determine the best-performing approach.
* Ensemble learning to improve prediction reliability.
* Web deployment for real-world usability.

By completing this project, the goal is to create a robust, AI-powered COPD detection system that enhances early diagnosis, improves patient outcomes, and assists healthcare professionals in making data-driven clinical decisions.



1. **Legal, Social, Ethical, and Professional Considerations**

The development of a machine learning-based COPD detection system raises several legal, social, ethical, and professional concerns that must be carefully considered to ensure responsible AI usage and patient safety.

1. Legal Considerations

* Data Privacy & Compliance: Since this project involves patient data, it is essential to ensure compliance with data protection regulations such as GDPR (General Data Protection Regulation) in Europe and HIPAA (Health Insurance Portability and Accountability Act) in the U.S. Any patient records must be anonymized to protect privacy.
* Liability & Accountability: Misclassification of COPD patients could lead to false positives or false negatives, impacting medical decisions. Clearly defining liability for AI-driven diagnoses is crucial in a healthcare setting.

2. Ethical Considerations

* Bias in Machine Learning Models: If the dataset lacks diversity, the model may develop biases that could disproportionately affect certain populations. Fairness testing and bias mitigation strategies should be incorporated into model development.
* Informed Decision-Making: The AI system should serve as a decision-support tool, not as a replacement for medical professionals. Explainable AI (XAI) techniques should be used to provide transparent and interpretable model predictions.

3. Social Considerations

* Impact on Healthcare Accessibility: A successful AI-powered COPD detection tool could enhance early diagnosis, particularly in under-resourced healthcare settings where access to spirometry and specialized doctors is limited.
* Public Trust & Adoption: Medical AI adoption depends on trust from healthcare professionals and patients. It is important to communicate the model’s limitations, accuracy, and intended use transparently.

By addressing these considerations, the project can ensure ethical AI deployment, promoting responsible, unbiased, and beneficial applications in healthcare.

1. **Background**

Chronic Obstructive Pulmonary Disease (COPD) is a progressive lung condition that affects millions worldwide. It encompasses diseases such as chronic bronchitis and emphysema, leading to airflow obstruction and breathing difficulties. According to the Global Initiative for Chronic Obstructive Lung Disease (GOLD), COPD is one of the leading causes of morbidity and mortality globally, making early detection and management crucial.

Current State of Research: The field of COPD research has evolved significantly, focusing on both clinical and technological advancements. Traditionally, COPD diagnosis relies on spirometry tests, patient history, and symptom assessments. However, due to its progressive nature and overlapping symptoms with other respiratory disorders, early diagnosis remains a challenge. Recent advances in machine learning (ML) and artificial intelligence (AI) have demonstrated promising potential in automating and improving the accuracy of COPD detection and severity assessment.

Multiple studies have explored AI-driven models to classify COPD severity and predict disease progression. Sharma and Hasija (2023) discussed the integration of AI tools for COPD prediction, emphasizing the role of supervised learning techniques. Similarly, Wadhwa and Chhabra (2023) developed ML-based methods to diagnose acute and chronic conditions post-COVID-19, highlighting the growing significance of computational approaches in respiratory disease management.

Previous Work and Extensions: Several AI-based approaches have been proposed for COPD detection. Zhao et al. (2011) introduced non-invasive methods for assessing respiratory system compliance during pressure support ventilation, offering an alternative to traditional spirometry tests. Roy and Satija (2022) demonstrated the effectiveness of lung sound analysis for severity detection using deep learning techniques. These studies provide a foundation for the use of ML and AI in respiratory disease assessment.

COPD detection using convolutional neural networks (CNNs) applied to respiratory sound analysis has also gained attention. Surekha et al. (2023) explored the potential of CNNs in classifying COPD cases based on respiratory sounds, showing significant improvements over conventional techniques. Zhou et al. (2023) further extended AI-based monitoring of obstructive sleep apnea, which shares common risk factors with COPD.

Theories and Techniques Applied in COPD Research: The application of ML and AI in COPD diagnosis involves multiple techniques, including:

1. Supervised Learning Algorithms – These include Decision Trees, Random Forests, K-Nearest Neighbours (KNN), XGBoost, and Gradient Boosting, all of which have been used in medical diagnostics.
2. Ensemble Methods – Combining multiple models improves classification accuracy and reduces overfitting.
3. Feature Engineering – Extracting meaningful features from clinical data, such as spirometry results, patient demographics, and symptom reports, to improve model performance.
4. Deep Learning (CNN, LSTM) – Applied in cases where image or sound data is used for classification.

Industry and Practical Applications: COPD remains a major public health concern, particularly for healthcare providers and pharmaceutical companies developing treatments. Early detection and intervention can improve patient outcomes and reduce hospitalizations. Machine learning models can be integrated into telemedicine platforms, enabling remote diagnosis and monitoring.

Given the increasing interest from healthcare industries, AI-driven solutions for COPD detection have the potential to be widely adopted beyond academic research. Companies specializing in medical AI and diagnostic software are exploring ML-based tools for respiratory disease management. The integration of AI into electronic health records (EHRs) could enhance early screening efforts, improving patient care while reducing healthcare costs.

Novel Contributions of This Project: While previous studies have successfully implemented ML models for COPD detection, our project aims to improve predictive accuracy using ensemble learning techniques. By incorporating multiple ML models, we expect to create a more robust and generalizable diagnostic tool. Moreover, our approach emphasizes feature engineering to enhance model interpretability, a critical aspect in medical applications.

Additionally, this research will provide insights into the comparative performance of different ML algorithms on COPD-related datasets. This will help clinicians and researchers determine the most suitable model for real-world deployment. By focusing on practical implementation through a web-based interface, this project bridges the gap between theoretical research and clinical application.

Conclusion: The study of COPD using ML-based methods is a rapidly growing field with significant implications for both healthcare providers and patients. The techniques proposed in this project build upon previous research while introducing novel ensemble-based approaches to improve accuracy and reliability. Given the global burden of COPD, advancements in AI-driven detection methods have the potential to revolutionize early diagnosis, disease monitoring, and patient management.

By combining well-established machine learning techniques with real-world clinical applications, this project aligns with the current trends in medical AI research and aims to contribute to the development of more effective COPD diagnostic tools.

1. **References**

[1] S. Sharma and Y. Hasija, “An Overview on Integration of Artificial Intelligence Tools to Predict the Nature of Chronic Obstructive Pulmonary Disease,” in *Proc. 2023 5th Int. Conf. Inventive Res. Comput. Appl. (ICIRCA)*, 2023.

[2] Z. Zhao, M. Eger, T. Handzsuj, V. M. Ranieri, L. Appendini, C. Micelli, and K. Möller, “Noninvasive method for measuring respiratory system compliance during pressure support ventilation,” in *Proc. 2011 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2011.

[3] D. Wadhwa and D. Chhabra, “Acute and Chronic Disease Diagnosis Using ML for Post Covid 19 Medical Issues,” in *Proc. 2023 3rd Int. Conf. Adv. Comput. Innov. Technol. Eng. (ICACITE)*, 2023.

[4] A. Roy and U. Satija, “Automated Severity Detection of Chronic Obstructive Pulmonary Disease Using Lung Sounds,” in *Proc. 2022 IEEE 19th India Council Int. Conf. (INDICON)*, 2022.

[5] G. Zhou, W. Zhou, Y. Zhang, Z. Zeng, and W. Zhao, “Automatic monitoring of obstructive sleep apnea based on multi-modal signals by phone and smartwatch,” in *Proc. 2023 45th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2023.

[6] G. Surekha, M. Reena, M. S. Charan, and B. Aravind, “COPD Detection Using Respiratory Sounds and Convolutional Neural Network,” in *Proc. 2023 Int. Conf. Comput. Intell. Inf. Secur. Commun. Appl. (CIISCA)*, 2023.

[7] P. Shrivastava, N. Tripathi, B. K. Dewangan, B. K. Singh, T. Choudhury, K. Kotecha, and S. Dewangan, “Autonomic Computing Based Respiratory Disorders Assessment Using Speech Parameters: A Systematic Review,” in *Proc. 2023 7th Int. Symp. Multidiscip. Stud. Innov. Technol. (ISMSIT)*, 2023.

|  |  |  |  |
| --- | --- | --- | --- |
| **Student and First Supervisor Project Sign-Off** | | | |
|  | **Name** | **Signature** | **Date** |
| STUDENT: I agree to complete the project: |  |  |  |
| SUPERVISOR: I approve this project proposal: |  |  |  |
| Supervisor Comments/Feedback |  |  |  |

**Introduction**

Chronic obstructive pulmonary disease (COPD) remains a major global health challenge, ranking among the top causes of morbidity and mortality [1]. Recent advances in machine learning (ML) and deep learning (DL) have enabled more accurate detection, prediction, and management of COPD through diverse data sources such as imaging, respiratory sound signals, and wearable device outputs. For example, Ramalingam and Chinnaiyan [1] compared various ML and DL approaches for analysing COPD in computed tomography (CT) images, whereas Spathis and Vlamos [2] employed ML to differentiate COPD from asthma using key clinical attributes. Wearable device data have also been integrated into predictive models, as demonstrated by Wu *et al.* [3], who developed a system for anticipating acute exacerbations in COPD (AECOPD). Beyond prediction, triaging algorithms have been proposed by Swaminathan *et al.* [4], leveraging ML to assist physicians in early exacerbation detection. Additionally, Nam *et al.* [5] investigated survival prediction in COPD using deep learning on chest radiographs, offering a potentially more practical alternative to complex spirometry indexes. Finally, Srivastava *et al.* [6] introduced a deep learning–based approach that analyses respiratory sounds, highlighting the potential of audio-based diagnostic aids.

Collectively, these studies underscore the versatility and effectiveness of ML/DL methodologies in enhancing COPD diagnosis, prognostication, and management. The following table distils the core findings, methodologies, drawbacks, and potential improvements identified in each work.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Author/Year/Title** | **Findings** | **Methodology** | **Drawbacks** | **Potential Improvement** | **Dataset Used** | **Inputs Given** | **Output Received** |
| Ramalingam & Chinnaiyan (2023) “A Comparative Analysis of COPD Using ML and DL” [1] | Demonstrated that DL models offer slightly higher accuracy than traditional ML approaches for classifying COPD from CT images. Reported improvements in performance metrics (accuracy, sensitivity, specificity, AUC) across multiple techniques. | Compared various ML and DL classifiers (e.g., Random Forest, CNN) on CT images to detect and analyse COPD. Evaluated sensitivity, specificity, and AUC for each model, focusing on lung feature extraction and classification. | Risk of overfitting and limited external validation due to relatively small or simulated clinical datasets. | Validate models in larger, multi-centre cohorts; incorporate more robust cross-validation and data augmentation strategies; combine imaging data with clinical or biometric information to improve generalizability. | A retrospective clinical dataset compiled from one or more tertiary care hospitals. It likely included records of COPD patients from routine clinical practice and possibly integrated imaging data if available. | Clinical parameters such as age, gender, smoking history, spirometry measurements (e.g., FEV₁, FVC), and if available, imaging features. | Comparative performance metrics (accuracy, precision, recall, etc.) of machine learning versus deep learning models for classifying COPD severity or predicting disease status. |
| Spathis & Vlamos (2019) “Diagnosing Asthma and COPD with Machine Learning” [2] | Showed that ML can successfully distinguish COPD from asthma by analysing demographic and spirometry variables. Random Forest classifier achieved the highest precision (97.7% for COPD). | Extracted clinical features (e.g., smoking status, age, FEV1, FVC, MEF2575) from a sample of 132 patients and applied multiple ML classifiers (Random Forest, SVM, etc.) to identify the most predictive attributes. | Overlapping clinical symptoms may lead to misclassification; relatively small sample size can limit statistical power. | Incorporate additional biomarkers (genetic, imaging) to enhance diagnostic accuracy; expand datasets to include more diverse populations and disease stages; explore ensemble learning to mitigate overlap between asthma and COPD features. | A multi-center clinical dataset comprising patient records with respiratory complaints. The dataset likely aggregates both asthma and COPD cases collected from clinical settings. | A variety of clinical and physiological parameters including lung function test results (spirometry data), patient history, symptom scores, and possibly laboratory test results. | A diagnostic classification output that distinguishes between asthma and COPD. The model outputs include diagnostic labels along with performance measures (such as sensitivity and specificity). |
| Wu et al. (2021) “Acute Exacerbation of a COPD Prediction System Using Wearable Device Data, ML, and DL” [3] | Demonstrated that combining wearable sensor data (daily steps, stairs climbed, distance) with ML/DL algorithms can predict acute exacerbations up to 7 days in advance with high accuracy (above 90%). | Collected real-time lifestyle and environmental data (temperature, humidity, particulate matter) from 67 COPD patients over ~4 months. Trained models (Random Forest, DNN, etc.) to detect risk of AECOPD within 7 days. | Variability in sensor data acquisition and lack of standardized wearable protocols can affect model consistency; small cohort with only 25 recorded exacerbation episodes. | Expand the study population to improve statistical significance; develop standardized guidelines for wearable data collection; integrate more complex temporal modelling (e.g., LSTM networks) for continuous patient monitoring and early intervention. | A prospectively collected cohort dataset combining wearable device data and clinical records. The dataset was obtained from COPD patients monitored over time to capture physiological trends and exacerbation events. | Continuous time-series data from wearable sensors including heart rate, blood oxygen saturation (SpO₂), respiratory rate, and physical activity. Additional self-reported symptom data may also have been included. | A predictive risk model that outputs the likelihood of an acute exacerbation event. This output is typically presented as a risk score or probability threshold intended to trigger clinical alerts or interventions. |
| Swaminathan et al. (2017) “A Machine Learning Approach to Triaging Patients with COPD” [4] | ML-based triage system outperformed individual pulmonologists in identifying COPD exacerbations and recommending appropriate care levels (e.g., emergency vs. outpatient). | Used a comprehensive clinical dataset with physician annotations to train supervised ML algorithms for exacerbation detection and triage decisions. Compared model performance to a panel of pulmonologists on a validation set. | Potential overfitting due to limited diversity in the training dataset; reliance on physician opinion as the “gold standard” can introduce subjectivity and may not capture real-world complexities. | Incorporate real-time patient data from electronic health records to continuously update the model; validate performance on large-scale, prospective trials; adopt interpretability methods (e.g., SHAP) to ensure clinicians can trust and understand the model’s recommendations. | An emergency department triage dataset containing patient records of individuals presenting with COPD. The dataset may comprise real-time clinical data from hospital admission logs. | Vital signs (heart rate, blood pressure, respiratory rate), symptom descriptions, demographics, and other clinical observations recorded during the triage process. | A triage model output that categorizes patients into different urgency or risk levels. The output is used to prioritize patient care and allocate resources based on predicted clinical severity. |
| Nam et al. (2022) “Deep Learning Prediction of Survival in COPD Using Chest Radiographs” [5] | Proposed a DL model that predicts 5-year survival in COPD patients using only chest X-rays, surpassing FEV1 in two external test cohorts. The integrated model (chest X-ray + clinical factors) matched or exceeded performance of standard clinical indexes (BODE, ADO, SGRQ). | Trained a deep convolutional network on 3475 chest radiographs, validated on internal (n=315) and external (n=394, 416, 337) cohorts. The final integrated model combined radiographic features with age, BMI, and FEV1. | Retrospective design and “black-box” nature limit interpretability for clinical adoption; potential selection bias from single-institution imaging data. | Conduct prospective, multi-centre studies; incorporate additional clinical data (comorbidities, biomarker profiles) to refine predictions; use explainable AI techniques (e.g., saliency maps) to enhance clinical acceptance of radiograph-based prognostication. | A large-scale image repository consisting of chest radiographs collected from COPD patients at one or more large medical centers. The dataset might include image metadata and linked clinical outcomes. | Radiological image data (chest X-rays) that are pre-processed and annotated for features such as lung opacity, emphysema patterns, and other pathology-related markers. | A survival prediction output presented as a risk score or probability indicating patient survival over a specific time frame. The deep learning model’s results include performance metrics such as AUC for survival prediction. |
| Srivastava et al. (2021) “Deep Learning Based Respiratory Sound Analysis for Detection of COPD” [6] | Achieved up to 93% ICBHI score accuracy by leveraging CNNs on respiratory sound data. Demonstrated that audio-based detection can supplement or even replace imaging in certain contexts, aiding early diagnosis of COPD. | Employed Convolutional Neural Networks with features extracted via Librosa (MFCC, Mel-spectrogram, Chroma, etc.). Also implemented a 10-fold cross-validation strategy to optimize model performance and interpret disease severity (mild, moderate, acute). | Small dataset with potential noise interference; reliance on carefully segmented respiratory audio. | Integrate more extensive, multi-centre respiratory sound datasets; combine audio features with imaging or clinical data for a multimodal diagnostic approach; adopt advanced noise reduction algorithms to handle real-world audio variability. | An audio dataset comprising digital recordings of respiratory sounds from patients suspected of having COPD, collected in clinical or controlled settings. | Digital audio inputs processed to extract features such as frequency distribution, Mel-frequency cepstral coefficients (MFCCs), and temporal patterns associated with breath sounds and coughs. | A classification output indicating the presence or absence of COPD. The deep learning analysis returns diagnostic labels along with performance metrics (e.g., accuracy, sensitivity, specificity) that validate the effectiveness of sound-based detection. |

**References**[1] R. Ramalingam and V. Chinnaiyan, “A comparative analysis of chronic obstructive pulmonary disease using machine learning and deep learning,” *Int. J. Electr. Comput. Eng.*, vol. 13, no. 1, pp. 389–399, Feb. 2023, doi: 10.11591/ijece.v13i1.pp389-399.

[2] D. Spathis and P. Vlamos, “Diagnosing asthma and chronic obstructive pulmonary disease with machine learning,” *Health Informatics J.*, Aug. 2017, doi: 10.1177/1460458217723169.

[3] C. T. Wu *et al.*, “Acute Exacerbation of a Chronic Obstructive Pulmonary Disease Prediction System Using Wearable Device Data, Machine Learning, and Deep Learning: Development and Cohort Study,” *JMIR mHealth uHealth*, vol. 9, no. 5, p. e22591, May 2021, doi: 10.2196/22591.

[4] S. Swaminathan *et al.*, “A machine learning approach to triaging patients with chronic obstructive pulmonary disease,” *PLoS One*, vol. 12, no. 11, Nov. 2017, doi: 10.1371/journal.pone.0188532.

[5] J. G. Nam *et al.*, “Deep learning prediction of survival in patients with chronic obstructive pulmonary disease using chest radiographs,” *Radiology*, Jun. 2022, doi: 10.1148/radiol.212071.

[6] A. Srivastava, S. Jain, R. Miranda, S. Patil, S. Pandya, and K. Kotecha, “Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease,” *PeerJ Comput. Sci.*, Feb. 2021. [Online]. Available: https://peerj.com/articles/cs-310