# AI Approaches for Pulmonary Disease Forecast

Awakash Mishra
*Maharishi School of Engineering & Technology, Maharishi University of Information Technology*
*Uttar Pradesh, India*
trapty@muit.in

Saurav Ganguly
*Department of Electronics and Communication Engineering, Presidency University*
*Karnataka, India*
saurav.ganguly@presidencyuniversity.in

Sandeep Kumar C*Department of Genetics, School of Sciences, JAIN (Deemed-to-be University),*
*Karnataka, India*
kc.sandeep@jainuniversity.ac.in



***Abstract*—There is a significant risk to the healthcare system that respiratory diseases will become the leading cause of mortality globally by 2030. Recent advancements in artificial intelligence (AI) techniques intended to identify and categorize a range of respiratory ailments have drawn considerable interest from the scientific community. This comprehensive study aims to provide an overview of the most current machine learning and deep learning techniques especially developed for the identification of airway illnesses. It also forecasts future developments in this discipline by examining present problems and suggesting possible lines of inquiry. Our thorough analysis comprises 155 publications that describe various respiratory illnesses, highlighting the crucial role that machine learning methods play in forecasting certain ailments. As a result, this study outlines the challenges that now exist and the opportunities for future advancement while also highlighting the potential benefits of machine learning and artificial intelligence (AI) in enhancing the efficacy of apps for diagnosing respiratory illnesses.**

***Keyword*—Artificial Intelligence, Airway Disorders, Diagnosis, Machine Learning, Deep Learning, Healthcare, Prediction, Respiratory Illnesses, Systematic Review and Medical Research**

## I. INTRODUCTION

Airway diseases are generally regarded as a significant public health issue because they affect the vital airways that transport oxygen and various other substances from the mouth to the lungs [1]. People who suffer from these conditions often experience stiffness or obstruction of these breathing passages, which causes them to feel as if they are "trying breathing via a straw." Respiratory infections are a substantial cause of illness and mortality, affecting an astonishing 65 million people annually and killing 3 million of them. Therefore, airway diseases rank as the third leading cause of death worldwide.

there were 3.32 million fatalities related to chronic respiratory disorders in 1990 and 3.91 million in 2017. Despite a 2.41% annual decline in the average age of death for persistent respiratory conditions (from 2.28% to 2.55%), there remain significant geographical variations in deaths attributable to these illnesses between 195 different countries [2]. The regions with lowest Socio-demographic indicators also have the highest mortality rates. The major cause of COPD deaths in areas with low Socio-demographic index is particle pollution, which encompasses ailments like asthma and pneumonia.

These airway illnesses influence certain conditions as well [3]. For instance, 1.4 million TB-related fatalities were reported in 2019, there are an estimated 40,000 people living with cystic fibrosis in the United States, and there are an estimated 10 million tuberculosis cases that have been documented. On the other hand, lung cancer is the biggest cause of cancer-related mortality, accounting for almost 350 fatalities per day in the US alone.

Healthcare professionals and specialists throughout the globe have committed their efforts to predicting, detecting, and diagnosing these disorders due to the growing incidence and severity of these diseases [4]. They have used both traditional and AI-based methods, which are discussed in more detail in the following sections.

### *A. Usual Techniques for Identifying Airway Disorders*

The tiny airways, which in healthy lungs account for 10% to 25% of total airway resistance, assume a much greater significance in a variety of airway illnesses. These tiny airways, which lack cartilage and have an interior diameter of less than 2 mm, provide unique diagnostic difficulties. lists several conventional techniques used for identifying individuals with airway diseases, along with a short discussion of each technique's advantages and disadvantages.

### *B. AI Methods for Forecasting Airway Diseases*

The variety of applications for artificial intelligence in the healthcare industry, which have achieved impressive strides in this field [5]. After examining the drawbacks of conventional diagnostic methods for airway disorders, artificial intelligence has shown to be a useful and highly effective technique for automated picture classification. AI has shown to be capable of identifying a variety of airway disorders by using a variety of machine and deep learning techniques. The use of AI in detecting airway disorders using models with the ability to learn and make decisions based on massive input data is supported by growing evidence from researchers from a variety of fields. In contrast to traditional machine learning, these learning models build abstract representations of input, acquire data autonomously, and provide more accurate results.

Researchers from a variety of disciplines have been gathering data to support the use of AI in the diagnosis of airway illnesses [6]. A variety of machine and deep learning

979-8-3503-4438-7/23/$31.00 ©2023 IEEE

methods have been used in thorough analyses aiming at recognizing different airway abnormalities, as described in Section 5. The urgent need for better detection, diagnosis, and treatment of airway illnesses, which pose a serious threat to world health, has sparked a boom in the use of AI in medicine, notably in the field of pulmonology as shown in Fig. 1.

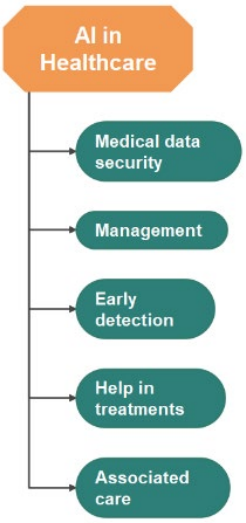


Fig. 1 Revolutionizing Healthcare: The Transformative Role of Artificial Intelligence in Predicting Pulmonary Diseases

Deep learning algorithms have a place in medical diagnostics. For instance, a deep learning system identified fibrotic lung illness with findings that were almost on par with a team of thoracic radiologists. For smokers with chronic obstructive pulmonary disease (COPD), a model was also effective in detecting and forecasting acute respiratory illness episodes and death [7]. It is crucial to remember that AI algorithms must be flexible enough to work with data of different quality levels and need a significant amount of well-structured data for development and validation. Additionally, practitioners must have a deep grasp of the potential applications of AI in the context of diseases like asthma and COPD. As a result, the integration of AI into medical practice has a bright future and will benefit both medical personnel and patients.

In the field of robotics, for instance, approximately 200,000 robots are installed every year to assist with treatments like prostate surgery, head and neck surgery, and administrative activities like updating patient data. On the other side, NLP is crucial for understanding, classifying, and analyzing unstructured clinical paperwork and for enabling patient-doctor interactions through chatbots powered by AI.

## II. LITERATURE REVIEW

Three subcategories of study have been done to predict different airway diseases, emphysema and pneumoconiosis [8]. In order to make comparative analysis easier, study conducted start by describing the goals of each research. Following an overall analysis, study conducted provide information on the datasets, techniques, results, conclusions, and limits in the form.

### *A. The Function of AI in COVID-19, Pulmonary Edoema, and Pulmonary Embolism Prediction*

Medical imaging is essential for making an accurate first diagnosis because it may spot certain pathological abnormalities that could otherwise go undiscovered. However, it has been difficult to compare and define the best detection algorithms since there aren't any publicly accessible datasets or benchmark studies. Islam et al.used several datasets to evaluate the efficacy of models for a variety of illnesses, including pulmonary edoema, in attempt to address this problem. To detecting abnormalities in chest X-rays, they used trained classifiers. A fundamental challenge in the processing of medical pictures, according to Liao et al, is the scarcity of ground truth labels.

### *B. The Role Artificial Intelligence in Mesothelioma and Lung Cancer Prediction*

By combining important prognostic signals, this framework intends to assist medical professionals and healthcare specialists in the early diagnosis and more successful treatment of malignant mesothelioma. Malignant Mesothelioma risk variables were looked at in research by Latif et al [9]. The dataset employed included both healthy people and Mesothelioma patients, with an emphasis on Mesothelioma patients' symptom identification.

### *C. The Function of AI in Cystic Fibrosis, Tuberculosis, and Asthma Prediction*

Multiple learning models were used by Awal et al.to study the characteristics defining asthma diagnosis and prediction. They developed a brand-new machine learning method for identifying asthma called BOMLA. Focused on research applying supervised approaches, such as classifier algorithms for data mining, to determine asthma control thresholds. By analyzing how time-series / time sequence variations in daily medical information influenced the determination of the asthma control level among patients, their major objective was to boost the accuracy of classifications techniques (Fig. 2).

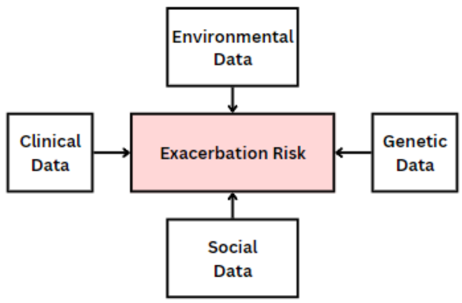


Fig. 2 Precision Medicine at Scale: Unveiling Innovative Approaches to Predict Asthma Exacerbation in Patients Using AI Techniques

*D. Emphysema and pneumoconiosis Prediction using AI*

The use of CT scans and tests for lung function are necessary for the early detection of emphysema, a condition that makes breathing difficult [10]. The challenges with certain diagnostic approaches have resulted in the rise of computer-assisted procedures. Deep learning networks were employed by Mondal et al. to increase detection accuracy for mechanized a pulmonary emphysema diagnosis. Bortsova et al. examined a weakly labelled approach comparable to multiple instances learning for assessing emphysema using labels and utilizing pre-existing label-related information. They developed a technique that learns intervals using a distinct loss function and a label portion (LPP) topology that is problem specific.

In the realm of deep learning, for instance, employed the Fleischner technique to identify copd via chest CT image analysis. Their major objective was to see if employing a system based on deep learning, identifying participant-level fibrosis patterns could foretell mortality and disability. To help medical professionals identify COP, or chronic ob. COPD, performed a detailed and meticulous study of acute respiratory audio data. The authors employed 10 divisions of K-fold cross-validation to enhance the effectiveness of pre-existing deep neural network networks for N samples.

## III. PROPOSED METHODOLOGY

The PRISMA standards, which include four crucial steps for choosing research papers, were followed in this systematic literature review.

*A. Identification*: In the first step, relevant records are found by searching via different sources and repositories. The systematic review procedure is built upon it.

*B. Screening*: In this step, publications are openly chosen after careful consideration of choices made during the systematic review. This step is essential for making sure the right research articles are included.

*C. Eligibility*: During this stage, all full-length papers are thoroughly evaluated in accordance with the predetermined standards and regulations. Papers that satisfy the predetermined criteria go forward to the next round.

*D. Inclusion*: In this last stage, the review paper includes the chosen articles that have complied with the requirements [11]. The requirements of PRISMA have been crucial in improving disclosure and documentation excellence of reviews and meta-analyses. Readers can now see how the authors chose the final article picking by learning how they used keywords, being published years, languages, and other factors.

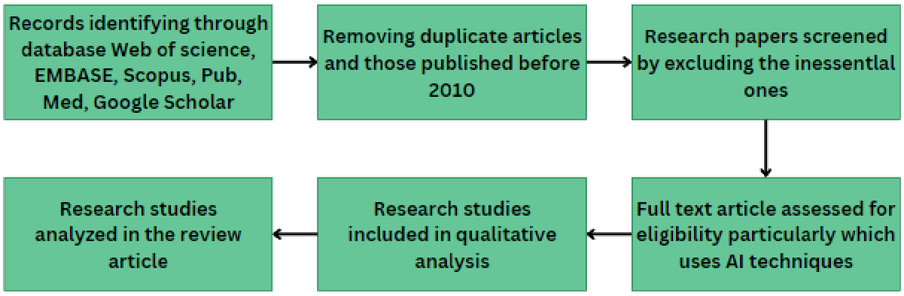


Fig. 3 A Systematic Approach to Research: Navigating the PRISMA Flowchart in the Study of AI Approaches for Pulmonary Disease Forecasting

## IV. FRAMEWORK FOR MULTIPLE AIRWAY DISEASE PREDICTION

This section provides a diagrammatic overview of the numerous stages involved in forecasting and categorizing airway disorders as shown in Fig. 3.

*A. Dataset*: The first and most important phase is gathering photos from multiple sources or datasets so that the system can successfully learn and identify them. A diversified and large picture collection must be gathered for reliable categorization. Multiple data sources, including X-rays, CT scans (Computed Tomography), histopathological pictures, and more, have been used to gather information on airway disorders.

*B.Pre-processing*:After data collection, thorough pre-processing is necessary to ensure the accuracy and reliability of the training data [12]. There might be issues with activity, blurriness, or concealed information in images from various sources. With such images, classification reliability may be compromised, thereby jeopardizing assessments of one's health. The quality of the photos is increased using pre-processing techniques including CLAHE for higher contrast, geometric adjustments, image filtering, and more. Using methods for data augmentation to expand the dataset without the need for additional data collection reduces the risk of overfitting. Data normalizing includes arranging the data in such a way as to guarantee that images have a constant distribution of statistical significance about their height of value of pixels. It also makes it possible for pixel intensity levels to be uniformly distributed across various images.

Since neural networks often need inputs of uniform size, image scaling is an important stage in the pre-processing process. Therefore, before being input into the convolutional neural network (CNN), all pictures are scaled to a standard dimension. After performing feature extraction on the training photos, the learning models may use these characteristics as inputs to more easily identify or predict the class of fresh images. Images utilized during the training phase are altered or changed because of this process.

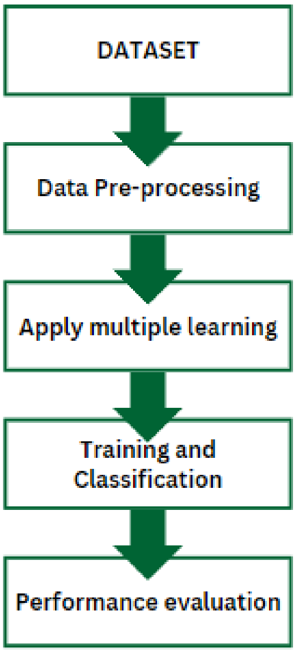


Fig. 4 Enhancing Disease Prediction: Exploring the Utility of Multiple Learning Models in Forecasting Airways Diseases through AI

*B. Learning Frameworks:* When modern systems use machine learning and deep learning techniques to enable computers to grasp tasks based on constantly changing datasets, they are referred to as artificially intelligent systems. Deep learning models may now be used to solve a variety of prediction issues thanks to recent improvements in learning algorithms and computing performance Convolutional Neural Networks (CNNs) stand out among them as a potent technique for identifying patterns in pictures, which makes them perfect for image classification jobs and a variety of other image-related applications (Fig. 4). CNNs are composed of neurons with trainable weights and biases that receive numerous inputs and are inspired by neural networks in the human brain. An activation function is used after these inputs are weighted, and the result is an output. Another effective strategy is transferring learning, which makes use of pretrained models like the Visual Geometry Group's (VGG16), VGG19, MobileNetV2, and ResNet50, which have been trained on big datasets and are considered industry standards in computer vision. These models may either be used directly for task prediction or incorporated into a model's training procedure. In addition, methods like ensemble learning and transfer learning can shorten training times, increase classification accuracy, and minimize modelling mistakes in the system (fig. 5).

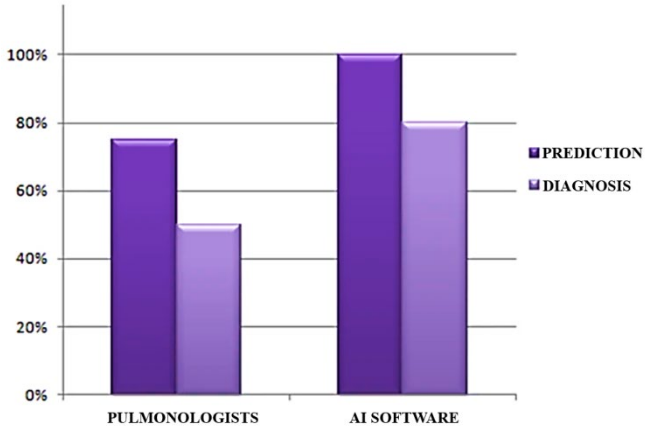


Fig. 5 Bridging the Gap: An In-depth Analysis of the Synergy between Pulmonologists and AI Techniques in Predicting Pulmonary Diseases

*C. Category:* Asthma, mesothelioma, TB, COVID-19, pulmonary edoema, emphysema, cystic fibrosis, pneumoconiosis, and lung cancer are among the classes of conditions that the trained model must classify pictures into at this crucial phase. The model should accurately categorise an image as a normal lung picture if it does not relate to any of these illnesses. Using certain measures, the model's accuracy in categorizing pictures is assessed in order to gauge the system's performance. These metrics provide a thorough evaluation of the model's performance in illness categorization.

## V. DISCUSSION

The successful introduction of AI into company operations depends on resolving these obstacles since they have a substantial negative influence on performance. Even while open-source datasets covering a variety of airway illnesses may not have the best quality data, they should nevertheless be considered for future research. There is a lot of promise in using publicly accessible data sets as organizational resources. Investigating the creation of generated data is a practical way to improve data integrity and confidentiality. Furthermore, to increase dataset size without requiring significant data gathering, data supplementation procedures must be put into practice. When there is a large amount of training information available, the strategic use of methods for transfer learning is beneficial.

Underfitting, when it comes to modelling mistakes, is the result of insufficient training of the model or generalization to new information sets, which produces inadequate results. Reducing modelling mistakes requires the deliberate use of strategies such as models that have been pretrained to avoid underestimating.

Improving the model's performance gradually is still a difficult task. Even if engineers are skilled in creating AI-based models, it's essential to thoroughly assess their effectiveness. A thorough evaluation and review process, enough data, and suitable approaches are all necessary for optimizing the efficiency of the model. Using learning transfer techniques may significantly improve AI-based predictive models over time.

## VI. FUTURE SCOPE

Even though the present emphasis of our study is on using AI to predict lung disorders, there is a lot of room for growth and improvement down the road. Refining artificial intelligence (AI) to grasp and anticipate the true causes of illness is a critical field of inquiry that will enable more accurate prognostic capacities.

One of the most important ways to improve the predicted accuracy and resilience of AI-driven lung disorder prediction models is to investigate the integration of diverse sources of info, including hereditary, image processing, and medical information. Furthermore, it becomes crucial to investigate clarified AI techniques to promote trust and comprehension between doctors and patients.

AI's application to personalized treatment plans and continuous surveillance is still an interesting development. Another direction for future research is to create artificial intelligence-driven support systems for decisions that customize therapies based on unique histories of patients and real-time data streams.

Future developments in this field are expected to take a more complex, AI-driven approach that might completely alter how pulmonary illnesses are predicted and treated. By exploring these possibilities, study conducted want to set the stage for a more advanced, precise, and customised method of predicting pulmonary illnesses, which will have a big influence on the results for patients and medical procedures.

## VII. CONCLUSION

Six online digital libraries were used to perform an extensive evaluation of 155 papers. This paper explores how machine learning and deep learning methods may be used to analyses airway problems. The study examines the research projects of several academics, emphasizing how machine learning and deep learning models have been crucial in the identification and classification of different airway diseases. The results of the systematic review are presented in tabular style, explaining the datasets, research methods, conclusions, and constraints that were found.

To find prospective directions for improving system performance in the future, an investigation has been undertaken into the challenges encountered by researchers in forecasting airway problems. The lasso algorithm, CNN, Decision Tree, GAP Net, and other algorithms were examined by researchers. The limits of these models to discriminate between illnesses or uncover abnormalities in the data were nonetheless shown by their results. Many of these models were found to have performance issues due to their use of small datasets, poor data preparation, and incomplete localization information for the final picture. Additionally, owing to modelling concerns like overfitting and underfitting, certain studies using methods like logistic regression and random forest showed reduced prediction accuracy.

Researchers also struggled with the use of methods including DFD-Net, Fuzzy Particle Swarm Optimization, CNN, Inception V3, ResNet, and DNN for the diagnosis of pulmonary emboli, and Canny Edge for the detection of TB. It is interesting that these investigations limited the generalizability of their results to a variety of airway-related problems by concentrating only on one or two airway illnesses for prediction.

Although artificial technology has many benefits, its drawbacks might limit its applicability, especially in the healthcare industry. There is, nevertheless, a lot of space for advancement in this area. Prioritizing concerns like selecting and improving models to get better results and integrating optimization approaches into networks to optimize results should be done. The appropriate categorization of data is a crucial component of using learning models, necessitating the adoption of approaches that improve the classification model's characteristics to provide accurate results. To allow its application to datasets for the quick identification of different airway disorders, deep learning methods' effectiveness should also be increased. Additionally, as they have the potential to improve airway disease detection models, traditional learning techniques should also be considered for future development.

## REFERENCES


[1] Koul A, Bawa RK, Kumar Y (2022) Artificial intelligence in medical image processing for airway diseases. In: Connected e-Health (pp 217–254). Springer, Cham. https://d oi.org/1 0.1007/978-3- 030-97929-4_ 10

[2] Li X, Cao X, Guo M, Xie M, Liu X (2020) Trends and risk factors of mortality and disability adjusted life years for chronic respiratory diseases from 1990 to 2017: systematic analysis for the Global Burden of Disease Study 2017. BMJ 368:10. https://doi. org/ 10. 1136/ bmj. m234

[3] Leitão Filho FS, Hang Chen H, Ngan DA, Tam A, Kirby M, Sin DD (2016) Current methods to diagnose small airway disease in patients with COPD. Expert Rev Respir Med 10(4):417–429. https:// doi. org/ 10. 1586/ 17476 348. 2016. 11554 55

[4] CF Foundation Estimates Increase in CF Population (2022) Cystic Fibrosis Foundation. https:// www. cff. org/ news/ 2022- 07/cf- found ation- estimates- increase- cf- population

[5] Chakaya J, Khan M, Ntoumi F, Aklillu E, Fatima R, Mwaba P, Zumla A (2021) Global tuberculosis report 2020: reflections on the global TB burden, treatment and prevention efforts. Int J Infect Dis 113:S7–S12. https://d oi.org/1 0.1016/j.i jid.2021.0 2.107

[6] Nascimento Maia P et al (2022) Correlation of digital flow peak with spirometry in children with and without asthma. J Asthma. https:// doi. org/ 10. 1080/ 02770 903. 2022. 20453 08

[7] Vaishya R, Javaid M, Khan IH, Haleem A (2020) Artificial intelligence (AI) applications for COVID-19 pandemic. Diab Metab Syndr 14(4):337–339. https://d oi.org/1 0.1016/j.d sx.2020.0 4.012

[8] Bakker JT, Klooster K, Bouwman J, Pelgrim GJ, Vliegenthart R, Slebos DJ (2022) Evaluation of spirometry-gated computed

tomography to measure lung volumes in emphysema patients. ERJ Open Res 8(1):10. https:// doi. org/ 10. 1183/ 23120 541.00492- 2021

[9] Si X, Xi JS, Talaat M, Donepudi R, Su WC, Xi J (2022) Evaluation of impulse oscillometry in respiratory airway casts with varying obstruction phenotypes, locations, and complexities. J Respir 2(1):44–58. https:// doi. org/ 10. 3390/ jor20 10004

[10] Ahmed Z, Mohamed K, Zeeshan S, Dong X (2020) Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. Database. https:// doi. org/ 10. 1093/ database/ baaa01 0

[11] Kanna, R. Kishore, V. Subha Ramya, Asraa Ahmed Khafel, Kadim A. Jabbar, Mustafa Al-Tahee, and Raed Khalid. "Cognitive Disability Prediction & Analysis using Machine Learning Application." In 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp. 1359-1364. IEEE, 2023.

[12] Dransfield MT, Kunisaki KM, Strand MJ, Anzueto A, Bhatt SP, Bowler RP, Make BJ (2017) Acute exacerbations and lung function loss in smokers with and without chronic obstructive pulmonary disease. Am J Respir Crit Care Med 195(3):324–330. https:// doi. org/ 10. 1164/ rccm. 201605- 1014OC