REVIEW ARTICLE

Artificial Intelligence Techniques to Predict the Airway Disorders Illness: A Systematic Review

Apeksha Koul[1] · Rajesh K. Bawa[2] · Yogesh Kumar[3]

Abstract

Airway disease is a major healthcare issue that causes at least 3 million fatalities every year. It is also considered one of the foremost causes of death all around the globe by 2030. Numerous studies have been undertaken to demonstrate the latest advances in artificial intelligence algorithms to assist in identifying and classifying these diseases. This comprehensive review aims to summarise the state-of-the-art machine and deep learning-based systems for detecting airway disorders, envisage the trends of the recent work in this domain, and analyze the difficulties and potential future paths. This systematic literature review includes the study of one hundred fifty-five articles on airway diseases such as cystic fibrosis, emphysema, lung cancer, Mesothelioma, covid-19, pneumoconiosis, asthma, pulmonary edema, tuberculosis, pulmonary embolism as well as highlights the automated learning techniques to predict them. The study concludes with a discussion and challenges about expanding the efficiency and machine and deep learning-assisted airway disease detection applications.

1 Introduction

The airway diseases have been called the worldwide source of sickness, as they damage the air tubes that carry oxygen and other gases into and out of the lungs. These diseases generally induce a constriction or obstruction of the airways. People with airway illnesses often remark that they are "trying to breathe out through a straw." Airway disease is a fundamental cause of mortality and disability as it touches at least 65 million people and kills 3 million individuals yearly. Hence it is ranked as the 3rd most prevalent reason for death around the globe overall [1]. As shown in Fig. 1, chronic respiratory disorders fatalities rose from 3.32 million in 1990 to 3.91 million in 2017. Annually, the age-standardized mortality rate for chronic respiratory diseases fell by 2.41% (2.28% to 2.55%), but in 195 nations, the annual fatality rates attributable to chronic respiratory illnesses have shown a great variation. The highest mortality was seen in regions with a low Socio-demographic index [2]. Particulate matter pollution, such as pneumoconiosis and asthma, has been the leading cause of COPD fatalities in areas with a low Socio-demographic index [3]. In addition, one of the studies also states that 40,000 people have been suffering from cystic fibrosis in the US [4], 10 million tuberculosis cases have been estimated, and 1.4 million are reported TB deaths in 2019 [5]. In contrast, lung cancer has been termed the leading cause of cancer for approx 350 deaths each day in the US [6].

After considering these numbers, several physicians, clinicians, and medical experts from all over the world have tried their best possible way to predict, detect and diagnose these dreadful diseases, for which they used conventional and AI-based approaches that are described in the upcoming subsections.

✉ Rajesh K. Bawa
rajesh.k.bawa@gmail.com

Apeksha Koul
apekshakoulo9@gmail.com

Yogesh Kumar
yogesh.arora10744@gmail.com;
yogesh.kumar@sot.pdpu.ac.in

1 Department of Computer Science and Engineering, Punjabi University, Patiala, Punjab, India

2 Department of Computer Science, Punjabi University, Patiala, Punjab, India

3 Department of Computer Science and Engineering, School of Technology, Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Fig. 1 Global impact of airway diseases [2]

Fig. 1 Global impact of airway diseases [2]
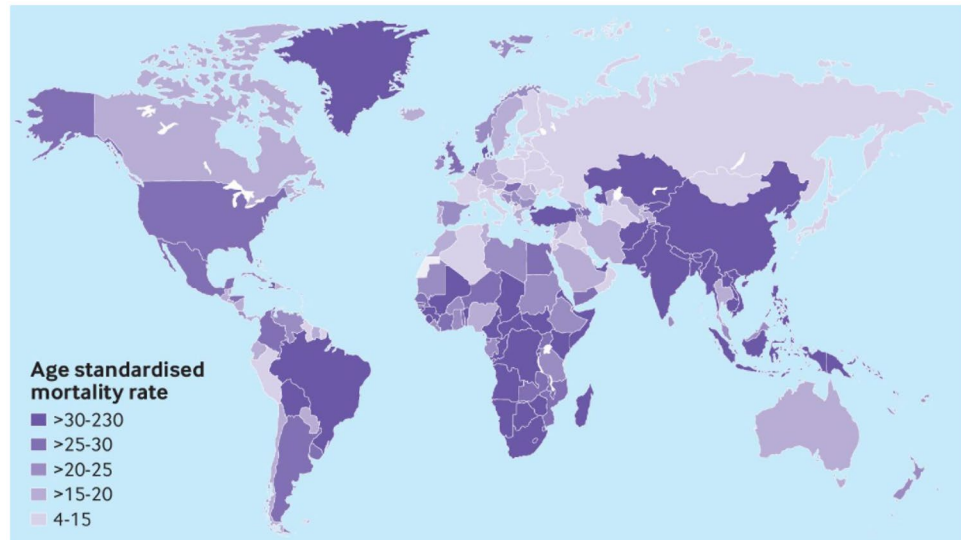
1.1 Conventional Methods to Diagnose Airway Diseases

Small airways account for between 10 and 25% of total airway resistance in healthy lungs, with their contribution to total airway resistance increasing significantly in multiple airway disorders. It is because the internal diameter of the tiny airways is less than 2 mm, and there is no cartilage [5]. Hence in Table 1, a few traditional ways such as spirometry, body plethysmography, impulse oscillometry, and washout tests used for diagnosing airway disorders in patients are shown, along with their brief description and drawbacks.

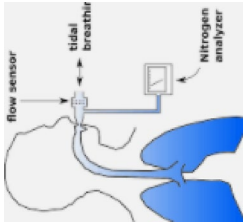## 1.2 AI Techniques to Predict Airway Diseases

Artificial intelligence has shown tremendous growth in the health sector in today's era via its various applications, as shown in Fig. 2. After having a brief idea about the drawbacks of traditional approaches to diagnosing airway diseases, artificial intelligence has demonstrated its efficiency and excellent performance in automatic image categorization using multiple machine and deep learning algorithms to detect numerous airway illnesses [10]. Researchers from various disciplines are steadily amassing evidence to support the use of AI in diagnosing airway sickness using models that can learn and make choices utilizing massive input data [11]. The fundamental rationale for using learning models is that these approaches learn by constructing a more abstract representation of input (unlike classical machine learning) in which the model collects information automatically and produces more accurate results [12].
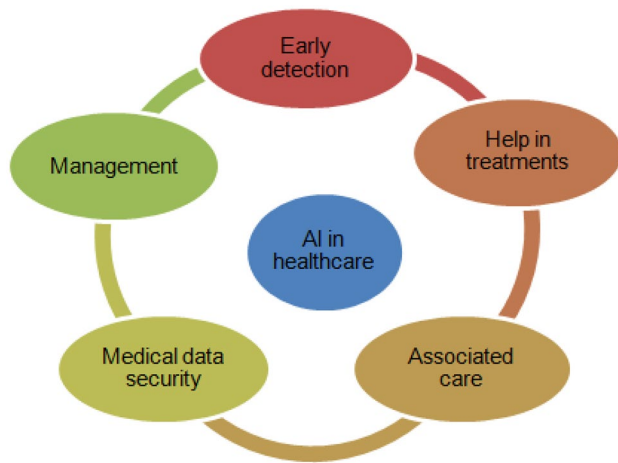
Researchers from many fields are progressively accumulating data to support the usage of AI to diagnose airway illness. Researchers have applied various algorithms of the machine and deep learning, as discussed in Sect. 5, to conduct an extensive assessment to detect multiple airway diseases. A corresponding spike has been seen in the medical applications of AI, particularly in pulmonology. Applications of AI to the global problem of airway illnesses can meet the most priority-based requirements highlighted for accurately detecting, diagnosing, and providing the best treatments [13]. A deep learning system generated results nearly equal to a group of thoracic radiologists identifying fibrotic lung disease in research [14]. Another study [15] found that a neural network developed by Google scientists was just as good as radiologists in detecting cancerous lung nodules. A similar model [16] could identify and predict acute respiratory illness episodes and death in smokers along with chronic obstructive pulmonary disease (COPD). Large amounts of well-structured data are required to create and validate AI algorithms, and the algorithms must operate with data of varying quality [17]. Clinicians must grasp how AI works in the context of multiple disorders such as asthma and chronic obstructive pulmonary disease. As a result, it would be interesting to witness the pros of AI (Artificial Intelligence) developed for doctors and patients due to its usage in medical practice in the future [18]. Other AI tools have also been used tremendously in medical science, such as robotics, natural language processing, expert systems, etc. In robotics, more than 2 lakh robots have been installed annually and are used to perform prostate surgery, head-neck surgery, update patient records, etc. Likewise, natural language processing (NLP) helps to analyze, understand and classify the unstructured format of clinical documentation and the interaction of patients with doctors via bots [19].

Thusly, to fully understand the role of techniques using artificial intelligence concepts to predict and diagnose

Table 1 Traditional approaches to diagnose airway diseases

Technique	Image	Description	Drawbacks
Spirometry [7]		The most common pulmonary function test is spirometry. It assesses lung function, precisely the amount of air that can be inhaled and expelled and the rate at which it can be done	Patient feels dizziness and shortness of breath for a moment after the test
Body plethysmography [8]		A pulmonary (lung-related) function test that evaluates how much air is in your lungs when you take a deep breath is called body plethysmography. It also determines how much air is left in your lungs after exhaling as much as possible	Technically demanding and time consuming
Impulse oscillometry [9]		A new technique for measuring airway resistance and reactance is the impulse oscillation system (IOS). It's a sort of forced oscillation in which oscillating sound waves of various frequencies, typically 5 and 20 Hz, are conveyed along the bronchial tree	The impulse used in the test is forceful which causes slight change in lung mechanism
Washout tests [10]		Nitrogen washout (also known as Fowler's method) is a test that measures anatomic dead space in the lungs and other factors related to airway closure throughout a breathing cycle	It consumes a lot of time and also fails to estimate the lung areas which are poorly ventilated

Fig. 2 Role of AI in healthcare

2 Materials and Methods

Full-text archives in the English language of six different publication databases have been searched between 2010 and 2022 such as (ScienceDirect (https://www.sciencedir ect.com), Google Scholar (https://scholar.google.co.in), Scopus (https://www.scopus.com), Web of Science (http://isiwebofknowledge.com), PubMed (http://www.ncbi.nlm.nih.gov/pubmed), EMBASE (https://www.embase.com). The paper has been searched using the keywords "airway diseases", "cystic fibrosis"," pneumoconiosis"," pulmonary embolism", "pulmonary edema", "Mesothelioma", "lung cancer ", "tuberculosis", "asthma ", "covid-19"," emphysema", "machine learning" "artificial intelligence," "deep learning," as well as combinations of these keywords. In addition to this, the articles have also been chosen based on inclusion and exclusion parameters (Table 2) which are based on various metrics such as time duration, scholarly articles from which the paper can be accessed, research to highlight the problem, a comparator for analyzing and comparing the researchers' work, methodologies to demonstrate the strategies that had been used in their articles, and finally research design to analyze the results.

In this systematic literature review, PRISMA (Preferred reporting items for systematic Reviews and Meta-Analyses) guidelines have been applied in which four phases have been used to select the research papers (Fig. 3), i.e., *Identification-* in which the identification of records is carried out by accessing various repositories, *Screening-* in which the papers are selected transparently by assaying the decisions that are made at different stages of the systematic review, *Eligibility-* in which all full-length articles are evaluated [20] and finally *Included-* in which the final selected articles are included to write the review paper. PRISMA is chosen as it helps improve the reporting and transparency of systematic reviews and meta-analyses. It is also useful for the readers to understand how the authors have filtered out the selected papers by using keywords, year of publication, language, etc., to frame any article.

airway diseases, a systematic literature review (SLR) has been conducted based on the AI learning models to predict such illnesses. The study is structured into various sections. Initially, a brief description of airway diseases and their global impact has been briefed in Sect. 1. This section also includes the traditional approaches as well as the role of AI in diagnosing airway diseases. Section 2 presents the qualities of the paper used in the study, while Sect. 3 presents the framework to predict airway diseases using AI-based techniques. Section 4 provides a comparison of various methods based on their dataset, applied algorithms with their results in terms of different metrics such as precision, accuracy, F1-score, etc., and the limitations of multiple airway diseases such as Mesothelioma, cystic fibrosis, emphysema, pulmonary edema, pneumoconiosis, lung cancer, pulmonary embolism, tuberculosis, asthma, and covid-19. The discussion segment, which answers the research questions from Sect. 2, is handled in Sect. 5. Finally, Sect. 6 concludes the study that assists researchers in determining the optimal technique for detecting disorders and the future scope.

Table 2 Inclusion standards and exclusion standards

| S. no.	Attributes	Inclusion standards	Exclusion standards
1	Duration	Research work that had been carried out between 2010 and 2022	Published articles before 2010
2	Exploration	Research work concentrating on (a) the findings, (b) the benchmark dataset, and (c) the research goal	Research work that focus on other diseases and not on airway diseases
3	Comparability	Research studies aims at the prediction of airway diseases	Research studies that work on other than airway diseases
4	Techniques	Research articles that mostly focus on machine and deep learning methods including few traditional ones	Research articles that apply the methods other than machine and deep learning models
5	Research design	Original articles that comprise of experimental results	Case studies, Language other than English, Patents
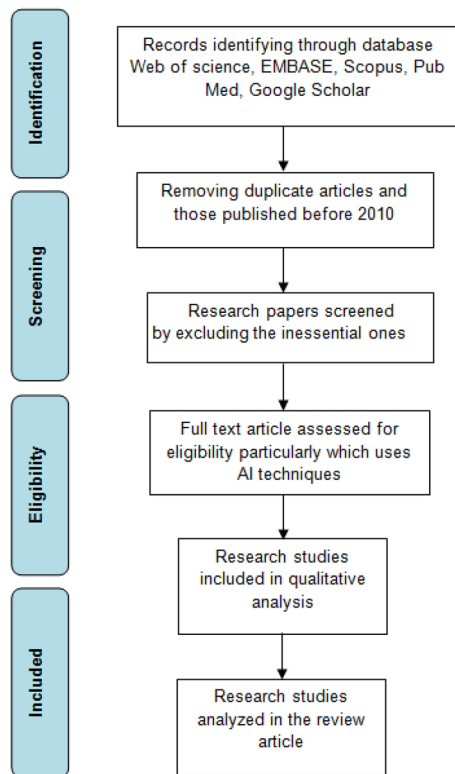
Fig. 3 PRISMA flow chart

As far as this paper is concerned, for a better understanding of the state of research on machine and deep learning in airway disease detection, peer-reviewed papers stated that these algorithms had played a vital role in predicting such disorders. In addition to this, a few research questions have also been framed that were investigated in the study:

RQ 1: Year wise analysis of predicting multiple airway diseases using AI based techniques.

RQ 2: How doctors are being helped by deep and machine learning techniques in detecting the airway diseases?

RQ 3: Which ML and DL techniques are broadly applied to predict airway diseases?

RQ 4: Name the characteristics that manipulate the quality of prediction models based on deep and machine learning?

3 Framework to Predict Multiple Airway Diseases

In this section, various phases to predict and classify airway diseases have been mentioned and diagrammatically shown in Fig. 4.

Dataset: The foremost step is to gather the images from various repositories or datasets so that the system can classify them by learning them. It is essential to feed the system with many images for better classification. The data for predicting airway diseases like pulmonary embolism, emphysema, tuberculosis, pulmonary edema, pneumoconiosis, cystic fibrosis, asthma, covid 19, and lung cancer has been collected from multiple data sources X-ray, CTs-can (Computed Tomography), histopathology image, etc. Table 3 shows the detailed description of the dataset for all the respective airway diseases.
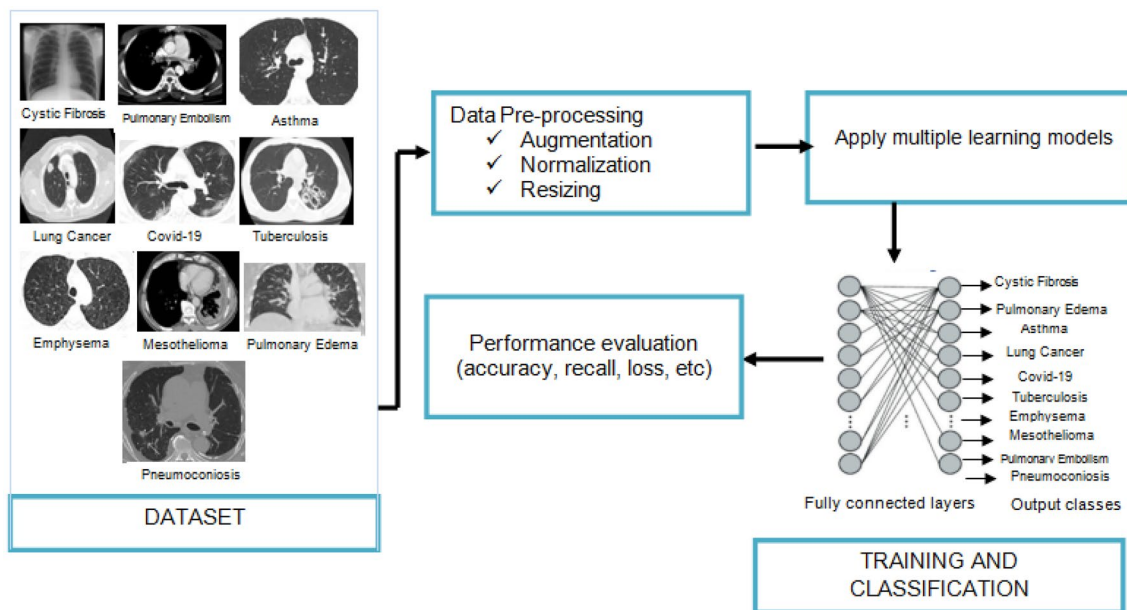

Fig. 4 Predicting airway diseases using multiple learning models

Table 3 Dataset of multiple airway diseases

References	Diseases	Dataset name	Description	URLs
[21]	Cystic fibrosis (CF)	Cystic fibrosis data	The dataset contains one file and 25 columns	https://www.kaggle.com/ukveteran/cystic-fibrosis-data
[22]		Cystic fibrosis registry in the United Kingdom	The dataset contains the geographical information about the cystic fibrosis patients	https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry
[23]		European cystic fibrosis registry	The database includes data from more than 49,000 CF people from 38 countries, and from 2008 to 2018	https://www.ecfs.eu/projects/ecfs-patient-registry/intro
[24]		Cystic fibrosis dataset	BioGPS has 10 datasets of cystic fibrosis on 5 species	http://biogps.org/dataset/tag/cystic%20fibrosis/
[25]		Cystic fibrosis data and statistics	State of residency, weight, height, sexuality, race, respiratory function test results, pancreatic enzyme usage, duration of hospitalizations, home IVs, and CF-related problems are among the information collected	https://www.health.ny.gov/statistics/diseases/chronic/cystic_fibrosis/
[26]	Pulmonary embolism	RNSA STR (pulmonary embolism detection)	The dataset includes the training and testing images, which each include 17 data fields related to PE	https://www.kaggle.com/c/rsna-str-pulmonary-embolism-detection
[27]		Pulmonary embolism in CT images	The dataset comprises of CT scan images taken from 35 different individuals	https://www.kaggle.com/andrewmvd/pulmonary-embolism-in-ct-images
[28]		Pulmonary embolism_codelist	The ICD-10 codes for pulmonary embolism diagnosis are included in the dataset	https://datacompass.lshtm.ac.uk/id/eprint/734/
[29]		Chinese Clinical Trials Registry Center	The dataset, which includes both APE and non-APE, was chosen retrospectively based on medical diagnosis	http://www.chictr.org/en/
[30]		FUMPE dataset	The dataset has 35 patients suffering from pulmonary embolism and fall within the age range of 24 to 82	https://figshare.com/authors/Moj-taba_Masoudi/5215238
[31]	Asthma	Informatica	The data set covered all visits by the asthma patient cohort within Intermountain Healthcare between 2005 and 2018	https://www.informatica.com/in/about-us/customers/customer-success-stories/intermountain-healthcare.html
[32]		Taiwan National Health Insurance Research Database	The dataset includes 1,000,000 samples randomly from the Registry of Beneficiaries (ID) in 2010, which contained around 27.38 million people	https://nhird.nhri.org.tw/en/
[33]		Data.world	A repository containing 10 asthma datasets	https://data.world/datasets/asthma
[34]		CHIS open data	This dataset contains the estimated %age of Californians with asthma (asthma prevalence)	https://data.chhs.ca.gov/dataset/01f456c3-db34-44f2-a52c-6811bef8ba6d/resource/13d6472c-1b35-4d4e-9c66-e643941de459/download/current-asthma-prev]
[35]		Asthma dataset	BioGPS has 24 datasets of asthma on 5 species	http://biogps.org/dataset/tag/asthma/

Table 3 (continued)

References	Diseases	Dataset name	Description	URLs
[36]	Lung Cancer	LIDC-IDRI	The dataset has computed tomography (CT) scans of lung cancer having marked-up annotated lesions	https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI
[37]		Lung cancer dataset	The data described pathological lung cancers in its 3 types	https://archive.ics.uci.edu/ml/datasets/lung+cancer
[38]		Lung cancer dataset	The dataset has an information about the diagnosis of each lung cancer at every trial	https://cdas.cancer.gov/datasets/nlst/
[39]		Chest X-ray 8	The dataset comprises of 108,948 frontal-view X-ray images of 32,717	https://paperswithcode.com/dataset/chestx-ray8
[40]		Dataset for Lung cancer diagnosis	CT and PET-CT DICOM images of lung cancer is in the dataset. It also has an XML Annotation files that locate the position of tumor with bounding boxes	https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=70224216
[41]	Covid-19	Novel coronavirus 2019 dataset	This dataset contains daily statistics on the number of patients impacted, fatalities, and recovery from the 2019 new coronavirus	https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset
[42]		Covid-19 dataset	The repository contains the data of COVID-2019 from January 22, 2020, to April 1, 2020,	https://coronavirus.jhu.edu/map.html
[43]		Covid-chest X-ray dataset	The details of the patients who are positive or suspected of Covid19 are in this public open dataset	https://github.com/ieee8023/covid-chestxray-dataset
[44]		Covid-19 radiography database	There are 3616 COVID-19 positive instances in the database, as well as 10,192 normal, 6012 non-COVID lung infection, and 1345 viral pneumonia photos	https://www.kaggle.com/tawsifurrahman/covid19-radiography-database
[45]	Tuberculosis	Shenzhen dataset	662 frontal CXR images are available in the dataset in which 335 are TB positive and 327 TB negative	https://www.kaggle.com/raddar/tuberculosis-chest-xrays-shenzhen
[46]		Tuberculosis Chest X-ray dataset	The database contains 700 Tuberculosis (TB) positive chest X-ray images as well as 2800 Normal images	https://www.kaggle.com/tawsifurrahman/tuberculosis-tb-chest-xray-dataset
[47]		Montgomery County X-ray Set	This collection has 138 x rays posterior-anterior side, out of which normal are 80 and abnormal are 58 that shows TB signs	https://lhncbc.nlm.nih.gov/LHC-publications/pubs/TuberculosisChestXrayImageDataSets.html
[48]		ImageCLEF tuberculosis	The dataset contains the CT images of TB affected patients	https://www.imageclef.org/2018/tuberculosis
[49]		Tuberculosis datasets	BioGPS has 16 datasets of tuberculosis on 5 species	http://biogps.org/dataset/tag/tuberculosis/

Table 3 (continued)

References	Diseases	Dataset name	Description	URLs
[50]	Emphysema	Exasens dataset	This repository presents a new dataset for the categorization of four types of respiratory diseases: COPD, bronchitis, infections, and Healthy Controls (HC)	https://archive.ics.uci.edu/ml/datasets/Exasens
[51]		COPD dataset	BioGPS has 10 datasets of emphysema on 5 species	http://biogps.org/dataset/tag/copd/
[52]		COPD Gene	The data identifies phenotypes on chest CT in COPD patients, such as emphysema, air trapping, and airway clot formation	http://www.copdgene.org/study-design
[53]		Danish Lung Cancer Screening Trial	In 4104 smokers and former smokers between the ages of 50 and 70, the dataset includes yearly CT screening for lung cancer compared to no screening	https://clinicaltrials.gov/ct2/show/NCT00496977
[54]		Computed tomography emphysema database	115 high-resolution CT (HRCT) scans and 168 square areas are carefully labeled in a sample of the slices to make up the database	https://lauge-soerensen.github.io/emphysema-database/
[55]	Mesothelioma	Mesothelioma disease dataset	The dataset contains three hundred and twenty-four patients with Mesothelioma. All of the samples in the collection have 34 characteristics	https://archive.ics.uci.edu/ml/datasets/Mesothelioma%C3%A2%E2%82%AC%E2%84%A2s+disease+data+set+
[56]		BioGPS	BioGPS has 1 dataset of mesothelioma on 5 species	http://biogps.org/dataset/tag/malignant%20mesothelioma/
[57]		Mesobank	MesobanK collects fresh Mesothelioma tumour samples, stored for 24 h in RNA later and then frozen at − 80 degrees	https://www.mesobank.com/samples-and-research/?lang=en
[58]		ICCR mesothelioma	The dataset is made up of parts that include Mesothelioma histology, clinical care, grading, and mortality	http://www.iccr-cancer.org/datasets/published-datasets/thorax/mesothelioma
[59]		Mesothelioma disease dataset	The dataset has RNA-Seq data for Mesothelioma	https://ega-archive.org/datasets/EGAD00001001915
[60]	Pulmonary edema	MIMIC-CXR database	The collection includes 377,110 pictures from 227,835 radiography tests completed at Boston's Beth Israel Deaconess Medical Center	https://physionet.org/content/mimic-cxr-pe-severity/1.0.1/
[61]		MIMIC-CXR dataset	473,064 chest X-ray pictures and 206,574 clinical information from 63,478 edema patients make up the dataset	https://lcp.mit.edu/mimic
[62]		Chest X-ray dataset	There are 112,120 unknown CXR front view pictures in this public collection, obtained from 30,805 patients	https://clinicalcenter.nih.gov/
[63]		MIMIC-CXR dataset	A large publically labelled chest radiographs	https://sciwheel.com/work/signin?targetUrl=%2Fwork%2F%23%2Fitems%2F7368080

Table 3 (continued)

References	Diseases	Dataset name	Description	URLs
[64]	Pneumoconiosis	ChestX-ray8 database	This database includes CXR classifications depending on the presence or exclusion of 14 radiological abnormalities	https://nihcc.app.box.com/v/ChestXrayNIHCC
[65]		Pneumoconiosis radiograph dataset	The data is collected from Chongqing CDC's electronic health records, which is a complete image dataset	https://cloud.tsinghua.edu.cn/f/d8324c2-5dbb744b183df/
[66]		Chest X-ray dataset	The dataset has images and diagnostic labels associated with it	https://www.cdc.gov/niosh/index.htm
[67]		Chest X-ray dataset	There are posterior-anterior (PA) radiography images in the collection, some of which are totally digital and others which are digitized films	https://www.ilo.org/safework/info/WCMS_108548/lang--en/index.htm
[68]		Pneumoconiosis dataset	The dataset contains the chest-X ray information	https://github.com/liyu10000/pneumoconiosis

Pre-processing: After collecting the images from various repositories, it is our prime duty to pre-process them before training our system with them. It is because the images could be blurred, noisy or their features are not visible, and training the system with such images can affect the accuracy performance and generate the wrong output, which would be risky for people's health. There are various techniques by which the image can be pre-processed, such as Contrast Limited Adaptive Histogram Equalization (CLAHE) for increasing the image contrast, geometric transformations, image filtering, etc. [69]. Data pre-processing also includes data augmentation to increase the dataset size without collecting the new data and reduce overfitting. Data normalization is the process of organizing the data in a structured way. Multiple images are put into a common statistical distribution in terms of size and pixel values so that there can be a change in the range of pixel intensity values [70]. Resizing images is a critical pre-processing step since neural networks receive inputs of the same size. All images need to be resized to a fixed size before sending them as an input to the convolution neural network. After this, feature extraction could be conducted on the training images so that those features will feed the learning models to identify or predict the class of any new image. This process produces images that have been altered or modified and will be utilized in the training phase [71].

Learning models: Modern systems are deemed artificially intelligent when they use the machine and deep learning methods, which allows the computer (the machine) to understand tasks from an ever-changing dataset. Thanks to recent breakthroughs in learning algorithms and processing speed machines, deep learning models have become possible for many prediction problems [72]. Hence in this phase, machine learning algorithms, deep learning algorithms, transfer learning, and ensemble learning models can be selected based on various factors such as the size of the dataset, complexity of the data, etc. CNN (convolution neural network), which is excellent for finding image patterns, is a perfect algorithm for classifying things (and many other tasks involving images). CNN's, like neural networks in the real brain, are made up of neurons with trainable biases and weights that receive various inputs [73]. The information's weighted sum is then calculated. After that, the weighted total is input into an activation function, which results. As far as transfer learning is concerned, VGG16 (Visual Geometry Group), VGG19, MobileNetV2, ResNet50 (Residual Neural Network), and other pretrained models are frequently trained on massive datasets, a standard benchmark in the computer vision area. These models can be used directly to predict new tasks or as part of a model's training process. Moreover, these techniques, such as transfer and ensemble techniques, are also used to minimize training time, enhance the accuracy of the classification and prevent modeling errors in the system [74].

Classification: In this phase, the trained model will determine which class a picture belongs to, such as emphysema,

cystic fibrosis, pneumoconiosis, pulmonary embolism, pulmonary edema, asthma, mesothelioma, tuberculosis, covid-19, and lung cancer. If the image does not correspond to either of these disorders, the model should suggest that it is a normal lung image. To verify the system's performance, the model that correctly classifies the image should be evaluated using specific evaluative metrics such as accuracy, loss, precision, recall, F1 score, and so on (Table 4) [75–79].

4 Background

The work done by researchers to forecast the numerous airway diseases has been discussed in three subsections, i.e., pulmonary edema, pulmonary embolism, and covid 19; Mesothelioma and lung cancer; asthma, tuberculosis, and cystic fibrosis; emphysema and pneumoconiosis. To undertake comparative analysis, the goal or objective of their study has been initially briefed. Further, the datasets, procedures, outcomes, findings, and limits are given in the tabular form, i.e., from Table 5, 6, 7, 8, followed by the overall analysis.

4.1 Role of AI to Predict Pulmonary Edema, Pulmonary Embolism, Covid-19

Medical pictures are one of the first diagnoses since they can disclose specific undetected pathologic changes. Still, the absence of publicly available datasets and benchmark studies makes it impossible to compare and define the best detection systems. To solve this, Islam et al. [79] employed several datasets to analyze the effectiveness of models on various disorders, including pulmonary edema. To find the anomalies in chest X-rays, the investigator's employed trained classifiers. According to Liao et al. [80], one major obstacle in analyzing medical pictures is limited ground truth labels. As

Table 4 Evaluation metrics to test system performance

Parameters	Symbols	Formulae
Accuracy	Acc	$\frac{TrueNegative+TruePositive}{TruePositive+FalsePositive+TrueNegative+FalseNegative}$
Loss	Loss	$\frac{(ActualValue-PredictedValue)^2}{Numberofobservations}$
Area under the curve	AUC	$\lim_{x\to\infty}\sum_{i=1}^{n}f(x)$
Sensitivity	St	$\frac{TruePositive}{TruePositive+FalsePositive}$
Specificity	Sp	$\frac{TrueNegative}{TrueNegative+FalseNegative}$
Recall	Re	$\frac{TruePositive}{TruePositive+FalseNegative}$
Precision	Pr	$\frac{TruePositive}{TruePositive+FalsePositive}$
F1 Score	F1	$\frac{2*Precision*Recall}{Precision+Recall}$
Root mean square error	RMSE	$\sqrt{\sum_{i=1}^{n}\frac{(\hat{y}_i-y_i)^2}{n}}$;

a result, they designed and tested a semi-supervised learning system for estimating pulmonary edema to aid therapeutic choices in congestive heart failure. To tackle the problem, they created a Bayesian model (Eq. 1) that learns the probabilistic feature representations from the complete picture collection and uses them to forecast the degree of edema

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \tag{1}$$

where $P(c|x)$ is the posterior probability, $P(x|c)$ is the likelihood, $P(c)$ is the class prior probability, and $P(x)$ is the predictor prior probability,

The authors maximized the log probability of the data to construct the probabilistic feature representation (Eq. 2) that has been learned from all images in predicting pulmonary edema.

$$\log p(x,y;\theta) = \sum_{i=1}^{N_i} \log p(x_i,y_i;\theta) + \sum_{i=N_i+1}^{N} \log p(x_i;\theta) \tag{2}$$

Here θ is the parameters, N is the total number of images, y is the edema severity label, x represent the single image.

According to Hong et al. [81], patients with pulmonary edema have many consequences for the rest of their lives. Physicians must deal with stress, concern, and inconvenience when treating such patients because they may suffocate if they are not treated. As a result, the author advocated a comprehensive investigation on the application of auditory classification algorithms for the automated identification of excessive lung water that engorges the alveolar beds to address these issues. The authors developed a unique approach using recursive feature eliminations, logistic regression (Eq. 3), and principal component analysis (Eqs. 4–6) to validate the learned data with supplemented samples from local hospitals.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{3}$$

Here Y_i is dependent variable, β_0 is population y-intercept, β_1 is population slope coefficient, X_i is an independent variable, and ε_i is random term error. Given a dataset of N centered observations in a d-dimensional space

$$X = \{x_1, x_2, x_3, \ldots\ldots\ldots, x_N\}, \sum_{k=1}^{N} x_k = 0, x_k \in R^d \tag{4}$$

PCA diagonalizes the covariance matrix where C is a covariance matrix,

$$C = \frac{1}{N}\sum_{K=1}^{N} X_K X_K^T \tag{5}$$

To solve the Eq. (5) using eigen values

Table 5 Comparative analysis of the techniques used to predict pulmonary edema, pulmonary embolism, covid-19

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[79]	Pulmonary edema	Indiana chest X-ray dataset, JSRT dataset, Shenzen dataset	Alex Net, VGG16, VGG19, ResNet50, Adam Optimizer, ResNet101, ResNet152	Accuracy = 52% AUC = 0.94 Sensitivity = 96% Specificity = 96%	The authors found that their network could localize the abnormalities such as Pulmonary edema successfully	The system generated degraded results when the rule based features were concatenated with the features extracted from DCN
[80]		330,000 chest X-ray images collected from Beth Israel Deaconess Medical Center	EM (electron microscopy), DGM(Deep Galerkin Method)	RMSE = 0.66 Pearson CC (correlation coefficient) = 0.52	Their findings promised to help physicians provide better treatment by allowing them to quantify the degree of pulmonary edema using chest X-ray pictures	Limited ground truth labels had been used for medical image analysis
[82]		in situ biochemical investigation dataset	Principal Component Analysis, Random Forest	Sensitivity = 97.3% Accuracy = 96.5% Specificity = 95.5%	The findings showed that FTIR micro spectroscopy in conjunction with chemometrics might be a useful tool recognizing it pulmonary edema	The model should be able to determine the contaminative effects of breakdown components in FTIR (Fourier transform infrared spectroscopy) spectroscopy measurements
[83]		MIMIC-CXR dataset	DenseNet, random minority oversampling, ResNet50	AUC = 79.1%	The authors found DL as a capable approach to estimate pulmonary edema from chest radiographs	Class imbalance
[84]		40 normal images taken from Show Chwan Memorial Hospital, Changhua, Taiwan	Support vector machine, Gabor filter	AUC (area under curve) = 0.999	The study was able to distinguish between normal and pulmonary edema lung images	The algorithms had been applied only on limited dataset which needs to be improved
[85]		CXR images from NIH clinical centre	CNN DenseNet model, Adam optimizer, K-fold cross validation	AUROC = 0.9164 Sensitivity = 71.493% Specificity = 10.011%	The design was able to distinguish precise information with localisation in the form of a heat map, making it simpler to discover and characterise pathologic anomalies associated with acute pulmonary edema	The system's performance needed to be improved to enhance the evaluation values
[86]		1.5 M frontal (PA) CXR studies obtained from adults over 18 years of age	RadBot-CXR(chest X-ray)	AUC = 93.6%	The authors reached a level of automatic interpretation of chest X-rays for pulmonary edema that was comparable to that of an expert	The architecture needed to address the localization of disease identified on the radiographic image

Table 5 (continued)

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[87]	Pulmonary embolism (PE)	CT scan Dataset	Deep Learning, Image Classification, Long Short Term Model	IoU (Intersection over union) threshold = 50% Accuracy = 91% Precision = 68%	The model provides a quick fix by combining categorization and detection of object algorithms to increase the performance of pulmonary embolism detection	The authors used CTPA-Scans from a single CT-Scan system (software and hardware) which varied from one system to another hence decreased the accuracy
[88]		2800 CTPA-Scan (CT pulmonary angiogram) Dataset	Linear Regression	Sensitivity = 93% Specificity = 95.5% F1 score = 86%	The authors found that the AI prototype algorithm had a greater degree of diagnostic performance for detecting PE on CTPAs	Testing for pulmonary embolism that might influence performance measurements were incorrectly categorised by the system
[98]		85 CTA lung images collected from November 2016 to May 2017 at the Wuhan Central Hospital	Generational Advanced Network	Sensitivity = 90.9% Susceptibility = 92%	The suggested computer-aided diagnostic technique successfully increased the diagnosis rate of PE	The algorithm had a significant probability of misdetection, indicating that more in-depth training datasets were required for embolism detection below grade 3
[99]		DICOM images of CTPA taken since January 2019	Deep Learning algorithm	Sensitivity = 79.6% Specificity = 95.0%	The model assessed the impact of DL-based PE identification on patient care parameters	Despite their excellent performance, the authors were unable to produce meaningful effects on medical performance metrics
[89]		590 patients (460 with APE and 130 without APE)	Deep Learning, Convolution Neural Network, U-Net	Sensitivity = 94.6% Specificity = 76.5%	The AUC of DL-CNN for the identification of pulmonary embolism was high, and it may help doctors to minimize their strain	The DL-CNN model could not be trained due to a lack of big training sets
[100]		Stanford dataset	Convolution Neural Network, PENet	AUROC = 84% Accuracy = 81% Specificity = 82% Sensitivity = 75%	The findings worked on the hard part of diagnosing pulmonary embolism with no the need of time and cost consuming methods	Failed in recognizing the other pathogens of pulmonary embolism
[90]		1427 people of an Italian National Hospital "Ospedali Riuniti di Ancona"	Machine Learning, Artificial Network, Q Analysis	Accuracy = 86%	The proposed framework was able to study partial as well as incomplete data of pulmonary embolism	The S[B]-paradigm had to be used to characterise final negative and positive diagnoses in the system

Table 5 (continued)

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[91]	Covid-19	Cohen's dataset	CNN, HOG (Histogram of Oriented Gradients)	Accuracy = 92.95% Recall = 85% Specificity = 82% Precision = 91.5%	The suggested CNN approach has a high detection rate and is quick and easy to use	The study analyzed the resilience of the systems by reacting to real-world circumstances using restricted datasets from multiple sources
[92]		200 chest X-ray and 180 Covid-19 images	SVM, CNN model, ResNet50	Accuracy: 92%	Local descriptors were improved by deep learning algorithms. Deep features as well as the SVM classification algorithm, in particular, outperformed the other techniques	The study should incorporate different imagistic patterns of Covid 19
[93]		Data collected from Jan 22, 2020 to Apr 1 2020 at Johns Hopkins University	Support Vector Machine, Deep Neural Network, Long Short Term Memory, Polynomial Regression	RMSE Confirmed: 455.92 Death: 117.94 Recovered: 809.71	In anticipating the COVID-19 transmission, the findings demonstrated that polynomial regression (PR) produced the lowest root mean square error (RMSE) score when compared to other methodologies	The study should work on more algorithms to enhance the RMSE score
[94]		OSR dataset of 1624 patients	Logistic regression, Naïve Bayes, KNN(K nearest neighbour), Random forest, SVM	Accuracy = 88% Sensitivity = 89% Specificity = 91% AUC = 90%	The ML algorithms provided in their research performed similarly to, but not as well as, RT-PCR for COVID-19 diagnosis	The technique only worked for those people who are covid negative. For positive patients, the model failed to detect covid in them
[95]		Real world data of 337 patient images	nConv net, Deep learning	Accuracy: 97.6%	This approach might aid hospital administrators and medical professionals in taking the essential actions to handle COVID-19 patients following their rapid diagnosis	System worked on the small dataset
[96]		Dataset collected from Joseph Paul Cohen and Paul Morrison Lan Dao	Mamta Ray Foraging Optimization, Fractional Multichannel Exponent Moments	Accuracy: 98.09%	By picking the most important traits, the suggested technique was capable of achieving both high efficiency and low resource usage	The system dealt with resource limitations and high CPU time
[97]		1065 CT pathogenic images	CNN, GraphNet	Accuracy: 89.5% Specificity: 0.88 Sensitivity: 0.87	The findings showed the achievement of rapid and accurately diagnosis of COVID-19	The performance of deep learning models was hampered due to signal-to-noise ratio

Table 6 Comparative analysis of the techniques used to predict Mesothelioma and lung cancer

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[101]	Mesothelioma	Data collected from the Diyarbakir district of southeast Turkey (324 patients)	Apriori method, recursive feature elimination method	Lift = 1.0–1.6 Support = 0.5–1.0 Confidence = 0.5–1.0	This study came to some important results on MM prognostic variables. Their findings revealed that histopathological variables have a role in the development of MM	Large dataset would escalate the execution time
[102]		UCI repository dataset	Apriori Algorithm	Support 75% confidence 90%	Asbestos exposure, its length, and duration of symptoms all had a significant influence on the frequency of MM, according to the findings	To extract the rules, the model did not examine the various association mining techniques
[103]		Dicle University, Turkey	Random forest, Clojure classifier (CC), decision tree model, kernel logistic regression (KLR)	Accuracy = 71.29%	The findings from biopsy as well as radiological testing are good predictors of mesothelioma, according to the authors	Overfitting was a problem with decision tree models, particularly random forest, which lowered the system's accuracy
[104]		UCI repository	Apriori Algorithm, Association Rule Mining, Data Normalization	Confidence Support = 75%	The data showed that the length of symptoms, together with exposure to asbestos, has had a significant impact on the MM (malignant Mesothelioma) rate	If memory usage for the quantity of transitions was limited, the Apriori algorithm revealed an issue of incompetence
[105]		Dataset collected from UCI machine learning repository	SMOTE, ADASYN, Artificial neural network, principal component analysis	Accuracy = 96%	Their study highlighted the significant input features of malignant Mesothelioma problem	Only few factors related to Mesothelioma were studied

Table 6 (continued)

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[107]	Lung Cancer	LUNA 16 and Kaggle Data Science Bowl (KDSB)	DFD-Net, denoising model, Convolution Neural Network	Accuracy = 87.8% Specificity = 89.1%	The method considerably enhanced the performance of model	The method addressed the issue of class imbalance in the information while using Kaggle dataset
[108]		Computed Tomography Scan Images	Fuzzy Particle Swarm Optimization(FPSO),Convolution Neural Network,	Accuracy = 94.97% St = 96.68% Sp = 95.89%	This study used an input lung picture to detect malignant lung nodules and to categorise lung cancer and its degree	Incorrect classification of cancer as benign or malignant
[109]		Hubei Taihe Hospital (110 patients)	Naïve Bayes Classifier,Fast Correlation Based Filter,	AUC = 98.9% Sensitivity = 98.1% Specificity = 100%	The ability of metabolic indicators to identify lung tumours early had been proven	Smaller size of data used to carry the research
[110]		Chest X-ray dataset	Convolution Neural Network	Acc = 84.02% Specificity = 85.34% Sensitivity = 82.71%	The training strategy that had been proposed was performing superior than the already existed transfer learning method	Accuracy needed to improved by working on its features
[111]		Data collected from cancer imaging archive (CIA) dataset	Ensemble classifier	Accuracy = 97.6% Precision = 97.2% Recall = 97.5% F1 score = 97.5%	The system recognized the cancer with maximum accuracy	-
[113]		LIDC-IDRI dataset	Convolution neural network	AUC = 0.967	The system gave a summary of the most common methods for nodule categorization as well as lung cancer prediction using CT imaging data	The author mentioned that the system efficiency for the training and testing data sets employed should be considered

Table 7 Comparative analysis of the techniques used to predict asthma, tuberculosis, cystic fibrosis

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[114]	Asthma	Asthma dataset	Decision Tree, Naïve Bayes, K Nearest Neigbour	Accuracy = 96.52%	The suggested model provided a low-cost, user-friendly instrument for detecting and classifying asthma in its early stages	Limited dataset used for testing
[115]		Clinical data of 2870 patients	Feature based time series classification	Recall = 87% Accuracy = 92% Precision = 89% F measure = 88% Specificity = 94% AUC = 87%	The researchers looked at how time—series data dynamics as well as temporal sequences affected daily asthma symptoms	The models should be trained with large sample of dataset
[116]		Data of 2010 patients	Principal component Analysis, Naïve Bayes	AUC = 85% Sensitivity = 90% Specificity = 83%	The authors had shown the capability to progress the early detection of asthma exacerbations when compared to traditional paper-based action plans	The data was collected manually which could have been prone to inaccuracies
[117]		Clinical data of 1225 patients	Normalization, Orthogonal array, Mahalanobis distance	Accuracy = 94.15%	The study developed an asthma detection algorithm that accurately detects illness	The study did not included a precise asthma risk score or a list of reference doctors
[118]	Tuberculosis	Data collected from Shenzhen china dataset and montgomery country dataset	CLAHE method, Deep convolution neural network, UNet architecture	Accuracy = 97.1% Specificity = 96.2% Sensitivity = 97.9%	The model provided inexpensive easily accessible, highly accurate solutions for the low-income countries who are suffering from TB disease	Even if the model achieved good sensitivity and specificity, it isn't easy to compare model performance to human performance because it hasn't been tested in the field yet
[119]		microscopic images of 22 sputum smear	CNN	Recall = 97.1% Precision = 78% F-score = 86.7%	The model helped the physicians to detect the disease in a small amount of time to improve the clinical outcome	–
[120]		100 TB ctscan images	ResNet	Accuracy = 85.2%	This research was expected for the substantial part to the discipline and encourages the use of ML approaches in the medical area	Traditional 3D CNN architectures tend to be less successful due to the limitation of the datasets as well as the features of comparable aberrant patterns of TB among the five severity levels
[121]		Montgomery and Shenzhen datasets	Ensemble classifier, Canny Edge Detector, Deep Learning	Accuracy = 93.59%, Specificity = 94.87% Sensitivity = 92.31%	The findings show that employing ensemble classifiers trained on a variety of characteristics derived from different types of photos improves detection performance	The model needed to be expanded to categorize chest X-rays on the basis of severity of tuberculosis, and more features needed to be investigated to improve the classifiers' performance

Table 7 (continued)

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[122]		501 Computed Tomography Scan images	Convolution Neural Network, Deep Learning	Recall=98.7% Precision=93.7%	The detection of pulmonary TB was successfully researched in the article, and a quantitative diagnosis report was developed after researching a large number of relevant publications and methodologies	The system failed to correctly classify pulmonary tuberculosis
[123]		Taipei Medical university	RF, ANN	Accuracy=88.67% Specificity=90.4% Sensitivity=80%	The model gave a chance for doctors to conduct preventative actions before ATDH occurred	The models were not trained with large data sample
[124]		Real time dataset collected in the form CXray images	Linear regression, KNN, Naïve Bayes, Decision Trees, Random Forest, SVM, MLP	Accuracy=98.57% Precision=99.58% Sensitivity=91.50% Specificity=99.50%	The authors filled a significant vacuum in the literature by comparing and contrasting multiple algorithms for predicting tuberculosis	To enhance model hyper-parameter tuning, the authors intended to investigate single class classification approaches and assessed the use of, deep learning ensembles
[125]		Montgomery dataset Shenzen dataset	InceptionNet model	Accuracy=87.5% AUC=0.92 Sensitivity=0.76 Specificity=0.95 Accuracy=91.7% AUC=0.96 Sensitivity=0.89 Specificity=0.93	The model offers a quick and accurate method for mass tuberculosis screening in low-resource areas across the world	Limited dataset alone gave insight into their model's intrinsic capabilities
[126]	Cystic fibrosis	Sixteen clinical data of patients whose age ranged from 19–59 years	HASTE technique (Half Fourier Single Shot Turbo spin-Echo), Lasso method	Recall=0.68 Precision=0.016	The researchers were able to define small intestines morphofunctional abnormalities in individuals with cystic fibrosis	Objective indicators of therapeutic responsiveness to the novel CFTR modulators, as well as gastrointestinal outcome assessments, were required
[128]		CT scans of 194 patients	Deep Learning, Cascade Network, binary and multiclass classification	Accuracy=94%	The findings revealed that the suggested technique may detect abnormalities and assign a grade to each illness in the initial stages of CF respiratory illness	The model needed to work on the detection of cystic fibrosis
[129]		To produce synthetic data, a second order ODE mathematical model of the lung was employed	Support vector machine, Naïve Bayes classifier, logistic regression,	Accuracy=99%	The model gave real-time assistance to clinicians in making diagnostic decisions	Their work was limited by the fact that all of the training and testing was done with simulated data and that all of the classes used the similar profile

Table 7 (continued)

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[130]		1000 distinct peaks were extracted from 277 perspiration samples	Gaussian based decision tree model	Accuracy = 98% Recall = 96% Precision = 94% Specificity = 98%	The model looked at the relationship between lipid profiles and moderate and severe CFTR gene mutations	Enrolling additional patients to enhance the sample size of a genetically varied CF community might improve this research in the future

$$\lambda v = Cv => \lambda\left(X_k.v\right) = \left(X_k.Cv\right), where\, k = 1, 2, \ldots., N \quad (6)$$

Lin et al. [82] investigated the diagnosis and progression of illnesses to integrate bio-fluid-based infrared spectroscopy into the clinical area. The authors looked at using Fourier transform infrared micro spectroscopy to detect abrupt cardiac death. Assessing the degree of pulmonary edema is complex, and detecting it in chest radiographs allows doctors to make prompt treatment decisions for patients. Using the large-scale clinical MIMIC-CXR (Chest X-ray Scan) database, Kumar et al. [83] tested different supervised and semi-supervised deep learning approaches to detect the degree of edema from radiological pictures. Furthermore, the authors evaluated three ways to alleviate class imbalance during implementation: weighted cross entropy loss (Eqs. 7–8), class aware sampling, and random minority oversampling. Kumar et al. [84] reported a texture analysis of chest X-ray to detect pulmonary edema in chest X-rays automatically. To put it another way, the authors were able to tell the difference between a chest X-ray with indications of pulmonary edema and a normal chest X-ray.

$$l_i = -\sum_{c \in C} y_i, c\log(p_i, c) \quad (7)$$

$$l^w = \sum_i m_i l_i \quad (8)$$

where l_i is the ith window's component of the multiclass cross-entropy loss, l^w is the weighted cross-entropy function, p is the probability that the window i belongs to class c as predicted by the given model, m_i is the weight applied to the ith component. According to Hayat et al. [85], quick screening of pulmonary edema patients is required such that radiologists can make a prediction as soon as feasible. However, depending on specialists' knowledge of reasoning impedes the diagnostic process. As a result, the author created a deep learning-based architectural model to detect the presence of acute pulmonary edema in chest X-ray pictures. Brestel et al. [86] sought to deliver expert-level information to every chest X-ray image right away. To achieve expert-level automatic interpretation of regular chest X-rays, they used a machine learning approach and discussed the results using a robust approach of clinical validation.

Kiourt et al. [87] used transfer learning methodologies to adopt and analyze some of the most common convolutional neural network designs to get decent model accuracy for detecting pulmonary embolism in CT scans. Deep convolutional neural network (DCNN) exhibited good results in identifying critical abnormalities in CT images, including intracranial haemorrhage, acute brain ischemia, and essential abdomen findings, according to Weikert et al. [88]. An end-to-end, fully convolutional network

(i.e., U-Net) was created by Liu et al. [89] to segment clots and determine clot volume in CTPA (CT pulmonary angiogram). This study aimed to evaluate U-ability Net's to identify clots in terms of efficiency and accuracy and compute the APE clot load (acute pulmonary embolism). Rucco et al. [90] suggested CAD (computer-aided drugs) for the mathematical theory of Hyper-networks and Q-analysis, allowing for a dimensional patient dataset description. The findings were utilized for feature selection in the artificial neural network training stage. The authors were able to identify and diagnose pulmonary embolism in that manner.

Since we know that the number of Covid-19 cases has risen dramatically in the last year, it's become even more critical to track and identify healthy and infected persons quickly and precisely. Many existing detection approaches are ineffective in detecting viral patterns. As a result, Chen [91] developed an excellent classification approach for detecting COVID-19 viral sequences. Ismael et al. [92] employed an AI-based technique that effectively monitors various lung illnesses. For deep feature extraction, the authors operated pre-trained CNN models with SVM (support vector machine) classifiers with different kernel functions such as linear, quadratic, cubic, and Gaussian, used for covid-19 classification.

Similarly, according to Punn et al. [93], artificial intelligence experts have concentrated their professional knowledge on constructing mathematical models for assessing the epidemic condition utilizing state-wide shared data. As a result, the authors advocated a few AI models to analyze everyday exponential behaviour and the projection of future corona virus reach-ability across countries using real-time data. The data science community has developed many machine learning (ML) models to improve Covid-19's diagnostic capabilities. Most are based on computed tomography (CT) images or chest X-rays. Cabitza et al. [94] used various classifiers by including machine learning techniques for blood-test results, which are generally available in clinical practices. Panwar et al. [95] focused on observational lockdown analysis and said that artificial intelligence technologies were necessary to defeat such a solution. With this in mind, the scientists presented a CNN-based algorithm, emphasizing that lockdown isn't the only way to combat the covid19 epidemic. Elaziz et al. [96] demonstrated a technique for accurately classifying covid-19 chest X-ray pictures. The characteristics of orthogonal moment features and feature selection approaches determine the classification strategy. The authors created a novel feature selection approach based on several assessment techniques to improve the behavior of Mamta ray foraging optimization. Several aspects may be used to identify viral infections based on imaging patterns. Wang et al. [97] hypothesized that CNN might aid in identifying distinctive characteristics that would be difficult to detect using visual recognition alone.

4.2 Role of AI to Predict Mesothelioma, Lung Cancer

Alam et al. [101] aimed to look for clinical, radiological, and histological variables in malignant Mesothelioma. The authors suggested a novel framework for identifying prognostic indicators utilizing non-invasive and cost-effective methods based on various techniques. According to the authors, their suggested framework would aid medical professionals and healthcare experts in detecting malignant Mesothelioma early and treating it more effectively by including crucial prognostic markers. In their study, Latif et al. [102] looked at the risk factors for malignant Mesothelioma. The scientists employed a dataset that included healthy people and Mesothelioma patients, but only Mesothelioma patients were chosen for symptom identification. According to the authors, these findings will aid in managing MM-related co-morbidities such as cardiovascular disease, cancer-related mental distress, diabetes, anemia, and hypothyroidism. Choudhry et al. [103] employed artificial intelligence-based algorithms to offer the best system for Malignant Pleural Mesothelioma early identification and prognosis (MPM). According to the authors, decision tree models, random forest, have a risk of overfitting; hence, they created a model that can diagnose with or without pricey biopsy data to overcome the flaws outlined above.

Similarly, Alam et al. [104] concentrated on investigating MM risk variables. The scientists included ill and healthy people in their study, resulting in a larger dataset. The dataset has a class imbalance problem, with the number of malignant Mesothelioma patients being much lower than the number of healthy people. Furthermore, the numerical attributes were categorized as nominal attributes, and association rules were created in the dataset. To detect malignant Mesothelioma, Gupta et al. [105] compared numerous machine learning algorithms with different feature sets to address the class imbalance problem. To achieve this, the authors used three sampling techniques: resampling, synthetic minority oversampling technique (SMOTE), and adaptive synthetic sampling (ADASYN). They also used other dimension reduction approaches, such as the ordinary least square method (OLS), principal component analysis (PCA), and random forest feature selection (RFFS), as well as genetic algorithms, to determine the exact collection of features. MesoNet was developed by Courtiol et al. [106]. It successfully predicted survival rates of Mesothelioma patients using whole-slide digitized images with no need for pathologist-provided locally labeled regions. The researchers also confirmed that the method was more suitable for estimating the patient's survival rate than conventional pathological approaches.

Table 8 Comparative analysis of the techniques used to predict emphysema, pneumoconiosis

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[131]	Emphysema	Benchmark dataset and manual dataset of 39 and 19 patients respectively	Improved red deer algorithm, Fuzzy C means, Adaptive local ternary pattern	Benchmark Dataset Accuracy = 94.99% F1 Score = 89.45% Sensitivity = 85.66% Manual Dataset Accuracy = 95.56% Sensitivity = 91.6% F1 score = 95.25%	Their research looked at how a deep learning technology may be used to diagnose pulmonary emphysema automatically	On the generated picture, the approach was unable to offer localization information
[132]		Data from the Danish Lung Cancer Screening Trial, which included 1990 people	Proportion Net, GAP Net	AUC = 0.96	The model's possible to identify the target was excellent enough to accurately characterize the geographical distribution of emphysema	With human-level accuracy, the geo-location was excellent enough to categorize the geographical distribution of emphysema
[133]		7143 COPD Gene Participants	Deep Learning, Convolution Neural Network, Cox proportional hazard models	Confidence Interval = 95%	The technology produced a decipherable result to recognize patients who are at higher risk of death	The model might be impacted by the unique CT methodology because it was trained using just COPD Gene data
[134]		126 input recordings taken from respiratory sounds database	CNN, MFCC, Librosa machine learning features, K fold cross validation	Specificity = 0.93 Sensitivity = 0.93 ICBHI score = 0.93	Medical specialists are able to use the algorithm to diagnose COPD using breathing sounds	They can expand its functions in the future to assist physicians in diagnosing numerous other ailments and their severity
[135]		The Danish Lung Cancer Screening Study used 600 low-dose CT images	Multiple instance learning approach	AUC (scan level prediction) = 0.82 AUC (region level prediction) = 0.88	The study give accurate predictions of emphysema occurrence at both the scan and area levels	The method resulted in a higher ratio of false to true detections in the region, resulting in worse search performance
[136]		From COLIBRI-COPD, 1409 COPD patients were recruited	SMOTE, Random forest	Accuracy = 84% PPV = 87% Sensitivity = 59%	The system was able to detect COPD at all stages of severity	Accuracy needs to be improved
[137]		9,925 CT scans collected from COPD Gene multi-center	Convolutional Neural Network	Pearson co-relation coefficient = 0.940	Biomarkers from images were learned directly by the deep-learning regression architecture and also simplified the development of biomarker extraction algorithms	The suggested method's shortcoming was that it required two-dimensional depth of field of the structures where biomarkers were calculated, which was not practical at the time
[138]		HRCT scans collected from Frederikshavn (Fre) and Aalborg (Aal) datasets	MILES classifier, Cross validation, miSVM-Q classifier	miSVM-Q AUC = 95% MILES AUC = 78.8%	Their research revealed two novel multiple instance based classifiers that can detect emphysema patches in COPD people without any need for manual annotation	The quantity and proportion of the datasets were the study's key limitations

Table 8 (continued)

References	Diseases	Dataset	Techniques	Outcome	Findings	Challenges/remarks
[139]	Pneumoconiosis	1881 digital images of chest X-ray	DCNN, Inception V3	AUC = 87.8%	The authors found that deep leaning solution could produce a comparatively superior classification performance	Limited dataset and more time complexity
[65]		Chongqing CDC, China collected image data from August 2016 to June 2017	ResNet 34, DenseNet40, DenseNet64, DenseNet53	Accuracy = 88.6% Precision = 83.3% Recall = 53.6% F1 score = 65.2%	The authors suspected pneumoconiosis in radiographic images	The authors needed to work on the performance of the system to enhance its accuracy
[140]		5424 chest radiographic images	CAD algorithm, McNemar test	Sensitivity = 0.89 to 0.98 Specificity = 0.68 to 0.86 Kappa values = 0.57 to 0.84	The findings of investigation showed that CAD can significantly increase performance of pneumoconiosis diagnosis	The research only covered a limited amount of patients, and the outcomes were inconsistent
[141]		405 pneumoconiosis patients	Convolution Neural Network	Accuracy = 97.3% Sensitivity = 98.1% Specificity = 97%	In terms of pneumoconiosis grading accuracy, their study beat two categories of radiologists	The model was unable to distinguish pneumoconiosis from those other lung illnesses with comparable pathologies
[142]		Original pulse signal	Support vector machine, RBF kernel function	Precision = 100% Recall = 86.85% F-measure = 92.96% Accuracy = 88.31%	The researchers' proposed technique of detecting coal miners' pulse signals will remind coal workers of the need of verifying and preventing illness progression	In the future, the authors would need to optimize the goal function to make the entire system greater realistic, taking into account the various consequences of misinterpretation of diagnosis
[143]		Pneumoconiosis data	Transfer learning, CNN	Specificity = 87% Sensitivity = 95% AUC = 94%	Both transfer learning strategies outperformed beginning from scratch with insufficient training data, according to the findings of the experiments	It was difficult to satisfy the need of information using deep learning since chest X-rays utilised in pneumoconiosis was limited due to sensitive patient data
[144]		Dataset taken from National Institute for Occupational Safety and Health website (NIOSH)	Autoencoder, SVM, CheXNet, Multilayer perceptron	Sensitivity = 93.3% Specificity = 88.46% Accuracy = 90.24%	When training datasets are skewed or lack variety, the cascaded machine learning architecture might be employed in various medical image analysis	Future research included a pilot study in which the approach is tested in a clinical environment using human readers

Several researchers have proposed several CNN strategies for lung cancer detection. These models, however, could not give the expected detection accuracy. Sori et al. [107] utilized a retraining technique, i.e., multi-phase CNN. The initial training was identical to the usual one. It was based on fine-tuning, which just used the selected parts of the model to describe and learn distinct morphology of lung nodule characteristics for contextual information, as shown in Fig. 5. Much work has gone into developing computer-assisted diagnosis and detection methods to increase the diagnostic quality for lung cancer detection categorization. As a result, Asuntha et al. [108] attempted to identify malignant lung nodules in the input picture and categorize lung cancer according to severity. The authors employed several optimal feature extraction approaches such as local binary pattern (LBP), scale-invariant feature transform (SIFT), and others to extract textural, geometric, and intensity features. Machine learning models improve the model's performance by learning from previous experiences. These models also seek to identify practical factors and their relationships. As a result, Xie et al. [109] discovered clinical metabolic markers that demonstrated significant differences between lung tumor patients and healthy persons. According to the authors, biomarkers were also employed to distinguish between histological subtypes and illness degrees.

Because they had a restricted lung cancer image dataset that did not work out using typical transfer learning and data augmentation techniques, Ausawalaithong et al. [110] implemented the transfer learning model twice to improve its performance. The first occurred when the model was transferred from the domain of public image to the chest X-ray. The model was then applied to lung cancer for the second time. According to the authors, multi-transfer learning solved the sample's limited size and produced greater results on the test than classic transfer learning. Shakeel et al. [111] introduced an intelligent machine learning approach to improvise the lung detection process. The CT scan-based lung pictures were continually evaluated using a multilayer brightness-preserving method that reduced image noise and improved lung image quality. Due to the segmentation process's relevance, the scientists adopted a multilayer augmented deep neural network technique to extract the cancer-affected area. Pradhan et al. [112] researched to predict various illnesses to make a judgment on lung cancer prediction. The authors also provided a comprehensive analysis of several machine learning algorithms to assess their competence and performance in predicting lung cancer and, as a result, detecting lung cancer with IoT integration. In a retrospective study of the USA National Lung Screening Trial (NLST), Kadir et al. [113] indicated the influence of lung imaging reporting and data system (lung RADS). Although lung RADS has been shown to lower the total number of benign nodules during screening, its categorization task has proven complicated.

As a result, the authors addressed the issue by recommending that radiologists and pulmonary medicine specialists use computer-assisted technology as a tool. As a result, the authors analyzed their study's progress in developing and validating the lung cancer predictive model and nodule categorization.

4.3 Role of AI to Predict Asthma, Tuberculosis, Cystic Fibrosis

Awal et al. [114] applied multiple learning models to investigate the parameters that characterize asthma diagnosis and prediction. BOMLA (Bayesian Optimisation-based Machine Learning Framework for Asthma), a new machine learning technique, had been developed for identifying asthma. Khasha et al. [115] set out to identify asthma control levels, and they only looked at research that used supervised approaches like classification models in data mining. Their primary goal was to improve the effectiveness of classification algorithms by investigating the impact of daily clinical data's time-series/time-sequences dynamics on asthma control level detection in patients. Zhang et al. [116] identified severe asthma exacerbations based on freely accessible daily monitoring data. Compared to previously published models, the authors hypothesized that a predictive algorithm created utilizing machine learning techniques and an extensive training dataset of daily monitoring data would yield greater accuracy for identifying asthma exacerbations, as shown in Fig. 6.

Several studies have employed Mahalanobis–Taguchi System (MTS) for intelligent illness detection with reasonable accuracy, according to Zhan et al. [117]. Their research aimed to see if MTS could be used to diagnose asthma based on regular blood measurements from healthy people and people living with asthma.

Tuberculosis can be cured by diagnosing it at its early stages, and to make it happen, the main requirement is to use the diagnostic technologies properly. Hence, Dasanayaka et al. [118] presented a model that can detect TB using deep generative adversal network. The chest X-ray images were chosen based on the subjective and objective quality assessment metrics. The objective quality assessment metrics were selected as peak signal-to-noise ratio (PSNR) (calculated by Eqs. 9–10), and the radiologists performed a subjective evaluation.

$$PSNR = 10log_{10}\left(\frac{MAX^2}{MSE}\right) \tag{9}$$

And to calculate (mean square error) MSE, we have an Eq. 10 where I and K are the observed and predicted values, m &n are the data points respectively
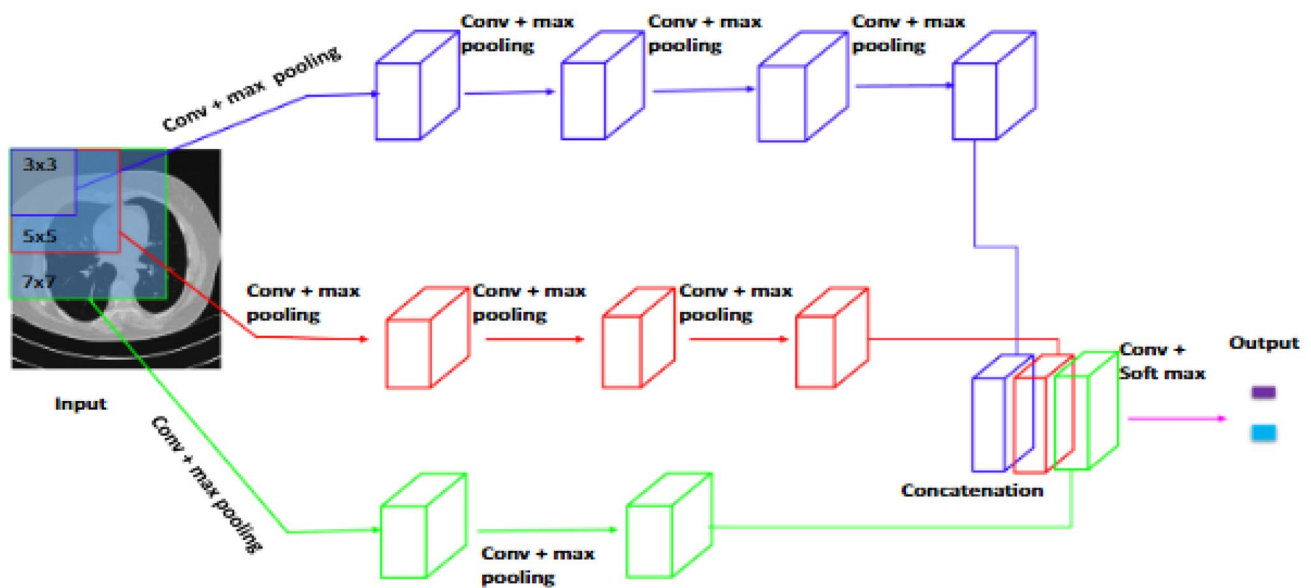
Fig. 5 Architecture based on multi-pathway CNN to detect lung cancer [107]

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - k(i,j)] \tag{10}$$

According to Panicker et al. [119], traditional microscopic sputum smear screening for TB (tuberculosis) diagnosis is time-consuming and error-prone. To make the diagnostic procedure more accessible, the scientists employed an image binarization approach and a modified convolution neural network to detect TB by identifying the pixels in the picture corresponding to bacilli. Gao et al. [120] concentrated on using cutting-edge deep learning approaches to analyze CT pulmonary images. The authors identified five categories of severity for tuberculosis to track therapy effectiveness. The latter was the subject of Hwa et al. [121]. They detected tuberculosis using contrast-enhanced canny edge detected (CEED-Canny) X-ray images which generated edge detection in lung X-ray pictures. Li et al. [122] developed a new model for generating quantitative computed

According to Lai et al. [123], machine learning technologies identify and diagnose various disorders. The effectiveness of artificial neural networks and support vector machine in predicting the development of prostate, ovarian, breast, and liver cancer were outstanding. As a result, the researchers utilized machine learning-based three algorithms to predict anti-tuberculosis drug-induced hepatotoxicity. According to Barros et al. [124], determining the severity of illnesses is necessary to improve patient quality of life, efficiently manage health resources, and so on. As a result, they focused on evaluating nine machine learning models, including KNN (calculated by Eq. 11), Nave Bayes, and decision trees, on improving TB prognosis to predict the chance of mortality using patient geographical, medical, and laboratory data. To summarise, the authors used feature selection techniques to identify the most relevant fields, used randomized search techniques to choose the optimal hyper-parameters of the machine learning model, and proposed an ensemble learning model to achieve better results.

$$d(p,q) = d(q,p) = \sqrt{(q_1 - p_1)^2 + (q_2 + p_2)^2 + \ldots + (q_n + p_n)^2} = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2} \tag{11}$$

tomography images in a clinic for diagnosing pulmonary TB. Based on CT data, the scientists used a fine-tuned 3D CNN model to classify pulmonary TB lesion areas. Their mission was to digitally assess the spatial position of each lesion, the confidence of each infection, the presence of calcification, lesion type classifications, overall infection likelihood, and practical volume of the left and right lungs.

Here the two points in Euclidean n-space are p and q. According to Das et al. [125], a chest X-ray is a potential signal for detecting TB. However, the shortage of competent radiologists in limited resource areas exacerbates the problem. As a result, the authors wanted to employ an end-to-end deep learning TB screening tool based on chest X-ray images so that their system could adjust to changed data over time.

The cystic fibrosis transmembrane conductance regulator is widely distributed in the colon and plays a crucial function in controlling gut secretion viscosity and pH. As a result, Malagelada et al. [126] evaluated gut function using imaging methods often used to detect anatomical abnormalities of the digestive system. The authors also predicted that combining internal (endoluminal) and exterior (MRI) imaging gives a new viewpoint on the relationship between cystic fibrosis anatomical and functional findings. Zucker et al. [127] investigated if a deep CNN model might aid automated Brasfield rating of radiographic images of a chest for cystic fibrosis sufferer that would be analogous to that of a radiologist. Marques et al. [128] developed a texture classification-based technique to detect cystic fibrosis anomalies. The researchers used convolutional neural networks, as shown in Fig. 7, in two ways, in which the first one detected aberrant tissues, and the next determined the sort of structural abnormalities. The authors also presented a network that solely used patch-wise annotations to compute pixel-wise heatmaps of present irregularities.

Dio et al. [129] stated that physicians were given tools to recognize airway disorders in an automated and quick manner. Their goal was to produce a proof-of-concept and the first favorable results that might lead to rapid, reliable, and automatic detection of such disorders, similar to how the Sweat Chloride Test was used to diagnose Cystic Fibrosis (CF). Likewise, Zhou et al. [130] proposed a novel technique for cystic fibrosis diagnosis based on combining desorption electrospray ionization mass spectrometry with gradient boosted decision trees (calculated by Eqs. 12, 13) that divided the set of data into sub-sets, where S, X is a discrete random variable and P is a probability.

$$E(S) = \sum_{i=1}^{c} -\mathrm{p}_i log_2 p_i \qquad (12)$$

and entropy using the frequency table of two attributes is given as

$$E(S, X) = \sum_{c \in X} P(c)E(c) \qquad (13)$$

to analyze perspiration samples.

4.4 Role of AI to Predict Emphysema, Pneumoconiosis

Emphysema is a condition that causes difficulties in breathing and needs to be detected early with computed tomography scans and primary function testing. On the other hand, the challenges involved with specific diagnostic methods have prompted additional computer-assisted treatments to flourish. As a result, Mondal et al. [131] used deep learning networks to conduct automated pulmonary emphysema
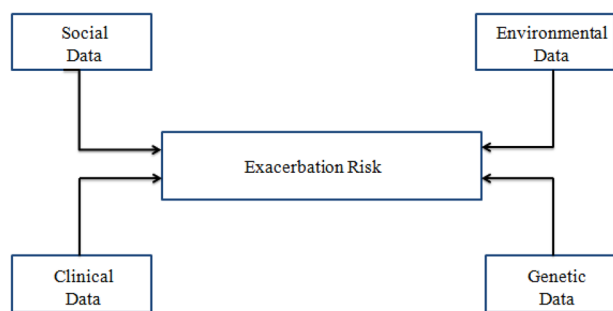
Fig. 6 Ways to predict asthma exacerbation patients

diagnosis, resulting in increased detection accuracy. Bortsova et al. [132] investigated a weakly labeled strategy comparable to multiple instance learning. Using this method, the authors measured emphysema by leveraging %age labels and applying existing information on the nature of these labels. The authors developed an approach that used a custom loss to learn the intervals and a label proportion problem-specific architecture (LPP). Deep learning has made significant progress in various complex tasks of processing images. In light of this, Humphries et al. [133] used the Fleischner technique to classify emphysema using chest CT image processing. The authors' primary goal was to see if emphysema patterns at the participant level may predict disability and death when identified using a deep learning algorithm. Srivastava et al. [134] employed neural networks to help medical experts diagnose Chronic Obstructive Pulmonary Disease by providing a complete and systematic assessment of clinical pulmonary audio data. The authors used ten splits of K-fold Cross-Validation (defined by Eq. 14) to escalate the presentation of the existing deep neural networks for N observations.

$$let k = \{1 \dots N\} \rightarrow \{1 \dots L\} \qquad (14)$$

which is an indexing function that indicates the division to which randomization assigns observation 'i'. The fitted function that is computed with the kth part of the removed data is denoted by $\hat{f}^{-k(i)}(x)$. The estimate prediction error using cross validation (CV) is shown by Eq. (15)

$$CV(f) = \frac{1}{N} \sum_{i=1}^{N} L(y_i, \hat{f}^{-k(i)}(x_i)) \qquad (15)$$

According to Nyboe et al. [135], automatic evaluation of emphysema presence might produce highly steady forecasts at a far cheaper cost, so it could be a viable alternative for an expert to do the assessments. As a result, they proposed a multiple instance learning (MIL) technique for emphysema detection. They also looked at whether emphysema at the scan level is sufficient to train the system to predict emphysema occurrence at the area level. Improving

Fig. 7 **Architecture of cascade approach for the detection of cystic fibrosis** [128]

the detection of extreme inactivity (EI) in COPD patients can reduce morbidity and death. Except for patients with evident EI, detecting such conduct in a real-life session is impossible. As a result, the authors such as Aguilaniu et al. [136] presented a machine learning technique to test for excessive inactivity. They created a prediction system that could accurately identify EI patients who would benefit the most from therapies like pulmonary rehabilitation. Gonzalez et al. [137] introduced a biomarker estimated approach based on deep learning techniques that relied on a regression network. In this system, the input was supplied to algorithm photos having the structure where the biomarker was generated, and the output was shown directly as the biomarker value. The two indicators used to demonstrate the suggested regression architectures were emphysema and bone mineral density (BMD), which are medically necessary. Pena et al. [138] described a technique for autonomously estimating emphysema areas in patients with chronic obstructive pulmonary disease (COPD) using High-Resolution Computed Tomography (HRCT) pictures that don't require manually labeled records for training. The authors wanted to use HRCT images without local annotations to detect emphysema sites in COPD patients.

Wang, X et al. [139] wanted to see if deep learning could be used to consider pneumoconiosis on digital chest radiographs as well as contrast its performance to that of qualified radiologists. The scientists used a conventional deep convolutional neural network to analyze chest X-ray pictures and verify them with various parameters. Furthermore, the scientists requested two trained radiologists to assess the testing dataset separately and compare their results to the computerized system. Hao et al. [65] presented an electronic health record-based pneumoconiosis radiography dataset. According to the authors, recent research has focused on machine learning for computer-aided detection. These have attained remarkable precision, with the artificial neural network (ANN) doing exceptionally well.

Nevertheless, wide use in clinical practice has been challenging due to unbalanced samples and a lack of readability. Hence to address such issues, the authors initially created a pneumoconiosis-based radiograph dataset, which included both un-favorable and favorable representatives. Secondly, deep convolutional diagnostic methodologies

were examined in identifying pneumoconiosis, and balanced training was used to improve recall. Pneumoconiosis diagnosis is based mainly on chest radiographic images, according to Wang et al. [140], and there is substantial disagreement across clinicians. Zhang et al. [141] set out to create an artificial intelligence (AI)-based model that would aid physicians in pneumoconiosis diagnosis as well as grading using chest radiographic images. The chest radiograph system was created with the help of a training cohort and validated with the help of an independent assessment cohort. Their groundbreaking research evaluated the possibility and effectiveness of AI-assisted radiography diagnosis and screening in the field of occupational lung disease. To surmount the issues of insufficient annotation pneumoconiosis data and increase the accuracy of pneumoconiosis diagnostics, Zheng et al. [142] presented two transfer learning models. They also demonstrated various pre-processing techniques for improving the quality and accuracy of X-rays, such as segmentation of lungs and amplification of data. Many computer-aided investigations on pneumoconiosis classification algorithms have been offered, according to Zhang et al. [143], but most of them were based on lung pictures. As a result, the authors provided a technique for diagnosing pneumoconiosis using wrist pulse signals gathered from non-pneumoconiosis individuals and pneumoconiosis patients. Machine learning approaches were utilized to process and assay the pulses of non-pneumoconiosis persons and the pneumoconiosis patient. Other than specialist radiologists, Wang et al. [144] noted a lack of sequential, automatic, and primary procedures for identifying and analyzing the evolution of pneumoconiosis in every coal miner. Consequently, the authors presented the most recent research findings from a study to address the challenges described by creating Computer-Aided Diagnosis (CAD) tools to identify pneumoconiosis using chest X-rays automatically.

4.5 Overall analysis

The best techniques have been filtered out based on their respective accuracy in Table 9 after observing the performance of the models that have been used for the detection and diagnosis of multiple airway diseases such as

pulmonary edema, pulmonary embolism, cystic fibrosis, pneumoconiosis, lung cancer, asthma, covid-19, Mesothelioma, tuberculosis, and emphysema.

5 Discussion

5.1 RQ 1: Year Wise Analysis of Predicting Multiple Airway Diseases Using AI Techniques

As demonstrated in Fig. 8, one hundred fifty-five (155) papers that range over 12 years, i.e. from (2010 to 2022)* have been studied for identifying different types of airway diseases using AI techniques which include pulmonary edema, cystic fibrosis, emphysema, Mesothelioma, pneumoconiosis, pulmonary embolism, lung cancer, covid19, tuberculosis, and asthma.

One hundred papers have been taken from recent years i.e. 2019 to 2022, and the remaining fifty-five papers have been taken from 2010 to 2018. This research intends to create a broader sense of the various types of AI and AI-derived techniques, such as machine and deep learning techniques for detecting and diagnosing multiple airway diseases.

- (As per the sources, the papers from the past ten years can be considered for SLR, but in this paper, we had to include some other information, such as a dataset of any particular disease, etc. for which we studied the papers atmost from the past 12 years.)

5.2 RQ 2: How Doctors are Being Helped by Deep and Machine Learning Techniques in Detecting the Airway Diseases?

The goal of an AI system is to organically learn a function and continue to improve without being actively taught. Machine learning and deep learning are artificial intelligence technologies that allow systems to detect patterns and connections between data and desired outputs [145]. These techniques connect multiple chunks of information discovered from the facts without requiring explicit human description [146]. At its most basic level, a machine and deep learning e-based strategy lead to a better diagnosis by analyzing a more excellent range of data than a physician.

We know that medical treatment is substantially affected by the increasing amount of healthcare data which makes it very difficult for pulmonologists to handle and analyze it for treating their patients manually. Hence, to improvise such a patient-doctor relationship, clinicians use AI techniques to predict, classify, and diagnose diseases efficiently. The stats have been shown in Fig. 9, which clearly defines the accuracy of AI techniques in predicting and analyzing

the diseases as compared to the traditional methods used by pulmonologists.

One advantage of employing AI to interpret diagnostic exams such as imaging is the ability to evaluate tests done in geographically isolated or underserved places [147]. This may result in a more early and accurate diagnosis, as well as a referral to expert treatment at an earlier stage of the disease, potentially influencing the prognosis Radiographs from these centers, on the other hand, can be remotely submitted and analyzed by a single central system using AI [148]. In Fig. 10, it has been shown that artificial intelligence techniques have successfully proven to be better than pulmonologists in predicting airway disorders.

5.3 RQ 3: Which ML and DL Techniques are Broadly Applied to Predict Airway Diseases?

Patients who require early diagnosis and treatment can benefit from AI-driven sickness detection models that help medical companies produce improved diagnostic tools. As a result, the literature mentions the techniques for diagnosing airway disorders such as cystic fibrosis, pneumoconiosis, pulmonary edema, pulmonary embolism, covid-19, emphysema, tuberculosis, lung cancer, asthma, and mesothelioma. Linear regression, support vector machine (SVM), random forest (RF), decision tree (DT), nave Bayes (NB), fuzzy particle, logistic regression, and ensemble learning are the most commonly used machine learning models in the literature. Convolutional Neural Networks (CNN) are the most often utilized deep learning models for illness diagnosis. Furthermore, artificial neural networks (ANN) and transfer learning techniques such as VGGNet, ResNet, and others have been extensively used as shown in Fig. 11.

5.4 RQ 4: Name the Characteristics that Manipulate the Quality of Prediction Models Based on Deep and Machine Learning?

AI algorithms in the medical industry are essential, notably for identifying a disease from the medical database. Many firms employ these approaches to predict illnesses early and better medical diagnostics. Irrespective of their continual progress, there are specific problems that remain. On the one hand, machine and deep learning algorithms can handle any complicated situation, but on the other hand, they also demand more research effort for practical implementations [149]. Hence, certain limitations occur while utilizing machine and deep learning models to diagnose airway illnesses gathered from research gaps and described below.

Data Paucity One of the most pervasive issues in artificial intelligence is a shortage of high-quality data. Every business

Table 9 Overall comparison of models

Ref	Diseases	Techniques	Accuracy
[87]	Pulmonary Embolism	Deep Learning, Image Classification, Long Short Term Model	91%
[82]	Pulmonary Edema	Principal Component Analysis, Random Forest	96.5%
[129]	Cystic fibrosis	Support vector machine, Naïve Bayes classifier, logistic regression	99%
[141]	Pneumoconiosis	Convolution Neural Network	97.3%
[111]	Lung cancer	Ensemble classifier	97.6%
[114]	Asthma	Decision Tree, Naive Bayes, K Nearest Neighbour	96.52%
[96]	Covid-19	Mamta Ray Foraging Optimization, Fractional Multichannel Exponent Moments	98.09%
[105]	Mesothelioma	SMOTE, ADASYN, Artificial neural network, principal component analysis	96%
[118]	Tuberculosis	CLAHE method, Deep convolution neural network, UNet architecture	97.1%
[131]	Emphysema	Improved red deer algorithm, Fuzzy C Means, Adaptive local ternary pattern	95.56%

will encounter this difficulty throughout the AI deployment process. Thus, in the future, open-source datasets of multiple airway diseases should be considered, though they occasionally lack quality but represent promising solutions for organizations [150]. In addition, synthetic data should be created to increase data security and privacy; data augmentation should be considered to enhance the size of the dataset without accumulating additional data; and finally, transfer learning techniques should be used when we have enough training data. *Modeling errors:* Two main restricting errors are overfitting and underfitting. This is called overfitting, when a model learns the information and noise in the training dataset to the point that it degrades its performance on a new dataset. In underfitting, models cannot train or generalize new datasets, thus impairing the system's performance [151]. Therefore, algorithms like pre trained models should be incorporated that work on system modeling errors in the future. *Perpetual improvement of models:* At times, optimizing a model's performance might be difficult. This is because optimizing a model's performance makes it more accurate in predicting and the most dependable and acceptable in artificial intelligence. Developing an AI-based model is not difficult for engineers while verifying its performance is critical to obtaining accurate and trustworthy results [152]. Hence to improve it, in the future, the appropriate amount of data should be utilized, the proper algorithms should be used, and models should be verified and evaluated appropriately. Transfer learning can also enhance the performance of AI-based models [153]. *Class imbalance:* An unbalanced classification issue is when the distribution of instances across recognized classes is uneven or biased [154]. Imbalanced classifications provide a problem to predict the class because most machine learning methods for classification are built on the premise of an equal number of instances for each class. So, in the future, Re-sampling, K fold cross-validation, or pre-trained models can be used to deal with class imbalance issues [155].

6 Conclusion and Future Work

Between 2010 and 2022, 155 studies were chosen from six digital libraries that can be accessed online, and four questions were investigated after going through them. Researchers analyzed a variety of technical developments that might be used to enhance AI-based models in the field of pulmonology. This article discusses the influence of machine learning and deep learning techniques on analyzing the information related to airway disorders. Furthermore, the paper details several academics' rigorous research efforts to demonstrate how machine learning and deep learning models help to detect or categorize various airway illnesses. The literature's systematic review is presented in a tabular format, with each column indicating the dataset, methodology, and outcomes used by the researchers and their limitations. After accumulating the constraints encountered by researchers for the prediction of airway problems, an attempt has been made to determine the usage of the latest models that may be incorporated in the future to enhance the system's performance. The lasso method, CNN, Decision Tree, GAP Net, and other algorithms were examined by certain researchers. However, their findings revealed that the models could not distinguish between diseases or detect anomalies in their data. Aside from that, several models only worked with limited data sets, failed to pre-process data, and refused to provide localization information for the final image. The model's overall performance is hampered due to these limitations, which must be addressed. Several approaches used in research publications, such as random forest and logistic regression, have the lowest prediction accuracy because of modeling flaws such as overfitting and underfitting. Many researchers have also struggled to categorize data using DFD-Net, Fuzzy Particle effectively, Swarm Optimization to detect lung cancer, CNN to detect pneumoconiosis, Dense Net, Inception V3, ResNet, and DNN to diagnose
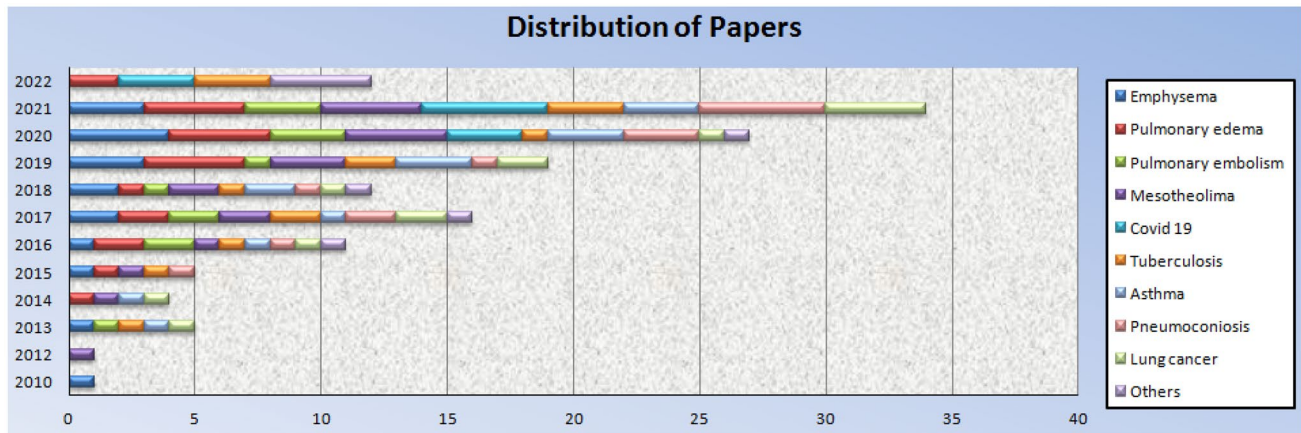
Fig. 8 Distribution of papers

**Fig. 9** Analysis between Pulmonologists and AI techniques in predicting the diseases
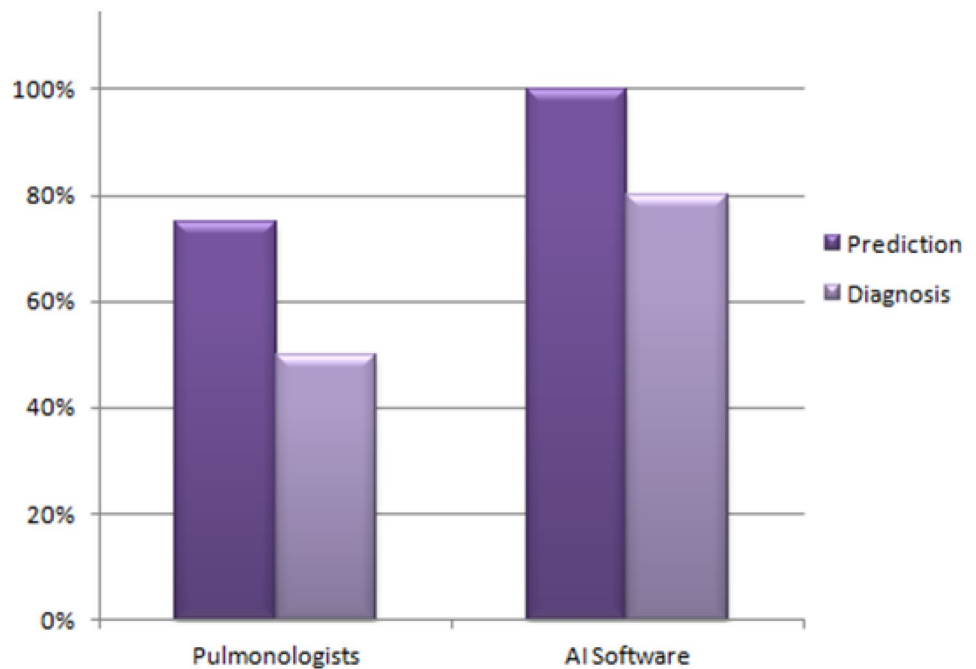
pulmonary embolism, and Canny Edge to detect tuberculosis. It's also worth noting that researchers have only focused on one or two airway disorders for prediction, which limits users' capacity to distinguish between other types of airway difficulties.

Although artificial technology has innumerable benefits, its flaws and limitations may restrict its applications, notably in the healthcare business. But there is also scope for improvement in this field to overcome the issues such as the best selection of models which can be further improved or innovated to provide better output, and optimization techniques should be added to the network to provide the optimal results. The most important task while using the learning models is correctly classifying the data. Hence, such a methodology that connects the data's features to improve the classification model and generates correct output should be included. Besides this, deep learning techniques should be improved to be applied to a dataset where they can detect various airway diseases using less processing time. In addition, the conventional learning methods should also be considered in the future as they can be used to improve the airway disease detection model.

Fig. 10 Airway Diseases Prediction by AI and Pulmonologists
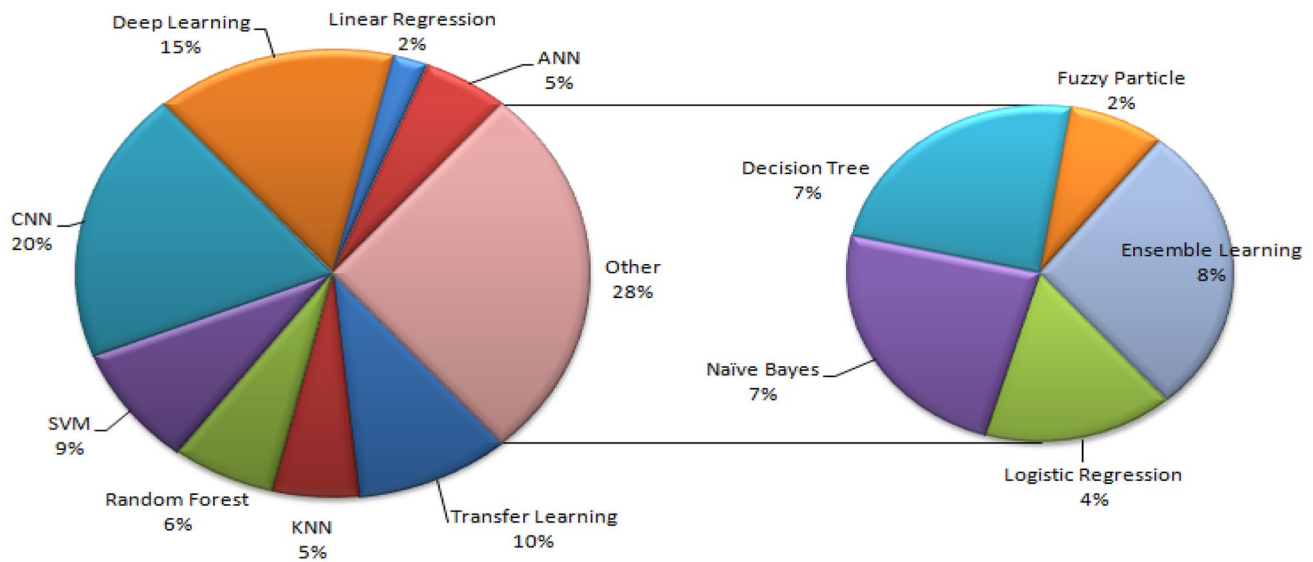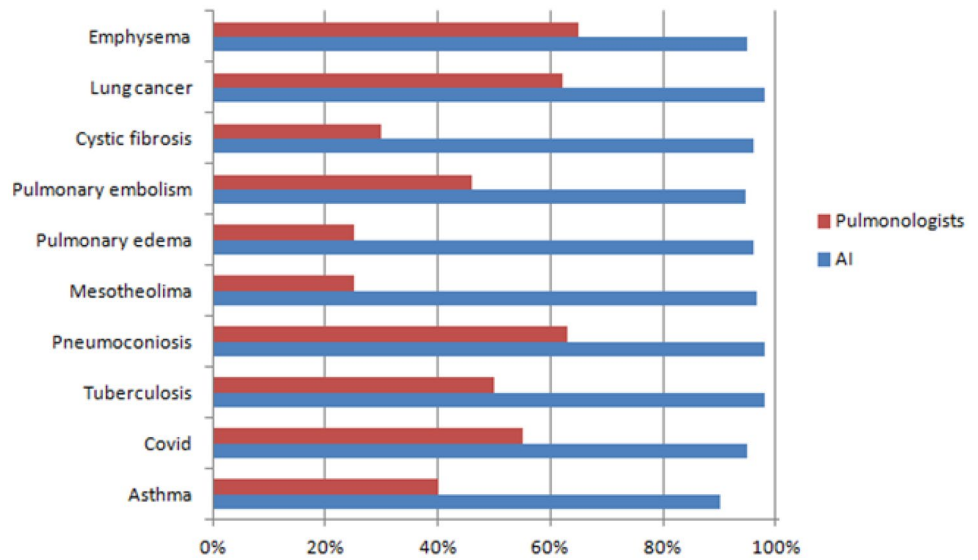
Fig. 11 Machine and deep learning based prediction models

Declarations

Conflict of interest The authors declare no conflict of interest.

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://doi.org/10.1007/978-3-030-97929-4_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/10.1586/17476348.2016.1155455
4. CF Foundation Estimates Increase in CF Population (2022) Cystic Fibrosis Foundation. https://www.cff.org/news/2022-07/cf-foundation-estimates-increase-in-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://doi.org/10.1016/j.ijid.2021.02.107
6. Siegel RL, Miller KD, Jemal A (2019) Cancer statistics, 2019. CA: Cancer J Clin 69(1):7–34. https://doi.org/10.3322/caac.21708

7. Nascimento Maia P et al (2022) Correlation of digital flow peak with spirometry in children with and without asthma. J Asthma. https://doi.org/10.1080/02770903.2022.2045308

8. 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://doi.org/10.1016/j.dsx.2020.04.012

9. 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/10.1183/23120541.00492-2021

10. 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/jor2010004

11. 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/baaa010

12. Ardila D, Kiraly AP, Bharadwaj S et al (2019) End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med 25:954–961. https://doi.org/10.1038/s41591-019-0447-x

13. 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

14. Sandvik RM, Gustafsson PM, Lindblad A, Buchvald F, Olesen HV, Olsen JH, Nielsen KG (2022) Contemporary N2 and SF6 multiple breath washout in infants and toddlers with cystic fibrosis. Pediatr Pulmonol 57(4):945–955. https://doi.org/10.1002/ppul.25830

15. Puttur F, Gregory LG, Lloyd CM (2019) Airway macrophages as the guardians of tissue repair in the lung. Immunol Cell Biol 97(3):246–257. https://doi.org/10.1111/imcb.12235

16. Ahmed ST, Kadhem SM (2021) Using machine learning via deep learning algorithms to diagnose the lung disease based on chest imaging: a survey. Int J Interact Mob Technol 15(16):10. https://doi.org/10.3991/ijim.v15i16.24191

17. Peng Y, Liu E, Peng S, Chen Q, Li D, Lian D (2022) Using artificial intelligence technology to fight COVID-19: a review. Artif Intell Rev. https://doi.org/10.1007/s10462-021-10106-z

18. Mirbabaie M, Stieglitz S, Frick NR (2021) Artificial intelligence in disease diagnostics: a critical review and classification on the current state of research guiding future direction. Heal Technol. https://doi.org/10.1007/s12553-021-00555-5

19. Davenport T, Kalakota R (2019) The potential for artificial intelligence in healthcare. Fut Healthc J 6(2):94–98. https://doi.org/10.7861/futurehosp.6-2-94

20. Kaul S, Kumar Y (2020) Artificial intelligence-based learning techniques for diabetes prediction: challenges and systematic review. SN Comput Sci 1(6):1–7. https://doi.org/10.1007/s42979-020-00337-2

21. Wong A, Lu J, Dorfman A, McInnis P, Famouri M, Manary D, et al (2021) Fibrosis-net: a tailored deep convolutional neural network design for prediction of pulmonary fibrosis progression from chest CT images. arXiv:2103.04008. https://doi.org/10.48550/arXiv.2103.04008

22. García-Caballero A, de Dios Caballero J, Maruri A, Serrano-Tomás MI, Del Campo R, Morosini MI, Cantón R (2022) Evaluation of different phenotypic methods to detect methicillin resistance in *Staphylococcus aureus* isolates recovered from cystic fibrosis patients. Diagn Microbiol Infect Dis. https://doi.org/10.1016/j.diagmicrobio.2021.115559

23. Khare S, Desimone M, Kasim N, Chan CL (2022) Cystic fibrosis-related diabetes: prevalence, screening, and diagnosis. J Clin Transl Endocrinol. https://doi.org/10.1016/j.jcte.2021.100290

24. Wang P, Shu W, Yin F, Zhang Q, Yin Y, Deng Q et al (2022). Bioinformatic analyses of potential key genes in the pathogenesis of idiopathic pulmonary fibrosis. https://doi.org/10.21203/rs.3.rs-1646754/v1

25. Ashkenazi M, Sity S, Sarouk I, Aluma BEB, Dagan A, Bezalel Y et al (2018) Omalizumab in allergic bronchopulmonary aspergillosis in patients with cystic fibrosis. JAA 11:101. https://doi.org/10.2147/JAA.S156049

26. Curation D (2021) The RSNA pulmonary embolism CT dataset. Radiology. https://doi.org/10.1148/ryai.2021200254

27. Masoudi M et al (2018) A new dataset of computed-tomography angiography images for computer-aided detection of pulmonary embolism. Sci Data. https://doi.org/10.1038/sdata.2018.178

28. Matthews A, Bhaskaran K (2018) Clinical code list: pulmonary embolism. [Data Collection]. London School of Hygiene & Tropical Medicine, London. https://doi.org/10.17037/DATA.00000734

29. Weifang L, Liu M, Xiaojuan G, Peiyao Z, Zhang L, Rongguo Z, et al (2020) Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning. Eur Radiol 30(6):3567–3575. https://doi.org/10.1007/s00330-020-06699-8

30. González G, Jimenez-Carretero D, Rodríguez-López S, Cano-Espinosa C, Cazorla M, Agarwal T et al (2020) Computer aided detection for pulmonary embolism challenge (CAD-PE). arXiv:2003.13440. https://doi.org/10.1007/978-3-540-73273-0_52

31. Luo G, Johnson MD, Nkoy FL, He S, Stone BL (2020) Automatically explaining machine learning prediction results on asthma hospital visits in patients with asthma: secondary analysis. JMIR Med Inform. https://doi.org/10.2196/21965

32. Wu FY, Liu JF (2019) Asthma medication increases dental caries among children in Taiwan: an analysis using the National Health Insurance Research Database. J Dent Sci 14(4):413–418. https://doi.org/10.1016/j.jds.2019.08.002

33. Larose DT, Larose CD (2014) Discovering knowledge in data: an introduction to data mining (vol 4). Wiley. https://www.stevens.edu/sites/stevens_edu/files/CareCenter/UTC/Discovering_Knowledge_in_Data.pdf

34. Garcia E, Urman R, Berhane K, McConnell R, Gilliland F (2019) Effects of policy-driven hypothetical air pollutant interventions on childhood asthma incidence in southern California. Proc Natl Acad Sci 116(32):15883–15888. https://doi.org/10.1073/pnas.1815678116

35. Li Y, Liu X, Chen H, Xie P, Ma R, He J, Zhang H (2021) Bioinformatics analysis for the role of CALR in human cancers. PLoS ONE. https://doi.org/10.1371/journal.pone.0261254

36. Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F (2013) The cancer imaging archive (TCIA): maintaining and operating a public information repository. J Digit Imaging 26(6):1045–1057. https://doi.org/10.1007/s10278-013-9622-7

37. Faisal MI, Bashir S, Khan ZS, Khan FH (2018) An evaluation of machine learning classifiers and ensembles for early stage prediction of lung cancer. In: 2018 3rd international conference on emerging trends in engineering, sciences and technology (ICEEST) (pp. 1–4). IEEE. https://doi.org/10.1109/ICEEST.2018.8643311

38. Heuvelmans MA, van Ooijen PM, Ather S, Silva CF, Han D, Heussel CP, et al (2021) Lung cancer prediction by deep learning to identify benign lung nodules. Lung Cancer 154:1–4. https://doi.org/10.1016/j.lungcan.2021.01.027

39. Tilve A, Nayak S, Vernekar S, Turi D, Shetgaonkar PR, Aswale S (2020) Pneumonia detection using deep learning approaches. In: 2020 international conference on emerging trends in information

technology and engineering (ic-ETITE) (pp 1–8). IEEE. https://doi.org/10.1109/ic-ETITE47903.2020.152

40. Angeline R, Kanna SN, Menon NG, Ashwath B (2022) Identifying malignancy of lung cancer using deep learning concepts. In: Artificial intelligence in healthcare (pp. 35–46). Springer, Singapore. https://doi.org/10.1007/978-981-16-6265-2_3

41. Dey SK, Rahman MM, Siddiqi UR, Howlader A (2020) Analyzing the epidemiological outbreak of COVID-19: a visual exploratory data analysis approach. J Med Virol 92(6):632–638. https://doi.org/10.1002/jmv.25743

42. Talukdar D, Tripathi V (2021) COVID-19 forecast for 13 Caribbean countries using ARIMA modeling for confirmed, death, and recovered cases. Research 10:1068. https://doi.org/10.12688/f1000research.73746.1

43. Maguolo G, Nanni L (2021) A critic evaluation of methods for covid-19 automatic detection from X-ray images. Inf Fusion 76:1–7. https://doi.org/10.1016/j.inffus.2021.04.008

44. Rahman T, Khandakar A, Qiblawey Y, Tahir A, Kiranyaz S, Kashem SBA, et al (2021) Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. Comput Biol Med. https://doi.org/10.1016/j.compbiomed.2021.104319

45. Jaeger S, Candemir S, Antani S, Wáng YX, Lu PX, Thoma G (2014) Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. Quant Imaging Med Surg 4(6):475–477. https://doi.org/10.3978/j.issn.2223-4292.2014.11.20

46. Rahman T, Khandakar A, Kadir MA, Islam KR, Islam KF, Mahbub ZB, Ayari MA, Chowdhury MEH (2020) Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. IEEE Access 8:191586–191601. https://doi.org/10.1109/ACCESS.2020.3031384

47. Candemir S, Antani S (2019) A review on lung boundary detection in chest X-rays. Int J Comput Assist Radiol Surg 14(4):563–576. https://doi.org/10.1007/s11548-019-01917-1

48. Sun J, Chong P, Tan YXM, Binder A (2017) ImageCLEF 2017: imageCLEF tuberculosis task-the SGEast submission. In: CLEF (working notes). http://ceur-ws.org/Vol-1866/paper_130.pdf

49. Bobak CA, Titus AJ, Hill JE (2019) Comparison of common machine learning models for classification of tuberculosis using transcriptional biomarkers from integrated datasets. Appl Soft Comput 74:264–273. https://doi.org/10.1016/j.asoc.2018.10.005

50. Zarrin PS, Roeckendorf N, Wenger C (2020) In-vitro classification of saliva samples of COPD patients and healthy controls using machine learning tools. IEEE Access 8:168053–168060. https://doi.org/10.1109/ACCESS.2020.3023971

51. Suzuki M, Cole JJ, Konno S, Makita H, Kimura H, Nishimura M, Maciewicz RA (2021) Large-scale plasma proteomics can reveal distinct endotypes in chronic obstructive pulmonary disease and severe asthma. Clin Transl Allergy. https://doi.org/10.1002/clt2.12091

52. Sakornsakolpat P, Morrow JD, Castaldi PJ, Hersh CP, Bossé Y, Silverman EK, et al (2018) Integrative genomics identifies new genes associated with severe COPD and emphysema. Respir Res 19(1):1–13. https://doi.org/10.1186/s12931-018-0744-9

53. Rasmussen JF, Siersma V, Malmqvist J, Brodersen J (2020) Psychosocial consequences of false positives in the Danish Lung Cancer CT Screening Trial: a nested matched cohort study. BMJ Open. https://doi.org/10.1136/bmjopen-2019-034682

54. Sorensen L, Shaker SB, De Bruijne M (2010) Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE Trans Med Imaging 29(2):559–569. https://doi.org/10.1109/TMI.2009.2038575

55. Er O, Tanrikulu AC, Abakay A, Temurtas F (2012) An approach based on probabilistic neural network for diagnosis of Mesothelioma's disease. Comput Electr Eng 38(1):75–81. https://doi.org/10.1016/j.compeleceng.2011.09.001

56. Xie D, Luo X (2021) Identification of four methylation-driven genes as candidate biomarkers for monitoring single-walled carbon nanotube-induced malignant transformation of the lung. Toxicol Appl Pharmacol. https://doi.org/10.1016/j.taap.2020.115391

57. Roy S, Galateau-Sallé F, Le Stang N, Churg A, Lyons MA, Attanoos R, Dacic S (2021) Molecular characterization of pleomorphic mesothelioma: a multi-institutional study. Mod Pathol. https://doi.org/10.1038/s41379-021-00900-z

58. Churg A, Attanoos R, Borczuk AC, Chirieac LR, Galateau-Salle F, Gibbs A, Henderson D, Roggli V, Rusch V, Judge MJ, Srigley JR (2017). Mesothelioma in the pleura and peritoneum histopathology reporting guide 2nd edition. International Collaboration on Cancer Reporting; Sydney. ISBN: 978-1-925687-03-3. https://www.iccr-cancer.org/datasets/published-datasets/thorax/mesothelioma/

59. Churg A, Nabeshima K, Ali G, Bruno R, Fernandez-Cuesta L, Galateau-Salle F (2018) Pathologic separation of benign from malignant mesothelial proliferations and histologic/molecular analysis of malignant mesothelioma subtypes. Lung Cancer 124:95–101. https://doi.org/10.1016/j.lungcan.2018.07.041

60. Liao R, Chauhan G, Golland P, Berkowitz S, Horng S (2021) Pulmonary edema severity grades based on MIMIC-CXR (version 1.0.1). PhysioNet. https://doi.org/10.13026/rz5p-rc64

61. Johnson AE, Pollard TJ, Berkowitz SJ, Greenbaum NR, Lungren MP, Deng CY, et al (2019) MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. Sci Data 6(1):1–8. https://doi.org/10.1038/s41597-019-0322-0

62. Wu JX, Chen PY, Li CM, Kuo YC, Pai NS, Lin CH (2020) Multilayer fractional-order machine vision classifier for rapid typical lung diseases screening on digital chest X-ray images. IEEE Access 8:105886–105902. https://doi.org/10.1109/ACCESS.2020.3000186

63. Horng S, Liao R, Wang X, Dalal S, Golland P, Berkowitz SJ (2021) Deep learning to quantify pulmonary edema in chest radiographs. Radiology. https://doi.org/10.1148/ryai.2021190228

64. Singh R, Kalra MK, Nitiwarangkul C, Patti JA, Homayounieh F, Padole A, et al (2018) Deep learning in chest radiography: detection of findings and presence of change. PLoS ONE. https://doi.org/10.1371/journal.pone.0204155

65. Hao C, Jin N, Qiu C, Ba K, Wang X, Zhang H, et al (2021) Balanced convolutional neural networks for pneumoconiosis detection. Int J Environ Res Public Health 18(17):9091. https://doi.org/10.3390/ijerph18179091

66. Halldin CN, Hale JM, Weissman DN, Attfield MD, Parker JE, Petsonk EL, et al (2019) The national institute for occupational safety and health B reader certification program—an update report (1987 to 2018) and future directions. J Occup Environ Med 61(12):1045. https://doi.org/10.1097/JOM.0000000000001735

67. Şener MU, Şimşek C, Özkara Ş, Evran H, Bursali I, Gökçek A (2019) Comparison of the International classification of high-resolution computed tomography for occupational and environmental respiratory diseases with the International labor organization International classification of radiographs of pneumoconiosis. Ind Health 57(4):495–502. https://doi.org/10.2486/indhealth.2018-0068

68. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2017) Chestx-ray8: hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: Proceedings of the IEEE conference on computer vision and pattern recognition (pp 2097–2106). https://doi.org/10.48550/arXiv.1705.02315

69. Kumar D, Wong A, Clausi DA (2015) Lung nodule classification using deep features in CT images. In: 2015 12th conference on computer and robot vision (pp 133–138). IEEE. https://doi.org/10.1109/CRV.2015.25

70. Bhatia S, Sinha Y, Goel L (2019). Lung cancer detection: a deep learning approach. In: Soft computing for problem solving (pp 699–705). Springer, Singapore. https://doi.org/10.1007/978-981-13-1595-4_55

71. Kohli R, Garg A, Phutela S, Kumar Y, Jain S (2021) An improvised model for securing cloud-based E-healthcare systems. In: IoT in healthcare and ambient assisted living (pp 293–310). Springer. https://doi.org/10.1007/978-981-15-9897-5_14

72. Wang X, Schwab E, Rubin J, Klassen P, Liao R, Berkowitz S et al (2019) Pulmonary edema severity estimation in chest radiographs using deep learning. In: International conference on medical imaging with deep learning--extended abstract track. https://openreview.net/forum?id=rygZBfCVqE

73. Raso R, Tartarisco G, Matucci Cerinic M, Pioggia G, Picano E, Gargani L (2015) A soft computing–based B-line analysis for objective classification of severity of pulmonary edema and fibrosis. JACC Cardiovasc Imaging 8(4):495–496. https://doi.org/10.1016/j.jcmg.2014.06.022

74. Yang F, Tang ZR, Chen J, Tang M, Wang S, Qi W, et al (2021) Pneumoconiosis computer aided diagnosis system based on X-rays and deep learning. BMC Med Imaging 21(1):1–7. https://doi.org/10.1186/s12880-021-00723-z

75. Naik A, Edla DR (2021) Lung nodule classification on computed tomography images using deep learning. Wireless Pers Commun 116(1):655–690. https://doi.org/10.1007/s11277-020-07732-1

76. Blain M, Kassin MT, Varble N, Wang X, Xu Z, Xu D, et al (2021) Determination of disease severity in COVID-19 patients using deep learning in chest X-ray images. Diagn Interv Radiol 27(1):20. https://doi.org/10.5152/dir.2020.20205

77. Kim Y, Hyon Y, Jung SS, Lee S, Yoo G, Chung C, Ha T (2021) Respiratory sound classification for crackles, wheezes, and rhonchi in the clinical field using deep learning. Sci Rep 11(1):1–11. https://doi.org/10.1038/s41598-021-96724-7

78. Goralski JL, Stewart NJ, Woods JC (2021) Novel imaging techniques for cystic fibrosis lung disease. Pediatr Pulmonol 56:S40–S54. https://doi.org/10.1002/ppul.24931

79. Islam MT, Aowal MA, Minhaz AT, Ashraf K (2017) Abnormality detection and localization in chest X-rays using deep convolutional neural networks. arXiv:1705.09850. http://arxiv.org/abs/1705.09850

80. Liao R, Rubin J, Lam G, Berkowitz S, Dalal S, Wells W et al (2019) Semi-supervised learning for quantification of pulmonary edema in chest X-ray images. arXiv:1902.10785. http://arxiv.org/abs/1902.10785

81. Hong KJ, Essid S, Ser W, Foo DCG (2018) A robust audio classification system for detecting pulmonary edema. Biomed Signal Process Control 46:94–103. https://doi.org/10.1016/j.bspc.2018.07.004

82. Lin H, Luo Y, Sun Q, Deng K, Chen Y, Wang Z, Huang P (2020) Determination of causes of death via spectrochemical analysis of forensic autopsies-based pulmonary edema fluid samples with deep learning algorithm. J Biophotonics 13(4):1–13. https://doi.org/10.1002/jbio.201960144

83. Kumar A, Wang Y, et.al (2014) Pulmonary edema detection in chest X-ray using machine learning. Int J Clin Med Images 1(7). https://www.imagejournals.org/articles/pulmonary-edema-detection-in-chest-xray-using-machine-learning-144.html

84. Kumar A, Wang YY, Liu KC, Tsai IC, Huang CC, Hung N (2014) Distinguishing normal and pulmonary edema chest X-ray using Gabor filter and SVM. In: 2014 IEEE international symposium on bioelectronics and bioinformatics (IEEE ISBB 2014) (pp 1–4). IEEE. https://doi.org/10.1109/ISBB.2014.6820918

85. Hayat C (2021) DenseNet-CNN architectural model for detection of abnormality in acute pulmonary edema. Khazanah Informatika Jurnal Ilmu Komputer Dan Informatika 7(2):73–79. https://doi.org/10.23917/khif.v7i2.13455

86. Brestel C, Cohen-sfaty M (2018) RadBot-CXR : classification of four clinical finding categories in chest X-ray using deep learning. Midl 1–9. https://openreview.net/forum?id=B1tMhcIDM

87. Kiourt C, Feretzakis G. Dalamarinis K, Kalles D, Pantos G, Papadopoulos I et al (2021) Pulmonary embolism identification in computerized tomography pulmonary angiography scans with deep learning technologies in COVID-19 patients. arXiv:2105.11187. https://doi.org/10.48550/arXiv.2105.11187

88. Weikert T, Winkel DJ, Bremerich J, Stieltjes B, Parmar V, Sauter AW, Sommer G (2020) Automated detection of pulmonary embolism in CT pulmonary angiograms using an AI-powered algorithm. Eur Radiol 30(12):6545–6553. https://doi.org/10.1007/s00330-020-06998-0

89. Liu W, Liu M, Guo X, Zhang P, Zhang L, Zhang R, Kang H, Zhai Z, Tao X, Wan J, Xie S (2020) Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning. Eur Radiol 30(6):3567–3575. https://doi.org/10.1007/s00330-020-06699-8

90. Rucco M, Sousa-Rodrigues D, Merelli E, Johnson JH, Falsetti L, Nitti C, Salvi A (2015) Neural hypernetwork approach for pulmonary embolism diagnosis. BMC Res Notes 8(1):1–11. https://doi.org/10.1186/s13104-015-1554-5

91. Chen JI-Z (2021) Design of accurate classification of COVID-19 disease in X-ray images using deep learning approach. J ISMAC 2(2):132–148. https://doi.org/10.36548/jismac.2021.2.006

92. Ismael AM, Şengür A (2021) Deep learning approaches for COVID-19 detection based on chest X-ray images. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2020.114054

93. Punn NS, Sonbhadra SK, Agarwal S (2020) COVID-19 epidemic analysis using machine learning and deep learning algorithms. MedRxiv. https://doi.org/10.1101/2020.04.08.20057679

94. Cabitza F, Campagner A, Ferrari D, Di Resta C, Ceriotti D, Sabetta E, Colombini A, De Vecchi E, Banfi G, Locatelli M, Carobene A (2021) Development, evaluation, and validation of machine learning models for COVID-19 detection based on routine blood tests. Clin Chem Lab Med 59(2):421–431. https://doi.org/10.1515/cclm-2020-1294

95. Chang T, Wu J, Chang L (2020) Imported SARS-CoV-2 V501Y. V2 variant (B.1.351) detected in travelers from South Africa and Tanzania to India. Travel Med Infect Dis 10:1–4. https://doi.org/10.1016/j.tmaid.2021.102023

96. Elaziz MA, Hosny KM, Salah A, Darwish MM, Lu S, Sahlol AT (2020) New machine learning method for image-based diagnosis of COVID-19. PLoS ONE. https://doi.org/10.1371/journal.pone.0235187

97. Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J, et al (2021) A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). Eur Radiol 31(8):6096–6104. https://doi.org/10.1101/2020.02.14.20023028

98. Li X, Wang X, Yang X, Lin Y, Huang Z (2021) Preliminary study on artificial intelligence diagnosis of pulmonary embolism based on computer in-depth study. Ann Transl Med 9(10):838–838. https://doi.org/10.21037/atm-21-975

99. Schmuelling L, Franzeck FC, Nickel CH, Mansella G, Bingisser R, Schmidt N, Stieltjes B, Bremerich J, Sauter AW, Weikert T, Sommer G (2021) Deep learning-based automated detection of pulmonary embolism on CT pulmonary angiograms: No significant effects on report communication times and patient turnaround in the emergency department nine months after technical implementation. Eur J Radiol. https://doi.org/10.1016/j.ejrad.2021.109816

100. Huang SC, Kothari T, Banerjee I, Chute C, Ball RL, Borus N, Huang A, Patel BN, Rajpurkar P, Irvin J, Dunnmon J, Bledsoe J, Shpanskaya K, Dhaliwal A, Zamanian R, Ng AY, Lungren MP (2020) PENet—a scalable deep-learning model for automated diagnosis of pulmonary embolism using volumetric CT imaging. NPJ Digital Med 3(1):1–9. https://doi.org/10.1038/s41746-020-0266-y

101. Alam TM, Shaukat K, Hameed IA, Khan WA, Sarwar MU, Iqbal F, Luo S (2021) A novel framework for prognostic factors identification of malignant mesothelioma through association rule mining. Biomed Signal Process Control. https://doi.org/10.1016/j.bspc.2021.102726

102. Latif MZ, Shaukat K, Luo S, Hameed IA, Iqbal F, Alam TM (2020) Risk factors identification of malignant mesothelioma: a data mining based approach. In: 2nd international conference on electrical, communication and computer engineering, ICECCE 2020, June, 12–13. https://doi.org/10.1109/ICECCE49384.2020.9179443

103. Choudhury A (2021) Predicting cancer using supervised machine learning: mesothelioma. Technol Health Care 29(1):45–58. https://doi.org/10.3233/THC-202237

104. Alam TM, Shaukat K, Mahboob H, Sarwar MU, Iqbal F, Nasir A, Hameed IA, Luo S (2021) A machine learning approach for identification of malignant mesothelioma etiological factors in an imbalanced dataset. Comput J. https://doi.org/10.1093/comjnl/bxab015

105. Gupta S, Gupta MK, Kumar R (2021) A novel multi-neural ensemble approach for cancer diagnosis. Appl Artif Intell. https://doi.org/10.1080/08839514.2021.2018182

106. Courtiol P, Maussion C, Moarii M, Pronier E, Pilcer S, Sefta M, Manceron P, Toldo S, Zaslavskiy M, Le Stang N, Girard N, Elemento O, Nicholson AG, Blay JY, Galateau-Sallé F, Wainrib G, Clozel T (2019) Deep learning-based classification of mesothelioma improves prediction of patient outcome. Nat Med 25(10):1519–1525. https://doi.org/10.1038/s41591-019-0583-3

107. Sori WJ, Feng J, Godana AW, Liu S, Gelmecha DJ (2021) DFD-Net: lung cancer detection from denoised CT scan image using deep learning. Front Comput Sci. https://doi.org/10.1007/s11704-020-9050-z

108. Asuntha A, Srinivasan A (2020) Deep learning for lung cancer detection and classification. Multimed Tools Appl 79(11–12):7731–7762. https://doi.org/10.1007/s11042-019-08394-3

109. Xie Y, Meng WY, Li RZ, Wang YW, Qian X, Chan C, Yu ZF, Fan XX, Pan HD, Xie C, Wu QB, Yan PY, Liu L, Tang YJ, Yao XJ, Wang MF, Leung ELH (2021) Early lung cancer diagnostic biomarker discovery by machine learning methods. Transl Oncol. https://doi.org/10.1016/j.tranon.2020.100907

110. Ausawalaithong, W., Thirach, A., Marukatat, S., & Wilaiprasitporn, T. (2019). Automatic Lung Cancer Prediction from Chest X-ray Images Using the Deep Learning Approach. *BMEiCON 2018 - 11th Biomedical Engineering International Conference.* https://doi.org/10.1109/BMEiCON.2018.8609997

111. Shakeel PM, Burhanuddin MA, Desa MI (2020) Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. Neural Comput Appl. https://doi.org/10.1007/s00521-020-04842-6

112. Pradhan K, Chawla P (2020) Medical Internet of things using machine learning algorithms for lung cancer detection. J Manag Anal 7(4):591–623. https://doi.org/10.1080/23270012.2020.1811789

113. Kadir T, Gleeson F (2018) Lung cancer prediction using machine learning and advanced imaging techniques. Transl Lung Cancer Res 7(3):304. https://doi.org/10.21037/tlcr.2018.05.15

114. Awal MA, Hossain MS, Debjit K, Ahmed N, Nath RD, Habib GMM, Khan MS, Islam MA, Mahmud MAP (2021) An early

115. Khasha R, Sepehri MM, Taherkhani N (2021) Detecting asthma control level using feature-based time series classification. Appl Soft Comput. https://doi.org/10.1016/j.asoc.2021.107694

116. Zhang O, Minku LL, Gonem S (2021) Detecting asthma exacerbations using daily home monitoring and machine learning. J Asthma 58(11):1518–1527. https://doi.org/10.1080/02770903.2020.1802746

117. Zhan J, Chen W, Cheng L, Wang Q, Han F, Cui Y (2020) Diagnosis of asthma based on routine blood biomarkers using machine learning. Comput Intell Neurosci. https://doi.org/10.1155/2020/8841002

118. Dasanayaka C, Dissanayake MB, Dasanayaka C, Dissanayake MB (2020) Computer methods in biomechanics and biomedical engineering: imaging & visualization deep learning methods for screening pulmonary tuberculosis using chest X-rays deep learning methods for screening pulmonary tuberculosis using chest X-rays. Comput Methods Biomech Biomed Eng. https://doi.org/10.1080/21681163.2020.1808532

119. Panicker RO, Kalmady KS, Rajan J, Sabu MK (2018) Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. Biocybern Biomed Eng 38(3):691–699. https://doi.org/10.1016/j.bbe.2018.05.007

120. Gao XW, James-Reynolds C, Currie E (2020) Analysis of tuberculosis severity levels from CT pulmonary images based on enhanced residual deep learning architecture. Neurocomputing 392:233–244. https://doi.org/10.1016/j.neucom.2018.12.086

121. Hwa SKT, Bade A, Hijazi MHA, Jeffree MS (2020) Tuberculosis detection using deep learning and contrast-enhanced canny edge detected X-ray images. IAES Int J Artif Intell 9(4):713–720. https://doi.org/10.11591/ijai.v9.i4.pp713-720

122. Li X, Zhou Y, Du P, Lang G, Xu M, Wu W (2021) A deep learning system that generates quantitative CT reports for diagnosing pulmonary tuberculosis. Appl Intell 51(6):4082–4093. https://doi.org/10.1007/s10489-020-02051-1

123. Lai NH, Shen WC, Lee CN, Chang JC, Hsu MC, Kuo LN, Yu MC, Chen HY (2020) Comparison of the predictive outcomes for anti-tuberculosis drug-induced hepatotoxicity by different machine learning techniques. Comput Methods Programs Biomed. https://doi.org/10.1016/j.cmpb.2019.105307

124. Lino Ferreira da Silva Barros MH, Oliveira Alves G, et al (2021) Benchmarking machine learning models to assist in the prognosis of tuberculosis. In: Informatics (vol 8, no. 2, p. 27). https://doi.org/10.3390/informatics8020027

125. Das D, Santosh KC, Pal U (2021) Inception-based deep learning architecture for tuberculosis screening using chest X-rays. In: 2020 25th international conference on pattern recognition (ICPR) (pp 3612–3619). IEEE. https://doi.org/10.1109/ICPR48806.2021.9412748

126. Malagelada C, Bendezú RA, Seguí S, Vitrià J, Merino X, Nieto A, Sihuay D, Accarino A, Molero X, Azpiroz F (2020) Motor dysfunction of the gut in cystic fibrosis. Neurogastroenterol Motil 32(9):1–9. https://doi.org/10.1111/nmo.13883

127. Zucker EJ, Barnes ZA, Lungren MP, Shpanskaya Y, Seekins JM, Halabi SS, Larson DB (2020) Deep learning to automate Brasfield chest radiographic scoring for cystic fibrosis. J Cyst Fibros 19(1):131–138. https://doi.org/10.1016/j.jcf.2019.04.016

128. Marques F, Dubost F, Kemner-van de Corput M, Tiddens HA, de Bruijne M (2018) Quantification of lung abnormalities in cystic fibrosis using deep networks. In: Medical imaging 2018: image processing (vol 10574, pp 365–371). SPIE. https://doi.org/10.1117/12.2292188

129. Di Dio R, Galligo A, Mantzaflaris A, Mauroy B (2021) Spirometry-based airways disease simulation and recognition using Machine Learning approaches. In: International conference on

learning and intelligent optimization (pp 98–112). Springer, Cham. https://doi.org/10.1007/978-3-030-92121-7_8

130. Zhou Z, Alvarez D, Milla C, Zare RN (2019) Proof of concept for identifying cystic fibrosis from perspiration samples. Proc Natl Acad Sci 116(49):24408–24412. https://doi.org/10.1073/pnas.1909630116

131. Mondal S, Sadhu AK, Dutta PK (2021) Adaptive local ternary pattern on parameter optimized-faster region convolutional neural network for pulmonary emphysema diagnosis. IEEE Access 9:114135–114152. https://doi.org/10.1109/ACCESS.2021.3105114

132. Bortsova G, Dubost F, Ørting S, Katramados I, Hogeweg L, Thomsen L, Wille M, de Bruijne M (2018) Deep learning from label proportions for emphysema quantification. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), vol 11071 LNCS. Springer International Publishing. https://doi.org/10.1007/978-3-030-00934-2_85

133. Humphries SM, Notary AM, Centeno JP, Strand MJ, Crapo JD, Silverman EK, Lynch DA (2020) Deep learning enables automatic classification of emphysema pattern at CT. Radiology 294(2):434–444. https://doi.org/10.1148/radiol.2019191022

134. Srivastava A, Jain S, Miranda R, Patil S, Pandya S, Kotecha K (2021) Deep learning based respiratory sound analysis for detection of chronic obstructive pulmonary disease. PeerJ Comput Sci 7:1–22. https://doi.org/10.7717/PEERJ-CS.369

135. Ørting SN, Petersen J, Thomsen LH, Wille MM, De Bruijne M (2018) Detecting emphysema with multiple instance learning. In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018) (pp 510–513). IEEE. https://doi.org/10.1109/ISBI.2018.8363627

136. Aguilaniu B, Hess D, Kelkel E, Briault A, Destors M, Boutros J, Zhi Li P, Antoniadis A (2021) A machine learning approach to predict extreme inactivity in COPD patients using non-activity-related clinical data. PLoS ONE. https://doi.org/10.1371/journal.pone.0255977

137. González G, Washko GR, Estépar RSJ (2018) Deep learning for biomarker regression: application to osteoporosis and emphysema on chest CT scans. In: Medical imaging 2018: image processing (vol 10574, p 105741H). International Society for Optics and Photonics. https://doi.org/10.1117/12.2293455

138. Pino Peña I, Cheplygina V, Paschaloudi S, Vuust M, Carl J, Weinreich UM, et al (2018) Automatic emphysema detection using weakly labeled HRCT lung images. PLoS ONE. https://doi.org/10.1371/journal.pone.0205397

139. Wang X, Yu J, Zhu Q, Li S, Zhao Z, Yang B, Pu J (2020) Potential of deep learning in assessing pneumoconiosis depicted on digital chest radiography. Occup Environ Med 77(9):597–602. https://doi.org/10.1136/oemed-2019-106386

140. Wang Z (2020) Deep learning for computer-aided diagnosis of pneumoconiosis. Res Square 1–14. https://doi.org/10.21203/rs.3.rs-460896/v1

141. Zhang L, Rong R, Li Q, Yang DM, Yao B, Luo D, Zhang X, Zhu X, Luo J, Liu Y, Yang X, Ji X, Liu Z, Xie Y, Sha Y, Li Z, Xiao G (2021) A deep learning-based model for screening and staging pneumoconiosis. Sci Rep 11(1):1–7. https://doi.org/10.1038/s41598-020-77924-z

142. Zheng R, Zhang L, Jin H (2021) Pneumoconiosis identification in chest X-ray films with CNN-based transfer learning. CCF Trans High Perf Comput 3(2):186–200. https://doi.org/10.1007/s42514-021-00067-8

143. Zhang S, Wu Y, Wang M (2018) Pulse signal analysis for pneumoconiosis detection with SVM. In: 2018 international symposium on computer, consumer and control (IS3C), pp 221–224. https://doi.org/10.1109/IS3C.2018.00063

144. Wang D, Arzhaeva Y, Devnath L, Qiao M, Amirgholipour S, Liao Q, et al (2020) Automated pneumoconiosis detection on chest X-rays using cascaded learning with real and synthetic radiographs. In: 2020 digital image computing: techniques and applications (DICTA) (pp 1–6). IEEE. https://doi.org/10.1109/DICTA51227.2020.9363416

145. Kumar Y, Koul A, Singla R, Ijaz MF (2022) Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-021-03612-z

146. Xu T, Cheng I, Long R, Mandal M (2013) Novel coarse-to-fine dual scale technique for tuberculosis cavity detection in chest radiographs. EURASIP J Image Video Process 2013(1):1–18. https://doi.org/10.1186/1687-5281-2013-3

147. Kumar Y, Sood K, Kaul S, Vasuja R (2020) Big data analytics and its benefits in healthcare. In: Big data analytics in healthcare (pp 3–21). Springer, Cham. https://doi.org/10.1515/jib-2017-0030

148. Sun J, Liao X, Yan Y, Zhang X, Sun J, Tan W, et al (2022) Detection and staging of chronic obstructive pulmonary disease using a computed tomography-based weakly supervised deep learning approach. Eur Radiol. https://doi.org/10.1007/s00330-022-08632-7

149. Sills MR, Ozkaynak M, Jang H (2021) Predicting hospitalization of pediatric asthma patients in emergency departments using machine learning. Int J Med Inf. https://doi.org/10.1016/j.ijmedinf.2021.104468

150. Ram S, Hoff BA, Bell AJ, Galban S, Fortuna AB, Weinheimer O, et al (2021) Improved detection of air trapping on expiratory computed tomography using deep learning. PLoS ONE. https://doi.org/10.1371/journal.pone.0248902

151. Auffermann WF (2021) Quantifying pulmonary edema on chest radiographs. Radiology 3(2):e210004. https://doi.org/10.1148/ryai.2021210004

152. Palaniappan R, Sundaraj K, Ahamed NU (2013) Machine learning in lung sound analysis: a systematic review. Biocybern Biomed Eng 33(3):129–135. https://doi.org/10.1016/j.bbe.2013.07.001

153. Zhang J, Xie Y, Li Y, Shen C, Xia Y (2020) Covid-19 screening on chest X-ray images using deep learning based anomaly detection. arXiv:2003.12338. https://www.researchgate.net/publication/340271344_COVID-19_Screening_on_Chest_X-ray_Images_Using_Deep_Learning_based_Anomaly_Detection

154. Nielsen AL, Thomassen A, Hess S, Alavi A, Høilund-Carlsen PF (2013) Deep venous thrombosis and pulmonary embolism detected by FDG PET/CT in a patient with bacteremia. Clin Nucl Med 38(4):276–277. https://doi.org/10.1097/RLU.0b013e3182817aaf

155. Harish Paruchuri MMAAGSV (2021) Medical diagnosis using deep learning techniques: a research survey. Ann Rom Soc Cell Biol 25(6):5591–5600