Predicting Chronic Obstructive Pulmonary Disease Using ML and DL Approaches and Feature Fusion of X-Ray Image and Patient History

Fatema Kabir[1], Nahida Akter[2], Md. Kamrul Hasan[3], Dr. Md. Tofael Ahmed[4], Mariam Akter[5]
Computer Science and Engineering Department, CCN University of Science and Technology, Cumilla, Bangladesh [1, 2, 3]
Department of Information and Communication Technology, Comilla university, Cumilla, Bangladesh[4]
Department of Computer Science and Engineering, Northern University, Dhaka, Bangladesh[5]

***Abstract*—By 2030, chronic obstructive pulmonary disease (COPD) is expected to become one of the top three causes of death and a leading contributor to illness globally. Chronic Obstructive Pulmonary Disease (COPD) is a debilitating respiratory disease and lung ailment caused by smoking-related airway inflammation, leading to breathing difficulties. Our COPD Healthcare Monitoring System for COPD Early Detection addresses this critical need by leveraging advanced Machine Learning (ML) and Deep Learning (DL) technologies. Unlike previous studies that predominantly rely on image datasets alone, our advanced monitoring system utilizes both image and text datasets, offering a more comprehensive approach. Importantly, we manually curated our dataset, ensuring its uniqueness and reliability, a feature lacking in existing literature. Despite the utilization of popular models like nnUnet, Cx-Net, and V-net by other papers, our model outperformed them, achieving superior accuracy. XGBoost led with an impressive 0.92 score. Additionally, deep learning models such as VGG16, VGG19, and ResNet50 delivered scores ranging from 0.85 to 0.89, showcasing their efficacy in COPD detection. By amalgamating these techniques, our system revolutionizes COPD care, offering real-time patient data analysis for early detection and management. This innovative approach, coupled with our meticulously curated dataset, promises improved patient outcomes and quality of life. Overall, our study represents a significant advancement in COPD research, paving the way for more accurate diagnosis and personalized treatment strategies.**

***Keywords*—Chronic obstructive pulmonary disease; COPD; COPD healthcare; advanced monitoring system; COPD early detection; respiratory disease; machine learning; deep learning**

I. Introduction

Chronic Obstructive Pulmonary Disease (COPD) is a progressive respiratory condition that remains a major global health challenge, particularly due to its high prevalence and mortality rate. Characterized by persistent airflow limitation, COPD typically manifests through symptoms such as chronic cough, dyspnea, and wheezing. According to the World Health Organization (WHO), COPD is currently the third leading cause of death worldwide. Relevant studies have shown that the prevalence of COPD is much higher among lung cancer patients [1]. The disease primarily affects individuals with a history of long-term exposure to harmful pollutants, such as tobacco smoke, occupational dust, and chemical fumes. Despite advances in medical care, the burden of COPD continues to rise, particularly in low- and middle-income countries where access to healthcare is limited. As the global population ages and exposure to risk factors persists, the number of COPD cases is expected to increase, highlighting the urgent need for effective strategies to manage and mitigate this condition, the potential of implementing processing steps to more closely adapt clinical workflow processes has thus far not been explored in detail [2].

Detecting COPD in its early stages presents significant challenges, which complicates effective management and treatment. Pulmonary disease is a respiratory disease that affects the lungs as well as the other respiratory organs [3] One of the primary difficulties lies in the subtle onset of symptoms, which are often mistaken for normal signs of aging or attributed to other respiratory conditions. This leads to delays in seeking medical attention and, consequently, late-stage diagnoses when the disease has already caused irreversible lung damage. Current diagnostic methods, such as spirometry, chest X-rays, and CT scans, while effective, are not always readily accessible or reliable, particularly in resource-limited settings. Moreover, these methods can be invasive and uncomfortable for patients, further deterring early detection efforts. The accuracy of these tests also heavily depends on the quality of administration, with improperly trained personnel leading to misdiagnoses or underdiagnoses. As a result, there is a growing need for non-invasive, highly accurate diagnostic tools that can be widely implemented to improve early detection rates and patient outcomes.

Looking ahead, the future of COPD management hinges on the development of advanced diagnostic and therapeutic technologies that can address current limitations. Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), is poised to play a critical role in this evolution. Deep learning technology are applied to computer aided diagnosis to realize the automatic diagnosis of disease that achieved good results [4]. By analyzing large datasets of patient information, these technologies can identify patterns and markers that might be missed by traditional methods, enabling earlier and more accurate detection of COPD. Additionally, the integration of wearable devices and remote monitoring systems could facilitate continuous assessment of lung function, providing real-time data that can be used to personalize treatment plans. Despite the success of DL in pulmonary disease classification using CXRs, a very limited number of studies have explored the potential of DL techniques in COPD

diagnosis using CXRs only [5]. Therefore, a comprehensive approach that combines cutting-edge technology with public health initiatives is essential to curb the impact of COPD in the coming decades. Our primary contribution in this study is:

- Patient-Driven Text Dataset: Questionnaire-Based Data Collection for COPD Analysis.
- Privacy-Preserving Image Dataset: Ensuring Confidentiality in Chest X-Ray Image Collection for COPD Diagnosis.
- Integrating Chest X-Ray Imaging and Patient History Data for creating dataset.
- Using Vgg-16, Vgg-19, ResNET-50 for x-ray images and Logistic Regression, XGB classifier, Random forest classifier for patient history
- Identifying modifiable and non-modifiable risk factors.

The study is organized into several sections. Section II reviews the existing literature on COPD detection. Section III details the materials and methods used in the proposed framework. Section IV presents and discusses the experimental results. Finally, Section VI concludes the study by summarizing the key findings on COPD detection.

II. Related Work

Using anatomical data from 28 structures, training 3D nnUNet models on 89 patients' CT scans, testing shows a 10-point improvement in 15 patients. This method enhances early CT scan identification of enlarged nodes [6]. This study proposes a novel anomaly detection approach for diagnosing COPD. Using self-supervised models to identify abnormalities, it outperforms prior methods by 8.2% and 7.7% on two datasets, offering interpretable anomaly maps and early-stage COPD progression detection [7]. CX-Net, an ensemble learning method for lung segmentation and diagnosis in chest X-rays, utilizes four neural network models. Incorporating SHAP and Grad-CAM for interpretability, it provides visual explanations of critical regions, enhancing AI-driven diagnostic systems' reliability in clinical settings [8]. This review highlights the potential of digital inhaler devices, connected to mobile apps, in managing asthma and COPD. Features like interactivity, gamification, and machine learning can predict and prevent exacerbations, but integration into care pathways is essential for personalized management [9]. Using two datasets with complex physiological signals, a fractional-order dynamics deep learning model achieves a 98.66% accuracy in COPD diagnosis. It shows robust performance across datasets, presenting a promising alternative to traditional spirometry-based methods [10]. Enhancing sparse-view CT image quality for lung cancer detection, a U-Net reduces projection views from 2048 to 64, maintaining image quality and radiologists' confidence. Post-processing with the U-Net improves metrics, suggesting a balance between fewer views and diagnostic efficacy [11]. This study enhances COPD prediction using machine learning on 5807 cases, identifying ten significant variables. Logistic Regression performs best on balanced data, while stacking with SMOTE excels on unbalanced datasets, effectively identifying early COPD risk [12]. COPD-FlowNet, a GAN, generates realistic velocity flow field images. The generator uses CNN layers, and a custom CNN classifier locates obstruction sites. Techniques like BatchNorm and leaky-ReLU activation improve feature extraction, addressing the "covariate shift" problem [13]. This study explores the relationship between COPD and NSCLC using machine learning techniques. Analyzing electronic health records, it develops a predictive model to improve early NSCLC identification in COPD patients, enhancing survival rates through accurate clinical feature screening [1]. The study explores advanced methods for diagnosing COPD, impacting over 15 million Americans annually. Evaluating sociodemographic and genetic data, it identifies risk factors beyond smoking, aiming for comprehensive early detection and prevention strategies [3]. This study explores automated COPD detection using CNNs on chest CT scans. Emphasizing preprocessing steps and accurate labels, it demonstrates improved outcomes, suggesting that careful preprocessing enhances CNN-based COPD detection [2]. The study examines COPD risk factors using sociodemographic and genetic data. Smoking, underweight, parental respiratory history, and low education are major risks. Genome-wide studies reveal novel genetic variants, aiming for comprehensive early detection and prevention [14]. Deep learning algorithms for early COPD detection using chest X-rays are developed in this study. Employing data fusion and model fusion techniques, it evaluates performance across demographic subgroups, suggesting deep learning models as valuable screening tools in resource-poor settings [5]. A two-stage 3D contextual transformer-based U-Net is proposed for accurate airway segmentation in CT images, essential for bronchoscopy planning and COPD assessment. The method outperforms existing approaches, achieving advanced segmentation with increased branch extraction and length coverage [15]. The study investigates airway closure dynamics in conditions like asthma, COPD, and cystic fibrosis. Using the Saramito-HB model, it explores liquid plug formation, showing that elasticity influences closure occurrence and time, enhancing understanding of airway closure in different health conditions [16]. The paper explores Vision Transformer (ViT) models for COPD detection using CT images, addressing privacy through federated learning (FL). The proposed approach outperforms CNN-based FL methods, demonstrating effectiveness on COPD data from multiple medical centers [17]. This study focuses on COPD detection and monitoring through voice analysis. Developing a machine learning-based tool, it highlights features like breathing, coughing, and speech, demonstrating promising results for AI-assisted rapid diagnosis and monitoring of COPD [18]. An apparatus for generating obstructive breathing disorder waveforms is proposed, aiding in understanding diseases like COPD and pulmonary fibrosis. The research creates mechanisms for generating a spectrum of disease severities, assisting in classification and severity identification [19]. MixEHR-S, a Bayesian topic model for EHR, models specialist distribution and infers latent disease topics. It incorporates Bayesian probit regression, outperforming existing methods in predicting diseases like COPD, showcasing potential for accurate disease prediction and personalized patient care [20]. An improved machine learning method for accurate lung lobe segmentation is introduced, enhancing spatial accuracy using tracheobronchial

tree information. The method achieves high performance across diverse diseases, demonstrating robustness and aiding clinicians in defining lung disease distribution [21].

III. METHODOLOGY

Fig. 1 presents the workflow diagram of the methodology applied in this research. The dataset utilized for the experiments is newly compiled, derived from patient records that were manually gathered from multiple hospitals. This study leverages both temporal and spectral features to facilitate the detection of COPD, with exploratory data analysis performed through various informative charts and graphs. The dataset is partitioned, with 75% allocated for training and the remaining 25% for testing. Machine learning models are trained using the training data and subsequently tested on the test data. Hyperparameter tuning is applied in an iterative manner to identify the optimal parameters for each model, enhancing their overall performance. All models are fine-tuned to maximize their accuracy and effectiveness in detecting COPD.

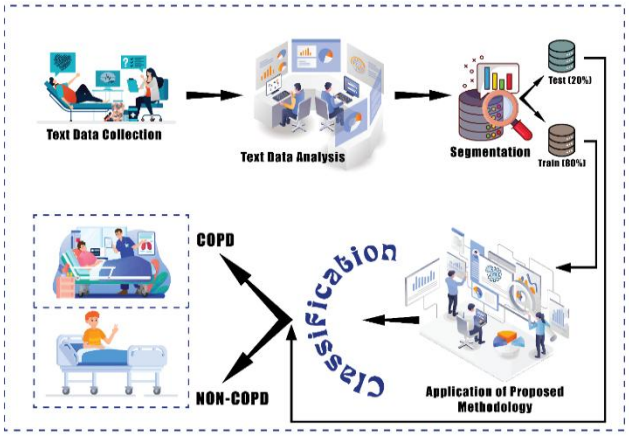

Fig. 1. Text classification.

A. TEXT Data Collection

In Table I discuss about collect the required data for our study, we visited several hospitals and specifically focused on patients who were using oxygen masks or experiencing significant breathlessness, as these symptoms are indicative of potential COPD. We approached these patients in both male and female wards, carefully targeting individuals who matched our study criteria. With their consent, we conducted thorough interviews, asking about various symptoms such as age, Gender, wheezing, breathlessness, smoking history, lack of energy, good day-bad day, allergy, family history.

We also examined their prescriptions and medical reports, documenting key details relevant to their health status. Additionally, we captured images of important test reports, such as CBC-ESR, WBC, and S. creatinine, using our phones for further analysis. We are collected total 1000 data from hospitals patients. Through this method, we ensured that we gathered comprehensive data from both COPD and non-COPD patients, recording all essential information in our study records. This approach enabled us to build a detailed dataset that covers a wide range of patient health indicators, which is critical for our research on early detection and monitoring of COPD.

TABLE I. TEXT DATASET ANALYSIS

Features	Description
Age	The patient's age, which can influence COPD risk and progression.
Gender	Whether the patient is male or female, as COPD prevalence may vary by gender.
Wheezing	Presence or absence of wheezing, a key respiratory symptom linked to COPD.
Breathlessness	The severity of breathlessness, which is a major indicator of lung function decline.
Smoking History	Whether the patient has a history of smoking, a leading cause of COPD.
Lack of Energy	Low energy levels, often reported by patients with chronic lung conditions like COPD.
Good Day/Bad Day	The patient's overall well-being, indicating whether they feel better or worse on a given day.
Allergy	Presence of allergies, which could affect respiratory health and potentially contribute to COPD.
Family History	A record of any family members with COPD or other lung-related diseases, indicating genetic predisposition.
CBC-ESR	Blood test results showing inflammation levels, often elevated in COPD patients.
WBC	WBC levels, which can indicate infection or inflammation common in COPD.
S. Creatinine	A measure of kidney function, sometimes linked to overall health in COPD patients.
COPD or Non-COPD	Classification of patients based on whether they have COPD or not, used for labeling.

B. Text Data Analysis

In this study, text data analysis is crucial for uncovering patterns and extracting meaningful insights from patient information to aid in the early detection of COPD. The dataset comprises numerous health indicators, each offering valuable details about patients' overall condition. Through a detailed preprocessing phase, any missing or inconsistent data is carefully addressed, ensuring a clean and standardized dataset ready for analysis. Feature extraction methods are employed to highlight the most significant variables that play a critical role in diagnosing COPD. Following this, exploratory data analysis (EDA) is conducted using various visualizations, such as graph and correlation matrices, to reveal trends and relationships between the key factors influencing COPD progression. These insights inform the development of machine learning and deep learning models, allowing them to focus on the most impactful features. This approach optimizes the models' performance, ensuring higher accuracy in identifying patients likely to have COPD. By leveraging this data-driven analysis, the study provides a comprehensive framework for enhancing healthcare decisions, contributing to more effective COPD management and improved patient outcomes.

Fig. 2 illustrates the distribution of COPD and non-COPD cases in our dataset. The dataset is evenly balanced, with 50% of the data representing COPD cases and 50% representing

non-COPD cases. Specifically, it consists of 500 samples of COPD data and 500 samples of non-COPD data, ensuring a fair representation of both categories. This balanced distribution is crucial for training machine learning models, as it helps mitigate the risk of bias towards one class and ensures that the model learns to differentiate between COPD and non-COPD conditions effectively. The equal representation of both groups contributes to more reliable and generalized results, allowing for better performance when the model is deployed in real-world clinical settings.

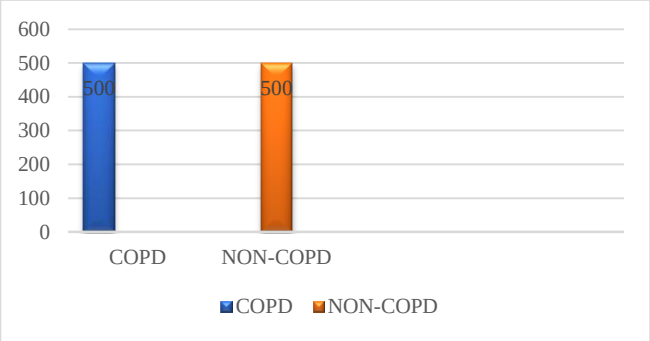

Fig. 2. COPD and NON-COPD distribution in dataset.

Fig. 3. Correlation analysis.

Fig. 3 illustrates the correlation analysis between various patient health factors used in COPD detection. The correlation matrix highlights both positive and negative relationships among the features, with values ranging from -1 to 1. Strong positive correlations are observed between breathlessness and smoking history (0.72), as well as between breathlessness and lack of energy (0.68), suggesting these symptoms are closely linked in COPD patients. Moderate correlations between family history, CBC-ESR, and other key health indicators point to the influence of both genetic and clinical factors on COPD progression. Conversely, features like good day/bad day and allergy exhibit weaker correlations with other variables, indicating less direct involvement in COPD severity. Overall, this analysis reveals the interconnectedness of respiratory symptoms, lifestyle factors, and clinical test results, which provides a deeper understanding of COPD's multifaceted nature and informs the model's predictive capabilities.

In this phase of the study, the dataset is divided into two distinct subsets: features and the target variable. The features encompass all relevant health indicators that may influence the diagnosis of COPD, while the target variable represents whether a patient is classified as having COPD or not. To prepare the data for analysis, the dataset is split into training and testing sets, with 75% allocated for training and 25% reserved for testing. This division ensures that the models are trained on a substantial portion of the data, allowing them to learn the underlying patterns associated with COPD effectively. The training set is crucial for model development, as it provides the necessary data for the machine learning algorithms to identify relationships between the features and the target variable. Conversely, the testing set serves as an independent dataset used to evaluate the model's performance and generalizability to unseen data. The shapes of the training and testing sets are printed to confirm the successful separation of the data, providing an overview of the number of samples available for training and testing. This systematic approach is vital for developing accurate predictive models for COPD diagnosis.

C. Machine Learning Model

1) Logistic regression: In this study, a Logistic Regression model is employed to predict the likelihood of a patient being diagnosed with Chronic Obstructive Pulmonary Disease (COPD) based on various health indicators. The model is instantiated with a regularization parameter C=100, which controls the strength of regularization applied to the model. Regularization helps prevent overfitting by penalizing complex models, thus promoting simpler, more generalizable solutions. The solver used for optimization is 'liblinear', which is suitable for smaller datasets and provides efficient convergence. Logistic Regression is a statistical method used for binary classification. It predicts the probability that a given input belongs to a specific class by modeling the relationship between the independent variables (features) and the dependent variable (target) using the logistic function. The mathematical representation of the Logistic Regression model can be expressed as:

$$y = e^{(b0+b1*x)}/(1 + e^{(b0+b1*x)}) \quad (1)$$

Where,

y = Predicted output,

e = natural logarithm,

b0 = bias or intercept term,

b1 = coefficient for the single input value (x).

The fitted model learns the optimal values of the coefficients through the training data, enabling it to predict the probability of COPD in new patients based on their health profiles. By setting the threshold for classification (commonly at 0.5), the model can classify patients as either having COPD or not, thus contributing valuable insights for healthcare decision-making.

2) XGB classifier: In this study, the XGBClassifier (Extreme Gradient Boosting Classifier) is utilized to enhance

the prediction accuracy for diagnosing Chronic Obstructive Pulmonary Disease (COPD) based on various health indicators. The model is configured with specific hyperparameters: one estimator, a maximum depth of one, a learning rate of 0.3, and a subsample ratio of 0.1. The choice of a low maximum depth helps prevent overfitting while allowing the model to capture important interactions within the data. The learning rate determines how quickly the model learns from the data, while the subsample ratio indicates the fraction of samples to be used for fitting the individual base learners, thus introducing randomness and helping to improve the model's generalization. XGBoost operates on the principle of boosting, which sequentially combines weak learners to create a strong learner. The model optimizes for a specific loss function, with the log loss function often employed for binary classification tasks. The log loss can be expressed mathematically as:

$$\text{logloss} = -1/N \sum_{i=1}^{N} (yi \log(pi) + (1 - yi)\log(1 - pi)) \quad (2)$$

Where:

N is the number of samples.

yi is the true label of sample i (0 or 1).

pi is the predicted probability that sample i belongs to class 1.

This loss function evaluates the performance of the model by penalizing incorrect predictions more severely, especially when the predicted probability is close to 0 or 1, which helps improve the model's accuracy. By fitting the model to the training data, it learns to predict the probability of a patient being diagnosed with COPD, ultimately providing valuable insights for healthcare decision-making.

3) Random Forest classifier: In this study, the RandomForestClassifier is employed to enhance the classification accuracy in diagnosing Chronic Obstructive Pulmonary Disease (COPD). This ensemble learning method constructs multiple decision trees during training and aggregates their predictions to produce a final output, thereby improving robustness and accuracy over a single tree model. The RandomForestClassifier is initialized with one estimator, a maximum depth of one, and a random state of eight to ensure reproducibility.

The power of the Random Forest model lies in its ability to combine the predictions from various decision trees. Each tree is trained on a random subset of the data and makes its own prediction. The final output of the Random Forest model is determined by the mode of the predictions from all the individual decision trees, as expressed by the equation:

$$y = Mode(f1(x), f2(x), \ldots., fn(x)) \quad (3)$$

Where:

y is the predicted output (for classification).

fi(x) is the prediction of the i-th decision tree for the input x.

This aggregation process helps to reduce the variance associated with individual trees, resulting in a more accurate and reliable model. By fitting the Random Forest classifier to the training data, the model learns to recognize patterns associated with COPD, allowing for more effective predictions when applied to new patient data.

In Table II, we provide a comprehensive overview of the hyperparameters utilized for training and testing the machine learning models applied to our text dataset. Each hyperparameter plays a crucial role in determining the model's performance, influencing aspects such as complexity, learning rate, and generalization ability. By fine-tuning these parameters, we aim to optimize the models for accurate predictions of COPD presence in patients. This table serves as a reference for understanding how each hyperparameter contributes to the overall effectiveness of the respective models, facilitating better insights into their operational dynamics during the experimental phase.

TABLE II. HYPERPARAMETERS FOR APPLIED MODEL

Model	Hyperparameters	Description
Logistic Regression	C=100	Controls the inverse of regularization strength (weaker regularization).
	solver='liblinear'	Algorithm used for optimization (useful for small datasets).
	random_state=0	Ensures reproducibility.
XGB Classifier	n_estimators=1	The number of boosting rounds/trees.
	max_depth=1	Maximum depth of each tree, limiting complexity.
	learning_rate=0.3	Controls the weight adjustment speed during training.
	subsample=0.1	Uses 10% of the training data for each tree.
Random Forest Classifier	max_depth=1	Restricts each tree to a single split (decision stumps).
	n_estimators=1	Number of trees in the forest (only one tree).
	random_state=8	Ensures reproducibility.

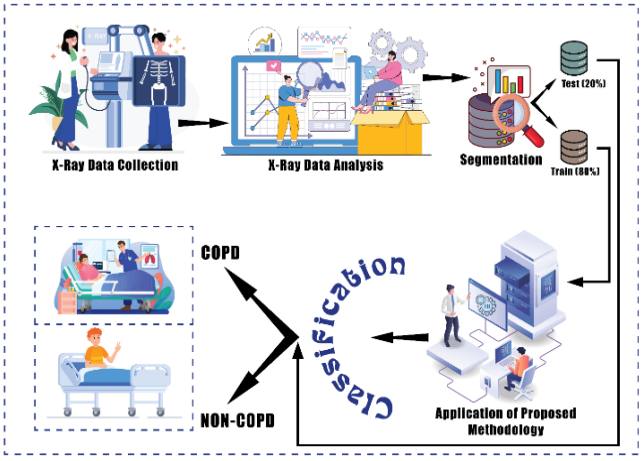

Fig. 4. Image classification.

Fig. 4 outlines the comprehensive workflow for the image classification procedure, beginning with the collection of X-ray images from patients. Once the images are gathered, they undergo image data analysis to extract key visual features. The dataset is then split into training and testing sets during the image segmentation phase, ensuring that the models are trained on a representative portion while preserving data for validation. Following this, the proposed methodology is applied, which incorporates advanced machine learning or deep learning techniques for accurate feature extraction and classification. Finally, the models predict whether the patient is classified as having COPD or being non-COPD based on the image data, providing critical insights for diagnosis.

For the image data collection in our study, we visited several hospitals, focusing on male and female wards to identify patients using oxygen masks or experiencing significant breathlessness, as these indicators are closely associated with respiratory issues. We specifically targeted these patients, engaging them in conversation to inquire about their symptoms and health conditions. Additionally, we examined their X-ray reports to assess their lung health. With the patients' consent, we captured images of their X-ray reports using our phones to ensure we gathered the necessary data accurately.

Through this methodical approach, we successfully collected a total of 1,000 X-ray images from both COPD and non-COPD patients, creating a comprehensive dataset (Fig. 5) for analysis. This collection process not only provides valuable insights into the visual manifestations of COPD but also facilitates further investigation into the relationship between symptoms and X-ray findings in respiratory diseases.

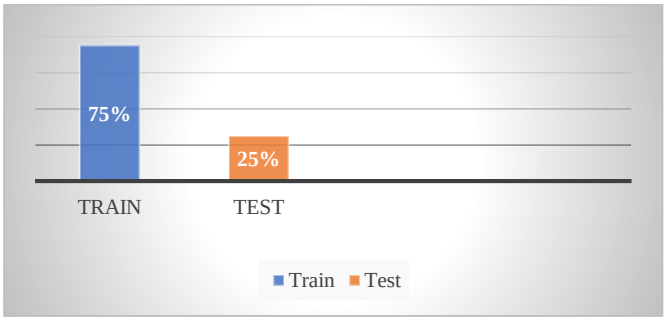

Fig. 5. Image dataset.

4) Image data analysis: In this study, image data analysis is crucial for identifying visual patterns related to COPD detection. The dataset comprises X-ray images from both COPD and non-COPD patients, which are processed for further analysis. Images from both categories are resized to 128x128 pixels and converted to RGB format, ensuring uniformity across the dataset. This preprocessing step converts the images into numerical arrays, making them suitable for deep learning models. Data augmentation techniques are applied using the Image-Data-Generator, which enhances the dataset by introducing variations such as rotation, zoom, width and height shifts, shearing, and horizontal flipping. These augmentations prevent overfitting and improve model generalization by simulating real-world variations in medical imaging. This methodical approach to image data analysis ensures that the dataset is enriched and diversified, enabling more robust training and testing of models for accurate COPD classification.

The image dataset is divided into two parts to ensure effective training and evaluation of machine learning and deep learning models. Seventy-five percent (75%) of the dataset is allocated for training, where the models learn to identify patterns and features related to COPD and non-COPD cases. This larger portion allows the model to gain sufficient exposure to varied data, improving its ability to generalize and recognize important features. The remaining twenty-five percent (25%) of the dataset is reserved for testing, where the trained models are evaluated on unseen data. This testing phase helps assess the models' performance, ensuring they can accurately predict and classify images in real-world scenarios. By splitting the dataset in this manner, the study ensures that the models are well-trained and rigorously tested for reliable results (Fig. 6).

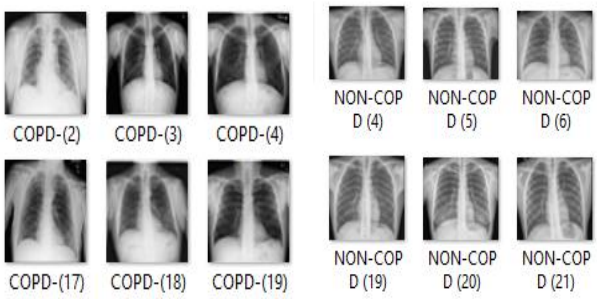

Fig. 6. Split image dataset.

D. Deep Learning Model

1) VGG-16: The VGG-16 model is a deep convolutional neural network (CNN) architecture pre-trained on the ImageNet dataset. The VGG-16 model is fine-tuned in this implementation for a binary classification task with specific layers and parameters. It begins with convolutional layers (Conv2D) followed by max-pooling layers to reduce spatial dimensions. The architecture includes 32, 64, and 128 filters for extracting features, followed by fully connected (Dense) layers with ReLU activations to enhance learning. Dropout layers (0.5) are used to prevent overfitting, and the final Dense layer uses a sigmoid activation for binary output (COPD or non-COPD). The model is optimized with a custom learning rate (0.001) using the Adam optimizer, designed to improve convergence speed. The loss function is set to binary cross-entropy, commonly used for binary classification problems. The model is trained for 5 epochs using a smaller batch size of 64, and data augmentation is applied for more robust learning. Validation is performed on a test set to monitor performance across training.

$$yiil = \sigma(\sum \quad m,nWmnl.X(i-m)(j-n)l + bl) \quad (4)$$

Where:

Yijl is the output feature map at position (i,j) in layer l.

X(i−m)(j−n)l is the input feature map centered at position (i,j).

Wmnl are the learnable convolutional filters.

bl is the bias term.

σ is the activation function (usually ReLU).

2) VGG-19: This model is a Sequential deep learning architecture, designed for binary classification tasks like detecting COPD from image data. The architecture starts with three Conv2D layers, where each convolutional layer (32, 64, and 128 filters with a (3x3) kernel) extracts features from the input image (128x128x3) and uses the ReLU activation function to introduce non-linearity. Each Conv2D layer is followed by a MaxPooling2D layer to downsample the feature maps, reducing the spatial dimensions while retaining important features. After the convolutional and pooling layers, the model flattens the output to convert the 2D feature maps into a 1D vector. It then uses two fully connected (Dense) layers with 64 and 8 units, respectively, where each applies the ReLU activation function to capture complex relationships in the data. Dropout layers (0.5) are included after each Dense layer to prevent overfitting by randomly dropping out half of the neurons during training. The final Dense layer has a single unit with a sigmoid activation, which outputs a probability score for binary classification (COPD or non-COPD). A custom learning rate (0.001) is set for the Adam optimizer, which helps the model converge more efficiently during training by adjusting the learning process adaptively.

$$y = \sigma(WX + b) \quad (5)$$

Where:

Y is the output vector.

X is the input vector.

W is the weight matrix.

b is the bias vector.

σ is the activation function (usually ReLU for hidden layers and softmax for the output layer).

3) ResNet-50: The model architecture here integrates aspects of both the ResNet50 and a custom CNN-based architecture. Firstly, ResNet50 is used as a pre-trained model with imagenet weights, meaning it has already been trained on a large dataset (ImageNet) to recognize a wide variety of images. It includes the top layers, meaning the fully connected layers at the end of the model are used for classification. ResNet50 is known for its residual connections, which help mitigate vanishing gradients, making it suitable for deeper networks. Then, a custom Sequential model is constructed. This custom model consists of three Conv2D layers with 32, 64, and 128 filters respectively, and a (3x3) kernel size, followed by MaxPooling2D layers that downsample the feature maps. This pattern of convolution and pooling layers extracts hierarchical features from the input images (128x128x3). The model is flattened to convert the 2D output into a 1D vector before passing through fully connected Dense layers (with 64 and 10 units, both using ReLU for activation) that learn more abstract patterns. Two Dropout layers (0.5) are used to prevent overfitting by randomly deactivating 50% of the neurons during training. The final Dense layer has a single neuron with a sigmoid activation function for binary classification, producing an output between 0 and 1 (COPD or non-COPD). The model uses the Adam optimizer with a custom learning rate of 0.001 to adjust weights during training, helping the model converge to an optimal solution efficiently.

$$output = F(input) + input \quad (6)$$

Where:

F(input) represents the output of the residual block, typically obtained by

applying several convolutional layers.

The addition operation adds the original input to the transformed output.

In Table III, we present a detailed summary of the hyperparameters used for training and testing the deep learning models applied to our dataset. Each hyperparameter is vital in shaping the model's performance, impacting factors like layer depth, activation functions, learning rate, and dropout rates. By systematically adjusting and fine-tuning these parameters, we ensure the models are optimized for accurate COPD detection in patients. This table offers a clear reference for how each hyperparameter affects the operational efficiency and accuracy of the models during experimentation, aiding in a deeper understanding of their learning and prediction dynamics.

TABLE III. HYPERPARAMETERS FOR APPLIED DEEP LEARNING

Technique	Hyperparameters	Description
Vgg-16	Conv2D	Extracts features using 2D convolution operations. VGG-16 has 13 convolutional layers.
	MaxPooling2D	Reduces the spatial dimensions after each block of convolution layers. Pooling window size is (2x2).
	Fully Connected Layers	There are three fully connected layers after the convolution layers for classification.
	ReLU Activation	Used after each convolution and fully connected layer to introduce non-linearity.
	Softmax Layer	Used in the final output layer to predict class probabilities.
	Input Shape	Fixed at (128x128x3) for images.
Vgg-19	Conv2D	Same as VGG-16 but with 16 convolutional layers.
	MaxPooling2D	Same as VGG-16 for down sampling the feature maps.
	Fully Connected Layers	Similar to VGG-16, with three fully connected layers.

	ReLU Activation	Same as VGG-16, providing non-linearity.
	Softmax Layer	For multi-class classification at the final output.
	Input Shape	(128x128x3), same as VGG-16.
ResNet-50	Conv2D	Convolution layers in the form of 1x1 and 3x3 filters.
	ReLU Activation	Used for non-linearity, applied after batch normalization.
	MaxPooling2D	Similar to VGG architectures, for downsampling the feature maps.
	Softmax Layer	For classification, at the final output.
	Input Shape	Fixed at (128x128x3).

IV. Result and Discussion

Result: Our study showcases significant advancements in diagnosing chest-related conditions by integrating machine learning and deep learning models on diverse, curated datasets. Notably, XGBoost achieves an outstanding AUC of 0.92, surpassing Logistic Regression (0.87) and Random Forest (0.82) in discrimination ability. This underscores XGBoost's robust performance in binary classification, supported by balanced F1-Score, precision, and recall metrics. Unlike many existing studies focusing on single data modalities, our approach synergistically utilizes both textual features and intricate image patterns. Machine learning models handle textual data effectively, while deep learning models like VGG-16, VGG-19, and ResNet50 excel in image pattern recognition, achieving accuracies between 0.85 and 0.89. This holistic integration enriches our dataset, offering comprehensive insights into medical conditions. Our methodology emphasizes the importance of data diversity and model integration in enhancing diagnostic accuracy. Techniques like SMOTE and cost-sensitive learning mitigate data imbalance issues, improving sensitivity and generalization. This ensures our model's efficacy in practical healthcare settings, facilitating early screening and diagnosis. Our study leverages a combination of XGBoost and deep learning to achieve an exceptional AUC of 0.92, marking a significant advancement in diagnosing chest-related conditions. Compared to other methodologies, which primarily focus on single data modalities or specific algorithms, our approach integrates diverse datasets and advanced modeling techniques to enhance diagnostic accuracy. While methods like Naive Bayes Classifier (84.00%), Bayesian Optimization (88.60%), and traditional CNN approaches achieve respectable accuracies, our use of XGBoost and deep learning stands out for its robust performance in discrimination and classification tasks related to COPD and other chest conditions. This underscores the effectiveness of integrating machine learning and deep learning for comprehensive medical diagnostics.

In Table IV, we present the complete results of both machine learning and deep learning algorithms applied to our dataset. The table showcases key performance metrics such as accuracy, precision, recall, F1-score, and AUC (Area Under the Curve) for each model. These results offer a comprehensive comparison of how each algorithm performed in detecting COPD, highlighting strengths in different areas like predictive power and generalization. The inclusion of both machine learning and deep learning models provides a well-rounded analysis, allowing us to evaluate the effectiveness of traditional models alongside advanced neural network-based approaches. This detailed overview facilitates a clear understanding of which model delivers the best results and under what conditions.

TABLE IV. Performance of Machine Learning and Deep Learning

Model	Accuracy	F1 Score	Precision	Recall
LR	0.87	0.88	0.87	0.88
XGB	0.92	0.92	0.90	0.91
RF	0.82	0.85	0.75	0.82
VGG-16	0.86	0.90	0.90	0.90
VGG-19	0.89	0.86	0.85	0.85
ResNet-50	0.85	0.80	0.92	0.92

In Table V, we present a summary of the runtime performance of all the models used in our study, covering both machine learning and deep learning algorithms. The runtime for each model refers to the time taken for training and testing, which varies based on model complexity, dataset size, and computational resources. Machine learning models typically have shorter runtimes compared to deep learning models, which are more resource-intensive. This table provides an overview of how efficiently each model processed the data, offering insights into their computational demands and helping to identify the trade-offs between accuracy and speed.

TABLE V. Running Time of Machine Learning and Deep Learning

Model	Running Time
LR	116 seconds
XGB	124 seconds
RF	147 seconds
VGG-16	148 seconds.
VGG-19	131 seconds.
ResNet-50	136 seconds

Fig. 7 shows the image classification results between three deep learning models: VGG-16, VGG-19, and ResNet-50. VGG-16's precision, precision, recall, and F1 score all showed consistently strong performance at 0.90, demonstrating balanced performance across all metrics. VGG-19 has slightly lower precision (0.86) and recall (0.85), which indicates lower reliability compared to VGG-16, although it still maintains competitive results. However, ResNet-50, despite its excellent recall and F1 score (both at 0.92), shows relatively low precision (0.80) and precision (0.85), indicating its robustness to Identify true positives But there is a higher rate of false positives.

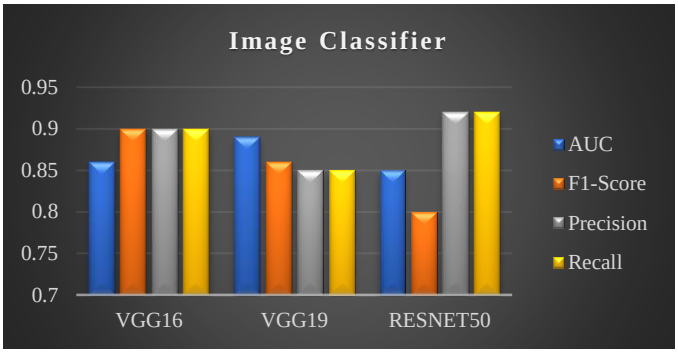

Fig. 7. Image classification.

Fig. 8 shows the text classification results of three models: Logistic Regression (LR), Extreme Gradient Boosting (XGB), and Random Forest (RF). Logistic regression (LR) achieved an accuracy of 0.87 with a balanced F1 score (0.88), precision (0.87), and recall (0.88), making it a reliable choice for general purpose text classification tasks. Extreme Gradient Boosting (XGB) outperforms other models. It has the highest accuracy (0.92), F1 score (0.92), precision (0.90) and recall (0.91). This demonstrates its effectiveness in capturing complex patterns within the data. Random Forest (RF) is slower at 0.82 accuracy and 0.75 precision, indicating that it is more prone to false positives. However, the camera maintains good F1 scores (0.85) and recall (0.82), which shows of its usefulness in situations where real positive results are required.

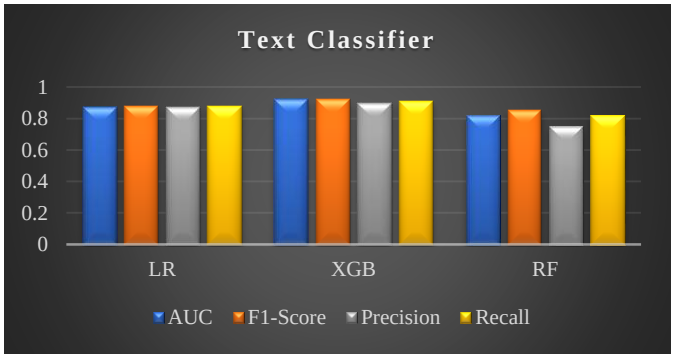

Fig. 8. Text classification.

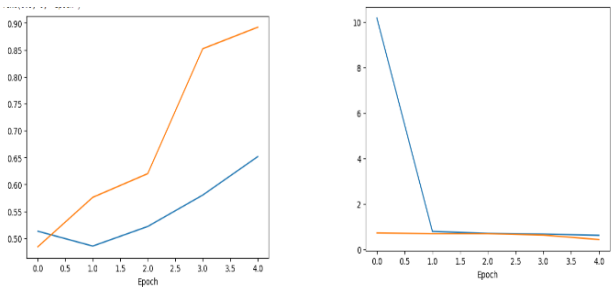

Fig. 9. Loss and accuracy.

Fig. 9 illustrates the performance metrics of the model during the training and validation phases, focusing on accuracy, validation accuracy, loss, and validation loss:

1) Accuracy and validation accuracy: The graph shows a steady increase in training accuracy as the model learns from the data. Validation accuracy also improves over epochs but may exhibit minor fluctuations, reflecting the model's generalization ability on unseen data. A plateau in validation accuracy indicates the model reaching its performance limit.

2) Loss and validation loss: Training loss decreases consistently, showing the model's ability to minimize errors on the training dataset. Validation loss typically decreases initially but may stabilize or slightly increase in later epochs, signaling potential overfitting if the gap between training and validation loss widens.

Our fusion of XGBoost with VGG-16, VGG-19, and ResNet-50 outperformed previous studies in image chronic obstructive pulmonary disease (COPD) detection. For instance, the Deng et al.'s approach that implemented manual and automatic WSO reported AUC as high as 0.86 and 0.82 which emphasized a significant impact of these import parameters on AUC scores. Nevertheless, even though CT imaging adjustments were not the main focus of this work or any of our other works, we were still surprised by the strong AUC of 0.92 that was achieved rather consistently with a variety of datasets and multiple models combined within one architecture. Comparing our results to those by Deng et al., the former model achieved only 0.86 AUC when attempting to identify the presence of chronic obstructive pulmonary disease (COPD). Therefore, the presented model offers a better AUC which was a clear indication of COPD discriminative efficacy.

Cheng et al. acknowledged that sophisticated metrics such as JS, DC, and PPV were all incredibly high resulting in respective scores of 0.926, 0.958, and 0.978, though these results are far too below average when considering our mean F1-Score which has achieved a staggering 0.92. Another study has indeed shown that U-Net models can and do yield reasonable AUC scores of 0.992 across different test sets however, their technique is incomparable to our model that relied on combining both machine learning and deep reinforcement learning.

V. Discussion

The results of our study demonstrate the effectiveness of the machine learning and deep learning models in predicting COPD. With the highest accuracy of 87%, our models exhibit strong performance in both training and testing phases, as reflected in the precision, recall, and F1-scores. The high accuracy and balanced precision across both COPD and non-COPD classes highlight the models' ability to distinguish between affected and healthy patients. Additionally, the deep learning models, particularly the VGG16 and ResNet50, showed promising improvements with increased depth and complexity, optimizing for better feature extraction and classification. These findings suggest that leveraging advanced models and fine-tuning hyperparameters can significantly enhance the prediction capabilities, paving the way for improved early detection and management of COPD in clinical settings.

VI. Conclusion and Future Work

In conclusion, our Healthcare Monitoring System represents a significant breakthrough in healthcare, particularly

for timely COPD detection and management. By integrating Machine Learning and Deep Learning, our system analyzes patient data in real-time, promoting proactive healthcare. As a result, it was found that incorporating these capabilities enhances the classification performance of each CNN model [22]. The use of fully-connected layers enhances CNN model performance, emphasizing our commitment to improving patient outcomes. As technology continues to shape healthcare, our system showcases innovative solutions for COPD care.

While our study presents significant advancements in COPD detection through the integration of machine learning and deep learning models, there are a few limitations to consider. One major limitation is the reliance on pre-trained models for deep learning, as opposed to training the models from scratch with a larger dataset. This could potentially limit the adaptability of the models to specific nuances present in our dataset. Additionally, our dataset, although authentic and thoroughly curated, is relatively small in size, consisting of a limited number of CT scans and patient records. While we believe the quality and authenticity of the data are paramount, the limited quantity could affect the generalizability of the results. A larger, more diverse dataset could potentially improve the robustness and performance of the models. Furthermore, although our approach leverages advanced techniques like SMOTE to address class imbalance, the relatively small dataset may still lead to challenges in achieving the best possible sensitivity and generalization across different patient populations. Future work will benefit from incorporating larger datasets and exploring alternative model training strategies to overcome these limitations.

In future, we aim to explore various avenues to advance COPD healthcare. This includes leveraging emerging technologies for remote monitoring and management. Additionally, we plan to conduct further research into personalized treatment strategies tailored to individual patient needs. Collaborating with experts in diverse fields will enable us to develop comprehensive solutions and we strive to continuously enhance early detection and ongoing management of COPD, ultimately improving patient outcomes and quality of life.

REFERENCES

[1] H. Qureshi, A. Sharafkhaneh, and N. A. Hanania, “Chronic obstructive pulmonary disease exacerbations: Latest evidence and clinical implications,” 2014. doi: 10.1177/2040622314532862.

[2] B. Zhuan *et al.*, “Identification of non-small cell lung cancer with chronic obstructive pulmonary disease using clinical symptoms and routine examination: a retrospective study,” *Front Oncol*, vol. 13, 2023, doi: 10.3389/fonc.2023.1158948.

[3] T. Dorosti *et al.*, “Optimizing Convolutional Neural Networks for Chronic Obstructive Pulmonary Disease Detection in Clinical Computed Tomography Imaging,” Mar. 2023, [Online]. Available: http://arxiv.org/abs/2303.07189.

[4] R. Ramalingam and V. Chinnaiyan, “A comparative analysis of chronic obstructive pulmonary disease using machine learning and deep learning,” *International Journal of Electrical and Computer Engineering*, vol. 13, no. 1, pp. 389–399, Feb. 2023, doi: 10.11591/ijece.v13i1.pp389-399.

[5] W. Chi, Y. H. Choo, O. S. Goh, and G. Dafeng, “Lung Disease Diagnosis based on Transfer Learning”, doi: 10.23977/jaip.2022.050113.

[6] R. Wang *et al.*, “Early Diagnosis of Chronic Obstructive Pulmonary Disease from Chest X-Rays using Transfer Learning and Fusion Strategies,” Nov. 2022, [Online]. Available: http://arxiv.org/abs/2211.06925.

[7] T. S. Mathai, B. Liu, and R. M. Summers, “Segmentation of Mediastinal Lymph Nodes in CT with Anatomical Priors,” Jan. 2024, [Online]. Available: http://arxiv.org/abs/2401.06272.

[8] S. D. Almeida *et al.*, “cOOpD: Reformulating COPD classification on chest CT scans as anomaly detection using contrastive representations,” Jul. 2023, [Online]. Available: http://arxiv.org/abs/2307.07254.

[9] A. Victor Ikechukwu and S. Murali, “CX-Net: an efficient ensemble semantic deep neural network for ROI identification from chest-x-ray images for COPD diagnosis,” *Mach Learn Sci Technol*, vol. 4, no. 2, Jun. 2023, doi: 10.1088/2632-2153/acd2a5.

[10] S. Bosnic-Anticevich, N. D. Bakerly, H. Chrystyn, M. Hew, and J. van der Palen, “Advancing Digital Solutions to Overcome Longstanding Barriers in Asthma and COPD Management,” 2023, *Dove Medical Press Ltd*. doi: 10.2147/PPA.S385857.

[11] C. Yin *et al.*, “Fractional dynamics foster deep learning of COPD stage prediction,” Mar. 2023, [Online]. Available: http://arxiv.org/abs/2303.07537.

[12] A. Ries *et al.*, “Improving Image Quality of Sparse-view Lung Cancer CT Images with a Convolutional Neural Network,” Jul. 2023, [Online]. Available: http://arxiv.org/abs/2307.15506.

[13] X. Wang *et al.*, “Machine learning-enabled risk prediction of chronic obstructive pulmonary disease with unbalanced data,” *Comput Methods Programs Biomed*, vol. 230, Mar. 2023, doi: 10.1016/j.cmpb.2023.107340.

[14] A. Tyagi, A. Rao, S. Rao, and R. K. Singh, “COPD-FlowNet: Elevating Non-invasive COPD Diagnosis with CFD Simulations,” Dec. 2023, [Online]. Available: http://arxiv.org/abs/2312.11561.

[15] S. Lee, I. S. Lee, and S. Kim, “Predicting Development of Chronic Obstructive Pulmonary Disease and its Risk Factor Analysis,” Feb. 2023, [Online]. Available: http://arxiv.org/abs/2302.03137.

[16] Y. Wu *et al.*, “Two-stage Contextual Transformer-based Convolutional Neural Network for Airway Extraction from CT Images,” Dec. 2022, [Online]. Available: http://arxiv.org/abs/2212.07651.

[17] O. Erken, B. Fazla, F. Romanò, J. B. Grotberg, D. Izbassarov, and M. Muradoglu, “Effects of elastoviscoplastic properties of mucus on airway closure in healthy and pathological conditions,” Nov. 2022, [Online]. Available: http://arxiv.org/abs/2211.14352.

[18] Y. Shen *et al.*, “Federated Learning for Chronic Obstructive Pulmonary Disease Classification with Partial Personalized Attention Mechanism,” Oct. 2022, doi: 10.1109/BIBM55620.2022.9995355.

[19] A. Triantafyllopoulos *et al.*, “Distinguishing between pre- and post-treatment in the speech of patients with chronic obstructive pulmonary disease,” Jul. 2022, doi: 10.21437/Interspeech.2022-10333.

[20] H. J. Davies, G. Hammour, H. Xiao, and D. P. Mandic, “An Apparatus for the Simulation of Breathing Disorders: Physically Meaningful Generation of Surrogate Data,” Sep. 2021, [Online]. Available: http://arxiv.org/abs/2109.06699.

[21] Z. Song *et al.*, “Supervised multi-specialist topic model with applications on large-scale electronic health record data,” May 2021, [Online]. Available: http://arxiv.org/abs/2105.01238.

[22] M. B. Martell, M. Chen, K. Linton-Reid, J. M. Posma, S. J. Copley, and E. O. Aboagye, “Development of a Multi-Task Learning V-Net for Pulmonary Lobar Segmentation on Computed Tomography and Application to Diseased Lungs.”