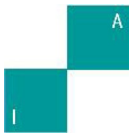
Inteligencia Artificial 24(74) 133-151
doi: 10.4114/intartif.vol24iss74pp133-151



INTELIGENCIA ARTIFICIAL

http://journal.iberamia.org/

Leveraging Transfer Learning for Efficient Diagnosis of COPD Using CXR Images and Explainable AI Techniques

Agughasi Victor Ikechukwu
Department of Computer Science & Engineering
Maharaja Institute of Technology Mysore, Karnataka, 571477 India
Email: victor.agughasi@gmail.com

Abstract
Chronic Obstructive Pulmonary Disease (COPD) is a predominant global health concern, ranking third in mortality rates, yet frequently remains undiagnosed until its advanced stages. Given its prevalence, the need for innovative and widely accessible diagnostic tools has never been more paramount. While spirometry tests serve as conventional diagnostic benchmarks, their reach remains limited, especially in regions with constrained medical resources. The presented research harnesses deep learning algorithms to facilitate early-stage COPD detection, specifically targeting Chest X-rays (CXRs). The clinically annotated VinDR-CXR dataset provides the primary foundation for model training, complemented by incorporating the ChestX-ray14 dataset for initial model pre-training. Such a dual-dataset strategy augments model generalization and adaptability. Among several explored Convolutional Neural Network (CNN) architectures, the Xception model emerges as a frontrunner. Through transfer learning methodologies, this model produces a noteworthy recall rate of 98.2%, markedly surpassing the metrics of the ResNet50 model. Recognizing the imperative for transparency in AI applications in medical imaging, the research integrates essential explainability approaches viz: Gradient Class Activation Mapping (Grad-CAM) and SHapley Additive exPlanations (SHAP). These techniques elucidate the AI's decision-making process, offering invaluable visual and analytical insights for fostering trust among medical professionals. In essence, this study not only underscores the potential of integrating AI with medical imaging for COPD detection but also accentuates the pivotal role of transparency in AI-driven medical interventions.

Keywords: COPD Diagnosis, Chest Radiography (CXR), Pre-trained Models, eXplainable AI, Grad-CAM and SHAP

1 Introduction

Chronic Obstructive Pulmonary Disease (COPD) has evolved into a major health issue worldwide, posing significant challenges to healthcare systems [1]. Hence, detecting COPD early for efficient disease management and improving patient prognosis is pivotal [2]. Traditional diagnostic approaches for COPD encompass clinical evaluation, pulmonary function tests, and imaging techniques, notably Chest X-ray (CXR) images [3]. However, the manual analysis of CXR images is labour-intensive and prone to errors, particularly when diagnosing COPD in its early stages [4]. Clinically, pulmonary function tests (PFT) and spirometry are essential for confirming COPD [5]. However, these tests may not be effective in detecting the early stages of COPD, leading to asymptomatic patients being less frequently tested [6]. The high cost and limited availability of spirometry, particularly in lower-income regions, further delay diagnosis [7]. In contrast, chest radiographs (CXR) are more affordable and widely accessible. This has spurred interest in utilizing CXRs to develop early diagnostic tools for COPD. Such tools could guide individuals toward timely interventions, including lung cancer screening and smoking cessation programs [8].

With advancements in artificial intelligence (AI), deep learning (DL) techniques have proven reliable in numerous medical imaging tasks, including the analysis of CXR images [9]. For example, ResNet50, a DL architecture, performs well in computer vision, and has been applied to various medical imaging problems [10]. In addition, its ability to mitigate the vanishing gradient problem through skip connections makes it an attractive choice for diagnosing COPD from CXR images [11]. Researchers have recently explored using DL models for

ISSN: 1137-3601 (print), 1988-3064 (on-line)

diagnosing COPD from medical images [12], [13]. However, challenges persist in model interpretability, which is crucial for adopting these models in clinical settings [14]. To address this issue, explainable artificial intelligence (XAI) techniques such as Grad-CAM(Gradient Class Activation Maps)[15] have been proposed, providing visual explanations of model predictions and enhancing clinicians' trust in the models.

The recent literature highlights the significant potential of convolutional neural networks (CNNs) for categorizing radiographic findings in chest X-rays (CXR), with several studies achieving expert-level performance in identifying common chest anomalies [16]. Despite extensive research on 14 prevalent radiographic findings [17]-[19], the application of CNNs in detecting Chronic Obstructive Pulmonary Disease (COPD) using CXRs remains underexplored, with only a limited number of studies focusing on integrating CXRs and electronic health records (EHRs) to develop diagnostic models for COPD [20]. Transfer learning, particularly using pre-trained models like ResNet50, has proven to be a formidable method in medical imaging, enhancing the ability of models to discern intricate features from diverse image datasets which are then fine-tuned for specific tasks such as COPD diagnosis [21]-[23]. For instance, Figure 1 illustrates the use of bounding box annotations on selected CXR images, highlighting the disparity between the annotated labels (in red) and model predictions (in blue).

Deep learning's role in improving the detection of various lung diseases through radiography has been thoroughly analyzed by Nasser and Akhloufi [24], who underscored the potential of models like VGG and ResNet, alongside ensemble learning techniques. Their review critically discusses the challenges in the field, particularly the need for greater model interpretability and explainability to facilitate clinical applications. In a similar vein, Chetoui et al. [25] have developed a model that successfully localizes lung diseases from CXRs, demonstrating high accuracy in identifying COVID-19 cases with an AUC of 99%, suggesting its utility in hospital settings for patient triage and isolation while supporting clinical decision-making. Likewise, Brunese et al. [26] present a novel three-step architecture for lung disease detection from CXRs that distinguishes between pneumonia and COVID-19 and identifies symptomatic areas within the X-rays. Their method, tested across 6,523 chest X-rays, showcases a remarkable accuracy rate of 97% and a quick detection time of approximately 2.5 seconds.

Further addressing the challenge of diagnosing lung diseases such as tuberculosis and pneumonia, Nahiduzzaman et al. [27] introduced an innovative approach combining a lightweight CNN model with an extreme learning machine (CNN-ELM). This method demonstrated superior performance compared to existing models, achieving an impressive AUC of 97% across 17 lung disorders. Meanwhile, the survey by Nazir et al. [28] reflects on the broader implications of AI in healthcare, particularly the reluctance to adopt these technologies due to their opaque 'black box' nature, which complicates trust and regulatory compliance. They advocate for the development of explainable AI (XAI) techniques to demystify AI operations and enhance trust among healthcare professionals and patients alike.

In this study, the aim was to harness the capabilities of deep learning and XAI, using Xception and ResNet50 models tailored for diagnosing COPD through CXR images. By integrating advanced pre-processing, explainability methods, and transfer learning, this research strives to improve both the performance and the interpretability of diagnostic models, potentially transforming COPD detection practices.

Survey Findings:

COPD is a significant global health challenge, with early detection crucial for effective treatment strategies. While clinical evaluations and pulmonary function tests traditionally serve as primary diagnostic tools, imaging techniques such as CXR are essential. However, manual analysis of these CXRs, especially in the early stages of COPD, can be prone to errors.

Some of the drawbacks of the existing system include the following:

- **Random Forest with Manual Feature Selection**: While practical for smaller feature sets, this method requires significant domain expertise for feature engineering and selection, making it less scalable and adaptable.
- **Shallow Neural Networks:** These offer some level of automated feature learning but often lack the depth required for capturing more complex patterns in medical images.
- **SVM with PCA:** Effective in dimensionality reduction but lacks the capability for automated feature learning and can struggle with larger, more complex datasets.

Deep learning techniques have made significant advancements in medical imaging. An approach that combines the strengths of the Xception and ResNet50V2 models, both known for their exceptional image-processing performance, was investigated. By leveraging transfer learning, this model benefits from insights gained from large

datasets like ImageNet and ChestX-ray14. These integration allow the model to offer both rapid and accurate diagnostic insights.

Main Contributions are as Follows:

i. The implementation of an explainable model leveraging Grad-CAM and SHAP on ResNet50 and Xception architectures for accurate and interpretable diagnosis of COPD.
ii. The incorporation of transfer learning techniques to improve model performance using pre-trained weights from ImageNet.
iii. Comprehensive evaluation of the model's effectiveness on a large dataset of CXR images, demonstrating its potential to improve patient outcome and assist clinicians in making informed decisions.

This study is in five subsections. Section one sheds light on the background of the study, highlighting the significance and relevant survey. Section 2 highlights the dataset description, pre-processing approaches, and the methodology. Section 3 focuses on the results obtained. Section 4 discusses the results and comparison with various studies. Finally, the conclusion and future scope are presented in section 5.

2 Materials and Methods

2.1 Dataset Description

Two datasets were used for this research, namely: The National Institutes of Health (NIH) Chest X-ray (CXR) dataset, a publicly available collection of more than 110,000 PA-CXR images from 30,805 individuals that have been labelled with one of fourteen frequent disease categories was used[30]. Six additional illnesses of the thorax, including edema, emphysema, fibrosis, pleural thickening, and a hernia, are included in this update of ChestX-ray8. The images are annotated with up to 14 thoracic diseases, including COPD. This comprehensive annotation provides an excellent opportunity to train and evaluate models for detecting specific diseases or multiple conditions simultaneously. In addition, the dataset contains a balanced mix of healthy patients and those with COPD, ensuring a representative sample for developing and evaluating the model. Furthermore, the dataset's diversity in patient demographics, disease severity, and imaging equipment used for acquisition contributes to the robustness of the models trained on it, enhancing their generalization capabilities and potential for real-world clinical applications.

The second dataset employed for validating the models is the "VinDR-CXR" dataset [31], sourced from two major medical institutions in Vietnam. It consists of over 100,000 CXR images, with a subset of approximately 18,000 images meticulously annotated by 17 experienced radiologists. These annotations include 22 specific rectangular labels for localized irregularities and six broad labels identifying potential diseases. The publicly available portion of this dataset includes 30,000 entries, with 15,000 used for model training and 3,000 designated for testing, as depicted in Figure 1. Figure 2 illustrates a selection of images with detailed bounding-box annotations. Notably, during the training phase, each X-ray was independently labeled by three radiologists, ensuring a rich diversity of interpretations, while the testing phase employed a consensus approach among five radiologists for each image. Both the training and validation subsets, along with all anonymized images, are provided in the DICOM format, adhering to medical imaging standards and ensuring consistency and reliability in model training and evaluation.

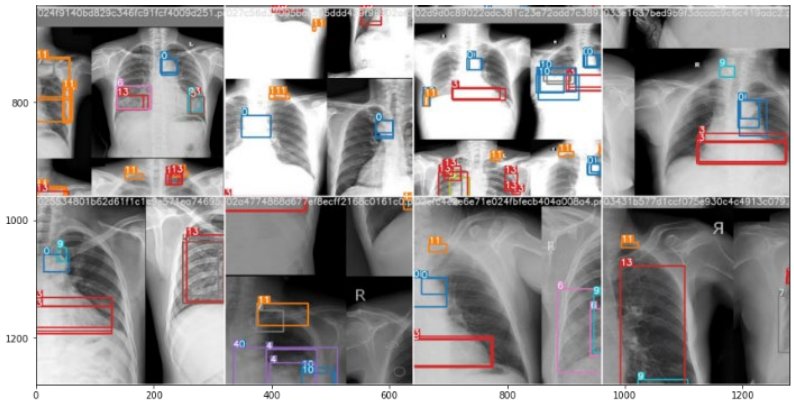


Fig. 1: Typical Examples of Abnormalities with Bounding Boxes.

In Figure 1, multiple bounding boxes, each representing predictions made by different radiologists, offer a visual representation of the inherent challenges faced in medical imaging interpretations. These bounding boxes highlight the variability in diagnostic opinions and emphasize the difficulties that exist within medical image interpretation. The disparities in these predictions underscore the complexities of handling inter-class predictions. With their unique training, experiences, and perspectives, each radiologist might perceive and interpret certain features differently. This variability is a testament to the multifaceted nature of medical diagnostics and raises questions about the standardization of interpretations. It further stresses the importance of developing advanced AI models that can account for such variabilities, ensuring that the technology is not just replicating human biases but is moving towards a more consistent and accurate diagnostic approach. Moreover, this visual disparity serves as a reminder of the critical role that continuous training and feedback play in refining human expertise and AI algorithms in medical imaging.

2.2 Proposed Methodology and Architecture

To foster the widespread acceptance and integration of AI in medical imaging, a novel approach for diagnosing COPD from CXR images has been devised, as illustrated in Figure 2. The model seamlessly combines the robust capabilities of ResNet50 and Xception, both known for their efficacy in image classification tasks. A distinct feature of this model is its reliance on the transfer learning approach. This methodology allows the model to utilize pre-trained data and knowledge from one task to improve performance on a related but distinct task. This is particularly advantageous in medical imaging, where obtaining large volumes of labeled data can be challenging.The architectural design of the model is divided into three steps:

The first, image pre-processing, focuses on enhancing the quality of the CXR images, ensuring that noise and artifacts were removed while preserving the essential features for accurate diagnosis. This step is crucial as the quality and clarity of the input image can significantly influence the diagnostic outcome. The second phase, methodology with transfer learning, capitalizes on the inherent strengths of ResNet50 and Xception. The model was able to achieve enhanced performance levels even with limited training data by harnessing the power of these pre-existing architectures and their learned features, ensuring a faster and more accurate diagnostic process. The final step underscores the importance of transparency and interpretability in AI models, especially in medical settings. By employing Gradient Class Activation Mapping (Grad-CAM)[32] and SHapley Additive exPlanations (SHAP)[33], the model provides visual and intuitive explanations for its predictions. Such insights bolster medical professionals' trust in the AI system and offer valuable feedback loops for further refinement and understanding of the model's decision-making process.

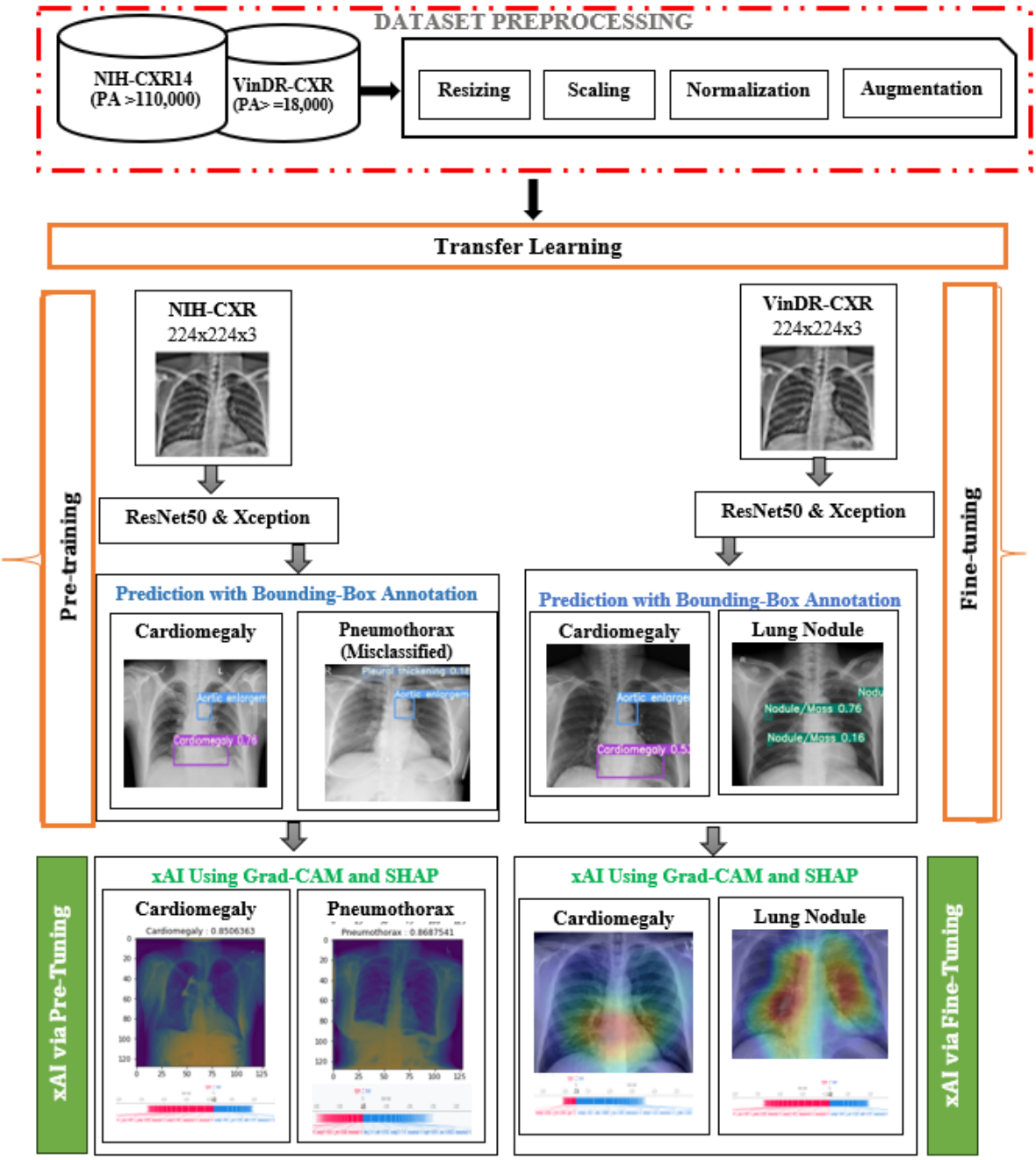


Fig. 2: The proposed architecture for xAI model for COPD Diagnosis from CXR images

2.3 Pre-processing Techniques

In the image processing phase, raw CXR images are pre-processed to improve image quality and accelerate the pertinent diagnostic features for COPD. This step entails resizing, normalization, and data augmentation techniques to ensure input images were compatible with the neural network, and to increase the diversity of the dataset. Following pre-processing, the ResNet-based model was trained using transfer learning. Using ImageNet weights as a backbone, the model took advantage of ResNet's robust feature extraction [29] capabilities and adapted them to the specific task of COPD diagnosis [34]. This method permits faster convergence and performs better than models trained from scratch.

2.4 Adapted Technique via Transfer Learning

This leverages two CNN-based models; ResNet50 and Xception, with ImageNet pre-trained weights from the TensorFlow Keras library. During the pre-training phase, these foundational models undergo training to classify 14 radiographic labels using the NIH ChestX-ray14 dataset, facilitating the learning of pulmonary features. Three primary radiographic labels - "Cardiomegaly", "Lung Nodule", and "Pneumothorax" were chosen from the fourteen based on their frequency in the ChestX-ray14 dataset.

In the subsequent fine-tuning phase, these pre-trained models were adapted to the VinDR-CXR dataset specifically for the COPD detection task. To expedite this fine-tuning, the weights for the initial 20% layers are kept constant, while adjustments are made to the subsequent layers. Given the uneven distribution of COPD cases, a class weights parameter was introduced to assign differential weights to the COPD and non-COPD classes. This weighting approach ensures that the loss function is regularized, attributing a higher significance to the COPD class loss and a lesser one to the non-COPD.

Data augmentation techniques were used to enhance model robustness and prevent overfitting by applying random transformations to the images (rotations, scaling, and horizontal flips). This process helps the model generalize better to new, unseen images by teaching it to recognize diagnostic features of COPD irrespective of variations in lung orientation. Essentially, data augmentation exposes the model to a wider array of illness manifestations, strengthening its diagnostic capabilities.

For both the pre-training and fine-tuning stages, a learning rate of 0.0001 is established, which undergoes a 7% decay with each epoch, and a batch size of 64. The choice of parameters in the model was based on extensive experimentation and prior research in the field[50],[51]. The Adam optimizer is employed in conjunction with the binary cross-entropy loss function. Training of the model was halted when no obvious improvement in the validation loss over five successive epochs were reached, thanks to the integration of a patience factor (an early termination parameter). Based on early stopping criteria, the model was trained for 50 epochs, ensuring sufficient training without overfitting.

2.5 Model Architecture: ResNet50

ResNet50 is designed to handle the vanishing gradient problem in deep neural networks through the use of residual connections. These connections act as shortcuts, allowing the input of a layer to be added to its output, forming a residual block. This approach helps the model learn the difference between the input and output, rather than the output itself. ResNet50 comprises of 50 layers, including convolutional layers, batch normalization, activation functions, and fully connected layers. For this specific application, the final layer was modified to a fully connected layer with two output nodes, enabling binary classification of COPD presence or absence. To enhance the model's interpretability, Layerwise-Grad-CAM[35] and SHAP[36] techniques were employed. Figure 3a illustrates the residual block, while Figure 3b shows the overall architecture, highlighting the input, functional sequential layers, and the final output layer designed for COPD classification.

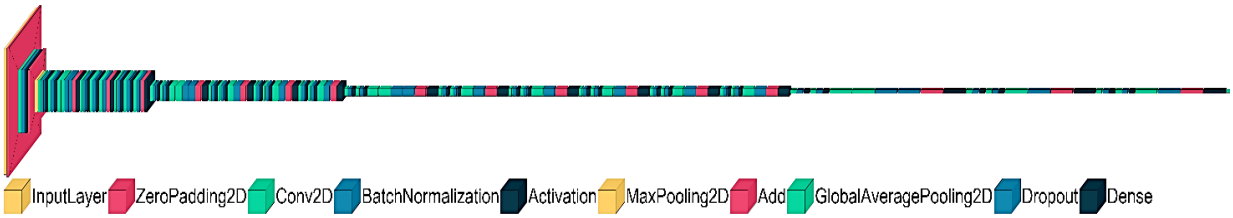


Fig. 3a: The Adapted ResNet-50 Model Architecture

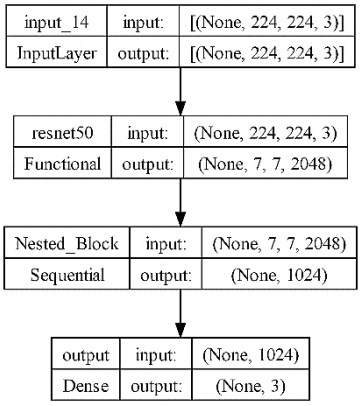


Fig. 3b: The ResNet50 model block depicting important layers

2.6 Model Architecture: Xception

Xception is a model that uses depthwise separable convolutions, a more efficient variant of regular convolutions. A depthwise separable convolution splits the input into channels, and applies a single filter to each channel, followed by a pointwise convolution that combines the outputs of the previous step, as depicted in Figure 4a. Originally proposed by Francois Chollet[46], the mastermind behind Keras and a renowned Google engineer, Xception stands as an "extreme" iteration of the Inception module. Xception has 36 layers, organized into three flows: entry, middle, and exit. Figure 4b depicts the functional blocks of the fine-tuned model.

The architecture is organized into three main flows. The entry flow consists of four blocks that progressively reduce the spatial dimensions of the input image. This is followed by the middle flow, composed of eight identical blocks designed to maintain these spatial dimensions while processing deeper features. The exit flow, comprising three blocks that reduces the spatial dimensions further while enhancing the model's ability to capture more complex patterns. Finally, the architecture culminates in a two-node fully connected layer tailored for binary classification, effectively determining the presence or absence of COPD. This structured design ensures efficient feature extraction and classification, crucial for accurate diagnostics.

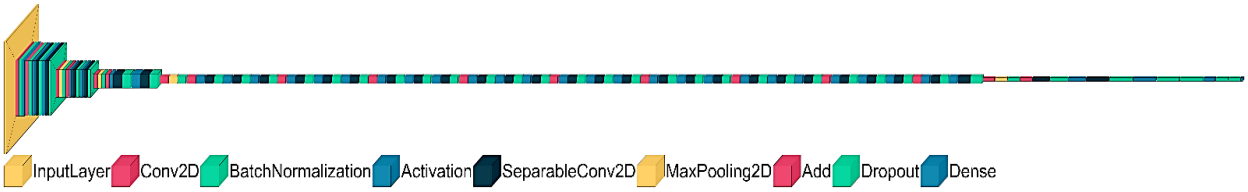


Fig. 4a: The Adapted Xception Model Architecture for COPD Diagnosis from CXR images

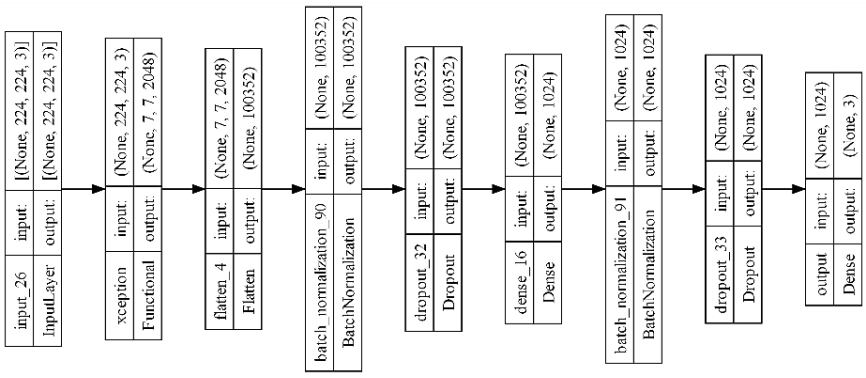


Fig. 4b: The Xception Model Block Depicting Important Layers

2.7 Training and Validation

The NIH-CXR14 dataset was split using the 80:20 rule corresponding to training, and testing respectively, and trained for 50 epochs. For the validation set, 20% of the VinDR-CXR dataset was utilized for hyperparameter tuning and improving the models performance. The model was trained on a Windows 11 PC with an Intel(R) Pentium(R) Core i7 8th Generation CPU clocked at 2.30GHz and a 6GB GeForce GTX 1060 graphics accelerator card, and the hyperparameters are presented in Table 1.

Table 1: Parameters for Training

Parameter	Value
Learning Rate	0.0001
Patience Factor	5 epochs
Batch Size	64
Learning Rate Decay	7% per epoch
Momentum	0.9
Optimization Function	Adam
Epochs	50

2.8 Explainable Techniques

The decision-making process comprises of two adapted techniques:

i. GradCAM: This technique uses the gradient information of a CNN model to produce a heatmap that shows the most relevant regions of an input image for a given prediction. Grad-CAM computes the importance of each neuron in the last convolutional layer of the CNN by multiplying the gradient of the class score by the neuron activation. The importance weights are then used to combine the activation maps of the last convolutional layer into a single heatmap. A ReLU function is applied to the heatmap to focus on the positive contributions, discarding negative values.

The formula for generating the Grad-CAM heatmap Hc{Grad-CAM} for a class *c* is presented in equation 1:

$$H_c^{Grad-CAM} = ReLU(\sum_n(\beta_c^n B^n)) \quad (1)$$

Where:

$H_c^{Grad-CAM}$ = Grad-CAM heatmap for class c

B_n = the activation map of feature n in the last convolutional layer

B^c_n = importance weight of feature n for class c, calculated as shown in equation 2.

$$\beta_c^n = \frac{1}{W}\sum_{k=l}\frac{\partial y_c}{\partial B_{kl}^n} \quad (2)$$

W = a normalization factor

y_c = the class score for class c

ReLU is the rectified linear unit activation function

Grad-CAM highlights salient features of an input image associated with a specific class prediction, providing insight into the decision-making process of the CNN model.

ii. SHAP: This method assigns an importance score to each feature of an input based on how much it contributes to a model's prediction. SHAP is based on the concept of Shapley values, which are derived from game theory. Shapley values measure the marginal contribution of each feature by averaging over all possible subsets of features[37]. To calculate the SHAP value for a feature, it is essential to examine the difference between the model prediction with and without the presence of that feature.

Additionally, interactions between features and the sequence in which they are introduced or eliminated should be considered. This involves using a weighted average over all potential combinations of features. The weights are determined based on the size of the subset and the total number of features.

The formula for computing the SHAP value for feature i is based on equation 3:

$$\phi_i(x) = \sum_{z \subseteq M \setminus \{i\}} i \frac{|z|!(|M|-|z|-1)!}{|M|!} (f(h(z \cup \{i\})) - f(h(z))) \quad (3)$$

Where:

- M = the set of all features
- x = the input instance
- f = the model function
- h = a function that maps a binary vector z to an input sample by removing the features set to 0 in z from x
- $|\cdot|$ = the cardinality operator that counts the number of non-zero elements in a set
- z = a binary vector that represents the presence or absence of each feature
- $\phi i(x)$ = the SHAP value for feature i

KernelSHAP, which combines LIME and Shapley values[49], was used to approximate the SHAP values. It trains a linear model on data points sampled from the simplified input space $\{0,1\}$M, using $f \circ h$ outputs as labels. The optimal coefficients of the linear model are then used as the SHAP values. The data points are weighted by a kernel function that depends on their distance from the original input x.

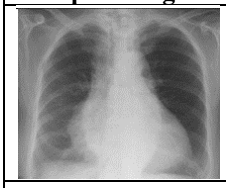
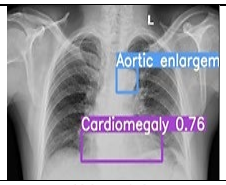
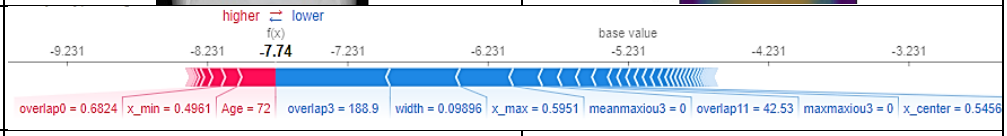
3 Results

The results of the proposed explainable model (xAI) that incorporates Grad-CAM and SHAP features for COPD diagnosis is presented in this section. Tables 2-4 illustrate the outcome of the explainable model on two select cases of COPD: Cardiomegaly and Pneumothorax, using the "pre-tuning" approach. In contrast, the result of "fine-tuning" the model is highlighted in Table 4 for the third case (Lung Nodule).

Case 1: Cardiomegaly:

Evident from a chest radiograph, Cardiomegaly is diagnosed when the cardiothoracic ratio (CTR) exceeds 50% on a posteroanterior (PA) view. Specific chambers responsible for the heart's enlargement can be identified by correlating with additional radiographic signs in the thoracic region.

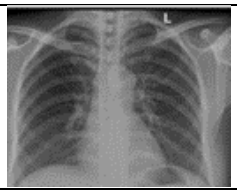
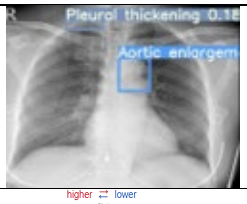
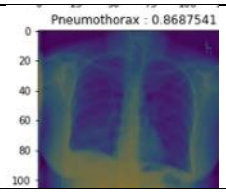
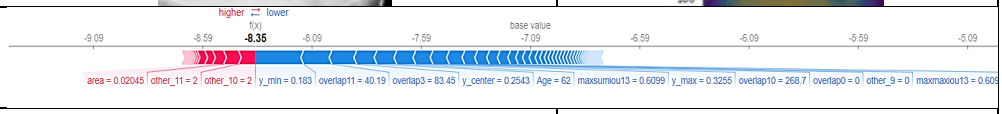
Table 2: xAI (Grad-CAM) ouput with a BBox Acc = 76%, & xAI Acc = 85%

Input Image	BBox Prediction	xAI (GradCAM+SHAP)
		
		
Case 1: Cardiomegaly	The model prediction of the bounding box (BBox) with 76% accuracy.	The result of **Pre-tuning** xAI with 85% confidence using Grad-CAM and the corresponding SHAP values

Case 2: Pneumothorax:

Termed medically as Pneumothorax refers to the accumulation of gas, typically air, within the pleural cavity. A severe variant, tension pneumothorax, arises when this gas accumulation expands, exerting pressure on the mediastinal structures. Initial radiographic evaluations might overlook a concealed pneumothorax, primarily when conducted in a supine or semi-upright position.

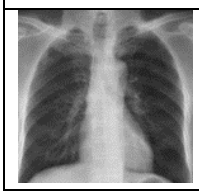
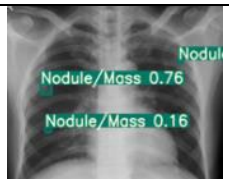
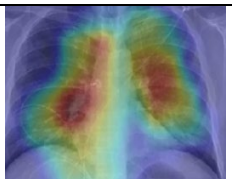
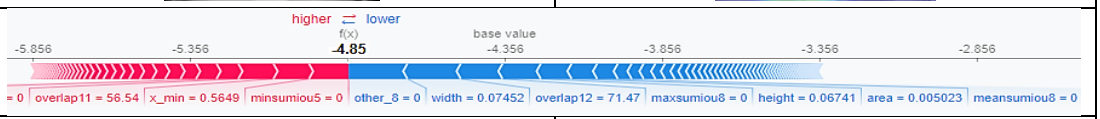
Table 3: xAI (Grad-CAM) ouput with a BBox Acc = 73%, & xAI Acc = 86.7%

Input Image	BBox Prediction	xAI (GradCAM+SHAP)
		
		
Case 2: Pneumothorax	**PS:** This was a special case of misclassification, where the model predicted Aortic enlargement with 73% accuracy, while the ground-truth image with xAI approaches correctly identified Pneumothorax	The result of **Pre-tuning** xAI with 86.7% confidence using Grad-CAM and the corresponding SHAP values.

Case 3: Lung Nodule:

A lung nodule is a small, rounded growth or spot visible on a chest imaging study, typically a radiograph or CT scan. While many nodules are benign and may result from infections or scars, some could be early indicators of lung cancer. Monitoring their size and characteristics over time is crucial to determine their nature and potential clinical implications. Regular imaging follow-ups are often recommended to track the nodule's appearance changes.

Table 4: xAI (Grad-CAM) ouput with a BBox Acc = 76%, & xAI Acc = 93%

Input Image	BBox Prediction	xAI (GradCAM+SHAP)
		
		
Case 3: Lung Nodule	The model prediction of the bounding box (BBox) with 76% accuracy.	The result of **fine-tuning** xAI with 93% confidence using Grad-CAM and the corresponding SHAP values

Trained for 50 epoch, the graph for the training and validation is highlighted in Figures 5 and 6. With a validation accuracy of 95% for the Xception model via fine-tuning, the proposed approach was better at discriminating COPD from CXR images.

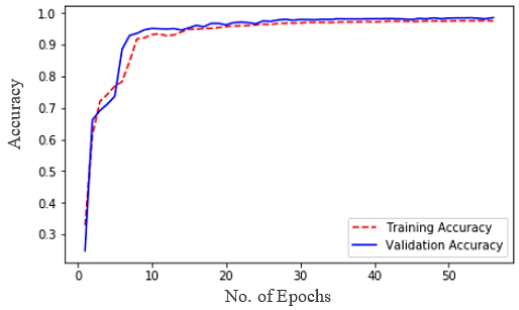


Fig. 5: Accuracy Curve after 50 Epochs

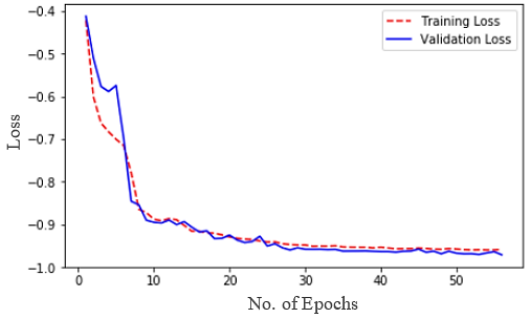


Fig. 6: The Loss Curve after 50 Epochs

3.1 Evaluation Metrics

The performance of the xAI model was evaluated on the validation set using standard metrics as illustrated in equations 4-7. These metrics comprehensively assessed the model's ability to diagnose COPD from chest X-ray images. Table II highlights the results obtained.

$$Accuracy(Acc) = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

$$Precsion(P)_{(x,y)} = \frac{TP_{(x,y)}}{TP_{(x,y)}+FP_{(x,y)}} \quad (5)$$

$$Recall(R)_{(x,y)} = \frac{TP_{(x,y)}}{TP_{(x,y)}+FN_{(x,y)}} \quad (6)$$

$$F-Score_{(x,y)} = {2|P*R|}/{|P|+|R|} \quad (7)$$

Where:Acc = Accuracy
TP = True Positives
TN = True Negatives
FP = False Positives
FN = False Negatives
$P_{(x,y)}$ = Precision for classes x and y
$R_{(x,y)}$ = Recall for classes x and y
$F\text{-}Score_{(x,y)}$ = F-Score for classes x and y

3.2 Performance Analysis

The performance of these architectures (Xception and ResNet50), across both pre-training and fine-tuning approaches, are presented in Table 5.

Table 5: Model Performance via Pre-Training and Fine-Tuning

Metrics	Pre-Training		Fine-Tuning	
	Xception	ResNet50	Xception	ResNet50
Accuracy(%)	94.7	93.0	95.0	92.5
Recall(%)	96.0	94.2	98.2	96.3

Based on the findings in Table 5, the following insights were drawn:

Accuracy:
In the pre-training phase, Xception exhibits a slightly higher accuracy of 94.7% compared to ResNet50's 93.0%, and in the fine-tuning phase, Xception further improves its Accuracy to 95.0%, while ResNet50 sees a slight dip to 92.5%. This indicates that Xception benefits more from fine-tuning than ResNet50 in terms of Accuracy.

Recall:
For the pre-training approach, Xception achieves a recall of 96.0%, which is marginally better than ResNet50's 94.2%. In the fine-tuning phase, both models show an improvement in Recall. Xception reaches an impressive 98.2%, whereas ResNet50 also improves to 96.3%. This suggests that fine-tuning enhances the models' ability to identify positive cases correctly.

The Xception model consistently outperforms ResNet50 in Accuracy and recall metrics across pre-training and fine-tuning stages. However, it is worth noting that both architectures demonstrate commendable performance, with

only slight differences between them. The fine-tuning process appears to be particularly beneficial for enhancing the Recall of both models. Further, the performance on select cases of COPD is presented in Table 6.

Table 6: Model Performance on Select COPD Cases via Fine-tuning

Cases	Acc (%)	Precision (%)	Recall (%)	F-Score (%)
Cardiomegaly	85.0	92.5	91.0	91.7
Pneumothorax	86.7	93.0	92.0	92.5
Lung Nodule	**93.0**	**97.0**	**96.5**	**96.7**

The efficacy of the special cases (Case 1 to Case 3) is presented in Table 6. The best performance was for Case 3 (Lung Nodule), as highlighted in bold figures. The summary of the findings are:

Case 1: Cardiomegaly

The model correctly classifies cardiomegaly cases 85.0% of the time. Of all the predicted cardiomegaly cases, 92.5% are actual cardiomegaly cases. And of all the actual cardiomegaly cases, the model identifies 91.0% of them. The model achieves a balanced score of 91.7%, indicating a good harmony between precision and Recall.

Case 2: Pneumothorax

The model has an accuracy of 86.7% for detecting Pneumothorax. 93.0% of the predicted pneumothorax cases are true cases. The model captures 92.0% of all actual pneumothorax cases. With an F-Score of 92.5%, the model demonstrates a well-balanced performance between precision and recall for pneumothorax detection.

Case 3: Lung Nodule

The model's lung nodule detection accuracy is a commendable 93.0%. A high precision of 97.0% suggests that most of the predicted lung nodule cases are accurate. The model successfully identifies 96.5% of all true lung nodule cases. Achieving an F-Score of 96.7% indicates that the model maintains an excellent balance between precision and Recall for lung nodule classification.

Overall, The model demonstrates strong performance across all three cases, with Lung Nodule detection being the best in precision, Recall, and overall Accuracy. The results suggest that the model is particularly adept at identifying and classifying lung nodules, followed closely by its capability in detecting COPD from CXR images.

Other applications of the proposed algorithm is to various medical imaging tasks beyond COPD diagnosis. For instance:

-**Pneumonia Detection:** Using CXR images, similar methodologies can enhance the detection and classification of different pneumonia types.

-**Lung Cancer Screening**: The algorithm can assist in identifying early-stage lung nodules from CXR or CT images.

- **Tuberculosis Detection:** The same approach can be utilized for detecting tuberculosis, particularly in resource-limited settings where advanced diagnostic tools are scarce.

- **Cardiomegaly Diagnosis:** As demonstrated in the study, the algorithm is effective in diagnosing heart enlargement conditions.

These applications demonstrate the versatility and robustness of the proposed approach across different medical imaging challenges.

3.3 Analysis with Similar Research

This section comprehensively explains how the proposed method stands compared to existing solutions using several state-of-the-art techniques for benchmarking.

CheXNet [38]: This method employs a 121-layer CNN that inputs a chest X-ray image and outputs the probability of pneumonia along with a heatmap localizing the areas of the image most indicative of pneumonia. The method achieves radiologist-level performance on pneumonia detection and outperforms previous state-of-the-art approaches on all 14 diseases in the ChestX-ray14 dataset. The method does not require manual feature selection or domain expertise but learns the features directly from the data.

DenseNet with CAMs [39]: A Deep CNN for COVID-19 Detection on Chest X-Rays: This method employs a 121-layer CNN that inputs a chest X-ray image and outputs the probability of COVID-19 along with a heatmap localizing the areas of the image most indicative of COVID-19. The method achieves high Accuracy and sensitivity in detecting COVID-19 from CXRs, and outperforms other methods that use shallow CNN architectures or handcrafted features.

A Neural Network for Other Pneumonia Detection on Chest Images [40]: This method employs a neural network that inputs a chest image and outputs the probability of lung disorders. The method uses feature reuse residual block and depthwise dilated convolutions to handle high-dimensional data and extract features from both CT and X-ray chest images. The method achieves high Accuracy and sensitivity in detecting Covid-19 and other pneumonia cases from chest images, outperforming other methods using SVM with PCA or different neural network architectures.

Ensemble Methods using AdaBoost [41-43]: These techniques leverage ensemble learning through AdaBoost to improve the classification performance. While the method shows resilience against overfitting, it lacks the capability for automatic feature learning, which can be crucial for capturing complex patterns in medical images.

Deep Belief Networks (DBN) with Manual Feature Fusion [43-45]: This deep learning technique uses a stack of restricted Boltzmann machines to learn high-level features from multiple data sources, such as textual medical records and images. The features from different sources are then manually concatenated and fed into a classifier, such as a softmax layer. This technique can learn complex data representations but has some drawbacks, such as low model interpretability and high computational cost. These comparisons can be found in Table 7.

Table 7: Comparison with Relevant Authors

S/No	Author/Year	Dataset	Recall (%)	Accuracy (%)
[38]	Rajpukar et al. 2017	ChestX-ray14	NA	85.0
[39]	Kikkisetti et al. 2020	Pneumonia CXR	79.0	79.0
[40]	Gaffari Celik 2023	Covid-19 CXR	NA	**99.8**
[41]	Sharma & Guleria 2023	Lung CXR	93.1	92.2
[42]	Alshmrani et al. 2023	Lung CXR	93.8	96.5
[43]	Xue et al. 2023	CXR & Lung-CT	NA	99.0
[44]	Wang et al. 2023	MIMIC-IV CXR	NA	72.5
[45]	Wen et al. 2023	NIH CXR & Pediatric CXR	84.0	80.3
xAI	Agughasi et al. 2023	VinDR-CXR	**96.5**	93.0

4 Discussion

Table 7 compares the proposed model and several other deep learning models developed for analyzing chest X-ray (CXR) images, as published by various authors from 2017 to 2023. The primary metrics used to compare the models are Recall (which highlights the portion of actual positives correctly identified) and Accuracy (the percentage of total correct predictions). It was observed that xAI, proposed by the author, achieves the best Recall of 96.5% and an accuracy of 93.0% on the VinDR-CXR validation dataset. This performs better than most studies presented in both Recall and Accuracy, indicating that the model correctly identifies positive cases (patients with COPD) and makes correct predictions overall.

The model proposed by Celik[40] using the Covid19 X-ray dataset, demonstrates similar Recall and accuracy scores. However, it is worth noting that the application differs significantly from that of xAI, focusing on Covid-19 X-ray data.

The high performance of xAI suggests that the ensemble approach combining a pretrained ResNet50 model with an Xception model is a successful strategy for this task. It further indicates that the specific preprocessing steps and architectural adaptations made to tailor the model for xAI diagnosis are beneficial.

However, it is essential to remember that the reliability and comparability of these results depend on the dataset used, how it was collected, and the context of the predictions. It is also crucial to consider other metrics such as specificity and AUC-ROC to fully evaluate these models' performance. While the results demonstrate promising performance for the xAI model, it is still vital to carry out further testing, particularly on independent and diverse datasets, to ensure its generalizability and reliability.

4.1 Explainability Using SHAP Values

Case 1: Cardiomegaly

The SHAP values for case 1 offer a comprehensive view of the model's decision-making process in diagnosing Cardiomegaly from the provided CXR image. The f(x) value of -7.74 is a central indicator of the model's prediction, suggesting a leaning towards a specific classification direction.
On the left side of the SHAP chart, highlighted in red, attributes contribute towards pushing the model's prediction away from a positive classification of Cardiomegaly. Notably: - "overlap0" with a value of 0.6824, and "x_min" with a value of 0.4961 represent the PA-CXR of a 72-year-old patient.
These values imply factors or regions in the image that might be inconsistent with typical manifestations of Cardiomegaly or other aspects that reduce the likelihood of a positive diagnosis.
In contrast, attributes on the right side, highlighted in blue, push the model's prediction towards a positive classification of Cardiomegaly. Among these, an "overlap3" stands out with a significant value of 188.9. Further pointers were: "width" with a value of 0.09896, "x_max" at 0.5951, "overlap11" with a value of 42.53, "area" at 0.005023, and "x_center" with a value of 0.5456
Attributes like "meanmaxiou3" and "maxmaxious", both at 0, might not have played a decisive role in this prediction.

Case 2: Pneumothorax

The SHAP values for the CXR image serve as a window into the model's decision-making process in diagnosing Pneumothorax. The central f(x) value of -8.35 indicates the model's inclination for this diagnosis, hinting at a particular direction in the classification.
On the left side of the SHAP chart (highlighted in red), some attributes that contribute towards the model's prediction away from a positive classification of Pneumothorax an "area" with a value of 0.002045, "other_11" and "other_10", both with a value of 2
These values signify factors or regions in the image that might be incongruent with classic manifestations of Pneumothorax or other aspects that lower the likelihood of a positive diagnosis.
In contrast, attributes on the right side, highlighted in blue, steer the model's prediction towards a positive classification of Pneumothorax. Among these are "y_min" with a value of 0.183, an "overlap3" with a significant value of 83.45, and an "maxumiou3" with a value of 0.6099, all for a 62-year-old patient. These attributes present factors or regions in the CXR image that align with typical indications of Pneumothorax, pushing the model towards recognizing its presence.
While the attributes on the right (blue) suggest factors supporting the presence of Pneumothorax, the overall negative f(x) value of -8.35 implies the characteristics on the left (red) have a more significant influence on the model's decision, leading to a negative prediction for Pneumothorax in this instance. This intricate interplay of factors captured by the SHAP values underscores the value of interpretability tools, ensuring the validity and transparency of AI-driven predictions in medical diagnostics.

Case 3: Lung Nodule

The SHAP values provide an insight into the model's decision-making process for a patient diagnosed with a Lung Nodule. On the left side of the SHAP chart (Table 4 (Appendix B), highlighted in red, the attributes contribute towards pushing the model's prediction in the opposite direction of the positive class (i.e., indicating the absence of a Lung Nodule). The most significant of these attributes is "overlap11" with a SHAP value of 56.54, followed by "x_min" with a value of 0.5649. Notably, the attribute "minsumious5" is at 0, implying it might not be a decisive factor in this prediction.
Conversely, on the right side of the chart (highlighted in blue), the attributes contribute towards a positive classification (i.e., indicating the presence of a Lung Nodule). Among these attributes, "overlap12" with a SHAP value of 71.47 is a significant contributor. Other attributes such as "width", "height", and "area" have SHAP values of 0.07452, 0.06741, and 0.005023, respectively, suggesting their influence on the positive prediction.

Attributes like “other_8”, “maxumious8”, and “meanumious” are at 0, indicating they might not have influenced this specific prediction.

In summary, the performance metric on select COPD cases is presented in Figure 7. It is essential to understand that making *f(x)* positive for a specific instance by tweaking the model was not the best approach since it compromises the Accuracy and introduces biases. These SHAP values offer a granular understanding of the model’s prediction for this patient. The attributes on the left (red) seem to counteract or negate the presence of a Lung Nodule, while the ones on the right (blue) support its presence. Given the central f(x) value of -4.85, it suggests the attributes opposing the positive classification might have a stronger influence, leading to the model’s decision. However, significant attributes on both sides imply a complex interplay of factors, and a careful evaluation might be necessary for clinical applications. Since the core aim was improving the model’s general performance and interpretability, incorporating Grad-CAM was necessary to validate the predictions made.

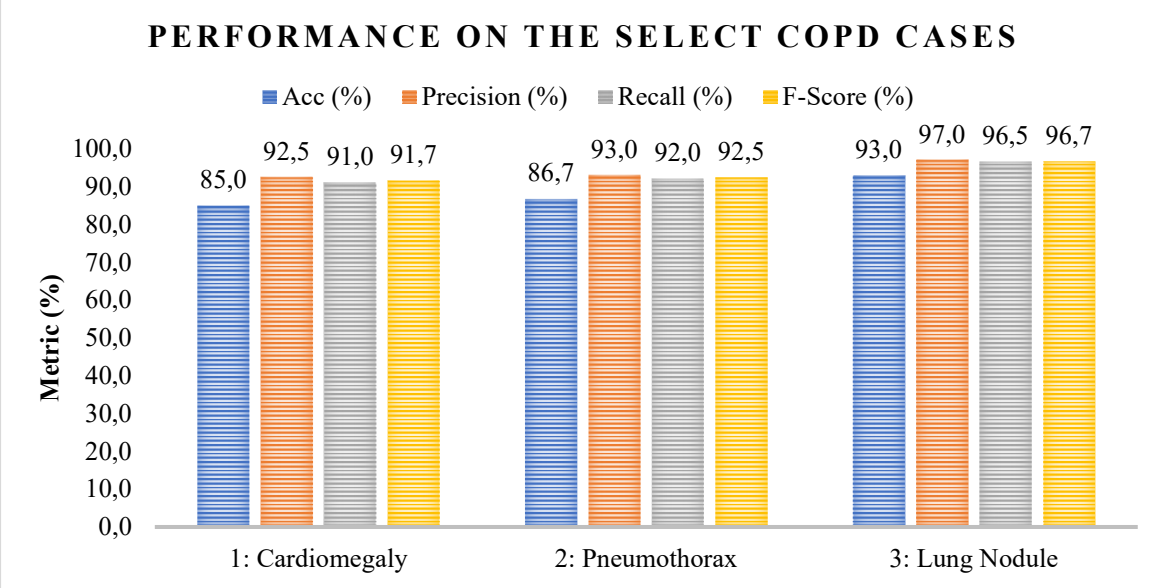


Fig. 7: The Performance Metric on 3 Select COPD Cases

4.2 EXPLAINABILTY USING GRAD-CAM

Grad-CAM heatmaps of true positive cases were visualized on the Xception base model for both NIH-CXR and the VinDR-CXR test datasets. As presented in Tables 2-4, these heatmaps capture the actual positive cases of the Xception model. An averaged heatmap was derived by averaging the heatmaps with predicted probabilities exceeding 0.75. This heatmap reveals that the Xception base model predominantly concentrates on the upper right side of the lungs, specifically between the upper lung and the trachea. This observation underscores the association of COPD with the upper airway. The findings reveal that the model could capture the prevalence of Pneumothorax with a greater accuracy. Thus, the model could be integrated into existing medical imaging systems that can assist radiologists and improve diagnostic Accuracy. Moreover, it could also be adapted for other types of respiratory diseases, extending its utility and impact.

5 Conclusion and future work

The research introduces a novel approach to diagnosing Chronic Obstructive Pulmonary Disease (COPD) using Chest X-ray (CXR) images, specifically focusing on posteroanterior (PA-CXR) views. By leveraging advanced deep learning frameworks such as ResNet50 and Xception, the model was tailored to the needs of specific demands of COPD detection, achieving optimal outcomes. The incorporation of interpretability tools like Grad-CAM and SHAP enhances the model’s relevance, fostering trust and confidence among medical professionals.

However, while the model demonstrates impressive recall rates, particularly for pneumothorax detection, it is crucial to acknowledge its limitations. The primary focus on PA-CXR images inadvertently excludes other vital perspectives, such as anteroposterior (AP-CXR) and lateral views, potentially overlooking the full spectrum of COPD manifestations. Additionally, although CT imaging is often considered the gold standard for visualizing lung regions, the model's adaptability to diverse real-world scenarios and further clinical validation remains a significant research area.

Future work should broaden the research scope by evaluating the model's performance across a wider array of CXR perspectives, including AP and lateral views, to ensure robustness and versatility. Integrating additional imaging modalities, such as computed tomography (CT) and lung ultrasonography, could provide a more comprehensive understanding of pulmonary conditions, thereby enhancing diagnostic capabilities.

While the current research has laid a solid foundation, the journey toward a holistic and universally adaptable model for COPD diagnosis is ongoing. Adequate plans for real experimental validation are underway, despite the challenges associated with using real-patient datasets. These plans include testing the algorithm on a diverse set of patient data from various medical institutions, comparing the model's predictions with clinical diagnoses made by expert radiologists, and evaluating the model's performance in real-time clinical workflows to ensure its practical applicability and reliability.

By pursuing these avenues in future investigations, the xAI framework stands to play a pivotal role in the proactive identification and treatment of COPD, ultimately enhancing patient care.

References

[1] Singh, Dave, Alvar Agusti, Antonio Anzueto, Peter J. Barnes, Jean Bourbeau, Bartolome R. Celli, Gerard J. Criner, "Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease: the GOLD science committee report 2019," Eur. Respir. J., vol. 53, no. 5, 2019. https://doi: 10.1183/13993003.00164-2019

[2] Patel AR, Patel AR, Singh S, Singh S, Khawaja I. Global Initiative for Chronic Obstructive Lung Disease: The Changes Made. Cureus. 2019 Jun 24;11(6):e4985. https://doi: 10.7759/cureus.4985. PMID: 31453045; PMCID: PMC6701900.

[3] Sarkar M, Bhardwaz R, Madabhavi I, Modi M. Physical signs in patients with chronic obstructive pulmonary disease. Lung India. 2019 Jan-Feb;36(1):38-47. doi: 10.4103/lungindia.lungindia_145_18. PMID: 30604704; PMCID: PMC6330798.

[4] A. V. Ikechukwu and S. Murali, "xAI: An Explainable AI Model for the Diagnosis of COPD from CXR Images," in 2023 IEEE 2nd International Conference on Data, Decision and Systems (ICDDS), Dec. 2023, pp. 1–6. doi: 10.1109/ICDDS59137.2023.10434619.

[5] J. M. Haynes, D. A. Kaminsky, and G. L. Ruppel, "The Role of Pulmonary Function Testing in the Diagnosis and Management of COPD," Respir Care, vol. 68, no. 7, pp. 889–913, Jul. 2023, doi: 10.4187/respcare.10757.

[6] A. Prasse, C. Quartucci, G. Zissel, G. Kayser, J. Müller-Quernheim, and B. C. Frye, "Interstitial Lung Diseases of Occupational Origin," in Orphan Lung Diseases, V. Cottin, L. Richeldi, K. Brown, and F. X. McCormack, Eds., Cham: Springer International Publishing, 2023, pp. 641–669. doi: 10.1007/978-3-031-12950-6_37.

[7] D. Iglesias and A. M. Malchuk, "Health Inequities and Structural Racism," in Chronic Illness Care, T. P. Daaleman and M. R. Helton, Eds., Cham: Springer International Publishing, 2023, pp. 567–579. doi: 10.1007/978-3-031-29171-5_42.

[8] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique," IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1153–1159, May 2016. https://doi: 10.1109/TMI.2016.2553401

[9] Huiyan Jiang, Zhaoshuo Diao, Tianyu Shi, Yang Zhou, Feiyu Wang, Wenrui Hu, Xiaolin Zhu, Shijie Luo, Guoyu Tong, Yu-Dong Yao, A review of deep learning-based multiple-lesion recognition from medical images: classification, detection and segmentation,Computers in Biology and Medicine, ISSN 0010-4825,https://doi.org/10.1016/j.compbiomed.2023.106726.

[10] A. V. Ikechukwu, M. S, and H. B, "COPDNet: An Explainable ResNet50 Model for the Diagnosis of COPD from CXR Images," in 2023 IEEE 4th Annual Flagship India Council International Subsections Conference (INDISCON), Mysore, India: IEEE, Aug. 2023, pp. 1–7. doi: 10.1109/INDISCON58499.2023.10270604.

[11] Zafar, Imran, Shakila Anwar, Faheem Kanwal, Waqas Yousaf, Fakhar Un Nisa, Tanzeela Kausar, Qurat Ul Ain, (2023) Reviewing methods of deep learning for intelligent healthcare systems in genomics and biomedicine, Biomedical Signal Processing and Control, Volume 86, Part B,https://doi.org/10.1016/j.bspc.2023.105263.
[12] Hasenstab, Kyle A., Nancy Yuan, Tara Retson, Douglas J. Conrad, Seth Kligerman, David A. Lynch, Albert Hsiao, and for the COPDGene Investigators. "Automated CT Staging of Chronic Obstructive Pulmonary Disease Severity for Predicting Disease Progression and Mortality with a Deep Learning Convolutional Neural Network." Radiology: Cardiothoracic Imaging 3, no. 2 (2021): e200477. https://doi.org/10.1148/ryct.2021200477..
[13] Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., Kieseberg, P. & Holzinger, A., (2018). Explainable AI: The New 42?. In: Holzinger, A., Kieseberg, P., Tjoa, A., Weippl, E. (eds) Machine Learning and Knowledge Extraction. CD-MAKE 2018. Lecture Notes in Computer Science(), vol 11015. Springer, Cham. https://doi.org/10.1007/978-3-319-99740-7_21
[14] Sebastian Bach, Alexander Binder, Grégoire Montavon, Frederick Klauschen, Klaus-Robert Müller, Wojciech Samek, "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation," PLoS ONE, vol. 10, no. 7, Jul. 2015, Art. no. e0130140. https://doi: 10.1371/journal.pone.0130140
[15] Ramprasaath R. Selvaraju,Michael Cogswell,Abhishek Das,Ramakrishna Vedantam,Devi Parikh, Dhruv Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," in Proc. of the IEEE International Conf. on Computer Vision (ICCV), Venice, Italy, 2017, pp. 618-626. https://doi: 10.1109/ICCV.2017.74
[16] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, Li Fei-Fei , "ImageNet Large Scale Visual Recognition Challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211-252, Dec. 2015. https://doi: 10.1007/s11263-015-0816-y
[17] S. I. Nafisah, G. Muhammad, M. S. Hossain, and S. A. AlQahtani, "A Comparative Evaluation between Convolutional Neural Networks and Vision Transformers for COVID-19 Detection," Mathematics, vol. 11, no. 6, p. 1489, Mar. 2023, doi: 10.3390/math11061489.
[18] S. Iqbal, A. N. Qureshi, J. Li, and T. Mahmood, "On the Analyses of Medical Images Using Traditional Machine Learning Techniques and Convolutional Neural Networks," Arch Computat Methods Eng, vol. 30, no. 5, pp. 3173–3233, Jun. 2023, doi: 10.1007/s11831-023-09899-9.
[19] Hoo-Chang Shin, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, Ronald M. Summers., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1285-1298, May 2016. https://doi: 10.1109/TMI.2016.2528162
[20] A. Victor Ikechukwu, S. Murali, R. Deepu, and R. C. Shivamurthy, "ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images," Global Transitions Proceedings, vol. 2, no. 2, pp. 375–381, Nov. 2021, doi: 10.1016/j.gltp.2021.08.027
[21] A. Victor Ikechukwu and M. S, "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, p. 025021, Jun. 2023, https://doi: 10.1088/2632-2153/acd2a5.
[22] A. V. Ikechukwu and S. Murali, "i-Net: a deep CNN model for white blood cancer segmentation and classification," IJATEE, vol. 9, no.95, Oct. 2022, https://doi: 10.19101/IJATEE.2021.875564.
[23] A. Victor Ikechukwu, P. Sreyas, A. Sena, H. Preetham, and K. Raksha, "Explainable Deep Learning Model for Covid-19 Diagnosis," IRJMETS, vol. 04, no. 07, pp. 3051–3059, Jul. 2022.vol. 10, no. 11, p.e0141287, 2015, doi: https://doi.org/10.1371/journal.pone.0141287.
[24] A. Ait Nasser and M. A. Akhloufi, "A Review of Recent Advances in Deep Learning Models for Chest Disease Detection Using Radiography," Diagnostics, vol. 13, no. 1, p. 159, Jan. 2023, https://doi: 10.3390/diagnostics13010159.
[25] M. Chetoui, M. A. Akhloufi, E. M. Bouattane, J. Abdulnour, S. Roux, and C. D. Bernard, "Explainable COVID-19 Detection Based on Chest X-rays Using an End-to-End RegNet Architecture," Viruses, vol. 15, no. 6, p. 1327, Jun. 2023, https://doi: 10.3390/v15061327.
[26] L. Brunese, F. Mercaldo, A. Reginelli, and A. Santone, "Explainable Deep Learning for Pulmonary Disease and Coronavirus COVID-19 Detection from X-rays," Computer Methods and Programs in Biomedicine, vol. 196, p. 105608, Nov. 2020, https://doi: 10.1016/j.cmpb.2020.105608.

[27] Md. Nahiduzzaman, Md. Omaer Faruq Goni, Rakibul Hassan, Md. Robiul Islam, Md Khalid Syfullah, Saleh Mohammed Shahriar, Md. Shamim Anower, Mominul Ahsan, Julfikar Haider, Marcin Kowalski, "Parallel CNN-ELM: A multiclass classification of chest X-ray images to identify seventeen lung diseases including COVID-19," Expert Systems with Applications, vol. 229, p. 120528, Nov. 2023, https://doi: 10.1016/j.eswa.2023.120528.
[28] S. Nazir, D. M. Dickson, and M. U. Akram, "Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks," Computers in Biology and Medicine, vol. 156, p. 106668, Apr. 2023, https://doi: 10.1016/j.compbiomed.2023.106668.
[29] V. I. Agughasi and M. Srinivasiah, "Semi-supervised labelling of chest x-ray images using unsupervised clustering for ground-truth generation," AET, vol. 2, no. 3, pp. 188–202, Sep. 2023, doi: 10.31763/aet.v2i3.1143.
[30] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "ChestX-ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases".
[31] Ha Q. Nguyen, Khanh Lam, Linh T. Le, Hieu H. Pham, Dat Q. Tran, Dung B. Nguyen, Dung D. Le, Chi M. Pham, Hang T. T. Tong, Diep H. Dinh, Cuong D. Do, "VinDr-CXR: An open dataset of chest X-rays with radiologist's annotations," arXiv:2012.15029 [eess], Mar. 2022, Accessed: Mar. 28, 2022. [Online]. Available: http://arxiv.org/abs/2012.15029
[32] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 618-626. https://doi: 10.1109/ICCV.2017.74
[33] M. Mamalakis, K. Dwivedi, M. Sharkey, S. Alabed, D. Kiely, and A. J. Swift, "A transparent artificial intelligence framework to assess lung disease in pulmonary hypertension," Sci Rep, vol. 13, no. 1, p. 3812, Mar. 2023, doi: 10.1038/s41598-023-30503-4
[34] Nawaz, Marriam, Tahira Nazir, Jamel Baili, Muhammad Attique Khan, Ye Jin Kim, and Jae-Hyuk Cha. 2023. "CXray-EffDet: Chest Disease Detection and Classification from X-ray Images Using the EfficientDet Model" Diagnostics 13, no. 2: 248. https://doi.org/10.3390/diagnostics13020248.
[35] Shagun Sharma, Kalpna Guleria(2023), A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks, Procedia Computer Science,Volume 218,2023,https://doi.org/10.1016/j.procs.2023.01.018.
[36] A. A. Huang and S. Y. Huang, "Dendrogram of transparent feature importance machine learning statistics to classify associations for heart failure: A reanalysis of a retrospective cohort study of the Medical Information Mart for Intensive Care III (MIMIC-III) database," PLoS ONE, vol. 18, no. 7, p. e0288819, Jul. 2023, doi: 10.1371/journal.pone.0288819.
[37] A. Victor Ikechukwu and M. S, "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, p. 025021, Jun. 2023, https://doi: 10.1088/2632-2153/acd2a5.
[38] Rajpurkar, P., Irvin, J., Ball, R. L., Zhu, K., Yang, B., Mehta, H., & Lungren, M. P. (2018). Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS medicine, 15(11), e1002686.
[39] Kikkisetti, S., Zhu, J., Shen, B., Li, H., & Duong, T. Q. (2020). Deep-learning convolutional neural networks with transfer learning accurately classify COVID-19 lung infection on portable chest radiographs. PeerJ, 8, e10309.
[40] Celik, G. (2023). Detection of Covid-19 and other pneumonia cases from CT and X-ray chest images using deep learning based on feature reuse residual block and depthwise dilated convolutions neural network. Applied Soft Computing, 133, 109906.
[41] Goram Mufarah M. Alshmrani, Qiang Ni, Richard Jiang, Haris Pervaiz, Nada M. Elshennawy (2023), A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images, Alexandria Engineering Journal, Volume 64, https://doi.org/10.1016/j.aej.2022.10.053.
[42] S. Bhimshetty and A. V. Ikechukwu, "Energy-efficient deep Q-network: reinforcement learning for efficient routing protocol in wireless internet of things," Indonesian Journal of Electrical Engineering and Computer Science, vol. 33, no. 2, Art. no. 2, Feb. 2024, doi: 10.11591/ijeecs.v33.i2.pp971-980.
[43] Bin Wang, Yuanxiao Li, Ying Tian, Changxi Ju, Xiaonan Xu, Shufen Pei, (2023) Novel pneumonia score based on a machine learning model for predicting mortality in pneumonia patients on admission to the intensive care unit, Respiratory Medicine, Volume 17, https://doi.org/10.1016/j.rmed.2023.107363.
[44] Ru Wen, Peng Xu, Yimin Cai, Fang Wang, Mengfei Li, Xianchun Zeng & Chen Liu (2023), A Deep Learning Model for the Diagnosis and Discrimination of Gram-Positive and Gram-Negative Bacterial Pneumonia

for Children Using Chest Radiography Images and Clinical Information, Infection and Drug Resistance, 16:, 4083-4092, DOI: 10.2147/IDR.S404786

[45] M. U. Alam, J. R. Baldvinsson, and Y. Wang, "Exploring LRP and Grad-CAM visualization to interpret multi-label-multi-class pathology prediction using chest radiography," in 2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS), Shenzen, China: IEEE, Jul. 2022, pp. 258–263. https://doi: 10.1109/CBMS55023.2022.00052.

[46] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI: IEEE, Jul. 2017, pp. 1800–1807. doi: 10.1109/CVPR.2017.195.

[47] Elsa D Angelini, Jie Yang, Pallavi P Balte, Eric A Hoffman, Ani W Manichaikul, Yifei Sun, Wei Shen, John H M Austin, Norrina B Allen, "Pulmonary emphysema subtypes defined by unsupervised machine learning on CT scans," thorax, p. thoraxjnl-2022-219158, Jun. 2023, doi: 10.1136/thorax-2022-219158.

[48] Ø. Johannessen, F. Uthaug Reite, R. Bhatnagar, T. Øvrebotten, G. Einvik, and P. L. Myhre, "Lung Ultrasound to Assess Pulmonary Congestion in Patients with Acute Exacerbation of COPD," COPD, vol. Volume 18, pp. 693–703, Apr. 2023, doi: 10.2147/COPD.S396855.

[49] A. Zern, K. Broelemann, and G. Kasneci, "Interventional SHAP Values and Interaction Values for Piecewise Linear Regression Trees," AAAI, vol. 37, no. 9, pp. 11164–11173, Jun. 2023, doi: 10.1609/aaai.v37i9.26322.

[50] V. I. Agughasi, "The Superiority of Fine-tuning over Full-training for the Efficient Diagnosis of COPD from CXR Images," Inteligencia Artificial, vol. 27, no. 74, pp. 62–79, 2024.

[51] A. V. Ikechukwu, S. Bhimshetty, D. R, and M. V. Mala, "Advances in Thermal Imaging: A Convolutional Neural Network Approach for Improved Breast Cancer Diagnosis," in 2024 International Conference on Distributed Computing and Optimization Techniques (ICDCOT), Mar. 2024, pp. 1–7. doi: 10.1109/ICDCOT61034.2024.10515323.