



## A Dual-Stage Deep Learning Framework for Pneumothorax Detection and Segmentation in Chest X-rays

Ahmed Kawther Hussein<sup>ID</sup>

Department of Computer Science, College of Education, Mustansiriyah University, Baghdad 00964, Iraq

Corresponding Author Email: [ahmedkawther@uomustansiriyah.edu.iq](mailto:ahmedkawther@uomustansiriyah.edu.iq)

Copyright: ©2025 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.301119>

### ABSTRACT

**Received:** 7 October 2025

**Revised:** 8 November 2025

**Accepted:** 17 November 2025

**Available online:** 30 November 2025

**Keywords:**

*Xception, U-Net, XGBoost, pneumothorax detection, deep learning, chest X-ray, image segmentation, InceptionResNetV2*

This study aims to investigate the impact of automated deep learning-based detection and segmentation approaches on pneumothorax diagnosis in chest radiographs, addressing a critical gap in computer-aided diagnosis systems. The article seeks new insights into combining segmentation and classification methodologies, contributing to a deeper understanding of automated medical image analysis for emergency radiological conditions. The study employs a quantitative approach, utilizing a dual-stage deep learning framework that combines U-Net architectures for segmentation with XGBoost classification. Data were collected from chest X-ray images with annotated pneumothorax regions and analyzed using PyTorch and scikit-learn frameworks, implementing various backbone architectures including ResNet34, VGG16, InceptionResNetV2, and Xception. The study present that the Xception-based U-Net achieves superior segmentation performance with an IoU of 0.6085 and accuracy of 0.9920. This research presents an innovative approach to pneumothorax diagnosis by integrating semantic segmentation with classification, yielding significant discoveries that enhance the existing knowledge of automated medical image analysis. emphasizing future research opportunities in multi-task medical image processing and prospective clinical applications.

### 1. INTRODUCTION

Effective treatment of pneumothorax, a potentially deadly condition caused by air accumulation in the pleural cavity between the lung and chest wall, necessitates fast and accurate diagnosis. Chest radiographs continue to be the principal diagnostic technique for identifying pneumothorax, but being widely available, they can be difficult to interpret, even for skilled radiologists [1, 2]. In the study by Pedrosa et al. [3], the importance of prompt diagnosis and the mounting strain on healthcare systems around the world have prompted research into automated diagnostic solutions, especially those that make use of deep learning and artificial intelligence.

In the last few years, automatic pneumothorax detection has progressed dramatically, ranging from basic computer vision techniques to complex deep learning architectures. Urooj et al. [4] presented the potential of Deep Convolutional Neural Networks (DCNN) enhanced with Network in Network (NIN) preprocessing, achieving AUC of 0.9844. This foundation was more upon by incorporating transfer learning approaches, as evidenced by developments in segmentation techniques. Bondarenko and Syryh [5] represented an improved UNet++ architecture that achieved a Dice similarity coefficient of 88.31% and IOU of 83.1%, marking a significant improvement in the precision of pneumothorax localization.

In order to get better local and global information in chest radiographs, Sanati et al. [1] supposed a combination for

convolutional and transformer network. By integrating U-Net segmentation with DenseNet structures, Manikandan et al. [2] proposed this hybrid technique with increasing the accuracy to 98.5 percent. frameworks called Lesion-aware categorization were recently introduced by Pedrosa et al. [3] highlighting the significance of interpretability in clinical situations.

While Urooj et al. [4] demonstrated the effectiveness of computer-assisted systems by acquiring 97.77 percent accuracy by DenseNet-169. Bondarenko and Syryh [5] enhance the field by providing approaches for predicting illness development. Kancherla et al. [6] improved segmentation accuracy with residual connections; Deng et al. [7] provided a way to accurately identify pneumothorax illnesses by testing different models and hyperparameter configurations.

Although these significant improvements, the current pneumothorax detection method still has substantial shortcomings. While individual models have demonstrated impressive performance in either segmentation or classification tasks, frameworks that effectively integrate both abilities are clearly lacking. Current methods often fail to utilize the supplementary data offered by both raw images and segmentation masks in their final diagnostic decisions.

Our suggested dual-stage strategy to solve the shortcomings of current methods, is motivated by these gaps in the present research environment.

## 2. LITERATURE REVIEW

Initial applications of deep learning emerged in 2020 when Groza and Kuzin [8] demonstrated the potential of basic U-Net architectures, achieving a promising Dice score of 0.8821. This foundational work was quickly advanced by Wang et al. [9], who enhanced the approach by introducing a Deep Convolutional Neural Network (DCNN) coupled with Network in Network (NIN) preprocessing, pushing the performance boundaries to an AUC of 0.9844.

As the field matured into 2021, researchers began exploring transfer-learning approaches to leverage pre-trained models' knowledge. Pedrosa et al. [3] made significant strides by combining U-Net with various sophisticated backbones, including ResNet34, SE-ResNext50, SE-ResNext101, and DenseNet121, achieving a notable Dice coefficient of 0.8574. This progress was thoroughly documented in a comprehensive review by Iqbal et al. [10], which highlighted the growing success of transfer learning models in achieving classification AUC scores of 88.87%.

A deep learning model has been built for pneumothorax identification by integrating two extensive open-source chest radiograph datasets: ChestX-ray14 and CheXpert [11] to determine the generalizability of a deep learning model for pneumothorax diagnosis between datasets from various external institutions and examine patient outcomes. Nine boxes have been generated for pneumothorax localization by Cho et al. [12] in order to assess the diagnostic performance attained through using fully connected tiny artificial neural networks (ANNs). The sensitivity and specificity were 80.6% and 83.0%, respectively.

Xu et al. [13] build efficient nomograms to predict delayed pneumothorax following microwave ablation in patients with lung cancer. The primary finding is that the nomograms can accurately forecast the probability of both delayed and acute pneumothorax following MWA.

Abedalla et al. [14] employ a variety of training and prediction strategies, including test-time augmentation, data augmentation, and stochastic weight averaging. With a mean Dice similarity coefficient of 0.8608 on the test set, the suggested segmentation network ranks in the upper 1% of systems in the Kaggle competition. Jones et al. [15] found a major concern is the generalizability of machine learning algorithms that have been trained. Differences in imaging infrastructure, patient population characteristics, disease distributions, and excessive fitting could lead models to perform well in one environment and poorly in another.

The year 2022 marked a shift toward hybrid architectures as researchers sought to overcome the limitations of single-model approaches. Zhang et al. [16] introduced the innovative *ResNeSt – UNet ++* architecture, incorporating spatial and channel squeeze excitation modules to achieve an impressive 88.31% Dice similarity. This advancement was followed by Sanati et al. [1], who proposed a novel convolutional-transformer architecture designed to capture both local and global features effectively. Kaur et al. [17] performed EfficientNet B0 with transfer learning, achieving 83% accuracy while maintaining computational efficiency.

To address vanishing gradients in the pneumothorax recognition model, Luo et al. [18] incorporated residual blocks. Hillis et al. [19] evaluated the precision of an AI model against Consensus thoracic radiology interpretations to recognize pneumothorax and tension pneumothorax. The parts

under the receiver were the main points.

A novel system was proposed by Iqbal et al. [20] and called (VDV), which is a model-level ensemble (MLE) of several data-level ensembles (DLE). Sensitivities and specificities for the identification of tension and pneumothorax were the secondary outcomes.

Recent developments in 2023 and early 2024 have focused increasingly on clinical applicability and real-world implementation. Manikandan et al. [2] achieved a remarkable 98.5% accuracy through the integration of DenseNet201 with U-Net segmentation. This was complemented by Pedrosa et al. [3], who introduced a lesion-aware classification approach incorporating object detection frameworks. Further advancing the field, Bondarenko and Syryh [5] proposed methods for predicting disease progression, emphasizing the importance of longitudinal analysis in clinical settings. Kumar et al. [21] presented PneumoNet, which is an ensemble deep learning network that addresses the problem of class disparity by using data augmentation to generate synthetic images and a segmentation algorithm to find dark regions.

Upasana et al. [22] constructed a model that includes an attention module and an xception network to identify pneumothorax in chest X-ray pictures. Using 2,597 chest X-ray images, the recommended model showed training accuracy of 99.18% and validation accuracy of 87.53%. Lin et al. [23] assumed that the framework enhanced subfigure segmentation is sophisticated. The CXR dataset from the NIH method for large chest x-rays. Images of human lungs are individually upgraded by Chutia et al. [24], using Contrast-Limited Adaptive Histogram Equalization (CLAHE) and Discrete Wavelet Transform (DWT) to eliminate noise and boost image quality. The proposed design performs better than the existing model.

Ikechukwu and Murali [25] proposed approach exhibits robust generalization abilities on the dataset for semantic lung segmentation via semi-supervised localization. The author suggests a comprehensible ensemble learning method for lung segmentation and the diagnosis of lung diseases utilizing CXR images [26]. Olayiwola et al. [27] sought to develop a convolutional neural network (CNN)-based model for the classification of lung diseases such as MobileNetV2, ResNet-50, ResNet-101, and AlexNet.

Despite these significant advancements, several critical gaps remain in the current research landscape. Most notably, while individual models have shown promising results in either segmentation or classification tasks, there is a conspicuous lack of research effectively combining both approaches. Because these activities in clinical diagnosis are related, this integration gap is very significant. Additionally, in their final classification conclusions, current methods frequently fall short of making the most of the rich information available from both raw images and segmentation masks.

A further difficulty is optimizing model architectures. Numerous architectures have been tested and proposed, but nothing is known about the best way to combine various loss functions and how they should be weighted. Given the complexity of medical image processing, where various mistake kinds may have diverse clinical ramifications. By integrating U-Net segmentation with a ResNet 34 backbone in a dual-stage architecture with DenseNet-based classification enhanced by XGBoost, the technique eliminates the critical integration gap.

### 3. METHODOLOGY

The proposed approach offers a two-phase strategy for classifying and detecting pneumothorax in chest X-ray images. There are two main parts to the framework:

(1) ***U-Net*** based segmentation network that uses various backbone architectures to figure out the pneumothorax's spatial extent.

(2) **Hybrid Classification System** combines deep learning feature extraction with gradient enhancement classification.

Clinicians can obtain both localization and diagnostic information with this integrated technique and that enables both pixel-wise detection of pneumothorax zones and overall diagnostic categorization. Through preprocessing processes, unique loss functions, and comprehensive data augmentation strategies to handle the fundamental complications of chest X-ray interpretation, the methodology solves challenges of medical image analysis.

#### 3.1 Problem formulation

It is essential to formulate the complicated computer vision problem of automatically detecting and classifying pneumothorax in chest radiographs as an integrated optimization problem, which includes binary classification and semantic segmentation. This dual-objective challenge necessitates both pixel-by-pixel detection of pneumothorax and a global assessment of its occurrence.

Let  $\mathcal{D} = \{(I_i, M_i, y_i)\}_{i=1}^N$  represent dataset of  $N$  image samples, where:

- $I_i \in \mathbb{R}^{H \times W}$  denotes the input chest X-ray image
- $M_i \in \{0, 1\}^{H \times W}$  represents the binary segmentation mask
- $y_i \in \{0, 1\}$  indicates the presence of pneumothorax

Our objective is to learn two mapping functions:

- Segmentation Function:  $f_\theta: \mathbb{R}^{H \times W} \rightarrow [0, 1]^{H \times W}$  where  $\theta$  represents the learnable parameters of the segmentation network.
- Classification Function:  $g_\phi: \mathbb{R}^{H \times W} \times [0, 1]^{H \times W} \rightarrow [0, 1]$  where  $\phi$  represents the learnable parameters of the classification network.

The optimization problem can be formulated as:

$$\min_{\theta, \phi} \mathcal{L}_{seg}(f_\theta(I), M) + \lambda \mathcal{L}_{cls}(g_\phi(I, f_\theta(I)), y) \quad (1)$$

where,  $\lambda$  is a weighting parameter balancing the two objectives.

This formulation encapsulates the interdependence between segmentation and classification tasks, where the segmentation output  $f_\theta(I)$  serves as an additional input to the classification function  $g_\phi$ , allowing the classifier to leverage both global image features and localized pneumothorax information. The joint optimization ensures that the segmentation network learns to identify relevant regions that contribute to accurate classification while maintaining pixel-wise accuracy in delineating pneumothorax boundaries.

#### 3.2 Theoretical framework

The proposed solution is grounded in a comprehensive theoretical framework that integrates concepts from deep learning, computer vision, and statistical learning theory. Our approach builds upon the Universal Approximation Theorem

for neural networks, which guarantees that sufficiently wide neural networks can approximate any continuous function on compact subsets of  $\mathbb{R}^n$ . This theoretical foundation is particularly relevant for medical image analysis, where complex patterns must be learned from high-dimensional data while maintaining robustness and generalizability.

Let  $\mathcal{X} = \mathbb{R}^{H \times W}$  be our input space and  $\mathcal{Y} = \{0, 1\}^{H \times W} \times \{0, 1\}$  be our output space. The learning problem can be viewed as finding functions in the hypothesis space  $\mathcal{H}$  that minimize the expected risk:

$$\mathcal{R}(h) = \mathbb{E}_{(x, y) \sim \mathcal{D}}[\mathcal{L}(h(x), y)] \quad (2)$$

where,  $h \in \mathcal{H}$  and  $\mathcal{L}$  is our composite loss function. Given the finite nature of our training set, we minimize the empirical risk while maintaining generalization through appropriate regularization techniques.

##### 3.2.1 Data image preprocessing

The preprocessing stage is crucial for ensuring optimal model performance and convergence. The experiments were conducted using the SIIM-ACR Pneumothorax Segmentation Dataset, containing 12,047 annotated chest X-ray images [26]. We establish a formal preprocessing framework that addresses the specific challenges of medical image analysis, including intensity normalization, spatial standardization, and data augmentation for improved generalization.

Define the preprocessing pipeline  $\mathcal{P}$  as a composition of transformations:

The preprocessing methodology is designed to satisfy several key theoretical properties:

##### Invariance Properties:

- Translation invariance:  $\mathcal{P}(T_\Delta(I)) \approx \mathcal{P}(I)$  for small translations  $\Delta$
- Scale invariance:  $\mathcal{P}(S_\alpha(I)) \approx \mathcal{P}(I)$  for scale factors  $\alpha$  near 1
- Rotation invariance:  $\mathcal{P}(R_\theta(I)) \approx \mathcal{P}(I)$  for small angles  $\theta$

##### Statistical Properties:

- Normalized intensity distribution:  $\mathbb{E}[\mathcal{P}(I)] = 0$ ,  $\text{Var}[\mathcal{P}(I)] = 1$
- Preserved spatial correlations for anatomical structures
- Controlled noise characteristics

##### Information Preservation:

- Minimal loss of diagnostically relevant information
- Preservation of edge and texture information critical for pneumothorax detection
- Maintenance of spatial relationships between anatomical structures

$$\mathcal{P} = T_n \circ T_{aug} \circ T_r \quad (3)$$

where:

- $T_r$  is the resizing transformation.
- $T_{aug}$  represents the augmentation function defined as.
- $T_n$  is the normalization function.

#### 3.3 Segmentation architecture

The segmentation component of our framework is built upon deep convolutional neural networks, specifically leveraging the U-Net architecture with various backbone

networks. The architectural design is motivated by the need to capture both fine-grained local features crucial for boundary detection and global contextual information necessary for understanding anatomical relationships in chest X-rays. Our framework extends the traditional U-Net architecture by incorporating modern deep learning advances and domain-specific optimizations for medical image analysis.

Let  $\mathcal{F} = \{f_\theta: \mathcal{X} \rightarrow \mathcal{Y} | \theta \in \Theta\}$  represent the class of learnable segmentation functions, where  $\Theta$  denotes the parameter space. The segmentation network implements a mapping:

$$f_\theta: \mathbb{R}^{H \times W} \rightarrow [0, 1]^{H \times W} \quad (4)$$

Through a sequence of nested function compositions that form the encoder-decoder architecture with skip connections:

$$f_\theta = g_d \circ h_{skip} \circ g_e \quad (5)$$

where,  $g_e$  is the encoder pathway,  $g_d$  is the decoder pathway, and  $h_{skip}$  represents the skip connection mechanism.

### 3.3.1 U-Net framework

The U-Net architecture implements a multi-scale feature extraction and synthesis framework that can be formalized as follows:

- **Encoder Pathway:** Let  $E_k$  represent the encoding operation at level  $k$ :

$$E_k: \mathbb{R}^{C_k \times H_k \times W_k} \rightarrow \mathbb{R}^{C_{k+1} \times H_{k+1} \times W_{k+1}} \quad (6)$$

where the feature transformation at each level is defined as:

$$E_k(X) = Pool \left( BN \left( \sigma \left( Conv_{k,2} \left( BN \left( \sigma \left( Conv_{k,1}(X) \right) \right) \right) \right) \right) \right) \quad (7)$$

with:

- $Conv_{k,i}$  representing the  $i$ -th convolution at level  $k$
- $BN$  denotes batch normalization
- $\sigma$  being the ReLU activation function
- $Pool$  implements max pooling with stride 2
- **Skip Connections:** The skip connection mechanism  $S_k$  at level  $k$  is defined as:

$$S_k: (X_{d,k}, X_{e,k}) \mapsto \text{Concat} (X_{d,k}, \text{Crop}(X_{e,k})) \quad (8)$$

where,  $X_{d,k}$  and  $X_{e,k}$  are decoder and encoder features respectively.

- **Decoder Pathway:** The decoder operation  $D_k$  at level  $k$  can be expressed as:

$$D_k: \mathbb{R}^{C_{k+1} \times H_{k+1} \times W_{k+1}} \times \mathbb{R}^{C_k \times H_k \times W_k} \rightarrow \mathbb{R}^{C_k \times H_k \times W_k} \quad (9)$$

- **Multi-Scale Feature Integration:** The final prediction integrates features across all scales:

$$\hat{Y} = \sigma \left( Conv_{1 \times 1} \left( \sum_{k=1}^K w_k F_k \right) \right) \quad (10)$$

where:

- $F_k$  represents features at scale  $k$
- $w_k$  are learnable scale weights
- $Conv_{1 \times 1}$  is the final projection to output space

The U-Net architecture implements an encoder-decoder framework with skip connections. For an input tensor  $X \in \mathbb{R}^{C \times H \times W}$ , the encoder pathway  $E$  consists of  $k$  blocks:

$$E_k(X) = Pool \left( \sigma \left( Conv_k \left( \sigma \left( Conv_k(X) \right) \right) \right) \right) \quad (11)$$

where,  $Conv_k$  represents convolutional layers at level  $k$ .

### 3.3.2 Loss function formulation

The optimization of our segmentation network is guided by a carefully designed composite loss function that addresses multiple aspects of the segmentation task. Given the inherent challenges in medical image segmentation, including class imbalance, varying region sizes, and the critical importance of boundary accuracy, we propose a weighted combination of complementary loss terms. Each component of our loss function is designed to address specific challenges in pneumothorax segmentation while maintaining stable training dynamics.

The combined loss function  $\mathcal{L}_{total}$  is defined as:

$$\mathcal{L}_{total} = \alpha \mathcal{L}_{BCE} + \beta \mathcal{L}_{Dice} + \gamma \mathcal{L}_{Focal} \quad (12)$$

where:

$$\mathcal{L}_{BCE} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (13)$$

This term provides pixel-wise supervision and encourages probabilistic prediction calibration.

$$\mathcal{L}_{Dice} = 1 - \frac{2|X \cap Y| + \epsilon}{|X| + |Y| + \epsilon} \quad (14)$$

where,  $\epsilon = 1 \times 10^{-5}$  is a smoothing factor. The Dice loss contributes:

- Region-based optimization: Focus on spatial overlap
- Scale invariance: Performance independent of region size
- Gradient characteristics:  $\frac{\partial \mathcal{L}_{Dice}}{\partial X} \propto \frac{Y}{(|X| + |Y|)^2}$
- **Focal Loss:**

$$\mathcal{L}_{Focal} = -\alpha_t (1 - p_t)^\gamma \log(p_t) \quad (15)$$

where,  $p_t$  is the predicted probability and  $\gamma = 2$  is the focusing parameter. This component provides:

- Dynamic weighting:  $w(p_t) = (1 - p_t)^\gamma$
- Class balancing: Through  $\alpha_t$  parameter
- Gradient modulation:  $\frac{\partial \mathcal{L}_{Focal}}{\partial p_t} \propto (1 - p_t)^{\gamma-1}$

The weights  $\alpha$ ,  $\beta$ , and  $\gamma$  are determined through empirical validation to optimize the trade-off between different loss components.

## 3.4 Classification framework

### 3.4.1 Feature extraction

The feature extraction stage implements a dual-stream architecture that processes both the original chest X-ray images and their corresponding segmentation masks. This approach enables the model to leverage both global image characteristics and localized pneumothorax indicators.

The feature extraction process can be formalized as:

$$\Phi(I, M) = [\phi_{CNN}(I); \phi_{CNN}(M)] \quad (16)$$

where,  $\phi_{CNN}$  represents the DenseNet feature extractor and  $[;]$  denotes concatenation.

The DenseNet feature extractor implements dense connectivity patterns defined by:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (17)$$

where,  $H_l$  represents the composite function of operations:

### 3.4.2 XGBoost classification

The XGBoost classifier extends traditional gradient boosting by incorporating second-order gradients and regularization. The model employs additive training to combine weak learners.

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in \mathcal{F} \quad (18)$$

where,  $\mathcal{F}$  is the space of regression trees.

## 3.5 Training dynamics

Adaptive optimization strategy implements by the training process a with dynamic learning rate adjustment and early holding techniques to avoid overfitting and guarantee strong convergence.

### 3.5.1 Learning rate schedule

Implements a plateau-based reduction strategy with momentum correction. The learning rate  $\eta_t$  at epoch  $t$  follows:

$$\eta_t = \eta_0 \cdot \prod_{i=1}^k \alpha_i \quad (19)$$

$\alpha_i = 0.5$  if plateau is detected,  $k$  is the number of learning rate reductions.

### 3.5.2 Early stopping criterion

Mechanism implements a patience-based approach with validation loss monitoring. Training terminates when:

$$\mathcal{L}_{val}^t > \min_{i \in [t-p, t-1]} \mathcal{L}_{val}^i \quad (20)$$

where,  $\mathcal{L}_{val}^t$  is the validation loss at epoch  $t$ .

## 3.6 Performance metrics

To evaluate the dual-stage pneumothorax detection approach, a comprehensive metric system that evaluates the pixel-wise segmentation accuracy and classification reliability is needed. We employ a comprehensive evaluation strategy

that considers both spatial and category correctness, as well as the therapeutic significance of different sorts of errors.

### 3.6.1 Segmentation metrics

For evaluating segmentation performance, we employ a suite of complementary metrics that capture different aspects of spatial accuracy. Let  $M_{pred}$  and  $M_{true}$  represent the predicted and ground truth segmentation masks respectively.

The Intersection over Union (IoU), also known as the Jaccard Index, is calculated as:

$$IoU = \frac{|M_{pred} \cap M_{true}|}{|M_{pred} \cup M_{true}|} \quad (21)$$

### 3.6.2 Classification metrics

The classification performance is evaluated through a comprehensive set of metrics derived from the confusion matrix. Let TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives respectively.

For binary classification metrics:

$$\begin{aligned} Precision &= \frac{TP}{TP + FP} \\ Recall &= \frac{TP}{TP + FN} \\ F1 &= 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \end{aligned} \quad (22)$$

additional classification metrics include:

**Accuracy:**

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (23)$$

where, TPR is the true positive rate and FPR is the false positive rate.

## 4. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we presented the experimental results and analysis of our proposed dual-stage framework for pneumothorax detection and segmentation in chest X-rays. The evaluation encompasses two main phases: (1) the segmentation performance using various backbone architectures in the U-Net framework, and (2) the classification performance using XGBoost with feature extraction. We first describe our experimental design and setup, followed by detailed analysis of the results from both phases.

Our experiments systematically evaluate the effectiveness of different architectural choices and their impact on overall system performance. The results demonstrate the superiority of the *Xception*-based architecture for segmentation tasks, while also highlighting the benefits of hyperparameter optimization in the classification phase. Through comprehensive analysis of multiple performance metrics, we provide the strengths and limitations of each approach, as well as potential areas for future improvement.

### 4.1 Experimental design

The experimental framework implemented using *PyTorch* framework and executed on *CUDA - enabled GPUs*. To ensure reproducibility, we maintained float 32 precision and

enabled deterministic operations throughout all experiments.

The training configuration utilized input dimensions of  $512 \times 512$  pixels with a batch size of 32. The initial learning rate was set to  $1 \times 10^{-3}$ , and training proceeded for 25 epochs with an early stopping mechanism implemented with a patience of 5 epochs. A learning rate scheduler was employed with a reduction factor of 0.5 and a patience of 3 epochs.

Our experimental results revealed several significant findings. The *Xception*-based *U-Net* demonstrated optimal performance for segmentation, effectively balancing both accuracy and boundary precision. The hyperparameter tuning of the *XGBoost* classifier successfully improved the detection of pneumothorax cases while maintaining high

performance for normal cases.

## 4.2 Analysis of results

Our work presented a comprehensive analysis of our dual-stage framework's performance in both segmentation and classification tasks. We evaluate multiple backbone architectures for the segmentation phase followed by an analysis of the classification performance with *XGBoost*.

We evaluated four different backbone architectures for the *U-Net* segmentation model: ResNet34, VGG16, InceptionResNetV2, and *Xception*. Table 1 presents the comparative performance metrics for these architectures.

**Table 1.** Comparison of segmentation performance across different architectures

Architecture	Loss	IoU	Accuracy
ResNet34	0.2761	0.4733	0.9885
VGG16	0.2811	0.5504	0.9928
InceptionResNetV2	0.2718	0.2238	0.9943
<i>Xception</i>	0.2674	0.6085	0.9920

**Table 2.** Classification performance before and after hyperparameter tuning

Stage	Class	Precision	Recall	F1 Score
Before Tuning	Normal (0)	0.90	0.97	0.93
	Pneumothorax (1)	0.85	0.64	0.73
After Tuning	Normal (0)	0.93	0.95	0.94
	Pneumothorax (1)	0.81	0.76	0.78

The *Xception*-based model demonstrated superior performance with the lowest loss value (0.2674) and highest IoU score (0.6085), while maintaining high accuracy (0.9920). While the InceptionResNetV2 achieved marginally higher accuracy (0.9943), its significantly lower IoU score (0.2238) indicates potential issues with precise boundary detection.

Following the segmentation phase, we implemented an *XGBoost* classifier with features extracted from both original images and predicted masks. The classification performance evaluates before and after hyper parameter tuning using random search optimization in Table 2.

The hyperparameter tuning process resulted in improved performance, particularly for pneumothorax detection (Class 1), where recall increased from 0.64 to 0.76. This improvement in recall was achieved while maintaining strong precision for normal cases (Class 0).

## 5. DISCUSSION

This study presents a comprehensive dual-stage framework for pneumothorax detection and segmentation in chest radiographs, introducing several novel contributions to the field of automated medical image analysis. The experimental results offer valuable insights into both the technical aspects of deep learning approaches for medical image analysis and their potential clinical implications.

An interesting result was discovered from the comparison of backbone architectures: InceptionResNetV2 had the lowest Intersection over Union score (0.2238) despite having the highest overall accuracy (0.9943), indicating the drawbacks of using accuracy as the only evaluation metric in extremely unbalanced segmentation tasks. The difference emphasizes how important it is to employ boundary-sensitive measurements, such as Intersection over Union (IoU), for

medical segmentation tasks, especially when it comes to situations like pneumothorax.

All models, including the top-performing *Xception* based architecture, occasionally had difficulty recognizing small or small *pneumothoraces*, especially those that presented atypically or in cases with poor quality images, according to the qualitative analysis of segmentation results. This restriction highlights the inherent difficulties in interpreting chest *X-ray* and points to possible areas for development by incorporating attention mechanisms made especially to improve focus on areas with subtle intensity variations.

A significant difference between sensitivity and specificity was identified in the stage of classification by comparing the performance metrics of the *XGBoost* model before and after hyperparameter changes. By improving the ability to detect positive instances, the optimization procedure raised the recall for pneumothorax cases from 0.64 to 0.76. This is especially useful in emergency situations where failing to detect a pneumothorax could have serious clinical repercussions. However, this improvement came with a slight decrease in precision (from 0.85 to 0.81), reflecting the inherent challenge in balancing false positives and false negatives in clinical diagnostic systems.

The dual-stage approach combining segmentation with classification demonstrates several advantages over single-stage systems. Using segmentation to first localize pneumothorax areas, and then classification using both segmentation masks and original image attributes, the framework efficiently integrates spatial and contextual information. This integrated in our approach reduced false negatives if we compared to image-only classification methods reported in prior studies such as Kaur et al. [17], who stated that their respective recall ratings for pneumothorax detection were 0.71 and 0.73.

## 6. SUMMARY AND CONCLUSION

Through comprehensive experimentation and evaluation, we determined that the *Xception*-based U-Net achieved superior segmentation performance with an *IoU* of 0.6085 and accuracy of 0.9920. The subsequent classification phase demonstrated strong discriminative capability, with precision and recall values of 0.93 and 0.95 respectively for normal cases, and 0.81 and 0.76 for pneumothorax cases after hyperparameter optimization.

The superior performance of our dual-stage approach compared to previous methods reported in the literature validates the hypothesis that combining spatial information from segmentation with contextual features for classification enhances diagnostic accuracy. This integrated approach represents a significant advancement in automated pneumothorax detection, offering potential benefits for emergency radiological practice through improved accuracy, interpretability, and workflow efficiency.

In conclusion, this article demonstrates that advanced deep learning architectures, when effectively combined in a multi-task framework, can achieve high-performance pneumothorax detection and segmentation in chest radiographs. The findings contribute to the growing body of evidence supporting the clinical applicability of artificial intelligence in medical imaging while highlighting important considerations for future research and implementation.

## ACKNOWLEDGMENT

The author would like to thank Mustansiriyah University ([www.uomustansiriyah.edu.iq](http://www.uomustansiriyah.edu.iq)) Baghdad-Iraq, for its support in the present work.

## REFERENCES

- [1] Sanati, A., Dashtestani, M.A., Rostami, H., Azad, S.T. (2023). A novel convolutional-transformer neural network architecture for diagnosis of pneumothorax. In 2023 28th International Computer Conference, Computer Society of Iran (CSICC), Tehran, Iran, Islamic Republic of, pp. 1-5. <https://doi.org/10.1109/CSICC58665.2023.10105407>
- [2] Manikandan, J., Shyni, S.A., Dhanalakshmi, R., Akshaya, S.V., Dharshini, S. (2023). Segmentation and detection of pneumothorax using deep learning. In 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, pp. 468-473. <https://doi.org/10.1109/ICICCS56967.2023.10142364>
- [3] Pedrosa, J., Sousa, P., Silva, J., Mendonça, A., Campilho, A. (2023). Lesion-aware chest radiography abnormality classification with object detection framework. In 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS), Aquila, Italy, pp. 806-813. <https://doi.org/10.1109/CBMS58004.2023.00324>
- [4] Urooj, F., Akbar, S., Hassan, S.A., Gull, S. (2023). Computer-aided system for pneumothorax detection through chest X-ray images using convolutional neural network. In 2023 International Conference on IT and Industrial Technologies (ICIT), Chiniot, Pakistan, pp. 1-6. <https://doi.org/10.1109/ICIT59216.2023.10335810>
- [5] Bondarenko, G.O., Syryh, A.S. (2024). Predicting pneumothorax progression: A methodology using lung mask comparison. In 2024 XXVII International Conference on Soft Computing and Measurements (SCM), Russian Federation, pp. 440-443. <https://doi.org/10.1109/SCM62608.2024.10554174>
- [6] Kancherla, D.S.V., Mannava, P., Tallapureddy, S., Chintala, V., P, K., Iwendi, C. (2023). Pneumothorax: Lung segmentation and disease classification using deep neural networks. In 2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), Erode, India, pp. 181-187. <https://doi.org/10.1109/ICSSAS57918.2023.10331853>
- [7] Deng, L.Y., Lim, X.Y., Luo, T.Y., Lee, M.H., Lin, T.C. (2023). Application of deep learning techniques for detection of pneumothorax in chest radiographs. Sensors, 23(17): 7369. <https://doi.org/10.3390/s23177369>
- [8] Groza, V., Kuzin, A. (2020). Pneumothorax segmentation with effective conditioned post-processing in chest X-ray. In 2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops), Iowa City, IA, USA, pp. 1-4. <https://doi.org/10.1109/ISBIWorkshops50223.2020.9153444>
- [9] Wang, Y.Q., Sun, L.L., Jin, Q. (2021). Enhanced diagnosis of pneumothorax with an improved real-time augmentation for imbalanced chest X-rays data based on DCNN. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 18(3): 951-962. <https://doi.org/10.1109/TCBB.2019.2911947>
- [10] Iqbal, T., Shaukat, A., Akram, M.U., Mustansar, Z., Khan, A. (2021). Automatic diagnosis of pneumothorax from chest radiographs: A systematic literature review. IEEE Access, 9: 145817-145839. <https://doi.org/10.1109/ACCESS.2021.3122998>
- [11] Thian, Y.L., Ng, D., Hallinan, J.T.P.D., Jagmohan, P., et al. (2021). Deep learning systems for pneumothorax detection on chest radiographs: A multicenter external validation study. Radiology: Artificial Intelligence, 3(4): e200190. <https://doi.org/10.1148/ryai.2021200190>
- [12] Cho, Y., Kim, J.S., Lim, T.H., Lee, I., Choi, J. (2021). Detection of the location of pneumothorax in chest X-rays using small artificial neural networks and a simple training process. Scientific Reports, 11: 13054. <https://doi.org/10.1038/s41598-021-92523-2>
- [13] Xu, S., Qi, J., Li, B., Bie, Z.X., Li, Y.M., Li, X.G. (2021). Risk prediction of pneumothorax in lung malignancy patients treated with percutaneous microwave ablation: Development of nomogram model. International Journal of Hyperthermia, 38(1): 488-497. <https://doi.org/10.1080/02656736.2021.1902000>
- [14] Abedalla, A., Abdullah, M., Al-Ayyoub, M., Benkhelifa, E. (2021). Chest X-ray pneumothorax segmentation using U-Net with EfficientNet and ResNet architectures. PeerJ Computer Science, 7: e607. <https://doi.org/10.7717/peerj-cs.607>
- [15] Jones, C.M., Buchlak, Q.D., Oakden-Rayner, L., Milne, M., Seah, J., Esmaili, N., Hachey, B. (2021). Chest radiographs and machine learning - Past, present and future. Journal of Medical Imaging and Radiation Oncology, 65(5): 538-544. <https://doi.org/10.1111/1754-9485.13274>
- [16] Zhang, X., Liu, Z.Q., Wang, Q.F., Chen, B. (2022). Pneumothorax segmentation of chest X-rays using

improved UNet++. In 2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Guangzhou, China, pp. 24-28. <https://doi.org/10.1109/MLISE57402.2022.00013>

[17] Kaur, G., Sharma, N., Chauhan, R., Garg, A., Gupta, R. (2023). Early detection of pneumothorax using EfficientNet B0 transfer learning model. In 2023 2nd International Conference on Futuristic Technologies (INCOFT), Belagavi, Karnataka, India, pp. 1-4. <https://doi.org/10.1109/INCOFT60753.2023.10425769>

[18] Luo, J.X., Liu, W.F., Yu, L. (2022). Pneumothorax recognition neural network based on feature fusion of frontal and lateral chest X-ray images. *IEEE Access*, 10: 53175-53187. <https://doi.org/10.1109/ACCESS.2022.3175311>

[19] Hillis, J.M., Bizzo, B.C., Mercaldo, S., Chin, J.K., et al. (2022). Evaluation of an artificial intelligence model for detection of pneumothorax and tension pneumothorax in chest radiographs. *JAMA Network Open*, 5(12): e2247172. <https://doi.org/10.1001/jamanetworkopen.2022.47172>

[20] Iqbal, T., Shaukat, A., Akram, M.U., Muzaffar, A.W., Mustansar, Z., Byun, Y.C. (2022). A hybrid VDV model for automatic diagnosis of pneumothorax using class-imbalanced chest X-rays dataset. *IEEE Access*, 10: 27670-27683. <https://doi.org/10.1109/ACCESS.2022.3157316>

[21] Kumar, V.D., Rajesh, P., Geman, O., Craciun, M.D., Arif, M., Filip, R. (2023). “Quo Vadis diagnosis”: Application of informatics in early detection of pneumothorax. *Diagnostics*, 13(7): 1305. <https://doi.org/10.3390/diagnostics13071305>

[22] Upasana, C., Tewari, A.S., Singh, J.P. (2023). An attention-based pneumothorax classification using modified Xception model. *Procedia Computer Science*, 218: 74-82. <https://doi.org/10.1016/j.procs.2022.12.403>

[23] Lin, M.Q., Hou, B.J., Mishra, S., Yao, T.Y., Huo, Y.K., Yang, Q., Wang, F., Shih, G., Peng, Y.F. (2023). Enhancing thoracic disease detection using chest X-rays from PubMed Central Open Access. *Computers in Biology and Medicine*, 159: 106962. <https://doi.org/10.1016/j.combiomed.2023.106962>

[24] Chutia, U., Tewari, A.S., Singh, J.P. (2024). Collapsed lung disease classification by coupling denoising algorithms and deep learning techniques. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 13: 1. <https://doi.org/10.1007/s13721-023-00435-0>

[25] Ikechukwu, A.V., Murali, S. (2023). CX-Net: An efficient ensemble semantic deep neural network for ROI identification from chest-x-ray images for COPD diagnosis. *Machine Learning: Science and Technology*, 4(2): 025021. <https://doi.org/10.1088/2632-2153/acd2a5>

[26] Kaggle. Chest X-ray images with pneumothorax masks. <https://www.kaggle.com/datasets/vbookshelf/pneumothorax-chest-xray-images-and-masks>.

[27] Olayiwola, J.O., Badejo, J.A., Okokpuije, K., Awomoyi, M.E. (2023). Lung-related diseases classification using deep convolutional neural network. *Mathematical Modelling of Engineering Problems*, 10(4): 1097-1104. <https://doi.org/10.18280/mmep.100401>