Journal of Medicine and Medical Studies

Research Article

A Novel Approach to Chest X-ray Lung Segmentation Using U-net and Modified Convolutional Block Attention Module

Mohammad Ali Labbaf Khaniki[1], Marzieh Mirzaeibonehkhater[2]*, Mohammad Manthour[3] and Katayoon Faraji[4]

[1]Faculty of Electrical Engineering, K.N. Toosi University of Technology, Tehran, Iran

[2]Department of Electrical and Computer Engineering, Indiana University-Purdue University

[3]Faculty of Electrical and Electronic Engineering, Shahed university, Tehran, Iran

[4]Martin J. Whitman School of Management, Syracuse University, Syracuse, NY, USA

A B S T R A C T

Lung segmentation in chest X-ray (CXR) images is a crucial task in medical image analysis, aiding in accurate disease diagnosis and treatment planning. This study presents an improved U-Net architecture by integrating a Convolutional Block Attention Module (CBAM) to enhance segmentation performance. The proposed CBAM unifies three attention mechanisms-Channel Attention, Spatial Attention and Pixel Attention-to refine feature extraction and improve focus on critical image regions. The Channel Attention mechanism emphasizes inter-channel dependencies, the Spatial Attention mechanism enhances localization accuracy by highlighting spatial correlations and the Pixel Attention mechanism refines segmentation precision at the pixel level. By incorporating CBAM into the U-Net framework, the model achieves superior performance in lung segmentation, as evaluated using the Dice coefficient and Jaccard similarity index. Experimental results demonstrate that the proposed approach significantly improves segmentation accuracy, making it a promising advancement in medical image processing for lung disease assessment.

Keywords: Attention Mechanism, Chest X-ray (CXR), Lung Segmentation, Medical Image Processing, U-Net

1. Introduction

Analyzing chest X-rays is complex and time-consuming, often demanding the identification of multiple abnormalities. Radiologists typically perform this task manually, straining healthcare resources. The complexity of the chest anatomy in these images, combined with the subjective nature of interpretation, can result in inconsistent and potentially biased diagnoses[1]. This highlights the need for automated systems to improve efficiency and accuracy in chest X-ray analysis. Furthermore, image quality and data-related issues can hinder accurate interpretation. To assist doctors with this, computer-aided detection (CAD) systems are designed to aid in the analysis of medical images[2]. These systems analyze digital medical visuals, pinpointing characteristic patterns and highlighting potentially problematic regions, like disease indicators, to support diagnostic decisions. CAD systems integrate artificial intelligence, computer vision and medical image processing techniques. In CAD systems, a crucial step is segmentation, which accurately isolates areas of concern, such as tumors, from normal tissue. This precise

separation enhances the reliability of subsequent analyses, like measuring tumor size or tracking disease progression[3].

Essentially, Machine Learning (ML), a type of Artificial Intelligence (AI), allows computers to learn from data independently, without requiring explicit instructions. This capability enables them to automatically improve their performance through experience, minimizing human involvement. This learning process is achieved through algorithms that identify patterns and relationships within datasets. These algorithms can then be applied to new, unseen data to make predictions or decisions. ML algorithms excel at identifying patterns, handling multiple objectives and generating predictions[4,5]. Deep learning, a subset of ML, has gained significant traction across diverse areas, including defect detection[6] and virtual reality applications[7,8]. The increasing need to incorporate advanced AI and ML techniques for image classification and segmentation is fueled by technological progress[9]. For example, research is being conducted on how to understand the inner workings of deep learning applied to error-correcting codes, examining their design, decoding processes and benefits compared to conventional methods. This type of research is crucial for building trust and transparency in AI-driven medical applications. Study[10] demonstrates that deep learning, when combined with methods that improve image contrast, can effectively automate the identification of white matter lesions in MRI scans of multiple sclerosis patients. This automation has the potential to significantly improve the speed and accuracy of diagnosis, leading to better patient outcomes. Research[11] has shown that the bladder's ability to expand easily when first filling and its efficient emptying (over 90%), is due to large folds in its dome, not small mucosal rugae as previously thought.

Deep learning has revolutionized image segmentation by providing numerous methods that greatly improve accuracy and speed. A key example is U-Net, a convolutional neural network designed for biomedical image segmentation. Its U-shaped design, featuring an encoder and decoder, allows it to achieve precise segmentations even with limited training data. This makes U-Net particularly valuable in medical imaging, where obtaining large, annotated datasets is often challenging. Furthermore, its ability to capture both local and global contextual information contributes to its superior segmentation performance[12]. Unet++ enhances medical image segmentation by using nested, interconnected pathways between the encoder and decoder components. These improved connections minimize differences in the information being processed by the encoder and decoder, making it easier for the learning algorithm to optimize the segmentation. This results in more accurate and detailed segmentation outcomes, particularly in complex medical images[13,14]. This method improves upon the standard U-Net by adding residual blocks, which help prevent the vanishing gradient problem and allow for the creation of more complex, deeper networks. This work introduced a 3D U-Net design specifically for identifying and isolating lung tumors in both CT scans and X-ray images[15].

Attention mechanisms, modeled after how humans visually focus, have demonstrated significant effectiveness in various image processing and natural language processing tasks. They allow models to selectively concentrate on the most relevant parts of the input data, improving performance and interpretability.

This method accurately captures the connections between words or events, even when they are far apart in a sequence. This is particularly useful in tasks where long-range dependencies are important, such as summarizing lengthy documents or understanding complex narratives[16]. surveys the role of positional encoding in transformer-based time series models, highlighting various encoding methods, their effectiveness and open challenges in the field. These mechanisms are used in a wide range of applications, including image classification, object detection, semantic segmentation, video analysis, image creation, 3D vision, multi-modal tasks and self-supervised learning. Their versatility stems from their ability to dynamically weigh the importance of different input features, allowing models to adapt to diverse and complex data patterns. Study[17] improves facial recognition precision by intelligently merging features from several models, utilizing attention mechanisms and information bottleneck principles[18]. This work presents a novel one-stage pedestrian detection system that integrates channel and spatial attention mechanisms into CNN architecture[19]. This work suggested a U-Net architecture improved by using multiple encoders for better feature extraction and adding attention mechanisms within the decoders to accurately focus on important features[20]. This method enhances the U-Net architecture by adding multi-scale spatial attention and dilated convolutions, allowing it to efficiently gather contextual information.

This study advances lung segmentation in chest X-rays by fusing U-Net with a combined attention module (CBAM), boosting accuracy through integrated channel, spatial and pixel focus.

- **Improved U-Net:** Integrating CBAM into U-Net allows the model to grasp broader context and concentrate on key areas, resulting in richer feature understanding and superior image segmentation. This enhanced focus translates to more precise and detailed segmentation results.

- **CBAM with triple attention:** Combining channel, spatial and pixel attention substantially improves the model's ability to pinpoint important details in X-ray images. Channel attention highlights key feature channels, spatial attention focuses on crucial locations and pixel attention emphasizes individual pixels, all contributing to more accurate and precise segmentation. This multi-faceted attention approach allows the model to learn complex, hierarchical representations of the image data. By focusing on the most informative elements, the model minimizes the impact of irrelevant information and noise, leading to more robust segmentation results.

Integrating CBAM with U-Net represents a notable advancement in medical imaging, potentially leading to more accurate diagnoses and improved patient care. The effectiveness of this technique is evaluated using metrics like the Dice coefficient and Jaccard similarity, which are crucial for measuring segmentation accuracy by comparing predicted and actual anatomical boundaries, as supported by research[21]. These metrics provide a quantitative measure of how well the model's predictions align with ground truth segmentations, ensuring reliable performance assessment.

This paper is structured as follows: Section II describes the Chest X-ray dataset and preprocessing steps. Section III explains the proposed method, detailing the integration of CBAM into U-Net. Section IV presents the simulation results, including

training and validation details. Finally, Section V concludes the paper with a summary of the key findings and contributions.

2. Description of the Chest X-ray Lung Segmentation Dataset

This section details the dataset and preprocessing steps used to train and evaluate our lung segmentation model. We utilized a publicly available Chest X-ray dataset from Kaggle, supplemented with data augmentation techniques to enhance model robustness and generalization. This comprehensive approach ensures that our model is trained on a diverse range of images, improving its applicability to real-world clinical scenarios.

2.1. Dataset description

To train models for automatic lung identification in X-rays, researchers utilized a dataset from Kaggle, consisting of chest X-ray images and their corresponding lung masks[22]. This dataset is valuable for medical research, particularly in automated tuberculosis screening. It contains X-ray images with segmentation masks, though some masks may be missing, requiring users to verify mask availability for each image. The dataset includes 360 normal and 344 abnormal X-ray images, all labeled by radiologists. **(Figure 1)** displays sample X-ray images and their masks from the training and validation sets.
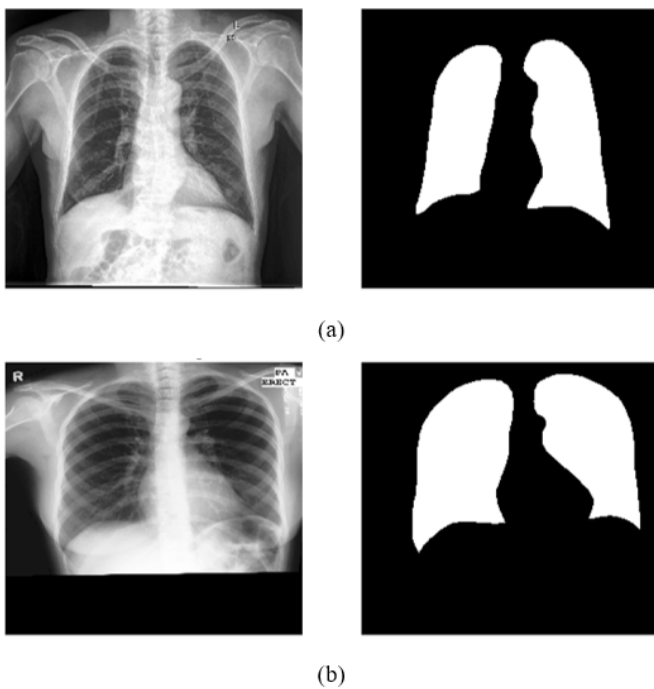

(a)

(b)

Figure 1: Showing chest X-ray images, alongside the lung masks created by expert radiologists, used for both training and validating the model.

This dataset provides a wide spectrum of lung abnormalities, including effusions and miliary patterns, making it a valuable tool for creating algorithms that identify and segment lung diseases in chest X-rays. Its diverse collection of normal and abnormal images offers a robust foundation for analysis. This dataset bridges medical expertise and AI, fostering advancements in automated diagnostics. The careful data collection and preparation make it essential for researchers pushing the boundaries of medical image analysis. Its utilization promotes the development of more accurate and efficient diagnostic tools, ultimately improving patient outcomes.

2.2. Image augmentation and preparation

In order to optimize neural network training for lung segmentation in chest X-rays, the initial dataset was significantly enlarged through a series of data augmentation techniques. Key methods included contrast adjustment, Gaussian blurring, random rotations, horizontal flips and their subsequent combinations. Contrast enhancement improved feature visibility, while blurring mitigated noise and prevented overfitting. Rotations and flips ensured the model's adaptability to varied image orientations, addressing potential biases related to patient positioning and anatomical variations. This comprehensive augmentation strategy, resulting in a sixfold increase in dataset size, effectively simulated diverse imaging conditions, thereby enhancing the network's robustness and accuracy in real-world clinical applications. Furthermore, these augmentations helped the model learn to recognize lung features under challenging conditions, such as varying lung sizes, shapes and textures, which are commonly encountered in clinical practice. The goal was to create a model that could generalize well to unseen data, ensuring reliable performance across a diverse patient population. **(Figure 2)** illustrates the augmented images and their associated masks, showcasing the effects of the applied enhancement and augmentation techniques.
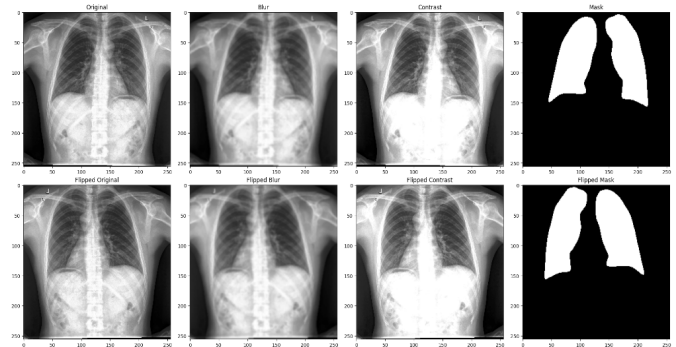

Figure 2: Visual representation of augmented images with their corresponding masks, utilizing the specified augmentation technique.

3. Methodology

This section begins by outlining the U-Net architecture as applied to lung segmentation in X-ray images. We then describe the Convolutional Block Attention Module (CBAM) and introduce our proposed enhanced U-Net model, which incorporates CBAM to improve segmentation accuracy.

3.1. U-Net architecture

U-Net, a convolutional neural network, is specifically designed for biomedical image segmentation. Its "U" shape comes from its symmetrical encoder-decoder structure. The encoder compresses the input image into a detailed feature map, reducing spatial size while increasing feature complexity via convolution and pooling. The decoder then reconstructs this map, using transposed convolutions to increase spatial dimensions for precise localization. Skip connections between encoder and decoder layers transfer contextual information, improving segmentation accuracy. This combination of context and localization makes U-Net highly effective for medical imaging. **(Figure 3)** depicts the U-Net architecture for lung segmentation.
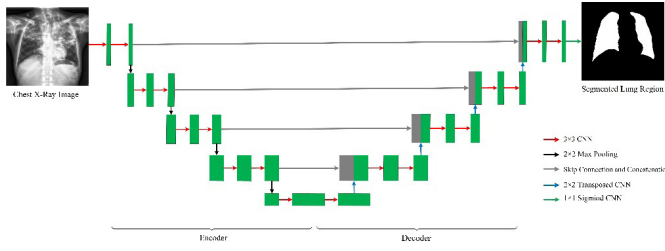
Figure 3: Schematic representation of the U-Net architecture as implemented for lung segmentation in chest X-ray images, showcasing the encoder-decoder structure and skip connections.

3.2. CBAM model

The CBSM is engineered to improve the accuracy of lung segmentation in chest X-rays using the U-Net architecture, especially when training data is scarce, unlike standard CNNs. By integrating channel, spatial and pixel attention, it significantly enhances the model's ability to concentrate on relevant features in X-ray images.

- **Channel attention:** This focuses on relationships between feature channels, allowing the model to prioritize the most informative channels and enhance feature identification. This is achieved by learning to assign different weights to each channel, effectively highlighting the most relevant feature maps.

- **Spatial attention:** This directs the model's focus to critical spatial locations, improving localization accuracy by emphasizing spatial feature correlations. By generating a spatial attention map, the model can selectively attend to specific regions of the input image, ignoring irrelevant background information.

- **Pixel attention:** This enables the model to concentrate on individual pixels, refining focus and boosting segmentation accuracy by prioritizing the most informative pixels. This fine-grained attention allows for precise boundary delineation and detailed feature extraction, particularly important in medical image analysis.

These mechanisms work together to create a richer feature representation, improving image segmentation performance. They enable the model to better capture global context and focus on specific regions. Let's consider a feature map F with dimensions H×W×C, where H is the height, W is the width and C is the number of channels. The CBSM dynamically adjusts weights to pinpoint significant regions in complex scenes. It employs a 1D channel attention map $M\_C \in R^{(1\times1\times C)}$, a 2D spatial attention map $M\_S \in R^{(H\times W\times1)}$ and a 2D pixel attention map $M\_P \in R^{(H\times W\times1)}$. The CBSM refines the input data sequentially using M_C, M_S and M_P. Therefore, the entire process of the enhanced CBSM can be represented as:

The channel attention-refined feature map is:

$$F\_C=(M\_C (F)+1)\times F. \qquad (1)$$

The spatial attention-refined intermediate feature map is:

$$F\_S=(M\_S (F\_C)+1)\times F\_C. \qquad (2)$$

The final feature map, refined by pixel attention, is:

$$F\_P=(M\_P (F\_S)+1)\times F\_S, \qquad (3)$$

Where × indicates element-wise multiplication and + denotes element-wise addition. The attention maps M_C, M_S

and M_P are broadcasted to match the dimensions of the feature maps they refine. The final output, F_P, represents the feature map sequentially refined by channel, spatial and pixel attention, providing a more focused and detailed representation for chest X-ray lung segmentation. This process enables finer control over pixel-level attention, potentially enhancing segmentation accuracy. The CBSM architecture is illustrated in **(Figure 4)**.
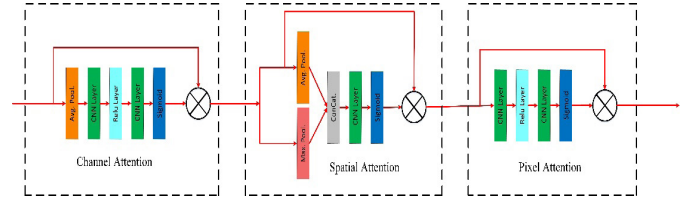

Figure 4: Visualization of the channel, spatial and pixel attention mechanisms in the CBSM model.

Based on Figure 4, the mathematical expressions for the attention mechanisms are as follows:

$$M\_C=\sigma(CNN\_2 (ReLU(CNN\_1 (GP\_avg (x))))), \qquad (1)$$

$$M\_P=\sigma(CNN\_2 (ReLU(CNN\_1 (x)))), \qquad (2)$$

$$M\_S=\sigma(CNN(concat(GP\_max (x),GP\_avg (x)))), \qquad (3)$$

Here's a breakdown of the components:

- **x:** The input feature map to the attention mechanism.

- **GP_avg:** Global average pooling, which reduces spatial dimensions while preserving channel information.

- **GP_max:** Global max pooling, which also reduces spatial dimensions while preserving channel information.

- **CNN_1 and CNN_2:** Convolutional neural network layers used to learn channel-wise dependencies.

- **ReLU:** The Rectified Linear Unit activation function, which introduces non-linearity.

- **σ:** The Sigmoid activation function, which normalizes the output to a range between 0 and 1.

Incorporating the CBSM after each down-sampling and up-sampling stage within the U-Net architecture allows the network to concentrate on the most critical features at each processing level. This is achieved by refining feature maps through the CBSM, which selectively amplifies salient information across channel, spatial and pixel domains. This enhancement is particularly beneficial when dealing with limited training data, as it enables the network to maximize information utilization by highlighting the most informative regions of the input images. **(Figure 5)** visually represents the U-Net architecture augmented with CBSM for lung segmentation in chest X-ray images.
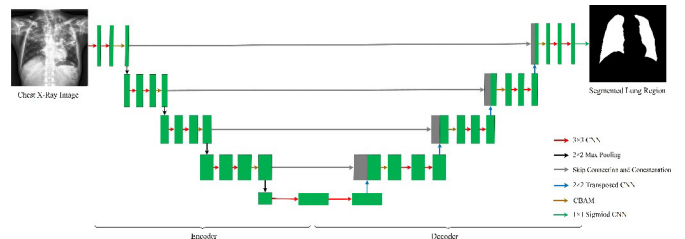

Figure 5: Diagram of the U-Net architecture integrated with the CBSM for lung segmentation in chest X-ray images, showcasing the strategic placement of CBSM after each down-sampling and up-sampling step.

4. Simulations

This section evaluates the proposed method's performance on chest X-ray segmentation. We utilize the Dice similarity coefficient and Jaccard index for assessment, comparing predicted segmentation masks with ground truth data. Additionally, precision, recall and accuracy are used to comprehensively evaluate segmentation performance.

4.1. Effectiveness of the proposed method using dice similarity coefficient and jaccard index

Semantic segmentation or pixel-wise classification, is a crucial technique where each image pixel is assigned to a specific category. This is essential in fields like medical imaging for tissue delineation, remote sensing for land cover classification and autonomous driving for road scene understanding. The goal is to label each pixel, ensuring pixels with the same label share attributes. Model performance is evaluated using the Jaccard index and Dice coefficient, which measure segmentation accuracy. These metrics rely on true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN). TP and TN represent correctly identified tuberculosis and normal images, respectively, while FP and FN represent incorrectly identified normal and tuberculosis images. The Jaccard index or Intersection over Union (IoU), measures the overlap between predicted and actual labels, calculated as the intersection divided by the union:

$$IoU=TP/(TP+FP+FN) \quad (4)$$

The Dice coefficient or Dice similarity coefficient, measures the overlap between two samples. It's calculated as twice the intersection of the predicted and true labels, divided by the sum of their sizes. The formula is:

$$Dice=(2×TP)/(2×TP+FP+FN) \quad (5)$$

These metrics are particularly useful in semantic segmentation because they quantify the overlap between predicted and actual segmentations. The Dice Similarity Coefficient and Jaccard Index results are shown in **(Figures 6 and 7)**.
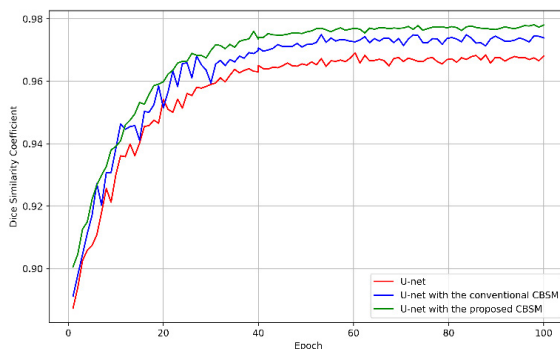

Figure 6: Dice similarity coefficient results comparing U-Net, U-Net with the conventional CBSM[21] and U-Net with the proposed CBSM.

The referenced figures likely illustrate the comparative performance of three U-Net architectures for chest X-ray image segmentation, as measured by the Dice similarity coefficient and Jaccard index. The U-Net without CBSM, acting as the baseline, lacks attention mechanisms and thus processes all features uniformly, resulting in the lowest performance. The U-Net with conventional CBSM integrates channel and spatial attention, allowing it to differentially weight channels and concentrate on pertinent image regions, leading to improved performance over the baseline. The U-Net with proposed CBSM further enhances this by incorporating pixel attention, enabling fine-grained focus on individual pixels, which is crucial for detailed segmentation tasks and results in the highest Dice similarity coefficient.
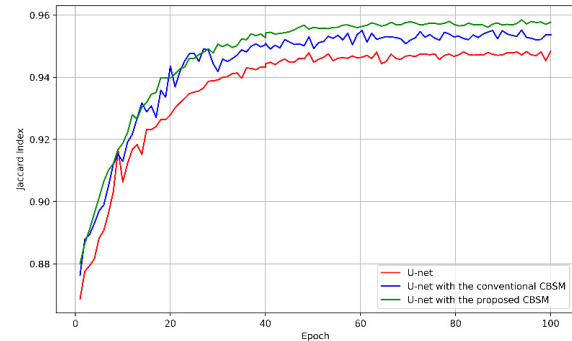

Figure 7: Jaccard Index (IoU) comparison between U-Net, U-Net with conventional CBSM[21] and U-Net with the proposed CBSM.

4.2. Comparison between the ground truth and segmentation masks

Visual comparison between automated segmentation masks and manual annotations is vital for evaluating accuracy, validating quantitative metrics, identifying algorithmic limitations, supporting clinical decisions and improving education and communication in medical imaging. **(Figure 8)** provides a visual comparison of chest X-ray segmentation results from various U-Net architectures.
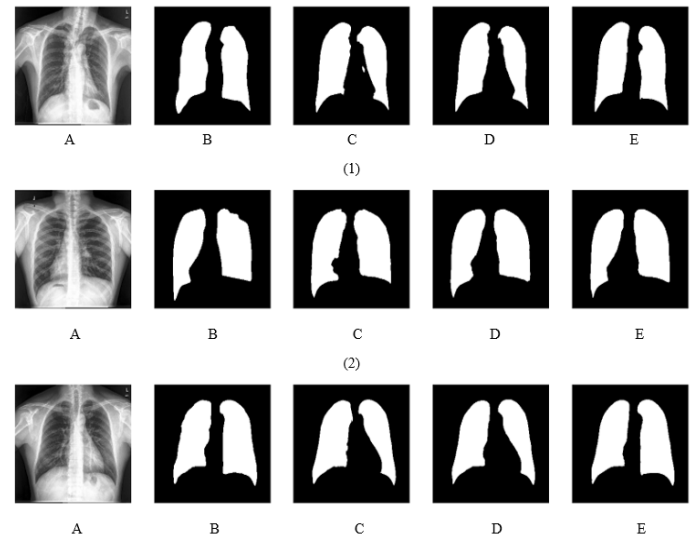

Figure 8: Visual comparison of lung segmentation results across three sample chest X-ray images. (A) Original chest X-ray image. (B) Manually annotated ground truth lung mask. (C) Segmentation result from the standard U-Net architecture. (D) Segmentation result from the U-Net architecture with conventional CBSM. (E) Segmentation result from the U-Net architecture with the proposed CBSM.

The sequential transition from panels C to E in Figure 8 effectively demonstrates the improved segmentation accuracy attained by integrating progressively advanced attention mechanisms into the U-Net model. In particular, the introduced CBSM enhances conventional channel and spatial attention

by incorporating pixel-level refinement, allowing for a more detailed and precise analysis of chest X-ray images. This results in the most optimized segmentation performance.

4.3. Segmentation performance evaluation using various metrics

In this section, we conduct an in-depth evaluation of additional performance metrics to assess the effectiveness of our proposed segmentation approach. These metrics are determined using the following formulas:

- **Accuracy:** The ratio of correctly identified cases (including both true positives and true negatives) to the total number of cases analyzed.

  Accuracy= (TP+TN+FP+FN)/(TP+FN) (6)

- **Recall:** The ratio of correctly detected positive cases to the total number of actual positive cases.

  Recall = TP/(TP+FN) (7)

- **Specificity:** The ratio of correctly detected negative cases to the total number of actual negative cases.

  Specificity= TN/(FP+TN) (8)

- **Precision:** The ratio of correctly identified positive cases to the total number of predicted positive cases.

  Precision = TP/(TP+FP). (9)

- **F1 Score:** The F1-score, which represents the harmonic mean of precision and sensitivity, balances both metrics, particularly in scenarios where one may hold greater significance than the other.

  F1-score =2 (Precision ×Recall)/(Precision +Recall) . (10)

The presented metrics collectively provide a comprehensive evaluation of the deep learning model's effectiveness in pixel-level classification for chest X-ray images. **(Table 1)** outlines these performance indices.

Table 1: The performance metrics for U-Net, U-Net with the conventional CBSM and U-Net with the proposed CBSM.

Method	Accuracy (%)	Recall (%)	Specificity (%)	Precision (%)	F1-score (%)
U-net	96.2	95.30	93.54	96.68	95.98
U-net with the conventional CBSM[21]	97.8	95.57	95.81	97.14	96.34
U-net with the proposed CBSM	98.8	97.50	97.64	97.68	97.58

(Table 1) showcases the incremental performance gains realized through the progressive integration of attention mechanisms within the U-Net architecture. While the baseline U-Net demonstrates commendable results, notably a 96.2% accuracy and 95.98% F1-score, its specificity indicates a potential for improvement in accurately discerning non-relevant data. The introduction of the conventional CBSM, incorporating channel and spatial attention, yields a notable enhancement across all metrics, culminating in a 96.34% F1-score. Further refinement is achieved with the proposed CBSM, which likely leverages pixel-level attention, resulting in a peak accuracy of 98.8%, a 97.58% F1-score and a significantly improved specificity of 97.64%. These results underscore the efficacy of the proposed

CBSM in delivering a more precise and nuanced analysis of chest X-ray images, thereby achieving superior segmentation outcomes.

5. Conclusion

This study presents a novel approach to lung segmentation in chest X-rays by enhancing the U-Net architecture with an innovative CBAM. This module effectively combines three distinct attention mechanisms-channel, spatial and pixel attention-to refine the model's structure, leading to significant improvements in performance. Each attention mechanism contributes to the overall segmentation accuracy: channel attention emphasizes important feature channels, spatial attention focuses on key regions and pixel attention targets the most relevant pixels, resulting in a more accurate and detailed segmentation. The improvements in feature representation and segmentation performance have been thoroughly validated through rigorous assessments using well-established metrics like the Dice coefficient and Jaccard similarity index, demonstrating the method's superiority over traditional models. Additionally, a comparative analysis of pixel classification metrics across different U-Net variations for chest X-ray segmentation shows a clear, consistent improvement as attention mechanisms are progressively integrated. In conclusion, the proposed CBAM combined with the U-Net architecture marks a significant advancement in medical image analysis, providing a more accurate and reliable tool for clinical applications.

Future research could enhance the proposed CBAM-integrated U-Net model by exploring multi-scale attention mechanisms to capture features at different resolutions, improving segmentation for varying lung sizes and pathologies. Expanding the model's training on diverse chest X-ray datasets could increase its generalizability across different populations and imaging conditions. Additionally, incorporating other advanced attention mechanisms, like dynamic or multi-head attention, may further refine performance. Integrating temporal analysis for dynamic imaging could aid in monitoring disease progression over time, while combining multi-modal imaging (e.g., CT and MRI) could provide richer diagnostic information. Optimizing the model for real-time clinical deployment and improving the interpretability of attention mechanisms would further support its practical use in clinical settings, helping radiologists make more informed decisions and ultimately contributing to better patient outcomes.

6. Conflicts of Interest Declaration

The authors affirm that they have no competing financial or non-financial interests that could influence the content of this manuscript.

7. References

1. Castellino RA. Computer aided detection (CAD): an overview. Cancer Imaging, 2005;5(1): 17.

2. Kumar DN, Joseph MK. Fused Feature Vector and Dual FCM for Lung Segmentation from Chest X-Ray Images. Int J Intell Eng Syst, 2024;17(2).

3. Ajami H, Nigjeh MK, Umbaugh SE. Unsupervised white matter lesion identification in multiple sclerosis (MS) using MRI segmentation and pattern classification : a novel approach with CVIPtools. 2023;12674: 1-6.

4. Birashk A, Kordestani JK, Meybodi MR. Cellular teaching-learning-based optimization approach for dynamic multi-objective problems. Knowledge-Based Syst, 2018;141(Feb): 148-177.

5. Samii A, Karami H, Ghazvinian H, et al. Comparison of DEEP-LSTM and MLP Models in Estimation of Evaporation Pan for Arid Regions. J Soft Comput Civ Eng, 2023;7(2): 155-175.

6. Monjezi V, Trivedi A, Tan G, et al. Information-theoretic testing and debugging of fairness defects in deep neural networks. 2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE), IEEE, 2023;19(July): 1571-1582.

7. Salehi M, Javadpour N, Beisner B, et al. Innovative Cybersickness Detection: Exploring Head Movement Patterns in Virtual Reality, 2024;1(Feb): 18.

8. Sanaei M, Gilbert SB, Javadpour N, et al. The Correlations of Scene Complexity, Workload, Presence and Cybersickness in a Task-Based VR Game, 2024;1(March): 58.

9. Devroye N, Mohammadi N, Mulgund A, et al. Interpreting deep-learned error-correcting codes. 2022 IEEE International Symposium on Information Theory (ISIT), IEEE, 2022;10(June): 2457-2462.

10. Nigjeh MK, Ajami H, Umbaugh SE. Automated classification of white matter lesions in multiple sclerosis patients ' MRI images using gray level enhancement and deep learning, 2023;12674(Oct): 6.

11. Azari F, Robertson AM, Tobe Y, et al. Elucidating the high compliance mechanism by which the urinary bladder fills under low pressures, 2025;1(Jan): 10312.

12. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18, Springer, 2015;9351(Nov): 234-241.

13. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018 and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, P Springer, 2018;11045(Sept): 3-11.

14. Agrawal T, Choudhary P. ReSE-Net: Enhanced UNet architecture for lung segmentation in chest radiography images. Comput Intell, 2023;39(3): 456-477.

15. Le VL, Saut O. RRc-UNet 3D for lung tumor segmentation from CT scans of Non-Small Cell Lung Cancer patients. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023;1(Dec): 2308-2317.

16. Irani H, Metsis V. Positional Encoding in Transformer-Based Time Series Models: A Survey, 2025;1(Feb): 15.

17. Akyash M, Zafari A, Nasrabadi NM. Trading-off Mutual Information on Feature Aggregation for Face Recognition, 2023;2(March): 2044-2051.

18. Zhang Y, Yi P, Zhou D, et al. CSANet: Channel and spatial mixed attention CNN for pedestrian detection. IEEE Access, 2020;8(April): 76252.

19. Aboussaleh I, Riffi J, El Fazazy K, et al. Efficient U-Net architecture with multiple encoders and attention mechanism decoders for brain tumor segmentation. Diagnostics, 2023;13(5): 872.

20. Amer A, Lambrou T, Ye X. MDA-unet: a multi-scale dilated attention U-net for medical image segmentation. Appl Sci, 2022;12(7): 3676.

21. Gaál G, Maga B, Lukács A. Attention U-net based adversarial architectures for chest X-ray lung segmentation. CEUR Workshop Proc, 2020;2692(March): 1-7.

22. Rahman T, Khandakar A, Kadir MA, et al. Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. IEEE Access, 2020;8(July): 191586-191601.