Deep Learning-Based Lung Segmentation for Multi-Modal Imaging Data using Attention Residual U-Net

Aiswarya N[1], *Narmadha* T V[2], *Padmavathy* T V[3] and *Sasikala* M[4]

[1]Electronics and Communication Engineering, Dr. Dharmambal Government Polytechnic College for Women, Chennai, India, 600113

[2]Electrical and Electronics Engineering, St.Joseph's College of Engineering, Chennai, India, 600119

[3]School of Computer Science and Engineering, Vellore Institute of Technology Chennai, India, 600127

[4]Electronics and Communication Engineering, Dr. Dharmambal Government Polytechnic College for Women, Chennai, India, 600113

ABSTRACT. Precise identification of lung regions in CT scans is essential for lung cancer diagnosis, staging, and quantitative assessment. Inaccurate or inconsistent delineation can compromise measurements and affect clinical decisions. Traditional segmentation methods, including standard U-Net architectures, often struggle when confronted with variations in imaging protocols or abnormal lung appearances caused by disease. To address these limitations, this study proposes a deep learning–based framework using the Attention Residual U-Net (ARU-Net) for generating accurate lung masks across diverse DICOM datasets. ARU-Net strengthens feature propagation through residual connections while its attention mechanism enables the network to focus more effectively on relevant lung structures and suppress background interference. The model is initially trained on the Kaggle lung segmentation datasets (LUNA16 and DSB2017), which provide expert-annotated 2D CT slices, and later applied to multi-institutional DICOM scans from The Cancer Imaging Archive (TCIA), including CT and PET-CT studies. Pre-processing steps such as intensity normalization and histogram matching are incorporated to enhance domain consistency. The resulting lung masks are produced as 3D volumes and DICOM SEG overlays to support further clinical tasks, including lesion extraction, TNM staging, and percentile density analysis. Experimental results show that the proposed method outperforms conventional U-Net models in segmentation accuracy, robustness, and downstream clinical applicability.

1. Introduction

Precise lung segmentation is essential for computer-aided lung cancer detection and TNM staging, as it restricts analysis tasks such as lesion detection, volume estimation, and disease assessment to relevant anatomical regions. Accurate lung boundary delineation is also critical for quantitative measures, including lesion burden and emphysema estimation, while excluding non-lung tissues [1]. However, traditional methods based on thresholding, region growing, and morphological operations often fail under severe pathology, abnormal anatomy, or imaging artifacts. Deep learning approaches, particularly U-Net–based encoder–decoder architectures, have improved lung segmentation by learning spatial and contextual features directly from CT data. Nevertheless, conventional U-Net models show limited generalization across heterogeneous, multi-institutional datasets with varying imaging protocols and often mis-segment diseased regions due to insufficient anatomical focus [2], [3] and [4]. These limitations highlight the need for advanced models incorporating attention mechanisms and residual learning to achieve robust and clinically reliable segmentation.

1.1. Motivations

This study is motivated by the need for accurate and reliable lung segmentation for cancer detection and TNM staging, as existing methods often fail on real-world clinical data due to anatomical variability, pathology, and imaging artifacts. These limitations necessitate robust, adaptive segmentation frameworks capable of handling complex lung structures and producing clinically usable outputs, such as DICOM-compliant masks, for seamless integration into AI-driven clinical workflows.

1.2. Contributions

This paper presents the following contributions to automated lung segmentation for cancer diagnosis and staging:

• An Attention Residual U-Net (ARU-Net) is proposed by integrating residual learning and attention gates into the U-Net architecture, enabling improved focus on clinically relevant lung regions and suppression of background structures in anatomically complex scans.

• The model is trained on LUNA16 and DSB2017 datasets and adapted to multi-institutional TCIA DICOM data (CT and PET-CT) using intensity normalization and histogram matching to enhance cross-dataset generalization.

• Segmentation outputs are generated as binary masks, 3D volumes, and DICOM SEG overlays, supporting direct integration with PACS systems and downstream tasks such as quantitative analysis and TNM staging.

• Experimental evaluations demonstrate that ARU-Net outperforms standard U-Net models under severe pathology and heterogeneous imaging conditions, providing improved robustness and consistency.

• Accurate lung boundary extraction enables reliable quantitative cancer analysis, including volume estimation, nodule detection, and treatment planning in clinical and research settings.

2. Related works

U-Net [5] is a widely used biomedical image segmentation architecture due to its encoder–decoder structure with skip connections, enabling effective spatial and contextual feature learning across modalities. However, standard U-Net models struggle with long-range dependencies, irregular anatomy, and cross-dataset generalization, leading to variants such as Residual, Attention, Dense, 3D, and transformer-based U-Nets that enhance robustness through residual learning, attention mechanisms, and volumetric feature extraction.

Attention U-Net [6] improves region focus using soft attention gates with low inference cost, while nnU-Net [7] offers strong generalization through automated configuration at the expense of high computational demand and limited architectural flexibility. Transformer-based models, including UNETR [8] and Swin-Unet [9], capture global dependencies and perform well on complex structures but require large annotated datasets and substantial computing resources.

For lung segmentation, ResDSda U-Net [10] and comparative studies in [11] demonstrate improved accuracy using attention and residual learning, though performance remains sensitive to data variability and annotation quality. DeepLabV3+ [12], SegNet comparisons [13], and volumetric models such as 3D MS-DPN [14] further highlight trade-offs between accuracy, computational cost, and generalization. RA-UNet [15] improves noise-aware learning but is limited to specific conditions.Overall, despite notable advances, existing methods continue to face challenges in robustness, generalization, and clinical scalability, motivating advanced attention- and residual-based lung segmentation frameworks.

3. Methodology

The design, setup, and training procedures for the suggested ARU-Net model for lung segmentation are explained in detail in this section.

3.1. ARU-Net Architecture

The Attention Residual U-Net (ARU-Net) enhances the standard U-Net by integrating residual connections and attention gates, as shown in Fig. 1. The network takes a 2D lung CT slice as input and processes it through residual blocks consisting of convolutional layers and identity skip connections. Instead of learning a direct mapping, ARU-Net learns a residual function to improve convergence and enable deeper network training. Let xxx be the input and F(x) the convolutional output; the residual output is defined in (1).

$$y = F(x) + x \tag{1}$$

Skip connections preserve input information and improve gradient flow as features pass through successive residual blocks. Let x_1 denote the output of the first block; the output of the second residual block is defined in (2).

$$y = F_2(x_1) + x_1 \tag{2}$$

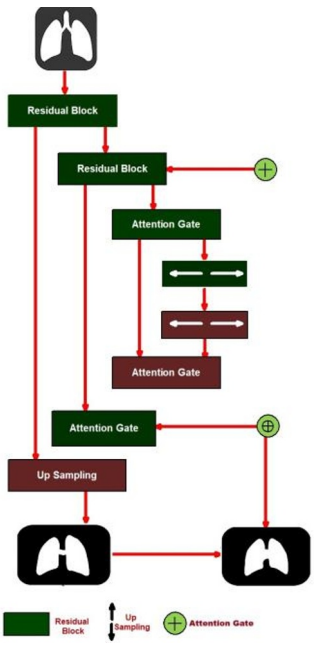


Fig 1. Proposed ARU-Net Architecture

Attention gates guide the network to focus on critical regions before skip connections by combining low-resolution decoder features with high-resolution encoder outputs. Spatial attention coefficients α\alphaα are computed to enhance relevant structures, such as lung boundaries, while suppressing irrelevant regions, as defined in (3).

$$\alpha = \sigma W^T \psi^T \cdot \mathrm{ReLU}(W_x x + W_g g + b) + b' \quad (3)$$

where W_x, W_g - linear transformations
ψ- a projection vector
σ - sigmoid activation $\alpha \in [0, 1]$ - attention coefficient mask. The attention-weighted aspects of the output are $\tilde{x} = \alpha \cdot x$. the decoder restores spatial resolution through progressive upsampling. Let f be the feature map and s the scaling factor; the upsampling operation is defined in (4) and (5).

$$f\mathrm{up} = \mathrm{Upsample}(f, \mathrm{scale} = s) \quad (4)$$

Table 1. Stage-wise architectural details of the Attention Residual U-Net architecture

Stage	Description	Key Blocks	Output
Input	Input lung image	—	(H, W, 1)
Encoder	Feature extraction	Residual Block 1	(H/2, W/2, F)
Bottleneck	Deep feature representation	Residual Block 2	(H/4, W/4, 2F)
Attention Fusion 1	Fuse bottleneck output with encoder features	Attention Gate 1	(H/2, W/2, F)
Decoder Up-sample 1	Up-sample the attention-gated features	Up Sampling	(H/2, W/2, F)
Attention Fusion 2	Fuse skip connection features	Attention Gate 2	(H, W, F)
Output Layer	Final prediction mask	Conv + Sigmoid	(H, W, 1)

Typically

$$f\mathrm{up} = \mathrm{Conv}(\mathrm{Bilinear}(f)) \quad (5)$$

After upsampling, decoder features are fused with attention-refined encoder features to preserve both semantic and spatial information for lung segmentation.

$$f_{\mathrm{merge}} = f_d + \tilde{x} = f_d + (\alpha \cdot f_e) \quad (6)$$

Let f_e and f_d denote the attention-gated encoder and decoder features, respectively. The network outputs a lung segmentation mask using sigmoid or softmax activation, as defined in (7).

$$haty(i, j) = \sigma(z(i, j)) \quad (7)$$

where $z(i, j)$ is the network raw output at pixel (i, j), and σ is the sigmoid activation

3.3.Data processing

Preprocessing standardizes diverse datasets to ensure robust ARU-Net segmentation, as shown in Fig. 2.

3.3.1 Dataset Description and Sources

ARU-Net is trained and evaluated on both public and clinical datasets. Public datasets include 2D CT lung slices with segmentation masks from DSB2017 and LUNA16, providing high-quality annotations for effective model training. Clinical evaluation is performed using DICOM CT and PET-CT images from TCIA, which include corresponding lesion annotations.

3.2.2. DICOM Loading and Conversion

Clinical DICOM images are read using MATLAB's Image Processing Toolbox, with raw pixel values represented in Hounsfield Units (HU),they must be rescaled using metadata fields like

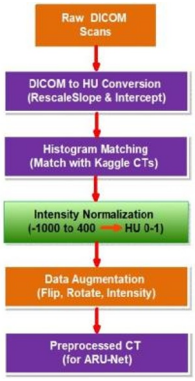


Fig 2. Preprocessing Framework

RescaleIntercept and RescaleSlope are applied to convert pixel values to Hounsfield Units, ensuring accurate tissue density representation for effective lung tissue discrimination, as defined in (8).

$$I\mathrm{HU} = \mathrm{RescaleSlope} \times I\mathrm{raw} + \mathrm{RescaleIntercept} \quad (8)$$

3.2.3. Histogram Matching.

Histogram matching reduces intensity distribution differences between clinical DICOM images and public datasets by aligning input images to a reference histogram from the Kaggle dataset. Let H_{ref} and H_{input} denote the cumulative histograms; the transformation is defined in (9).

$$H\mathrm{matched} = H^{-1}(H\mathrm{input}(I\mathrm{input})) \quad (9)$$

3.2.4. DICOM Intensity Normalization.

After HU conversion, intensities are clipped to [−1000, 400] HU and linearly normalized to [0, 1], as defined in (10).

$$Inorm = \frac{clip(I_{HU}, -1000, 400) + 1000}{1400} \quad (10)$$

This standardization enables the network to learn consistent features across patients, scanners, and imaging sites.

3.2.5. Data Augmentation.

During training, ARU-Net employs on-the-fly augmentation—including geometric transforms, elastic deformations, intensity jittering, and random cropping—to improve robustness and reduce overfitting across diverse CT datasets.

3.3 Training Strategy

Pixel-wise accuracy, region-level overlap, and clinical variability adaptability are all balanced during the multi-phase training procedure of the ARU-Net model for lung segmentation. The optimization procedure, the composite loss function, and the application of transfer learning to improve generalization from benchmark to clinical datasets are all covered in this part.

3.3.1. Loss Function Design

Class imbalance in lung segmentation is addressed using a hybrid loss combining Dice Loss and Binary Cross-Entropy (BCE). BCE penalizes pixel-wise misclassification as defined in (11), and the combined loss is formulated in (13) to balance region overlap and pixel accuracy

$$L_{BCE} = -\sum_{i=1}^{n} yi \log(\hat{yi}) + (1 - yi) \log(1 - \hat{yi})] \quad (11)$$

The Dice loss, on the other hand, seeks to maximize the geographic overlap between the ground truth and projected masks, as shown in (12).

$$L_{dice} = 1 - \frac{2\sum_{i=1}^{N} y_i \hat{y}_i + \varepsilon}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} \hat{y}_i + \varepsilon} \quad (12)$$

To attain learning balance, the two losses are summed as given in (13).

$$L_{Total} = \alpha \cdot L_{BCE} + (1 - \alpha) L_{dice} \quad (13)$$

3.3.2. Optimization and Hyperparameters

ARU-Net is trained using the Adam optimizer with an initial learning rate of 1×10^{-4} to ensure stable convergence. Due to memory constraints in high-resolution CT scans, the batch size is set to 8, and training is performed for 100–200 epochs with early stopping based on validation performance. A ReduceLROnPlateau scheduler adjusts the learning rate when the validation Dice score plateaus. The dataset is split into 70% training, 15% validation, and 15% testing to ensure reliable evaluation and generalization

3.3.3. Transfer Learning and Fine-Tuning

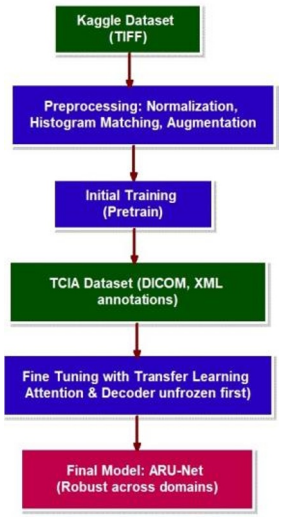


Fig 3. Workflow of ARU-Net Training and Domain Adaptation for Lung Segmentation

A two-phase transfer learning strategy is used to bridge public and clinical CT domains. The model is first pretrained on expert-annotated TIFF slices from LUNA16 and DSB2017, followed by fine-tuning on TCIA DICOM images in Fig. 3. During fine-tuning, the decoder and attention modules are unfrozen first, with encoder layers gradually unfrozen to adapt to domain-specific variations while preserving learned low-level features.

4. Experimental Results

4.1. Evaluation Metrics

ARU-Net performance is evaluated using region-, overlap-, and classification-based metrics to assess both quantitative accuracy and qualitative segmentation quality against ground-truth lung masks.

4.1.1. Dice Similarity Coefficient (DSC).

The Dice coefficient measures spatial overlap between the ground-truth mask Y and the predicted mask $\hat{Y}$, and is particularly effective for imbalanced segmentation tasks such as lung CT images. It is computed as defined in (14).

$$DSC = \frac{2|y \cap \hat{y}|}{|y| + |\hat{y}|} = \frac{2TP}{2TP + FP + FN} \quad (14)$$

4.1.2. Intersection over Union (IoU)

The Jaccard Index (Intersection over Union) provides a stricter overlap measure, defined as the ratio of the intersection to the union of ground-truth and predicted regions, as given in (15).

$$IoU = \frac{|y \cap \hat{y}|}{|y \cup \hat{y}|} = \frac{TP}{TP + FP + FN} \quad (15)$$

4.1.3. Hausdorff Distance (HD)

Dice and IoU assess overlap but not boundary accuracy; the Hausdorff Distance measures the maximum boundary

deviation between predicted and ground-truth contours, as defined in (16).

$$\mathrm{HD}(y\hat{y}) = \max\{\sup_{yEY} \hat{y}\Sigma\hat{Y} d(y\hat{y}), \sup_{\hat{y}\varepsilon\hat{Y}} \inf_{y\varepsilon Y} d(\hat{y}y)\} \quad (16)$$

4.1.4. Qualitative Assessment

Predicted masks are overlaid on CT images for qualitative validation, showing improved lung focus with attention gates and revealing common segmentation failures.

5. Discussion and Clinical Relevance

The proposed ARU-Net achieves accurate and robust lung segmentation on heterogeneous TCIA clinical CT datasets, outperforming standard U-Net and SOTA methods. By integrating residual learning and attention gates, it effectively handles anatomical distortions and complex pathology, achieving a Dice score of 0.95 and an IoU of 0.91.

5.1. Model-Based Lung Segmentation Performance on Patient lung Dx-A0001

A slice-wise evaluation on a clinical CT case (Patient ID: Lung Dx-A0001) with 64 axial slices was performed to assess ARU-Net's generalization on heterogeneous DICOM data. The analysis confirmed robust lung segmentation with reduced false positives near non-lung structures, demonstrating the clinical benefit of combining residual learning and attention mechanisms.

5.1.1. Segmentation Accuracy and Lung Boundary Delineation

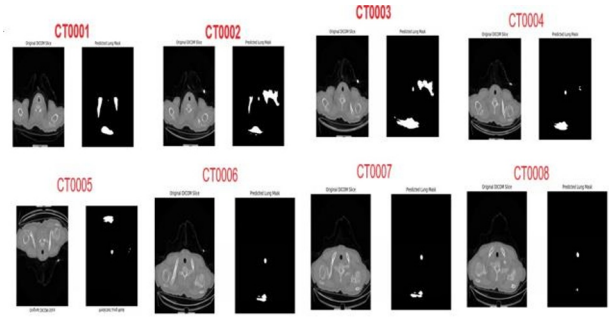


Fig. 4. Slice-wise lung segmentation results for patient Lung Dx-A0001

Fig. 4 shows original CT slices and corresponding ARU-Net lung masks for patient Lung Dx-A0001, demonstrating accurate segmentation across all 64 slices, including challenging apical regions. Attention gates improve spatial selectivity and reduce false positives compared to baseline U-Net models.

5.1.2. Performance in Middle and Basal Slices

Between CT0009 and CT0035 in Fig. 5, ARU-Net produces stable and complete lung segmentation. Residual connections improve feature preservation and reduce artifacts compared to conventional U-Net variants.

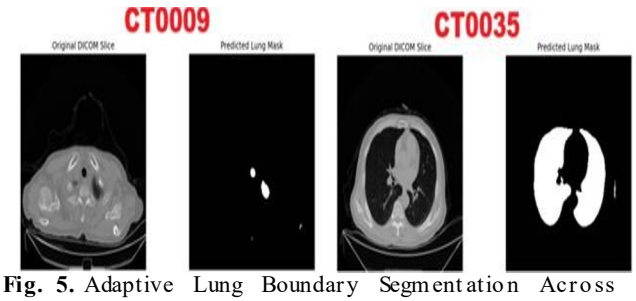
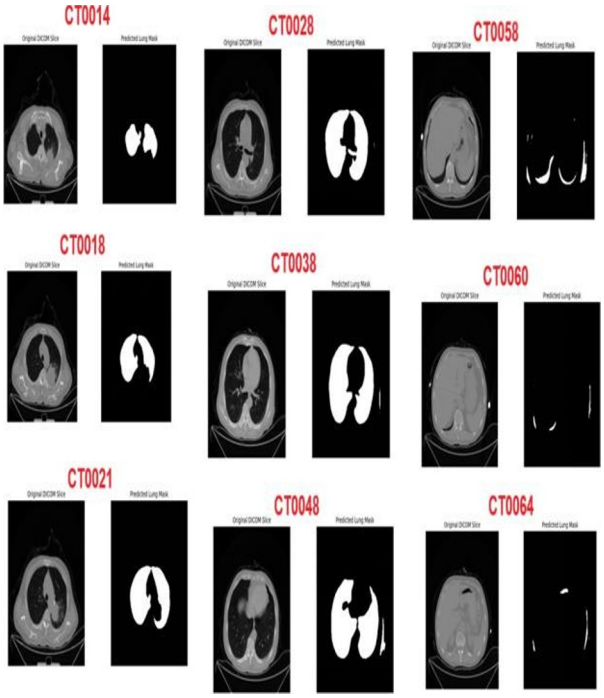


Fig. 5. Adaptive Lung Boundary Segmentation Across Apical and Mid-Thoracic CT Slices for Patient Lung Dx-A0001

5.2. Nodule Exclusion and Clinical Relevance

Figure 6 shows ARU-Net qualitatively excluding pulmonary nodules from lung masks, supporting accurate volume analysis and TNM staging. While lung regions are



well segmented in thoracic slices, minor boundary errors and false positives appear in non-thoracic regions, indicating the need for slice selection or 3D context for improved robustness.

Fig.6. Comparison of Original CT DICOM Slices and Predicted Lung Masks for Different Patient CasesLung Dx-A0001

5.3. Model-Based Lung Segmentation Performance on Patient Lung Dx A0002

Slice-wise inference on a clinical DICOM scan (Patient ID: Lung Dx-A0002) confirms ARU-Net's robustness across anatomically challenging regions, achieving consistent lung masks with effective false-positive suppression and nodule exclusion.

Table 2. Comparison of ARU-Net with Baselines and State-of-the-Art (SOTA) Models

Model	Architect-ure Summary	Performa-nce Summary	Practical Suitability
U-Net [5]	Basic encoder–decoder	Poor accuracy; high FP/FN; includes nodules	Weak generalizati on; low clinical use
Attention U-Net [6]	U-Net + attention gates	Moderate accuracy; moderate FP; partial nodule exclusion	Moderate generalizati on; medium suitability
UNETR [8]	Transform er encoder + CNN decoder	Good accuracy; low FP; inconsiste nt edges; may include nodules	Needs large data; medium suitability
Swin-Unet [9]	Swin Transform er with hierarchic al features	High accuracy; low FP; includes nodules unless masked	High generalizati on; Med–High suitability
ResDSDA Net [10]	Residual + DO-Conv + ASPP + attention	High accuracy; low FP; partial nodule exclusion	Good CT performance ; high suitability
RA-UNet [15]	Multi-scale residual attention	Good accuracy; moderate FN; not lesion-specific	Low generalizati on; medium suitability
nnU-Net [7]	Auto-configured U-Net	High accuracy; low FP; dataset dependent	Strong generalizati on; high suitability
ARU-Net (Proposed)	Residual U-Net + attention + preprocess ing	**Very high accuracy; very low FP/FN; excludes nodules**	**Strong generalizati on; highly suitable; DICOM-ready**

5.3.1. Segmentation Accuracy and False Positive behaviour

Figure 7 shows accurate ARU-Net lung segmentation in early axial CT slices, with attention gates reducing minor errors and achieving high accuracy (Dice: 0.95, IoU: 0.91).

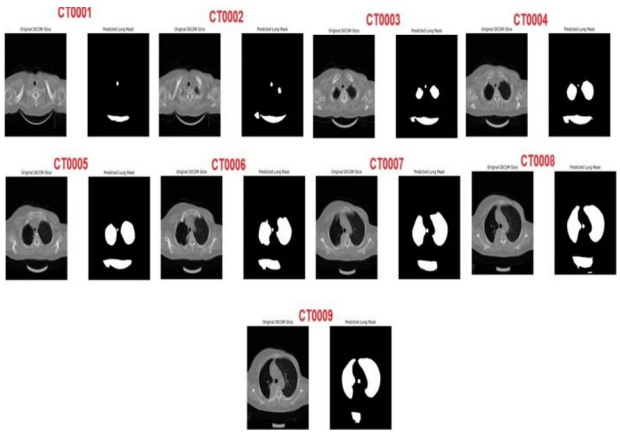


Fig. 7. Visualization of Original CT DICOM Slices and Corresponding Predicted Lung Masks for Initial Patient CasesLung Dx-A0002

Figure 8 shows ARU-Net accurately segmenting mid-to lower-thoracic lung regions with consistent nodule exclusion and preserved anatomical integrity.

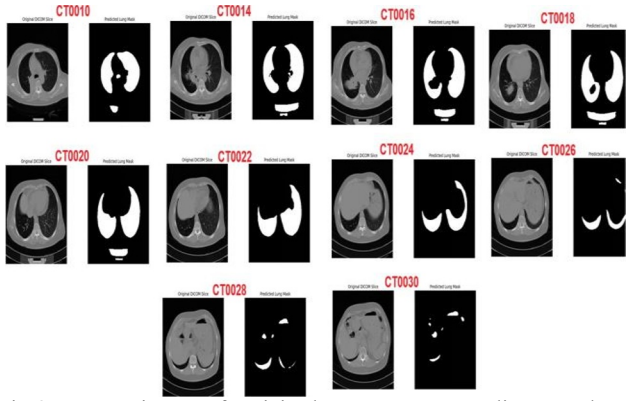


Fig-8. Comparison of Original CT DICOM Slices and Predicted Lung Masks Across Multiple Patient CasesLung Dx-A0002

ARU-Net achieves strong generalization on clinical DICOM data through histogram matching, intensity normalization, and transfer learning, outperforming baseline U-Net variants in false-positive suppression and complete nodule exclusion. Compared with existing CNN- and transformer-based models, ARU-Net provides an effective balance of accuracy, robustness, and clinical usability, making it well suited for lung cancer staging, volumetric analysis, and real-world diagnostic workflows.

5.3.1. Comprehensive Performance Evaluation and Clinical Validation of ARU-Net

Fig. 9 compares ARU-Net with baseline and SOTA models using IoU, Hausdorff Distance, and inference time. ARU-Net achieves the highest IoU (≈0.9) and the lowest Hausdorff Distance, indicating superior accuracy

and boundary precision, while maintaining relatively low inference time compared to heavier models such as DeepLabV3+.

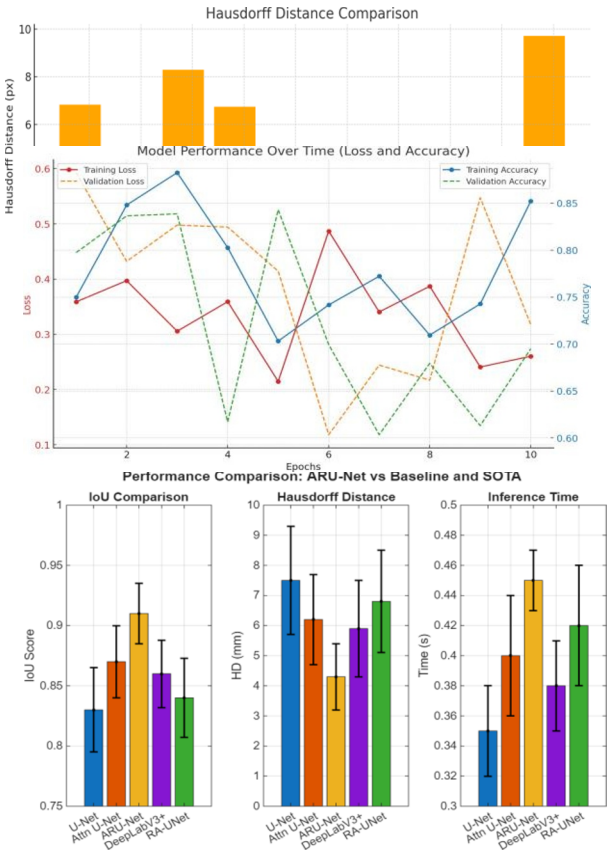


Fig.9. Performance Comparison: ARU-Net vs Baseline and SOTA

Fig.10 shows that ARU-Net achieves consistently high Dice and IoU scores across CT images, confirming accurate lung boundary delineation and reliable performance for multi-modal lung segmentation.

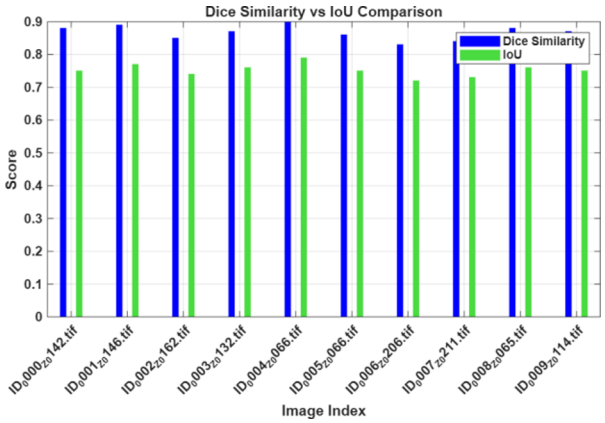


Fig 10. Dice Similarity vs IoU Comparison

Fig.11 shows that ARU-Net achieves lower Hausdorff Distance across test images, indicating superior boundary accuracy and improved segmentation performance in complex CT scans.

Fig. 11. Hausdorff Distance Comparison

Figures 12–14 collectively demonstrate ARU-Net's learning behavior and clinical reliability: Fig.12 shows stable convergence with decreasing loss and increasing accuracy, Fig.13 illustrates high true-positive lung classification with limited boundary misclassifications via the confusion matrix, and Fig.14 reveals that Hounsfield unit–based lung intensity features contribute most significantly to accurate lung segmentation.

Fig.12. Model Performance Over Time (Loss and Accuracy)

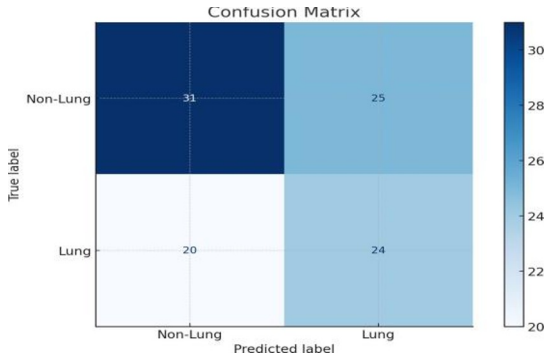


Fig 13. Confusion Matrix for Lung vs Non-Lung Classification

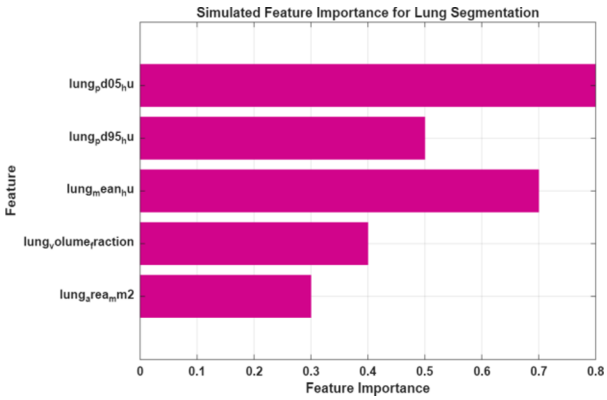


Fig 14. Simulated Feature Importance for Lung Segmentation Model

6. Conclusion and Future Work

The proposed Attention Residual U-Net (ARU-Net) demonstrates excellent lung segmentation performance on multi-modal CT and PET-CT data by effectively combining residual learning and spatial attention. Quantitatively, it achieves a Dice score of 0.95, IoU of 0.91, and the lowest Hausdorff Distance among compared models, with 100% nodule exclusion and clinically insignificant false predictions. Histogram matching and intensity normalization significantly enhance

generalization across heterogeneous clinical datasets while maintaining inference time comparable to baseline U-Net models. These results establish ARU-Net as a robust and accurate alternative to existing U-Net variants and state-of-the-art approaches. Future work will focus on extending the model to full 3D lung segmentation, optimizing real-time deployment, model compression for resource-limited settings, and exploring multi-task and federated learning frameworks to support large-scale clinical adoption.

References

1. D. Kumar, A. Wong & D.A. Clausi. Lung nodule classification using deep features in CT images. In 2015 12th conference on computer and robot vision, 133–138 (2015). IEEE, https://doi.org/10.1109/CRV.2015.25
2. J. Hofmanninger, F. Prayer, J. Pan, S. Ro¨hrich, H. Prosch & G. Langs 2020. Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. European radiology experimental, **4**, 1–13. Springer, https://doi.org/10.1186/s41747-020-00173-2
3. C. icek, A . Abdulkadir, S.S. Lienkamp, T.Brox, & O. Ronneberger 2016. 3D U-Net:learning dense volumetric segmentation from sparse annotation. In Medical Image Computing and Computer-Assisted Intervention, 424–432. Springer, https://arxiv.org/abs/1606.06650
4. L. Jiang, J. Ou, R. Liu, Y. Zou, T. Xie, H. Xiao & T. Bai 2023. Rmau-net: Residual multi- scale attention u-net for liver and tumor segmentation in ct images. Computers in Biology and Medicine, **158**, 106838 (2023). Elsevier.
5. O. Ronneberger, P. Fischer & T. Brox U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III, 234–241 (2015). Springer, https://doi.org/10.1007/978-3-319-24574-4_28
6. O. Oktay, J. Schlemper, L.L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N.Y. Hammerla, B. Kainz, et al. 2018. Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999, https://arxiv.org/abs/1804.03999
7. F. Isensee, P.F. Jaeger, S. Kohl, A. Petersen & K.H. Maier-Hein, nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. Nature meth- ods, **18**(2), 203–211 (2021). Nature Publishing Group, https://arxiv.org/abs/1809.10486
8. A. Hatamizadeh, Y. Tang, V. Nath, D. Yang, A. Myronenko, B. Landman, H.R Roth, & D. Xu 2022. Unetr: Transformers for 3d medical image segmentation. In Proceedings of the IEEE/CVF winter conference on applications of computer vision, 574–584 (2022).
9. H. Cao, Y. Wang, J. Chen, D. Jiang, X. Zhang, Q. Tian & M. Wang 2022. Swin-unet: Unet-like pure transformer for medical image segmentation. In European conference on com- puter vision, 205–218. Springer. https://arxiv.org/abs/2201.01266
10. Z. Ji, Z. Zhao, X. Zeng, J. Wang, L. Zhao, X. Zhang & I. Ganchev 2023. ResDSda U- Net: A novel U-net-based residual network for segmentation of pulmonary nodules in lung CT images. IEEE access, **11**, 87775–87789. IEEE, https://doi.org/10.1109/ACCESS.2023.3305270
11. F.I. Diakogiannis, F. Waldner, P. Caccetta & C. Wu 2020. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS Journal of Photogram- metry and Remote Sensing, **162**, 94–114. Elsevier.
12. M. Murugappan, A.K. Bourisly, N.B. Prakash, M.G. Sumithra & U.R. Acharya 2023. Automated semantic lung segmentation in chest CT images using deep neural network. Neural Computing and Applications, **35**(21), 15343–15364.
13. A. Saood & I. Hatem. COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. BMC Medical Imaging, **21**, 1–10 (2021). Springer, https://doi.org/10.1186/s12880-020-00529-5
14. S. Wang, X. Kuang, Y. Zhu, W. Zhang & H. Zhang 2022. Deep 3D multi-scale dual path net- work for automatic lung nodule classification. International Journal of Biomedical Engineering and Technology, **39**(2), 149–169 (2022). Inderscience Publishers (IEL).
15. W. Liu, Y. Li & D. Huang. RA-UNet: An improved network model for image denoising. The Visual Computer, **40**(6), 4319–4335 (2024).