JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 1 AUGUST 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i1.6735

# PERFORMANCE COMPARISON OF DEEP CNN ARCHITECTURES FOR LUNG AREA SEGMENTATION IN CHEST IMAGING

**Larissa Navia Rani[1*]; Mardison[1]; Agus Perdana Windarto[2]**

Information Systems Study Program[1]
Faculty of Computer Science, Universitas Putra Indonesia YPTK Padang, Indonesia[1]
https://upiyptk.ac.id[1]
larissa_navia_rani@upiyptk.ac.id*; mardison@upiyptk.ac.id

Master's Program in Informatics[2]
STIKOM Tunas Bangsa, Pematangsiantar, Indonesia[2]
https://stikomtunasbangsa.ac.id/web/[2]
agus.perdana@amiktunasbangsa.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)



***Abstract***— *Lung area segmentation is a critical preprocessing step in computer-aided diagnosis systems for respiratory diseases such as lung cancer and pneumonia. Accurate segmentation enhances the detection and monitoring of pathological conditions but manual delineation is time-consuming and subject to variability. This research aims to identify the most effective convolutional neural network (CNN) architecture for automated lung segmentation by comparing three models: U-Net, DeepLab, and a proposed hybrid model combining U-Net with ResUNet_Light. The models were trained and evaluated using a publicly available chest CT dataset under identical experimental settings, including preprocessing steps, training parameters, and standard evaluation metrics: Dice Similarity Coefficient (DSC), Intersection over Union (IoU), Precision, and Recall. Results show that the proposed U-Net + ResUNet_Light model achieves the best performance across all metrics (DSC: 0.6767, IoU: 0.5652, Precision: 0.8480, Recall: 0.7920), outperforming both U-Net and DeepLab. These improvements are attributed to the integration of residual blocks, which enhance feature propagation and gradient flow, enabling better generalization and segmentation accuracy, especially along complex lung boundaries. In contrast, while DeepLab performs well in capturing contextual information, its higher complexity may hinder real-time applicability. U-Net, though efficient, showed limitations in accurately segmenting irregular regions. The findings demonstrate the potential of the proposed model for clinical deployment, where both accuracy and efficiency are critical. This study contributes to the development of more robust deep learning-based segmentation methods and highlights the importance of architectural enhancements in CNN design for medical image analysis.*

***Keywords***: *DeepLab, lung area segmentation, lung segmentation, ResUNet_Light, U-Net*

**Intisari**—*Segmentasi area paru-paru merupakan langkah praproses penting dalam sistem diagnosis berbantuan komputer untuk penyakit pernapasan seperti kanker paru-paru dan pneumonia. Segmentasi yang akurat meningkatkan deteksi dan pemantauan kondisi patologis, tetapi penggambaran manual memakan waktu dan rentan terhadap variabilitas. Penelitian ini bertujuan untuk mengidentifikasi arsitektur jaringan saraf tiruan konvolusional (CNN) yang paling efektif untuk segmentasi paru-paru otomatis dengan membandingkan tiga model: U-Net, DeepLab, dan model hibrida yang diusulkan yang menggabungkan U-Net dengan ResUNet_Light. Model-model tersebut dilatih dan dievaluasi menggunakan dataset CT dada yang tersedia untuk umum dengan pengaturan eksperimen yang identik, termasuk langkah-langkah praproses, parameter pelatihan, dan metrik evaluasi standar: Koefisien Kesamaan Dice (DSC), Intersection over Union (IoU), Presisi, dan Recall. Hasil menunjukkan bahwa model U-Net + ResUNet_Light yang diusulkan mencapai kinerja terbaik di semua metrik (DSC: 0,6767, IoU: 0,5652, Presisi: 0,8480, Recall:*

*0,7920), mengungguli U-Net dan DeepLab. Peningkatan ini dikaitkan dengan integrasi blok residual, yang meningkatkan propagasi fitur dan aliran gradien, memungkinkan generalisasi dan akurasi segmentasi yang lebih baik, terutama di sepanjang batas paru-paru yang kompleks. Sebaliknya, sementara DeepLab berkinerja baik dalam menangkap informasi kontekstual, kompleksitasnya yang lebih tinggi dapat menghambat penerapan waktu nyata. U-Net, meskipun efisien, menunjukkan keterbatasan dalam segmentasi daerah yang tidak teratur secara akurat. Temuan ini menunjukkan potensi model yang diusulkan untuk penerapan klinis, di mana akurasi dan efisiensi sangat penting. Studi ini berkontribusi pada pengembangan metode segmentasi berbasis pembelajaran mendalam yang lebih kuat dan menyoroti pentingnya peningkatan arsitektur dalam desain CNN untuk analisis citra medis.*

***Kata Kunci**: DeepLab, segmentasi area paru-paru, segmentasi paru-paru, ResUNet_Light, U-Net.*

## INTRODUCTION

Lung diseases are among the leading causes of morbidity and mortality worldwide [1], [2]. Early diagnosis and monitoring of lung conditions heavily rely on chest imaging interpretation, particularly thoracic radiography (chest X-ray) and computed tomography (CT) [3], [4]. In clinical practice, lung area segmentation is a critical step for isolating important parts of the image and focusing analysis on relevant areas. However, manual segmentation is highly dependent on the radiologist's expertise and is vulnerable to inter-observer variability [5], [6]. Several convolutional neural network (CNN) architectures have been developed for image segmentation tasks, especially in the biomedical domain [7], [8].

The U-Net architecture is among the most widely adopted due to its symmetric encoder-decoder structure and skip connections, which facilitate precise localization and efficient training even with limited data [9], [10]. However, U-Net often struggles to capture broader contextual information, which may limit its effectiveness in complex anatomical regions [11], [12], [13]. To address this limitation, the DeepLab architecture introduced atrous (dilated) convolutions and Atrous Spatial Pyramid Pooling (ASPP), which allow the model to capture multi-scale context and enhance boundary delineation [14], [15].

Despite its superior segmentation accuracy in various scenarios, DeepLab requires intensive computation and may not be ideal for real-time environments or settings with limited resources [16], [17], [18]. Meanwhile, architectures like ResUNet and its variant, ResUNet_Light, incorporate residual connections to ease the training of deeper networks and mitigate vanishing gradient problems. These models offer a good balance between performance and efficiency, although their performance may be suboptimal in capturing fine-grained structures unless paired with appropriate multi-scale feature extraction mechanisms [19], [20]. Therefore, selecting or designing an optimal architecture often involves balancing accuracy, computational cost, and generalizability across various modalities and imaging conditions.

In previous research conducted by Adnan Saood and Iyad Hatem [21], the study investigated the use of two deep learning methods, SegNet and U-Net, for segmenting COVID-19 infected lung tissue on CT scan images. The strength of this paper lies in the clear comparative approach between the two models, with results showing that SegNet excelled in binary segmentation (accuracy 0.95), while U-Net performed better in multi-class segmentation (accuracy 0.91). The study also made a practical contribution by providing an automatic solution for COVID-19 diagnosis and severity assessment, which was highly relevant during the pandemic.

However, its weakness lies in the relatively small dataset (100 images) and class imbalance, especially for class C3 (pleural effusion), which caused suboptimal model performance in segmenting that class. Additionally, the paper did not elaborate in detail on how data augmentation or other techniques were used to address class continuity issues. Meanwhile, another study by (Turk & Kılıçaslan, 2025)[22] investigated the use of three deep learning models, namely U-Net, V-Net, and Seg-Net, enhanced for lung image segmentation to detect tuberculosis (TB) in chest X-ray images.

The main advantage of this study was the enhancement of U-Net and V-Net architectures with attention mechanisms and non-local blocks, which resulted in more accurate segmentation performance, with Dice coefficients reaching 96.43% and 96.42%, respectively. Moreover, the study used advanced preprocessing techniques like adaptive filtering and histogram equalization to improve image quality, along with rigorous validation via five-fold cross-validation. However, the limitation of this study was the limited dataset size (a combination of Shenzhen and Montgomery datasets) and class imbalance, particularly for the


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 1 AUGUST 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i1.6735

rarer TB cases. Although Seg-Net performed worse than U-Net and V-Net, it remained competitive in classification tasks. The article also did not discuss in detail the implementation challenges in clinical environments or the high computational resource requirements for these models. Overall, the study made a significant contribution to the automation of TB diagnosis but required further testing on more diverse datasets to ensure better generalization.

Based on the reviewed literature, with rapid advancements in deep learning, Convolutional Neural Network (CNN) architectures, particularly encoder-decoder architectures, have become the dominant approach in medical image segmentation tasks [14], [23]. One of the most widely used architectures is U-Net, specifically designed for biomedical image segmentation, which has shown excellent results in various studies. However, basic U-Net has limitations in capturing broader spatial contexts and fine structural details, especially in complex structures like the lungs, which have variable shapes between individuals and are affected by artifacts or diseases [24], [25].

As a solution to U-Net's limitations, architectures such as DeepLab have been introduced. DeepLab uses techniques like Atrous Convolution and Spatial Pyramid Pooling to better capture multi-scale information and global context. This architecture is theoretically superior in understanding more complex spatial structures, but it comes at the cost of higher computational requirements [26].

In an effort to combine the strengths of several approaches, this study also proposes a new hybrid method, U-Net + ResUNet_Light. This architecture combines the efficient segmentation capability of U-Net with the residual connection strength of ResUNet_Light, aiming to improve gradient flow and deepen the network without significantly increasing complexity [10], [27].

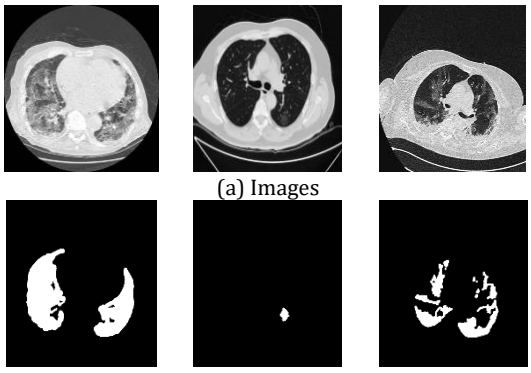
Although each architecture has its own strengths, few studies have comprehensively compared the performance of basic U-Net, DeepLab, and hybrid approaches like U-Net + ResUNet_Light, especially in the context of lung area segmentation in chest images [28]. Therefore, this study is conducted to evaluate and compare the performance of these three architectures using relevant segmentation metrics, such as Dice Coefficient and Intersection over Union (IoU), to determine which approach is optimal for this task [29]. By making this comparison, it is hoped that this study will contribute to the development of more accurate and efficient image-based diagnostic support systems.

## MATERIALS AND METHODS

This section outlines the materials and experimental procedures used in this study. It begins with a description of the dataset used to evaluate the segmentation model performance. Next, we introduce the deep learning architectures being compared—basic U-Net, DeepLab, and the proposed hybrid model, U-Net + ResUNet_Light—highlighting their structural differences and theoretical strengths. Finally, we detail the experimental design, including data preprocessing, training settings, evaluation metrics, and performance comparison, to ensure a fair and reproducible comparison across all models.

### Research Dataset

Lung area segmentation requires high-quality annotated datasets to ensure accurate model training and validation. In this study, we use a publicly available chest imaging dataset that is widely used for segmentation tasks. A sample from the dataset used in this study is shown in Figure 1..



(a) Images

(b) Masks

Source : ([https://www.kaggle.com/datasets/beosup/lung-segment](https://www.kaggle.com/datasets/beosup/lung-segment) 2025)

Figure 1. Research Dataset

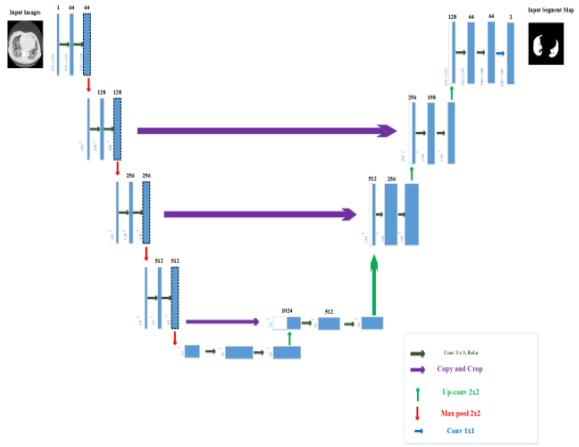
Figure 1 shows a sample from the lung cancer image segmentation dataset used in this study. The dataset consists of a large number of chest X-ray (CXR) or CT images along with corresponding lung masks, which are annotated by medical experts to ensure ground truth accuracy..

### Proposed Model Architecture

To comprehensively evaluate the effectiveness of various CNN architectures for lung segmentation, three different models are implemented: basic U-Net, DeepLab, and the proposed hybrid model combining U-Net with ResUNet_Light.

VOL. 11. NO. 1 AUGUST 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v11i1.6735

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)



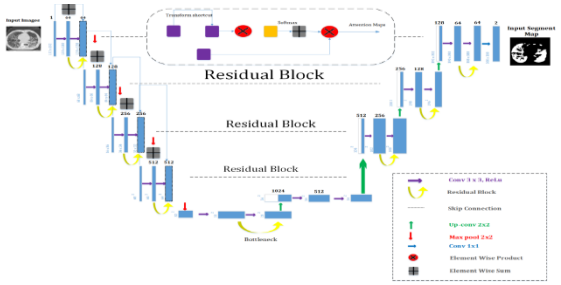
Source: (Saood and Hatem 2021[21])

Figure 2. U-Net Architecture

Figure 2 illustrates the U-Net architecture. The basic U-Net uses the classic encoder-decoder design with skip connections, which directly transfer feature maps from the encoder to the decoder, enabling precise localization. This model is lightweight and easy to train, making it a popular choice for biomedical image segmentation. However, it often lacks the ability to capture global context, especially in cases where lung boundaries are ambiguous or disrupted by pathology.

The DeepLab model, particularly DeepLabv3+, introduces advanced modules such as atrous convolutions and Atrous Spatial Pyramid Pooling (ASPP) to capture contextual features at various scales. This allows the model to better handle irregular shapes and boundaries. However, the trade-off is a significant increase in computational complexity and memory usage, making it less ideal for real-time applications or environments with limited resources.

The proposed model, U-Net + ResUNet_Light, integrates the residual learning capabilities of ResUNet_Light into the U-Net architecture. Residual blocks facilitate gradient flow in deeper networks, allowing for better feature extraction without significantly increasing computational load.




Source: (Research Result, 2025)

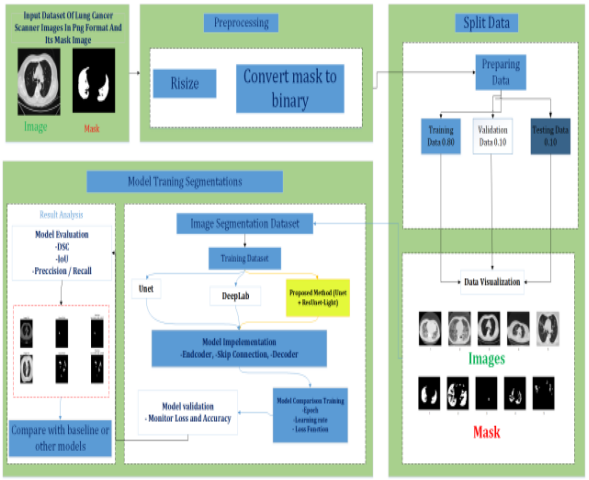
Figure 3. The proposed architecture

Figure 3 illustrates the integration aimed at maintaining U-Net's localization strength while enhancing depth and representational power through residual connections. This hybrid approach seeks a balance between segmentation accuracy and model efficiency. The proposed U-Net + ResUNet_Light architecture combines U-Net's efficient localization with ResUNet_Light's robust feature extraction by incorporating residual blocks at each encoder and decoder level.

The model follows a symmetric encoder-decoder structure with skip connections to preserve spatial information. Standard convolutional layers are replaced by residual blocks, allowing direct gradient flow across layers, addressing vanishing gradient issues in deeper networks. Max pooling is used for downsampling, while transposed convolutions are applied for upsampling in the decoder path. This integration aims for high segmentation accuracy with improved convergence and generalization, particularly in challenging chest imaging scenarios where lung boundaries may be irregular or unclear.

**Research Design**

To ensure an objective and reproducible comparison between models, a controlled experimental design was applied. Each model was trained using the same training and validation datasets, with identical preprocessing, augmentation techniques, and training configurations (e.g., batch size, optimizer, and learning rate). Models were trained for a fixed number of epochs or until convergence, and the best-performing checkpoint was selected based on validation loss.




Source: (Research Result, 2025)

Figure 4. Research Design



Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 1 AUGUST 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i1.6735

Figure 4 illustrates the research design, which outlines the complete process of lung cancer image segmentation using deep learning methods. The performance of each model is evaluated using standard metrics, including Dice Similarity Coefficient (DSC), Intersection over Union (IoU), Precision, Recall, and inference time per image to assess both accuracy and efficiency. Statistical tests are also conducted to determine the significance of performance differences among the models.

The process begins with data input, using a CT scan dataset in PNG format, including the corresponding binary masks that label the cancer areas. The data undergoes preprocessing steps such as resizing the images for uniformity and converting the masks to binary format (1 for cancer regions and 0 for background). The dataset is then split into three parts: 80% for training, 10% for validation, and 10% for testing. Visualizations are created to confirm the accuracy of the data.

In the model training phase, the preprocessed dataset is used to train three models: U-Net, DeepLab, and the proposed hybrid method (U-Net + ResUNet_Light). All models are based on an encoder-decoder architecture, utilizing feature extraction and segmentation reconstruction. Skip connections are incorporated to preserve spatial details, and training parameters, such as the number of epochs, learning rate, and loss function, are adjusted and compared to identify the best-performing model.

The testing model conducted in this study is a comparative experimental test that aims to evaluate and compare the performance of three CNN architectures—U-Net, DeepLab, and U-Net + ResUNet_Light—in the task of segmenting lung areas in chest CT images. Testing was conducted using test data amounting to 10% of the total dataset, which was not previously used in the training process. Each model was tested based on evaluation metrics including Dice Similarity Coefficient (DSC),

Intersection over Union (IoU), Precision, Recall, Accuracy, and Loss to assess segmentation accuracy and efficiency. The test results show that the proposed model U-Net + ResUNet_Light provides the best performance, proving that the integration of residual connections in U-Net is able to improve segmentation accuracy and generalization compared to conventional architectures.

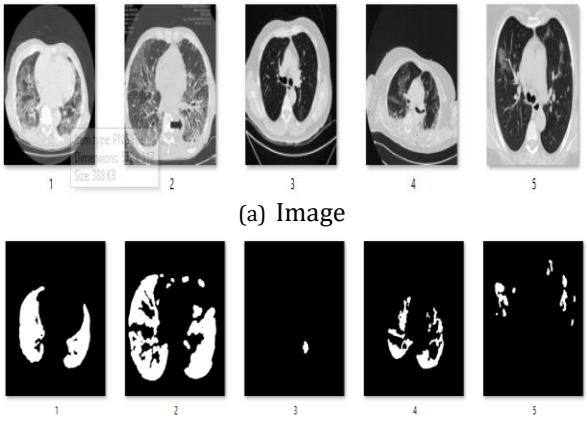
Finally, the results are analyzed using the aforementioned evaluation metrics to compare the predicted segmentation masks with the ground truth. Performance comparison is conducted between the proposed method and the baseline models.

## RESULTS AND DISCUSSION

This section presents the results obtained from the implementation and evaluation of three CNN-based models—U-Net, DeepLab, and the proposed hybrid model, U-Net + ResUNet_Light—for lung area segmentation in chest imaging. Each model was trained and validated on a consistent dataset with identical preprocessing and training configurations. Their performance was evaluated using standard segmentation metrics, including the Dice Similarity Coefficient (DSC), Intersection over Union (IoU), and Precision/Recall, and further compared to assess the relative strengths and limitations of each approach.

### Model Training and Validation

All models were trained using the same training dataset, obtained after preprocessing and splitting the original lung image dataset into 80% training, 10% validation, and 10% testing.



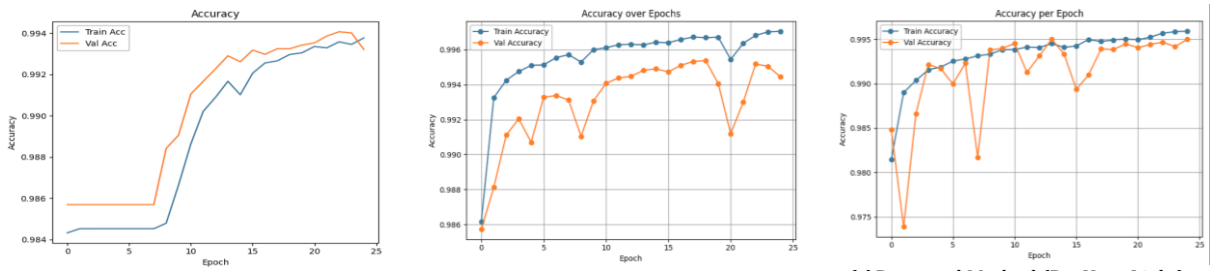
(a) Image

(b) Mask

Source: (Research Result, 2025)

Figure 5. Preprocessing and Data Augmentation

Figure 5, before model training, all images were first processed by resizing them to a uniform resolution, normalizing pixel intensity values, and applying data augmentation techniques such as rotation, flipping, and contrast adjustment to improve model generalization. The dataset was split into training, validation, and testing sets using a stratified approach to ensure a balanced representation of cases across subsets. The training process was monitored using validation loss and accuracy across epochs. Figures 6 and 7 display the accuracy and loss results during the training of each tested model.




(a) Unet (b) DeepLab (c) Proposed Method (ResUnet Light)

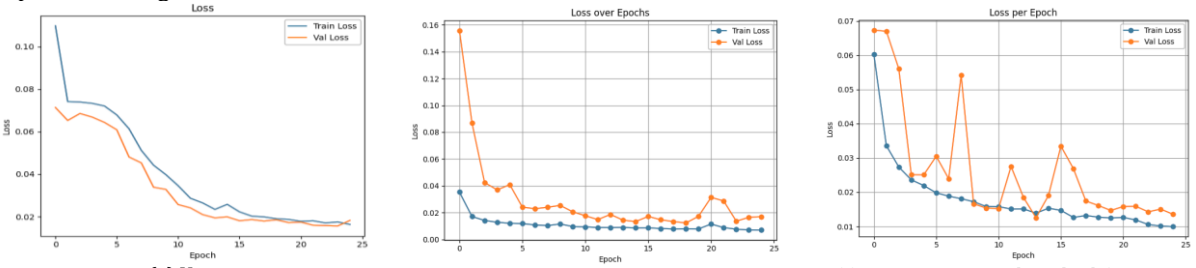
Source: (Research Result, 2025)

Figure 6. Comparison of Accuracy for Each Model

For U-Net (Figure 6a), the training accuracy increases steadily and reaches a high value of approximately 0.994 by the end of the 25th epoch. However, the validation accuracy lags behind the training accuracy, gradually increasing with a less pronounced growth rate. This indicates that while the model is learning well from the training data, it may be starting to overfit as the gap between the train and validation accuracy widens. The overfitting is especially evident in the later epochs, which may limit the model's ability to generalize to unseen data.

In the case of DeepLab (Figure 6b), the training accuracy also increases over time, but with noticeable fluctuations in the middle epochs. The accuracy plateaus toward the end, suggesting that the model has reached its learning capacity, and additional training may not lead to significant improvements. The validation accuracy for DeepLab shows more variability, with peaks and dips, indicating that the model is more sensitive to changes in the validation set. These oscillations suggest that DeepLab, with its more complex architecture, may be prone to overfitting or instability when applied to the validation data, although it still achieves high performance.

The Proposed Method (ResUNet Light) (Figure 6c) demonstrates a more stable and consistent performance across epochs. The training accuracy increases steadily and reaches approximately 0.995 by the 25th epoch, similar to the other models. However, what sets this model apart is the smoother progression of validation accuracy, which remains relatively close to the training accuracy. This indicates that the ResUNet Light method is better at generalizing to unseen data, with fewer signs of overfitting. The validation curve's stability suggests that this model strikes a better balance between learning from the data and generalizing effectively, making it potentially more robust for deployment in real-world applications..




(a) Unet (b) DeepLab (c) Proposed Method (ResUnet Light)

Source: (Research Result, 2025)

Figure 7. Comparison of Loss for Each Model

For U-Net (Figure 7a), the training loss decreases rapidly in the early epochs, quickly stabilizing at a low value around 0.02 by the 25th epoch, indicating that the model effectively learns from the training data. However, the validation loss, although also decreasing, remains consistently higher than the training loss, reaching a stable value around 0.04. This discrepancy suggests that the model might be overfitting, as it fits the training data well but struggles to generalize to the unseen validation set. The higher validation loss reinforces the earlier observation that U-Net may not perform as well when applied to real-world, unseen data.

In the case of DeepLab (Figure 7b), the training loss decreases sharply in the first few epochs, approaching a low value near 0.02, and then stabilizes with slight fluctuations between epochs. However, the validation loss behaves erratically, experiencing large peaks and dips


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 1 AUGUST 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i1.6735

throughout the training process before eventually stabilizing at around 0.02. These fluctuations suggest that the model is highly sensitive to the validation data and may have difficulty generalizing effectively. The erratic behavior of the validation loss indicates that DeepLab, with its more complex architecture, is prone to overfitting and instability, limiting its practical application in real-world settings.

The Proposed Method (ResUNet Light) (Figure 7c) shows a similar trend, with the training loss decreasing quickly in the initial epochs and stabilizing at a low value of around 0.02 by the 25th epoch. The validation loss also decreases rapidly in the early stages but, unlike DeepLab, experiences fewer fluctuations. It stabilizes around 0.01 by the end of training, indicating a more consistent and stable performance. The smooth decrease in validation loss suggests that ResUNet Light is better at generalizing to unseen data, with fewer signs of overfitting or instability. This stability in both training and validation loss indicates that ResUNet Light strikes the best balance between fitting the training data and maintaining robust generalization.

Table 1. Model Validation Comparison

| Model | Accuracy | Loss |
|---|---|---|
| U-Net | 0,9938 | 0,0163 |
| DeepLab | 0,9862 | 0,0353 |
| U-Net + ResUNet_Light (Proposed) | 0,9959 | 0,0100 |

Source: (Research Result, 2025)

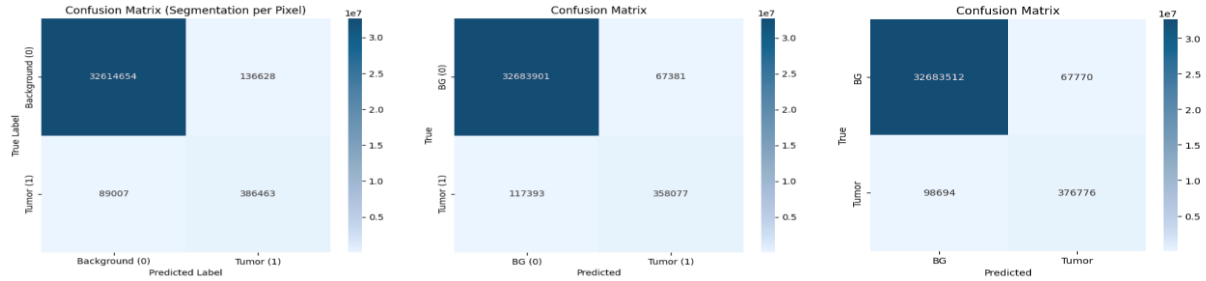
Table 1 provides a comparative analysis of the validation performance for three models: U-Net, DeepLab, and the proposed U-Net + ResUNet_Light. The comparison is based on two key metrics: Accuracy and Loss.

The U-Net model achieved an accuracy of 99.38%, indicating strong performance on the validation set. However, its validation loss of 0.0163 is relatively higher compared to the other models, suggesting that while the model performs well, it still has some room for improvement in terms of minimizing errors and generalizing to unseen data.

DeepLab, on the other hand, achieved a lower accuracy of 98.62% and the highest validation loss of 0.0353 among the three models. This higher loss value, along with the lower accuracy, aligns with the earlier observations of fluctuating validation loss, suggesting that DeepLab faces more challenges in terms of stability and generalization.

The Proposed Model (U-Net + ResUNet_Light) outperforms both U-Net and DeepLab, achieving the highest accuracy of 99.59% and the lowest validation loss of 0.0100. This demonstrates its superior ability to not only achieve high accuracy but also minimize errors during validation, indicating that it is more effective at generalizing to new data without overfitting.

In summary, the proposed U-Net + ResUNet_Light model delivers the best overall performance, achieving both the highest accuracy and the lowest loss, making it the most robust and reliable model for lung segmentation tasks. While U-Net provides strong results, it is outperformed by the proposed model, and DeepLab shows more variability in performance, making it less reliable in comparison.




(a) Unet (b) DeepLab (c) Proposed Method (ResUnet Light)

Source: (Research Result, 2025)

Figure 8. Confusion Matrix Comparison for Each Model

The confusion matrix analysis for the three models—U-Net, DeepLab, and the proposed U-Net + ResUNet_Light—reveals insights into the performance of each model in terms of correctly classifying background and tumor pixels.

For U-Net (Figure 8a), the model correctly identifies a large number of background pixels (32,614,654) and tumor pixels (386,463). However, there are noticeable errors, with 136,628 background pixels misclassified as tumors and 89,007 tumor pixels misclassified as background.

These errors indicate that while U-Net performs well, there is room for improvement in distinguishing between background and tumor pixels, particularly in reducing false positives and false negatives.

DeepLab (Figure 8b) shows similar performance but with some improvements in the classification of background pixels, correctly identifying 32,683,901 background pixels. However, it still misclassifies 67,381 background pixels as tumors and 117,393 tumor pixels as background. This results in a significant number of false negatives, where tumors are misclassified as background, showing that DeepLab still faces challenges in tumor detection, despite performing better than U-Net in some areas.

The Proposed Method (ResUNet Light) (Figure 8c) outperforms both U-Net and DeepLab, correctly identifying 32,683,512 background pixels and 376,676 tumor pixels. The number of false positives (67,770) and false negatives (98,694) is lower compared to the other two models, indicating that ResUNet Light has better precision and recall. This model is more effective at detecting tumors and classifying background pixels, making it the most reliable approach among the three for lung segmentation tasks.

In conclusion, while all models show strong performance, the Proposed Method (ResUNet Light) demonstrates the best overall accuracy, with the highest true tumor detection and lowest false negatives and false positives. This suggests that it provides the most robust and reliable segmentation, particularly for distinguishing between background and tumor areas.

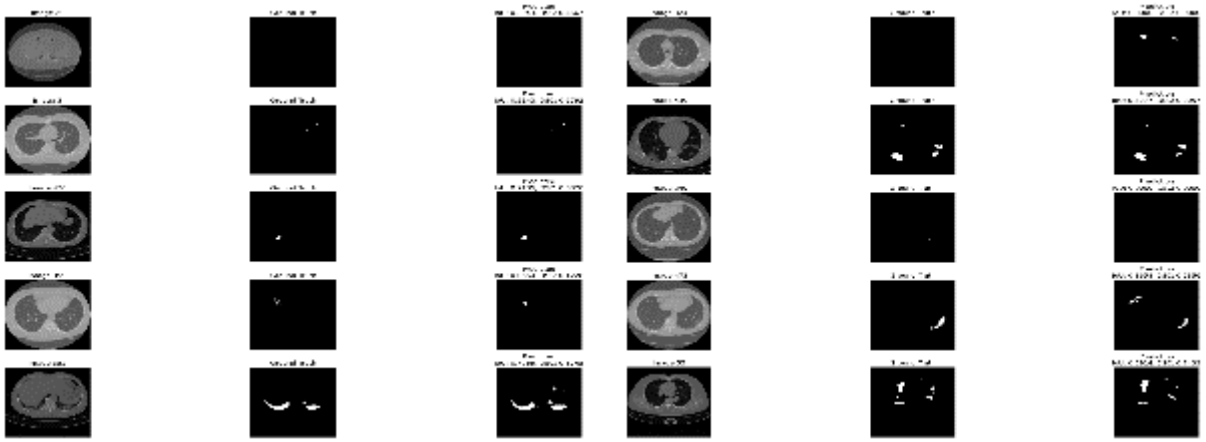
**Segmentation Performance**

The visual comparison of segmentation results further validates the quantitative findings. The U-Net model occasionally fails to accurately segment peripheral lung regions, particularly in cases with irregular morphologies. This indicates that while U-Net performs well in many instances, it struggles with complex structures and shapes, leading to imperfect segmentations in more challenging areas.

In contrast, DeepLab generates smoother segmentation contours, offering more refined boundaries. However, its broader receptive field sometimes results in over-segmentation, where the model includes regions beyond the actual lung tissue, causing errors in the segmented areas. This suggests that while DeepLab benefits from its ability to capture broader contexts, it is less precise in its delineation of lung boundaries.

The Proposed Method (U-Net + ResUNet_Light) achieves the most accurate segmentation, especially around complex boundaries. This model demonstrates superior performance in handling irregular lung structures, effectively generalizing across different lung shapes and morphologies. The results indicate that the proposed method is highly capable of managing the complexity of lung segmentation, achieving precise delineation of both peripheral and central lung areas.

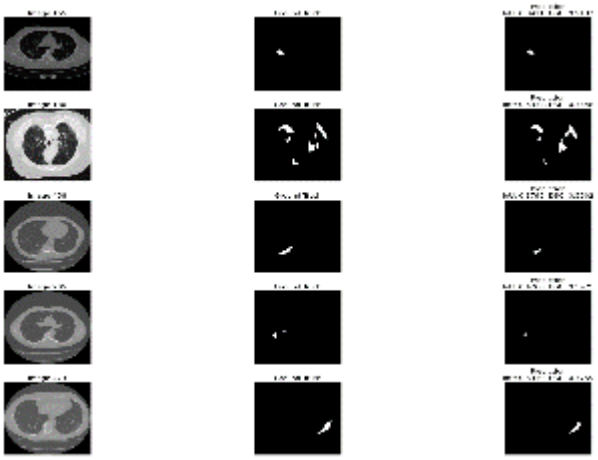
Quantitative evaluation was performed on the test dataset using three metrics, and the visual samples from the proposed model (ResUNet Light) showcase its superior segmentation accuracy, reinforcing its reliability for lung tissue delineation.



(a) Unet (b) DeepLab


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 1 AUGUST 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i1.6735



(c) Proposed Method (ResUnet Light)

Source: (Research Result, 2025)

Figure 9. Example of Test Results for All Models

The visual comparison of segmentation results across the three models—U-Net, DeepLab, and the proposed U-Net + ResUNet_Light—demonstrates the varying degrees of accuracy and reliability in lung segmentation tasks. The U-Net model performs (Figure 9a) decently in simpler cases, but it struggles with more complex lung structures, particularly the peripheral regions. For example, in Image 20, the predicted segmentation is incomplete, especially near the lung boundary, with a Dice Similarity Coefficient (DSC) of 0.4672, indicating significant misalignment with the ground truth. In other cases, such as Image 353, while U-Net achieves a DSC of 0.7136, there are still noticeable under-segmentation issues. These results suggest that while U-Net is capable, it needs further refinement, particularly in handling challenging cases with irregular lung structures.

DeepLab (Figure 9b) performs better than U-Net, with smoother segmentation contours. For instance, in Image 185, DeepLab's prediction is much closer to the ground truth, with a DSC of 0.6957. However, it still encounters problems with under-segmentation in peripheral lung regions, as seen in Image 136, where the DSC drops to 0.6342. Additionally, in more complex images like Image 418, where DeepLab produces some over-segmentation, the DSC reaches 0.7104. This indicates that while DeepLab offers smoother boundaries, it sometimes struggles with both over-segmentation and missing smaller details, particularly in more intricate regions of the lung.

The Proposed Method (U-Net + ResUNet_Light) (Figure 9c) stands out as the best-performing model, consistently achieving high accuracy in all cases. For example, in Image 185, it achieves an impressive DSC of 0.9443, with nearly perfect segmentation of the lung area. In more complex cases like Image 418, the model performs exceptionally well with a DSC of 0.9404, showcasing its ability to generalize effectively across different lung structures. The model also handles peripheral lung areas with remarkable precision, as seen in Image 473, where it reaches a DSC of 0.9729, further validating its robustness and reliability.

In conclusion, while U-Net performs reasonably well on simpler cases, it struggles with under-segmentation in more complex regions. DeepLab provides smoother contours but encounters issues with over-segmentation and missed details. The Proposed Model (ResUNet Light) excels in all aspects, offering the most precise and reliable segmentation across various lung cases. Its superior performance in both peripheral and central lung regions, as demonstrated by its consistently high DSC scores, makes it the most effective model for lung segmentation tasks.

Table 2. Model Evaluation Comparison

| Model | DSC | IoU | Precision | Recall |
|---|---|---|---|---|
| U-Net | 0,6507 | 0,5246 | 0,739 | 0,813 |
| DeepLab | 0,6544 | 0,5457 | 0,842 | 0,753 |
| U-Net + ResUNet_Light (Proposed) | 0,6767 | 0,5652 | 0,848 | 0,792 |

Source: (Research Result, 2025)

Table 2 presents a comprehensive evaluation of three models—U-Net, DeepLab, and the Proposed Model (U-Net + ResUNet_Light)—using key segmentation metrics: Dice Similarity Coefficient (DSC), Intersection over Union (IoU), Precision, and Recall. In terms of Dice Similarity Coefficient (DSC), the Proposed Model (U-Net + ResUNet_Light) outperforms both U-Net and

DeepLab, achieving the highest score of 0.6767. This indicates that the proposed model provides the best overlap between predicted and ground truth lung regions, highlighting its superior segmentation accuracy. Both DeepLab and U-Net show similar performance, with DSC scores of 0.6544 and 0.6507, respectively. While these scores are respectable, they fall short compared to the performance of the proposed model.

Looking at the Intersection over Union (IoU), which measures the ratio of the intersection to the union of the predicted and true regions, the Proposed Model again leads with a score of 0.5652, closely followed by DeepLab at 0.5457. U-Net scores the lowest in this metric with 0.5246. This further reinforces the advantage of the proposed method in accurately delineating the lung regions with minimal overlap errors.

When considering Precision, which measures the accuracy of the predicted positive regions, the Proposed Model excels with a precision score of 0.848. This indicates that it has the highest ability to correctly classify lung regions without many false positives. DeepLab follows closely with a precision of 0.842, while U-Net lags behind at 0.739. This suggests that while U-Net has strong recall, its precision suffers compared to the other two models.

In terms of Recall, which measures how well the model captures all the true positive regions, U-Net achieves the highest recall of 0.813, demonstrating that it is the best at identifying all relevant lung areas. However, it comes at the cost of lower precision. The Proposed Model ranks second with a recall score of 0.792, indicating that it captures almost as many true lung regions as U-Net, but with fewer false positives. DeepLab has the lowest recall score of 0.753, suggesting that it misses some true lung areas, which is consistent with its lower performance in this metric.

**Discussion**

the Proposed Model (U-Net + ResUNet_Light) outperforms both U-Net and DeepLab across all key evaluation metrics. It strikes an optimal balance between precision and recall, making it the most reliable and accurate model for lung segmentation tasks. While U-Net excels in recall, it lacks the precision required for high-quality segmentation. DeepLab, although strong in precision, struggles with recall, making the Proposed Model the most effective and robust solution for accurate lung segmentation.

The computational details for this study include the following hardware: NVIDIA RTX 3060 GPU (12GB VRAM), Intel Core i7-11700F CPU, and 32GB DDR4 RAM. These details are included to support replication and evaluate the computational efficiency of each model.

## CONCLUSION

Based on the experimental evaluation results, this study shows that the proposed U-Net + ResUNet_Light model provides the best performance in lung area segmentation in chest CT images compared to U-Net and DeepLab. The proposed model recorded the highest scores in all evaluation metrics, namely DSC of 0.6767, IoU of 0.5652, Precision of 0.848, and Recall of 0.792, outperforming both U-Net and DeepLab. This achievement reflects the successful integration of residual blocks in improving segmentation accuracy and improving information flow in the U-Net architecture. Thus, the research objective of designing and testing a more effective architecture has been achieved, and the proposed model is proven to be technically superior and has the potential to be applied in clinical diagnosis support systems that require reliable and high-precision medical image segmentation. Future research could explore integrating attention mechanisms or transformer-based modules to further enhance segmentation performance, as well as testing the model on diverse and multi-center datasets to evaluate its generalizability in real-world clinical environments.

## REFERENCE

[1] V. Srivastava, “An enhanced texture-based feature extraction approach for classification of biomedical images of ct-scan of lungs,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 7, pp. 18–25, 2021, doi: 10.9781/ijimai.2020.11.003.

[2] “Retraction: An Efficient Model for Lungs Nodule Classification Using Supervised Learning Technique (Journal of Healthcare Engineering (2023) 2023 (8262741) DOI: 10.1155/2023/8262741),” *Journal of Healthcare Engineering*, vol. 2023, 2023, doi: 10.1155/2023/9767402.

[3] J. Liu, H. Shao, Y. Jiang, and X. Deng, “CNN-Based Hidden-Layer Topological Structure Design and Optimization Methods for Image Classification,” *Neural Processing Letters*, vol. 54, no. 4, pp. 2831–2842, 2022, doi: 10.1007/s11063-022-10742-8.

[4] İ. Atik, “Pneumonia detection on chest x-ray images using residual convolutional neural


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 1 AUGUST 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i1.6735

network," *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol. 39, no. 3, pp. 1719–1731, 2024, doi: 10.17341/gazimmfd.1271385.

[5] B. Hunter, "Radiomics-based decision support tool assists radiologists in small lung nodule classification and improves lung cancer early diagnosis," *British Journal of Cancer*, vol. 129, no. 12, pp. 1949–1955, 2023, doi: 10.1038/s41416-023-02480-y.

[6] P. Ajmera *et al.*, "A deep learning approach for automated diagnosis of pulmonary embolism on computed tomographic pulmonary angiography," *BMC Medical Imaging*, vol. 22, no. 1, pp. 1–9, 2022, doi: 10.1186/s12880-022-00916-0.

[7] A. P. Windarto, T. Herawan, and P. Alkhairi, "Early Detection of Breast Cancer Based on Patient Symptom Data Using Naive Bayes Algorithm on Genomic Data," in *Artificial Intelligence, Data Science and Applications*, Y. Farhaoui, A. Hussain, T. Saba, H. Taherdoost, and A. Verma, Eds., Cham: Springer Nature Switzerland, 2024, pp. 478–484.

[8] A. P. Windarto, I. R. Rahadjeng, M. N. H. Siregar, and P. Alkhairi, "Deep Learning to Extract Animal Images With the U-Net Model on the Use of Pet Images," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 8, no. 1, pp. 468–476, 2024.

[9] P. Rao, "Weight pruning-UNet: Weight pruning UNet with depth-wise separable convolutions for semantic segmentation of kidney tumors," *Journal of Medical Signals and Sensors*, vol. 12, no. 2, pp. 108–113, 2022, doi: 10.4103/jmss.jmss_108_21.

[10] F. Valeri, "UNet and MobileNet CNN-based model observers for CT protocol optimization: comparative performance evaluation by means of phantom CT images," *Journal of Medical Imaging*, vol. 10, 2023, doi: 10.1117/1.JMI.10.S1.S11904.

[11] W. Baccouch, S. Oueslati, B. Solaiman, and S. Labidi, "ScienceDirect ScienceDirect performance for for automatic automatic A comparative comparative study study of of CNN CNN and and U-Net U-Net performance segmentation of of medical medical images : images : application application to to cardiac cardiac MRI ," *Procedia Computer Science*, vol. 219, no. 2022, pp. 1089–1096, 2023, doi: 10.1016/j.procs.2023.01.388.

[12] L. I. Kesuma, "ELREI: Ensemble Learning of ResNet, EfficientNet, and Inception-v3 for Lung Disease Classification based on Chest X-Ray Image," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 5, pp. 149–161, 2023, doi: 10.22266/ijies2023.1031.14.

[13] A. Qayoom, J. Xie, and H. Ali, "Polyp segmentation in medical imaging: challenges, approaches and future directions," *Artificial Intelligence Review*, vol. 58, no. 6, p. 169, 2025, doi: 10.1007/s10462-025-11173-2.

[14] J. Lee *et al.*, "Unsupervised machine learning for identifying important visual features through bag-of-words using histopathology data from chronic kidney disease," *Scientific Reports*, vol. 12, no. 1, pp. 1–13, 2022, doi: 10.1038/s41598-022-08974-8.

[15] Y. Wang, F. Sibaii, K. Lee, M. J. Gill, and J. L. Hatch, "CluSA: Clustering-based Spatial Analysis framework through Graph Neural Network for Chronic Kidney Disease Prediction using Histopathology Images," *medRxiv*, vol. 1, no. 165, pp. 1–13, 2021.

[16] A. E. K. Gunawan, "Stock Price Movement Classification Using Ensembled Model of Long Short-Term Memory (LSTM) and Random Forest (RF)," *International Journal on Informatics Visualization*, vol. 7, no. 4, pp. 2255–2262, 2023, doi: 10.30630/joiv.7.4.1640.

[17] S. Liu, "New onset delirium prediction using machine learning and long short-Term memory (LSTM) in electronic health record," *Journal of the American Medical Informatics Association*, vol. 30, no. 1, pp. 120–131, 2023, doi: 10.1093/jamia/ocac210.

[18] B. Sabzipour, "Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment," *Journal of Hydrology*, vol. 627, 2023, doi: 10.1016/j.jhydrol.2023.130380.

[19] B. I. Recognition, "Rethinking Breast Cancer Diagnosis through Deep Learning Based Image Recognition," 2023.

[20] W. H. Rafi, M. D. Sulistiyo, S. Hadiyoso, and U. N. Wisesty, "Polyp Identification from a Colonoscopy Image Using Semantic Segmentation Approach," vol. 5, no. 2, pp. 423–431, 2023, doi: 10.47065/bits.v5i2.4083.

[21] A. Saood and I. Hatem, "COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet," *BMC*

*Medical Imaging*, vol. 21, no. 1, pp. 1–10, 2021, doi: 10.1186/s12880-020-00529-5.

[22] F. Turk and M. Kılıçaslan, "Lung image segmentation with improved U-Net, V-Net and Seg-Net techniques," *PeerJ Computer Science*, vol. 11, pp. 1–21, 2025, doi: 10.7717/PEERJ-CS.2700.

[23] I. Ben Ahmed, W. Ouarda, C. Ben Amar, and khouloud Boukadi, "DEES-breast: deep end-to-end system for an early breast cancer classification," *Evolving Systems*, vol. 15, no. 5, pp. 1845–1863, 2024, doi: 10.1007/s12530-024-09582-9.

[24] F. E. Alazemi, "An Efficient Model for Lungs Nodule Classification Using Supervised Learning Technique," *Journal of Healthcare Engineering*, vol. 2023, 2023, doi: 10.1155/2023/8262741.

[25] J. O. H. Engineering, "Retracted: An Efficient Model for Lungs Nodule Classification Using Supervised Learning Technique," *Journal of healthcare engineering*, vol. 2023, p. 9767402, 2023, doi: 10.1155/2023/9767402.

[26] A. Kumar, "An XNOR-ResNet and spatial pyramid pooling-based YOLO v3-tiny algorithm for Monkeypox and similar skin disease detection," *Imaging Science Journal*, vol. 71, no. 1, pp. 50–65, 2023, doi: 10.1080/13682199.2023.2175423.

[27] S. Buragadda, "HCUGAN: Hybrid Cyclic UNET GAN for Generating Augmented Synthetic Images of Chest X-Ray Images for Multi Classification of Lung Diseases," *International Journal of Engineering Trends and Technology*, vol. 70, no. 2, pp. 229–238, 2022, doi: 10.14445/22315381/IJETT-V70I2P227.

[28] M. Jannat *et al.*, "Lung Segmentation with Lightweight Convolutional Attention Residual U-Net," *Diagnostics*, vol. 15, no. 7, 2025, doi: 10.3390/diagnostics15070854.

[29] E. Chukwujindu, K. Faiz, A. De Sequeira, S. Chidom, and H. Faiz, "Improving medical image segmentation with SAM2: analyzing the impact of object characteristics and finetuning on multi-planar datasets.," *European Journal of Radiology Artificial Intelligence*, vol. 3, p. 100034, 2025, doi: https://doi.org/10.1016/j.ejrai.2025.100034.



Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri