SegNet-Based Framework For Robust Lung Tumor Segmentation From PET/CT Scans

Yessine Amri[1*], Amine Ben Slama[2], and Zouhair Mbarki[3]

[1] *Biochemistry Laboratory, Bechir Hamza Children's Hospital, Tunis, 1007, Tunisia*
[2] *Research Laboratory of Biophysics and Medical Technologies LRBTM (LR13ES07), Higher Institute of Medical Technologies of Tunis (ISTMT), University of Tunis El Manar, Tunis,1006, Tunisia*
[3] *RIFTSI Research Laboratory, ENSIT, University of Tunis, Tunis, 1008, Tunisia*
[*] *Corresponding author. E-mail: amri.yessine@yahoo.com*

Received: Jul. 25, 2025; Accepted: Oct. 12, 2025

Lung cancer remains one of the leading causes of cancer-related mortality worldwide, often due to late-stage diagnosis and the complexity of tumor localization in thoracic imaging. Accurate and automated segmentation of lung tumors from PET/CT images is essential for early diagnosis, treatment planning, and outcome monitoring. Manual segmentation is time-consuming and prone to observer variability, underscoring the need for reliable deep learning-based solutions. This study proposes an automated lung tumor segmentation framework using the SegNet architecture, a deep encoder-decoder convolutional neural network. A dataset of 1,425PET/CT images, manually annotated by expert radiologists, was utilized. Data augmentation techniques were applied to improve generalization. SegNet was trained to perform pixel-wise binary classification, and its performance was benchmarked against the widely used U-Net model. Evaluation metrics included Accuracy, Recall, Dice coefficient, and Intersection over Union (IoU).The proposed SegNet model achieved strong segmentation performance across independent experiments. Average results were: Accuracy of $92.24\% \pm 1.42$, Recall of $94.02\% \pm 1.287$, Dice coefficient of $93.47\% \pm 1.4$, and IoU of $93.03\% \pm 1.2$. Compared to U-Net (Dice: $92.18\% \pm 1.081$, IoU: $91.70\% \pm 1.287$), SegNet demonstrated improved spatial boundary accuracy, particularly for tumors located near complex anatomical structures. Statistical tests confirmed the significance of the performance difference ($p < 0.05$). The SegNet-based model provides accurate and robust segmentation of lung tumors in PET/CT images, outperforming U-Net under the same conditions. Its use of max-pooling indices enhances spatial precision, making it well-suited for clinical applications. Future work will explore 3D extensions, multi-class segmentation, and multi-center validation to enhance its applicability in real-world diagnostic workflows.

Keywords: Lung tumor, PET/CT, Deep learning, SegNet, Medical image segmentation, U-Net

http://dx.doi.org/10.6180/jase.202607_30.008

1. Introduction

Lung cancer remains one of the most prevalent and deadly malignancies worldwide, with an alarming rise in incidence and mortality. According to the Global Cancer Observatory, it accounts for approximately 1.8 million deaths annually, making early detection and precise localization of tumors critical for improving patient prognosis and survival rates [1, 2]. The low five-year survival rate-often below 20%-is largely due to delayed diagnosis and the inherent complexity in distinguishing lung tumors from surrounding thoracic structures in imaging modalities [3]. Effective segmentation of lung tumors is thus a crucial step in enabling timely therapeutic intervention and accurate

radiation therapy planning [4].

Positron emission tomography (PET) and computed tomography (CT) are routinely used in combination (PET/CT) to assess lung lesions. PET provides functional information, capturing metabolic activity through radiotracer uptake, while CT contributes anatomical detail with high spatial resolution [5, 6]. Although PET imaging facilitates the detection of hypermetabolic cancerous tissues, it suffers from limitations such as poor spatial resolution and motion artifacts. CT images, on the other hand, offer enhanced structural resolution but often fail to clearly distinguish tumors when adjacent tissues have similar intensities. This complementary nature of PET and CT underscores the importance of leveraging both modalities for improved diagnostic accuracy and tumor segmentation.

In clinical practice, PET/CT is widely used for staging, treatment planning, and monitoring response to therapy in lung cancer patients. However, differences in scanner models, acquisition protocols, and reconstruction algorithms may lead to heterogeneous image quality. These variations introduce challenges for automated segmentation, as changes in spatial resolution, noise levels, and tracer uptake can affect tumor delineation accuracy and reproducibility. Addressing these challenges motivates the development of robust algorithms capable of generalizing across diverse clinical scenarios [7–10]. Consequently, deep learning methods, particularly convolutional neural networks (CNNs), have emerged as powerful tools for medical image segmentation.

Among CNN-based architectures, encoder-decoder networks such as U-Net and SegNet have gained considerable attention for their ability to perform pixel-wise classification with high precision [11–13]. U-Net utilizes skip connections between corresponding encoder and decoder layers to preserve spatial information and enhance fine boundary detection. SegNet, derived from the VGG16 model, retains max-pooling indices from the encoder to guide the decoder in upsampling, which enhances edge localization and spatial coherence [14]. Both architectures have been successfully applied in various medical imaging tasks, including brain, prostate, and lung tumor segmentation.

However, challenges remain. Variability in tumor shape, size, and position across patients, as well as imaging artifacts and heterogeneity in datasets, necessitate robust and generalizable models. Data scarcity, particularly in annotated PET/CT datasets, further complicates training deep networks from scratch. To address these challenges, data augmentation techniques and optimized hyperparameter tuning have been widely employed [15, 16]. The goal is to improve generalization, reduce overfitting, and enable the model to learn invariant features across variations in orientation, intensity, and morphology.

This study presents a deep learning-based approach using the SegNet architecture for automated segmentation of lung tumors in PET/CT images. Our contributions include (1) the development of a fully automated segmentation pipeline tailored for hybrid PET/CT data, (2) performance evaluation across a large, expert-annotated dataset consisting of 1,425 images, and (3) comparative analysis with the U-Net model to assess the robustness and accuracy of the proposed approach. We demonstrate that SegNet achieves high segmentation quality as quantified by standard metrics including Accuracy, Recall, Dice Coefficient, and Intersection over Union (IoU), and offers significant advantages for clinical deployment in lung cancer diagnosis and treatment planning [17–19].

The remainder of this paper is organized as follows. Section 2 provides a detailed description of the dataset, data augmentation strategies, and the preprocessing pipeline. This section also outlines the proposed SegNet architecture, the segmentation methodology, training procedures, and performance evaluation metrics. Section 3 presents the experimental results, including both quantitative and qualitative assessments, as well as comparative analysis with the U-Net model. In Section 4, we offer a critical discussion of the findings, contextualize them within the existing literature, and address potential limitations. Finally, Section 6 concludes the study and highlights directions for future research and clinical application.

2. Materials and methods

2.1. Dataset and Image Acquisition

The dataset utilized in this study consists of PET/CT images acquired from the Cancer Imaging Archive [20]. A total of 956 PET/CT image slices were initially collected, each annotated with a ground truth segmentation mask prepared and validated by two senior radiologists (figure 1). These annotations served as the reference standard for evaluating the accuracy of the segmentation models. The PET/CT images had an in-plane resolution of 512×512 pixels with voxel sizes ranging from $0.97 - 1.25$ mm. Slice thickness varied between $3 - 5$ mm depending on acquisition protocol. The dataset included a heterogeneous distribution of tumor types, from small nodules (< 1 cm) to large masses (> 5 cm), with lesions located both centrally and peripherally. This diversity was essential to enhance the model's generalization ability across different anatomical contexts.

To expand the dataset and improve the learning capacity of the deep neural networks, a data augmentation strategy

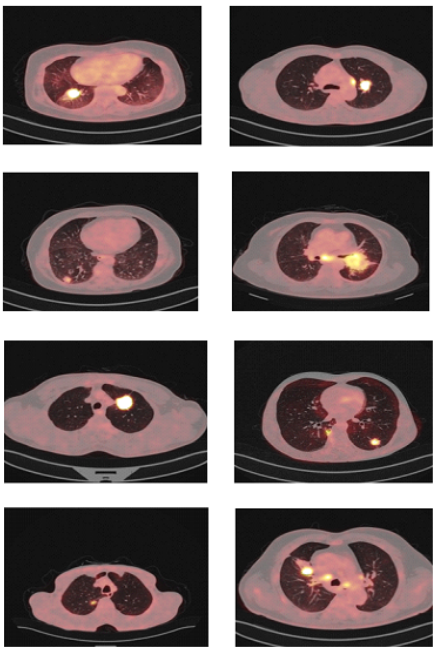

Fig. 1. Example of the used dataset.

was applied, increasing the total number of annotated images to 1,425. Each image reflects a cross-sectional view of the thorax, containing one or more lung lesions of varying size, shape, and intensity.

The dataset was randomly partitioned into three subsets to support supervised training, validation, and testing. The training set consisted of 70% of the total images, while the validation and test sets each contained 15%. Patient ages ranged from 28 to 74 years, with a median age of 52. All included patients had confirmed diagnoses of lung cancer and underwent PET/CT imaging for diagnostic or treatment planning purposes. Each image maintained a standardized resolution, and preprocessing ensured consistent image dimensions and intensity normalization across the dataset.

2.2. Data Augmentation Strategy

Given the relatively modest size of the initial dataset and the high variability of tumor appearances, data augmentation was employed to enhance the diversity of the training samples and reduce the risk of overfitting. Augmentation techniques included random rotations ($\pm 15^\circ$), horizontal/vertical flips (50% probability), translations (up to 10% of image size), and scaling ($0.9 - 1.1$ range). [21]. These transformations preserved the anatomical correctness of the lung structures while introducing sufficient variability for robust model training.

The augmentation was performed dynamically during training, meaning that different augmented versions of each image were generated at each training epoch. This process enabled the model to encounter a wide variety of tumor presentations and contexts, thereby improving its ability to generalize to unseen data. Only linear geometric transformations were used to ensure the augmented images remained anatomically plausible and clinically interpretable. Through this strategy, the effective size of the training set increased several fold, providing the model with an enriched learning environment.

2.3. Image Preprocessing and Tumor Localization

The first step in the segmentation pipeline involved isolating the lung region from surrounding thoracic structures to reduce noise and focus the analysis. Given the similarity in intensity values between tumors and adjacent tissues (e.g., mediastinum, diaphragm, or chest wall), this localization step was critical to minimize false detections. The original CT images were processed using histogram equalization to normalize intensity distributions and improve contrast between lung parenchyma and non-lung tissues [22].

Morphological operations, such as dilation and erosion, were used to refine the lung mask and eliminate spurious regions. The centroid of the largest connected component in the lung mask was computed to serve as a seed for further processing. Once the lung fields were segmented, the region of interest (ROI) was cropped and fed into the deep learning segmentation models for tumor delineation. This preliminary localization reduced computational complexity and enhanced the accuracy of the subsequent segmentation step.

2.4. SegNet Architecture for Tumor Segmentation

The core of the proposed method is a convolutional neural network based on the SegNet architecture, which is specifically designed for semantic segmentation tasks [23, 24]. SegNet is an encoder-decoder architecture that processes the input image at the pixel level, allowing for precise delineation of complex structures such as tumors in medical images. The encoder in SegNet mirrors the architecture of the VGG16 network, consisting of a series of convolutional layers followed by batch normalization and rectified linear unit (ReLU) activation functions. After every few convolutional blocks, a max-pooling operation is applied to reduce spatial resolution and extract abstract feature representations (figure 2).

A unique feature of SegNet lies in its decoder structure.

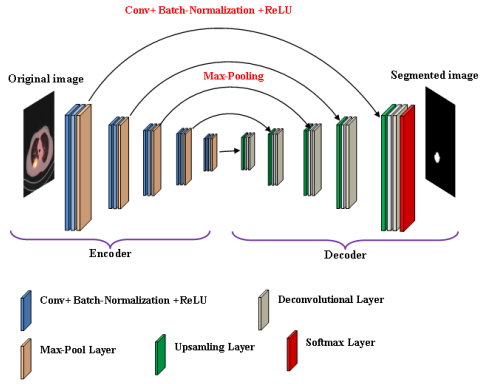

Fig. 2. Seg-Net model.

Unlike traditional architectures that use transposed convolutions or interpolation alone, SegNet decoders rely on the pooling indices obtained during the encoder's max-pooling operations. These indices are stored and later reused during the upsampling phase to restore the spatial dimensions of the feature maps. This strategy enables more accurate localization of features and reduces ambiguity in boundary reconstruction-an essential requirement for medical image segmentation where precision is critical.

The decoder performs upsampling by placing the maximal activation values at the locations specified by the encoder's pooling indices. This is followed by convolutional and batch normalization layers, progressively reconstructing the spatial structure of the image. The final layer of SegNet includes a sigmoid activation function that maps the feature maps to a binary segmentation mask, classifying each pixel as tumor or background. The architecture ensures that spatial detail is preserved during upsampling, which is particularly important when segmenting small or irregularly shaped lesions.

In this study, the SegNet model was modified to accept PET/CT inputs and trained to output binary segmentation masks matching the expert annotations. Each input image and its corresponding output mask had identical dimensions to facilitate direct pixel-wise comparison during training and evaluation.

2.5. Model Training and Hyperparameter Configuration

To train the SegNet model, we used a supervised learning approach with binary cross-entropy as the loss function, measuring the pixel-wise discrepancy between predicted and true segmentation masks. A linear learning rate decay was applied, starting at 0.0005 and reducing to 0.000062 over 500 epochs. Early stopping was triggered if validation loss did not improve for 20 consecutive epochs. The best-performing model (lowest validation loss) was saved and later used for testing [25].

A batch size of 4 was selected to balance memory usage and training stability. Weight initialization was conducted using a Gaussian distribution (mean 0 , standard deviation 0.01), and a learning rate decay schedule was implemented to gradually reduce the learning rate during training. Dropout layers were inserted after the deeper convolutional blocks to improve generalization, with a dropout rate of 0.3 for SegNet (Table 1 and 2).

For comparative purposes, a U-Net model was also trained using the same dataset and preprocessing pipeline. U-Net employs skip connections to transfer spatial information directly from the encoder to the decoder, enhancing performance in small lesion detection. The same hyperparameter settings were used to ensure a fair comparison between the two architectures [26].

2.6. Evaluation Metrics and Statistical Analysis

The performance of the segmentation models was evaluated using four standard quantitative metrics widely used in medical image segmentation: Accuracy, Recall, Dice Similarity Coefficient (DSC), and Intersection over Union (IoU) [27]. Accuracy measures the proportion of correctly classified pixels, while Recall (sensitivity) quantifies the ability to correctly identify tumor pixels. The Dice coefficient assesses the overlap between predicted and ground truth masks, and IoU evaluates the ratio of intersection to union between the predicted and true segmentation regions.

These metrics were calculated for each test image, and mean values were reported across multiple experimental runs. To assess statistical significance, a one-sided Mann-Whitney U test and Student's t-test were applied to compare the predicted segmentations with ground truth masks. A p-value threshold of 0.05 was used to determine statistical significance. All evaluations were performed using the same test set to ensure consistency in model comparison.

3. Results and discussion

This section presents a comprehensive evaluation of the proposed SegNet-based approach for lung tumor segmentation using PET/CT images. We report both quantitative and qualitative results, compare SegNet with the well-established U-Net architecture, and assess statistical significance through rigorous analysis. Each subsection provides insights into the model's performance in terms of segmentation accuracy, learning behavior, visual segmentation

Table 1. Optimum results according to the used hyper-parameters via Seg-Net and U-Net models

Method	Data augmentation	Learning rate	Momentum rate	Batch size	Weight initialization	Training methods	Learning rate decay	Dropout rates
Seg-Net	Linear	0.0005	0.99	4	Normal	SGD	Linear (0.000062)	0.30
U-Net	Linear	0.0005	0.99	4	Normal	SGD	Linear (0.000062)	0.10

Table 2. Tested hyper-parameters

Hyper-parameters	Data		Learning rate				Batch size					Momentum rate			Weight			Training methods			
Tests	L	nL	0.5	0.05	0.005	0.0005	2	4	8	16	32	0.8	0.9	0.99	N	U	GU	RMS	A	Ag	SGD

*Linear "L", No-linear "nL", Normal "N", uniform "U", Glorot uniform "GU", RMSprop "RMS", Adam "A", Adagrad "A", Stochastic gradient descent "SGD"

quality, and comparative assessment.

3.1. Quantitative Evaluation of SegNet Performance

To evaluate the segmentation performance of the proposed SegNet architecture, we conducted four independent experiments on different test splits of the dataset. Each experiment involved training the model from scratch using the training and validation sets, and testing it on a separate subset of 214 PET/CT images. The evaluation metrics included Accuracy, Recall, Dice coefficient, and Intersection over Union (IoU), which are widely used to assess medical image segmentation quality.

Experiments showed that including augmentation improved the Dice coefficient from 91.2% (without augmentation) to $93.47\% \pm 1.4(95\%CI : 92.788\%, 94.152\%)$ (with augmentation), confirming its role in enhancing generalization. On the other hand, the average across all trials was $92.24\% \pm 1.42$ (95% CI: 91.55%, 92.92%) for accuracy, $94.02\% \pm 1.287$ (95% CI: 93.405%, 94.641%) for recall, and $93.03\% \pm 1.2$ (95% CI: 92.412%, 93.648%) for IoU, and $93.2\% \pm 0.971$ (95% CI: 92.776%, 93.708%) for F1 score. These results demonstrate the robustness and high precision of the SegNet model in segmenting lung tumors across various patient cases and tumor morphologies.

3.2. Learning Behavior and Loss Convergence

During the training process, the SegNet model exhibited stable convergence behavior. The binary cross-entropy loss function was minimized progressively over 500 epochs. The final loss value for SegNet was approximately 0.048 , indicating a strong agreement between predicted segmentation masks and ground truth labels. Although the U-Net model reached a slightly lower loss of 0.032 , this did not translate into superior segmentation quality, as discussed in the following sections.

This observation highlights an important point: a lower training loss does not necessarily imply better generalization or superior spatial accuracy in the segmentation outputs, especially in complex clinical data such as PET/CT lung imaging.

3.3. Qualitative Evaluation and Visual Segmentation Accuracy

In addition to the numerical metrics, we performed a qualitative evaluation of the segmentation results. Representative samples of segmentation outputs produced by both SegNet and U-Net are presented in Figure 3. For each sample, the original PET/CT image, the ground truth mask, and the segmentation results of both models are shown for visual comparison.

Visual inspection revealed that SegNet consistently produced smoother and more anatomically accurate segmentation contours [28]. The use of pooling indices during the decoder upsampling process helped preserve spatial information, allowing the model to delineate even small and irregular tumor boundaries. This was particularly evident in cases where tumors were located near complex anatomical structures, such as the diaphragm or chest wall.

In contrast, U-Net occasionally produced over-segmented or under-segmented regions, especially in cases where tumor contrast was low or when the tumor was adjacent to non-lung tissues. These limitations suggest that while U-Net is powerful, SegNet's pooling-based upsampling provides a notable advantage in preserving shape consistency and structural detail.

SegNet achieved consistent delineation for both small nodules and large heterogeneous masses. Its advantage was particularly notable for peripheral lesions near the

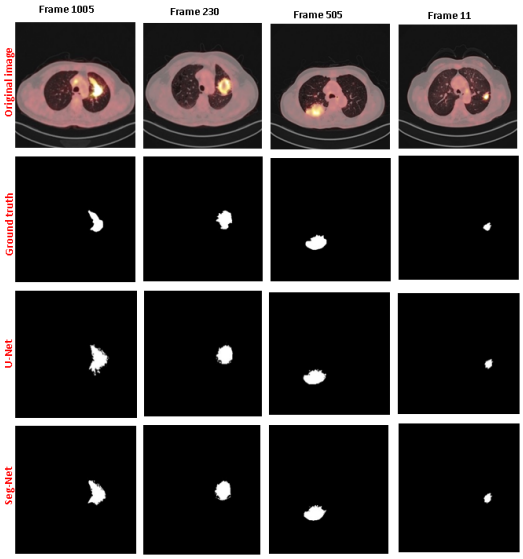

Fig. 3. Objective segmentation performance achieved with Seg-Net and U-Net architectures.

chest wall, where U-Net often failed. From a clinical perspective, accurate segmentation of small or irregular tumors is essential for radiotherapy planning, ensuring precise target definition and sparing healthy tissue.

3.4. Comparative Analysis with U-Net Architecture

To benchmark the performance of the proposed SegNet model, a direct comparison was made with the U-Net architecture under identical preprocessing, training, and evaluation conditions. The average Dice coefficient for U-Net was 92.18% $\pm$ 1.081 (95% CI: 91.671%, 92.709%), and its IoU averaged 91.70% $\pm$ 1.287 (95% CI: 91.082%, 92.318%). Although these scores are competitive, they remain slightly lower than those achieved by SegNet (Dice: 93.47% $\pm$ 1.4(95% CI: 92.788%, 94.152%), IoU: 93.03% $\pm$ 1.2(95% CI: 92.412%, 93.648%)).

Figure 4 presents the comparative results between SegNet and U-Net across all evaluated metrics. SegNet outperformed U-Net in recall and Dice similarity, which are particularly important for medical applications where missing tumor pixels (false negatives) must be minimized [29].

3.5. Statistical Analysis

To assess whether the differences in performance between the two models were statistically significant, we applied the Mann-Whitney U test and the Student's t-test to the Dice and IoU values. For each metric, the null hypothesis

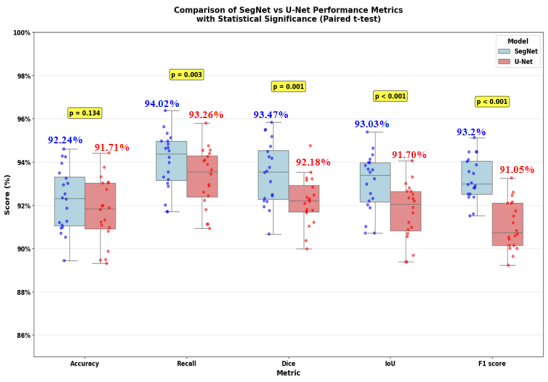

Fig. 4. Comparison of average performance between Seg-Net and U-Net.

assumed no difference between SegNet and U-Net. The SegNet model achieved statistically significant improvements over U-Net, with p -values below the 0.05 threshold in both tests [30].

For Dice coefficient comparison, the Mann-Whitney Ut test yielded a p-value of 0.001, indicating that the SegNet model produced results closer to the ground truth segmentations than U-Net. Similarly, IoU comparisons resulted in a p -value < 0.001. These findings confirm the superior performance of the proposed method and validate its clinical reliability in segmenting lung tumors from PET/CT images.

In summary, the proposed SegNet-based segmentation approach demonstrated consistent and reliable performance across multiple test scenarios. Its use of pooling indices in the decoder enhanced boundary accuracy, enabling superior delineation of lung tumor regions compared to the U-Net baseline. Quantitative results confirmed high levels of recall and overlap, and statistical tests validated the significance of the improvements. These results strongly support the applicability of SegNet for accurate and robust segmentation of lung tumors in clinical PET/CT imaging workflows [31].

The purpose of this study was to develop and evaluate an automated segmentation method for lung tumors in PET/CT images using a deep convolutional neural network based on the SegNet architecture. Lung cancer remains one of the most aggressive and fatal malignancies worldwide, often diagnosed at advanced stages due to the limitations of early detection strategies. Imaging technologies, particularly PET and CT, play a crucial role in the diagnostic and treatment planning pipeline. However, manual segmentation of tumors remains subjective, time-consuming, and prone to inter-observer variability. Therefore, the develop-

ment of reliable computer-aided segmentation tools is of considerable clinical importance.

In this study, we propose a segmentation framework based on the SegNet architecture, which demonstrates strong performance in accurately delineating lung tumors. The segmentation quality was assessed using standard metrics-Accuracy, Recall, Dice coefficient, and IoU across four distinct experiments. The mean results (Accuracy: 92.24% $\pm$ 1.42, Recall: 94.02% $\pm$ 1.287, Dice: 93.47% $\pm$ 1.4, IoU: 93.03% $\pm$ 1.2) highlight the high precision and robustness of our approach under various conditions. These outcomes are consistent with, and in some cases surpass, those reported in prior lung tumor segmentation studies (table 3).

In the work of Kumar et al. [32], authors developed a U-Net-based computer-aided detection (CAD) system for lung nodule segmentation using CT images from the LIDC-IDRI dataset. After preprocessing to isolate nodules and their masks, they trained the model with various optimizers. Their best performance, achieved using the Adam optimizer, yielded a Dice Similarity Coefficient (DSC) of 82.05% and an IoU of 75.39%, demonstrating U-Net's effectiveness in this domain.

Based on the integration of multimodal imaging, Alshmrani et al. [33] proposed a fusion-based U-Net architecture designed for lung tumor segmentation using PET/CT scans.

To handle the complexity of multimodal data, they explored several fusion strategies within the U-Net framework. Their hyper-dense VGG16 U-Net variant showed the highest performance, with a Dice score of 73%, highlighting the benefits of multimodal fusion for improved segmentation accuracy and early lung cancer detection.

In a related segmentation strategy, Park et al. [34] introduced a two-stage U-Net model tailored for [18F]FDG PET/CT scans. Their approach combines a 3D global U-Net for tumor localization with a 2D regional U-Net for detailed segmentation refinement on selected slices. Evaluated on a large cohort of 887 patients, this two-stage method outperformed conventional one-stage 3D U-Nets, achieving a Dice score of 75% and demonstrating superior delineation of tumor margins.

Finally, Zhou et al. [35] developed a multitask connected U-Net model that incorporates complementary data from both CT and PET modalities. Their architecture integrates PET guidance and a tumor area detection module to better localize lesions. Tested on four datasets, the model achieved an average Dice coefficient of 56%, outperforming several other CNN models.

Beyond the numerical improvements, one of the key strengths of our approach lies in its consistent performance across a wide variety of cases. PET/CT images often include tumors of varying sizes, locations, and uptake patterns. The SegNet model maintained high performance even in challenging scenarios, such as lesions adjacent to the chest wall or mediastinum, where contrast is low, and spatial boundaries are less well-defined. This suggests that SegNet's architecture is well-suited to handling the heterogeneous nature of clinical lung imaging data.

It is also worth noting that our comparative experiments with the U-Net model highlighted interesting architectural differences. While U-Net is renowned for its skip connections and has become a standard in medical image segmentation [36], our results showed that SegNet outperformed U-Net in this specific context. The mean Dice coefficient for U-Net was 92.18% $\pm$ 1.081, and the IoU was 91.70% $\pm$ 1.287, both slightly lower than the values obtained using SegNet. This supports the notion that architectural modifications-such as reusing max-pooling indices instead of relying solely on feature concatenation-can lead to significant gains in segmentation accuracy, particularly in PET/CT data where spatial coherence and edge information are crucial.

A possible explanation for SegNet's superior performance lies in its ability to more effectively encode spatial relationships within the image. The reuse of encoder pooling indices allows the decoder to focus on meaningful regions, improving the delineation of tumor boundaries. In contrast, U-Net's concatenation-based skip connections may result in information redundancy and occasionally misalignment in regions of high anatomical complexity. This finding aligns with other studies that emphasize the benefits of preserving spatial coherence in segmentation tasks involving organ boundaries and lesions [36–39].

Despite the encouraging results, several limitations must be acknowledged. First, while the dataset was augmented and diversified through geometric transformations, it was still based on a single imaging center and scanner protocol. This may limit the generalizability of the model to datasets from other institutions with differing acquisition parameters. A multi-institutional study involving varied scanner types and patient demographics would provide a more comprehensive evaluation of the model's robustness.

Second, the study focused solely on binary segmentation (tumor vs. background). In clinical practice, it may be necessary to segment multiple tissue types or tumor subregions (e.g., necrotic core vs. viable tumor), which would require multi-class segmentation strategies. Future work could expand the model's capabilities by integrating multi-class labels or by incorporating additional imaging

Table 3. Summary of some studies reporting some lung tumors segmentation approaches.

Literature	Year	Used techniques	Used Dataset	Dice (%)
Kumar et al. [32]	2021	U-net	CT	82.05
Alshmrani et al. [33]	2023	U-net	CT	73.00
Park et al. [34]	2023	Two-stage U-Net	PET/CT	75.00
Zhou et al. [35]	2024	Connected U-Net	PET/CT	56.00
Proposed method	-	Seg-Net	PET/CT	93.47

modalities such as MRI to improve tissue characterization.

Third, although the SegNet model achieved excellent results, further performance improvements may be possible through architectural enhancements such as attention mechanisms, dilated convolutions, or residual connections. These innovations have shown promise in other medical image segmentation tasks and could be explored in the context of PET/CT lung imaging [40].

Another avenue for improvement lies in the incorporation of temporal or volumetric information. While our model operates on 2D slices, tumors are inherently three-dimensional structures. Extending the approach to 3D segmentation using volumetric CNNs or 2.5D strategies (considering adjacent slices) could enhance contextual understanding and segmentation continuity across slices. However, such approaches would require significantly more computational resources and memory, which should be weighed against the expected performance gains.

Recent progress in PET/CT segmentation has also explored Transformer-based networks, lightweight architectures (e.g., LKAFormer), and hybrid models combining convolutional and attention mechanisms [41–43]. While these models show promise, they often require larger datasets and increased computational resources. In contrast, SegNet offers a balance of accuracy, interpretability, and efficiency, which motivated our choice for this study.

Finally, clinical validation is essential before this approach can be fully integrated into diagnostic workflows. While we performed a thorough evaluation using annotated ground truth from expert radiologists, a prospective study involving multiple radiologists and assessment of the impact on clinical decision-making would be needed to confirm its utility in practice.

Despite these limitations, the proposed SegNet-based model offers a valuable contribution to the field of medical image analysis. Its robust performance, relatively low complexity, and strong generalization potential make it a promising candidate for integration into automated diagnostic systems. Furthermore, its superior results compared to baseline models and previously published methods underscore its effectiveness in addressing the challenges inherent in lung tumor segmentation.

4. Conclusion

This study provides strong evidence supporting the application of SegNet for the segmentation of lung tumors in PET/CT images. Through a combination of architectural efficiency, rigorous evaluation, and high segmentation accuracy, the proposed method stands as a reliable tool for advancing precision in computer-aided diagnosis and radiotherapy planning. Future studies will focus on expanding the dataset, incorporating volumetric segmentation, and validating the model in real-world clinical settings.

Acknowledgements

We are deeply grateful to all those who contributed to the success of this research project.

Authorship contribution statement

Yessine Amri and Amine Ben Slama contributed equally to this work.

References

[1] B. Le Goff, J.-M. Berthelot, and Y. Maugars, (2015) *"Échographie du thorax antérieur [Ultrasound of the anterior thorax]"* **Revue du Rhumatisme Monographies** *82*(2): 83–87. DOI: 10.1016/j.monrhu.2015.02.005.

[2] C. Jani, D. C. Marshall, H. Singh, R. Goodall, J. Shalhoub, O. Al Omari, and C. C. Thomson, (2021) *"Lung cancer mortality in Europe and the USA between 2000 and 2017: An observational analysis"* **ERJ Open Research** *7*(4): DOI: 10.1183/23120541.00311-2021.

[3] P. B. Bach, J. R. Jett, U. Pastorino, M. S. Tockman, S. J. Swensen, and C. B. Begg, (2007) *"Computed tomography screening and lung cancer outcomes"* **JAMA** *297*(9): 953–961. DOI: 10.1001/jama.297.9.953.

[4] E. B. Jemia, H. Kamoun, S. Louhaichi, H. Smadhi, D. Greb, I. Akrout, and M. L. Megdiche, (2017) *"Variation du profil histologique du cancer du poumon durant les 25 dernières années dans un service de pneumologie à Tunis [Variation of the histological profile of lung cancer during the last 25 years in a pulmonology department in*

Tunis]" **Revue des Maladies Respiratoires *34***: A79. DOI: 10.1016/j.rmr.2016.10.168.

[5] Y. Men, Z. Hui, J. Liang, Q. Feng, D. Chen, H. Zhang, and L. Wang, (2016) *"Further understanding of an uncommon disease of combined small cell lung cancer: Clinical features and prognostic factors of 114 cases"* **Chinese Journal of Cancer Research *28***(5): 486. DOI: 10.21147/j.issn.1000-9604.2016.05.03.

[6] B. Padovani, J. Mouroux, L. Seksik, S. Chanalet, J. Sedat, C. Rotomondo, and J. J. Serres, (1993) *"Chest wall invasion by bronchogenic carcinoma: Evaluation with MR imaging"* **Radiology *187***(1): 33–38. DOI: 10.1148/radiology.187.1.8451432.

[7] N. Suzuki, T. Saitoh, and S. Kitamura, (1993) *"Tumor invasion of the chest wall in lung cancer: Diagnosis with US"* **Radiology *187***(1): 39–42. DOI: 10.1148/radiology.187.1.8451433.

[8] S. J. Swensen, R. W. Viggiano, D. E. Midthun, N. L. Müller, A. Sherrick, K. Yamashita, and A. L. Weaver, (2000) *"Lung nodule enhancement at CT: Multicenter study"* **Radiology *214***(1): 73–80. DOI: 10.1148/radiology.214.1.r00ja1473.

[9] D. F. Yankelevitz, A. P. Reeves, W. J. Kostis, B. Zhao, and C. I. Henschke, (2000) *"Small pulmonary nodules: Volumetrically determined growth rates based on CT evaluation"* **Radiology *217***(1): 251–256. DOI: 10.1148/radiology.217.1.r00oc33251.

[10] G. Ferretti, A. Jankowski, A. Calizzano, D. Moro-Sibilot, and J. P. Vuillez, (2008) *"Imagerie radiologique et TEP Scanner dans les cancers du poumon [Radiological imaging and PET scans in lung cancer]"* **Journal de Radiologie *89***(3): 387–402. DOI: 10.1016/S0221-0363(08)89016-6.

[11] N. A. Dewan, S. D. Reeb, N. C. Gupta, L. S. Gobar, and W. J. Scott, (1995) *"PET-FDG imaging and transthoracic needle lung aspiration biopsy in evaluation of pulmonary lesions: A comparative risk-benefit analysis"* **Chest *108***(2): 441–446. DOI: 10.1378/chest.108.2.441.

[12] J. E. Bibault, A. Oudoux, J. Durand-Labrunie, X. Mirabel, É. Lartigau, and H. Kolesnikov-Gauthier, (2015) *"TEP et radiothérapie stéréotaxique pulmonaire : rôles dans la préparation du traitement et le suivi de la maladie [Positron emission tomography and stereotactic body radiation therapy for lung cancer: From treatment planning to response evaluation]"* **Cancer/Radiothérapie *19***(8): 790–794. DOI: 10.1016/j.canrad.2015.05.027.

[13] S. S. Gambhir, J. Czernin, J. Schwimmer, D. H. Silverman, R. E. Coleman, and M. E. Phelps, (2001) *"A tabulated summary of the FDG PET literature"* **Journal of Nuclear Medicine *42***(5 suppl): 1S–93S.

[14] M. K. Gould, C. C. Maclean, W. G. Kuschner, C. E. Rydzak, and D. K. Owens, (2001) *"Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: A meta-analysis"* **JAMA *285***(7): 914–924. DOI: 10.1001/jama.285.7.914.

[15] C. Perrotin, P. Lemeunier, D. Grahek, T. Molina, A. Petino, M. Alifano, F. Bellenot, P. Magdeleinat, J. N. Talbot, and J. F. Regnard, (2005) *"Résultats de la TEP [18F]-FDG dans la stadification préopératoire des tumeurs pulmonaires [Results of FDG-PET scanning in the preoperative staging of broncho-pulmonary tumors]"* **Revue des Maladies Respiratoires *22***(4): 579–585. DOI: 10.1016/s0761-8425(05)85610-4.

[16] J. F. Vansteenkiste, S. G. Stroobants, P. J. Dupont, P. R. De Leyn, E. K. Verbeken, G. J. Deneffe, and M. G. Demedts, (1999) *"Prognostic importance of the standardized uptake value on 18F-fluoro-2-deoxy-glucose–positron emission tomography scan in non–small-cell lung cancer: An analysis of 125 cases"* **Journal of Clinical Oncology *17***(10): 3201–3206.

[17] M. Vallières, C. R. Freeman, S. R. Skamene, and I. El Naqa, (2015) *"A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities"* **Physics in Medicine & Biology *60***(14): 5471. DOI: 10.1088/0031-9155/60/14/5471.

[18] Z. Li, J. Zhang, T. Tan, X. Teng, X. Sun, H. Zhao, and G. Litjens, (2020) *"Deep learning methods for lung cancer segmentation in whole-slide histopathology images—the acdc@ lunghp challenge 2019"* **IEEE Journal of Biomedical and Health Informatics *25***(2): 429–440. DOI: 10.1109/JBHI.2020.3039741.

[19] H. Hu, Q. Li, Y. Zhao, and Y. Zhang, (2020) *"Parallel deep learning algorithms with hybrid attention mechanism for image segmentation of lung tumors"* **IEEE Transactions on Industrial Informatics *17***(4): 2880–2889. DOI: 10.1109/TII.2020.3022912.

[20] P. Li, S. Wang, T. Li, J. Lu, Y. HuangFu, and D. Wang. *A large-scale CT and PET/CT dataset for lung cancer diagnosis (Lung-PET-CT-Dx) [Data set]*. 2020. DOI: 10.7937/TCIA.2020.NNC2-0461.

[21] M. Wang, X. Zhou, M. Jin, Y. Zhang, L. Liu, and G. Huang, (2024) *"Multiroimix: A data augmentation method for PET/CT multimodal medical images"* **Journal of Medical and Biological Engineering *44***(3): 366–374. DOI: 10.1007/s40846-024-00862-y.

[22] A. Klaengkaew, S. Sutthigran, N. Thammasiri, K. Yuwatanakorn, C. Thanaboonnipat, S. Ponglowhapan, and N. Choisunirachon, (2021) *"The evaluation of non-anesthetic computed tomography for detection of pulmonary parenchyma in feline mammary gland carcinoma: a preliminary study"* **BMC Veterinary Research *17***(1): 237. DOI: 10.1186/s12917-021-02950-6.

[23] X. Zhang, G. Xu, X. Wu, W. Liao, L. Xiao, Y. Jiang, and H. Xing, (2024) *"Fast-SegNet: Fast semantic segmentation network for small objects"* **Multimedia Tools and Applications *83***(34): 81039–81055. DOI: 10.1007/s11042-024-18829-1.

[24] R. Pang, H. Tan, Y. Yang, X. Xu, N. Liu, and P. Zhang. "A Novel SegNet Model for Crack Image Semantic Segmentation in Bridge Inspection". In: *Advances in Knowledge Discovery and Data Mining. PAKDD 2024*. Ed. by D. Yang, X. Xie, V. Tseng, J. Pei, J. Huang, and J. Lin. ***14647***. Lecture Notes in Computer Science. Singapore: Springer, 2024, 26. DOI: 10.1007/978-981-97-2259-4_26.

[25] N. Şahin, N. Alpaslan, and D. Hanbay, (2022) *"Robust optimization of SegNet hyperparameters for skin lesion segmentation"* **Multimedia Tools and Applications *81***(25): 36031–36051. DOI: 10.1007/s11042-021-11032-6.

[26] Y. Zhang, (2025) *"Image denoising based on deep feature fusion and U-Net network"* **Journal of Applied Science and Engineering *28***(10): 2077–2085. DOI: 10.6180/jase.202510_28(10).0020.

[27] H. Chen and Z. Xu, (2021) *"A new end-to-end network model for medical image segmentation"* **Journal of Applied Science and Engineering *24***(2): 207–213. DOI: 10.6180/jase.202104_24(2).0009.

[28] G. Athanasiou, J. L. Arcos, and J. Cerquides, (2023) *"Enhancing medical image segmentation: Ground truth optimization through evaluating uncertainty in expert annotations"* **Mathematics *11***(17): 3771. DOI: 10.3390/math11173771.

[29] S. Alqazzaz, X. Sun, X. Yang, and L. Nokes, (2019) *"Automated brain tumor segmentation on multi-modal MR image using SegNet"* **Computational Visual Media *5***(2): 209–219. DOI: 10.1007/s41095-019-0139-y.

[30] G. Di Leo and F. Sardanelli, (2020) *"Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach"* **European Radiology Experimental *4***(1): 18. DOI: 10.1186/s41747-020-0145-y.

[31] J. M. Rogasch, N. Frost, S. Bluemel, L. Michaels, T. Penzkofer, M. von Laffert, and C. Furth, (2021) *"FDG-PET/CT for pretherapeutic lymph node staging in non-small cell lung cancer: a tailored approach to the ESTS/ESMO guideline workflow"* **Lung Cancer *157***: 66–74. DOI: 10.1016/j.lungcan.2021.05.003.

[32] S. N. Kumar, P. M. Bruntha, S. I. Daniel, J. A. Kirubakar, R. E. Kiruba, S. Sam, and S. I. A. Pandian. "Lung nodule segmentation using U-Net". In: *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*. ***1***. 2021, 420–424. DOI: 10.1109/ICACCS51430.2021.9441977.

[33] G. M. Alshmrani, Q. Ni, R. Jiang, and N. Muhammed, (2023) *"Hyper-dense_lung_seg: Multimodal-fusion-based modified U-Net for lung tumor segmentation using multimodality of CT-PET scans"* **Diagnostics *13***(22): 3481. DOI: 10.3390/diagnostics13223481.

[34] J. Park, S. K. Kang, D. Hwang, H. Choi, S. Ha, J. M. Seo, and J. S. Lee, (2023) *"Automatic lung cancer segmentation in [18F] FDG PET/CT using a two-stage deep learning approach"* **Nuclear Medicine and Molecular Imaging *57***(2): 86–93. DOI: 10.1007/s13139-022-00745-7.

[35] L. Zhou, C. Wu, Y. Chen, and Z. Zhang, (2024) *"Multitask connected U-Net: automatic lung cancer segmentation from CT images using PET knowledge guidance"* **Frontiers in Artificial Intelligence *7***: 1423535. DOI: 10.3389/frai.2024.1423535.

[36] G. Saimassay, M. Begenov, U. Sadyk, R. Baimukashev, A. Maratov, and B. Omarov, (2024) *"Enhanced U-Net architecture for lung segmentation on computed tomography and X-ray images"* **International Journal of Advanced Computer Science and Applications *15***(5): DOI: 10.14569/IJACSA.2024.0150594.

[37] A. Ben Slama, Y. Amri, A. Fnaiech, and H. Sahli, (2025) *"Automated ECG arrhythmia classification using hybrid CNN-SVM architectures"* **Journal of Electronic Science and Technology *23***(3): 100316. DOI: 10.1016/j.jnlest.2025.100316.

[38] Y. Amri, A. Ben Slama, Z. Mbarki, R. Selmi, and H. Trabelsi, (2025) *"Automatic glioma segmentation based on efficient U-Net model using MRI images"*

Intelligence-Based Medicine: 100216. DOI: 10.1016/j.ibmed.2025.100216.

[39] A. Ben Slama, Y. Amri, S. Barbaria, H. B. Rahmouni, and H. Trabelsi, (2025) *"Lung diseases classification using pre-trained based deep learning model and support vector machine"* **Polish Journal of Medical Physics and Engineering *31***(3): 178–194. DOI: 10.2478/pjmpe-2025-0021.

[40] I. Domingues, G. Pereira, P. Martins, H. Duarte, J. Santos, and P. H. Abreu, (2020) *"Using deep learning techniques in medical imaging: a systematic review of applications on CT and PET"* **Artificial Intelligence Review *53***(6): 4093–4160. DOI: 10.1007/s10462-019-09850-2.

[41] S. Yin, H. Li, L. Teng, and et al., (2024) *"Brain CT image classification based on mask R-CNN and attention mechanism"* **Scientific Reports *14***: 29300. DOI: 10.1038/s41598-024-78566-1.

[42] S. Yin, L. Wang, T. Chen, H. Huang, J. Gao, J. Zhang, M. Liu, P. Li, and C. Xu, (2025) *"LKAFormer: A lightweight Kolmogorov-Arnold transformer model for image semantic segmentation"* **ACM Transactions on Intelligent Systems and Technology**:

[43] Q. Shi, S. Yin, K. Wang, L. Teng, and H. Li, (2021) *"Multichannel convolutional neural network-based fuzzy active contour model for medical image segmentation"* **Evolving Systems *13***: 535–549. DOI: 10.1007/s12530-021-09392-3.