Polytechnic Journal

Polytechnic Journal

Volume 13 | Issue 2 Article 1

October 2023

Semantic Lung Segmentation from Chest X-Ray Images Using Seg-Net Deep CNN Model

Dathar Abas Hasan
Duhok Polytechnic University College of Health and Medical Technology-Shekhan Duhok - Iraq,
dathar.hasan@dpu.edu.krd

Umed Hayder Jader
Erbil Polytechnic University Soran Technical College, IT Department Erbil – Iraq

How to Cite This Article

Hasan, Dathar Abas and Jader, Umed Hayder (2023) "Semantic Lung Segmentation from Chest X-Ray Images Using Seg-Net Deep CNN Model," *Polytechnic Journal*: Vol. 13: Iss. 2, Article 1.
DOI: https://doi.org/10.59341/2707-7799.1712

Semantic Lung Segmentation from Chest X-Ray Images Using Seg-Net Deep CNN Model

Abstract

Implementing an accurate image segmentation to extract the lung shape from X-ray images is a vital step in designing a CAD system that diagnoses various types of chest diseases. Lung segmentation is a complex process due to the blurred regions that separate the lung area and the rest of the image. The conventional image segmentation techniques do not meet the ambitions to achieve precise lung segmentation. In this paper, we utilized the Seg-Net semantic segmentation model as a practical approach to distinguish the lung region pixels in X-ray images. The model involves an encoder network that extracts the data from the input images, and a corresponding decoder network that maps the low-resolution encoder feature maps to full input-resolution feature maps for pixel-wise classification. The model is trained and tested using 539 X-ray images from Shenzhen's publicly available dataset. The robust performance of the Seg-Net model is investigated in terms of global accuracy, dice coefficient, and intersection over union. The model achieved global accuracy (97.71%), dice coefficient (96.95%), and Jaccard index (94.08%). The experimental results indicate that the Seg-Net model is an effective tool for performing complicated segmentation tasks and extracting regions of interest such as lung area, eye vessels, lesions, and tumors from medical images.

Keywords

Semantic Segmentation – CNN – Seg-Net – Chest X-ray – Deep Learning

ORIGINAL ARTICLE

Semantic Lung Segmentation from Chest X-ray Images Using Seg-Net Deep CNN Model

Dathar Abas Hasan [a,*], Umed Hayder Jader [b]

[a] Duhok Polytechnic University, College of Health and Medical Technology, Shekhan, Duhok, Iraq
[b] Erbil Polytechnic University, Soran Technical College, IT Department, Erbil, Iraq

Abstract

Implementing an accurate image segmentation to extract the lung shape from X-ray images is a vital step in designing a CAD system that diagnoses various types of chest diseases. Lung segmentation is a complex process due to the blurred regions that separate the lung area and the rest of the image. The conventional image segmentation techniques do not meet the ambitions to achieve precise lung segmentation. In this paper, we utilized the Seg-Net semantic segmentation model as a practical approach to distinguish the lung region pixels in X-ray images. The model involves an encoder network that extracts the data from the input images, and a corresponding decoder network that maps the low-resolution encoder feature maps to full input-resolution feature maps for pixel-wise classification. The model is trained and tested using 539 X-ray images from Shenzhen's publicly available dataset. The robust performance of the Seg-Net model is investigated in terms of global accuracy, dice coefficient, and intersection over union. The model achieved global accuracy (97.71%), dice coefficient (96.95%), and Jaccard index (94.08%). The experimental results indicate that the Seg-Net model is an effective tool for performing complicated segmentation tasks and extracting regions of interest such as lung area, eye vessels, lesions, and tumors from medical images.

Keywords: Semantic segmentation, CNN, Seg-Net, Chest X-ray, Deep learning

1. Introduction

The WHO reports considered lung disorders as the third prominent reason for increased mortality around the world for five years. Lung diseases are mainly caused by virus infection, air pollution, smoking, and genetic considerations [1]. The early detection of the infected lung is an important procedure to reduce the deterioration of the patient's health. Nowadays, most lung diseases are detected and examined by radiologists using CT scans and Chest X-ray (CXR) images [2]. During the examination process, many mistakes may appear because of fatigue and exhaustion, or the inability to determine the size of the affected lung as a result of the blurring of the border between the lung and the rest of the image parts [3]. Therefore, an automated, accurate, and fast lung segmentation system is necessary to support physicians in making appropriate decisions and enhancing the healthcare system [4].

Semantic segmentation is a computer vision task that assigns semantic labels to each pixel in an image. Unlike image classification, which assigns a single label to the entire image, semantic segmentation aims to provide a detailed understanding of the image at the pixel level. It is commonly used in applications such as autonomous driving, medical image analysis, object recognition, and scene understanding [5]. The traditional computer vision techniques for segmentation often rely on low-level features, such as edges or textures, to separate objects in an image. However, these methods often struggle with complex scenes and suffer from limited accuracy. This is where deep learning and convolutional neural networks (CNNs) have made significant advancements in semantic segmentation [6].

Received 7 August 2023; accepted 28 August 2023.
Available online 9 October 2023

* Corresponding author.
E-mail addresses: dathar.hasan@dpu.edu.krd (D.A. Hasan), omid.jader@epu.edu.iq (U.H. Jader).

https://doi.org/10.59341/2707-7799.1712

Deep learning-based approaches for semantic segmentation utilize CNNs to learn hierarchical representations of image data. These networks can capture local and global contexts by using convolutional layers to extract features at different scales. The learned features are combined with upsampling and deconvolutional layers to generate dense pixel-wise predictions [7,8]. Deep learning has revolutionized the field of semantic segmentation by enabling automatic feature learning, end-to-end training, dense pixel-wise predictions, architectural advancements, and the utilization of large datasets [9]. The combination of deep learning techniques and semantic segmentation has significantly advanced the accuracy and applicability of computer vision tasks, leading to breakthroughs in various fields, including autonomous driving, medical imaging, and object recognition [10].

In this paper, we performed a semantic lung segmentation based on the Seg-Net deep CNN model to extract the lung area pixels from the CXR images. Seg-Net's architecture is memory efficient, primarily due to pooling indices in the decoder instead of storing all the feature maps. This feature enables to segmenting of high-resolution images with minimum memory requirements. This paper is organized as follows: in section 2, the most related works have been discussed then describing the structure of the Seg-Net architecture. In section 3, the segmentation results have been presented and discussed. In section 4, concluding remarks are stated.

2. Material and methods

2.1. Related works

Deep learning became the most effective and most used tool for many studies to design CAD systems for early detection and identification of various medical cases. B. Ait Skourt et al., employed the U-net architecture for lung segmentation in CT images. The proposed work can be trained end-to-end even with a limited number of images and surpasses the performance of numerous other methods in this context [10]. Pang T. et al., introduced a segmentation model that leverages radionics by combining hand-crafted characteristics and deep characteristics for interstitial lung disease (ILD) from high-resolution CT images. The ILD Database-MedGIFT is used consists of 128 patients with 108 annotated image series. The final experiment demonstrates a segmentation result with a DSC of 89.42% [11]. P. B. Chanda and S. K. Saekar proposed Adaptive Threshold and Gaussian Filter algorithms for lung segmentation in X-ray images. In addition, Gabor filtering is utilized for noise reduction, the Otsu Thresholding technique is used to enhance the brightness of the lighter zone and intensify the darkness of the darker zone as well and K-Mean and FCM algorithms are used as classification methods [12]. R. Selvan et al. address the challenge of compact lungs through the introduction of variational imputation. They propose a deep learning pipeline built upon variational auto-encoders. Different augmentation methods such as Standard 2, Block Masking, and Diffused Noise are utilized to train the model to handle missing data, data imputation, and record other characteristics like shape information [13]. Dathar H. and Abdulazeez AM, introduced an enhanced version of U-Net with reduced parameters and faster network training. In comparison to the original U-Net, this improved version decreases the number of convolution kernels per layer, resulting in a remarkable 75% reduction in parameters and a significant 70% reduction in network training process time [14]. Zhang X. et al., enhanced the U-Net network to achieve accurate segmentation of various types of lung nodules in CT images. They utilized the Otsu thresholding technique as a segmentation technique and then applied a-hull algorithm, to refine lung outline [6].

W. Liu et al. proposed an enhanced U-Net model that consists of 5 layers on both the encoder and decoder sides. They consumed efficient net-b4 in the encoder, as well as for the decoder part, Dropout layer (to solve overfitting issues), two Residual blocks (to prevent Gradient loss and better information spreading) and a LeakyReLU (as activation method) are utilized [15]. Chavan Mahesh et al. introduced ResUNet++ by mixing (fusing) U-Net and ResNet feature extraction to segment X-ray chest images. Atrous Spatial Pyramidal Pooling (ASPP) layer, Residual Block, Attention Block, Excitation Block, and Squeeze Blocks are utilized to improve the proposed model. In addition, the MC dataset that contains 138 X-ray images has been consumed in the model training and validation process, and compared with ResUNet and UNet architectures [16]. Zhakaw. H. H. and Taban F. M. proposed a U-Net-based model by improving a ground truth mask to partition lung X-ray images from three different datasets which were Covid-Qu, Montgomery & Shenzhen, and the third one created by themselves that contains 1300 images called KURD-Covid dataset. They increased one extra layer to the decoder side to learn the model faster and one layer to the decoder part for more accurate segmentation [9]. Murugappan M. et al. Presented

DeepLabV3 plus 4 existing deep neural networks (ResNet-50, MobileNet-V2, ResNet-18, and Xception). The proposed segmentation system is examined by two scenarios: couple-class (background and lung field) and quadruple-class (background, lung field, ground-glass opacities, and consolidation) [1]. Y. Rao et al. proposed a new threshold mechanism to accurately segment the Ground Glass Opacity (GGO) of the CT lung images for COVID-19 patients. The mechanism involves three operations to modify the threshold according to the image contrast. Besides, the model requires a lesser learning process and hardware specifications and comes up with faster execution compared to other deep learning algorithms [17].

2.2. Methodology

Seg-Net is one of the effective CNN models that is used to achieve semantic segmentation. Therefore, we utilized its robustness to extract the lung pixels from the CXR images. In this section, we focused on the used dataset, the preprocessing operations, the training environments and parameters, and finally the Seg-Net model description.

2.2.1. Dataset

The dataset is prepared by Shenzhen No.3 People's Hospital in Guangdong Medical College. The dataset includes CXR images from outpatient clinics which were captured as part of the daily routine using Philips DR Digital Diagnose systems. The dataset contains 539 CXR images in PNG format, 336 images are manifestations of tuberculosis and the rest are normal cases. The images are 8-bit channel width and multi-dimension size. Therefore, we preprocessed the input images using a 256 × 256 center cropping size to focus on the lung regions and reduce the background area. In addition, data augmentation processes such as reflection, rotation, and scaling have been used to increase the training set images and then increase the model generalization. We used 70% of the images for model training, 20% for testing, and 10% for validating.

2.2.2. Seg-Net Architecture

Seg-Net was proposed in 2015 by V. Badrinarayanan, A. Kendall, and R. Cipolla from the University of Cambridge [18]. It is one of the CNN architectures that is designed specifically for semantic segmentation tasks. Seg-Net focuses on producing dense pixel-wise predictions while being computationally efficient. Fig. 1 shows the Seg-Net architecture which includes four encoder-decoder stacks structure with skip connections.

Each encoder stack consists of several convolutional layers, a ReLU layer, and a 2 × 2 downsampling max-pooling layer to extract the features from the input images. The decoder network upsamples the input features using the trainable kernel via skip connections with the encoder network.

The encoder's first layer is the input layer, it accepts the CXR images with 256 × 256 × 1 dimensions and zero-center normalization. The next layer is the convolutional layer with 64 filters, 3 × 3 × 1 filter dimension, stride 1, and padding 1. These filters are passed over the CXR input image to perform dot product operation between the filter's kernels and image. This process extracts 64 various feature maps from the input images, these feature maps include edges, borders, patterns, …etc. Each convolution layer is followed by the ReLU activation function layer. If the input value is negative then the ReLU function maps it to 0, otherwise, it remains unchanged as shown below (Fig. 2).

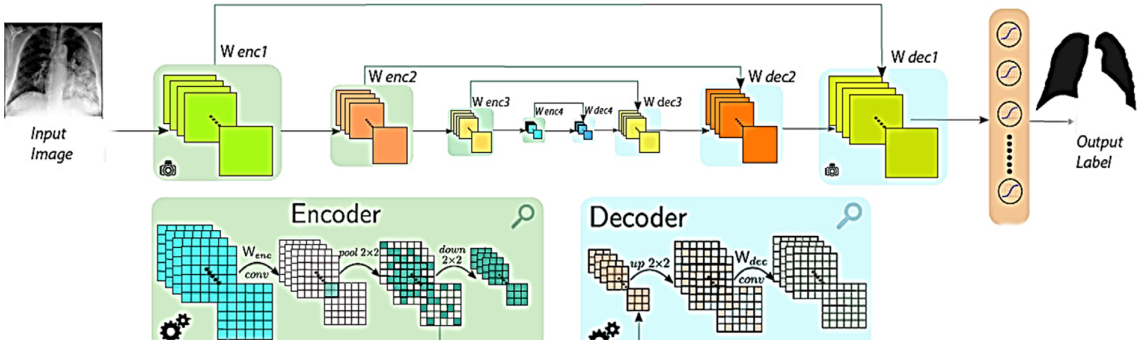

Fig. 1. Seg-Net architecture.

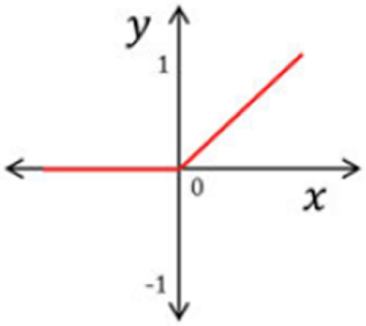

Fig. 2. ReLU activation function.

The convolution and ReLU layers are performed twice. It means that the 64 extracted features from the first stage are passed to a new convolution and ReLU layers to map new 64 output features. The next layer is a 2 × 2 max pooling layer with a stride of 2 and zero padding. The 2 × 2 max pooling window is passed over the 64 extracted features to compute the corresponding value in the output feature map. The size of the output feature map h_o x w_o is calculated using the formulas:

$$h_0 = \left[\frac{h - f + s}{s}\right] \quad \text{Eq.(1)}$$

$$w_0 = \left[\frac{w - f + s}{s}\right] \quad \text{Eq.(2)}$$

Where h and w are the height and width of the input feature map, s is the stride amount in the horizontal and vertical direction, and f is the size of the max pooling window which is 2.

This mechanism is performed at each encoder stack to extract more features and classify each pixel to its corresponding label.

On the decoder side, there are four deconvolution stacks, each stack consists of a max unpooling layer followed by a couple of convolution and batch normalization layers. The main task of the decoder is to reconstruct the segmentation map from the encoded features. It involves producing a compressed representation of the input image to the max un-pooling layers, allowing the network to extract key structural features and relationships within the input. Fig. 3 shows the max-pooling and max unpooling operations.

The convolution layers in the decoder side implement transposed convolution which involves applying a pre-learned convolutional kernel to the reduced size feature maps. In comparison with the convolutional layers on the encoder side, a kernel in the transposed convolution is applied to each location of the output feature map, with a stride that determines the distance between successive kernel positions. Each kernel output is then added to the corresponding location in the output feature map. The transposed convolution broadcasts input elements via the kernel, thereby producing an output that is larger than the input. Fig. 4 shows the transposed convolution operation between the 2 × 2 feature map and the 2 × 2 kernel.

The max up-sampling and transposed convolutional layers gradually restore the spatial dimensions to generate the image labels and reduce

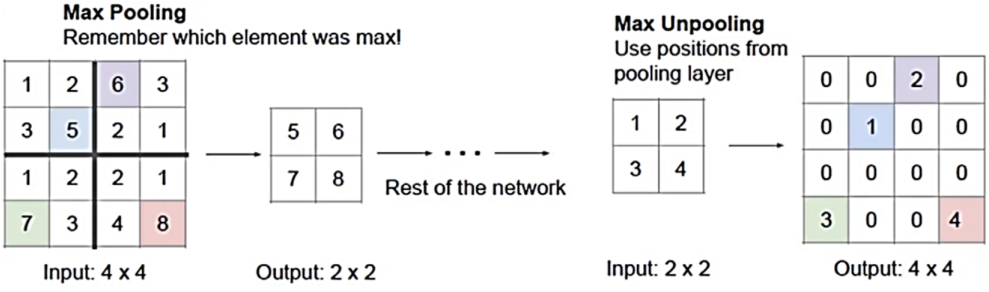

Fig. 3. Max pooling and the Corresponding Max unpooling operations.

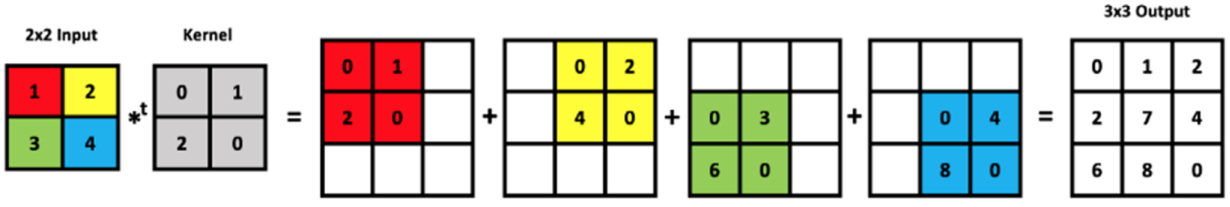

Fig. 4. Transposed Convolution operations.

the number of feature channels. Seg-Net incorporates skip connections that establish direct connections between corresponding encoder max pooling and decoder max unspooling layers. These connections allow the decoder to access lower-level feature maps from the encoder. By combining low-level and high-level features, Seg-Net can recover fine-grained details and improve segmentation accuracy. Seg-Net introduces a novel concept of storing pooling indices during the pooling operation in the encoder. These indices indicate the locations of maximum activations and are subsequently used in the decoder during up-sampling. By leveraging these indices, Seg-Net ensures precise up-sampling and avoids the loss of spatial information. At the end of the decoder, Seg-Net employs a soft-max layer to generate pixel-wise predictions. The soft-max layer assigns a class probability distribution to each pixel, indicating the likelihood of belonging to different semantic categories.

2.2.3. Model training

The training of the Seg-Net model involves using MATLAB framework with a single Nvidia MX110 GPU. The Adaptive Moment Estimation ADAM is used as an optimization algorithm to iteratively update the network weights during the learning process. ADAM produces unique adaptive learning rates for the network weights from estimates of the first and second moments of the gradients. In addition, the model is trained using 15 epochs and 5655 iterations with a constant learning rate of 10^{-3} and a minimum batch size of 1. In comparison with another semantic segmentation model, the Seg-Net model has only 517,632 updatable parameters due to the flat architecture. This architecture involves using the same number of filters (64 in our case) at each convolution layer to extract the feature maps. Therefore, the Seg-Net needs less time to update these parameters and train the model to perform accurate lung segmentation.

3. Results and discussion

To measure the performance of the Seg-Net semantic segmentation model, we used several evaluation metrics such as Global Accuracy (GA), Dice Similarity Coefficient (DSC), and Jaccard index. The global accuracy is usually expressed as the ratio between the truly predicted pixels for all labels in the testing dataset to the total number of pixels, as shown in the following expression:

$$GA = \frac{TN + TP}{TP + TN + FN + FP} \quad \text{Eq.(3)}$$

Where TP refers to the truly predicted pixels as a background, TN is the truly predicted pixels as the lung. FP is the background pixels predicted as lung pixels, and FN is the lung pixels predicted as background pixels. DSC is defined as an overlapping index between the predicted and actual labels as illustrated in the expression below:

$$DSC = \frac{2*TP}{(TP + FP) + (FN + TP)} \quad \text{Eq.(4)}$$

The Jaccard Index (JI) is defined as the similarity measurement between the ground truth and predicted labels. Jaccard index is called Intersection Over Union metric (IoU) and can be expressed as:

$$IoU = \frac{TP}{TP + FN + FP} \quad \text{Eq.(5)}$$

The Seg-Net model is trained using 377 CXR images, during the learning process the model is evaluated using 54 CXR images. We tested the performance of the Seg-Net model using 108 CXR images. Fig. 5 shows the actual and predicted labels for four CXR images.

As shown in Fig. 5, the predicted label in 4-b is very similar to the lung's shape in the CXR image of Fig. 5-a. In addition, we overlapped the actual and predicted labels in Fig. 5-c to distinguish the mismatched areas which are highlighted by green and violet colors. While the white and black colors refer to the pixels that are truly predicted for the lung and background respectively.

The experimental results of the Seg-Net model are listed in Table 1.

To investigate and prove the effectiveness and the ability of the Seg-Net model to perform accurate lung segmentation from the CXR images, we compare our experimental results with other related works as shown in Table 2.

The results shown in Table 2 indicate that most of the related studies employed the CXR images in the Shenzhen dataset to train and test their proposed models. In addition, the U-Net model was the most used technique to perform the segmentation process. As the U-Net has a large number of parameters that need to be updated during the learning process, it requires more time to train the model. In the Seg-Net model, the number of parameters is 90% less than in the U-Net model, therefore it needs less time to train the model. In comparison with the

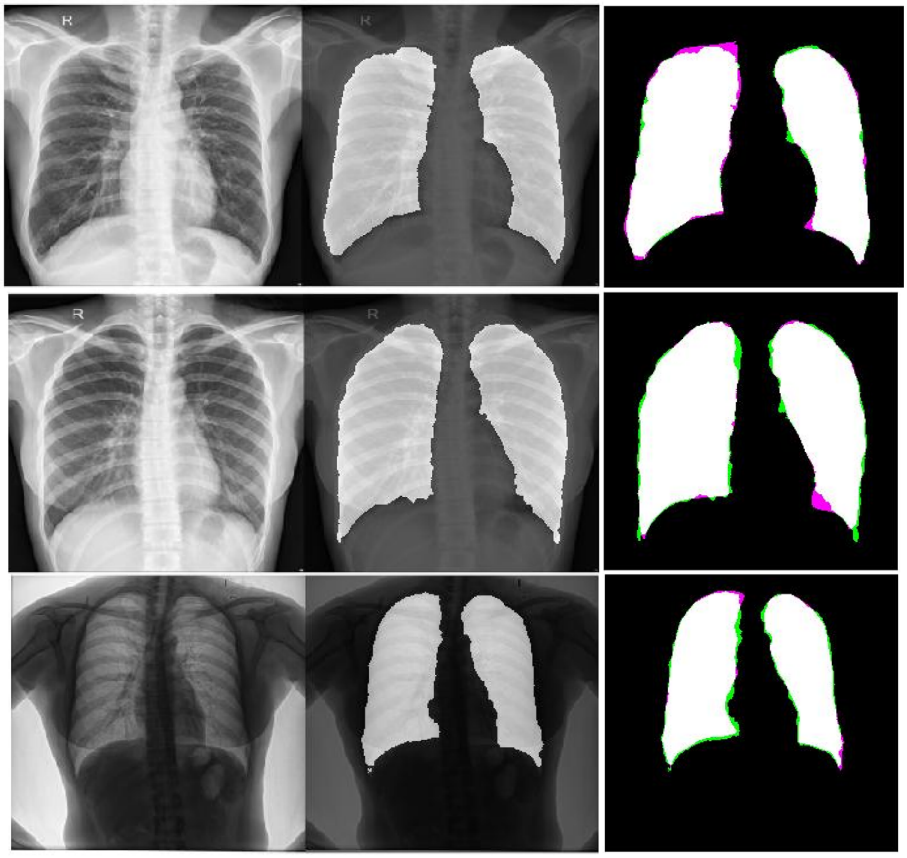

Fig. 5. Cxr samples with actual and predicted labels.

Table 1. Evaluation results of the Seg-Net model.

Evaluation Metric	Value
Global Accuracy	97.71%
DSC	96.95%
IoU	94.08%

results of the other related studies shown in Table, Seg-net achieved higher accuracy, higher DSC, and higher IoU with less parameter number and less training time.

4. Conclusion

In this study, we focused on the vital role of deep learning techniques in enhancing the healthcare sectors by designing a CAD system for the early detection of chest diseases. This system involves employing the Seg-Net model as one of the deep CNN models to identify and segment the lung shapes from the CXR images. Accurate lung segmentation is an important task for building an automated medical image analysis system. By

Table 2. Performance comparison of the Seg-Net model with other related studies.

Authors	Dataset	Model	Accuracy	IoT	DSC
Abas Hasan D and Mohsin Abdulazeez A [14]	Shenzhen	Modified U-Net	97.00%	93.37%	96.75%
Selvan R et al. [13]	Shenzhen	Modified U-Net	88.15%	–	85.03%
Hashem S and Kamil M [19]	Shenzhen	U-Net	91.47%	74.94%	–
Parra Garcia A. [20]	RSNA pneumonia detection dataset	SymBlockNet 5 Based on U-Net	–	89.5%	84.4%
ZH. Hamad and TM. Lung [9]	Shenzhen	U-Net	97.27%	91.70%	95.59%
Our work	**Shenzhen**	**Seg-Net**	**97.71%**	**94.08%**	**96.95%**

evaluating the Seg-Net model and other related state-of-the-art techniques on the Shenzhen data set, Seg-Net shows its ability to perform more accurate lung segmentation than other techniques. In addition, transfer learning can be used with the Seg-Net model to enhance the model performance. It can be used to segment the ROI for multi-class brain scans to extract tumor regions based on the severity.

Conflict of interest

Authors have no conflict of interest to declare.

References

[1] Murugappan M, Bourisly AK, Prakash NB, Sumithra MG, Acharya UR. Automated semantic lung segmentation in chest CT images using deep neural network. Neural Comput Appl 2023;35(21):15343–64. http://dx.doi.org/10.1007/s00521-023-08407-1.

[2] Reamaroon N, Sjoding MW, Derksen H, Sabeti E, Gryak J, Barbaro RP, et al. Robust segmentation of lung in chest x-ray: applications in analysis of acute respiratory distress syndrome. BMC Med Imag 2020;20(1):116. http://dx.doi.org/10.1186/s12880-020-00514-y.

[3] Kamble B, Sahu SP, Doriya R. A review on lung and nodule segmentation techniques. 2020. p. 555–65. http://dx.doi.org/10.1007/978-981-15-0694-9_52.

[4] Saood A, Hatem I. COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. BMC Med Imag 2021;21(1):19. http://dx.doi.org/10.1186/s12880-020-00529-5.

[5] COVID-19 diagnosis systems based on deep convolutional neural networks techniques: a review. In: Dino HI, Zeebaree SRM, Hasan DA, Abdulrazzaq MB, Haji LM, Shukur HM, editors. International conference on advanced science and engineering (ICOASE); 2020 23-24 Dec. 2020; 2020. http://dx.doi.org/10.1109/ICOASE51841.2020.9436542.

[6] Zhang X, Liu X, Zhang B, Dong J, Zhang B, Zhao S, et al. Accurate segmentation for different types of lung nodules on CT images using an improved U-Net convolutional network. Medicine 2021;100(40):e27491. http://dx.doi.org/10.1097/md.0000000000027491.

[7] Krithika Alias AnbuDevi M, Suganthi K. Review of semantic segmentation of medical images using modified architectures of UNET. Diagnostics (Basel) 2022;12(12). http://dx.doi.org/10.3390/diagnostics12123064.

[8] Agrawal T, Choudhary P. Segmentation and classification on chest radiography: a systematic survey. Vis Comput 2023; 39(3):875–913. http://dx.doi.org/10.1007/s00371-021-02352-7.

[9] Hamad ZH, Lung TM. Region segmentation using modified U-net architecture. Eurasian J Sci Eng 2022;8(3):14. http://dx.doi.org/10.23918/eajse.v8i3p25.

[10] Ait Skourt B, El Hassani A, Majda A. Lung CT image segmentation using deep neural networks. Procedia Comput Sci 2018;127:109–13. http://dx.doi.org/10.1016/j.procs.2018.01.104.

[11] Pang T, Guo S, Zhang X, Zhao L. Automatic lung segmentation based on texture and deep features of HRCT images with interstitial lung disease. BioMed Res Int 2019;2019: 2045432. http://dx.doi.org/10.1155/2019/2045432.

[12] Chanda PB, Sarkar SK. Effective and reliable lung segmentation of chest images with medical image processing and machine learning approaches. In: IEEE international conference on advent trends in multidisciplinary research and innovation (ICATMRI)2020; 2020. p. 1–6. http://dx.doi.org/10.1109/icatmri51801.2020.9398450.

[13] Selvan R, Dam E, Detlefsen N, Rischel S, Sheng K, Nielsen M, et al. Lung segmentation from chest X-rays using variational data Imputation. 2020. http://dx.doi.org/10.48550/arXiv.2005.10052.

[14] Abas Hasan D, Mohsin Abdulazeez A. A modified convolutional neural networks model for medical image segmentation. Test Eng Manag 2020;83:16798–808.

[15] Liu W, Luo J, Yang Y, Wang W, Deng J, Yu L. Automatic lung segmentation in chest X-ray images using improved U-Net. Sci Rep 2022;12(1):8649. http://dx.doi.org/10.1038/s41598-022-12743-y.

[16] Chavan M, Varadarajan V, Gite S, Kotecha K. Deep neural network for lung image segmentation on chest X-ray. Technologies 2022;10(5). http://dx.doi.org/10.3390/technologies10050105.

[17] Rao Y, Lv Q, Zeng S, Yi Y, Huang C, Gao Y, et al. COVID-19 CT ground-glass opacity segmentation based on attention mechanism threshold. Biomed Signal Process Control 2023; 81:104486. http://dx.doi.org/10.1016/j.bspc.2022.104486.

[18] Badrinarayanan V, Handa A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. 2015. http://dx.doi.org/10.48550/arXiv.1505.07293.

[19] Hashem S, Kamil M. Segmentation of chest X-ray images using U-net model. Mendel 2022;28:49–53. http://dx.doi.org/10.13164/mendel.2022.2.049.

[20] Parra Garcia A. Automated lung segmentation in chest X-ray images using deep learning, without encoder-decoder skip connections. 2022. http://dx.doi.org/10.13140/RG.2.2.21401.08802.