

COVID-19 CT Image Segmentation Based on Modified U-Net

Wenbo Zheng

Shandong University of Science and Technology, 579 Qianwangang Road, Huangdao District, Qingdao, Shandong, China

334454283@qq.com

Abstract. This paper presents a COVID-19 CT image segmentation method based on a modified U-Net model. In the study and treatment of COVID-19, the segmentation of CT images is crucial for understanding the virus's impact on lung tissues. Traditional image segmentation methods have limitations when dealing with the complex lung lesions caused by COVID-19. Therefore, we optimized the U-Net model's structure and introduced an efficient loss function along with image enhancement techniques to improve segmentation accuracy and computational efficiency. Experiments were conducted on the Medseg and Radiopaedia datasets. The results demonstrate that the modified U-Net model outperforms traditional methods in terms of segmentation accuracy and computational efficiency. Additionally, we discuss the advantages and limitations of the model and propose directions for future research.

Keywords: COVID-19, CT Image Segmentation, U-Net Deep Learning, Medical Image Processing, Loss Function Optimization.

1. Introduction

In the research and treatment of COVID-19, CT image segmentation plays a pivotal role. As a non-invasive imaging tool, CT scanning provides high-resolution images of the lungs, which is crucial for a detailed understanding of how COVID-19 affects lung tissues. By accurately segmenting the infected regions within CT images, medical researchers and clinicians can better pinpoint the specific areas of viral invasion. This precision allows for more accurate diagnosis and the development of tailored treatment plans.

However, the inherent complexity and variability of CT images present significant challenges for precise segmentation. COVID-19-induced lung lesions often display irregular shapes and varying sizes, with blurred boundaries between the infected areas and healthy tissue. This makes it difficult for traditional image segmentation methods to achieve the necessary accuracy and reliability. Precise segmentation is not only vital for the localization of the infected regions but also directly influences subsequent processes such as disease monitoring, lesion analysis, treatment planning, and evaluation of therapeutic efficacy.

Traditional image segmentation techniques, such as threshold-based methods, region-growing algorithms, and edge detection, have performed well in simpler image processing tasks. Yet, when confronted with the high heterogeneity of COVID-19 CT images, these methods often fall short. They struggle to capture sufficient image details in complex structures, leading to suboptimal segmentation accuracy. Consequently, this affects the effectiveness of follow-up research and clinical applications.

In recent years, deep learning models, particularly the U-Net architecture, have shown great promise in the field of medical image segmentation. U-Net is a convolutional neural network (CNN) architecture specifically designed for biomedical image segmentation tasks. Its distinctive U-shaped structure, which combines convolution and deconvolution operations, enables the model to effectively extract and integrate both low-level and high-level features from images. A key feature of U-Net is its skip connections, which introduce feature maps from the convolution stages into the deconvolution process. This mechanism preserves essential spatial information and fine details, thereby significantly enhancing segmentation accuracy. U-Net has been successfully applied in various medical image segmentation tasks, including tumor detection and organ segmentation, demonstrating its robustness and powerful capability.

The primary objective of this study is to further enhance the performance of the U-Net model for COVID-19 CT image segmentation. To achieve this, we introduce several structural optimizations to the original U-Net model and incorporate more efficient loss functions and image enhancement techniques. These improvements not only bolster the model's ability to capture complex lesion features but also increase the stability and consistency of the segmentation results. Our experimental results show that the modified U-Net model outperforms traditional methods across multiple key evaluation metrics, particularly excelling in terms of segmentation accuracy and computational efficiency. In addition to presenting our improved U-Net model, we also explore the advantages and potential limitations of our approach, providing insights into future research directions. By building on the strengths of deep learning in medical image analysis, we aim to contribute a more reliable and efficient solution to the ongoing challenges posed by COVID-19 CT image segmentation.

2. Previous works

In the field of medical image segmentation, traditional methods such as thresholding, region growing, edge detection, and watershed algorithms have been widely used as primary techniques. These methods rely on low-level image features, such as pixel color, brightness, texture, and edge information, and typically implement segmentation through a series of predefined rules and mathematical models. Although these methods have shown some effectiveness in simple image processing tasks, they have gradually revealed significant limitations when applied to real-world medical scenarios, especially those involving complex lesion structures.

Take the thresholding method as an example. It works by setting one or more thresholds to partition the pixels in an image into different regions. Due to its simplicity and ease of implementation, thresholding has been widely used in basic image processing tasks [1]. However, in medical imaging, factors like lighting conditions, noise, and differences in tissue characteristics often lead to highly unstable segmentation results. Specifically, variations in CT and MRI equipment, imaging conditions, and patient physiological differences can cause changes in the brightness values of the same tissue. This makes it difficult for thresholding to provide consistent segmentation. Additionally, noise and artifacts in the images can interfere with the process, resulting in either incorrect or missed segmentation, thus reducing the accuracy and practicality of the method.

Region growing methods use the similarity and connectivity between pixels to gradually expand from a set of initial seed points to adjacent pixels, achieving target region segmentation. While effective in handling simple images with clear boundaries, this approach often encounters issues of over-segmentation or under-segmentation when applied to large-scale medical images. It is also highly sensitive to the selection of the initial seed points and the growth parameters. Even slight errors can lead to inaccurate segmentation results. For complex lesion regions, such as those caused by COVID-19 in the lungs, region growing struggles to adapt to their diversity and irregularity, resulting in inconsistent segmentation outcomes.

Edge detection methods attempt to identify prominent edges in an image to achieve segmentation. However, in medical images, particularly CT and MRI scans, the boundaries of lesion areas are often blurred and complex in shape [2]. Traditional edge detection algorithms face challenges in accurately extracting these boundaries. For example, classic edge detection methods like Canny and Sobel can

easily miss or incorrectly detect edges when dealing with medical images containing fuzzy boundaries, intricate textures, and noise. COVID-19 lesions in lung CT images have irregular shapes and unclear boundaries, further highlighting the limitations of edge detection methods. As a result, they often fail to meet the high precision requirements of clinical applications.

The watershed algorithm is a morphology-based segmentation method that uses gradient information in the image to divide pixels into different "watershed" regions. In theory, it can handle details well. However, the algorithm is highly sensitive to noise and artifacts [3]. Even slight noise can lead to over-segmentation, generating numerous irrelevant regions and adversely affecting segmentation accuracy and efficiency. Since medical images often contain substantial noise and subtle variations, the watershed algorithm faces considerable challenges in practical applications.

these traditional segmentation methods address certain aspects of medical image segmentation but have apparent limitations. They typically rely on predefined rules and feature selection, lacking a comprehensive understanding of global image information. Consequently, they often fall short in handling complex and diverse medical images, such as COVID-19-induced lung lesions, which present irregular shapes, blurred boundaries, and high heterogeneity.

To overcome these challenges, researchers have recognized the limitations of traditional methods and started exploring more advanced segmentation techniques. In recent years, deep learning, particularly Convolutional Neural Networks (CNNs), has brought about a revolutionary change in medical image segmentation [4]. Deep learning-based segmentation methods can automatically learn and extract high-level features from images, eliminating the need for manual feature selection. This approach significantly enhances segmentation accuracy, robustness, and generalization capabilities. The introduction of the U-Net model has become a milestone in the field of medical image segmentation. U-Net adopts a unique U-shaped architecture with skip connections that fuse features at different levels. This enables the model to capture both detailed information and global features in the image, achieving high-precision segmentation. Since its inception, U-Net and its various improved versions have been widely applied in areas such as tumor detection, organ segmentation, and tissue recognition, yielding remarkable results.

The strength of deep learning models also lies in their ability to learn and extract complex feature representations from large datasets, adapting to various forms of lesions and tissue structures. Researchers have employed techniques such as attention mechanisms, improved loss functions, and data augmentation to further enhance model performance. These improvements have enabled deep learning models to excel in handling the complexity of medical images. Particularly in the segmentation of COVID-19 lung CT images, U-Net-based models have significantly improved segmentation accuracy by optimizing network structures and training strategies [5]. This provides robust technical support for precise clinical diagnosis and individualized treatment. Additionally, the development of deep learning technologies opens up new possibilities for real-time processing and automated diagnosis, paving the way for advancements in medical image analysis.

3. Dataset and Preprocessing

The dataset used in this study consists of two parts: the Medseg portion and the Radiopaedia portion. The Medseg part includes 100 axial CT images from more than 40 COVID-19 patients [6]. These images were initially in JPG format and were converted into standardized CT images through a public platform. Specifically, the Medseg dataset contains 100 training slices, each with a resolution of 512x512 pixels, along with the corresponding training masks. The masks have four channels, representing "ground glass", "consolidations", "lungs other", "background". Additionally, the dataset includes 10 test slices for model evaluation.

The Radiopaedia portion of the dataset is sourced from Radiopaedia and consists of 9 axial volumetric CT scans. This dataset is more complex, containing a total of 829 slices. Among these, 373 slices were assessed as positive by radiologists and were annotated accordingly. Similar to the Medseg part, these images were also converted and standardized. The corresponding masks for the Radiopaedia dataset have four channels, each representing different lung lesion types and background regions.

Before performing image segmentation, medical images and their corresponding mask files were loaded from the specified directory. These image files were converted into a floating-point format to facilitate processing. By visualizing these images and masks, the quality of the data and potential issues, such as noise or uneven lighting conditions, can be intuitively understood. Next, the pixel value distribution histogram of the original image data was plotted to analyze its distribution characteristics. This step is crucial as it helps to observe basic information, such as image illumination and contrast, which guides the subsequent normalization process. To reduce instability during model training and improve the model's adaptability to the images, a normalization process was applied. This typically involves adjusting the pixel values to a range between 0 and 1, making the processed data more suitable for deep learning models.

4. Model

In this study, the original U-Net architecture was modified to enhance its performance in medical image segmentation, particularly for the segmentation of COVID-19 lung CT scans. These modifications, along with the training process, hyperparameter settings, chosen loss function, and optimization techniques, are described in detail to illustrate the enhancements made to the U-Net model.

The original U-Net architecture was proposed by Ronneberger et al. [5]. Due to its symmetrical structure, it is widely recognized in the medical image segmentation field for precise localization. However, this architecture sometimes falls short when processing images with high variability and complex details, such as those depicting lung diseases affected by COVID-19. To address these issues, several modifications were introduced.

First, the depth of the network was increased by adding extra convolutional layers in both the contracting (downsampling) and expanding (upsampling) paths. This modification enables the model to capture more complex features, which is crucial for accurately segmenting the subtle characteristics of lung infections. Next, an advanced activation function was adopted. The traditional ReLU function was replaced with the LeakyReLU activation function, as it allows for a small non-zero gradient when the unit is inactive. This can help address the "neuron death" problem encountered with ReLU. Additionally, different padding techniques were employed to ensure that the initial image size matches the output size after passing through the U-Net. In this modified U-Net, 2x2 max pooling, 1x1 convolution, 2x2 up-convolution, and padding of 1 were used. This configuration allows the model to capture complex information while maintaining the original pixel dimensions of the input image. Binary cross-entropy was chosen as the loss function due to its effectiveness in binary classification tasks, such as distinguishing infected regions from non-infected regions in lung CT scans. Moreover, the intersection over union (IoU) metric was monitored as a secondary indicator to evaluate the accuracy of the segmentation boundaries.

5. Results

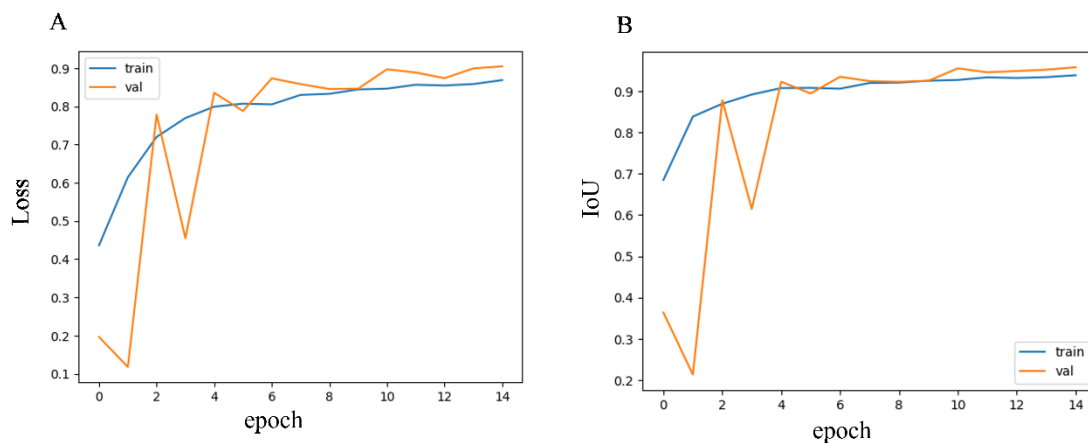


Figure 1. Loss and IoU curve.

From Figure 1, it can be observed that the model's accuracy (Acc) and intersection over union (IoU) curves indicate a gradual improvement in performance as training progresses. In the initial stages, due to weight initialization and the model's limited learning of data features, both Acc and IoU are relatively low. However, as the number of training epochs increases, the model becomes more effective at capturing key information in CT images. This is especially true when handling more complex lesion areas. Over time, the curves tend to stabilize, suggesting that the model has reached a state of convergence.

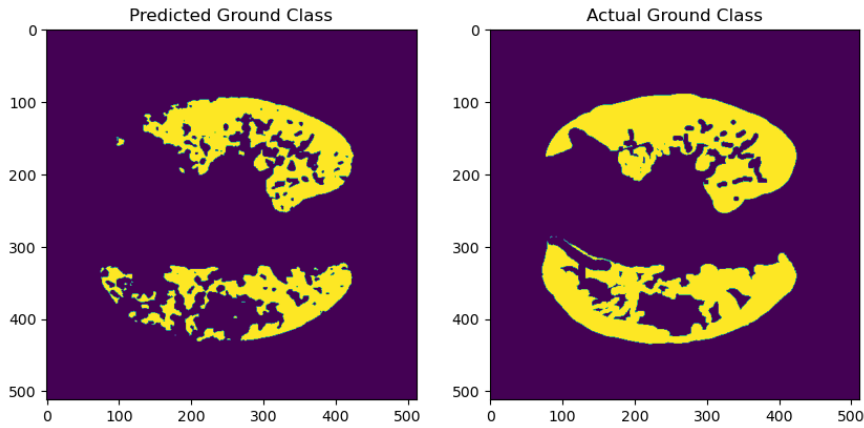


Figure 2. Samples with high IoU.

Figure 2 shows that, in some samples, the model accurately segments the lesion areas, achieving high Acc and IoU levels. This is particularly evident when the contours are clear, and the lesion regions are relatively concentrated. Analysis of these samples reveals that the model can effectively identify boundaries, demonstrating strong generalization ability for more regular lesion regions. However, even in these well-performing samples, small errors in tiny regions still exist. Future improvements could be made by increasing the model's complexity or introducing more fine-grained labels to enhance performance.

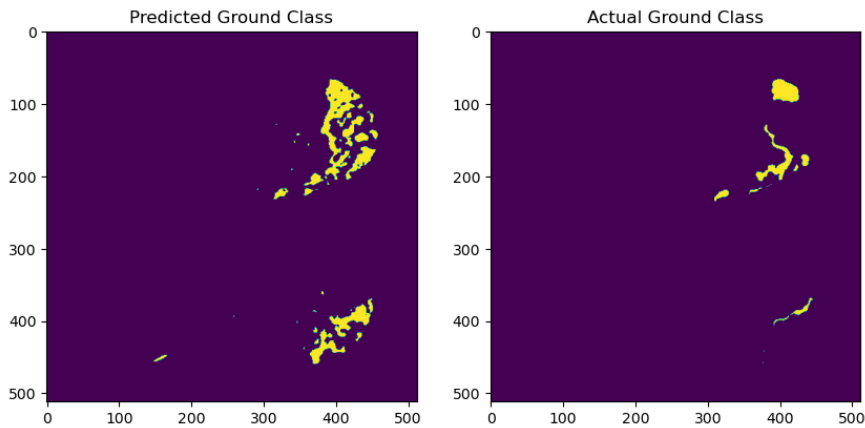


Figure 3. Samples with low IoU.

From Figure 3, it can be seen that in some samples, the Acc and IoU values are significantly lower. The model fails to effectively segment the lesion areas in these cases. These samples often have blurry lesion regions with unclear boundaries, especially when the contrast between the lesion and normal tissue is low, leading to misjudgments by the model. Further analysis indicates that this is related to the model's insufficient generalization ability for different types of lesions, particularly when handling

complex structures. To address this issue, more advanced image enhancement techniques, such as adaptive contrast enhancement, can be introduced in the future. Additionally, using more refined loss functions could help optimize the model's edge detection capabilities.

6. Discussion

The improved U-Net model demonstrates significant advantages in the field of medical image segmentation, particularly in the analysis of COVID-19 CT images. Its strengths include enhanced depth and complexity, enabling the model to capture more subtle lesion features. The addition of padding increases the stability of the training process, while precise image alignment techniques ensure consistency between input and output image dimensions, which is essential for the accurate localization of lesion areas. Furthermore, the use of the binary cross-entropy loss function improves model performance in binary classification tasks, particularly in distinguishing between infected and non-infected regions.

However, these improvements also bring about certain challenges. The model incurs higher computational costs, leading to longer processing times and the need for more advanced hardware resources. Additionally, the model's complexity increases the risk of overfitting, especially when the training dataset is not sufficiently large or diverse. This complexity further complicates hyperparameter tuning, making the optimization process more time-consuming. Despite these challenges, the enhanced U-Net model has the potential to achieve more efficient and accurate image analysis through the future development of automated hyperparameter tuning, multimodal data fusion, and real-time analysis techniques.

The impact of the improved U-Net model on COVID-19 diagnosis and treatment is substantial. Rapid and accurate lesion identification facilitates early diagnosis and timely intervention, which are critical in managing the spread and severity of the disease. Precise segmentation of lesion areas provides a scientific basis for the development of personalized treatment plans, allowing clinicians to tailor interventions based on the extent and characteristics of the infection. Additionally, by comparing segmentation results before and after treatment, physicians can more effectively assess the efficacy of therapeutic strategies and make necessary adjustments.

Looking to the future, several promising directions can be explored to further enhance the model's capabilities and expand its applications. One key area is the integration of multimodal data, such as combining CT images with other medical imaging modalities (e.g., MRI, PET) or incorporating clinical data (e.g., laboratory results, patient history). This multimodal approach could provide a more comprehensive analysis of disease patterns, thereby improving the model's segmentation accuracy and robustness. Another direction is the incorporation of advanced techniques like attention mechanisms and transformers, which have shown success in other fields of deep learning. These methods could further enhance the model's ability to focus on relevant regions within complex medical images.

Furthermore, the development of automated and adaptive hyperparameter tuning algorithms could significantly reduce the need for manual adjustments, making the model more accessible for broader clinical use. The use of federated learning also presents an opportunity to train the model on distributed datasets across different institutions without compromising patient data privacy. This approach would not only improve the generalizability of the model but also address the challenge of data scarcity in specific diseases.

Expanding the dataset to include a wider variety of medical image data, encompassing different diseases and diverse patient populations, would further enhance the model's generalization capabilities. Disease-specific model adjustments and customizations will be crucial to accommodate the unique characteristics of various conditions. For example, adapting the model to detect early-stage cancers, neurological disorders, or cardiovascular abnormalities could revolutionize the diagnostic process in these fields.

7. Conclusion

In conclusion, the improved U-Net model holds great promise for playing an essential role in a broader range of medical image segmentation tasks. Its application extends beyond COVID-19 to potentially transform how diseases are diagnosed, monitored, and treated. With ongoing research and development focusing on addressing current challenges, such as computational costs, overfitting, and data diversity, the model's effectiveness and versatility will continue to grow. Future advancements in this area could lead to more precise, real-time image analysis, ultimately enhancing patient outcomes and advancing personalized medicine.

References

- [1] Y. Guo, Y. Wang, K. Meng, and Z. Zhu, "Otsu Multi-Threshold Image Segmentation Based on Adaptive Double-Mutation Differential Evolution," *Biomimetics*, vol. 8, no. 5, Art. no. 5, Sep. 2023, doi: 10.3390/biomimetics8050418.
- [2] J. Zhang, W. Wang, and J. Wang, "Edge Detection in Colored Images Using Parallel CNNs and Social Spider Optimization," *Electronics*, vol. 13, no. 17, Art. no. 17, Jan. 2024, doi: 10.3390/electronics13173540.
- [3] M. Leena Silvester and R. M. S. Kumar, "Segmentation of Images Using Watershed and MSER: A State-of-the-Art Review," in *Recent Advances in Intelligent Systems and Smart Applications*, M. Al-Emran, K. Shaalan, and A. E. Hassanien, Eds., Cham: Springer International Publishing, 2021, pp. 463–480. doi: 10.1007/978-3-030-47411-9_25.
- [4] B. Kayalibay, G. Jensen, and P. van der Smagt, "CNN-based Segmentation of Medical Imaging Data," Jul. 25, 2017, arXiv: arXiv:1701.03056. doi: 10.48550/arXiv.1701.03056.
- [5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
- [6] Igor.Slinko., COVID-19 CT Images Segmentation, 2020. [Online]. Available: <https://kaggle.com/competitions/covid-segmentation>