Research on thyroid nodule segmentation algorithm based on improved U-Net model

Peng Zhou, *Zhangjing* Wu, *Yong* Yu*, *Yuxuan* Zhao, *Dan* Huang, and *Min* Zhang

School of Public Health, Hubei University of Medicine. Shiyan 442000. China

Keywords: Deep learning, thyroid nodule ultrasound image segmentation, U-Net, Attention mechanism.

Abstract. This research proposes an image segmentation model based on an improved U-Net network (rcU-Net) in response to the phenomena of misinterpretation and missed diagnosis in the process of artificial diagnosis and screening due to the variability of thyroid nodule size and unclear edges in ultrasound images. This paper uses the TN3K dataset as the experimental dataset. The superiority of the proposed model is validated through comparative experiments and ablation experiments. Experimental results show that the proposed model achieves an accuracy of 95.61%, an AUC of 90.67%, a specificity of 98.12%, and a Dice coefficient of 75.69% on the thyroid ultrasound image dataset. The deep learning model proposed in this paper performs well in the segmentation task of thyroid nodule ultrasound images, providing new solutions for the identification of small nodules, nodule edge segmentation, and noise interference in thyroid nodule segmentation tasks.

1 Background

Thyroid nodules, also known as thyroid occupying diseases, are prevalent in clinical practice, with benign and malignant types.Due to the lack of obvious symptoms in the early stages, patients often cannot be diagnosed through symptom observation, which can easily develop into serious diseases such as malignant tumors. According to a survey conducted by the Chinese Society of Endocrinology of the Chinese Medical Association[1]in 31 provinces, the prevalence of hyperthyroidism among Chinese people is 1.22%; hypothyroidism prevalence is 13.95%; the highest prevalence is thyroid nodules, accounting for 20.43%.Thyroid diseases affect approximately 40% of the adult population in China, with a total population exceeding 200 million.

Currently, computer-aided diagnosis has a certain research foundation, and the emergence of convolutional neural networks has shown more practical value.Such as U-Net[2], V-Net[3]and TransU-Net[4] are commonly used in medical image segmentation; while

* Corresponding author: yongyu@hbmu.edu.cn

models like ResNet[5] and GoogleNet[6]are commonly used in medical image classification. Ultrasound images often have low contrast and artifacts, such as shadows, echo interference, etc. Therefore, to address these issues, this paper focuses on the thyroid nodule segmentation models and methods, proposing a high-precision, high-efficiency automatic segmentation method, which can provide certain assistance for subsequent thyroid diagnosis and treatment.

2 Materials and methods

This paper proposes an improved U-Net network model for thyroid nodule ultrasound image segmentation, called rcU-Net. This model uses the U-Net model as the backbone network and introduces residual connections to increase the depth of the network and alleviate the vanishing gradient problem. At the same time, it introduces attention mechanisms to improve the recognition ability and segmentation accuracy of small nodules.

2.1 RCU -Net network structure

This paper introduces an improved U-Net model for thyroid nodule segmentation.The model is based on the U-Net framework, and its overall structure includes an encoder path, attention modules, residual blocks, and a decoder path.In the encoder path, each feature map undergoes two rounds of convolution, with each round having two channels.First, the feature map is processed through a 3×3 convolution, while also passing through a second-channel residual block, and converged through residual connections.Then, the feature map obtained from the previous step is processed through a 3×3 convolution in the first channel, while also introducing the attention mechanism from the second channel, and finally integrating the results through feature concatenation.This process completes hierarchical feature extraction, followed by pooling operations to enter the next level.

2.2 Dataset

The experimental data used in this study is the TN3K dataset, which includes 3493 ultrasound images from 2421 patients taken between January 2016 and August 2020[7].These images were selected from over 30,000 images provided by hospitals such as the Southern Medical University Zhujiang Hospital, based on the following criteria:(1) Each image contains at least one thyroid nodule region;(2) Exclusion of lymph images or images containing large colored regions;(3) Only one representative image is selected from multiple images in the same area or from the same viewpoint of the patient.The dataset is divided into training and testing sets, consisting of 2879 and 614 images respectively.

2.3 Introducing residual connection

To alleviate the gradient vanishing problem that occurs as the network becomes deeper[8],this paper combines the dual convolutional feature extraction module with the residual idea,adds dimension-matching convolutional layers to the input,and then connects the two parts to form the final output feature map.By introducing residual structures, the dual convolutional feature extraction module can better capture the details and contextual information of the image,while alleviating gradient vanishing, accelerating network convergence, and improving training efficiency.

2.4 Improved attention mechanism

The CBAM attention mechanism [9] adopts dual-pooling channels of average pooling and max pooling to extract features.In medical images, there is little grayscale difference between small nodules and thyroid parenchyma.Using average pooling may lead to a loss of information about small nodules, thereby reducing segmentation effectiveness.Therefore, we replace dual-channel max pooling and average pooling with single-channel max pooling.Max pooling can gather clues about unique object features to infer finer channel attention, highlighting small nodules, image noise, and pixel differences in the thyroid parenchyma, reducing the chance of false positives and improving segmentation accuracy.

2.5 Evaluation indicators

The network in this paper is both an object detection network and a binary classification network. Therefore, this paper adopts precision, recall, accuracy, area under the ROC curve (AUC), dice coefficient, specificity, and mean absolute error (MAE) as evaluation metrics for the network in this paper. The calculation formulas are shown in equations (1) to (6):

$$\text{accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$\text{recall} = \frac{TP}{TP+FN} \tag{2}$$

$$\text{precision} = \frac{TP}{TP+FP} \tag{3}$$

$$\text{MAE} = 1/n \sum_{i=1}^{n} |\widehat{y_i} - y_i| \tag{4}$$

$$\text{Dice} = \frac{2|X+Y|}{|X|+|Y|} \tag{5}$$

$$\text{specificity} = \frac{TN}{TN+FP} \tag{6}$$

Where |X∩Y| represents the number of elements in the intersection of X and Y, and |X| and |Y| represent the number of elements in sets X and Y, respectively.TP:True Positive, FP:False Positive, TN:True Negative, TN:True Negative, FN:False Negative.

3 Experimental results

This chapter proposes a thyroid nodule segmentation method based on an improved U-Net network and analyzes the experimental results.Firstly, detailed explanations are given for the experimental dataset and parameter settings.Next, an analysis of the segmentation results of the experiments is conducted.Finally, the analysis includes comparative experiments, ablation experiments, and visualization of experimental results.

3.1 Experimental parameter settings

The specific parameters of the network in the experiment are as follows: Batch size is set to 1, the initial learning rate is set to 0.001, the model optimizer is momentum, Dice-loss is used as the loss function, and the maximum number of iterations is 40.During the training process, the operating system is Windows 10, the programming language is Python 3.9, and all programs are implemented in the PyTorch framework.

3.2 Segmentation results and analysis

As shown in Figure 1, from columns (b), (c), (d), and (e), it can be observed that both SegNet and Unet++ networks exhibit obvious over-segmentation and under-segmentation phenomena.From columns (e), (f), (g), and (h), it can be seen that the segmentation performance of models with different attention mechanisms is superior to SegNet[11]and Unet++[10]networks. However, there still exists the phenomenon of excessive fine segmentation.Among them, the model with the Shuffle Attention mechanism[12] shows many small white dots in the segmentation results, indicating that the model's resistance to noise interference is weak after adding the Attention mechanism.

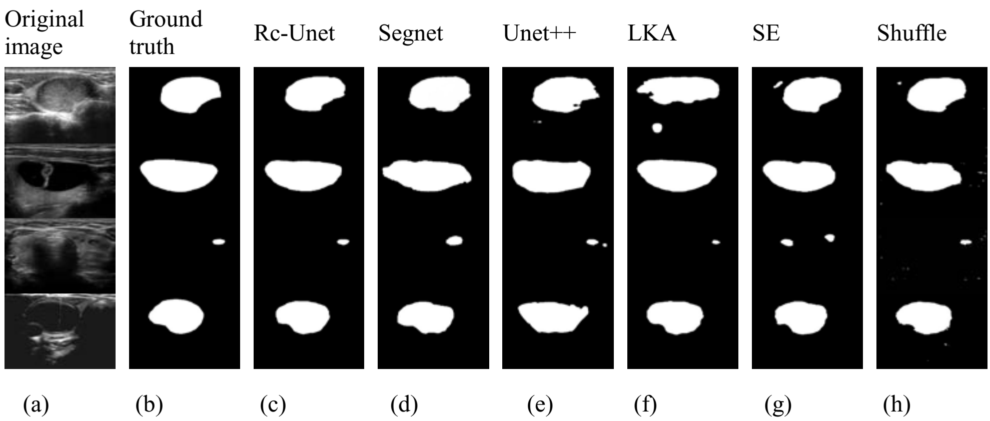

Fig. 1. Segmentation results of differentels and different attention mechanisms.

3.3 Comparative test analysis

Tables 1 and 2 present the specific metrics for network segmentation and compare them. Table 1 compares the metrics between different attention mechanisms, while Table 2 compares the segmentation effectiveness of different network models.

Table 1. Segmentation results of different attention mechanisms.

	Precision	recall	accuracy	auc	dice	specificity	mae
No attention mechanism	66.55	76.84	93.95	87.03	63.02	97.23	6.37
CBAM	75.36	77.04	94.79	87.51	68.45	97.99	5.47
LKA	74.72	79.53	95.63	88.62	69.90	97.74	6.47
Shuffle	73.79	82.40	95.63	90.19	74.00	97.98	5.00
SE	78.59	82.65	95.59	90.46	75.51	98.28	5.39
improved CBAM	77.01	83.21	95.61	90.67	75.69	98.12	4.42

From Table 1, it can be observed that after adding the improved attention mechanism, the model's performance metrics are significantly higher compared to the model without the attention mechanism. Additionally, compared to the original CBAM attention mechanism, the improved attention mechanism in this paper performs better across the mentioned metrics. Furthermore, compared to other attention mechanisms, the improved attention mechanism in this paper demonstrates good performance across most metrics. Looking at the Dice coefficient and Mean Absolute Error (MAE), the Dice coefficient of the improved CBAM attention mechanism is higher than other attention mechanisms and models without attention mechanisms, while the MAE is the lowest among all cases, indicating that the attention mechanism proposed in this paper has been effective. The segmentation accuracy of the model is higher than models with other attention mechanisms and without attention mechanisms. Regarding recall, precision, and AUC metrics, the metrics of the improved CBAM attention mechanism reach 83.21%, 95.61%, and 90.67%, respectively, indicating that the attention mechanism proposed in this paper can effectively capture long-distance dependencies in images to improve segmentation accuracy.

Table 2. Segmentation results of different network models.

	Precision	recall	accuracy	auc	dice	specificity	mae
Unet	66.55	76.84	93.90	87.03	63.02	97.23	6.37
Unet++	70.38	75.68	94.79	86.54	63.80	97.42	5.67
Sgunet	67.21	76.78	93.92	86.77	62.63	96.77	6.10
Segnet	68.92	74.10	94.24	85.50	62.53	96.92	5.77
Resnet	76.20	78.07	95.44	88.07	69.45	98.10	5.25
rcU-Net	77.01	83.21	95.61	90.67	75.69	98.12	4.42

Table 2 provides a detailed overview of the performance of different network models in segmenting thyroid nodule ultrasound images. From Table 2, it can be seen that, compared to other models, the model proposed in this paper outperforms others in the mentioned metrics, with the smallest absolute mean error.

In summary, by comparing the performance metrics of different network models and attention mechanisms in thyroid nodule ultrasound image segmentation tasks, we can conclude that the improved attention mechanism proposed in this paper performs better across most metrics. Our model exhibits good performance in multiple key metrics, with high precision, recall, accuracy, AUC value, Dice coefficient, and specificity, while maintaining a low average absolute error. These advantages enable our model to achieve higher accuracy and better robustness in thyroid nodule ultrasound image segmentation tasks.

3.4 Ablation experiment analysis

Table 3 presents the results of ablation experiments analyzing the network segmentation model under different module combinations. From the table, it can be seen that different combinations of modules have a significant impact on the model's performance.

Table 3. Ablation experiment analysis results.

	Precision	recall	accuracy	auc	dice	specificity	mae
Unet	66.55	76.84	93.95	87.03	63.02	97.23	6.37
backbone+ improved CBAM	70.38	75.68	94.89	86.54	63.80	97.42	6.67
backbone+res	76.20	78.07	95.44	88.07	69.58	98.10	5.25
rc-Unet	77.01	83.21	95.61	90.67	75.69	98.12	4.42

Firstly, compared to the basic U-Net model, the U-Net+improved CBAM model with improved CBAM module shows improvements in precision, recall, and accuracy, but there is a slight decrease in recall and AUC values, indicating that while the improved CBAM module enhances some metrics, it may sacrifice the performance of other metrics. Secondly, the performance of the U-Net+res model is comprehensively improved after incorporating the residual module, especially in precision, recall, accuracy, and Dice coefficient, indicating the effectiveness of the residual module in enhancing model performance. Finally, the rc-Unet model, combining the improved CBAM and residual modules, achieves the best performance in all metrics, especially in precision and recall. This suggests that the rc-Unet model can fully leverage the advantages of different modules when integrating them, optimizing performance. In summary, the ablation experiments analysis indicates that different combinations of modules have a significant impact on model performance. By judiciously selecting and optimizing module combinations, further improvements in the performance of network segmentation models can be achieved.

4 Conclusion

This paper addresses a series of issues in the segmentation of thyroid nodules in ultrasound images and completes the segmentation task of thyroid nodules using an improved U-Net network. The work includes introducing residual mechanisms and attention modules, improving attention modules, and utilizing the improved U-Net network for segmenting thyroid nodule ultrasound images, achieving good segmentation results. Experimental results demonstrate that the proposed method outperforms existing methods in segmentation performance, enabling more accurate segmentation of thyroid nodules in ultrasound images. Overall, this study has successfully achieved its intended objectives, but there are limitations in data acquisition. Due to the small scale of the dataset used, even increasing the number of training iterations cannot fully extract more feature information. Additionally, our current network models mainly focus on the segmentation task of thyroid nodules. Future research will further extend to the classification of benign and malignant nodules to achieve higher classification accuracy.

This study was funded by the Hubei Province College Student' Innovation and Entrepreneurship Training Project (No. S202210929038)

References

1. Endocrinology Branch of the Chinese Medical Association. Survey report on the prevalence of thyroid diseases in China [N]. Chinese Journal of Science, 2020-06-03 (Comprehensive 4th Edition).
2. Ronneberger O ,Fischer P ,Brox T .U-Net: Convolutional Networks for Biomedical Image Segmentation.[J].CoRR,2015,abs/1505.04597
3. Milletari F ,Navab N ,Ahmadi S .V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation.[J].CoRR,2016,abs/1606.04797
4. Chen J , Lu Y , Yu Q ,et al.TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation[J]. 2021.DOI:10.48550/arXiv.2102.04306.
5. He K , Zhang X , Ren S ,et al.Deep Residual Learning for Image Recognition[J].IEEE, 2016.DOI:10.1109/CVPR.2016.90.
6. Szegedy C ,0015 L W ,Jia Y , et al.Going Deeper with Convolutions.[J].CoRR,2014,abs/1409.4842
7. Gong H , Chen G , Wang R ,et al.Multi-task Learning for Thyroid Nodule Segmentation with Thyroid Region Prior[C]//International Symposium on Biomedical Imaging 2021.2021.DOI:10.1109/ISBI48211.2021.9434087.
8. Chang W Y , Lin T N .Vanishing Nodes: Another Phenomenon That Makes Training Deep Neural Networks Difficult[J]. 2019.DOI:10.48550/arXiv.1910.09745.
9. Woo S , Park J , Lee J Y ,et al.CBAM: Convolutional Block Attention Module[J].Springer, Cham, 2018.DOI:10.1007/978-3-030-01234-2_1.

10. Zhou Z , Siddiquee M M R , Tajbakhsh N ,et al.UNet++: A Nested U-Net Architecture for Medical Image Segmentation[J]. 2018.DOI:10.1007/978-3-030-00889-5_1.
11. Badrinarayanan V ,Kendall A ,Cipolla R .SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation.[J].CoRR,2015,abs/1511.00561
12. Yang Y B .SA-Net: Shuffle Attention for Deep Convolutional Neural Networks[J]. 2021.DOI:10.1109/ICASSP39728.2021.9414568.