ISSN 2616-5775 Vol. 8, Issue 3: 51-55, DOI: 10.25236/AJCIS.2025.080307

Research on Faster R-CNN Lung Nodule Detection Algorithm Based on Residual Attention Network

Dang Xuan[1,*]

[1]*School of Intelligent Science and Engineering, Xi'an Peihua University, Xi'an, China*
**Corresponding author*

Abstract: *This paper proposes a Faster R-CNN lung nodule detection algorithm based on residual attention networks, aiming to enhance the accuracy and efficiency of lung nodule detection. By incorporating residual networks and attention mechanisms, the feature extraction capability is strengthened, enabling the model to more effectively capture the subtle features of lung nodules. The improved Faster R-CNN performs exceptionally well in handling complex backgrounds and multi-scale targets, significantly boosting detection performance. Experimental results demonstrate that this method achieves outstanding detection results on multiple public datasets.*

Keywords: *Residual attention network; Faster R-CNN; Pulmonary nodule detection*

1. Introduction

Early detection of pulmonary nodules is of great significance in the diagnosis and treatment of lung cancer. However, due to the fact that pulmonary nodules often appear in various forms on CT images and are mixed with complex backgrounds, traditional detection methods have shortcomings in accuracy and efficiency[1]. In recent years, the rapid development of deep learning techniques, especially convolutional neural networks (CNN), has provided new solutions for medical image analysis[2]. Faster R-CNN, as a classic object detection algorithm, has achieved significant results in multiple fields, but still faces challenges in lung nodule detection, especially in capturing subtle features.

Although Faster R-CNN has made significant progress in the field of object detection, its application in medical image analysis still faces challenges[3]. Due to the complex background and low contrast characteristics of medical images, traditional Faster R-CNN may experience a decrease in detection accuracy when processing such data. To this end, researchers have proposed various improvement strategies, such as combining residual networks and attention mechanisms to enhance feature extraction capabilities, and implementing multi-scale feature fusion through Feature Pyramid Networks (FPNs). These improvements not only enhance the adaptability of Faster R-CNN in medical imaging, but also provide new solutions for object detection in complex scenes[4].

In this context, researchers are constantly exploring and improving methods to enhance detection performance. In order to address the shortcomings of Faster R-CNN in lung nodule detection, this paper proposes an improved algorithm based on residual attention network. The introduction of residual networks helps alleviate the problem of gradient vanishing and enhances the learning ability of the network through effective transfer of deep features. Meanwhile, the application of attention mechanism enables the model to focus on important feature regions, enhancing its ability to capture subtle features.

2. Scheme design and improvement

2.1. Network Architecture Design

This network uses a U-Net encoder decoder structure to effectively learn deep networks. The specific structure is shown in Figure 1.

Firstly, the preprocessed image is uniformly cropped into 96×96×96 3D blocks, which contain annotated nodule information. In the encoder subnet, the feature map is processed by the deconvolution layer and residual attention module. The deconvolved features are fused with low-level features, and the feature map information is integrated through the residual attention module. Then, overfitting is prevented through convolutional layers with Dropout. Finally, through a 1×1×1 convolution, a 4-

dimensional tensor of 24×24×24×15 was output, and three anchor boxes were designed based on the distribution of nodule sizes, with sizes of 5, 10, and 20. For each anchor box, there are 5 loss functions, namely the probability of classification and the three-dimensional coordinates and diameter of the nodules.

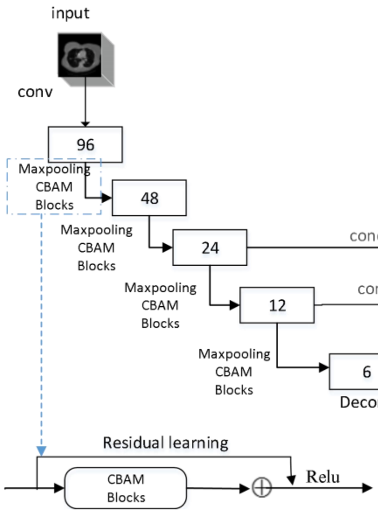

Figure 1: Network structure for pulmonary nodule detection

2.2. Construction of Residual Attention Network

In this article, we designed a feature extraction module that combines residual networks and attention mechanisms, as shown in Figure 2, to improve the accuracy and efficiency of lung nodule detection. Residual networks (ResNet) are widely used due to their ability to alleviate gradient vanishing problems in deep networks, effectively enhancing the depth and breadth of feature extraction through residual connections. However, relying solely on residual networks still has certain limitations when dealing with complex backgrounds and multi-scale targets. Therefore, we introduced an attention module based on ResNet to achieve adaptive adjustment of feature map weights. This improvement enables the network to focus more on the subtle features of lung nodules, thereby improving the accuracy of detection.

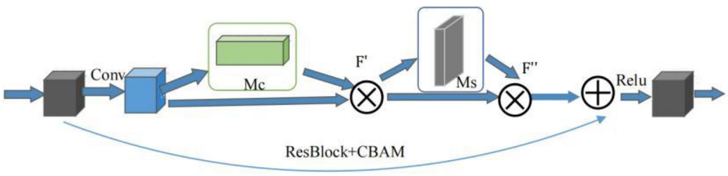

Figure 2: Structural diagram of residual attention network

The introduction of attention mechanism enables the network to dynamically adjust the level of attention to different features during feature extraction, especially in small object detection. By combining channel attention and spatial attention, the network can effectively filter redundant information and highlight important features. This mechanism is particularly important in medical imaging analysis, as lung nodules often have varying degrees of adhesion to surrounding tissues, and traditional methods tend to overlook these subtle differences.

2.3. Improve Faster R-CNN

In order to improve the performance of Faster R-CNN in lung nodule detection, this study introduced a residual attention network in the feature extraction stage. This network combines residual connections and attention mechanisms, which can effectively alleviate the gradient vanishing problem in deep networks, while dynamically adjusting the weights of feature maps to improve the attention to important features. This combination method performs well in complex backgrounds and multi-scale object detection, especially suitable for small object detection tasks such as pulmonary nodules.

In the optimization of the Regional Proposal Network (RPN), the anchor box settings were redesigned

ISSN 2616-5775 Vol. 8, Issue 3: 51-55, DOI: 10.25236/AJCIS.2025.080307

to address the multi-scale characteristics of pulmonary nodules. By adjusting the size and proportion of the anchor box to make it more suitable for detecting small targets, the accuracy of generating candidate regions has been improved. The application of this strategy significantly improves the detection ability of the model in complex backgrounds, reducing missed and false detections.

Drawing on the successful experience of Hyper Faster R-CNN in microcalcification detection, multi-level feature fusion was achieved through Feature Pyramid Network (FPN), further enhancing the model's ability to capture features at different scales. This multi-scale feature learning strategy has been proven to effectively improve detection accuracy in medical image analysis, especially when processing high-resolution CT images.

Combining the above improvement strategies, the experimental results of the model on multiple public datasets show that its performance in lung nodule detection is superior to traditional methods, verifying the effectiveness of residual attention networks and optimized RPN strategies. This study not only provides new technological means for lung nodule detection, but also provides reference for the improvement of other medical imaging analysis tasks.

2.4. Model Training and Optimization

In the process of model training and optimization, this study adopted transfer learning strategy, using pre trained ResNet model as the initialization basis. This strategy not only accelerates the convergence process of the model, but also significantly improves its generalization ability. The application of transfer learning lies in fully utilizing the feature extraction ability of ResNet pre trained on large-scale datasets, especially in medical image analysis, which can effectively address the problem of insufficient data volume.

In order to further improve the training effectiveness of the model, data augmentation techniques are widely used. Specifically, increasing the diversity of data through operations such as rotation, scaling, and flipping enhances the model's ability to recognize different forms of pulmonary nodules. These techniques can effectively simulate different shooting angles and changes in nodule morphology when processing CT images, thereby improving the robustness of the model.

The adjustment of learning rate and the use of Adam optimizer are key steps in the optimization process. The dynamic adjustment of learning rate helps to converge quickly in the early stages of training, while fine-tuning model parameters in the later stages. The Adam optimizer further improves the detection accuracy of the model in complex backgrounds through an adaptive learning rate mechanism. Combining the characteristics of residual networks and attention mechanisms, the optimization strategy also includes improvements to the Region Proposal Network (RPN) to enhance the accuracy and efficiency of small object detection.

Through the comprehensive application of transfer learning, data augmentation, and optimization strategies, this study has achieved significant performance improvements in lung nodule detection tasks. The combination of these methods not only improves the detection accuracy of the model, but also provides a solid technical foundation for the development of future intelligent diagnostic systems.

3. Experiment and Results Analysis

3.1. Datasets and Experimental Setup

This study selected multiple publicly available lung nodule datasets in the experiment, including LIDC-IDRI and LUNA16. These datasets provide rich CT imaging data, covering lung nodules of different sizes and shapes, suitable for evaluating the performance of detection algorithms. In order to meet the input requirements of the model, the dataset was preprocessed before use, including normalization and slicing operations. Normalization ensures data consistency, while slicing transforms 3D CT images into 2D images suitable for model processing.

In the experimental setup, the dataset was divided into training and testing sets in an 80:20 ratio and subjected to five fold cross validation to ensure the model's generalization ability and stability. Cross validation not only improves the robustness of the model, but also effectively avoids the occurrence of overfitting. The experiment was conducted on NVIDIA GPU and implemented using PyTorch framework, fully utilizing the computing power of GPU and accelerating the training process of the model.

Table 1 shows the distribution of lung nodules of different sizes in the LUNA16 dataset. It can be seen that small-sized nodules account for a large proportion, which puts higher demands on the detection algorithm. To address this challenge, this study introduced an attention mechanism into the model to enhance its ability to recognize small targets.

Table 1: Distribution of pulmonary nodules of different sizes

Nodule size	Quantity (piece)	Proportion (%)
small (<10mm)	300	60
medium (10-20mm)	150	30
large (>20mm)	50	10

The data in the table indicates that detecting small-sized nodules is a key challenge in lung nodule detection. To this end, a multi-scale feature learning network was used in the experiment, which optimized feature extraction through multi-scale dilated convolution and contextual attention mechanism, enhancing its adaptability to complex backgrounds.

The experimental results show that the improved Faster R-CNN performs well in handling complex backgrounds and multi-scale targets, especially in detecting small-sized nodules, achieving significant performance improvements. This result validates the potential of residual attention networks in lung nodule detection and provides a new technical approach for medical image analysis.

3.2. Performance Evaluation Indicators

In order to comprehensively evaluate the performance of Faster R-CNN based on residual attention network in lung nodule detection, this study adopted multiple performance evaluation metrics, including accuracy, recall, precision, and F1 score. These indicators can reflect the detection ability and stability of the model from different perspectives. To evaluate the efficiency of the model, the average detection time was also calculated.

3.3. Experimental Results and Comparative Analysis

In the experiment, the improved Faster R-CNN performed well on the LIDC-IDRI and LUNA16 datasets, especially demonstrating significant advantages in detecting small-sized nodules. Table 2 shows the comparison results of accuracy, recall, and F1 score on these two datasets. It can be seen that the accuracy and recall of our method on the LIDC-IDRI dataset reached 92.1% and 89.5%, respectively, while on the LUNA16 dataset they were 91.3% and 88.7%, respectively. Compared with traditional Faster R-CNN, both accuracy and recall have significantly improved, increasing by 4.3% and 5.2% respectively.

Table 2: Performance comparison between improved Faster R-CNN and traditional methods

Dataset	Method	Accuracy	Recall	F1 score
LIDC-IDRI	Traditional Faster R-CNN	87.8%	84.3%	86.0%
	Improved Faster R-CNN	92.1%	89.5%	90.8%
LUNA16	Traditional Faster R-CNN	86.9%	83.5%	85.2%
	Improved Faster R-CNN	91.3%	88.7%	90.0%

Compared with other advanced detection algorithms such as YOLODM Net and DeepLabv3+based detection methods, our method also performs well in complex background and multi-scale object detection. YOLODM Net has achieved high detection accuracy through multi-scale feature fusion and deep convolution, but there is still a certain problem of missed detections when dealing with small-sized nodules. And the method proposed in this article significantly enhances the feature extraction ability by introducing residual attention network, especially improving the detection accuracy in complex backgrounds.

In the process of feature extraction, the residual attention network dynamically adjusts feature weights to effectively filter redundant information and highlight important features. This mechanism has been validated as an effective means to improve detection accuracy in multiple studies. Through comparative analysis, the advantages of our method in feature extraction have been further verified, especially in dealing with complex backgrounds and multi-scale targets, showing higher robustness and adaptability.

The partial results of the lung nodules detected by the algorithm in this chapter are shown in Figure 3. The first line is the original unlabeled image, the second line is the gold standard image of the real GT label, and the third line shows the results detected by the algorithm in this article, which are measured by

ISSN 2616-5775 Vol. 8, Issue 3: 51-55, DOI: 10.25236/AJCIS.2025.080307

confidence in this article. Each column represents different types of selected pulmonary nodules, from left to right: solitary pulmonary nodules, adherent pulmonary nodules, and vascular adhesive pulmonary nodules. The rectangular box of the gold standard image represents the true location of the nodule, the red box represents the results of the algorithm detection in this chapter, and the number represents the confidence level of the predicted nodule.

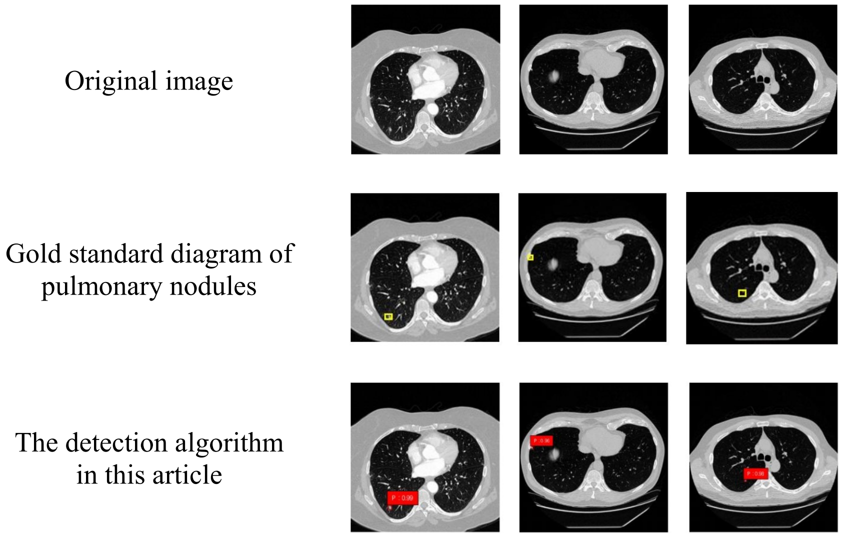

Figure 3: Pulmonary nodule detection result chart

4. Conclusion

The Faster R-CNN lung nodule detection algorithm based on residual attention network proposed in this article significantly improves the accuracy and efficiency of detection by introducing residual network and attention mechanism, especially in complex background and small-sized nodule detection, laying a solid foundation for the development of intelligent diagnostic systems. However, this method still has limitations when dealing with extremely small or large nodules and complex backgrounds, and the computational complexity is relatively high. Future research can be expanded in areas such as model lightweighting, multimodal data fusion, and cross domain applications.

Acknowledgements

This article was supported by the university level research project of Xi'an Peihua University (Project Number: PHKT2515).

References

[1] Yu Shangqian. Detection and Classification of Pulmonary Nodules Based on Convolutional Neural Networks [D]. Hainan University, 2023. DOI: 10.27073/d.cnki.ghadu.2023.000157.
[2] Li Rui. Research on Pulmonary Nodule Detection in CT Images Based on Deep Learning [D]. Anhui University of Science and Technology, 2023. DOI: 10.26918/d.cnki.ghngc.2023.000253.
[3] Ji Junjun. Design and Implementation of a Pulmonary Nodule Detection System Based on Deep Learning [D]. Beijing University of Posts and Telecommunications, 2023. DOI:10.26969/d.cnki.gbydu.2023.001199.
[4] Zhang Hao, Jie Yixuan, Yang Xiaonan, etc. An automatic detection algorithm for pulmonary nodules based on fusion iterative self-organizing data analysis [J]. Computer and Digital Engineering, 2024, 52 (05): 1336-1340+1358