Proceedings of Machine Learning Research vol 245:1–7, 2024 2024 International Conference on Machine Learning and Intelligent Computing

Brain Tumor Detection Algorithm Based on Improved YOLOv7

Jiajun Wu* 827099774@QQ.COM
Shanghai Maritime University, Shanghai, 201306, China

Maoting Gao
College of Information Engineering, Shanghai Maritime University, Shanghai, 201306, China

Editors: Nianyin Zeng and Ram Bilas Pachori

Abstract

A brain tumor detection algorithm based on improved YOLOv7 is proposed to address the problem of low image resolution in detecting small targets in brain tumor detection. By introducing the SPD module in the Backbone section, cross row convolution and pooling operations are eliminated, fine-grained feature learning is strengthened, and the accuracy of model detection is increased; Introducing CA attention mechanism to enhance the learning of more critical and effective features, further improving the efficiency and accuracy of the network model; And use the dynamic non monotonic frequency modulation loss function Wise-IoU to enhance the model's detection ability for low-quality samples. Overall improved YOLOv7 network model significantly improves the accuracy of low resolution samples and small object detection, and can be effectively applied to the detection of brain tumor images.

Keywords: YOLOv7, Brain tumor detection, Object detection

1. Introduction

Brain tumor refer to a group of abnormally proliferating cells formed in brain tissue, which continuously proliferate and compress surrounding normal brain tissue, leading to brain dysfunction and even endangering life. (Liu and Zhang, 2021) At present, the diagnosis of brain tumors mainly relies on magnetic resonance imaging (MRI), which generates a large amount of image data through scanning, which are examined and diagnosed by radiologists. Due to the strong heterogeneity of brain tumor imaging, as well as the large number and multiple sequences of brain tumor images with diverse features such as shape, size, and position, doctors need to spend energy on judgment and diagnosis, which can easily lead to errors in examinations.

At present, deep learning technology has been widely applied in the medical image analysis. After Krizhevsky et al. (2017) significantly improved the performance of deep learning networks, the academic community sparked a wave of deep learning. In terms of brain image analysis, Maleki et al. Reported for the first time the pioneering work of using deep learning for brain image analysis in detecting lesions in multiple sclerosis (MS). Abd Ellah et al. Used AlexNet's convolutional neural network and Error Correction Output Encoding Support Vector Machine (ECOC-SVM) to evaluate cranial images in the Response (Rider) Neuro MRI database. Kang et al. proposed a new reparameterized convolutional YOLO architecture RCS-YOLO based on channel shuffle on brain tumor detection, achieving a balance between accuracy and speed. Kang et al. combined two-level routing attention (BRA), generalized feature pyramid network (GFPN), and fourth detection head into YOLOv8 (Li et al., 2023) and developed a new BGF-YOLO architecture, achieving further improvement in accuracy. However, these models have not improved the feature extraction network for the low resolution characteristics of brain tumor CT images, and their ability to extract target features is often weak.

In summary, current research on brain tumor target detection still faces issues such as difficulty in detecting small targets due to low image resolution. This article combines practical application scenarios, based on the characteristics and task requirements of brain tumor detection, using the YOLOv7 algorithm as the basis, introduces the SPD module in the Backbone section to eliminate cross row convolution and pooling operations, strengthen the learning of fine-grained features, increase the accuracy of model detection, and introduce the CA attention mechanism (Hou et al., 2021) in the Head section to enhance the learning of more critical effective features, further improving the efficiency and accuracy of the network model. In addition, the model uses the loss function Wise-IoU (Tong et al., 2023) to enhance its detection ability for low-quality samples.

2. Brain tumor detection model based on improved YOLOv7

Brain tumor detection has a strong real-time demand, therefore, the basic model mainly considers the One Stage algorithm with fast detection speed. Comparative experiments conducted on large-scale image recognition datasets such as ImageNet and MS CO-CO show that YOLOv7 has a relatively high detection speed and accuracy in the current object detection field in the range of 5-160 frames/s. Although YOLOv8 has higher accuracy, its network structure is deeper and the detection speed is also lower. Therefore, considering detection accuracy and speed comprehensively, YOLOv7 is selected as the basic model.

In response to the problem of low image resolution and difficulty in detecting small targets in brain tumor detection, based on the YOLOv7 algorithm, in order to strengthen the learning of fine-grained and effective features, we will start from three aspects: 1) introducing an SPD module in the Backbone section to eliminate cross row convolution and pooling operations to increase the accuracy of model detection; 2) In the Head section, introduce CA attention mechanism to enhance the learning of more critical and effective features; 3) Use the dynamic non monotonic frequency modulation loss function Wise-IoU to enhance the detection ability for low-quality samples. The improved YOLOv7 structure is shown in Figure 1, where the improved modules are marked with triangular symbols.

2.1. SPD module

In the detection of low resolution images and small targets, due to the use of cross convolutional and pooling layers in existing convolutional neural networks, the good performance of convolutional neural networks is greatly reduced, resulting in fine-grained information loss and low learning efficiency of feature representation. To address this issue, a convolutional neural module SPD is introduced in the Backbone section to replace each cross convolutional and pooling layer, utilizing the structure of a Space to Depth (SPD) layer and a non step convolutional layer of the SPD module to eliminate the alternation between the original two layers. The SPD module structure is shown in Figure 2.

2.2. Coordinate Attention

As the number of layers in the network increases, the granularity of semantic features will gradually compress, causing some detailed features to be lost, resulting in high similarity of features in each

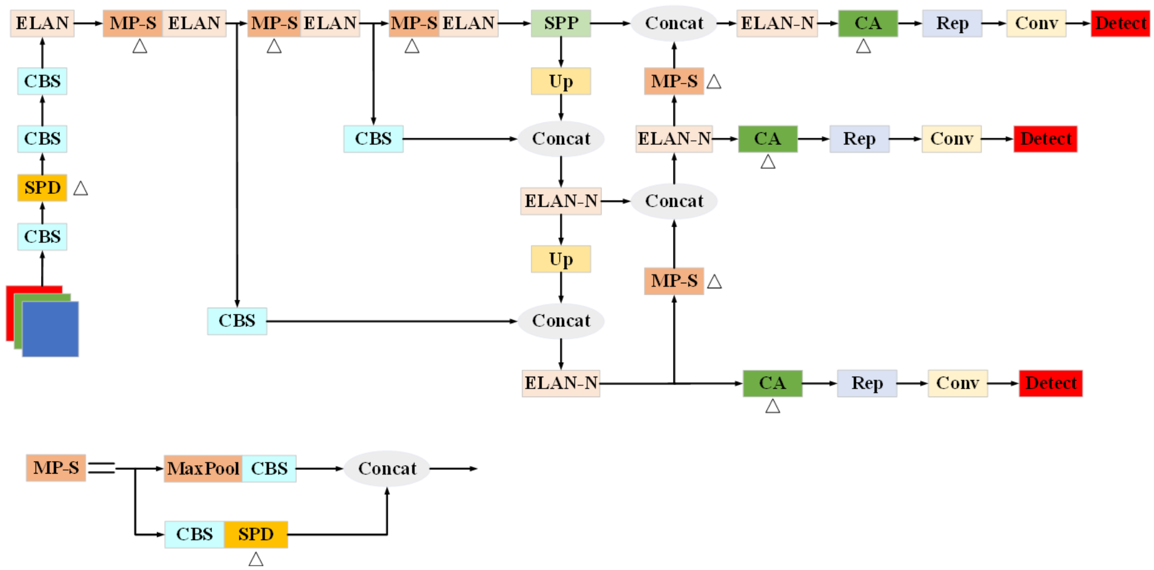

Figure 1: YOLOv7 improved model structure diagram.

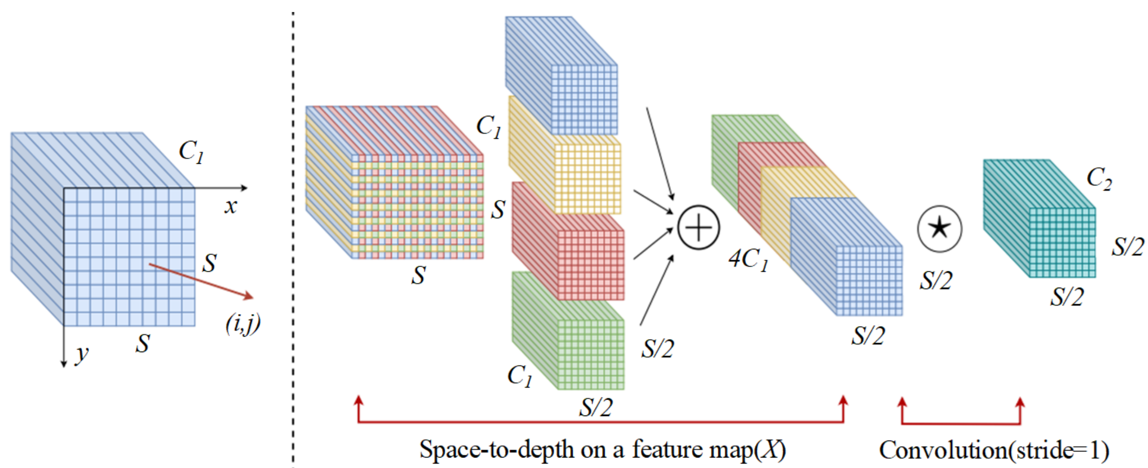

Figure 2: SPD Module structure diagram.

channel, which in turn leads to high model misjudgment rates. To this end, a residual structured CA (Coordinate Attention) module is introduced in the Head section. The CA module structure is shown in Figure 3.

2.3. Wise-IoU loss function

Considering that brain tumor CT data may contain many low-quality samples, in the YOLOv7 original model using the CIoU loss function, geometric metrics such as distance and aspect ratio will intensify the punishment for low-quality samples, leading to a decrease in the generalization performance of the model. Therefore, the Wise-IoU v3 loss function is used to replace the CIoU used in the YOLOv7 original model as the loss function for the brain tumor detection model.

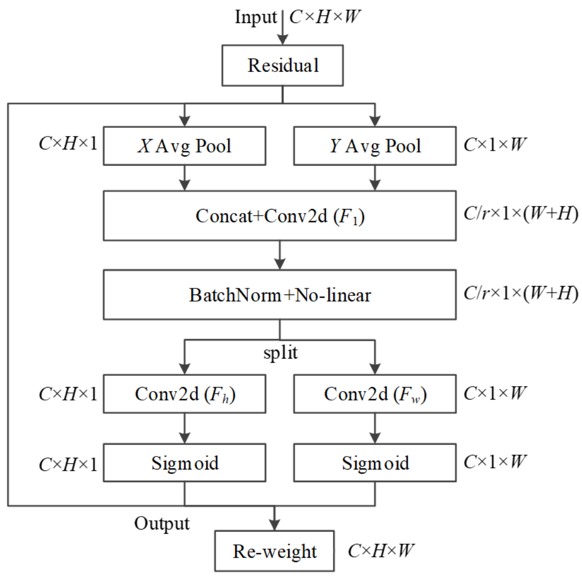

Figure 3: CA Module.

The goal is to adaptively penalize the geometric metrics of low-quality samples, and construct distance attention based on distance metrics, as shown in Equation 1:

$$L_{\mathrm{WIoUv3}} = rR_{\mathrm{WIoU}}L_{\mathrm{IoU}}, \quad r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{1}$$

Among them, R_{WIoU} is the penalty term, as shown in Equation 2.

$$R_{\mathrm{WIoU}} = \exp\left(\frac{(x - x_{\mathrm{gt}})^2 + (y - y_{\mathrm{gt}})^2}{\left(W_g^2 + H_g^2\right)^*}\right) \tag{2}$$

3. Experiment and Result Analysis

3.1. Experimental Environment and Datasets

Experimental hardware configuration: The CPU is Intel i7-10870H, and the GPU is NVIDIA GeForce RTX 2070 Super. Software configuration: Under Ubuntu 18.04 operating system, the Python framework uses version 1.7.0, Python uses version 3.8, and Cuda uses version 11.0.

Only excellent datasets can train detection models with high robustness and generalization. However, due to the fact that the existing brain tumor datasets are single class and have a small number of images, relevant datasets obtained from network sources are added to the existing publicly available brain tumor datasets such as Br35H, and integrated into a three class dataset for detection model experiments. The categories are glioma tumor, pituitary tumor, and meningioma tumor. Finally, 2908 usable brain tumor CT image samples were obtained, which were divided into training, validation, and testing sets in an 8:1:1 ratio.

3.2. Evaluating indicator

The experiment uses Precision (P), Recall (R), and Mean Average Precision (mAP) as the main evaluation indicators of the algorithm, as shown in Equations 3 to 5:

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{3}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{4}$$

$$\mathrm{mAP} = \frac{1}{m} \sum \mathrm{AP}\,(i) \tag{5}$$

In Equations 3 to 5, TP represents the number of correctly recognized detection boxes, and FP represents the number of incorrectly recognized detection boxes. FN is the number of undetected correct targets, m is the number of detected categories, and AP is the area under the PR curve, reflecting the effectiveness of the model in recognizing a certain category.

3.3. Ablation experiment

In order to enhance the robustness of the model, YOLOv7 with data augmentation was used as the baseline model. The impact of different improvement strategies on the baseline model is shown in Table 1.

Table 1: Comparison of ablation experiments.

+SPD	+CA	+Wise-IoU	P	R	mAP@0.5
			81.0%	83.3%	85.8%
√			82.1%	**84.7%**	87.6%
	√		81.3%	83.5%	86.1%
		√	85.8%	83.2%	86.4%
√	√		84.6%	83.2%	88.9%
√		√	86.3%	82.6%	87.1%
	√	√	85.5%	82.9%	86.6%
√	√	√	**87.9%**	83.7%	**89.3%**

The Table 1 shows that incorporating the SPD and CA modules, along with Wise IoU, significantly improved the mAP of the YOLOv7 model by 1.8%, 1.3%, and 0.4%, respectively, resulting in an overall enhancement of 3.5%. This suggests that the model has been greatly enhanced, making it particularly effective for tasks such as brain tumor detection.

3.4. Comparative experiments on different networks

In order to verify the advantages of the improved YOLOv7 algorithm and compare it with some mainstream object detection models, the currently widely used YOLOv5 algorithm and the Faster R-CNN algorithm that combines speed and accuracy were selected for comparison. Without losing generality, some other algorithms were also selected, such as earlier versions of the YOLO series and SSD algorithm. Set the confidence threshold uniformly to 0.5 and compare their accuracy, recall, mAP, and FPS. The results are shown in Table 2.

Table 2: Comparison of Algorithm Models.

Algorithm	P	R	mAP@0.5	FPS
Faster R-CNN	76.3%	71.3%	75.5%	26
SSD	79.6%	72.9%	82.5%	43
YOLOv1	77.2%	75.2%	79.7%	**84**
YOLOv3	79.3%	81.6%	83.2%	77
YOLOv5	79.2%	83.4%	85.3%	61
YOLOv7	81.0%	83.3%	85.8%	73
Ours	**87.9%**	**83.7%**	**89.3%**	71

From the experimental results in Table 2, The enhanced YOLOv7 model demonstrates superior detection performance for low-resolution and small objects, with improvements of up to 6.9% in accuracy, 0.4% in recall, and 3.5% in mAP over its predecessor. Despite its more intricate architecture leading to a slightly reduced frame rate per second (FPS)—down by just 2 frames from the original model—the trade-off maintains high detection speeds. Compared to other established models such as YOLOv5 and classical approaches like SSD, YOLOv1, YOLOv3, and YOLOv4, the refined YOLOv7 stands out with its notable gains in precision.

3.5. Comparison and Analysis of Results

To compare the detection performance of the improved YOLOv7 model and the original YOLOv7 model more intuitively and conveniently, three images from the test set were randomly selected for detection and visual comparison. The results are shown in Figure 4.

From Figure 4, it can be seen that there are still many false positives and missed detections in the YOLOv7 original model for brain tumor detection, and the confidence level of detection is not high. In the detection results of the improved YOLOv7 model, there are fewer false positives and missed detections, and the confidence level is higher, resulting in more accurate detection. This indicates that the improved YOLOv7 model has stronger feature extraction and analysis capabilities, higher confidence level in detection, and closer detection boxes to reality.

4. Conclusion

This article combines practical application scenarios, based on the characteristics and task requirements of brain tumor detection, introduces the SPD module, and introduce the CA attention mechanism in the Head section. In addition, the model uses the dynamic non monotonic frequency modulation loss function Wise-IoU to enhance its detection ability for low-quality samples. The overall improved network model significantly improves the accuracy of low-quality samples and small object detection. Subsequent research will focus on expanding the learning dataset, reducing the number of model parameters, and further improving the accuracy and speed of model detection.

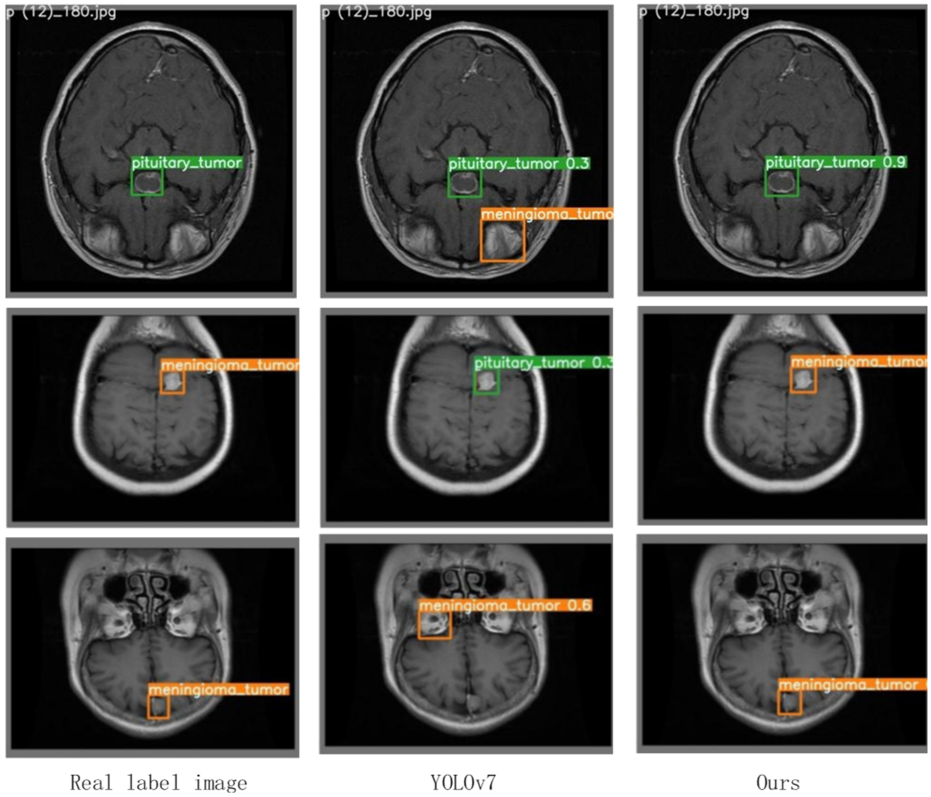

Figure 4: Comparison chart of detection results.

References

Qibin Hou, Daquan Zhou, and Jiashi Feng. Coordinate attention for efficient mobile network design, 2021.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Commun. ACM*, 60(6):84–90, may 2017. doi: 10.1145/3065386.

Yiting Li, Qingsong Fan, Haisong Huang, Zhenggong Han, and Qiang Gu. A modified yolov8 detection network for uav aerial image recognition. *Drones*, 7(5), 2023. doi: 10.3390/drones7050304.

MH Liu and G Zhang. Research progress in brain tumor imaging based on deep learning. *Imaging technology*, 33(5):50–55, 10 2021.

Zanjia Tong, Yuhang Chen, Zewei Xu, and Rong Yu. Wise-iou: Bounding box regression loss with dynamic focusing mechanism, 2023.