Heliyon 10 (2024) e27678

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Image detection of aortic dissection complications based on multi-scale feature fusion

Yun Tan [a], Zhenxu Wang [a], Ling Tan [b,*], Chunzhi Li [a], Chao Deng [b], Jingyu Li [b], Hao Tang [b], Jiaohua Qin [a]

[a] Central South University of Forestry and Technology, Hunan, China
[b] The Second Xiangya Hospital of Central South University, Hunan, China

A R T I C L E I N F O

Index Terms:
Target detection
Convolution neural network
Small target
Multiscale-feature

A B S T R A C T

Background: Aortic dissection refers to the true and false two-lumen separation of the aortic wall, in which the blood in the aortic lumen enters the aortic mesomembrane from the tear of the aortic intima to separate the mesomembrane and expand along the long axis of the aorta.
Purpose: In view of the problems of individual differences, complex complications and many small targets in clinical aortic dissection detection, this paper proposes a convolution neural network MFF-FPN (Multi-scale Feature Fusion based Feature Pyramid Network) for the detection of aortic dissection complications.
Methods: The proposed model uses Resnet50 as the backbone for feature extraction and builds a pyramid structure to fuse low-level and high-level feature information. We add an attention mechanism to the backbone network, which can establish inter-dependencies between feature graph channels and enhance the representation quality of CNN.
Results: The proposed method has a mean average precision (MAP) of 99.40% in the task of multi object detection for aortic dissection and complications, which is higher than the accuracy of 96.3% on SSD model and 99.05% on YoloV7 model. It greatly improves the accuracy of small target detection such as cysts, making it more suitable for clinical focus detection.
Conclusions: The proposed deep learning model achieves feature reuse and focuses on local important information. By adding only a small number of model parameters, we are able to greatly improve the detection accuracy, which is effective in detecting small target lesions commonly found in clinical settings, and also performs well on other medical and natural datasets.

1. Introduction

ORTIC dissection(AD) is a cardiovascular disease that can be caused by an intimal tear or rupture of the nutrient artery in the middle layer of the aorta. This leads to blood entering the middle layer and creating a hematoma that results in excessive pressure, which causes tearing of the intimal layer.The proximity of the false lumen to the native lumen(i.e,the true lumen) makes AD the most complex and dangerous cardiovascular disease [1].Normal vessels and vessels containing aortic dissection are shown in Fig. 1.

Aortic dissection has emerged as a prevalent critical illness and a frequently occurring disease which is marked by sudden onset,

* Corresponding author.
E-mail address: dr.tanling@csu.edu.cn (L. Tan).

https://doi.org/10.1016/j.heliyon.2024.e27678
Received 20 August 2023; Received in revised form 4 March 2024; Accepted 5 March 2024
Available online 15 March 2024
2405-8440/

rapid progression, diverse clinical manifestations, elevated misdiagnosis rate, and high mortality rate.Over the years, owing to the aging population and the escalating incidence of hypertension, the prevalence of aortic dissection has annually amplified [1–3].The detection rate of aortic diseases has significantly increased due to the rapid development of imaging technology.However,due to the complex clinical manifestations of acute aortic dissection, there is a high rate of miss and misdiagnosis.Untreated acute aortic dissection has a high mortality rate, with the mortality rates of 21%,37% and 74% on day 1,day 2 and week 1,respectively.The high mortality rate over a short period of time demonstrates the urgent need for faster and more up-to-date techniques and methods to support physicians in diagnosing the condition.

Aortic dissection is a complex disease that presents numerous grave complications.Some of the complications are identifiable from contrast-enhanced CT images, which include poor renal perfusion, pleural effusion, cyst, intimal tear, and aortic insufficiency,etc.The typical diagnostic approach involves the physician perusing the CT image data for the location of the dissection's origin and analyzing the range, depth, and any abnormal thoracic and abdominal images. The CT image sequences for aortic dissection usually cover hundreds of images from the upper layer of the thoracic cavity to the lower layer of the abdomen. This diagnostic process is time-consuming and labor-intensive.Even experienced and well-trained practitioners find it challenging to determine the location, size, dissection type, the presence of any complications, and the severity of those complications in hundreds of CT images.

In recent years, deep learning has made significant strides in the field of medicine, particularly in auxiliary diagnosis and treatment. For instance, Gardner and colleagues [4] proposed a conditional generative adversarial network that can detect and segment the total tumor volume in images acquired during radiotherapy.

While Gerard et al. [5] proposed Reg3DNet+ residual regression convolutional neural network to directly regress high-resolution images of local tissue volume changes from CT images.Through image registration between total vital capacity and lung volume under functional residual capacity, using tissue mass and structure-preserving registration algorithms, it was possible to obtain a direct estimation of regional lung volume changes from both paired and single CT images.Furthermore, Marios [6] developed a fast proof-of-concept workflow based on deep learning networks for assessing individual organ doses after chest computed tomography (CT) examinations.The study leveraged an independent dataset consisting of 19 patients for each network for organ dose prediction. Huo et al. [7] employed several machine learning algorithms to construct a classification and prediction model for early aortic dissection patients,to facilitate the identification of potential dissection cases among misdiagnosed patients.Among the models built, the Bayesian network achieved an 84.55% accuracy, with an Area Under Curve(AUC) of 85.7%.Singh et al. [8],on the other hand, developed a deep learning model to automatically assess aortic dissection in chest CT images.They trained the model with 4235 images from 30 patients (15 dissection cases and 15 non-dissection cases) and tested it on 3423 images obtained from 40 patients (20 dissection cases and 20 non-dissection cases),using a pretrained inception-v3 [9] network as the classification model.The study demonstrated the efficacy of deep convolutional neural networks in the recognition of aortic dissection, achieving an AUC of 97% and a sensitivity of 100%.In their research, Lee et al. [10] utilized a generative-discriminative learning approach to predict object boundaries in medical image datasets, applying it to true and false lumen segmentation in aortic dissection. Kovács et al. [11] employed De-scoteaux's method to measure the intimal flap in the segmented CT image after segmenting the aortic image. In another study, Duan et al. [12] used the GVF snake model for the segmentation of the descending aorta and utilized the Hessian matrix and the spatial continuity prior model based on Bayesian theory to extract the aortic dissection membrane. Gayhart et al. [13] developed an advanced CAD system for the evaluation of CT images of aortic diseases, who is one of the first researchers to study the issue of aortic dissection detection.Finally, Dehghan et al. [14] adopted an atlas-based approach for aortic segmentation, in which the segmented cross-sectional images were fine-tuned for flap detection and shape analysis.

Zhou and his colleagues [15] proposed a deep learning model as MOLS-Net, for the automatic and efficient segmentation and detection of aortic dissection in CT volumes. The method utilized the pyramid attention module to associate sequence featuresof CT image sequences of different scales, which guided the image segmentation. The model outperformed state-of-the-art methods on multiple datasets, providing an accurate and effective approach to aortic dissection diagnosis. Qin et al. [16] developed an advanced image classification technique that replaced fixed-size inputs with appropriately large ones and used residual modules with less computational cost and parameter inversion.This method performed better than existing methods, reducing network parameters and computational costs. Peng and his colleagues [17] proposed an innovative ROI(regions of interest) extraction algorithm, named RESI,

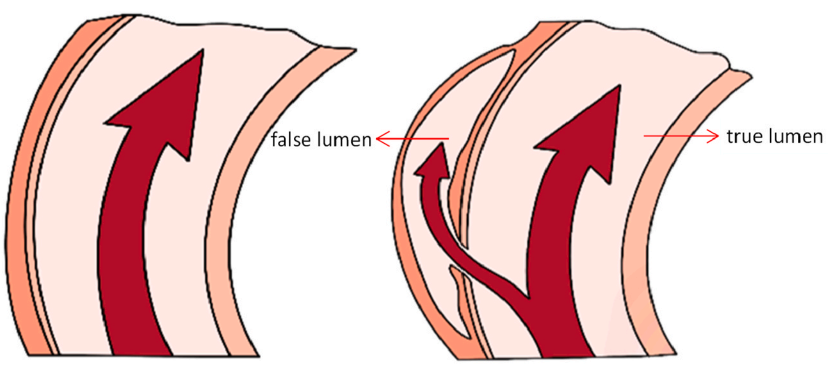

Fig. 1. The left side is a normal blood vessel, and the right side is a blood vessel containing an aortic dissection.

which utilized sequence information for high-precision ROI identification in the first stage and then applied DenseNet-121 for further diagnosis.This technique achieved slice-level accuracy of 97.41%,outperforming state-of-the-art methods.Tan et al. [18] presented a detection method for aortic dissection using CTA images.They extracted ROI through binarization and morphological opening operations, and applied the deep learning model DenseNet-121.The method achieved excellent results for sensitivity and specificity.

In summary, although the current deep learning detection of organs and foci in medical images, including aortic dissection, has reached a high level of accuracy, there is still a gap between it and the actual clinical application, mainly because most of the studies are based on a single coarctation detection, and there is no detection of concurrent multi-complications accompanying the coarctation. Although current methods have achieved promising results on most medical object detection datasets, detecting small lesions in clinical images remains challenging. In practice, the same image may contain dissection, complications and other aberrations, as shown in Fig. 2.Further research is needed to address this issue and improve clinical diagnostic accuracy. Therefore, in this study, we propose a MFF-FPN model for the detection of

aortic dissection and its associated complications using a multi-scale feature fusion and attention mechanism. The contributions of this study include: Firstly, a multi-scale feature fusion pyramid network is built to integrate shallow texture features and high-level semantic information. Secondly, the attention mechanism is introduced within the pyramid to suppress irrelevant information. Finally, the detection module is constructed and the feature map is utilized to identify the location and type of the target. The experimental results demonstrate that the proposed method outperforms existing techniques in the detection of aortic dissection and complications.

2. Related work

2.1. ResNet50 network

The ResNet network model is extensively used in various fields, such as medical image processing and remote sensing image analysis. It encompasses typical network models such as ResNet18, ResNet34, ResNet50, ResNet101and ResNet152. In this study, we use the resnet50 to build the backbone network. The principal structure of ResNet50 is depicted in Fig. 3.

2.2. Feature pyramid network

FPN (Feature Pyramid Network) [19] is aimed at detecting small targets.The high-level features contain rich semantic information, but the resolution is low. If the object detection algorithm only uses high-level features for prediction, the obtained target position is relatively rough. Assuming that in the deep network a pixel in the final high-level feature map may correspond to the 20 × 20 pixels area of the output image, then the features of objects that are smaller than 20 × 20 pixels have a high probability of being lost. At the same time, the low-level feature semantic information is relatively small, but the target position is accurate, which is helpful for small

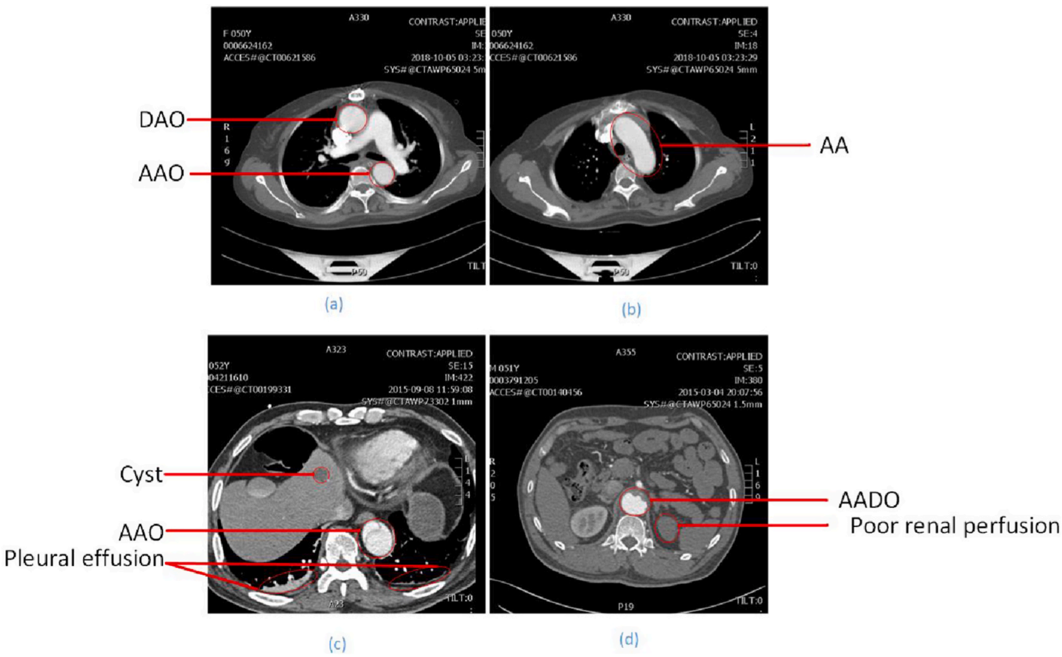

Fig. 2. (a) Image section containing ascending aorta (AAO) and descending aorta (DAO) (b) image section containing aortic arch (AA) (c) image section containing cyst, pleural effusion and AAO (d) image section containing poor renal perfusion and abdominal aorta (AADO).

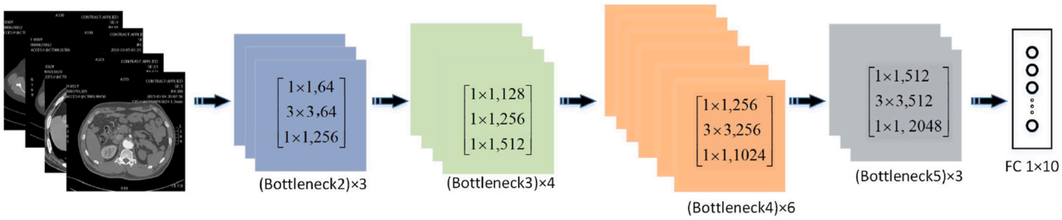

Fig. 3. ResNet50 network.

target detection. FPN fuses high-level features with low-level features, thereby simultaneously utilizing the high-resolution of low-level features and rich semantic information of high-level features, and independently predicting multi-scale features, which significantly improves the detection effect of small objects. In terms of specific operations, the traditional FPN uses C1 to C5 as input features, and the feature map size is reduced to half from the bottom to the top. For example, if the input resolution is 640 × 640, then C3 represents feature level 3 with a resolution of 80 × 80, and C5 represents feature level 5with a resolution of 20 × 20. Traditional FPN aggregates multi-scale features in a top-down manner:

$$P_5 = CONV(C_5) \tag{1}$$

$$P_4 = CONV(C_4 + Resize(P_5)) \tag{2}$$

$$P_2 = CONV(C_2 + Resize(P_3)) \tag{3}$$

Among them, *Resize* is usually an upsampling operation for resolution matching, and *CONV* is usually a convolution operation for feature processing. The traditional pyramid structure is shown in Fig. 4.

3. Proposed detection method for aortic

3.1. Dissection and complications based on multi-scale feature fusion

This study proposes a multi-scale feature fusion based feature pyramid network(MFF-FPN), as shown in Fig. 5, mainly includes input, FPN, ROI pooling layer, RPN module, fully connected layer and classification modules.

The MFF-FPN detection algorithm is mainly divided into two modules: one is the FPN module for feature extraction and the other is the detection module. The feature extraction part consists of a pyramid network, and the detection module is the same as Faster-RCNN [20]. The specific process is: images of any scale are input and image features are obtained through backbone networks subsequently. RPN extracts the feature blocks of the candidate frame from the feature map, and then the ROI pooling layer extracts the feature blocks of the candidate frame and the convolutional layer. The output feature map is finally sent to the fully connected layer, and then is classified and predicted.

3.2. Embedded attention module

According to the characteristics of the feature pyramid network, this study embeds a channel attention block behind the pyramid block to reduce the interference of redundant information of multi-semantic and multi-scale features, and effectively strengthens the network's ability to learn the discriminative features.

The Squeeze-and-Excitation Network(SENet) [21] is a classic attention mechanism and is widely used by industry expert systems.

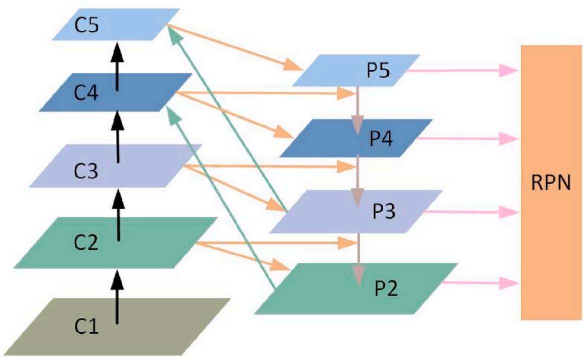

Fig. 4. Traditional feature pyramid network structure.

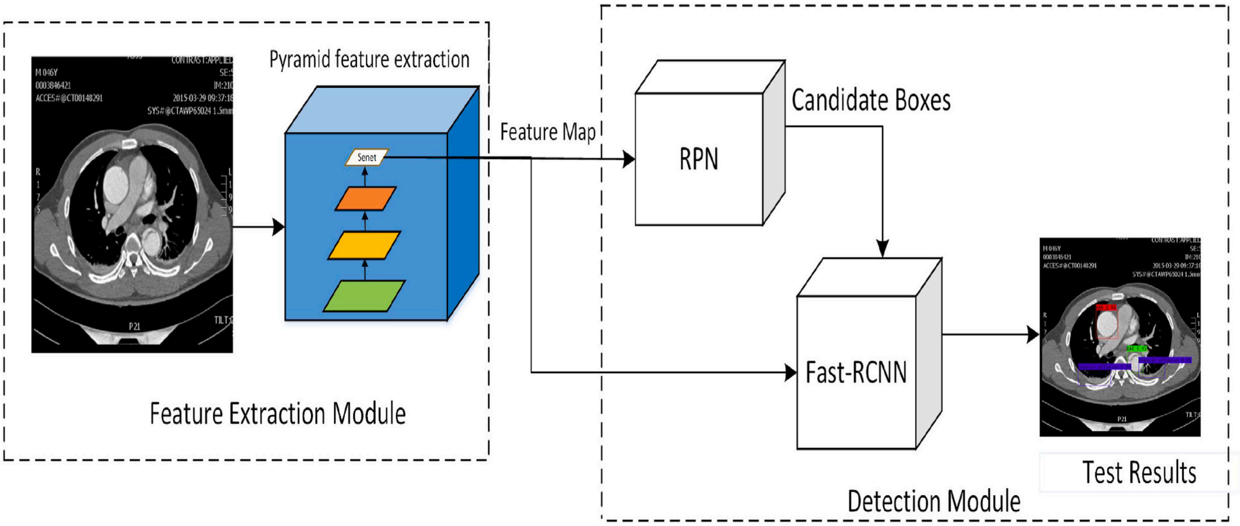

Fig. 5. The structure of MFF-FPN

The role of the SE attention mechanism is to enhance important features, suppress general features and

perform three-step operations on the feature map obtained by convolution, which includes sequeeze, excitation and scale, as shown in Fig. 6.

By compressing the feature map input to the C5 layer in the spatial dimension and performing global average pooling on it, the current global compressed features are obtained, and the overall perception of the aortic dissection image is realized.

Then a fully connected neural network is used to perform a linear transformation on the result after compression. In this way, the result is obtained by using the excitation as a weight and then is multiplied by the feature, which can be dynamically adjusted in the channel dimension and can better handle the mutual correlation of the Resnet-50 channel level. As can be seen from Fig. 6, the input is a C × H × W feature map, (H,W) is the size of the image and C is the number of channels of the image. Firstly, the feature extrusion operation is performed from the feature space dimension of the enhanced CT image, and the C × H × W feature is transformed to the C × 1 × 1 feature. At this time, the number of output feature channels is equal to the number of input feature channels, which describes the overall distribution of features over a channel. Therefore, the obtained one-dimensional feature vector also has a larger global view and contains more information. It is described as:

$$\text{Squeeze} : \text{Zc} = \text{F}_{\text{sq}}\left(\text{u}_c\right) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} u_c(i,j) \tag{4}$$

Where H,W represents the height and width of the input feature map, and $u_c(i,j)$ represents the feature value at the position (i,j).

The second step is excitation. By introducing a learnable W parameter to generate weights for each channel feature, and using a binary classification function to obtain a normalized weight between 0 and 1, the correlation between each feature channel is realized. The equation of the gating unit (the calculation method representing the eigenvector of 1*1*C in Fig. 5) is as follows:

$$\text{Excitation} : \text{s} = F_{ex}(z, W) = \sigma(W_2\delta(W_1 z)) \tag{5}$$

Where δ represents the ReLU activation function and σ represents the sigmoid activation function. $W_1 \in R^\wedge(C/r \times c)$, and $W_2 \in R^\wedge(C/r \times c)$ respectively represent the weight matrix of the two fully connected layers, and z is the image feature.

The last step is a scale operation, which regards the excitation output as the importance of each channel to the feature map, and then weights the previous features through point multiplication between channels. As a benchmark, the re-correction of channel weights is realized. It is described as:

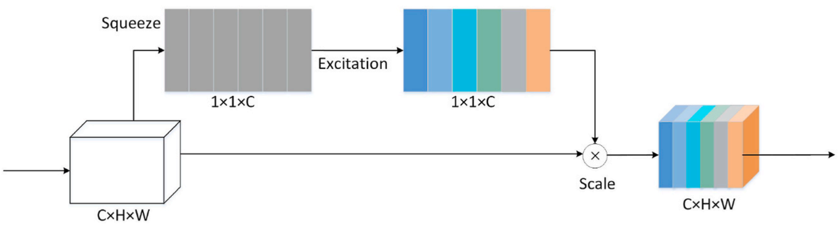

Fig. 6. SENet structure.

$$\text{Scale}: \widetilde{X}_c = \text{Fscale}(u_c, s_c) = s_c \cdot u_c \tag{6}$$

Where $\widetilde{X}_c$ is a featured channel belonging to $\widetilde{X}$, s_c is the generated channel weight, and u_c is the input of the convolutional layer.

3.3. Proposed multi-scale feature fusion

To improve the detection capability of small targets and to fuse more features without consuming more cost, we construct an FPN network based on ResNet50.

The ResNet50 network structure contains five stages, which are noted here as C1,C2,C3,C4,C5 and this study mainly focuses on the improvement of C4 and C5. Firstly, the feature pyramid network is used to extract the feature hierarchy. The P2 and P3 feature information of the bottom layer is down-sampled by a convolutional kernel of size 3×3. The acquired features are added to C4 and C5 layers respectively, and the number of channels is adjusted by 1×1 convolution to achieve the fusion of shallow features and high-level semantic features. The top-down

and bottom-up paths are fused into a module so that it can be repeatedly stacked to achieve higher-level feature fusion, which is the reuse of features. The pyramid aggregates the bottom-level features as:

$$P5 = CONV(\text{B5} + \text{Resize}(\text{P3})) \tag{7}$$

$$P4 = CONV(Upsampling(B5) + C4 + Resize(P2)) \tag{8}$$

Where *Upsampling* denotes the upsampling operation for resolution matching and *CONV* is the convolution operation for feature processing. The specific fusion method is shown in Fig. 7 below.

4. Experiments and result analysis

4.1. Dataset

The CT image dataset used in the experiment is provided by the cooperative hospital. The dataset consists of 105 patients and non-patients with angiographic contrast agent-enhanced sequential images. The CT image dataset includes Stanford type A dissection involving the ascending aorta, Stanford type B involving only the descending thoracic aorta and its distal end, and images of healthy patients without active dissection. The patients' ages range from 36 to 85 years old, with an average age of 46.7 years, and mainly concentrate between 40 and 60 years old. The CT image dataset includes 37 males and 35 females. We select 3356 images including the chest cavity and abdomen. The image resolution is 512×512, the plane pixel spacing is 0.5–0.86 mm, and the slice spacing is 3 mm. The dataset includes three complications of pleural effusion, cyst and poor renal perfusion. In addition, considering the detection of the ascending aorta, aortic arch, descending aorta and abdominal aorta can determine the type of aortic dissection and guide clinical treatment, we also detect the ascending aorta, aortic arch, descending aorta and abdominal aorta, which are seven types of targets in total. 70% is used for training, 15% is used for validation and 15% is used for test. We use the StandardScaler class to implement the normalization of the data, thus the preprocessing of feature data is realized, which improves the speed of training and accelerates its convergence.

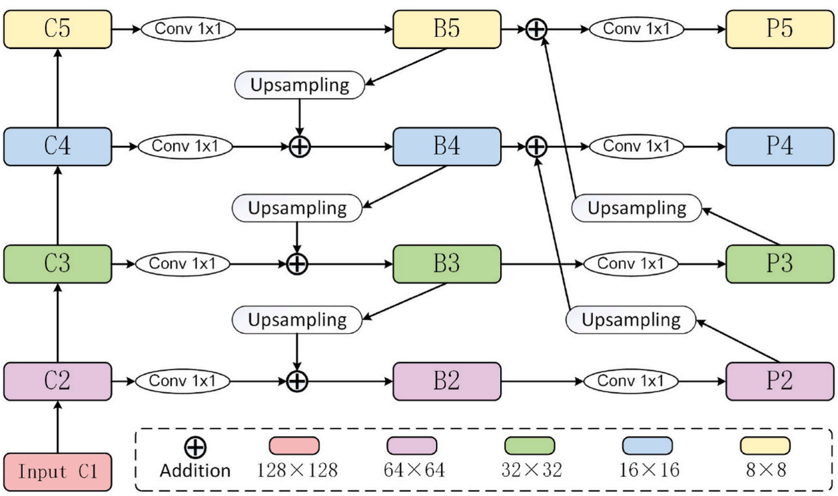

Fig. 7. Multi-scale feature fusion pyramid.

4.2. Experimental hardware and software environment

The operating system used in this experiment is Windows 10 system. The CPU model is AMD Ryzen 3 desktop Processor. The GPU is NVIDIA GeForce RTX3060 graphics card and the running memory size is 24 GB DDR4. This model is based on the PyTorch deep learning framework. The programming language is Python, using CUDA11.0 and CUDNN 8.0.5 to accelerate GPU.

4.3. Evaluation indicators

Target detection usually uses precision, recall, Average Precision (AP) and Mean Average Precision (mAP) as the evaluation indicators. The AP value is the area enclosed by the curve drawn with the precision rate on the abscissa and the recall rate on the ordinate. The definitions of precision rate and recall rate are shown in Equ.(9) and Equ.(10):

$$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{10}$$

Among them, TP (True Positive) is an positive example and is considered as positive, FP (False Positive) is a negative example but is considered as positive and FN (False Negative) is a positive sample but is considered as negtive.

AP and mAP are used to represent the average accuracy of a single category and all categories respectively, and mAP is the average value of various APs when the intersection and union ratio threshold of the real box and the actual box is set to 0.5. It is described as:

$$AP = \int_0^1 P(R)dR \tag{11}$$

$$mAP = \frac{1}{C}\sum_{i=1}^{C} AP_i \tag{12}$$

P represents the precision rate and R represents the recall rate. C represents the total number of categories and AP_i represents the AP value of the ith category.

4.4. Analysis of experimental results

1) *Accuracy:*Taking AP and mAP as the evaluation index of the model, to verify the advancement and effectiveness of the designed model, comparative experiments have been carried out on YoloV5 [22], YoloV7 [23], SSD [24] and the proposed MFF-FPN. The AP results are shown in Table 1, which includes seven types of targets (AAO,AA,DAO,AADO, Poor perfusion, Pleural effusion and Cysts).

Through the comparative experiments, it can be seen that the proposed network is very effective in improving the detection effect. Compared with the SSD network, the accuracy rate has increased by 3.1%, and it has increased by 2.3% compared with the classic YoloV5. Among them, the AP of cysts with a small area is 100%, which is 19.39% higher than the SSD network, and 16.20% higher than the classic YoloV5 model, which significantly improves the accuracy of small target detection and achieves good detection results. The validity of the proposed model in this study is confirmed.

2) *Model convergence:* The network training takes about 200 h in total and iterates 500 times. During this process, the accuracy and loss of model tends to be stable. It can be seen that the proposed network is effective for learning the features of aortic dissection complications. Fig. 8 shows the curves of the accuracy rate (mAP) on the left, while loss value and learning rate in the training and verification stages are shown on the right.

Table 1
AP of different models.

Networks	SSD	YoloV5	Yolo V7	MFF-FPN
AAO	98.98%	99.10%	**100%**	**100%**
AA	96.04%	99.60%	**100%**	**100%**
DAO	**100%**	99.30%	**100%**	**100%**
AADO	**100%**	99.50%	99.96%	**100%**
Poor perfusion	**99.12%**	99.10%	98.70%	97.22%
Pleural effusion	98.77%	99.50%	**100%**	**100%**
Cysts	80.61%	83.80%	94.70%	**100%**
mAP	96.30%	97.10%	99.05%	**99.40%**

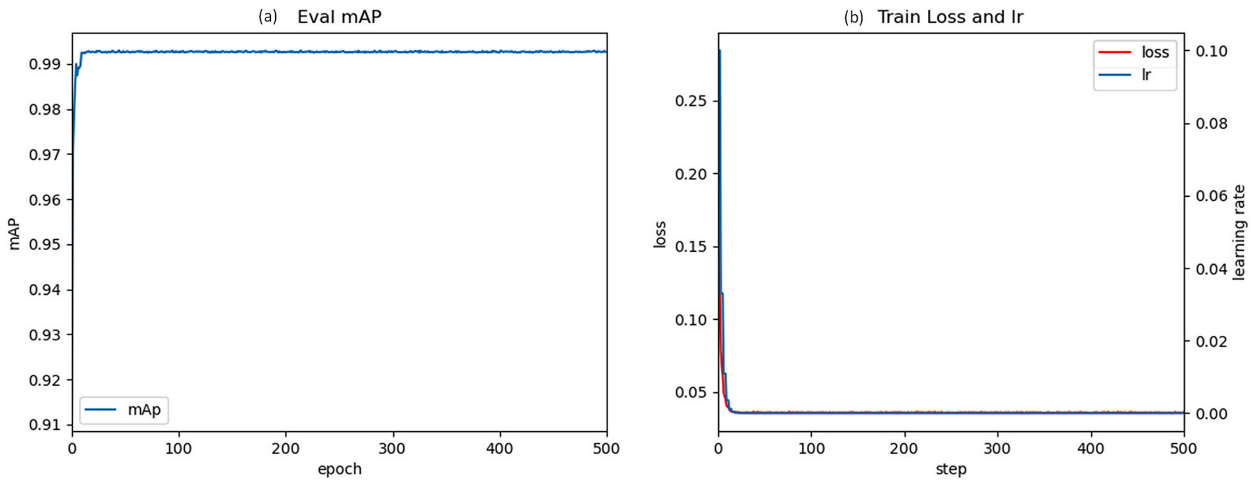

Fig. 8. The left side is the curve of mAP;the right side is the curve of loss and learning rate.

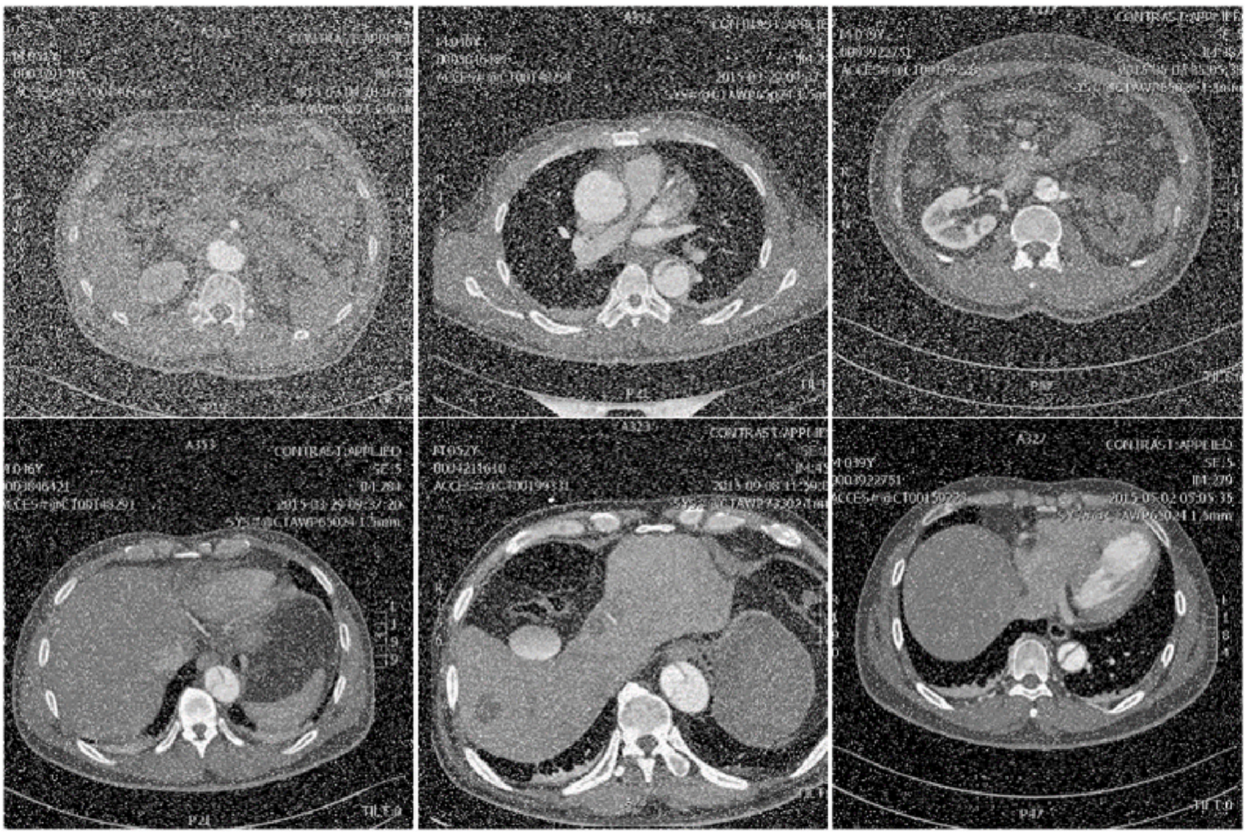

Fig. 9. Images after adding noise.

3) *Model generalization:*To demonstrate the effectiveness of the model on other medical dataset, experiments are conducted on the pulmonary nodule dataset [25]. After processing, 2000 labeled images of pulmonary nodules were obtained. The experimental results are shown in Table 2.

To demonstrate the extensive effects of the model, experiments are conducted on the PASCAL VOC2007 [26] dataset at the same time, and the experimental results are shown in Table 3.

Through the comparison of experiments, it can be seen that the proposed network not only has excellent detection ability in medical

Table 2
Experimental results on the pulmonary nodule dataset.

Networks	mAP(IOU = 0.50)
SSD	97.5%
YoloV5	97.2%
YoloV7	96.4%
MFF-FPN(ours)	**97.7%**

Table 3
Experimental results of VOC2007 dataset.

Networks	mAPIoU = 0.50
Faster R–CNN	69.9%
SSD300	68.0%
SSD512	71.6%
YoloV5	79.4%
YoloV7	82.3%
MFF-FPN(ours)	85.**2%**

images but also has obvious advantages in natural image processing.

4) *Ablation experiment:*The experimental results are shown in Table 4, in which Base means Faster-RCNN with FPN, Improvement 1 means only adding the attention mechanism, but not modifying the pyramid structure network and

Improvement 2 means only modifying the pyramid structure, but not adding the attention mechanism. The ablation comparison experiment is to verify the optimization effect of each improved module.

Because the sizes of cyst in the dataset images are all under 1 cm, we regard the cyst detection results as small target detection results. The detection results for small cyst targets are shown in Table 5. It can be seen that after adding the attention mechanism, the average precision has increased by 28.61%, indicating that the attention mechanism has effectively improved the accuracy of feature extraction. The average precision of feature fusion has increased by 32.95%. When these two improvements are added to the model, the average precision is increased by 37.79%, which greatly improves the detection of small and dense targets.

5) *Model robustness:*To verify the robustness of the pro-posed model, salt and pepper noise is added to the aortic dissection dataset as shown in Fig. 9. SSD, YoloV5 and YoloV7 are compared with the proposed MFF-FPN. In addition, to verify the advantages of the proposed model in small target detection, the detection results of cysts are also listed separately. The detection results after adding noise to the input images are shown in Table 6.

It is shown that the accuracy difference obtained by using the MFF-FPN before and after noise addition is 1.2%, the accuracy difference is 15.8% on YoloV5 and is 1.2% on YoloV7. The difference of the proposed model is significantly smaller than the other models and the accuracy is also significantly higher than YoloV7, indicating that the proposed model is more robust.

5. Conclusion

The MFF-FPN presented in this study addresses the issue of poor detection accuracy of small targets in clinical imaging diagnosis. It significantly enhances the detection accuracy of the network, particularly for small targets such as cysts and poor perfusion. It also effectively improves the recognition performance of the model between similar categories. Importantly, the proposed MFF-FPN not only excels in the detection of aortic dissection and complications but also demonstrates inspiring results in the detection of other clinical lesions and ordinary detection characteristics.

Although the proposed network can be applied to the detection of aortic dissection complications on enhanced CT, there are still some limitations. The dataset used to train and validate the model is from the same hospital, and the generalization of the network on images from different imaging devices has not yet been verified. In a follow-up study we will collect images on different imaging devices to validate the generalizability. At the same time, the proposed model has high accuracy but the speed of detection need further improvement.

Aortic dissection has a rapid onset, in order to meet the actual clinical requirements, deep learning will definitely combine the automatic generation of medical diagnostic reports and lesion detection in the future. Going forward, we intend to carry out research on faster, smarter models, and generating diagnostic report models in conjunction with detection results.

Data availability statement

Two public image datasets are used to support this study and are available at (PASCAL VOC2007:VOC/voc2007/index.html" title

Table 4
Results of ablation experiment.

Networks	attention mechanism	Feature Fusion	mAP/%
Base	×	×	92.90%
Improved model 1	✓	×	97.00%
Improved model 2	×	✓	98.80%
MFF-FPN(ours)	✓	✓	**99.40%**

Table 5
Detection results of small targets.

Networks	attention mechanism	Feature Fusion	Cysts mAP
Base	×	×	62.21%
Improved model 1	✓	×	90.82%
Improved model 2	×	✓	95.16%
MFF-FPN (ours)	✓	✓	**100%**

Table 6
Test results after adding salt and pepper noise.

Networks	mAP(IoU = 0.50)	Original image mAP	difference
SSD	94.0%	96.30%	−2.3%
Yolo v5	81.3%	97.10%	−15.8%
Yolo v7	97.8%	99.05%	−1.2%
MFF-FPN(ours)	98.3%	99.40%	−1.2%

= "http://host.robots.ox.ac.uk/pascal/VOC/voc2007/index.html">http://host.robots.ox.ac.uk/pascal/VOC/voc2007/index.html; pulmonary nodule dataset: Dolejsi M, Kybic J, Polovincak M et al. The Lung TIME—Annotated Lung Nodule Dataset and Nodule Detection Framework [C]//SPIE Conference on Medical imaging, 2009).

Ethics statement statement

All participants and patients provided informed consent to participate in the study.

CRediT authorship contribution statement

Yun Tan: Writing – review & editing, Conceptualization. **Zhenxu Wang:** Software. **Ling Tan:** Conceptualization. **Chunzhi Li:** Conceptualization. **Chao Deng:** Data curation. **Jingyu Li:** Resources. **Hao Tang:** Validation, Supervision. **Jiaohua Qin:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China (No. 62002392); in part by the Key Research and Development Plan of Hunan Province (No. 2019SK2022); in part by the Natural Science Foundation of Hunan Province (No. 2022JJ31019).

References

[1] D. Pacini, L. Di Marco, D. Fortuna, L.M.B. Belotti, D. Gabbieri, C. Zussa, F. Pigini, A. Contini, M.C. Barattoni, R.a. De Palma, Acute aortic dissection: epidemiology and outcomes, Int. J. Cardiol. 167 (2013) 2806–2812.
[2] J.S. Lindholt, S. Juul, H. Fasting, E.W. Henneberg, [Hospital costs and benefits of screening for abdominal aortic aneurysm. Results from a randomized screening trial], Eur. J. Vasc. Endovasc. Surg. 23 (2002) 55–60.
[3] J.A. Elefteriades, Thoracic aortic aneurysm: reading the enemy's playbook, Curr. Probl. Cardiol. 33 (2008) 203–277.
[4] M. Gardner, Y.B. Bouchta, A. Mylonas, M. Mueller, C. Cheng, P. Chlap, R. Finnegan, J. Sykes, P.J. Keall, D.T. Nguyen, Realistic CT data augmentation for accurate deep-learning based segmentation of head and neck tumors in kV images acquired during radiation therapy, Med. Phys. 50 (7) (2023) 4206–4219.
[5] S.E. Gerard, M.F. Chaudhary, J. Herrmann, G.E. Christensen, R.S.J. Estepar, J.M. Reinhardt, E.A. Hoffman, Direct estimation of regional lung volume change from paired and single CT images using residual regression neural network, Med. Phys. 50 (9) (2023) 5698–5714.

[6] M. Myronakis, J. Stratakis, J. Damilakis, Rapid estimation of patient-specific organ doses using a deep learning network[J], Med. Phys. 50 (11) (2023) :7236–7244.
[7] D. Huo, B. Kou, Z. Zhou, M. Lv, A machine learning model to classify aortic dissection patients in the early diagnosis phase, Sci. Rep. 9 (2019) 2701.
[8] V. Singh, R. Gorniak, A. Flanders, P. Lakhani, Abstract No. 606 Automated assessment of acute aortic dissection on thoracic CT using deep learning, J. Vasc. Intervent. Radiol. 30 (2019) S261–S262.
[9] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2818–2826.
[10] N. Lee, H. Tek, A.F. Laine, True-false lumen segmentation of aortic dissection using multi-scale wavelet analysis and generative- discriminative model matching, in: Medical Imaging 2008: Computer- Aided Diagnosis, 6915, SPIE, 2008, pp. 878–888.
[11] T. Kovacs, P. Cattin, H. Alkadhi, S. Wildermuth, G. Szekely, Au- tomatic segmentation of the aortic dissection membrane from 3D CTA images, in: Medical Imaging and Augmented Reality: Third International Workshop, 3, Springer, Shanghai, China, 2006, pp. 317–324. August 17-18, 2006 Proceedings.
[12] D. Xiaojie, S. Meichen, W. Jianming, Z. He, C. Dandan, Segmenta- tion of the aortic dissection from CT images based on spatial continuity prior model, in: 2016 8th International Conference on Information Technology in Medicine and Education (ITME), IEEE, 2016, pp. 275–280.
[13] M. Gayhart, H. Arisawa, Automated detection of healthy and diseased aortae from images obtained by contrast-enhanced CT scan, Comput. Math. Methods Med. 2013 (2013).
[14] E. Dehghan, H. Wang, T. Syeda-Mahmood, Automatic detection of aortic dissection in contrast-enhanced CT, in: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), IEEE, 2017, pp. 557–560.
[15] Q. Zhou, J. Qin, X. Xiang, Y. Tan, Y. Ren, MOLS-Net: multi-organ and lesion segmentation network based on sequence feature pyramid and attention mechanism for aortic dissection diagnosis, Knowl. Base Syst. 239 (2022) 107853.
[16] J. Qin, W. Pan, X. Xiang, Y. Tan, G. Hou, A biological image classification method based on improved CNN, Ecol. Inf. 58 (2020) 101093.
[17] H. Peng, Y. Tan, H. Tang, L. Tan, X. Xiang, Y. Wang, N.N. Xiong, Aortic dissection diagnosis based on sequence information and deep learning, Comput. Mater. Continua (CMC) 73 (2022) 2757–2771.
[18] Y. Tan, L. Tan, X. Xiang, H. Tang, J. Qin, W. Pan, Automatic detection of aortic dissection based on morphology and deep learning, Comput. Mater. Continua (CMC) 62 (2020) 1201–1215.
[19] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, S. Belongie, Feature pyramid networks for object detection, in: Proceedings of theIEEE conference on computer vision and pattern recognition, 2017, pp. 2117–2125.
[20] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: towards real-time object detection with region proposal networks, Adv. Neural Inf. Process. Syst. 28 (2015).
[21] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141.
[22] https://github.com/ultralytics/yolov5github.com.
[23] C.Y. Wang, A. Bochkovskiy, H.Y.M. Liao, YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, in: Proceedings Of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 7464–7475.
[24] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, A.C. Berg, Ssd: single shot multibox detector, in: Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, 14, Springer, 2016, pp. 21–37.
[25] M. Dolejsi, J. Kybic, M. Polovincak, et al., The lung TIME—annotated lung nodule dataset and nodule detection Framework, in: SPIE Conference On Medical Imaging, 2009.
[26] http://host.robots.ox.ac.uk/pascal/VOC/voc2007/index.html.