

## 基于改进Faster R-CNN的红外目标检测算法

汪西晨 彭富伦 李业勋 张俊举

### Infrared target detection algorithm based on improved Faster R-CNN

WANG Xichen, PENG Fulun, LI Yexun, ZHANG Junju

引用本文:

汪西晨, 彭富伦, 李业勋, 等. 基于改进Faster R-CNN的红外目标检测算法[J]. 应用光学, 2024, 45(2): 346–353. DOI: 10.5768/JAO202445.0202001

WANG Xichen, PENG Fulun, LI Yexun, et al. Infrared target detection algorithm based on improved Faster R-CNN[J]. *Journal of Applied Optics*, 2024, 45(2): 346–353.  
DOI: 10.5768/JAO202445.0202001

在线阅读 View online: <https://doi.org/10.5768/JAO202445.0202001>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于嵌入式GPU的红外弱小目标检测算法

Infrared weak small target detection algorithm based on embedded GPU

应用光学. 2020, 41(5): 1089–1095 <https://doi.org/10.5768/JAO202041.0506004>

#### 复杂背景下红外图像弱小目标检测

Dim-small targets detection of infrared images in complex backgrounds

应用光学. 2021, 42(4): 643–650 <https://doi.org/10.5768/JAO202142.0402002>

#### 基于双路多尺度金字塔池化模型的显著目标检测算法

Salient target detection algorithm based on dual-channel multi-scale pyramid pooling model

应用光学. 2021, 42(6): 1056–1061 <https://doi.org/10.5768/JAO202142.0602007>

#### 基于注意力机制与图卷积神经网络的单目红外图像深度估计

Depth estimation of monocular infrared images based on attention mechanism and graph convolutional neural network

应用光学. 2021, 42(1): 49–56 <https://doi.org/10.5768/JAO202142.0102001>

#### 一种基于2D-CNN的激光超声表面缺陷检测方法

Laser ultrasonic surface defects detection method based on 2D-CNN

应用光学. 2021, 42(1): 149–156 <https://doi.org/10.5768/JAO202142.0107002>

#### 多维特征点空间的红外弱小目标检测方法

Multi-dimensional feature point space infrared dim target detection method

应用光学. 2020, 41(6): 1268–1276 <https://doi.org/10.5768/JAO202041.0606002>



关注微信公众号，获得更多资讯信息

文章编号:1002-2082(2024)02-0346-08

# 基于改进 Faster R-CNN 的红外目标检测算法

汪西晨<sup>1</sup>, 彭富伦<sup>2</sup>, 李业勋<sup>3</sup>, 张俊举<sup>1</sup>

(1. 南京理工大学 电子工程与光电技术学院, 江苏 南京 210094; 2. 西安应用光学研究所, 陕西 西安 710065;  
3. 江苏北方湖光光电有限公司, 江苏 无锡 214100)

**摘要:** 为提升红外目标的检测精度, 提出了一种引入频域注意力机制的 Faster R-CNN 红外目标检测算法。首先, 针对红外图像边缘模糊和噪声问题, 设计了一种并行的图像增强预处理结构; 其次, 在 Faster R-CNN 中引入频域注意力机制, 设计了一种新型红外目标检测主干网络; 最后, 引入路径增强金字塔结构, 融合多尺度特征进行预测, 利用底层网络丰富的位置信息, 提升检测精度。在红外飞机的数据集上进行实验, 结果表明, 改进后的 Faster R-CNN 目标检测框架比以 ResNet50 为主干的算法的 AP 提升了 7.6%。此外, 与目前主流算法对比, 本文算法提高了红外目标的检测精度, 验证了算法改进的有效性。

**关键词:** 红外目标检测; 图像增强; Faster R-CNN; 频域注意力机制; 多尺度特征融合

中图分类号: TN206; TP391.4

文献标志码: A

DOI: [10.5768/JAO202445.0202001](https://doi.org/10.5768/JAO202445.0202001)

## Infrared target detection algorithm based on improved Faster R-CNN

WANG Xichen<sup>1</sup>, PENG Fulun<sup>2</sup>, LI Yexun<sup>3</sup>, ZHANG Junju<sup>1</sup>

(1. School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094,  
China; 2. Xi'an Institute of Applied Optics, Xi'an 710065, China; 3. Jiangsu North Huguang Photoelectric  
Co.,Ltd., Wuxi 214100, China)

**Abstract:** In order to improve the detection accuracy of infrared targets, a Faster R-CNN infrared target detection algorithm introducing a frequency domain attention mechanism was proposed. Firstly, a parallel image enhancement preprocessing structure was designed to address the issues of edge blur and noise in infrared images. Secondly, a frequency domain attention mechanism was introduced into Faster R-CNN, and a new infrared target detection backbone network was designed. Finally, a path enhanced pyramid structure was introduced to fuse multi-scale features for prediction, and the rich location information of the underlying network was utilized to improve detection accuracy. The experiment was conducted on a dataset of infrared aircraft. The results show that the AP of improved Faster R-CNN target detection framework is 7.6% higher than that of the algorithm with ResNet50 as the main stem. In addition, compared with current mainstream algorithms, the proposed algorithm improves the detection accuracy of infrared targets and verifies the effectiveness of the algorithm improvement.

**Key words:** infrared target detection; image enhancement; Faster R-CNN; frequency domain attention mechanism; multi-scale feature fusion

## 引言

由于红外成像技术依赖目标与环境间的红外

辐射差异, 不受天气和光照的影响, 能够在夜间、

浓雾等可见光设备无法准确成像的环境下对目标

收稿日期:2023-04-11; 修回日期:2023-08-10

基金项目:国家自然科学基金 (61971386)

作者简介:汪西晨 (1999—), 女, 硕士研究生, 主要从事深度学习图像研究。E-mail: [2241495795@qq.com](mailto:2241495795@qq.com)

通信作者:张俊举 (1979—), 男, 博士, 副教授, 主要从事光电信息探测、图像信号处理和光电成像器件及应用研究。

E-mail: [zj\\_w1231@163.com](mailto:zj_w1231@163.com)

进行成像,因此被广泛应用于军事领域对敌军的检测任务中。与可见图像相比,红外图像分辨率低、边缘模糊,且缺少颜色和纹理信息,给红外目标检测带来困难,因此传统依赖手工设计特征的目标检测算法并不适用于红外目标的检测任务。深度学习和卷积神经网络的发展有目共睹,为将可见光领域的目标检测算法移植到红外领域的研究打开了新思路。对红外领域深度学习算法的研究具有重要的实用价值。

2020年,曹红燕等<sup>[1]</sup>通过增加多尺度特征提取的方式,在原有YOLOv3算法基础上,提高了对远距离红外目标的识别精度,同时融合BN(batch normalization)层和卷积层,将目标检测精度从66%提升至88%。2021年,顾佼佼等<sup>[2]</sup>使用改进的Faster R-CNN算法在红外船舶数据集上进行实验,通过引入多尺度融合和改进损失函数等方式,使模型精度提升了3.95%。同年,蔡伟等<sup>[3]</sup>提出一种YOLO-FCSP目标检测框架,减少原有YOLO(you only look once)下采样次数,结合跨阶段局部网络、多路径聚合,对红外弱小目标的检测精度达到92.6%。

注意力机制的引入是卷积神经网络发展过程中又一关键技术。注意力机制可以解释为一种对输入重要性动态分配权重的过程,给较重要的信

息分配较大的权重。2022年,谌海云等<sup>[4]</sup>设计了一种有效空间通道混合注意力模块,引入至YOLOX算法检测框架中,同时改进特征提取模块,在FLIR红外数据集上进行实验,模型检测精度相较于基准YYOLOX提升了5.04%。

目前,根据是否需要提取区域候选预测框,将深度学习的目标检测框架分为两种:两阶段(two-stage)和单阶段。Faster R-CNN、Mask R-CNN、Cascade R-CNN<sup>[5-7]</sup>等是两阶段目标检测框架;单阶段目标检测框架代表为YOLO<sup>[8-10]</sup>系列、RetinaNet<sup>[11]</sup>等。Faster R-CNN目标检测框架是经典的两阶段目标检测框架。相比于单阶段检测算法,两阶段检测框架检测精度相对较高。

针对以上研究,采用Faster R-CNN两阶段红外目标检测框架,设计图像增强模块,引入注意力机制和多尺度融合改进算法网络结构,使算法适用于红外飞机数据集的检测。

## 1 Faster R-CNN 原理

两阶段检测框架也可称为基于区域候选的检测框架,包括区域预选框(Proposal)的提取和边界框回归预测两个阶段。Faster R-CNN算法原理如图1所示。

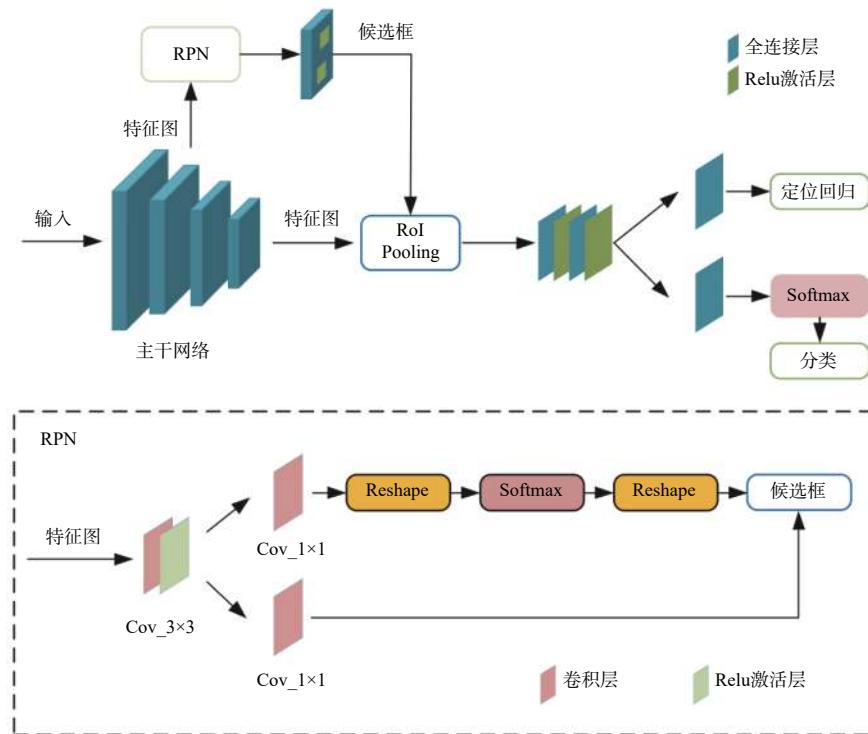


图1 Faster R-CNN 算法结构

Fig. 1 Structure diagram of Faster R-CNN algorithm

如图 1 所示, 输入图像经过主干网络提取特征得到特征图, 传入区域候选提取网络 (region proposal network, RPN), 输出得到候选框; 随后, 将生成的候选框和主干提取的特征图一齐输入到 RoI Pooling 层, 获得每个候选框的特征图, 并将这些特征图固定到  $7 \times 7$  像素的大小, 输入到全连接层获得最后的结果。

## 2 改进的 Faster RCNN 算法

针对红外目标检测, 对算法进行了以下设计。算法主要包括 4 个部分: 图像预处理、特征提取主干、多尺度特征提取颈部及检测头。首先, 算法对输入图片进行预处理; 随后输入到 FCA-ResNet-50 中提取特征图; 然后在颈部网络 PAFPN 中提取不同尺度的特征预测框; 接着使用 RoI Align<sup>[6]</sup> 将预测框的特征缩放到  $7 \times 7$  像素大小, 最后进行分类和回归。

### 2.1 图像预处理

红外图像大多伴随着分辨率不高、边缘模糊的问题, 尤其是飞机在云层中飞行时, 飞机的红外成像与云层的红外成像边缘难以辨别。针对以上问题, 在图片输入到特征提取网络之前设计一个如图 2 所示的图像预处理增强模块, 通过将输入图像分别经过并行的直方图均衡处理和伽马变换与原图拼接在一起, 构成与可见光形式相同的三通道图像, 随后输入到神经网络中。将图像预处理算法放置在检测算法整体设计中, 使得图像增强和检测训练能够端到端的实现。

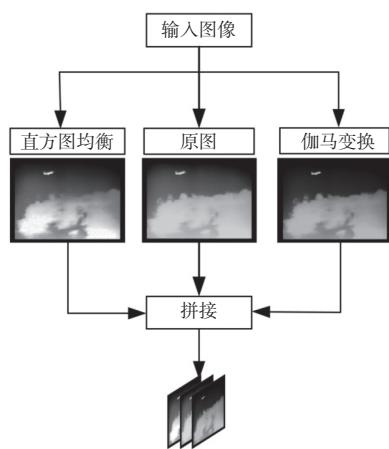


图 2 图像预处理结构

Fig. 2 Structure diagram of image preprocessing

如图 3(b) 所示, 将经过直方图均衡化的图像与原图进行对比, 可以看出, 经过直方图均衡处理

后, 图像对比度显著增强, 同时目标边缘更明确, 但目标对比度不自然。除了直方图均衡处理外, 采用伽马变换对原图进行处理, 如图 3(c) 所示。经过伽马变换后, 图像对比度提升, 图像更清晰, 边缘模糊的问题得到较大改善。视觉上直方图均衡让亮部变更亮, 伽马变换让暗部变更暗, 以此来增加图像对比度, 达到降噪和边缘模糊处理的目的。

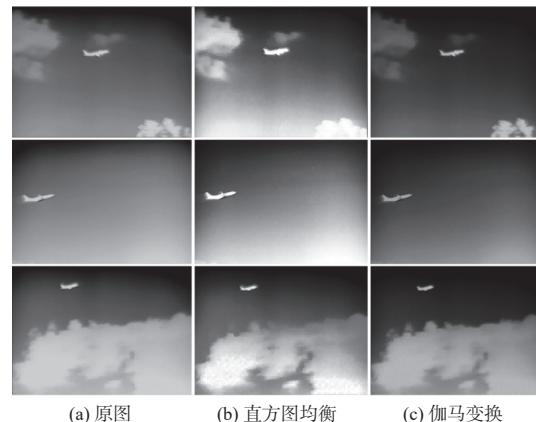


图 3 图像增强对比

Fig. 3 Image enhancement contrast

红外图片以单通道的形式输入到神经网络中, 所携带的信息有限, 通过并行的图像预处理设计, 将红外图像设计为图 2 所示的三通道图像, 输入到神经网络中能够学习到更丰富的特征。

### 2.2 改进的特征提取网络

如图 4 所示, 原始残差结构仅包含 3 个卷积层和 Relu 激活层, 将频域注意力 (FCA) 引入到残差网络结构中, 在原始残差映射的后面添加一个 FCA 注意力机制, 设计一种新的残差结构。输入  $X$  经过 3 层卷积后输出  $X'$ , 进入到注意力操作中;  $X'$  经过二维离散余弦变换 (DCT) 和 2 个线性全连

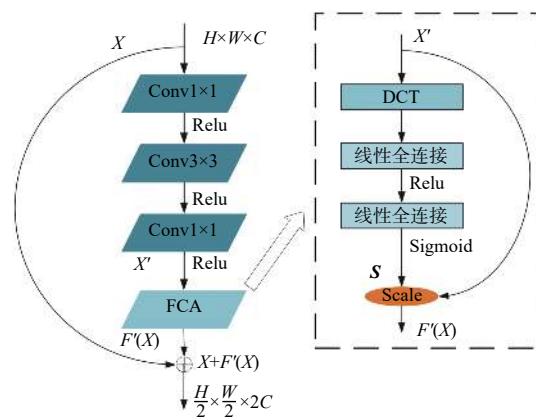


图 4 FCA-ResNet 残差结构

Fig. 4 Residual structure diagram of FCA-ResNet

接层后, 使用 Sigmoid 函数获得通道权值向量  $\mathbf{S}$ , 随后与  $X$  相乘, 获得带有通道注意的映射  $F'(X)$ , 最后与  $X$  相加, 经过残差映射后特征图大小变为原来的  $1/2$ , 通道数量变为原来的 2 倍。

### 2.3 FCA 注意力机制

SE 注意力机制<sup>[12]</sup> 通过“压缩-激励”操作获得一个通道权值向量, 代表各个通道的重要程度, 如图 5(a) 所示。SE 注意力机制简单地使用一个全局平均池化操作实现对输入的压缩, 并不能很好地捕获各种输入的复杂信息。

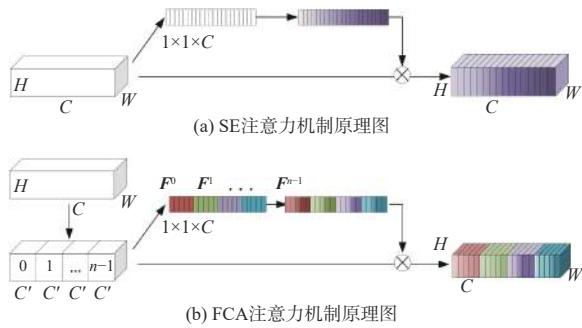


图 5 SE 与 FCA 注意力机制原理图

Fig. 5 Schematic diagram of SE and FCA attention mechanism

频域注意力机制 (FCANet)<sup>[13]</sup> 延用了 SE 注意力机制的思想, 将其扩展到频域上, 利用二维离散余弦变换 (2D DCT) 对通道进行压缩, 获得一个具有包含频域信息的通道权值向量。图 5(b) 展示了 FCANet 的思想。

通道注意力向量可通过式 (1) 获得:

$$\mathbf{a} = \text{Sigmoid}(f_c(\text{compress}(X))) \quad (1)$$

式中:  $\mathbf{a} \in \mathbb{R}^C$  表示注意力向量,  $C$  表示通道数量;  $X$  表示输入特征;  $\text{compress}$  表示二维离散余弦变换操作;  $f_c$  表示全连接层或一维卷积; Sigmoid 表示激活层操作。将通道权值向量与输入  $X$  在通道维度上相乘, 能获得施加通道注意力后的输出, 如式 (2):

$$\tilde{X}_i = \mathbf{a}_i X_i \quad \text{s.t. } i \in \{0, 1, \dots, C-1\} \quad (2)$$

式中:  $\tilde{X}_i$  表示单通道输出;  $X_i$  表示单通道输入;  $i$  表示通道的索引。

2D DCT 的基函数可以由式 (3) 表示:

$$B_{h,w}^{i,j} = \cos\left(\frac{\pi h}{H}(i + \frac{1}{2})\right) \cos\left(\frac{\pi w}{W}(j + \frac{1}{2})\right) \quad (3)$$

2D DCT 可被写成:

$$f_{h,w}^{2d} = \sum_{i=0}^H \sum_{j=0}^W x_{i,j}^{2d} B_{h,w}^{i,j} \quad (4)$$

$$\text{s.t. } h \in \{0, 1, \dots, H-1\}, w \in \{0, 1, \dots, W-1\}$$

式中:  $f_{h,w}^{2d}$  为 2D DCT 的输出;  $x_{i,j}^{2d}$  为输入。

如图 6 所示, 将原始图像进行二维离散余弦变换后, 可以看出, 红外图像的数据能量大部分分布在低频区域 (左上角), 但仍有部分能量分布在高频区。全局平均池化操作可以看作是  $h=0, w=0$  的一种特殊情况, 相当于保留最低频率分量的信息, 为了引入更多的信息, 利用 2D DCT 的多个频率分量对通道进行压缩。

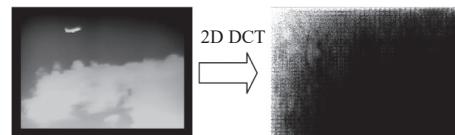


图 6 二维离散余弦变换

Fig. 6 Two-dimensional discrete cosine transform

频域通道注意力将通道分割成  $n$  份, 如图 5(b) 所示, 用  $X^i$  表示。对于每个部分都有与之对应的 2D DCT 的频率分量, 经过 DCT 后各部分的压缩结果表示为式 (5):

$$\mathbf{F}^i = 2D \text{ DCT}^{u_i, v_i}(X^i) = \sum_{h=0}^H \sum_{w=0}^W X_{h,w}^i B_{h,w}^{u_i, v_i} \\ \text{s.t. } i \in \{0, 1, \dots, n-1\} \quad (5)$$

式中:  $(u_i, v_i)$  为  $X^i$  的频率分量;  $\mathbf{F}^i$  表示各部分经过压缩后的权值向量。完整的压缩向量  $\mathbf{F}$  由  $\mathbf{F}^i$  拼接得到, 如式 (6) 所示, cat 表示拼接操作:

$$\mathbf{F} = \text{cat}([\mathbf{F}^0, \mathbf{F}^1, \dots, \mathbf{F}^{n-1}]) \quad (6)$$

### 2.4 多尺度特征融合预测

随着网络层数的加深, 网络特征中能提取到较多的语义信息, 但也会丢失细节信息, 相反底层网络中保留着大量的细节尤其是位置信息。利用多尺度预测能够充分利用底层信息, 有效提升算法的检测精度。目前大多数算法常用的 FPN<sup>[14]</sup> 金字塔结构进行多尺度特征提取。本文算法采用路径增强金字塔 (PAFPN)<sup>[15]</sup> 进行特征提取。与 FPN 相比, PAFPN 增加了一个自底向上的下采样对特征金字塔增强, 缩短底层信息的传递路径。

图 7 是一个以 FCA-ResNet50 为主干的 PAFPN 结构, 其中, 每个 stage 都是由图 5(b) 所示的 FCA-ResNet 残差结构堆叠而成。如图 7(a) 所示, 抽取不同尺度特征层 ( $C_2, C_3, C_4, C_5$ ), 通道维度分别为 256、512、1024、2048, 然后通过如图 5(b) 所示的自顶向下的上采样生成特征图 ( $P_2, P_3, P_4, P_5$ ), 接着通过如图 7(c) 所示的自底向上的下采样生成 ( $N_2, N_3, N_4, N_5$ ), 分别进行 Propsoal 的特征提取和缩放, 送入全连接层进行分类和回归。

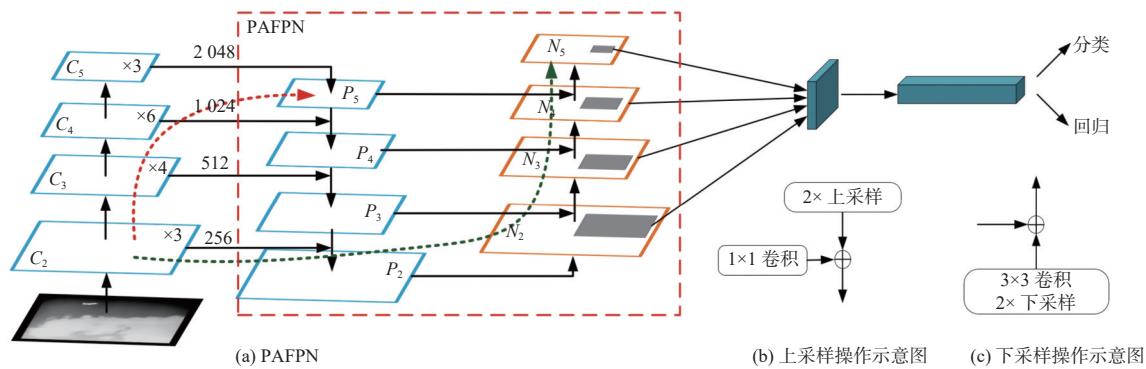


图 7 以 FCA-ResNet50 为主干的 PAFPN 结构

Fig. 7 PAFPN structure diagram with FCA-ResNet50 as main stem

### 3 实验与分析

#### 3.1 实验环境与数据集

实验在 Ubuntu18.04 操作系统上进行,采用两张 NVIDIA TITAN 显卡进行并行训练,深度学习环境为 Pytorch 1.10 和 cuda 10.2。实验采用在大型数据集上预训练得到的权重作为网络的初始化参数。

**数据集:** 实验所用红外飞机数据集部分来源于实验室拍摄图片,部分来源于网络公开红外弱小目标数据集<sup>[16]</sup>,其中包括 2057 张训练集图片和 540 张验证集图片。

**评价指标:** 以 MS COCO<sup>[17]</sup> 为基准引入 AP(average precision)作为网络性能的评价指标,采用浮点数(FLOPs)和参数量表示网络的复杂度。

#### 3.2 模型训练与测试

训练参数设置如表 1 所示。

表 1 参数设置

Table 1 Parameters setting

参数	设置
输入图像大小/像素	800×800
优化器	SGD
权重衰减系数	0.0001
初始学习率	0.005
批大小(batch_size)	4
训练周期/epoch	20
迭代次数/次	8000

实验采用“step”的学习率策略,初始学习率为设置 0.005,在第 5、8、11、16 个 epoch 处衰减,每次衰减为原来的 0.1。采用二值交叉熵损失计算分类损失,Smooth L1 计算边界框损失。

图 8 所示为训练过程中损失函数曲线,经过 8000

次迭代后,损失值保持在 0.1 以下,所以在 8000 次迭代以后停止训练,以防产生过拟合现象。

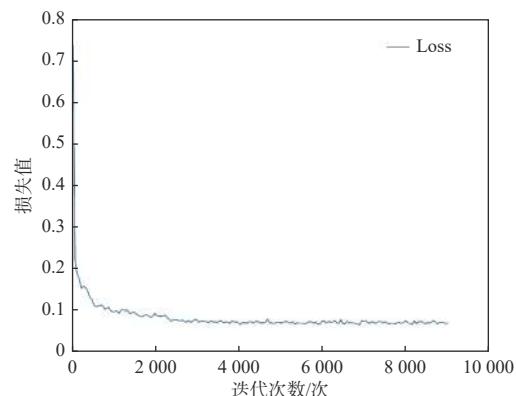


图 8 Loss 曲线

Fig. 8 Loss curve

如图 9 所示,使用不同尺寸、不同飞行姿态的红外图像对算法进行测试,可以看出,经过训练的模型能够很好地检测出红外飞机目标。图中类别后的数字代表置信度。



图 9 算法测试结果

Fig. 9 Detection results of algorithm

#### 3.3 对比分析

为进一步分析改进网络的有效性,设计了以下

对比实验进行分析: 1) 图像增强前后对比实验; 2) 模型改进对比试验, 其中模型改进包括主干网络和颈部网络; 3) 本文算法与目前主流算法检测结果对比试验。实验结果分别展示在表 2~表 4 中。

为了验证图像增强的有效性, 我们采用 ResNet50 作为检测主干和 FPN 金字塔结构, 在红外飞机数据集上进行对比实验, 如表 2 所示。

表 2 不同图像增强方式检测结果

Table 2 Detection results of different image enhancement methods

模型	图像增强		AP
	直方图均衡	伽马变换	
ResNet50+FPN			84.6
	√		84.9
		√	84.7
	√	√	85.1

由表 2 可知, 在单独使用直方图均衡和伽马变换进行图像增强预处理时, 网络性能分别提升了 0.3% 和 0.1%, 而通过将这两种图像预处理结合, 网络检测精度提升了 0.5%, 证明在红外图像输入网络前进行预处理能够提升算法的检测精度, 验证了并行结构设计的有效性。

实验 2 为验证模型改进的有效性, 基于 Faster R-CNN 的目标检测框架, 统一使用图像增强算法, 在红外飞机检测数据集上进行模型改进对比分析, 结果记录在表 3 中。

表 3 不同模型检测结果

Table 3 Detection results of different models

Backbone	Neck	AP	FLOPs/G	参数量/M
ResNet50	FPN	85.1	134.38	41.12
SE-ResNet50	FPN	88.7	134.45	43.63
FCA-ResNet50	FPN	90.0	134.38	43.61
FCA-ResNet50	PAFPN	92.7	149.87	47.15

由表 3 前 3 行可知, SENet 的引入给 ResNet50 的网络带来了 3.6% 的 AP 提升, FCANet 的引入给 ResNet50 的网络带来了 4.9% 的 AP 提升, 证明了频域注意力机制的有效性; 同时, 相较于 ResNet50, 引入注意力机制后, 并没有给网络带来很高的复杂度的提升。

将 3 组主干在  $P_2$ 、 $P_3$ 、 $P_4$  层特征上的热力图映射回原图, 见图 10。从  $P_2$  层热力图可以看出, 施

加特征注意力之后网络学习的特征更加集中于目标本体; 同时通过对 SE 注意力和 FCA 注意力的  $P_2$  层热力图, 由于 FCA 注意力考虑了不同频率分量的信息, 施加 FCA 注意力机制的网络对于目标本身的学习更加全面。通过对  $P_2$ 、 $P_3$ 、 $P_4$  层的热力图, 可以证明, 随着网络层数加深, 网络的感受野更大, 学习的范围更广, 这有利于大目标的学习, 而对于小目标, 底层特征信息显然尤为重要。

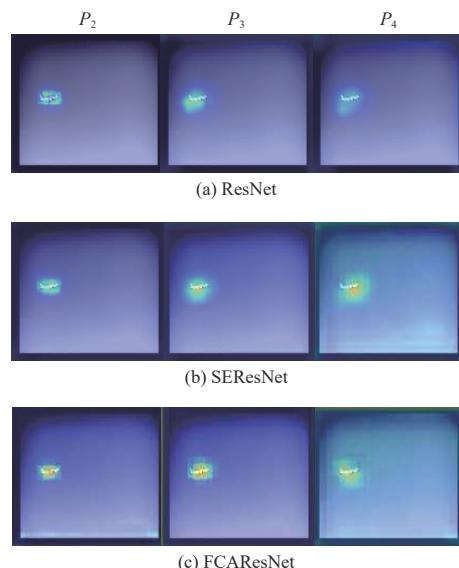


图 10 不同主干  $P_2$ 、 $P_3$ 、 $P_4$  层的热力图可视化

Fig. 10 Visualization of heat map of  $P_2$ ,  $P_3$  and  $P_4$  layers in different backbones

为验证多尺度特征融合检测的有效性, 以 FCA-ResNet50 做为主干, 引入 PAFFPN 结构进行分析。表 3 中对比了使用 FPN 结构和 PAFFPN 结构的检测结果。从表 3 后两行中可以看出, 使用 PAFFPN 结构取代 FPN 结构之后以 FCA-ResNet50 为主干的网络检测精度提升了 2.7%, 验证了路径增强的有效性, 进一步验证了注意力机制的有效性; 同时, 引入 PAFFPN 结构也没有给网络结构带来很高的复杂度提升。

为验证算法整体的有效性, 在红外飞机数据集上进行不同算法框架的对比实验。实验所采用的均是目前主流的算法框架, 实验结果记录在表 4。由表 4 可知, 经过改进的 Faster R-CNN 算法有最优的检测精度。根据表 4 中参数量和浮点数(FLOPs)可以看出, 通常单阶段的检测框架模型复杂度要比两阶段的模型复杂度低, 所以检测速度较快, 但因为缺少 RPN 结构提取预测框, 所以检测精度往往要比两阶段框架低。

表 4 主流算法框架的检测结果

Table 4 Detection results for mainstream algorithmic frameworks

Model	图像增强	Backbone	Neck	AP	FLOPs/G	Param/M
One-stage						
YOLOv3	×	DarkNet53	FPN	82.1	121.15	61.52
RetinaNet	×	ResNet50	FPN	83.2	127.82	36.1
Fcos	×	ResNet50	FPN	84.7	125.41	32.02
Two-stage						
Cascade R-CNN	√	ResNet50	FPN	86.7	162.18	68.93
Mask R-CNN	√	ResNet50	FPN	86.1	187.45	43.77
Faster R-CNN	√	ResNet50	FPN	85.1	134.38	41.12
本文算法	√	FCA-ResNet50	PAFPN	92.7	134.38	43.61

如图 11 所示, 将本文检测结果与其他 3 组检测结果进行对比。可以看出: 在目标物体尺寸较大时, 各组检测结果的差距并没有很大, 但是在针对小目标的检测, YOLOv3 和传统 Faster-RCNN 算法会出现漏检的情况; 而在复杂背景情况下, 本文算法和其他算法均会出现误检的情况, 但相比之下, 本文算法表现较好。

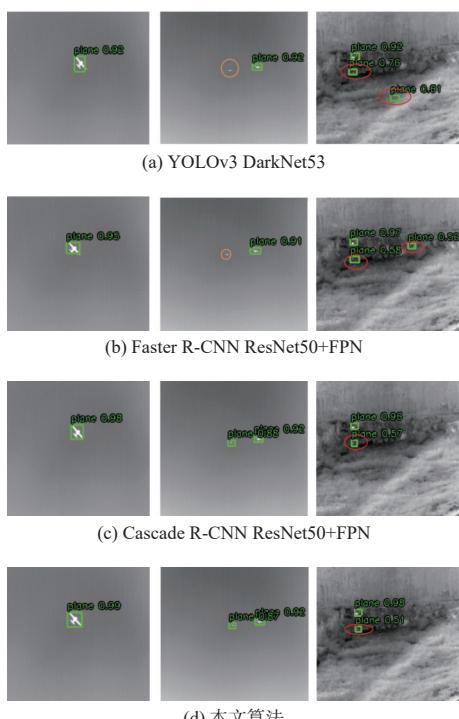


图 11 不同算法的检测结果对比

Fig. 11 Comparison of detection results of different algorithms

## 4 结论

本文构建了红外飞机目标检测的数据集, 采用并行图像增强的算法对输入图片进行增强。设计了一种改进的 Faster R-CNN 检测框架, 将其运用于红外目标的检测任务当中。利用频域注意力机制对 ResNet 网络增强, 设计了 FCA-ResNet50 主干结构, 结合路径增强金字塔网络结构, 在仅小幅度提升算法复杂度的情况下提升目标的检测精度, 相比于其他主流算法, 有较好的检测表现。实验证明, 本文算法能够很好地被运用于红外飞机目标检测任务中, 证明了将可见光领域的目标检测算法移植到红外领域的可行性。下一步, 计划对于红外目标检测算法进行进一步改进, 设计出具有实时检测能力的算法框架。

## 参考文献:

- [1] 曹红燕, 沈小林, 刘长明, 等. 改进的 YOLOv3 的红外目标检测算法 [J]. 电子测量与仪器学报, 2020, 34(8): 188-194.  
CAO Hongyan, SHEN Xiaolin, LIU Changming, et al. Improved YOLOv3 infrared target detection algorithm [J]. Electronic Measurement and Instrument, 2020, 34(8): 188-194.
- [2] 顾佼佼, 李炳臻, 刘克, 等. 基于改进 Faster R-CNN 的红外舰船目标检测算法 [J]. 红外技术, 2021, 43(2): 170-178.  
GU Jiaojiao, LI Bingzhen, LIU Ke, et al. Infrared ship target detection algorithm based on improved Faster R-CNN [J]. Infrared Technology, 2021, 43(2): 170-178.
- [3] 蔡伟, 徐佩伟, 杨志勇, 等. 复杂背景下红外图像弱小目标检测 [J]. 应用光学, 2021, 42(4): 643-650.  
CAI Wei, XU Peiwei, YANG Zhiyong, et al. Dim target detection in infrared image with complex background [J]. Applied Optics, 2021, 42(4): 643-650.
- [4] 谭海云, 余鸿皓, 王海川, 等. 基于改进 YOLOX 的红外目标检测算法 [J]. 电子测量技术, 2022, 45(23): 72-81.  
CHEN Haiyun, YU Honghao, WANG Haichuan, et al. Infrared target detection algorithm based on improved YOLOX [J]. Electronic Measurement Technology, 2022, 45(23): 72-81.
- [5] GIRSHICK R. Faster R-CNN[C]//2015 IEEE International Conference on Computer Vision. New York: IEEE, 2015: 1440-1448.
- [6] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-

- CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. New York: IEEE, 2017: 2961-2969.
- [ 7 ] CAI Z, VASCONCELOS N. Cascade R-CNN: high quality object detection and instance segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(5): 1483-1498.
- [ 8 ] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 779-788.
- [ 9 ] REDMON J, FARHADI A. Yolo9000: better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 7263-7271.
- [ 10 ] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO series in 2021[EB/OL]. [2023-03-20]. <https://ui.adsabs.harvard.edu/abs/2021arXiv210708430G/abstract>.
- [ 11 ] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. New York: IEEE, 2017: 2980-2988.
- [ 12 ] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 7132-7141.
- [ 13 ] QIN Z, ZHANG P, WU F, et al. Fcanet: frequency channel attention networks[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2021: 783-792.
- [ 14 ] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 2117-2125.
- [ 15 ] WANG K, LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 8759-8768.
- [ 16 ] 回丙伟, 宋志勇, 范红旗, 等. 红外序列图像中弱小飞机目标检测跟踪数据集 [DB/OL]. Science Data Bank, 2019 [2019-10-28]. <http://10.11922/csdata.2019.0074.zh>. HUI Bingwei, SONG Zhiyong, FAN Hongqi, et al. A dataset for dim-small target detection and tracking of aircraft in infrared image sequences [DB/OL]. Science Data Bank, 2019 [2019-10-28]. <http://10.11922/csdata.2019.0074.zh>.
- [ 17 ] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft coco: common objects in context[C]//Proceedings of 13th European Conference on Computer Vision–ECCV 2014, Part V 13. Zurich, Switzerland: Springer International Publishing, 2014: 740-755.