

文章编号 1004-924X(XXXX)XX-0001-11

采用多尺度特征融合 SSD 的遥感图像小目标识别

陈欣^{1,2}, 万敏杰^{1,2*}, 马超^{1,2}, 陈钱^{1,2}, 顾国华^{1,2}

(1. 南京理工大学 电子工程与光电技术学院, 江苏 南京 210094;

2. 南京理工大学 江苏省光谱成像与智能感知重点实验室, 江苏 南京 210094)

摘要:针对复杂背景下遥感小目标的识别问题,提出了一种改进型多尺度特征融合 SSD 方法。首先,设计了一种特征图融合机制,将分辨率高的浅层特征图与具有丰富语义信息的深层特征图进行融合,并在特征图间构建特征金字塔,对小目标特征进行增强。然后,引入通道注意力模块,通过构建权重参数空间,将注意力更多的集中在关注目标区域的通道,以减小背景干扰。最后,对先验框相对于原图的比例进行了调整,使其能够更好地适应遥感小目标尺度。利用采集的遥感飞机图像数据集对方法性能进行定性和定量测试,结果表明:改进方法检测精度相较 SSD 提高了 4.3%,其能够适应复杂场景下的遥感多尺度目标识别任务,降低小目标的漏检率。

关键词:深度学习;小目标检测;复杂背景;特征融合;通道注意力机制

中图分类号:TP751 **文献标识码:**A **doi:**10.37188/OPE.XXXXXXXX.0001

Remote sensing image small target recognition using multi-scale feature fusion-based SSD

CHEN Xin^{1,2}, WAN Min-jie^{1,2*}, MA Chao^{1,2}, CHEN Qian^{1,2}, GU Guo-hua^{1,2}

(1. College of Electronic and Optical Engineering, Nanjing Univ. of Science and Tech.,
Nanjing 210094, China;

2. Jiangsu Key Laboratory of Spectral Imaging and Intelligent Perception,
Nanjing Univ. of Science and Tech., Nanjing 210094, China)

* Corresponding author, E-mail: minjiewan1992@njust.edu.cn; gghnjjust@mail.njust.edu.cn

Abstract: For the recognition of small remote sensing targets in complex backgrounds, an improved multi-scale feature fusion-based single shot multi-box detector (SSD) method is proposed. First, a feature map fusion mechanism is designed to fuse shallow high-resolution feature maps and deep feature maps with rich semantic information, and feature pyramids between feature maps are built to enhance small target features. Then, the channel attention module is introduced to overcome the background interference by constructing a weight parameter space to put more attention on the channels that focus on the target region. Finally, the scale between the priori box and the original map is adjusted to better fit the remote sensing small target scale. Qualitative and quantitative tests based on remote sensing aircraft image datasets are im-

收稿日期:2021-03-27;修订日期:2021-04-16.

基金项目:国家自然科学基金青年项目(No. 62001234);江苏省自然科学基金青年项目(No. BK20200487);中国博士后科学基金面上项目(No. 2020M681597);江苏省博士后科研资助(No. 2020Z051);上海航天科技创新基金(No. SAST2020-071);中央高校基本科研业务费专项资金资助(No. JSGP202102)

plemented, and the results show that the proposed method improves the detection accuracy by 4.3% when compared with SSD, and can adapt to multi-scale remote sensing target recognition tasks in complex scenes and reduce the missing detection rate of small targets.

Key words: Deep learning; Small target detection; Complex background; Feature fusion; Channel attention mechanism

1 引言

遥感图像目标检测技术是目前计算机视觉领域重要研究方向之一,相对于其他类型遥感图像如合成孔径雷达而言,高分辨率的光学遥感图像具有其独特的优势。在民用和军事领域,光学遥感图像检测都得到了越加广泛的应用^[1]。传统的遥感图像目标检测方法,如尺度不变特征变换、方向梯度直方图(HOG)、形变部件模型(DPM)等,都依赖人工设计的特征,检测精度较低,实时性差。

自2012年Hinton等人提出的AlexNet在ILSVRC图像分类比赛夺冠后,深度学习迅速发展,并在计算机视觉领域得到了越来越广泛的运用^[2],越来越多学者将深度学习应用于遥感场景下的目标识别中^[3-4]。基于深度学习的目标检测算法主要分为两类:一是以Faster R-CNN^[5]系列为代表的基于候选区域的两阶段目标检测算法,该类方法基于原始图像生成一系列建议区域,并将建议区域和特征图输入感兴趣区域池化层,最终实现目标的分类与定位;二是以SSD^[6]、YOLO^[7]系列为代表的基于回归的单阶段目标检测算法,该类检测方法从回归的角度出发,直接在图像上回归出目标的边框位置和目标类别。文献[8]对Faster R-CNN不同层的特征图进行聚合并改进了其损失函数,在SAR图像中的舰船目标识别上取得较好检测效果。文献[9]提出了DSSD模型,采用ResNet-101并结合反卷积模块加入了上下文信息,提高了网络的特征提取能力,但其检测速度较慢。Liu等人^[10]提出了结合浅层的细节特征和高层的语义特征的FSSD算法,在保证精度的同时提高了检测速度,但其对小目标检测结果较差。文献[11]将YOLOv2算法与生成对抗网络、卷积密集连接相结合,提升了算法的鲁棒性。Wang等人^[12]针对遥感影像小目标检测,提出了一种结合特征金字塔(Feature

Pyramid Networks, FPN)^[13]的SSD改进算法,该算法一定程度上提高了检测速度和精度。文献[14]提出了一种多尺度卷积神经网络遥感目标检测框架——MSCNN,该算法对多尺度遥感目标检测结果较好,但其对复杂背景的适应性较弱。上述两阶段类目标检测算法的高计算成本限制了该类算法的目标检测速度,单阶段的目标检测算法则对多尺度目标特别是小目标的检测精度较差。

为了实现复杂背景下遥感小目标的高效检测,本项目提出了一种采用多尺度特征融合SSD的识别方法,该方法设计了特征融合模块,将主要负责小目标检测大尺度的浅层特征图与具有丰富语义信息的深层特征图相融合,采用金字塔特征结合方式,对浅层信息进行了增强,提高了其对小型目标的提取能力。同时针对遥感图像复杂的背景环境,增加通道注意力模块,增大目标区域通道的特征响应,减弱背景区域通道的特征响应。最后调整先验框相对于原图的比例,提高网络的小目标检测能力。本项目方法在遥感飞机数据集上进行分析实验,实验结果表明,本项目方法在满足实时检测的基础之上,对复杂背景下的遥感图像小目标具有更强的检测能力。

2 SSD算法模型

SSD(single shot multi-box detector)算法是Liu等人在2016年提出的一种在兼顾精度同时大幅提升检测速度的目标检测算法。SSD以VGG16作为基础的特征提取网络,在将其最后两层的全连接层变换成卷积层的基础上,额外添加了5个卷积结构形成其网络框架,其网络框架如图1所示。

SSD算法的核心主要有两点:一是提取不同尺度的特征图进行目标识别,使用大尺度高分辨率的特征图用来进行小尺度目标识别,而小尺寸

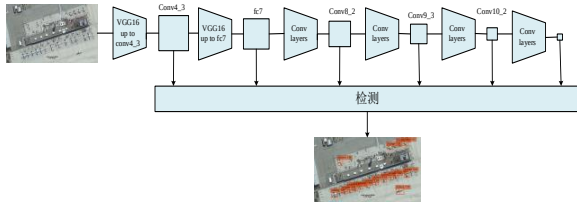


图1 SSD网络框架

Fig. 1 SSD network framework

语义信息丰富的特征图则用来识别大尺度目标。SSD分别提取了Conv4_3、Conv7、Conv8_2、Conv9_2、Conv10_2和Conv11_2共6个不同尺度的特征图送入后端进行目标的分类和回归,有效的提升了对不同尺度目标的识别能力。二是先验框的生成,针对待检测特征图,对其上的每一个像素点都生成了不同尺度和长宽比的先验框,预测时目标框以这些先验框为基准计算偏差,可在一定程度上降低训练的难度。

3 改进的SSD方法

SSD算法虽然提取了多尺度特征图进行目标检测,但是负责小目标检测的主要是浅层特征图Conv4_3,该层特征图语义信息较少,导致在对复杂背景下的遥感图像进行检测时,对遥感图像中小目标的检测能力较弱,同时对复杂背景的区别能力不强。

为解决上述问题,本项目采用了一种多尺度特征融合SSD方法,主要从以下三点进行改进:首先,利用特征融合机制将深层、浅层特征进行融合,采用特征金字塔实现特征的增强;然后,加入通道注意力模块(Channel Attention Module, CAM)增强了背景和目标之间的可区分度;最后,对先验框相对于原图的比例进行了调整以更好适应小目标识别问题。改进后的SSD网络结构如图2所示。

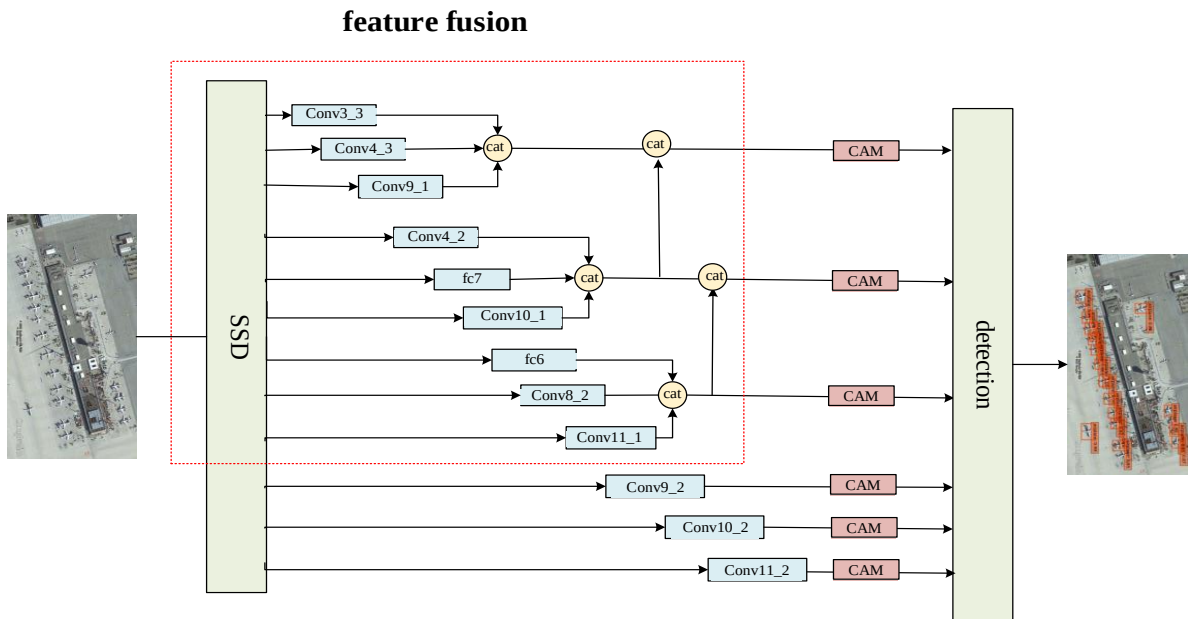


图2 改进后的SSD网络结构图

Fig. 2 The structure of the improved SSD network

3.1 多尺度特征融合

SSD算法提取多尺度特征图进行目标检测,负责小目标检测的主要是低层特征图Conv4_3,其分辨率高,拥有较为丰富的细节信息,但是其所包含的语义信息较为匮乏,导致其针对遥感图

像小目标的检测能力较弱。

针对以上问题,本项目设计了一种特征融合机制,将三个不同深度的特征图进行融合。将丰富的语义信息从深层特征图融合的浅层特征图的同时,也将浅层特征图丰富的细节信息融合到

了深层特征图,保留了更多的上下文信息。对 Conv4_3、Fc7、Conv8_2 三个不同尺度特征层进行特征增强。具体步骤如下:首先,采用 1×1 卷积核在提取到的三个不同深度的特征图上进行卷积操作,将各个深度特征图的通道数进行统一。其次针对提取到的不同深度特征图尺度不同的问题,使用 3×3 卷积核对其进行卷积,统一其输出尺度。最后在各个特征图输出的尺度和维度都一致之后,采用通道拼接 (Concatenate, cat) 的方式对其进行融合。

融合过后得到 Conv4_3_1、Fc7_1、Conv8_2_1 三个不同尺度特征增强的特征图。具体过程如图 3 所示图。

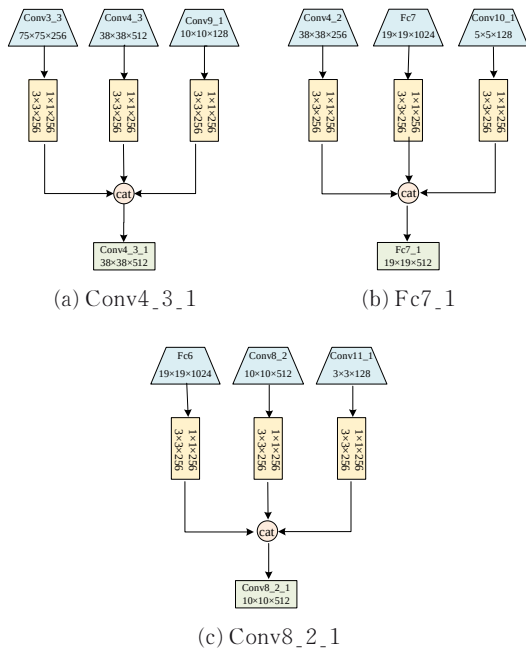


图3 特征融合增强示意图

Fig. 3 Schematic diagram of feature fusion enhancement

同时采用特征金字塔代替 SSD 原有的特征结合方式,算法共提取六个特征层进行检测,后三个检测层由 Conv11_2、Conv10_2 和 Conv9_2 构成,另外三个检测特征图由 Conv4_3_1、Fc7_1、Conv8_2_1 三个特征增强的特征图通过构建特征金字塔获得,即利用三个不同尺度特征图进一步地进行特征增强。首先,将 Conv8_2_1 特征图作为特征金字塔中第一个特征图。其次,对 Conv8_2_1 进行上采样增大其尺度,然后将其与

Fc7_1 经过 1×1 卷积降维后的特征图进行通道连接,将连接获得的特征图作为金字塔的第二个特征图。最后,将连接后的特征图上采样增大尺度后与 Conv4_3_1 降维后的特征图进行通道连接,将其作为金字塔的第三个特征图。流程图如图 4 所示。

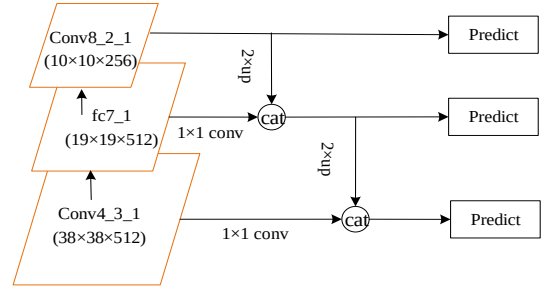


图4 FPN 结合示意图

Fig. 4 FPN bonding schematic

3.2 通道注意力机制

由于光学遥感图像背景复杂,背景噪声容易对目标的识别造成干扰,为了实现复杂背景下的目标准确高效识别,减小背景的干扰,在多尺度特征融合之后加入通道注意力模块,将模型注意力更多的转移到关注目标区域的通道,增大这一部分通道的特征响应,相应的减弱关注背景区域那一部分通道的特征响应,从而增大目标区域和背景区域的可区分性,以减少背景对目标识别的干扰。

通道注意力模块设计了一个新的权重参数空间,通过引入不同的权重来衡量不同通道之间的重要程度,为关注目标区域的通道附加更高的注意力权重,相应的减小关注背景区域通道的注意力权重,进而增大背景和目标的区分性,以减小遥感图像复杂背景的干扰。其具体流程如图 5 所示。

首先,对输入尺寸为 $H \times W \times C$ 的特征 x 进行最大池化和平均池化,分别得到两个 $1 \times 1 \times C$ 的一维特征向量,使用两种不同的池化方式在获得两种不同空间上下文的特征描述的同时,也可以适应不同深度特征图中特征提取的需求。其次,再经过两层全连接层先降低维度再升高维度,在得到每个通道权重大小的同时整合输入特征的每个通道信息。再次,将每个像素点的位置元素相加,经过 Sigmoid 激活函数将权重归一化

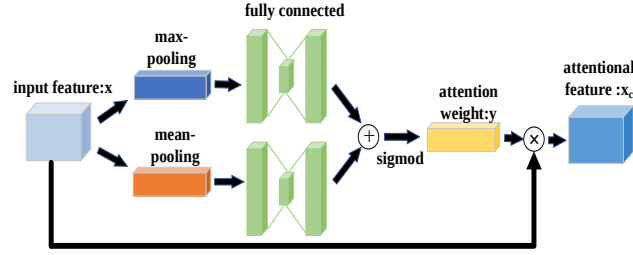


图5 通道注意力机制示意图

Fig. 5 Schematic diagram of channel attention mechanism

到0~1之间,得到注意力权重 y 。最后,将注意力权重与特征进行特征值重塑,得到注意力特征 x_c 如式(1)所示。

$$x_c = x \cdot y. \quad (1)$$

3.3 先验框尺寸调整

SSD算法将用于检测的特征图划分成网格结构,以每个网格的中心点为中心,生成不同大小不同比例的先验框,待检测目标的中心落在哪个网格,就由该网格负责该目标的检测,如图6所示。

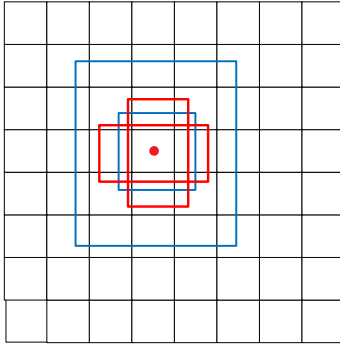


图6 SSD的先验框生成示意图

Fig. 6 Schematic diagram of a priori frame generation for SSD

特征图上生成的先验框的尺寸遵守一个线性递增规则:随着特征图的尺寸的减小,先验框尺寸线性增加,如式(2)所示。

$$s_k = s_{\min} + \frac{s_{\max} - s_{\min}}{m - 1} (k - 1), k \in [1, m], \quad (2)$$

式中, $m=5$, $s_{\min}=0.2$, $s_{\max}=0.9$, s_k 表示先验框尺寸相对原图片比例。

通过数据集中目标尺寸的统计,其中飞机目标最小的尺寸约为 20×20 ,而SSD算法中最浅层特征图上先验框的尺寸是 30×30 ,不适应飞机

目标的最小尺寸,容易造成漏检。故针对数据集中目标的尺寸信息调整比例,将 $s_{\min}$, $s_{\max}$ 的值调整为0.12, 0.88,使得在浅层特征图中先验框的尺寸减小为18,以适应数据集中飞机的最小尺度,避免因遥感图像目标尺寸小而导致在浅层特征图中先验框无法匹配到目标。尺寸调整前后的先验框大小如表1所示。

表1 改进后先验框尺寸

Tab. 1 Improved a priori frame dimension

特征图	原本先验框大小	改进后先验框大小
Conv4_3_1	30	18
Fc7_1	60	36
Conv8_2_1	111	93
Conv9_2	162	150
Conv10_2	213	207
Conv11_2	264	264

3.4 算法总结

方法1采用多尺度特征融合SSD的遥感图像小目标识别方法

输入: I: 光学遥感飞机图像训练样本集

D_L : 训练样本集标签: $D_L = \{x_i, y_i\}, i = 1, 2, \dots, L$

B: 真实框 b_i 的集合

D_T : 测试样本集: $D_T = \{x_i\}, i = 1, 2, \dots, T$

输出: 测试样本集的检测结果、检测精度

1: 构建改进型多尺度特征融合SSD模型

2: 训练网络

a: 初始化网络参数

随机初始化网络权重 θ , 设定最大迭代次数 M 、学习率 lr

b: 输入I, 利用 D_L 进行监督训练

for $i = 0:M$ do
 前向传播
 I 在图片上生成先验框 a_j , 计算 a_j 和 b_i 的
 IOU 值: $\text{IOU} = \frac{(a_j \cap b_i)}{(a_j \cup b_i)}$,
 II 根据 IOU 值划分正负样本
 III 计算损失函数: $L(x, c, l, g) = \frac{1}{N} (L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, l, g))$
 反向传播:
 利用损失函数导数优化网络权重: $\theta^{t+1} = \theta^t - \lambda \nabla_{\theta^t} L(\theta^t)$
 end for
 c: 得到训练好的网络模型
 3: 利用训练好的网络模型对测试集图片进行检测
 4: 输出检测结果 检测精度

4 实验与结果分析

4.1 数据集与参数设置

本项目在自己收集和标注的遥感飞机数据集上进行多尺度目标检测实验。该数据集共包含 420 幅光学遥感飞机图像, 平均尺度约为 $1\,300 \times 900$, 共标注飞机目标 3 324 个, 图像分辨率在 $0.8 \sim 2.0$ m 之间, 每一幅图像至少包含一个目标。

为了避免过拟合, 本项目采用旋转、翻转等变换进行数据扩充。在训练过程中, 本项目随机将数据集的 70% 用于训练, 剩下的 30% 用于测试。根据飞机数据集的尺度 ($1\,300 \times 900$) 以及目标分布信息统计, 定义了相应的边界框尺度量 (Small object: $\text{area} < 60^2$, Medium object: $60^2 < \text{area} < 120^2$, Large object: $\text{area} > 120^2$), 对所采集的遥感飞机数据集中的飞机目标尺度情况进行统计, 相应的统计信息如表 2 所示。由表 2 可知, 所采集数据集中小目标数量占总体目标数量 37.2%。

实验硬件环境 CPU 为 Inter E5-2680, 内存为 256G, GPU 为 NVIDIA TITAN RTX, 操作系统为 Ubuntu 16.04, 采用 Pytorch 作为深度学习框架进行训练和测试。本项目采用端到端训练方

表 2 基于实例尺寸分布的边界面积定义及目标数量

Tab. 2 Definition of boundary area and number of targets based on instance size distribution

	Small	Medium	Large
尺度	$(0, 60^2]$	$(60^2, 120^2]$	$(120^2, +\infty]$
目标数量	1 236	1 864	221

式, 初始学习率设置为 0.000 1, 批处理大小设置为 16, 优化方法为随机梯度下降, 动量设置为 0.9, 正则化系数设置为 0.000 5。

4.2 评价指标

本项目采用平均检测精度 (average precision, AP) 作为飞机目标检测结果的评价指标。AP 计算的是召回率在 $0 \sim 1$ 之间的平均精度, 即精度-召回率曲线所包络的面积, 因此 AP 值越高代表检测性能越好。精度 p 和召回率 r 指标可以表示为 (3)、(4) 式所示。

$$p = \frac{T_p}{T_p + F_p}, \quad (3)$$

$$r = \frac{T_p}{T_p + F_N}, \quad (4)$$

其中, T_p : 真正例, F_p : 假正例, F_N : 假反例。

4.3 实验结果与分析

在所采集的遥感飞机数据集上对不同的目标检测方法进行训练和测试, 检测结果如表 3 所示。

表 3 各方法在遥感飞机数据集上的实验结果

Tab. 3 Experimental results of each algorithm on remote sensing aircraft dataset

检测方法	基础网络	AP (IOU=0.5)	FPS
Faster R-CNN	VGG16	87.4%	5.2
Faster R-CNN	ResNet-101	88.6%	3.6
YOLO	GoogleNet	80.5%	33.5
SSD	VGG16	85.9%	30.2
DSSD	ResNet-101	86.2%	12
FSSD	VGG16	87.5%	21.2
ours	VGG16	90.2%	24.3

由表 3 可知, 基于候选框的双阶段检测算法 Faster R-CNN 以 VGG16 作为基础网络时检测精度为 87.4%, 将其基础网络替换为 ResNet-101 后, 精度提高 1.2%, 但同时, 其帧频 (FPS) 也有相应下降; 相较于 Faster R-CNN, YOLO 作为单

阶段目标检测算法其帧频大幅提升,但同时精度有所不足;SSD算法在兼顾检测速度的同时,取得了85.9%的检测精度;DSSD相较SSD采用反卷积特征融合提高了对多尺度目标的检测能力;FSSD在SSD基础上引入金字塔特征融合方式,增强了不同尺度目标特征的表达能力。相较于

其他典型的目標检测算法,本项目所提的改进方法取得了最优的平均检测精度,其精度较SSD算法提升了4.3%,同时其帧频也较高。实验证明,本项目所提方法能够在满足实时检测的基础上,实现对复杂背景下遥感小目标的高效检测。

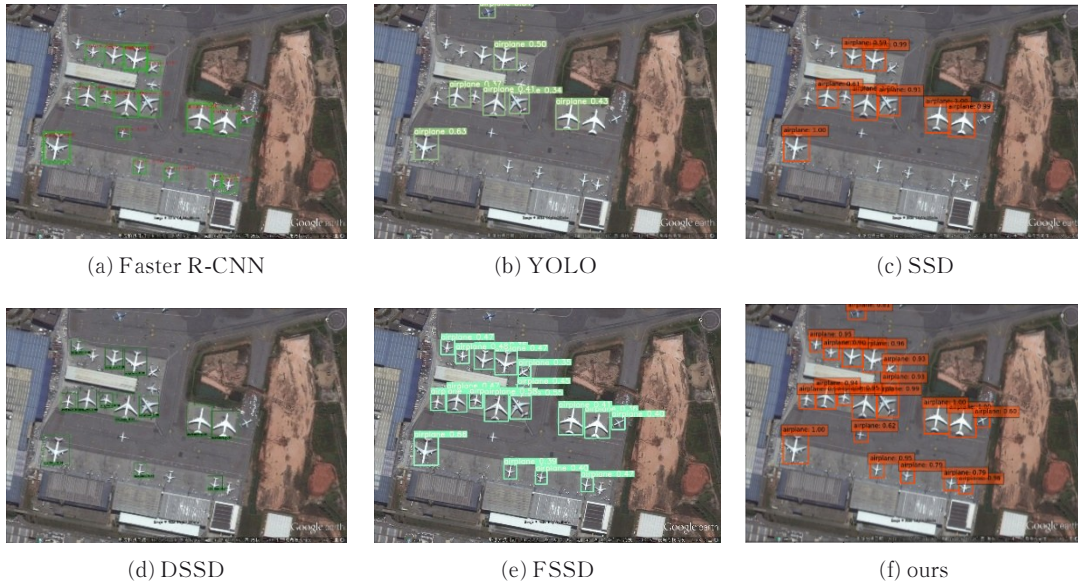


图7 各个方法检测结果

Fig. 7 Detection results of each algorithm

如图7所示,给出各方法在遥感飞机上的检测效果。从图7中Faster R-CNN、YOLO SSD、DSSD、FSSD以及本项目所提方法在遥感飞机上的检测效果可以看出,相较于其他方法,本项目所提出的方法对小目标的检测能力更强,能有效降低小目标的漏检率。同时给出SSD与改进方法在不同区域:复杂背景区域、多尺度目标区

域以及小目标集中区域的识别效果对比,如图8~图9所示。从图中可看出,相较于SSD算法,改进方法在复杂背景区域、多尺度目标区域、小目标密集区域的检测效果都明显提升,改进方法能够有效的进行复杂背景下多尺度目标的识别,降低小目标的漏检率。



图8 SSD算法检测结果

Fig. 8 SSD algorithm detection results

4.5 迁移实验

为了验证本项目所提算法的有效性、通用性以及可迁移性,除了在上述采集的遥感飞机数据集进行算法的性能对比之外,将在上述遥感飞机数据集上训练好的检测模型,迁移到不同姿态、不同类型的空中飞机目标检测中。收集了 35 张

以天空为背景的不同姿态不同机型的飞机图片,每张图片中平均含有 4~8 个飞机目标。以其为迁移数据集进行迁移实验,改进方法和 SSD 方法的在迁移数据集上的平均检测精度分别为 0.672 和 0.541,改进方法的迁移精度较原本方法提升了 0.131。检测结果对比如图 10 所示。

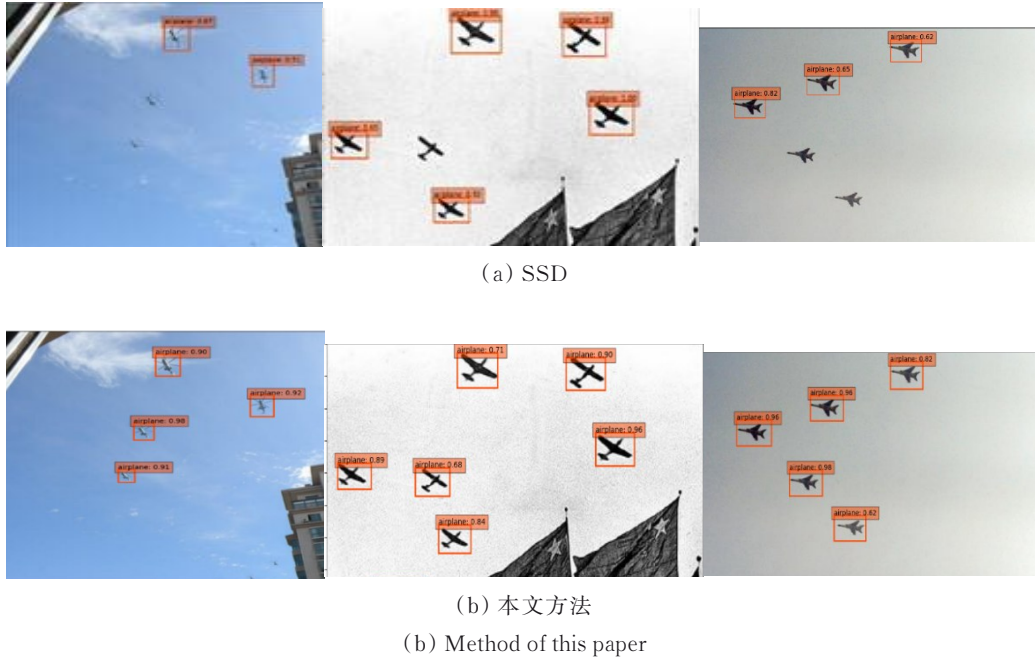


图 10 对空中不同姿态、不同类型的飞机目标检测结果

Fig. 10 Detection results for different attitudes and types of aircraft targets in the air

图 10 给出 SSD 和改进方法的检测模型对空中飞机目标的检测效果图,本项目所采集的遥感飞机数据集中的飞机目标皆为地面目标,机型主要为民航客机,姿态皆为俯拍。将在其上训练好的检测模型,用于空中的不同姿态、不同类型的飞机目标检测。其中的检测目标分别为:空中仰拍姿态直升机、空中仰拍姿态以及侧拍姿态的不同机型军事飞机。通过检测的对比结果可看出,相较于 SSD 算法,改进方法训练好的检测模型对天空背景下不同姿态、不同类型的飞机目标都能进行更加有效的检测,这是因为进行特征融合的过程中,不仅提高了特征的提取能力,也一定程度上增强了特征的泛化能力,使得模型具有较好的迁移能力。迁移实验的结果证明了本项目所提改进算法所学习到的检测模型具有一定的迁

移性和通用性。

5 结 论

本项目针对遥感图像复杂背景下小目标的识别问题提出了一种改进型多尺度特征融合 SSD 方法。首先,针对小目标特征弱的问题,设计了一种特征图融合机制,将分辨率高的浅层特征图与具有丰富语义信息的深层特征图进行融合,并在特征图间构建特征金字塔,增强小目标特征,其精度提升 2.8%。然后,针对遥感目标背景复杂的问题,增加通道注意力模块,将算法注意力更多的集中在目标区域的通道,增大目标区域通道的特征响应,减弱背景噪声区域通道的特征响应,减小背景的干扰,精度在原先基础上提

升1.1%。最后,针对小目标尺度,对先验框相对于原图的比例进行了调整,使算法能够较好地适应小尺度遥感目标识别问题,精度在原先基础上提升了0.4%。

性能实验证明了本文方法在识别精度上优

于主流算法,消融实验以及迁移实验证明了方法的完备性和通用性。实验结果表明,本项目所提方法能够有效增强小目标的表征能力,降低小目标的漏检率,能够在满足实时检测的基础上,实现对复杂背景下遥感图像中小目标的高效识别。

参考文献:

- [1] 范丽丽,赵宏伟,赵浩宇,等. 基于深度卷积神经网络的目标检测研究综述[J]. 光学精密工程, 2020, 28(5): 1152-1164.
FAN L L, ZHAO H W, ZHAO H Y, *et al.* Survey of target detection based on deep convolutional neural networks[J]. *Opt. Precision Eng.*, 2020, 28(5): 1152-1164. (in Chinese)
- [2] 方路平,何杭江,周国民. 目标检测算法研究综述[J]. 计算机工程与应用, 2018, 54(13): 11-18, 33.
FANG L P, HE H J, ZHOU G M. Research overview of object detection methods[J]. *Computer Engineering and Applications*, 2018, 54(13): 11-18, 33. (in Chinese)
- [3] 徐胜军,欧阳朴衍,郭学源,等. 多尺度特征融合空洞卷积 ResNet 遥感图像建筑物分割[J]. 光学精密工程, 2020, 28(7): 1588-1599.
XU S J, OUYANG P Y, GUO X Y, *et al.* Building segmentation in remote sensing image based on multiscale-feature fusion dilated convolution resnet [J]. *Opt. Precision Eng.*, 2020, 28(7): 1588-1599. (in Chinese)
- [4] 陈彦彤,李雨阳,吕石立,等. 基于深度语义分割的多源遥感图像海面溢油监测[J]. 光学精密工程, 2020, 28(5): 1165-1176.
CHEN Y T, LI Y Y, LÜ S L, *et al.* Research on oil spill monitoring of multi-source remote sensing image based on deep semantic segmentation [J]. *Opt. Precision Eng.*, 2020, 28(5): 1165-1176. (in Chinese)
- [5] REN S Q, HE K M, GIRSHICK R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [6] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: Single Shot MultiBox Detector[M]. *Computer Vision-ECCV 2016*. Cham: Springer International Publishing, 2016: 21-37.
- [7] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection [C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 27-30, 2016, Las Vegas, NV, USA. *IEEE*, 2016: 779-788.
- [8] 王博,董登峰,周维虎,等. 面向激光跟踪仪跟踪恢复的合作目标视觉检测[J]. 光学精密工程, 2020, 28(2): 271-282.
WANG B, DONG D F, ZHOU W H, *et al.* Visual detection of target ball for laser tracker target tracking recovery[J]. *Opt. Precision Eng.*, 2020, 28(2): 271-282. (in Chinese)
- [9] FU C Y, LIU W, RANGA A, *et al.* DSSD: deconvolutional single shot detector [EB/OL]. 2017
- [10] LI Z X, ZHOU F Q. FSSD: feature fusion single shot multibox detector [EB/OL]. 2017
- [11] 王建林,付雪松,黄展超,等. 改进 YOLOv2 卷积神经网络的多类型合作目标检测[J]. 光学精密工程, 2020, 28(1): 251-260.
WANG J L, FU X S, HUANG Z C, *et al.* Multi-type cooperative targets detection using improved YOLOv2 convolutional neural network [J]. *Opt. Precision Eng.*, 2020, 28(1): 251-260. (in Chinese)
- [12] 王俊强,李建胜,周学文,等. 改进的 SSD 算法及其对遥感影像小目标检测性能的分析[J]. 光学学报, 2019, 39(6): 0628005.
WANG J Q, LI J S, ZHOU X W, *et al.* Improved SSD algorithm and its performance analysis of small target detection in remote sensing images [J]. *Acta Optica Sinica*, 2019, 39(6): 0628005. (in Chinese)

-
- [13] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 936-944.
- [14] 姚群力, 胡显, 雷宏. 基于多尺度卷积神经网络
的遥感目标检测研究[J]. 光学学报, 2019, 39
(11): 1128002.
- YAO Q L, HU X, LEI H. Object detection in re-
mote sensing images using multiscale convolutional
neural networks[J]. *Acta Optica Sinica*, 2019, 39
(11): 1128002. (in Chinese)