

引用格式: SHI Wen-xu, TAN Dai-lun, BAO Sheng-li. Feature Enhancement SSD Algorithm and Its Application in Remote Sensing Images Target Detection[J]. *Acta Photonica Sinica*, 2020, 49(1):0128002

史文旭, 谭代伦, 鲍胜利. 特征增强 SSD 算法及其在遥感目标检测中的应用[J]. *光子学报*, 2020, 49(1):0128002

特征增强 SSD 算法及其在遥感目标检测中的应用

史文旭^{1,2}, 谭代伦³, 鲍胜利^{1,2}

(1 中国科学院成都计算机应用研究所, 成都 610081)

(2 中国科学院大学, 北京 100049)

(3 西华师范大学 数学与信息学院, 四川 南充 637009)

摘 要: 为了提高对复杂场景下多尺度遥感目标的检测精度, 提出了基于多尺度单发射击检测(SSD)的特征增强目标检测算法. 首先对 SSD 的金字塔特征层中的浅层网络设计浅层特征增强模块, 以提高浅层网络对小目标物体的特征提取能力; 然后设计深层特征融合模块, 替换 SSD 金字塔特征层中的深层网络, 提高深层网络的特征提取能力; 最后将提取的图像特征与不同纵横比的候选框进行匹配以执行不同尺度遥感图像目标检测与定位. 在光学遥感图像数据集上的实验结果表明, 该算法能够适应不同背景下的遥感目标检测, 有效地提高了复杂场景下的遥感目标的检测精度. 此外, 在拓展实验中, 文中算法对图像中的模糊目标的检测效果也优于 SSD.

关键词: 遥感图像; 深度学习; 目标检测; 多尺度特征; 特征金字塔

中图分类号: TP753

文献标识码: A

doi:10.3788/gzxb20204901.0128002

Feature Enhancement SSD Algorithm and Its Application in Remote Sensing Images Target Detection

SHI Wen-xu^{1,2}, TAN Dai-lun³, BAO Sheng-li^{1,2}

(1 *Chengdu Institute of Computer Application, Chinese Academy of Sciences, Chengdu 610081, China*)

(2 *University of Chinese Academy of Sciences, Beijing 100049, China*)

(3 *School of Mathematics and Information, China West Normal University, Nanchong, Sichuan 637009, China*)

Abstract: In order to improve the detection accuracy of multi-scale remote sensing ship targets in complex scenes, a feature enhancement single shot multi-scale detector is proposed. Firstly, the shallow feature enhancement module is designed to improve the feature extraction ability of the shallow network in the pyramid structure of Single Shot MultiBox Detector (SSD). Then the deep feature fusion module is designed to replace the deep network in the pyramid structure of SSD to improve the feature extraction ability of deep network. Finally, the image features are matched with candidate frames of different aspect ratios to adapt to remote sensing image targets of different scales. The experiments tested on the optical remote sensing image dataset demonstrate that the proposed method can adapt to target detection under different background and effectively improve the detection performance of multi-scale remote sensing targets in complex scenes. On the extended experiment, the proposed method performance over SSD in blurry target detection.

Key words: Remote sensing images; Deep learning; Target detection; Multi-scale feature; Feature pyramid

OCIS Codes: 280.0280; 100.3008; 100.2000; 110.2970

基金项目: 国家自然科学基金(No.11871059), 四川省新一代人工智能重大专项(No. 2018GZDZX0036), 四川省科技厅重点研发项目(No. 2018SZ0040), 四川省教育厅自然科学基金(Nos.18ZA0469, 15ZA0152), 四川省科技计划资助(No. 2019YFG0299)

第一作者: 史文旭(1995-), 男, 硕士研究生, 主要研究方向为深度学习, 遥感图像目标检测. Email: shiwenxu1995@126.com

导师(通讯作者): 鲍胜利(1973-), 男, 正研级高级工程师, 博士, 主要研究方向为深度学习、智能信息处理. Email: baohigh@casit.com.cn

收稿日期: 2019-07-29; 录用日期: 2019-10-22

<http://www.photon.ac.cn>

0 引言

光学遥感图像目标检测是计算机视觉领域重要的研究课题之一,在军用和民用领域都具有重要的应用价值.但遥感图像场景复杂多变,目标尺度形态各不相同且遥感图像容易受光照、大气颗粒等因素影响,使得遥感图像中的多尺度目标检测方法面临诸多挑战,因此,对于复杂场景下的多尺度遥感目标的检测研究具有重要的研究意义^[1].

近年来,随着 AlexNet 网络^[2]在 2012 年的 ImageNet 图像^[3]分类比赛中取得巨大的成功,基于深度学习的目标检测方法得到了迅猛的发展.目前基于深度学习的目标检测算法可以分为以 Fast R-CNN^[4], Faster R-CNN^[5] 为代表的双阶段目标检测算法和以统一实时目标检测算法 (You Only Look Once, YOLO)^[6], 多尺度单发射击检测算法 (Single Shot MultiBox Detector, SSD)^[7] 和 CornerNet^[8] 为代表的单阶段目标检测算法.双阶段目标检测算法流程分为候选区域的提取和目标定位与分类两个阶段,检测精度高但是效率较低.单阶段目标检测算法直接在输入图像上划分区域并进行目标的定位与分类,实现了检测精度与检测效率的兼顾,受到了更多学者的青睐,其中基于 YOLO 和 SSD 算法又相继衍生出了 YOLO V2^[9], YOLO V3^[10], 反卷积单发检测网络算法 (Deconvolutional Single Shot Detector, DSSD)^[11], 多尺度单发射击检测增强算法 (Enhancement of SSD, ESSD)^[12], 多尺度单发射击检测特征融合算法 (Feature-fused SSD, FFSSD)^[13] 等众多版本,在检测效率和检测精度上都得到了不同程度的提升.由于深度学习算法具有较强的特征提取能力,基于深度学习的目标检测方法被推广至各个领域,不少学者开始利用深度学习方法对遥感图像目标进行检测.文献^[14-15]基于 Faster R-CNN 算法并结合特征融合的思想,对遥感图像中舰船、油罐等目标进行验证实验,取得较好效果但检测效率较低.文献^[16]基于 SSD 算法并参考 DenseNet 网络^[17] 和特征金字塔 (Feature Pyramid Network, FPN) 网络^[18] 提出改进 SSD 算法并对遥感图像进行检测,该算法一定程度上提高了遥感图像目标检测的精度,但难以满足实时应用的需求.文献^[19-20]基于 SSD 算法提出了基于多尺度卷积神经网络的飞机目标检测算法,提高了遥感图像中的多尺度飞机目标的检测精度,但检测速度大幅降低,同时该方法对小目标的检测精度较差.深度学习方法虽然一定程度上提高了遥感图像目标检测的精度与效率,但依然没有达到预期的效果,可能有以下两个原因:1)与常规的自然场景不同,地面物体目标在高分遥感影像中尺度多样,从单一的特征层提取特征图用来目标检测忽略了上下文信息的关联性,尤其是对小目标的检测效果不佳;2)遥感图像场景复杂多变且容易受光照等因素影响噪声严重,在深度学习特征提取过程中容易出现因正负样本的不平衡导致的漏检错检等情况.

为实现对遥感图像目标的实时高效的检测,本文提出了一种特征增强的 SSD 算法 (Feature Enhancement Single Shot MultiBox Detector, FESSD),设计浅层特征增强模块 (Shallow Feature Enhancement Module, SFF),通过多分支卷积操作充分提取浅层网络特征,增强浅层网络对遥感图像小目标的特征提取能力.设计深层特征增强模块 (Deep Feature Enhancement Module, DFF),利用双路径网络提高深层网络对遥感图像目标的特征提取能力.设计新的损失函数,提高网络检测性能.所提方法在光学遥感数据集上进行分析实验,实验结果表明,本文所提的 FESSD 能够在满足实时检测的基础之上,实现对复杂背景下不同尺度遥感图像目标的高精度检测.

1 特征增强 SSD 算法 (FESSD)

1.1 SSD 算法

SSD 算法是 LIU W 等^[7]提出的一种目标检测算法,该算法采用金字塔结构的特征层组对目标进行分类和定位,算法框架结构如图 1 所示.在金字塔结构的特征层组中,浅层网络 (Conv4_3、FC7) 特征图感受野较小,其主要用来检测小目标物体,深层网络 (Conv8_2、Conv9_2、Conv10_2、Conv11_2) 特征图感受野较大用来检测大目标物体.该算法提取多层特征进行目标检测,缓解了单层目标检测算法对检测网络所带来的负担,提高了检测效率.

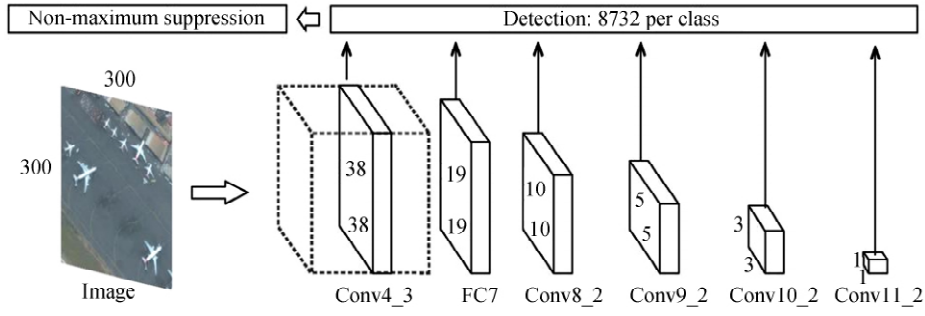


图 1 SSD 算法框架图

Fig.1 SSD overall framework

1.2 FESSD 算法

SSD 算法虽然采用多尺度特征对目标进行检测,但对遥感图像进行检测时由于浅层特征语义信息表达能力较差,造成遥感图像小目标物体检测精度差,此外,深层特征丢失了大量遥感图像的细节信息,不利于遥感目标检测的定位.

为解决上述问题,以 SSD 目标检测算法为基础,提出了一种用于遥感图像目标检测的 FESSD 算法,主要对 SSD 算法做以下两个方面的改进,一是参考多分支卷积 Inception 操作^[21]和残差网络 ResNet^[22]的思想设计浅层特征增强模块(Shallow Feature Enhancement Module, SFE)对 Conv4_3、FC7 层进行特征增强操作,二是参考 Dual Path Network^[23]的思想,设计深层特征增强模块(Deep Feature Enhancement Module, DFE)替换原始 SSD 算法中的 Conv8_2、Conv9_2、Conv10_2,并命名为 DFE_8、DFE_9、DFE_10.最后,将上述得到的不同尺度的特征送入检测网络,与不同纵横比的候选框进行匹配以适应不同尺度大小的遥感图像目标,并通过 Non-Maximum Suppression 操作^[24]进行优化处理,FESSD 算法的整体结构框架如图 2(a)所示.

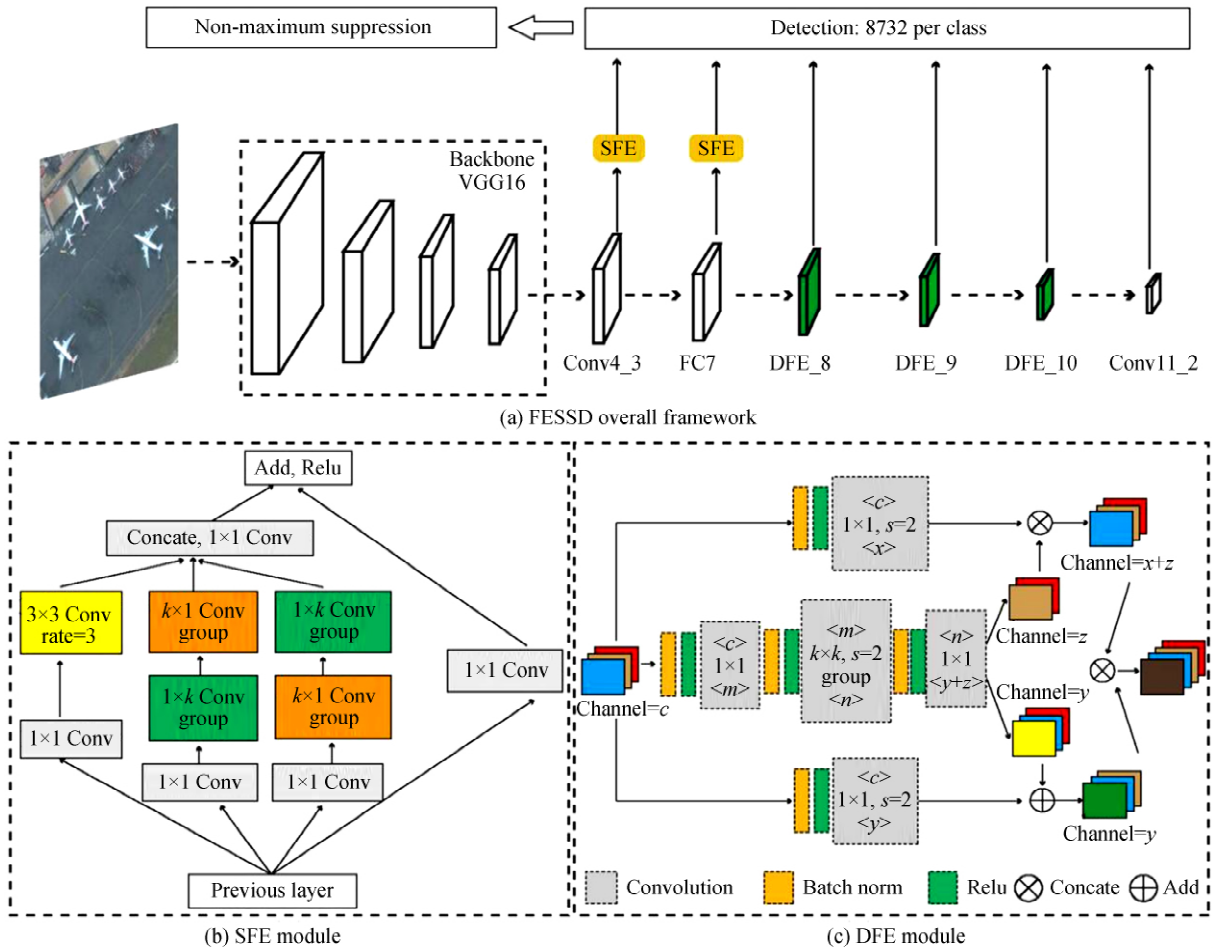


图 2 FESSD 算法框架图、浅层特征增强模块和深层特征增强模块图

Fig.2 FESSD overall framework, shallow feature enhancement module and deep feature enhancement module

FESSD 算法为充分利用浅层网络特征的细节表达能力,提高对遥感图像中的小目标物体的检测精度,设计浅层特征增强模块(Shallow Feature Enhancement Module,SFF),具体结构如图 2(b)所示.其基于多分支卷积操作^[21]的思想,使用 $1 \times 1, 3 \times 3$ 的卷积使得网络学习到更多的非线性关系的同时结合残差网络^[22]和空洞卷积使网络捕获更多的特征信息,拓宽特征的感受野,此外将 $k \times k$ 卷积分解为 $k \times 1$ 和 $1 \times k$,在保证感受野大小不变的同时使用分组卷积操作,节省该模块消耗的时间,另一个分支的不同之处在于将 $k \times 1$ 和 $1 \times k$ 卷积反向组合,将得到的特征结果进行 Concat 特征融合操作并与残差分支的 1×1 卷积结果进行 Add 特征融合操作.此外,FESSD 算法为提高深层网络的特征提取能力,增强网络对目标位置的回归能力,设计深层特征增强模块(Deep Feature Enhancement Module,DFE),具体结构如图 2(c)所示,该模块基于 Dual Path Network^[23]的思想,并结合残差网络 ResNet^[22]和密集网络 DesNet^[25]对特征提取的特点,不仅深化了整个网络,而且具有从浅层到深层的隐式的特征融合,使得深层网络的特征信息更加丰富,同时使用分组卷积加快模块特征处理效率.

图 3 为本文所提 FESSD 算法与 SSD 算法在 Conv4_3 层、Conv8_2(DFE8)层输出的特征对比图,图 3(b)和图 3(c)为 FESSD 算法输出的特征图,图 3(d)、图 3(e)为 SSD 算法输出的特征图.显然图 3(b)相比于图 3(d)舰船信息更加清晰,尤其是小目标舰船,图 3(b)仍高亮显示,具有更强的语义信息,图 3(c)相比于图 3(e)大目标舰船特征信息更为丰富,理论上在图 3(b)和图 3(c)的检测效果会明显优于图 3(d)、图 3(e).

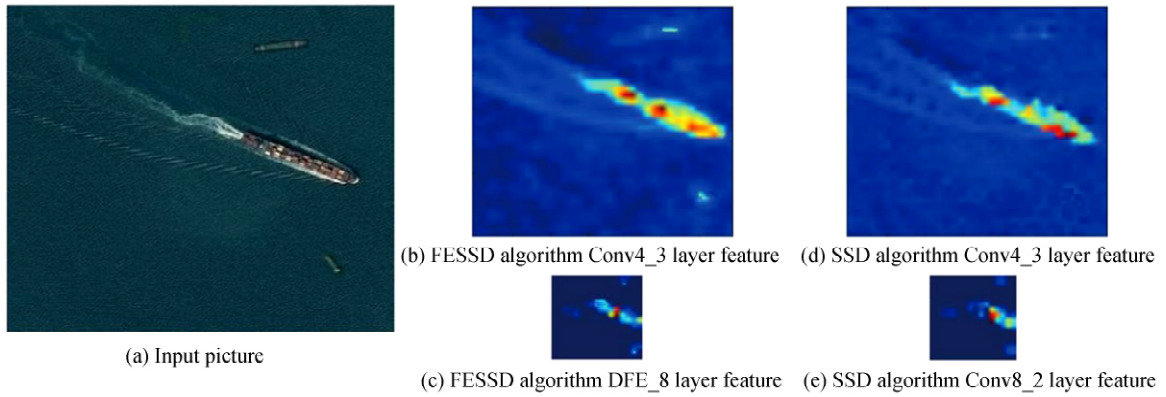


图 3 特征增强前后特征图对比

Fig.3 Comparison of feature maps before and after feature enhancement

相比于 SSD 算法,本文所提的 FESSD 算法针对原始 Conv4_3、FC7 层引入浅层特征增强模块(CFE),可以弥补 SSD 算法浅层特征语义信息表达能力较差问题,另一方面用深层特征增强模块(DFE)代替原始网络中的 Conv8_2、Conv9_2、Conv10_2 层,在提高深层网络特征提取能力的同时大幅度提高目标检测的效率.3.4 节的消融实验将详细分析上述两个模块的有效性.

1.3 候选框选择与损失函数

为实现对遥感图像中不同尺度的目标进行检测,设计不同纵横比的候选框进行匹配以适应不同尺度大小的遥感图像目标,假设选取 m 个层的卷积特征用于目标,则第 i 层的特征图候选框尺寸为

$$S_i = S_{\min} + \frac{S_{\max} - S_{\min}}{m-1}(i-1) \quad i \in [1, m] \quad (1)$$

式中, S_{\min} 和 S_{\max} 分别取值为 0.2 和 0.95,表示浅层和深层的候选框与特征图的尺度系数.候选框的宽和高的比取值为 $a_r \in \{1, 2, 3, \frac{1}{2}, \frac{1}{3}\}$,则候选框的宽为 $w_i = S_i \cdot \sqrt{a_r}$,候选框的高位 $h_i = S_i / \sqrt{a_r}$,候选框的中心坐标为 $(\frac{i+0.5}{|f_k|}, \frac{j+0.5}{|f_k|})$, $i, j \in [0, |f_k|]$,其中 $|f_k|$ 表示第 k 个特征的尺寸.

此外,针对模型训练过程中遥感图像正、负样本划分不平衡导致模型退化的问题,本文在 SSD 算法损失函数的基础之上结合聚焦分类损失函数(focal classification loss)^[26]对模型进行优化训练,表示为

$$L(x, c, p, l, g) = \frac{1}{N} [L_{fl}(x, c, p) + aL_{loc}(x, l, g)] \quad (2)$$

式中, N 表示与真实框相匹配的候选框的个数, x 为输入图像, c 为目标类别, p 为预测类别概率, l 为候选框, g 代表真实的标签框, a 为两者的权重. $L_{fl}(x, c, p)$ 和 $L_{loc}(x, l, g)$ 分别为聚焦分类损失和边框回归损失. 其中, $L_{loc}(x, l, g)$ 借鉴了 Faster R-CNN 的位置回归函数, 表示为

$$L_{loc}(x, l, g) = \sum_{i \in \text{Pos}} \sum_{m \in \{c_x, c_y, w, h\}} x_{ij}^k \text{smooth}_{L_1}(l_i^m - g_j^m) \quad (3)$$

式中, x_{ij}^k 代表针对类别 k , 第 i 个候选框与第 j 个真实框的结果是否保持一致, 若一致则取值为 1 否则取值为 0. l_i^m 表示候选框的值, g_j^m 为真实框的值.

式(2)中的 $L_{fl}(x, c, p)$ 结合聚焦分类损失函数采用交叉熵计算损失, 表示为

$$L_{fl}(x, c, p) = - \sum_{i \in \text{Pos}} x_{ij}^k p \log(c_i^k) - \sum_{i \in \text{neg}} (1 - p) \log(c_i^0) \quad (4)$$

式中, c_j^0 代表正确且类别为背景预测框的概率, c_i^k 为利用 Softmax 函数计算的概率值.

2 实验与分析

2.1 实验数据

为评估本文所提的 FESSD 算法的检测性能, 采集不同场景下含有飞机 (aircraft)、油罐 (oiltank)、舰船 (ship)、立交桥 (overpass)、操场 (playground) 共 5 类目标, 幅面大小约为 500 pixel \times 375 pixel 的光学遥感影像共计 1 554 张, 共含有实例样本 7 969 个, 其中 5 731 个目标作为训练集, 2 238 个作为测试集. 后续根据任务需求, 会继续扩充数据集, 增加样本实例的个数. 相比于有 24 640 个实例样本 20 类目标的 PASCAL VOC 数据集^[27], 本文数据集足够达到可以对相应遥感图像进行检测的需求. 此外, 依据公开数据集对目标大小划分的标准 (Small object: area $<$ 32²; Medium object: 32² $<$ area $<$ 96²; Large object: area $>$ 96²) 对采集的光学遥感数据进行划分, 分布情况如表 1 所示, 与目前经典的遥感影像数据集 NWPU VHR-10^[28], AID dataset^[29] 相比, 本文设计的数据主要以中小目标为主, 且样本分布密集增加了目标检测的难度.

表 1 自采集数据集统计

Table 1 Self-collected data set statistics

Data set	Class	# Image	# Instances	Target amount(percentage)		
				Small	Medium	Large
Training set	Aircraft	320	2 382	1 491	890	1
	Oiltank	220	1 735	884	845	6
	Ship	400	1 317	659	588	70
	Overpass	140	143	0	0	143
	Playground	140	154	0	8	146
Test set	Aircraft	101	970	571	398	1
	Oiltank	78	759	374	385	0
	Ship	70	415	210	195	10
	Overpass	36	40	0	7	33
	Playground	49	54	0	5	49

图 4 为本文数据集的样本示例, 从左到右分别为分辨率约 1 m 的靠近停机位的 20 架民用客机, 23 个油罐, 开阔海域的 16 艘船, 分辨率约 5 m 的立交桥 1 座立交桥和操场. 由于一些飞机和舰船目标在小分辨率的情况下只有极少的像素个数, 已很难通过人眼确定是否为飞机和舰船, 因此, 仅对像素个数超过 5 个的目标

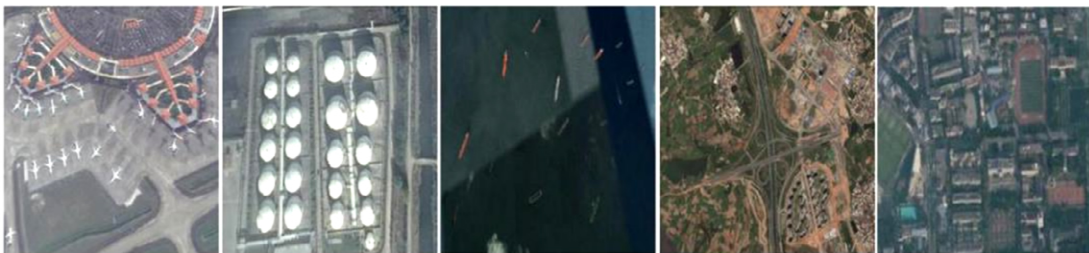


图 4 数据集部分样本示例

Fig.4 Samples of the dataset

进行标注.标注的软件为开源软件 LabelImg(版本:windows_v1.8.0),该软件很容易就能对遥感图像目标进行标记 (x, y, w, h) ,其中 (x, y) 为目标所在矩形框的左上角坐标, w 是矩形框的宽度, h 为矩形框的高度.

2.2 评估标准及训练方法

本文采用已被广泛应用于目标检测评估^[10,13,14]的平均精度均值(Mean Average Precision, mAP)和每秒帧率(Frame Per Second, FPS)作为衡量模型的标准.其中 mAP 是指多个类别精度均值(Average Precision, AP)的平均值,每个类别都可以根据准确率(precision)和召回率(recall)绘制一条曲线,其在 0 到 1 区间范围内绘制的曲线与坐标轴所围成的面积即为精度均值,具体可表示为 $AP = \int_0^1 p(r)dr$,准确率和召回率的定义为

$$\begin{cases} \text{precision} = \frac{TP}{TP+FP} \\ \text{recall} = \frac{TP}{TP+FN} \end{cases} \quad (5)$$

式中,TP 代表真正例, FN 代表假反例, FP 代表假正例.此外,每秒帧率(FPS)是目标检测的另一个重要的参考指标,其数值的大小直接反应了检测速度的快慢.

本文所提 FESSD 算法在 NVIDIA Geforce GTX1070Ti,采用 Pytorch 开源深度学习框架进行训练和测试,利用在 ImageNet 数据集上预训练的模型进行迁移学习,初始学习率设置为 0.001 并梯度衰减至 0.000 001,动量因子设置为 0.9,采用小批量 SGD 优化方式对损失函数式(2)进行优化,寻求最优解.

2.3 实验结果与分析

本文评估了 FESSD 算法的检测精度与检测速度,并与几种典型的多尺度目标检测算法进行对比,对比结果如表 2 所示.

表 2 各算法在遥感图像物体的检测精度和检测速度
Table 2 Detection accuracy and speed of objects in remote sensing images by various algorithms

Method	Backbone	Metric/%						FPS
		Aircraft	Oiltank	Ship	Over-pass	Play-ground	mAP(IOUS=0.5)	
Faster R-CNN ^[5]	VGG16	69.77	69.59	66.88	71.10	89.94	73.46	4.4
YOLO ^[6]	GoogleNet	60.26	58.47	59.14	62.21	87.57	65.53	50.6
YOLO V2 ^[9]	DarkNet19	70.12	71.24	67.25	69.41	87.50	73.10	25.3
SSD300 ^[7]	VGG16	69.91	70.73	66.39	70.90	91.87	73.96	50.4
DSSD321 ^[11]	ResNet-101	75.75	73.11	70.32	71.22	90.52	76.18	4.3
ESSD321 ^[12]	VGG16	77.26	73.26	71.06	70.85	90.87	76.66	24.6
FFSSD300 ^[13]	VGG16	75.94	73.42	70.56	70.96	90.47	76.27	28.7
FESSD(ours)300	VGG16	79.01	76.12	72.27	71.05	91.96	78.08	48.6
SSD512 ^[7]	VGG16	71.56	72.28	70.04	71.29	92.05	75.44	28.6
DSSD513 ^[11]	ResNet-101	76.48	75.71	71.62	72.05	91.25	77.42	2.1
ESSD321 ^[12]	VGG16	77.45	75.89	72.34	71.44	91.66	77.76	12.3
FFSSD512 ^[13]	VGG16	77.23	76.89	72.36	71.59	91.61	77.94	13.4
FESSD(ours)512	VGG16	80.96	78.22	73.56	71.89	92.16	79.36	26.3

由表 2 可知,本文所提的 FESSD 算法在平均精度均值(mAP)性能指标优于其他算法,每秒帧率(FPS)相对于 SSD 算法几乎没有损失.从检测精度上看,FESSD 算法对光学遥感图像数据检测的平均精度均值高于其他算法,尤其是当输入图片尺寸设定为 512×512 像素时,FESSD 算法平均精度均值为 79.36%,与 SSD,DSSD,FFSSD 算法相比分别提高了 4.12%,1.94%,1.42%,且 FESSD 算法对以中小目标为主的 aircraft,oiltank,ship 三类目标的检测精度相较于其它算法提升幅度更为明显.从检测速度上看,FESSD 算法的对光学遥感图像的检测速度较高且能够满足实时检测.实验结果表明,本文所提的 FESSD 算法能够在满足实时检测的基础之上有效的提高复杂背景下遥感图像的目标检测精度,尤其是对小目标物体的检测更具优势.

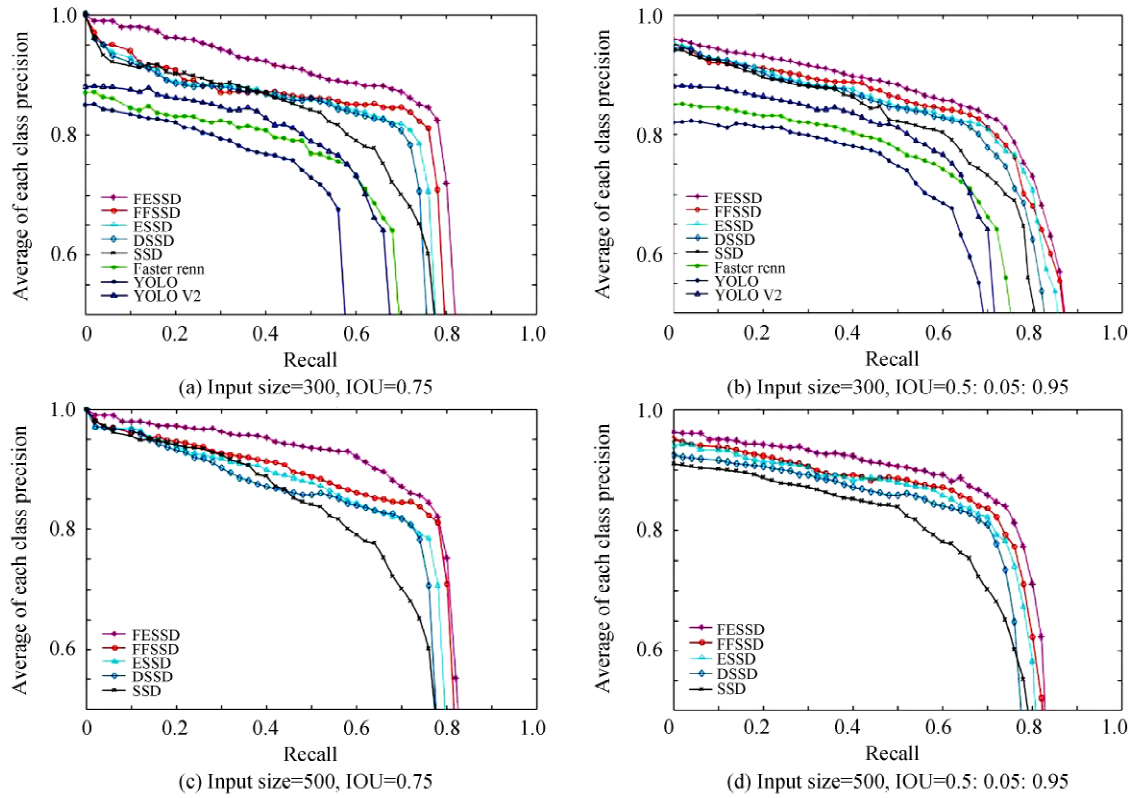


图 5 不同 IOU 情况下平均精度与召回率关系

Fig.5 Recall vs. average of each class precision graph in different IOU conditions

此外,平均精度均值(mAP)通常根据不同的 IOU(Intersection Over Union)阈值进行评估,IOU 阈值越高,对应的边框回归越准确,图 5 显示在 IOU=0.75 和 IOU=0.5 : 0.05 : 0.95 时各类别平均精度与召回率之间的关系.从检测结果可以看出,本文所提 FESSD 算法在平均精度均值(mAP)和召回率(recall)两个性能指标均优于其他方法,即 FESSD 算法对目标位置的回归能力优于其它算法.

在不同场景下,FESSD 算法对遥感图像数据集的部分检测结果示例如图 6 所示,在光照不同(图 6(a)、图 6(b))、密集分布(图 6(a)、图 6(b)、图 6(c)、图 6(d))以及小尺度遥感目标(图 6(a)、图 6(b)、图 6(c)、图 6(d)、图 6(e))等情况下,该算法均可以精确的检测出相关的目标物体,证明了本文所设计的 FESSD 算法能够有效的提高复杂背景下多尺度遥感图像中的目标物体检测能力.

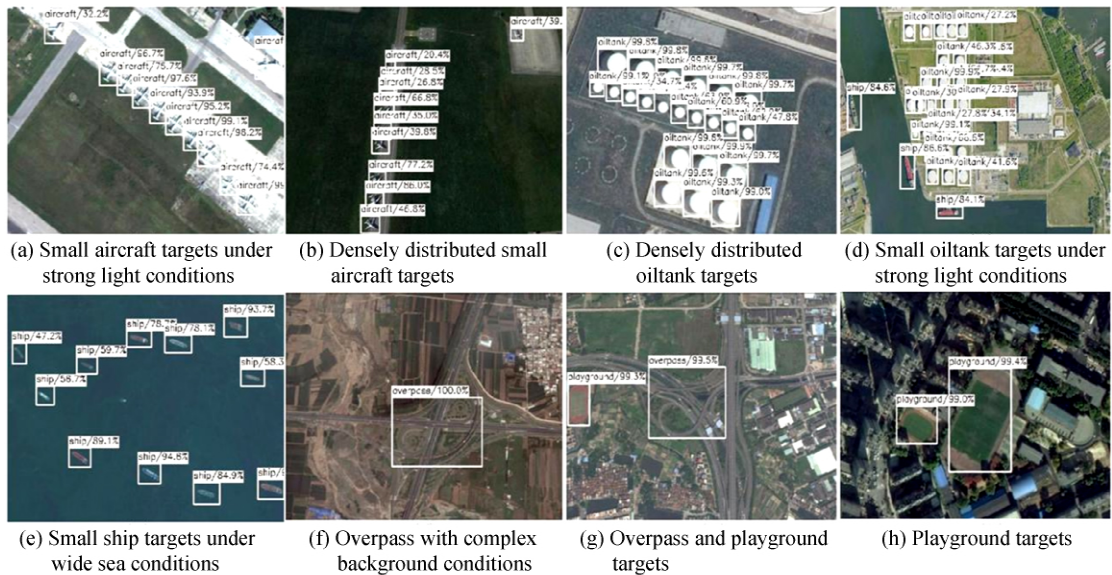


图 6 FESSD 算法对光学遥感图像检测示例

Fig.6 Examples of FESSD algorithm for optical remote sensing image detection

2.4 消融实验

为分析本文设计的浅层特征增强模块(CFE)、深层特征增强模块(DFE)对遥感图像目标检测的平均精度均值(mAP)的影响,实验设置了不同的特征增强组合方式,当输入尺度为 300×300 像素时,各模块实验效果对比如表 3 所示。

表 3 特征增强方法对光学遥感图像目标检测的影响

SSD	SFE	DFE	Metric/%					mAP	FPS
			Aircraft	Oiltank	Ship	Overpass	Playground		
✓			69.91	70.73	66.39	70.90	91.87	73.96	50.4
✓	✓		78.38	75.45	71.63	70.06	90.91	77.29	35.6
✓		✓	72.26	71.62	70.33	71.51	91.47	75.44	55.8
✓	✓	✓	79.01	76.12	72.27	71.05	91.96	78.08	48.6

由表 3 可知,在 SSD 算法的基础之上增加浅层特征融合模块(SFE),平均精度均值(mAP)由 73.96% 提升至 77.29%,整体提高了 3.33%,其中对以小目标为主的 aircraft、oiltank、ship 三类目标的检测精度分别提高了 8.47%、4.72%、5.24%,由于在 SSD 算法上增加了 SFE 模块,导致额外消耗了时间,检测速度由 50.4 FPS 降为 35.6 FPS;在 SSD 算法基础之上用深层特征增强模块(DFE)来替代原始 SSD 算法中的深层网络,平均精度均值由 73.96% 提升至 75.44%,各类别目标的检测精度得到提高的同时,由于 DFE 模块相较于原始深层网络结构效率更高,检测速度相较于 SSD 算法也得到了提升,达到 55.8 FPS;本文提出的 FESSD 算法以 SSD 算法为基础,增加上述两个模块,从检测精度上看,FESSD 算法的平均检测精度达到 78.08% 较 SSD 算法提升了 4.12%,各类别目标检测精度均得到提高,尤其是以小目标为主的 aircraft、oiltank、ship 三类目标提升幅度更为明显,从检测速度上看,FESSD 算法检测速度为 48.6 FPS,相较 SSD 算法几乎没有损失。

此外,表 4 对图 2(a)中各个独立模块的实验效果进行对比,根据各层网络所提取的特征大小可知:在 SSD 的 Conv4_3 层增加 SFE 模块能够使得 mAP 由 73.96% 提升至 76.57%,表明 SFE 模块能够有效地提高对小目标物体(Small object: Pixel area $< 32^2$)的检测精度;在 SSD 的 FC7 层增加 SFE 模块使得目标检测 mAP 提升至 77.86%,表明 SFE 模块能够有效地提高对中目标物体(Medium object: $32^2 < \text{Pixel area} < 96^2$)的检测精度;使用 DFE_8,DFE_9,DFE_10,DFE_11 代替原有 SSD 的 Conv8_2,Conv9_2,Conv10_2,Conv11_2 使得检测精度进一步提高,表明 DFE 模块能够提高对大目标物体(Large object: area $> 96^2$)的检测精度。

表 4 各独立模块对光学遥感图像目标检测的影响

mAP	+DFE_11	+DFE_10	+DFE_9	+DFE_8	+SFE2	+SFE1
73.96						
76.57						✓
77.29					✓	✓
77.45				✓	✓	✓
77.86			✓	✓	✓	✓
77.99		✓	✓	✓	✓	✓
78.08	✓	✓	✓	✓	✓	✓

上述消融实验的结果证明,本文设计的 FESSD 算法中的浅层特征增强模块(SFE)能够提高浅层网络的特征提取能力,提高对遥感图像中的小目标物体的检测精度,深层特征增强模块(DFE)能够在提高深层网络特征提取能力的同时大幅度减少网络中的参数量以提高检测速度。

3 拓展实验分析

为验证本文所提 FESSD 算法的有效性、通用性,以及对复杂背景下遥感图像目标的检测效果,除了在上述采集的训练集和测试集进行算法性能的对比之外,对于因云层、大气颗粒等问题导致的成像模糊的图像,本文将利用上述训练的模型(FESSD512, IOU=0.5)直接对国产高分辨率光学遥感卫星传回的影像数据进

行检测,并与 SSD 算法(SSD512,IOU=0.5)的检测效果进行对比,对比结果如图 7 所示。

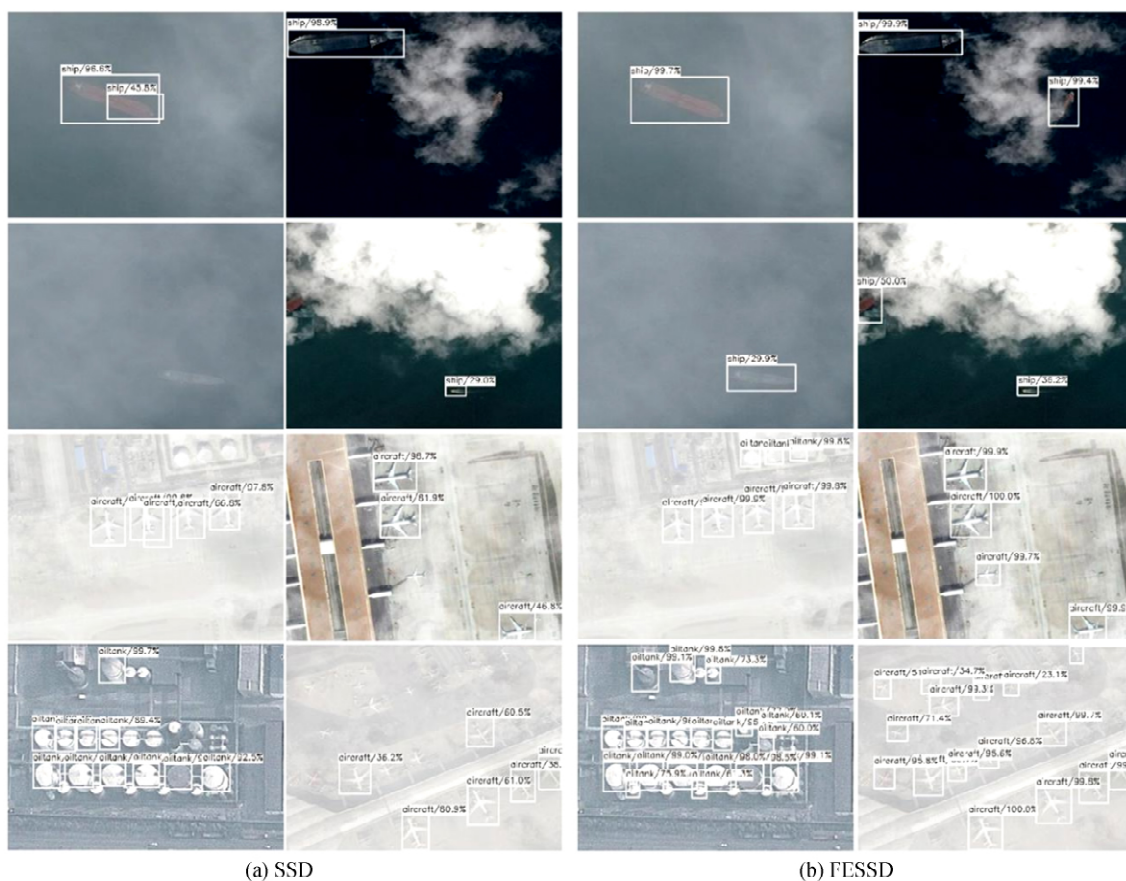


图 7 FESSD 算法和 SSD 算法拓展实验检测示例
Fig.7 Examples of FESSD and SSD algorithm migration experiment detection

如图 7 所示,本文给出了 8 张不同视点高度、不同天气情况下国产高分辨率遥感影像数据的检测实例。通过直观的对比可以看出,FESSD 算法对模糊目标的检测效果优于 SSD 算法,这是因为在进行特征增强的过程中在提高特征提取能力的同时,一定程度上增强了特征的泛化能力,增强了对模糊目标的响应能力,进而提高对模糊目标的检测性能.拓展实验结果证明本文所提算法提高了对模糊图片的特征提取能力,检测性能优于 SSD 算法。

4 结论

提出了一种特征增强的 SSD 算法(FESSD),旨在提高对复杂场景下多尺度遥感图像目标物体检测的精度.利用多分支卷积操作的思想设计浅层特征增强模块(SFE),弥补 SSD 算法浅层特征语义信息表达能力较差问题,增强了浅层网络的特征提取能力;利用双路径网络的思想设计深层特征增强模块(DFE),在提高深层网络特征提取能力的同时提高目标检测的效率.与当前多种深度学习算法相比,FESSD 算法的结构设计合理,检测精度更高且能够满足实时检测需求.从拓展训练的结果可以看出本文所提的 FESSD 算法的对模糊目标的检测效果优于 SSD 算法.在未来的研究中,我们将进一步扩展光学遥感图像数据库,优化检测算法,以进一步提高对遥感图像小目标物体的检测精度。

参考文献

- [1] BALL J E , ANDERSON D T , CHAN C S. A comprehensive survey of deep learning in remote sensing: theories, tools and challenges for the community[J]. *Journal of Applied Remote Sensing* , 2017, **11**(4): 042609.
- [2] KRIZHEVSKY A , SUTSKEVER I , HINTON G. ImageNet classification with deep convolutional neural networks[J]. *Advances in Neural Information Processing Systems* , 2012, **25**:1106-1114.
- [3] RUSSAKOVSKY O , DENG J , SU H , et al. ImageNet large scale visual recognition challenge[J]. *International Journal of Computer Vision* , 2015, **115**(3):211-252.
- [4] GIRSHICK R. Fast R-CNN[J].*Computer Science* , 2015, **6**: 1440-1448.

- [5] REN S, HE K, GIRSHICK R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks. [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, **39**(6):1137-1149.
- [6] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection[C]. Conference on Computer Vision and Pattern Recognition, 2016, **6**: 779 -788.
- [7] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: single shot MultiBox detector[C]. European Conference on Computer Vision, 2016, **6**: 21-27.
- [8] HEI L, JIA D. CornerNet: detecting objects as paired keypoints[C]. European Conference on Computer Vision, 2018, **9**: 734-750.
- [9] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. Conference on Computer Vision and Pattern Recognition, 2017, **7**: 6517-6525.
- [10] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. *arXiv*:1804.02767, 2018.
- [11] FU C, LIU W, RANGAA, *et al.* DSSD: Deconvolutional single shot detector[J]. *arXiv*:1701.06659, 2017.
- [12] LENG J, LIU Y. An enhanced SSD with feature fusion and visual reasoning for object detection[J]. *Neural Computing and Applications*, 2018, 1-10.
- [13] CAO G, XIE X, YANG W, *et al.* Feature-fused ssd: fast detection for small objects[J]. *arXiv*:1709.05054, 2017.
- [14] LI Jian-wei, QU Chang-wen, PENG Shu-juan, *et al.* Ship detection in SAR images based on convolutional neural network [J]. *Systems Engineering and Electronics*, 2018, **40**(9): 1953-1959.
李健伟, 曲长文, 彭书娟, 等. 基于卷积神经网络的 SAR 图像舰船目标检测[J]. 系统工程与电子技术, 2018, **40**(9): 1953-1959.
- [15] WANG Jin-chuan, TAN Xi-cheng, WANG Zhao-hai, *et al.* Faster R-CNN deep learning network based object recognition of remote sensing image[J]. *Geo-Information Science*, 2018, **20**(10): 1500-1508.
王金传, 谭喜成, 王召海, 等. 基于 Faster R-CNN 深度网络的遥感影像目标识别方法研究[J]. 地球信息科学学报, 2018, **20**(10): 1500-1508.
- [16] HUANG G, LIU Z, VAN DER MAATEN L, *et al.* Densely connected convolutional networks[C]. Conference on Computer Vision and Pattern Recognition, 2017, **7**: 2261-2269.
- [17] LIN T Y, DOLLAR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection[C]. Conference on Computer Vision and Pattern Recognition, 2017, **7**: 936-944.
- [18] WANG Jun-qiang, LI Jian-sheng, ZHOU Xue-wen, *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.
王俊强, 李建胜, 周学文, 等. 改进 SSD 算法及其在遥感影像小目标检测性能分析[J]. 光学学报, 2019, **39**(6): 0628005.
- [19] XING Peng, XU Yue-Lei, TANG Hong, *et al.* Fast airplane detection incorporating multi-layer features offully convolutional Networks[J]. *Acta Optica Sinica*, 2018, **38**(3): 0315003.
辛鹏, 许悦雷, 唐红, 等. 全卷积网络多层特征融合的飞机快速检测[J]. 光学学报, 2018, **38**(3): 0315003.
- [20] CHEN Z, ZHANG T, OUYANG C. End-to-end airplane detection using transfer learning in remote sensing images[J]. *Remote Sensing*, 2018, **10**(1): 139.
- [21] SZEGEDY C, LIU W, JIA Y, *et al.* Going deeper with convolutions[C]. Conference on Computer Vision and Pattern Recognition, 2015, **6**: 1-9.
- [22] HE K, ZHANG X, REN S, *et al.* Deep residual learning for image recognition[C]. Conference on Computer Vision and Pattern Recognition, 2016, 770-778.
- [23] CHEN Y, LI J, XIAO H, *et al.* Dual path networks[J]. *arXiv*:1707.01629, 2017.
- [24] QIU S, WEN G, DENG Z, *et al.* Accurate non-maximum suppression for object detection in high-resolution remote sensing images[J]. *Remote Sensing Letters*, 2018, **9**(3): 238-247.
- [25] HUANG G, LIU Z, LAURENS V D M, *et al.* Densely connected convolutional networks[C]. Conference on Computer Vision and Pattern Recognition, 2017, **7**: 4700-4708.
- [26] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, **99**: 2999-3007.
- [27] EVERINGHAM M, ESLAMI S M A, GOOL L V, *et al.* The pascal visual object classes challenge: a retrospective[J]. *International Journal of Computer Vision*, 2015, **111**(1): 98-136.
- [28] CHENG G, ZHOU P, HAN J. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, **54**, 7405-7415.
- [29] XIA G.S, HU J, HU F, *et al.* Aid: A benchmark dataset for performance evaluation of aerial scene classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, **55**, 3965-3981.

Foundation item: The National Natural Science Foundation of China (No.11871059), the New Generation of Artificial Intelligence Major Program in Sichuan Province(No.2018GZDZX0036), the Key Program of Sichuan Science and Technology Department(No.2018SZ0040), the Natural Science Foundation of Sichuan Education Department(Nos.18ZA0469, 15ZA0152), Sichuan Science and Technology Program (No.2019YFG0299)