

基于支持集特征增强的 Meta Faster R-CNN 小样本目标检测

马俊光, 文峰, 殷向阳

(沈阳理工大学 信息科学与工程学院, 沈阳 110159)

摘要: 小样本目标检测的元学习方法能快速适应少量训练样本, 较好解决现有常规检测模型泛化能力不强、适应新任务速度缓慢、鲁棒性差的问题, 具有较高的实际应用价值, 但该方法对支持集特征利用能力不足、检测精度不高。为此, 基于支持集特征增强, 针对元学习 SOTA 算法 Meta Faster R-CNN 进行改进, 从支持集背景抑制与目标特征增强两个角度出发, 削弱与待查询目标无关的背景信息并加强支持集内部特征之间的联系, 构建一种检测性能更高的小样本目标检测算法。实验结果表明: 在 PASCAL VOC Novel Set 数据集上的元测试阶段, 本文改进算法在 1-shot、2-shot、3-shot、5-shot、10-shot 下的平均精度均值 (mAP@0.5) 较原算法分别提升了 0.066%、12.038%、12.289%、10.073%、9.539%; 在元微调后的测试阶段, 本文改进算法的 mAP@0.5 较原算法有所提升或基本持平; 增强支持集特征能够有效提升小样本目标检测精度。

关键词: 小样本目标检测; 元学习; Meta Faster R-CNN; 背景抑制; 特征增强

中图分类号: TP391.4 **文献标志码:** A **DOI:** 10.3969/j.issn.1003-1251.2025.02.008

Meta Faster R-CNN Few-shot Object Detection Based on Support Set Feature Enhancement

MA Junguang, WEN Feng, YIN Xiangyang

(Shenyang Ligong University, Shenyang 110159, China)

Abstract: The meta-learning method for few-shot object detection can quickly adapt to a small number of training samples, better solve the problems that the existing conventional detection model has weak generalization ability, slow adaptation to new tasks, and poor model robustness. The method has high practical application value, but it has insufficient ability to utilize the features of the support set, and the detection accuracy is not high. In this case, from the view of background attenuation of the support set and enhancement of target features and based on the support set feature enhancement, Meta Faster R-CNN, a meta-learning SOTA algorithm is improved to weaken the background information that has nothing to do with the target to be queried and to strengthen the connection between features within the support set. In this way, a few-shot object detection algorithm with a higher detection performance is constructed. The experimental results show that in the meta-testing stage on the PASCAL VOC Novel Set dataset, the mean average precision (mAP@0.5) of the improved algorithm in 1-shot, 2-shot, 3-shot, 5-shot and 10-shot is improved by 0.066%, 12.038%, 12.289%, 10.073% and 9.539%, respectively; in the testing stage after meta-fine-tun-

ing, the mAP@0.5 of the improved algorithm in this paper also improves or is basically the same as that of the original algorithm; the enhancement of the support set feature can effectively improve the few-shot object detection accuracy.

Key words: few-shot object detection; meta-learning; Meta Faster R-CNN; background attenuation; feature enhancement

传统的目标检测方法依赖于大量标注数据,如 PASCAL VOC、MS COCO 等目标检测数据集,其每个对象类别均有成百上千个标注示例。然而,标注数据既昂贵又耗时,且由于真实数据存在固有的长尾分布,大部分对象的样本数量比较稀少^[1],特别是军事遥感检测、非常见疾病检测、工业生产中的残次品检测等,由于隐私、安全和高标记成本等因素,很难获得此类场景的大量标注图像^[2-3]。为解决上述问题,小样本目标检测(few-shot object detection, FSOD)应运而生,其在对基础类别的大量标记示例进行训练后,仅通过少数示例来识别新的对象类。然而,小样本目标检测技术尚处于学术探索阶段,其性能与大规模数据集下的目标检测技术仍存在较大差距^[2]。

早期小样本目标检测研究大都简单地将小样本分类器转换为检测器,未真正解决小样本目标检测中的定位问题,如 LSTD^[4]等。随着元学习方法的快速发展,越来越多的研究将元学习引入小样本目标检测领域。元学习模型在学习一系列任务的过程中掌握在新任务中快速学习和适应的能力,符合小样本目标检测的基本需求。早期的元学习方法以 YOLOv2、Faster R-CNN 等常规目标检测方法为基础进行改进,初步显示了其较好的检测效果,如 Meta-YOLO、Meta R-CNN 等。Wang 等^[5]在 2020 年提出了两阶段微调方法(TFA),其性能超过了同时期的元学习方法。此后,小样本目标检测领域涌现了大量新理论、新技术^[6-7]。针对目前提出的小样本目标检测方法,研究者们按照不同的标准对其进行分类^[1-2],其中典型的分类方法是将其分为基于元学习的方法和基于迁移学习的方法两类^[1]。

基于元学习的小样本目标检测方法通常遵循两阶段训练流程:第一阶段为在基类上进行的元训练阶段;第二阶段为包含基类和新类的元微调阶段。元学习器学习不同类别物体的元知识,通过特征重加权等方法使模型泛化到新类,通常特征提取器与元学习器联合训练,且学习过程遵循情节学习的方式。元学习方法下的小样本目标检测通常采用 N -way K -shot 的框架对问题进行求

解,其将整个任务分解为大量的元任务,每个元任务将数据集划分为支持集和查询集,其中支持集包含 N 个类别,每个类别有 K 个样本,查询集通常为单个样本,该方法又称为 N -way K -shot 目标检测,Meta-DETR^[8]、Meta Faster R-CNN^[9]、UNP^[10]等均属于该类元学习方法。

基于迁移学习的小样本目标检测方法普遍采用联合微调策略,首先在基类上训练,然后在一个类别平衡的小样本基类和新类训练集上微调整个模型,其效果在 TFA 提出前不及元学习方法。TFA 保持模型特征提取部分不变,仅微调最后一层便可提升检测精度,说明从基类学习到的特征表示能迁移到新类。Context-Transformer^[11]、FsDetView^[12]、FSSP^[13]、DMNet^[14]、TeSNet^[15]、BC-YOLO^[16]等均属于该类迁移学习方法。

迁移学习方法由于缺乏任务间知识共享,难以利用之前任务中的学习成果,存在对新任务适应能力差的问题,元学习方法则可实现对新任务的快速适应,较好解决现有常规检测模型泛化能力不强、鲁棒性差的问题。元学习方法中的 SOTA 算法 Meta Faster R-CNN 能够同时较好地保持基类和新类的检测精度,即使不进行元微调,也可在新类上取得较好的检测效果,且 Meta Faster R-CNN 算法采用残差网络(ResNet)作为骨干网络,比 Meta-DETR 等算法的参数数量更少,资源占用更低。

尽管 Meta Faster R-CNN 算法在小样本目标检测领域已经取得了一定的成果,但仍存在不能良好地利用支持集特征而导致检测精度不高的问题,故本文以 Meta Faster R-CNN 算法为基础,进行如下改进。

1) 从支持集背景抑制角度出发,引入背景抑制(background attenuation, BA)模块^[17],削弱与待查询目标无关的背景信息,较好地区分前景与背景,该模块在通道维度上进行特征加强与无用特征的抑制,增强支持集特征,从而提高检测精度。

2) 从目标特征增强角度出发,利用一种概念简单且有效的无参注意力模块(simple, parameter-free attention module, SimAM)同时在通道和空间

维度上对支持集特征图进行特征增强,加强支持集内部特征之间的联系,更好地利用支持集特征,从而提高检测精度。

1 基于改进 Meta Faster R-CNN 的小样本目标检测算法

Meta Faster R-CNN 算法主要包括两部分:基于度量学习的轻量化粗粒度非线性原型匹配网络

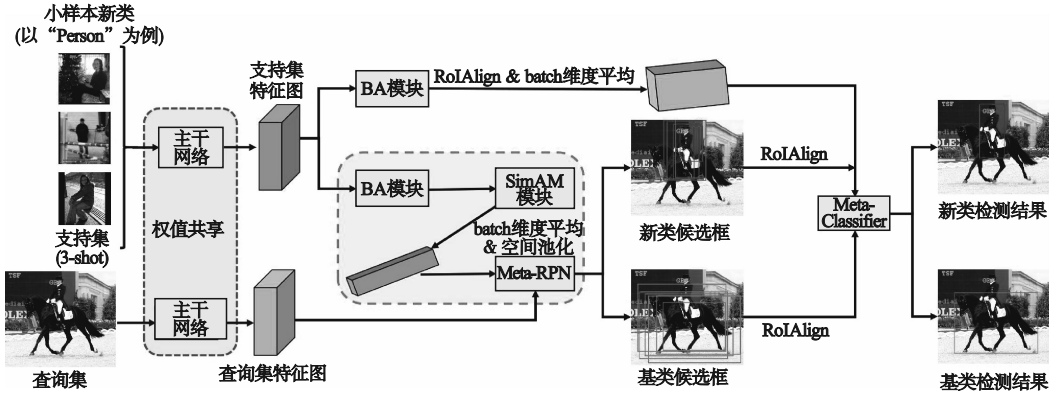


图1 改进后的 Meta Faster R-CNN 算法框架

Fig. 1 The framework of the improved Meta Faster R-CNN algorithm

本文算法框架分为支持集与查询集两个分支,采用 ResNet101 的前四个卷积模块 $\text{conv1}_x \sim \text{conv4}_x$ 作为特征提取主干网络,两个分支首先经过权值共享的主干网络,提取出 res_4 支持集特征与 res_4 查询集特征。 res_4 支持集特征图的处理过程也分为两个分支:一个分支经过 BA 模块后直接由感兴趣区域对齐模块 RoIAlign 对真实标注框所在位置进行特征对齐,并通过 ResNet101 的 conv5_x 模块对该位置特征进行进一步提取,得到 res_5 支持集特征,在批量 (batch) 维度上对 res_5 支持集特征进行平均后得到支持集特征图原型;另一个分支经过 BA 模块对无关背景进行抑制后,再经过 SimAM 模块对支持集特征进行增强,进一步在 batch 维度上对该特征进行平均后经过空间池化得到支持集类原型。将该支持集类原型与查询集特征送入 Meta-RPN 模块后生成与支持集类别一致的候选框,该候选框所在位置的查询集特征经 RoIAlign 模块特征对齐和 conv5_x 模块特征提取后,与支持集特征图原型共同输入到 Meta Classifier 模块进行查询匹配,输出类别置信度和预测框位置。

1.1 基于支持集背景抑制的小样本目标检测

为更好地利用支持集特征,首先从抑制与待查询目标无关的背景信息入手,引入 BA 模块以

Meta-RPN,可较好地计算类原型与锚框特征之间的相似性;细粒度原型匹配网络 Meta-Classifier,通过粗粒度到细粒度优化的过程产生特定新类的候选框。Meta Faster R-CNN 算法在支持集与查询集之间的互相关性方面进行了改进,但忽略了支持集的内在联系,本文通过引入 BA 模块与 SimAM 注意力模块,增强支持集的内在联系,更好地利用支持集特征,提高检测精度。改进后的 Meta Faster R-CNN 算法框架如图 1 所示。

较好地地区分前景与背景,该模块在通道维度上对支持集特征进行特征增强与无用特征的抑制。BA 背景抑制模块结构如图 2 所示。图中: Y 为 BA 模块支持集特征图; W_e 为可学习参数矩阵,用于学习不同通道的重要性, $W_e \in \mathbf{R}^{C_s \times 1}$; C_s 为支持集特征图的通道数; H_s 和 W_s 分别为 BA 模块支持集特征图的高度和宽度; Leaky ReLU 为非线性激活函数; Z 为支持集特征图 Y 经 BA 模块处理后得到的特征图。

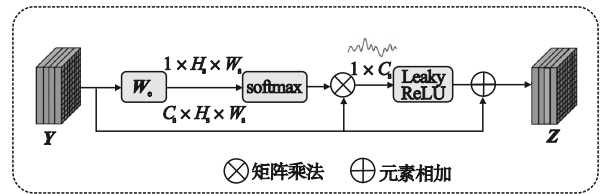


图2 BA 模块结构

Fig. 2 BA module structure

相关计算公式为

$$F_{BA}(y_i) = \sigma(W_e y_i) = \frac{\exp(W_e y_i)}{\sum_{j \in \Omega} \exp(W_e y_j)} \quad (1)$$

$$G = \sum_{i \in \Omega} F_{BA}(y_i) \cdot y_i \quad (2)$$

$$Z = Y + \alpha \cdot \text{LeakyReLU}(G) \quad (3)$$

式中: y_i 为 Y 中第 i 个像素的特征向量, $y_i \in$

$\mathbf{R}^{1 \times C_s}$; Ω 为所有像素索引的集合; σ 表示 softmax 函数; $F_{BA}(y_i)$ 表示特征向量 y_i 的权重; G 表示 Y 中最有价值的特征, $G \in \mathbf{R}^{1 \times C_s}$; α 为超参数。支持集特征图经过 BA 模块后, 无关背景信息被抑制, 前景信息更加突出。

1.2 基于支持集特征增强的小样本目标检测

为更好地利用支持集特征, 另一个角度是加强网络对待检测目标的注意。现有注意力模块主

要分为通道注意力模块和空间注意力模块, 但两者各有局限, 未综合考虑通道与空间注意力的结合, 卷积块注意力模块 (CBAM) 将空间注意力与通道注意力进行并行或串行组合, 但忽略了两种注意力之间的协同。因此, 本文采用统一权值的 SimAM 注意力模块, 其通过评估每个特征点的重要性, 更好地实现注意力机制。SimAM 注意力模块结构如图 3 所示。

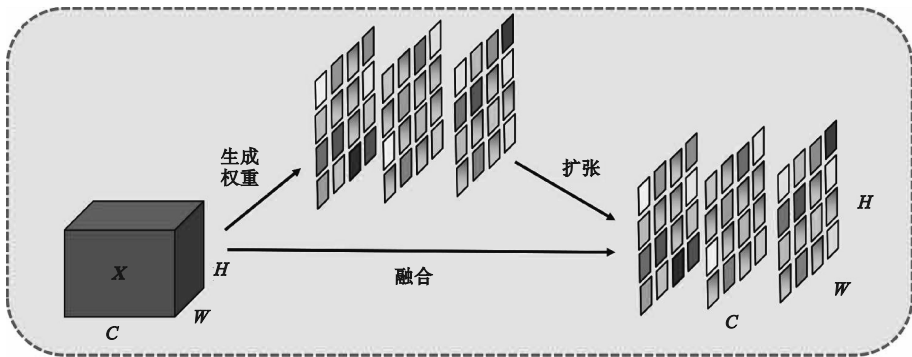


图3 SimAM 结构

Fig.3 SimAM structure

本文在 Meta Faster R-CNN 网络中引入 SimAM 注意力模块, 无需增加额外参数, 同时从空间和通道维度对支持集特征进行增强, 以提升小样本目标检测精度。SimAM 注意力模块通过能量函数评估每个神经元的重要性, 表达式为

$$\tilde{X} = \text{sigmoid}(1/E) \odot X \quad (4)$$

式中: E 为神经元能量函数, E 越低, 表示该神经元与周围神经元区别越大, 重要性越高, $1/E$ 表示每个像素级特征的重要性; X 为 SimAM 模块支持集特征图; \tilde{X} 为 X 特征增强后的特征图。能量函数 E 的计算式为

$$E = e_t(w_t, b_t, y, x_k) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{k=1}^{M-1} (y_o - \hat{x}_k)^2 \quad (5)$$

式中: \hat{t} 和 \hat{x}_k 均为像素级的特征值, $\hat{t} = w_t t + b_t$, 表示神经元 t 线性变换后的值, $\hat{x}_k = w_t x_k + b_t$, 表示 t 之外的神经元 x_k 线性变换后的值, w_t 和 b_t 分别为线性变换的权重和偏置; $M = H \times W$, 表示该通道上神经元的数量, H 和 W 分别表示 SimAM 模块特征图的高度和宽度; y 表示真实值, y_t 与 y_o 分别为神经元 t 和 o 的特征值。当 $\hat{t} = y_t$, $\hat{x}_k = y_o$ 时, E 取最小值, 对 y_t 和 y_o 采用二进制标签, $y_t = 1$, $y_o = -1$ 。将 $\hat{t} = w_t t + b_t$ 和 $\hat{x}_k = w_t x_k + b_t$ 代入式 (5) 并添加正则项得

$$e_t(w_t, b_t, y, x_k) = \frac{1}{M-1} \sum_{k=1}^{M-1} \left[-1 - (w_t x_k + b_t) \right]^2 + \left[1 - (w_t t + b_t) \right]^2 + \lambda w_t^2 \quad (6)$$

式中 λ 为正则项参数。式 (6) 的解析解为

$$w_t = -\frac{2(t - u_t)}{(t - u_t)^2 + 2\sigma_t^2 + 2\lambda} \quad (7)$$

$$b_t = -\frac{1}{2}(t + u_t)w_t \quad (8)$$

式中 u_t 和 σ_t^2 分别表示该通道中除 t 之外所有神经元的平均值和方差。

将式 (7) 和式 (8) 代入式 (6) 得到最小能量 E^* , 即

$$E^* = e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{u})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (9)$$

其中

$$\hat{u} = \frac{1}{M} \sum_{k=1}^M x_k \quad (10)$$

$$\hat{\sigma}^2 = \frac{1}{M} \sum_{k=1}^M (x_k - \hat{u})^2 \quad (11)$$

2 模型训练

2.1 数据集

定义类别集合 $C = C_{\text{base}} \cup C_{\text{novel}}$, 且 $C_{\text{base}} \cap C_{\text{novel}} = \emptyset$, 其中 C_{base} 为基类集合, 基类集合中的每

一个类别均包含大量的标注样本, C_{novel} 为新类集合, 新类集合中的每一个类别下均只有少量标注样本。定义样本集合 $D = D_{\text{base}} \cup D_{\text{novel}}$, 其中 D_{base} 为基类样本集合, D_{novel} 为新类样本集合。

元学习方法将整个训练任务分解为大量的元任务, 每个元任务中标注数据集 $T_m = S_m \cup Q_m = \{s_1^m, s_2^m, \dots, s_N^m\} \cup \{q_m\}$, 其中 S_m 表示第 m 个任务中的支持集, S_m 包含 N 个类别, s_n^m 表示第 m 个任务中支持集的第 n 个 ($n = 1, 2, \dots, N$) 类别标注样本的集合, 每个集合 s_n^m 包含 K 个标注样本, 记为 K -shot, Q_m 表示第 m 个任务中的查询集, q_m 表示第 m 个任务中的查询样本, 每个查询集中通常为单个样本, q_m 的类别为 S_m 中包含的类别, 但 $q_m \notin S_m$ 。

本文按照上述定义制作相关的小样本目标检测数据集, 以 TFA 中的划分标准对 PASCAL VOC 数据集进行以下划分:

1) 将数据集 PASCAL VOC2007 和 PASCAL VOC2012 的 train 和 val 数据合并用于模型训练, PASCAL VOC2007 的 test 数据用于模型测试;

2) 将 PASCAL VOC 中鸟、公交车、牛、摩托车和沙发共 5 个类别划分为新类数据, 记为 Novel Set, 其余 15 个类别划分为基类数据;

3) 将 PASCAL VOC 数据集中 K 的取值设置为 1、2、3、5、10, 将对应 K 值下的小样本目标检测任务记为 1-shot、2-shot、3-shot、5-shot、10-shot。

2.2 模型的训练与测试过程

模型的训练过程包含两个阶段: 元训练阶段, 每个元任务的训练数据集 T_m 在基类 D_{base} 中选取; 元微调阶段, 在 D_{base} 与 D_{novel} 中取相同数量 (K 个) 的样本形成集合 D' , 每个元任务的训练数据集 T_m 在 D' 中选取。

模型的测试过程包括元测试阶段和元微调后的测试阶段。元测试阶段在不对新类进行微调的情况下, 直接在元训练阶段后对基类和新类的小样本数据进行测试; 元微调后的测试阶段则是对新类数据进行微调后再对基类和新类的小样本数据进行测试。两个阶段的测试数据集 T_i 均在 D_{base} 与 D_{novel} 构成的集合 D'' 中选取, 其中 $D' \cap D'' = \emptyset$, $T_m \cap T_i = \emptyset$, 即训练集合与测试集合之间无重叠样本数据。

2.3 训练环境及相关参数设置

硬件设备: 32 GB \times 2 NVIDIA Tesla V100S 显卡, 16 核 1.90 GHz 的 Intel (R) Xeon (R) Bronze 3206R CPU。深度学习框架为 Detectron2v0.4、

PyTorch1.10.0, Python 版本为 3.7.16, CUDA 版本为 11.1, 操作系统为 Ubuntu 系统。

本文在 Pascal VOC 数据集上使用随机梯度下降法进行端到端训练。在元训练阶段, 采用双卡并行训练, 单卡的批次大小设置为 3, 动量设置为 0.9, 权重衰减系数设置为 0.0001, 首次训练共设置 20 000 次迭代, 其中前 15 000 次迭代学习率设置为 0.002, 后 5 000 次迭代学习率设置为 0.0002; 在引入不同模块进行训练时, 共设置 10 000 次迭代, 其中前 7 500 次迭代学习率设置为 0.001, 后 2 500 次迭代学习率设置为 0.0001。

查询集图像长边上限设置为 1 333 像素, 短边下限设置为 600 像素; 将支持集图像裁剪并对边缘进行零填充, 调整其大小为 320×320 像素。

2.4 损失函数及评价指标

总损失 L 计算式为

$$L = L_{\text{meta_rpn}} + L_{\text{meta_classifier}} \quad (12)$$

$$L_{\text{meta_rpn}} = L_{\text{cls}}^B + L_{\text{loc}} \quad (13)$$

$$L_{\text{meta_classifier}} = L'_{\text{cls}}^B + L'_{\text{loc}} \quad (14)$$

式中: $L_{\text{meta_rpn}}$ 和 $L_{\text{meta_classifier}}$ 分别表示 Meta-RPN 和 Meta-Classifier 两部分的损失; L_{cls}^B 和 L'_{cls}^B 分别为 Meta-RPN 和 Meta-Classifier 两部分的二进制交叉熵损失, 表示支持集与查询集之间的类别匹配关系; L_{loc} 为 Meta-RPN 部分候选框与真实框之间的 L_1 回归损失, L'_{loc} 为 Meta-Classifier 部分预测框与真实框之间的 L_1 回归损失, 两者均表示支持集与查询集之间的位置匹配关系。

为验证改进算法的有效性, 本文采用平均精度均值 (mAP) 对算法进行评价。将交并比 (IoU) 阈值设置为 0.5, 计算每个类别的平均精度, 再取平均得到评价指标 mAP@0.5。该指标能够全面反映算法对不同类别目标检测的性能表现, 可提供客观的性能衡量标准。

3 实验结果与分析

3.1 对比实验

将现有 SOTA 小样本目标检测算法与本文改进 Meta Faster R-CNN 算法在 PASCAL VOC Novel Set 数据集上进行对比实验, 结果如表 1 所示。

由表 1 可知, 本文改进 Meta Faster R-CNN 算法较原算法在元测试阶段 1-shot、2-shot、3-shot、5-shot、10-shot 下的 mAP@0.5 值分别提升了 0.066%、12.038%、12.289%、10.073%、9.539%, 在元微调后的测试阶段 1-shot、2-shot、

3-shot、10-shot 下的 mAP@0.5 值分别提升了 3.991%、3.930%、1.665%、1.301%，仅 5-shot 下的 mAP@0.5 值略有下降。本文提出的改进算法

较原算法的检测效果更好，相比其他算法也有较优的表现。

表 1 小样本目标检测算法对比实验结果

Table 1 Comparisons of experimental results based on few-shot object detection algorithm

算法名称	mAP@0.5/%				
	1-shot	2-shot	3-shot	5-shot	10-shot
	元测试阶段				
TFA ^[5]	32.4	22.1	23.1	31.7	35.7
Meta Faster R-CNN ^[9]	40.2	30.5	33.3	42.3	46.9
本文改进 Meta Faster R-CNN	40.266	42.538	45.589	52.373	56.439
	元微调后的测试阶段				
TFA w/cos ^[5]	39.8	36.1	44.7	55.7	56.0
TFA w/fc ^[5]	36.8	29.1	43.6	55.7	57.0
Meta-DETR ^[8]	35.1	49.0	53.2	57.4	62.0
UNP ^[10]	43.7	58.3	59.8	63.7	64.2
FsDetView w/cos ^[12]	26.9	35.7	42.3	48.9	57.8
FsDetView w/fc ^[12]	24.2	35.3	42.2	49.1	57.4
FSSP ^[13]	41.6	—	49.1	54.2	56.5
DMNet ^[14]	34.7	50.7	54.0	58.8	62.5
TeSNet ^[15]	35.6	42.7	44.1	54.4	57.8
BC-YOLO ^[16]	—	—	—	50.4	57.6
Meta Faster R-CNN ^[9]	43.0	54.5	60.6	66.1	65.4
本文改进 Meta Faster R-CNN	46.991	58.430	62.265	65.541	66.701

3.2 消融实验

本文改进 Meta Faster R-CNN 算法在 PAS-

CAL VOC Novel Set 数据集上的消融实验结果如表 2 所示。

表 2 本文改进算法的消融实验结果

Table 2 Ablation experiment results of improved algorithm in this paper

实验组号	Meta Faster R-CNN 算法	BA 背景抑制模块	SimAM 注意力模块	mAP@0.5/%				
				1-shot	2-shot	3-shot	5-shot	10-shot
	元测试阶段							
1	✓			40.2	30.5	33.3	42.3	46.9
2	✓	✓		38.039	41.555	43.310	52.341	54.964
3	✓		✓	37.692	41.393	42.059	49.830	54.512
4	✓	✓	✓	40.266	42.538	45.589	52.373	56.439
	元微调后的测试阶段							
5	✓			43.0	54.5	60.6	66.1	65.4
6	✓	✓		44.963	57.384	58.195	63.710	66.658
7	✓		✓	42.498	55.783	57.789	63.570	65.958
8	✓	✓	✓	46.991	58.430	62.265	65.541	66.701

由表 2 中元测试阶段的实验结果可知:本文改进 Meta Faster R-CNN 算法在 1-shot、2-shot、3-shot、5-shot、10-shot 下的 mAP@0.5 值均高于原算法,表明改进算法模型对于新任务的适应能力更强,泛化能力更佳;实验 2 与实验 3 在 1-shot 下的 mAP@0.5 值略低于原算法,其他情况下的 mAP@0.5 值均高于原算法,这是由于 1-shot 目标实例数量过少,改进过程中损失了部分目标的特征信息。将 BA 背景抑制模块与 SimAM 注意力

模块相结合,特征信息损失情况有所改善,故本文改进算法在 1-shot 下的 mAP@0.5 值高于原算法。

由表 2 中元微调后测试阶段的实验结果可知,本文改进 Meta Faster R-CNN 算法在所有情况下的 mAP@0.5 值均高于除原算法外的其他算法,在 1-shot、2-shot、3-shot、10-shot 下的 mAP@0.5 值均高于原算法,仅在 5-shot 下略低于原算法。

对比实验 2、实验 3 和实验 6、实验 7 可以发

现,BA背景抑制模块在mAP@0.5值的提升效果上发挥了更重要的作用,该模块与SimAM注意力模块相结合,使本文算法较原算法在检测精度上取得了较大的提升。单独使用BA背景抑制模块或单独使用SimAM注意力模块时,在元微调后的测试阶段中部分实验的性能表现有所下降。仅采用BA背景抑制模块抑制支持集背景信息而不增强目标特征,易导致目标特征信息不够突出;仅采用SimAM注意力模块则容易将前景和背景信息混淆,难以准确提取和增强目标特征;将两者的优势结合后,不仅可有效抑制支持集背景的干扰,还可显著增强待检测目标的特征表现,故检测精度有所提升。

4 结论

提出了一种改进的Meta Faster R-CNN小样本目标检测算法,改善了原算法在检测时仅关注支持集与查询集之间的联系而忽略了支持集内在联系的问题,通过引入BA背景抑制模块削弱了与待查询目标无关的背景信息,更好地区分了待检测图像的前景与背景信息,同时利用SimAM模块对支持集特征图进行特征增强,更好地加强了支持集的内在联系,合理利用支持集特征,提高了检测精度。实验结果表明,本文改进后的算法比原算法的检测精度有了显著提升,增强支持集特征能够有效提升小样本目标检测精度。

参考文献(References):

- [1] 刘春磊,陈天恩,王聪,等.小样本目标检测研究综述[J].计算机科学与探索,2023,17(1):53-73.
LIU C L, CHEN T E, WANG C, et al. Survey of few-shot object detection[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(1): 53-73. (in Chinese)
- [2] 史燕燕,史殿习,乔子腾,等.小样本目标检测研究综述[J].计算机学报,2023,46(8):1753-1780.
SHI Y Y, SHI D X, QIAO Z T, et al. A survey on recent advances in few-shot object detection[J]. Chinese Journal of Computers, 2023, 46(8): 1753-1780. (in Chinese)
- [3] 宋云凯,吴原顶,叶蕴瑶,等.基于特征加权与融合的小样本遥感目标检测[J].软件导刊,2024,23(4):150-156.
SONG Y K, WU Y X, YE Y Y, et al. Few-shot object detection on remote sensing images based on feature weighting and fusion[J]. Software Guide, 2024, 23(4): 150-156. (in Chinese)
- [4] CHEN H, WANG Y L, WANG G Y, et al. LSTD: a low-shot transfer detector for object detection[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2018, 32(1): 2836-2843.
- [5] WANG X, HUANG T E, DARRELL T, et al. Frustratingly simple few-shot object detection[C]//Proceedings of the 37th International Conference on Machine Learning. Vienna, Austria; PMLR, 2020: 9919-9928.
- [6] 张振伟,郝建国,黄健,等.小样本图像目标检测研究综述[J].计算机工程与应用,2022,58(5):1-11.
ZHANG Z W, HAO J G, HUANG J, et al. Review of few-shot object detection[J]. Computer Engineering and Applications, 2022, 58(5): 1-11. (in Chinese)
- [7] HUANG G, LARADJI I, VAZQUEZ D, et al. A survey of self-supervised and few-shot object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(4): 4071-4089.
- [8] ZHANG G J, LUO Z P, CUI K W, et al. Meta-DETR: image-level few-shot detection with inter-class correlation exploitation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(11): 12832-12843.
- [9] HAN G X, HUANG S Y, MA J W, et al. Meta Faster R-CNN: towards accurate few-shot object detection with attentive feature alignment[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(1): 780-789.
- [10] YAN B W, LANG C B, CHENG G, et al. Understanding negative proposals in generic few-shot object detection[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024, 34(7): 5818-5829.
- [11] YANG Z, WANG Y L, CHEN X Y, et al. Context-Transformer: tackling object confusion for few-shot detection[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12653-12660.
- [12] XIAO Y, LEPETIT V, MARLET R. Few-shot object detection and viewpoint estimation for objects in the wild[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(3): 3090-3106.
- [13] XU H H, WANG X Q, SHAO F M, et al. Few-shot object detection via sample processing[J]. IEEE Access, 2021, 9: 29207-29221.
- [14] LU Y, CHEN X Y, WU Z X, et al. Decoupled metric network for single-stage few-shot object detection[J]. IEEE Transactions on Cybernetics, 2023, 53(1): 514-525.
- [15] ZHAO X W, LIU X L, MA Y Q, et al. Temporal speciation network for few-shot object detection[J]. IEEE Transactions on Multimedia, 2023, 25: 8267-8278.
- [16] XIA R Y, LI G Q, HUANG Z W, et al. Bi-path combination YOLO for real-time few-shot object detection[J]. Pattern Recognition Letters, 2023, 165: 91-97.
- [17] CHEN T I, LIU Y C, SU H T, et al. Dual-awareness attention for few-shot object detection[J]. IEEE Transactions on Multimedia, 2023, 25: 291-301.

(责任编辑:宋颖韬)