



计算机科学

COMPUTER SCIENCE

基于特征融合的小样本目标检测

华杰, 刘学亮, 赵焯

引用本文

华杰, 刘学亮, 赵焯. 基于特征融合的小样本目标检测[J]. 计算机科学, 2023, 50(2): 209-213.

HUA Jie, LIU Xueliang, ZHAO Ye. [Few-shot Object Detection Based on Feature Fusion](#)[J]. Computer Science, 2023, 50(2): 209-213.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[融合循环划分的张量指令生成优化](#)

Tensor Instruction Generation Optimization Fusing with Loop Partitioning

计算机科学, 2023, 50(2): 374-383. <https://doi.org/10.11896/jsjcx.220300147>

[受人脑中记忆机制启发的增量目标检测方法](#)

Incremental Object Detection Inspired by Memory Mechanisms in Brain

计算机科学, 2023, 50(2): 267-274. <https://doi.org/10.11896/jsjcx.220900212>

[基于改进区域候选网络的场景文本检测](#)

Scene Text Detection with Improved Region Proposal Network

计算机科学, 2023, 50(2): 201-208. <https://doi.org/10.11896/jsjcx.211000191>

[基于深度学习的刚体位姿估计方法综述](#)

Survey of Rigid Object Pose Estimation Algorithms Based on Deep Learning

计算机科学, 2023, 50(2): 178-189. <https://doi.org/10.11896/jsjcx.211200164>

[基于数据增强的自监督飞行航迹预测](#)

Self-supervised Flight Trajectory Prediction Based on Data Augmentation

计算机科学, 2023, 50(2): 130-137. <https://doi.org/10.11896/jsjcx.211200016>

基于特征融合的小样本目标检测

华杰 刘学亮 赵焯

合肥工业大学计算机与信息学院 合肥 230601

(jihua@mail.hfut.edu.cn)

摘要 小样本目标检测旨在通过少量的样本学习来训练目标检测模型,现有的小样本目标检测方法大多基于经典的目标检测算法。在二阶段的检测方法中,由于新类别样本数量少,产生了许多无关的边界框,导致候选区域的准确率较低。为了解决这个问题,提出了一种基于特征融合的小样本目标检测算法 FF-FSOD。该方法采用特征融合的方法进行数据增强,对新类别样本进行补充,扩大样本的覆盖范围,同时引入 FPN 网络进行多尺度特征提取,再对 RPN 网络进行改进,引入支持集图像分支,计算支持集图像特征与查询集图像特征的深度互相关性,得到注意力特征图,进而获得更精确的候选框。所提模型的有效性在 MS COCO 和 FSOD 数据集上得到了验证,实验结果表明,该方法获得了更精准的候选框,进而提升了检测精度。

关键词: 小样本学习;目标检测;深度学习;特征融合;特征金字塔

中图法分类号 TP391

Few-shot Object Detection Based on Feature Fusion

HUA Jie, LIU Xueliang and ZHAO Ye

School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230601, China

Abstract Few-shot object detection aims to train target detection model through a small amount of sample learning. At present, most of the existing few-shot object detection methods are based on classical target detection algorithms. In the two-stage detection method, due to the small number of new class samples, many irrelevant border boxes are generated, resulting in low accuracy of candidate regions. To solve this problem, this paper proposes a few-shot object detection algorithm FF-FSOD based on feature fusion. It uses the feature fusion method to enhance the data, supplements the new category samples, increases the coverage range of the sample, and introduces the FPN network to extract multi-scale feature. Then, the RPN network is improved, and the support set image branch is introduced. The depth correlation between the support set image feature and the query set image feature is calculated, and the attention feature map is obtained, and the more accurate candidate box is obtained. The effectiveness of the proposed model is verified on MS COCO and FSOD datasets. Experimental results show that the proposed method obtains more accurate candidate boxes and improves the detection accuracy.

Keywords Few-shot learning, Object detection, Deep learning, Feature fusion, Feature pyramid

1 引言

近年来,随着卷积神经网络的发展,深度学习在图像领域取得了长足的进步。然而,要训练算法模型,首先需要标注大量样本,尽管目前最新的研究能够利用少量样本完成算法的训练,但这些算法主要是用来解决图像分类^[1]问题。然而,在更具挑战和应用价值的目标检测问题上,针对小样本学习的研究还停留在起步阶段^[2-3]。

小样本图像目标检测集小样本学习方法与目标检测算法于一体,其目的和小样本分类一样是利用少量的标记样本完成目标检测模型的训练,使得目标检测模型能够达到较好的泛化性能。小样本目标检测能够克服现有的目标检测方法依赖大量的标注数据的问题,近年来逐渐引起了学术界的广泛

关注。现有的小样本目标检测方法大多基于经典的目标检测算法,主要分为一阶段和二阶段两种检测方法。其中,一阶段检测方法的优势在于检测速度快,而二阶段检测方法着重于追求精度。小样本目标检测更多的是沿用二阶段目标检测方法,利用基类与新类样本集合对检测模型进行优化,实现对新类中图像目标的检测。比如文献[6-8]以 Faster R-CNN 为基础模型,有部分工作以 YOLO 和 SSD 等经典的一阶段检测方法为基础,文献[9-10]分别以 YOLOv2 和 SSD 为主干网络。本文沿用了二阶段检测方法,并在此基础上进行了改进。

小样本图像目标检测不仅要关注分类任务的高层语义信息,而且要提取低层图像语义信息,并给出目标在图像中的确切位置。目前二阶段检测方法在小样本目标检测上存在以下几个方面的问题:

到稿日期:2022-05-17 返修日期:2022-11-08

基金项目:国家重点研发计划(2018AAA0102002);国家自然科学基金(61976076,61632007)

This work was supported by the National Key R & D Program of China(2018AAA0102002) and National Natural Science Foundation of China(61976076,61632007).

通信作者:刘学亮(Liuxueliang1982@gmail.com)

(1) Faster R-CNN 依靠区域建议网络 (Region Proposal Network, RPN) 来区分背景与前景, 然后在候选区域上进行分类和回归, 进而得到最终的检测结果。然而, 在小样本情况下, 新的待检测的目标数量较少, 使得候选区域会出现较多的与目标类别无关的情况, 检测准确率也随之下降。

(2) 与基类样本相比, 新类别的目标样本数量较少, 有些小目标在原图像的占比过小, 特征比较模糊, 尺度空间存在差异。同时, 支持集图像和查询集图像之间的差异性也给小样本图像目标检测带来一定的挑战。

针对上述问题, 本文设计了一种基于特征融合的小样本目标检测算法 (Few-Shot Object Detection Based on Feature Fusion, FF-FSOD)。首先引入支持集图像分支, 特征融合模块对其进行特征融合, 再输入特征提取网络中获得支持集图像特征, 通过相同的网络得到查询集图像特征; 接着通过 FPN 网络分别获得查询集图像特征和支持集图像特征; 然后将特征图输入改进的 RPN 网络中, 得到候选框, 进而生成 RoI (Region of Interest) 特征图和 RoI 特征向量; 最后进行分类和定位, 完成最终的检测。针对新类别样本数量少的问题, 采用特征融合的方法进行数据增强, 对新类别样本进行补充, 增大样本的覆盖范围。针对样本类别的尺度存在差异的问题, 引入了 FPN (Feature Pyramid Network) 网络进行多尺度特征提取。针对小样本条件下较难获得高精度的候选框的问题, 对 RPN 网络进行改进, 引入支持集图像分支, 计算支持集图像特征与查询集图像特征的深度互相关性, 得到注意力特征图, 进而获得更精确的候选框。

2 相关工作

相比小样本图像分类任务, 小样本目标检测任务更具有挑战性。小样本目标检测根据其思想和模型结构可以分为: 基于数据增强的方法、基于度量学习的方法、基于模型结构的方法、基于微调的方法和基于元学习的方法。

在训练样本不够充足的情况下, 数据增强常被用来对样本进行扩充, 进而构建数量相对较多的样本, 在原有的图像检测算法上进行训练和测试。Wu 等^[4]通过构建多个尺度的正样本, 建立特征金字塔网络来增强数据多样性。Zhang 等^[11]利用虚拟样本增加样本的多样性, 在样本数量极少情况下使性能得到一定程度的提升。文献^[12]通过对目标样本进行复制实现数据的扩充, 语义嵌入方法^[13-14]和语义增强^[15]的方法也被应用在小样本目标检测中。

基于度量学习的方法是对支持图像的特征和目标图像的特征进行相同的转换, 映射到同一嵌入空间内, 根据距离或相似度来对目标区域进行进一步的分类, 从而实现小样本目标检测。文献^[16]提出了一种计算距离的度量学习方法; Zhang 等^[17]将对网络引入度量学习中, 以解决小样本图像目标检测问题; Xu 等^[18]在分类模型中增加了语义相似性度量, 提高了模型的泛化能力。

传统的监督学习在迭代训练过程中是以样本为基本单元, 而基于元学习的方法是在训练过程中将任务分成多个小任务, 选择一组能够在不同任务中都达到良好性能的参数, 一般进行少量的迭代即可在新任务中取得较好的结果。Kang 等^[10]将元特征学习和 YOLOv2 相结合, 使得检测器能够快速适应新类别。Yan 等^[5]在 Mask R-CNN 的基础上

进行研究, 提出了 Meta R-CNN。该网络将支持集图像得到的类别特征与兴趣区域特征进行融合, 得到新的预测特征, 再进行后续的分类和回归任务。Xiao 等^[19]在此基础上改进了融合网络, 取得了更好的检测性能。而文献^[8]提出的基于微调的方法冻结了特征提取器以及 RPN 的参数, 对检测器的头部的回归和分类网络进行随机初始化, 其结果超过了一些基于元学习的方法。Sun 等^[20]则仅对回归和分类网络进行微调, 以此来提高检测准确率。

一些研究者尝试构建新的目标检测模型结构来解决小样本目标检测问题, Fan 等^[7]把支持集的样本信息加入 RPN 网络中, 指导其生成更加相关的候选框, 从而提升了检测性能。Dong 等^[21]利用模型互通的方法, 将多模型学习与自主学习相结合, 先在简单样本上进行训练, 再由现有模型生成复杂的样本, 如此循环往复从而提升模型性能。Chen 等^[22]引入了新的注意力检测器, 结合了自上而下和自下而上的多种注意力, 利用混合训练策略, 获得了较好的检测性能。

3 提出模型

本节首先对小样本目标检测任务进行定义, 接着详细描述本文的基于特征融合的小样本目标检测网络结构, 再对其中的特征融合模块、多尺度 FPN 模块和改进的 RPN 模块进行阐述。

3.1 问题定义

首先给定基类集合 D_{base} 和新类集合 D_{new} , 其中基类集合 D_{base} 中的每个类别都有充足的标记样本, 新类集合 D_{new} 中的每个类别只有少量的标记样本 (通常数量少于 10 个), 而 D_{base} 和 D_{new} 中的类别没有交集。小样本目标检测的目的是通过基类集合 D_{base} 和新类集合 D_{new} 中的样本对检测模型进行迭代优化, 最终实现对新类中的图像目标的检测。在最新的研究工作中, 该问题一般在 N -way K -shot 的元学习框架下进行求解, 即在基类集合 D_{base} 中进行训练, 然后从新类 D_{new} 中随机抽取 N 个类, 再从每个类中抽取 K 个样本形成支持集 $S = \{x_i\}_{i=1}^m$ ($m = K \times N$), 新类剩余样本用作查询集 Q , 这样的问题被称为 N -way K -shot 目标检测。

3.2 模型概述

为了更好地解决小样本目标检测问题, 本文设计了一种基于特征融合的小样本目标检测网络 (FF-FSOD), 具体的算法结构如图 1 所示。首先, 引入支持集图像这一分支, 特征融合模块先对其进行特征融合, 特征提取网络负责获得支持集图像特征和查询集图像特征。由于小目标在原图像中的占比过小, 特征比较模糊, 需要 FPN 网络提取多尺度特征图, 再将此特征图输入改进的 RPN 网络中得到候选框, 进而生成 RoI 特征图和 RoI 特征向量, 再进行分类和定位, 完成最终的检测。类似于 Faster RCNN^[23], FF-FSOD 损失函数由 RPN 损失函数和 Fast R-CNN^[24] 损失函数组成, 这两个函数进行迭代训练, 完成模型的优化过程。RPN 损失函数包括回归损失和二分类交叉熵损失, 如式 (1) 所示:

$$L_{\text{RPN}} = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*) \quad (1)$$

其中, p_i 是边界框的预测分类概率; p_i^* 是真实标签值, 当 i 为正样本时, $p_i^* = 1$, 反之, $p_i^* = 0$; t_i 表示预测的边界框坐标, t_i^* 是真实边界框的坐标。

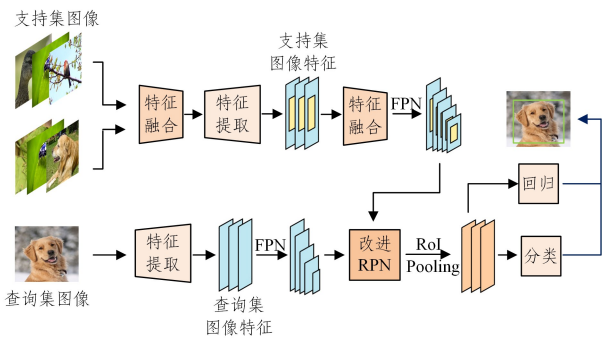


图 1 FF-FSOD 算法结构

Fig. 1 FF-FSOD algorithm structure

Fast R-CNN 损失函数包括分类损失和边界框回归损失,分类损失 L_{cls} 由 softmax 交叉熵损失组成,边界框回归损失 L_{loc} 是 smooth L1 损失,如式(2)所示:

$$L(p, u, t, v) = L_{cls}(p, u) + \lambda L_{loc}(t, v) \quad (2)$$

其中, p 是 softmax 交叉熵的概率分布, u 为目标的真实类别标签, t 为边界框回归器预测的类别回归参数, v 为真实目标的边界框回归参数。

针对小样本,本文的 FF-FSOD 模型使用了特征融合方法进行数据增强,同时引入 FPN 模块,从而生成多尺度的特征金字塔,增加了支持集图像支路,改进了 RPN 模块,从而生成更准确的候选框。

3.3 特征融合模块

特征融合模块本质上是一个混合块,如图 2 所示,它将两个输入组合成一个新的共享特征。本文对支持集图像进行特征融合,受到 MSDA^[25] 方法的启发,设计了特征融合变体模块,它并不局限于固定比例的融合,可表示为 $M(A_0, A_1) = 2[kA_0' + (1-k)A_1']$,其中 A_0 和 A_1 分别表示来自相同类别的不同样本, A_0' 和 A_1' 分别为对应的特征图, k 为可变融合参数,后续的相关实验也证明了 k 对检测性能的影响。对于小样本目标检测来说,数据增强是一种朴素并且直接的提升算法性能的方法。本文引入的特征融合模块分别在特征提取网络的前后进行特征融合,以获得更加丰富的图像信息以及特征信息,达到对样本数据进行增强、扩大样本覆盖范围的目的。特征融合算法的伪代码如算法 1 所示。

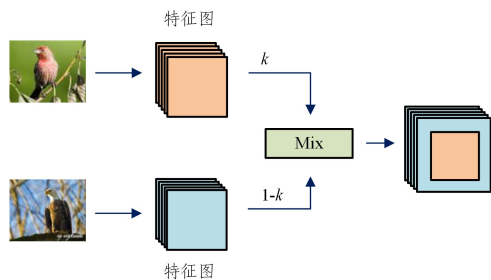


图 2 特征融合模块

Fig. 2 Feature fusion module

算法 1 特征融合算法

输入:待融合图像 A_0, A_1 , 可变融合参数 k

输出:融合图像 $M(A_0, A_1)$

1. 对 A_0 和 A_1 裁剪后进行零填充,再调整为 320×320 的图像;
2. 将 ResNet-50 作为骨干网络对输入图像进行特征提取,得到 A_0' 和 A_1' ;
3. 对 A_0' 和 A_1' 进行特征融合 $M(A_0, A_1) = 2[kA_0' + (1-k)A_1']$,得到融合图像。

3.4 多尺度 FPN 模块

与图像分类相比,目标检测任务相对复杂一些,它不仅需要完成分类而且还要实现目标定位。如果目标的尺度变化较大,尤其是图像中含有一些小目标,经过特征提取模块后得到的图像特征比较模糊,那么检测的效果也会受到影响。在样本数据比较充足的情况下,可以通过构建复杂的算法模型来达到对不同尺度目标进行检测的目的。但在小样本目标检测中,往往缺乏多种尺度的训练样本。因此,本文考虑到尺度变化的影响,将 FPN^[26] 模块引入查询集分支和支持集分支中,以此获得多尺度的特征信息。FPN 的结构如图 3 所示,主要包含自下而上和自上而下的路径以及横向连接。在自上而下的过程中进行上采样,获得多尺度的分辨率较高的特征图,在自下而上的过程中通过横向连接把相同尺寸的特征图相加,既利用了低层语义信息,又结合了高层语义信息,使得特征图的尺度空间更加丰富。

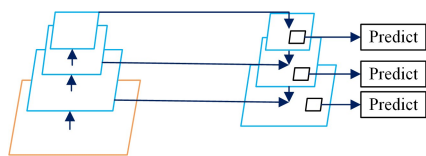


图 3 FPN 网络结构

Fig. 3 FPN network structure

3.5 改进的 RPN 模块

在原有的目标检测任务中,RPN 网络常被用来产生与目标相关的边界框,能否提供更精准的边界框对后续检测效果有直接的影响。在小样本目标检测任务中,RPN 网络会在大量的基类类别上进行训练,而需要检测的目标对象只包含少量的样本,因此 RPN 网络可能会产生许多与目标对象不相关的边界框,进而降低后续分类和检测的准确率。为了解决上述问题,本文对 RPN 网络进行改进,首先增加了支持集图像这一分支,支持集图像经过 FPN 网络后,经过平均池化和深度卷积形成一维的特征向量,再与查询集特征进行卷积操作,经过深度互相关的计算后,再输入 RPN 网络中,引导 RPN 网络产生与目标对象更加相关的候选框。

4 实验

本文在 FSOD 数据集和 MS COCO 数据集上进行实验,说明了有关实验设置的详细信息,对于融合比例的影响以及引入的 FPN 和改进的 RPN 模块,本文进行了相关的消融实验,并对结果进行了比较和分析。

4.1 数据集和设置

MSCOCO 数据集是微软构建的包含 123 287 张图片、约 886 000 个边界框,共有 80 个类别的数据集,包含了目标检测和语义分割等任务。该数据集的内容是从自然界和生活中的图片截取而来,包含了更多的图像类别,同时目标数量较多,而且图像中也包含了一些小目标。本文在 MS COCO 数据集上与其他小样本目标检测的经典算法进行了比较,包括 Meta YOLO^[10], Meta R-CNN^[5], MPSP^[4] 和 FSOD^[7] 等。FSOD 数据集是专门为小样本学习和评估模型而设计的,包含 1 000 个类别,800 个用于训练集,200 个用于测试集,大约 66 000 张图片和 182 000 个边界框。本文在 FSOD 数据集上进行了不同的 N 和 K 值的实验,来验证本文提出的模型的性能以及 k

值对检测结果的影响。

本文使用 Pytorch 框架, Pytorch 和 torchvision 的版本分别为 0.4.1 和 0.2.0, NVIDIA 的 CUDA 版本为 9.0。本文的模型使用随机梯度下降进行端到端训练, 在前 11 200 次迭代中, 学习率设置为 0.001, 后 800 次迭代中更新为 0.0001。与一般的目标检测一样, 本文固定了 Res1-3 层的权重, 对高层进行训练来避免过拟合。本文将查询集图像的长边上限设置为 1 000 像素, 短边调整为 600 像素, 对支持集图像裁剪后进行零填充, 最终调整为 320×320 像素的图像, 采用经典指标^[27] AP, AP_{50}, AP_{75} 进行评估。

4.2 MS COCO 数据集

在本小节中, 本文在 MS COCO 数据集上进行了对比实验, 采用标准 COCO 指标进行评价, 与经典的小样本目标检测算法的比较结果如表 1 所列。表现最好的方法用加粗表示, “—”表示参考文献中没有相关结果, YOLOv3 608×608 ^[28] 表示非小样本目标检测方法, 在一定程度上不具有公平性, 结果仅作对比参考。从表 1 可以看出, 本文提出的方法在 AP, AP_{50}, AP_{75} 的指标上, 比 MPSR 分别高出 1.6%, 3.1% 和 1.5%, 也略高于 FSOD, 分别高出 0.3%, 0.6% 和 0.6%。实验结果表明, 本文提出的改进算法在结合了不同尺度信息和引入了特征融合模块, 以及对 RPN 网络进行改进以后, 具有了更好的检测精度, 与其他经典的小样本目标检测方法相比, 本文提出的方法也具有出色的性能。

表 1 MS COCO 数据集实验结果

Table 1 Experimental results of MS COCO dataset

	(单位: %)		
	AP	AP ₅₀	AP ₇₅
YOLOv3 608×608 ^[28]	33.0	57.9	34.4
Meta YOLO ^[10]	5.6	12.3	4.6
Meta R-CNN ^[5]	8.7	19.1	6.6
TFA(w/fc) ^[28]	10.0	—	9.2
TFA(w/cos) ^[28]	10.0	—	9.3
FSOD ^[7]	11.1	20.4	10.6
FSCE ^[20]	11.1	—	9.8
UNIVERSAL PROTOTYPE ^[29]	11.0	—	10.7
FF-FSOD(Ours)	11.4	21.0	11.2

为了研究融合比例 k 值对检测精度的影响, 本文设置了不同的 k 值, 对比结果如表 2 所列。本文采用特征融合的方法进行数据增强, 对新类别样本进行补充, 扩大了样本的覆盖范围。从表 2 可以看出, 融合比例为 0.5:0.5 即 k 为 0.5 时效果最佳。 k 增大时, 检测效果降低, k 减小时亦是如此。这是因为融合比例过大或过小时, 会以某图像的信息作为主导, 没有充分利用图像的语义信息, 进而影响了检测效果。

表 2 融合比例 k 值对实验结果的影响

Table 2 Effect of fusion ratio k on experimental results

	(单位: %)		
	AP	AP ₅₀	AP ₇₅
FSOD ^[7]	11.10	20.40	10.60
FF-FSOD($k=0.7$)	11.21	20.62	10.82
FF-FSOD($k=0.6$)	11.30	20.71	10.91
FF-FSOD($k=0.5$)	11.45	21.06	11.24
FF-FSOD($k=0.4$)	11.31	20.69	10.90
FF-FSOD($k=0.3$)	11.19	20.58	10.77

4.3 FSOD 数据集

对于 N -way K -shot 检测问题, 本文在 FSOD 数据集上

设置了不同的 N 值和 K 值来验证本文模型的性能, 对比结果如表 3 所列, FSOD* 表示复现实验。由表 3 可以看出, 2-way 的训练结果比 5-way 略好, 在 2-way 5-shot 和 5-way 5-shot 中, 在融合比例 k 为 0.5 时, 效果最佳, 分别在 AP_{50} 和 AP_{75} 指标上达到了 65.96%, 42.47%, 66.26% 和 42.26%, 相比 FSOD 分别提升了 1.05%, 1.56%, 1.23% 和 0.59%。综合 4.2 节中表 2 的实验结果, 说明融合比例 k 值对检测精度的确有影响, 并且在 FSOD 和 MS COCO 数据集上的趋势相同, 说明融合比例 k 值的选择与支持数据集无关。表 3 的结果也说明了本文提出的改进方法在获得了更加丰富的图像信息以及特征信息之后, 对样本数据进行增强, 提升了小样本目标检测的精度。

表 3 FSOD 数据集实验结果

Table 3 Experimental results of FSOD dataset

	2-way 5-shot		5-way 5-shot	
	AP ₅₀	AP ₇₅	AP ₅₀	AP ₇₅
FSOD*	64.91	40.91	65.03	41.67
FF-FSOD($k=0.7$)	65.63	41.98	65.32	41.73
FF-FSOD($k=0.6$)	65.76	42.34	65.83	41.96
FF-FSOD($k=0.5$)	65.96	42.47	66.26	42.26
FF-FSOD($k=0.4$)	65.78	42.41	65.75	41.92
FF-FSOD($k=0.3$)	65.69	42.22	65.20	41.84

4.4 消融实验

与经典的小样本目标检测方法相比, 本文设计了 FF-FSOD 网络, 增加了特征融合模块、多尺度 FPN 模块, 并且改进了 RPN 模块。为了比较以上模块对实验结果的影响, 本文在 FSOD 数据集上对 2-way 5-shot 目标检测进行了相关的消融实验, 表 4 列出了实验结果。从表 4 可以看出, 相比 FPN 模块以及改进的 RPN 模块, 特征融合模块更加重要。在缺少特征融合模块而引入其他两个模块的情况下, 在 AP_{50} 和 AP_{75} 的指标上分别比引入 3 个模块的结果降低了 0.14% 和 0.1%; 只缺少 FPN 模块时, 其结果分别降低了 0.08% 和 0.09%; 没有对 RPN 模块进行改进而加入了其他两个模块时, 其结果分别降低了 0.11% 和 0.1%。总体而言, 本文引入的 3 个模块都对小样本目标检测的效果起到了积极有效的作用。

表 4 不同模块的消融实验结果

Table 4 Experimental results of ablation in different modules

	(单位: %)				
	特征融合 模块	FPN 模块	改进 RPN 模块	AP ₅₀	AP ₇₅
	✓			65.75	42.40
		✓		65.71	42.28
			✓	65.73	42.36
FF-FSOD ($k=0.5$)	✓	✓	✓	65.85	42.43
		✓	✓	65.82	42.37
	✓		✓	65.88	42.38
	✓	✓	✓	65.96	42.47

结束语 本文针对小样本目标检测问题, 提出了一种基于特征融合的小样本目标检测算法 FF-FSOD。该方法综合利用特征融合和多尺度 FPN 网络, 改进了 RPN 网络, 在 MS COCO 和 FSOD 数据集上的良好表现证明了本文提出的网络框架的有效性。本文方法对图像进行了数据增强, 提取了多尺度的图像特征, 获得了更精准的候选框, 进而提升了检测精度。小样本目标检测的一些问题还需进一步研究, 比如在

小样本情况下,在提取高语义信息的过程中如何获得更多的位置信息,增强目标关注信息,进一步抑制背景信息等。目前大部分工作都是采用二阶段的方式完成检测,对候选框进行改进的方法在保证有效的同时也增加了模型的复杂度,对设备内存也有较高的需求。为了进一步提高小样本目标检测的准确率和泛化能力,我们还将进行进一步研究。

参 考 文 献

[1] CAO J, QU X, LI X X. Few-shot image classification method based on sliding feature vectors[J]. Journal of Jilin University (Engineering and Technology Edition), 2021, 51(5): 1785-1791.

[2] CHEN H, WANG Y, WANG G, et al. LSTD: a low-shot transfer detector for object detection[C]// Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence. 2018:2836-2843.

[3] DONAHUE J, JIA Y, VINYALS O, et al. Decaf: A deep convolutional activation feature for generic visual recognition[C]// International Conference on Machine Learning. PMLR, 2014: 647-655.

[4] WU J, LIU S, HUANG D, et al. Multi-scale positive sample refinement for few-shot object detection[C]// European Conference on Computer Vision. Cham: Springer, 2020: 456-472.

[5] YAN X, CHEN Z, XU A, et al. Meta r-cnn: Towards general solver for instance-level low-shot learning[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019:9577-9586.

[6] XIAO Z, ZHONG P, QUAN Y, et al. Few-shot object detection with feature attention highlight module in remote sensing images[C]// 2020 International Conference on Image, Video Processing and Artificial Intelligence. SPIE, 2020: 217-223.

[7] FAN Q, ZHUO W, TANG C K, et al. Few-shot object detection with attention-RPN and multi-relation detector[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:4013-4022.

[8] WANG X, HUANG T, GONZALEZ J, et al. Frustratingly simple few-shot object detection[C]// International Conference on Machine Learning. PMLR, 2020:9919-9928.

[9] LI X, DENG J, FANG Y. Few-shot object detection on remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-14.

[10] KANG B, LIU Z, WANG X, et al. Few-shot object detection via feature reweighting[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019:8420-8429.

[11] ZHANG W, WANG Y X. Hallucination improves few-shot object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 13008-13017.

[12] RIOU K, ZHU J, LING S, et al. Few-Shot Object Detection in Real Life: Case Study on Auto-Harvest[C]// 2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSp). IEEE, 2020: 1-6.

[13] ZHU C, CHEN F, AHMED U, et al. Semantic relation reasoning for shot-stable few-shot object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:8782-8791.

[14] RAHMAN S, KHAN S, BARNES N, et al. Any-shot object detection[C]// Proceedings of the Asian Conference on Computer Vision. 2020.

[15] CHEN Z, FU Y, ZHANG Y, et al. Semantic feature augmentation in few-shot learning[J]. arXiv: 1804. 05298, 2018.

[16] KARLINSKY L, SHTOK J, HARARY S, et al. Repmet: Representative-based metric learning for classification and few-shot object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019:5197-5206.

[17] ZHANG T, ZHANG Y, SUN X, et al. Comparison network for one-shot conditional object detection[J]. arXiv: 1904. 02317, 2019.

[18] XU P B, SANG J T, LU D Y. Few shot image recognition based on class semantic similarity supervision[J]. Journal of Image and Graphics, 2021, 26(7): 1594-1603.

[19] XIAO Y, MARLET R. Few-shot object detection and viewpoint estimation for objects in the wild[C]// European Conference on Computer Vision. Cham: Springer, 2020:192-210.

[20] SUN B, LI B, CAI S, et al. Fscf: Few-shot object detection via contrastive proposal encoding[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021:7352-7362.

[21] DONG X, ZHENG L, MA F, et al. Few-example object detection with model communication[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 41(7): 1641-1654.

[22] CHEN X, JIANG M, ZHAO Q. Leveraging bottom-up and top-down attention for few-shot object detection[J]. arXiv: 2007. 12104, 2020.

[23] 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 and Machine Intelligence, 2016, 39(6): 1137-1149.

[24] GIRSHICK R. Fast r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision. 2015:1440-1448.

[25] HARRIS E, MARCU A, PAINTER M, et al. Fmix: Enhancing mixed sample data augmentation[J]. arXiv: 2002. 12047, 2020.

[26] 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. 2017: 2117-2125.

[27] TSUNGYL, PRIYA G, ROSS G, et al. Focal loss for dense object detection[C]// ICCV. 2017.

[28] REDMON J, FARHADI A. Yolov3: An incremental improvement[J]. arXiv: 1804. 02767, 2018.

[29] WU A, HAN Y, ZHU L, et al. Universal-prototype enhancing for few-shot object detection[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021:9567-9576.



HUA Jie, born in 1995, postgraduate. His main research interests include intelligent information retrieval and so on.



LIU Xueliang, born in 1981, Ph.D, professor, is a member of China Computer Federation. His main research interests include multimedia information retrieval and so on.