

## 基于特征对齐和特征融合的半监督目标检测算法

汤文兵, 李 菲

安徽理工大学 计算机科学与工程学院, 安徽 淮南 232001

**摘要:**目的 针对半监督目标检测导致数据特征表示不充分, 数据样本类不平衡等问题, 提出一种基于特征对齐和特征融合的半监督目标检测方法。方法 在常见的半监督目标检测框架中, 伪标签是完全根据分类分数生成的, 然而, 高置信度预测并不总是保证准确的 bbox 定位。为了解决定位不准确问题和特征表示不充分问题, 受 Consistent Teacher 中的 FAM-3D 算法启发, 考虑分类和定位的最优特征可能在不同尺度上, 引入 T-head 特征对齐头算法, 在 Unbiased Teacher V2 中成功地将分类和定位分支进行对齐, 并且引入 ASFF, 通过空间过滤冲突信息的方法来抑制不一致性, 从而提高了特征的尺度不变性, 实现特征在空间上的融合; 通过学习不同特征图之间的联系来解决特征金字塔内部的不一致性问题。结果 根据实验结果, 改进的算法在 COCO 数据集、VOC 数据集上都有一定的比例提升。结论 改进的算法可以有效减轻数据表示不充分和数据样本类不平衡问题, 同时也提高了算法的精度。

**关键词:** 目标检测; 半监督学习; 特征对齐; 特征金字塔; ASFF

中图分类号: TP391 文献标识码: A doi:10.16055/j.issn.1672-058X.2025.0001.005

### Semi-supervised Object Detection Algorithm Based on Feature Alignment and Feature Fusion

TANG Wenbing, LI Fei

School of Computer Science and Engineering, Anhui University of Science and Technology, Anhui Huainan 232001, China

**Abstract: Objective** In response to issues such as insufficient data feature representation and imbalanced sample classes in semi-supervised object detection, a semi-supervised object detection method based on feature alignment and feature fusion was proposed. **Methods** In common semi-supervised object detection frameworks, pseudo-labels are generated solely based on classification scores. However, high-confidence predictions do not always fully guarantee accurate bbox positioning. In order to solve problems of inaccurate positioning and insufficient feature representation, inspired by the FAM-3D algorithm in the Consistent Teacher, considering that the optimal features for classification and positioning may be at different scales, the T-head feature alignment head algorithm was introduced and the classification and positioning branches were successfully aligned in Unbiased Teacher V2. Additionally, ASFF was introduced to suppress inconsistency by spatially filtering conflict information, thereby improving the scale invariance of features and achieving spatial fusion of features. The internal inconsistencies within the feature pyramid were addressed by learning the connections between different feature maps. **Results** According to experimental results, the improved algorithm demonstrated certain performance improvements on the COCO dataset and VOC dataset. **Conclusion** The proposed algorithm effectively alleviates issues of insufficient data representation and imbalanced sample classes while also enhancing algorithm accuracy.

**Keywords:** object detection; semi-supervised learning; feature alignment; feature pyramid; ASFF

收稿日期: 2023-03-19 修回日期: 2023-05-21 文章编号: 1672-058X(2025)01-0035-07

作者简介: 汤文兵(1967—), 男, 高级工程师, 从事计算机控制研究。

通讯作者: 李菲(1998—), 女, 河南郑州人, 硕士研究生, 从事方向计算机视觉研究。Email: 1403038232@qq.com.

引用格式: 汤文兵, 李菲. 基于特征对齐和特征融合的半监督目标检测算法[J]. 重庆工商大学学报(自然科学版), 2025, 42(1): 35-41.

TANG Wenbing, LI Fei. Semi-supervised object detection algorithm based on feature alignment and feature fusion[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2025, 42(1): 35-41.

## 1 引言

近年来,随着深度学习的发展,在目标检测任务上取得了显著的成绩。但这种优异的表现很大程度上依赖于训练一个具有大量带有人工标注标签的图像网络。随着对更强大、更准确检测模型需求的不断增加,收集和标记更多数据的需求也在增加。然而,对于大规模数据集,手工标注类标签和边界框是一项非常昂贵和繁琐的工作。为了减少训练对象检测器的标签监督,提出了半监督目标检测(Semi-Supervised Object Detection, SSOD)算法,旨在利用有限的标记数据和更丰富的未标记数据来提高性能。

在全监督目标检测算法中,两阶段目标检测算法缺点是实现复杂、计算量大,相比一阶段算法推理速度较慢。在半监督目标算法下,想要继承一阶段算法推理速度快的优点,将其扩展到半监督目标检测算法中,首当其冲的是 Unbiased Teacher V2 算法,其算法是第一个将一阶段无锚框算法应用到半监督目标检测中的,因为它们更容易和有效地部署在资源有限的设备上,无需繁重的预处理/后处理。

目前,构建半监督目标检测的主流方式有两种,即伪标签学习<sup>[1]</sup>和一致性正则化<sup>[2]</sup>。伪标签学习利用标注好的小部分数据集先训练一个模型,然后根据模型对剩余标签进行推理,利用非极大值抑制(Non-Maximum Suppression, NMS)技术通过 Threshold 筛选出合适的标签。而一致性正则化核心思想是对于一个给定的输入,即使人为施加扰动,网络仍然可以产生趋向于原始期望的输出。相比之下,伪标签学习技术在目标检测领域更受关注,大多数都是基于 Faster RCNN<sup>[3]</sup>的两级锚点检测器,如 Soft Teacher<sup>[4]</sup>、Unbiased Teacher<sup>[5]</sup>等半监督目标检测算法。

如今,许多半监督算法在目标检测领域开展,主要使用伪标签学习,通过教师学生框架来进行半监督目标检测。首先,未标注的图片通过不同程度的增强,将其传送给学生和教师模型,通过 EMA (Exponential Moving Average, 指数滑动平均)<sup>[6]</sup>更新教师模型,教师模型产生伪标签,两者进行共同学习。受半监督分类算法 FixMatch<sup>[7]</sup>的启发,Sohn 等<sup>[8]</sup>提出了 STAC 框架,将半监督学习(Semi-Supervised Learning, SSL)算法从分类任务应用到检测任务,但存在伪标签无法实时更新等问题;针对这个问题,Liu 等随后提出 Unbiased Teacher 模型,该模型提出了一个 Student-Teacher 相互学习的半监督训练范式,引入 Focal loss 损失函数来解决类之间的不平衡,但其伪标签存在定位和分类不一致问题;Xu 等之后提出 Soft Teacher 模型,该模型提出了软教师机制和盒抖动方法。通过伪盒回归的学习,

来选择可靠的伪盒,但不能根据类自适应地选择伪标签。Unbiased Teacher V2<sup>[9]</sup>的动机是填补没有一个半监督框架可以应用在 anchor-free 检测器上的问题,Unbiased Teacher V2 因此选择将之前半监督技术用在 FCOS<sup>[10]</sup>这种 anchor-free 检测器上。但是,FCOS 等单级检测器缺乏 RPN 和 ROI Pooling/Align 来提取准确的目标信息用于定位质量估计,同时 FCOS 的预测中心度不能很好地代表定位质量。

针对以上半监督目标算法的不足,提出了一种基于特征对齐和特征融合的算法。通过优化目标分类和目标定位分支,增强两个任务之间的交互性,在该模型的头部添加 T-head<sup>[11]</sup>特征对齐模块。通过解决特征金字塔不同尺度特征之间的不一致性,在 FPN 模块加入了 ASFF<sup>[12]</sup>注意力机制。实验结果表明:改进的 Unbiased Teacher V2 算法能得到有效的提升,并且能够在一定程度上提高伪标签的质量。

## 2 目标检测算法的分析与改进

### 2.1 半监督学习框架

半监督学习有潜力提高使用未标记数据的机器学习模型的预测性能,更多地应用到半监督分类上。最近在半监督目标检测中取得了显著的进展,主要包括 3 个主要模型:PI 模型<sup>[13]</sup>、Temporal Ensembling<sup>[14]</sup>模型、Mean Teacher 模型<sup>[14]</sup>。采用 Mean Teacher 模型作为主要的半监督学习方法,其结构如图 1 所示。均值教师在半监督学习中引入两个网络:一个是主网络(Student Network),另一个是辅助网络(Teacher Network)。主网络通过已经标记的数据进行训练,而辅助网络则根据未标记数据生成伪标签,并将其用于训练。在训练过程中,学生模型的参数通过 EMA 对教师参数进行更新。均值教师的主要思想是通过引入辅助网络,以未标记数据的伪标签作为额外的监督信息来提升模型的泛化能力,减轻过拟合现象。这种方法利用了未标记数据中的信息并通过一致性训练来实现半监督学习的性能提升。

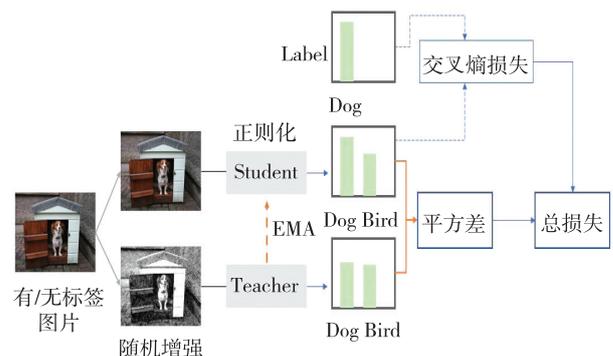


图 1 Mean Teacher 模型

Fig. 1 Mean Teacher model

## 2.2 FCOS 模型

在 Unbiased Teacher V2 中,已说明无锚检测器使用带有中心度的框选择不能很好地工作。FCOS 在实验中没有中心度分支时,它的平均精度、平均 IOU 值以及前 5 个检测结果的平均 IOU 值如表 1 所示。其中 w/o 表示在一个模型中没有某个模块。

表 1 FCOS 预测伪标签比较

Table 1 Comparison of predicting pseudo labels for FCOS

方法	AP	平均 IOU	Top-5 IOU	定位质量与分类相关分数
w/o centerness	25.2	0.369	0.585	0.235
FCOS	26.0	0.369	0.593	0.279

由表 1 中数据可知:无中心度分支的 FCOS 和 FCOS 检测结果的平均 IOU 均为 0.369,说明两者定位能力相差不大。无中心度分支 FCOS 分类得分选择的前 5 名在 IOU 上的表现仍然较差(0.585 比 0.593),这表明 FCOS 选择高质量伪标签的能力较弱。同时,对于分类分数与定位质量之间的相关性,没有中心度的 FCOS 的定位质量与分类相关分数小于 FCOS(0.235 比 0.279)。这些统计数据表明:FCOS 的分类与定位之间存在着更为严重的不一致。因此进行改进,提出了 T-head 特征对齐算法去解决这个问题。

## 2.3 改进的 Unbiased Teacher V2 算法

本节介绍了基于 Unbiased Teacher V2 的改进后的整体模型,其结构如图 2 所示。

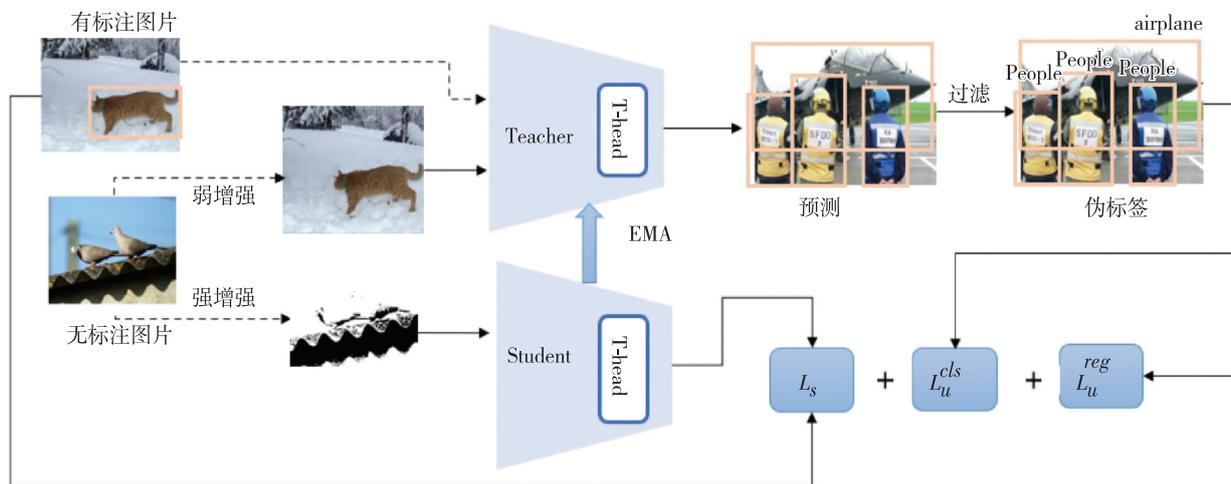


图 2 改进后的 Unbiased Teacher V2 整体框架

Fig. 2 Improved overall framework of Unbiased Teacher V2

Unbiased Teacher V2 是基于无锚框的一阶段半监督目标检测算法,采用 Student-Teacher 双分支结构,两者默认采用 FCOS 目标检测模型。

首先,教师主导的分支中,在 Burn In 阶段将有标注图片通过多个 epoch 有监督训练得到初始检测器,并将初始检测器得到的模型送入教师模型中。无标签数据分别通过不同程度的强增强和弱增强,将通过弱增强的数据送入教师模型,使用在初始阶段训练的模型,将无标签数据也送入教师模型,产生相应的伪标签,最后,通过 NMS 将置信度低于某值的伪标签过滤掉。

同时,将通过强增强生成的无标签数据送入学生模型,在训练过程中,学生模型的参数按照式(1)来对教师参数进行更新。上述操作反复进行,直到训练完成,最后取验证指标较高的 Teacher 或者 Student 模型作为可用模型:

$$\theta'_t = \alpha \theta'_{t-1} + (1-\alpha) \theta_t \quad (1)$$

其中: $\theta'_t$  是在训练步骤  $t$  中权值  $\theta$  的 EMA 值, $\alpha$  是平滑

系数。

## 2.4 ASFF 特征融合机制

为解决特征金字塔不同尺度之间的不一致问题,引入了 ASFF 自适应空间特征融合机制,学习了在空间上过滤冲突信息以抑制梯度反传时候不一致的方法,从而改善了特征的比例不变性,并且推理开销降低。首先利用  $1 \times 1$  卷积调整通道数,使其通道数一致,再进行插值操作调整到一样的大小;然后采用  $3 \times 3$  卷积操作,并在卷积的基础上添加一个步长为 2 的 maxpooling,这样调整后,P3 层就能和 P5 层尺寸一致。在图 3 中,以 ASFF-2 为例,展示了 ASFF 的具体结构。首先从 FPN 中产生特征  $x_1, x_2$ ,对应产生的权重为  $\alpha, \beta, x^{1-2}$  表示特征层从 1 层调整到 2 层后某处的特征向量,对应的权重和特征进行相乘相加,ASFF2 层结构公式如式 2 所示:

$$S_{ASFF2} = x^{1-2} \cdot \alpha^2 + x^{2-2} \cdot \beta^2 \quad (2)$$

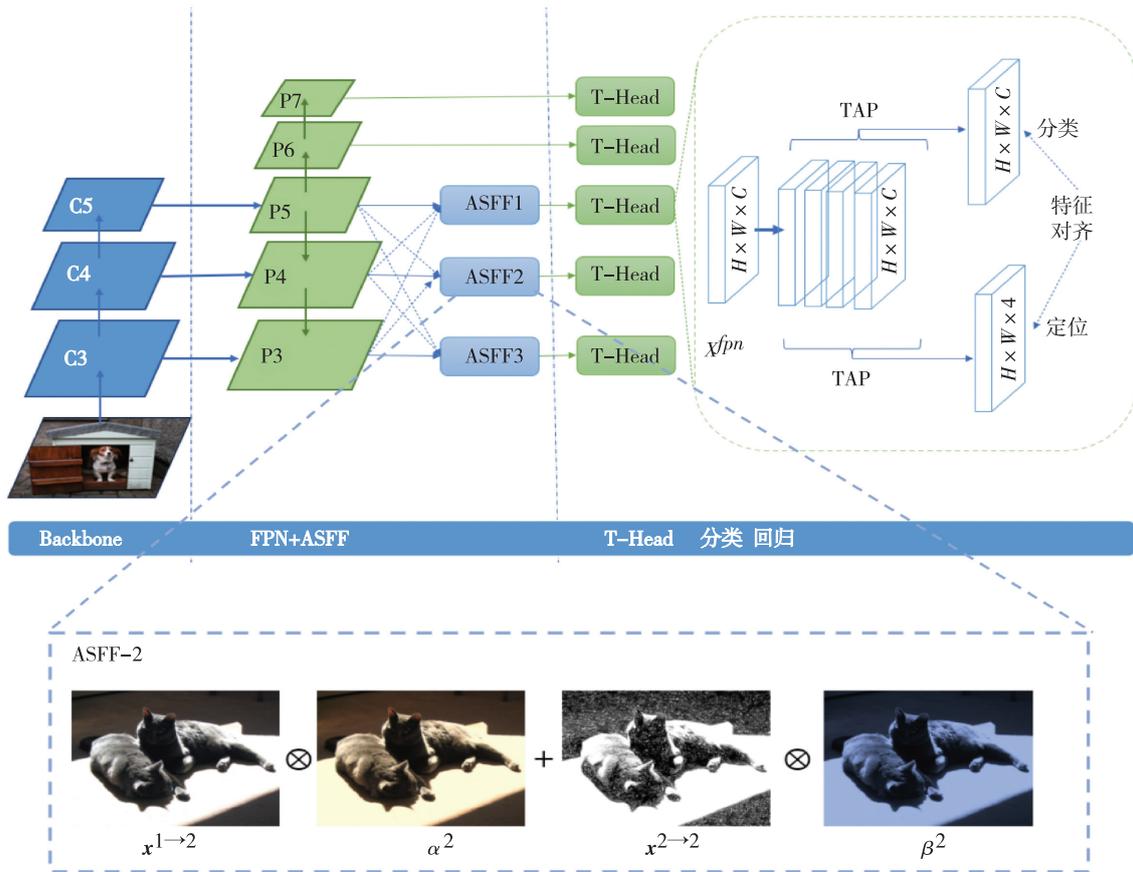


图 3 改进后 FCOS 整体结构

Fig. 3 Improved overall structure of FCOS

2.5 T-head 特征对齐

该算法是一种改进一阶段检测头的方法 T-head, 改进目标检测中的头部分支结构, 设计该算法来增强分类和定位分支之间的交互性, 以实现特征对齐。在 FCOS 中, 没有显式的定位分支, 分类和定位分支是通过共享权重的方式进行对齐。T-head 结构如图 3 所示: 首先, T-Head 对 FPN 特征进行预测; 其次, 使用预测来计算每个锚点处的任务对齐度量, 基于该度量, TAL 为 T-Head 产生学习信号; 最后, T-Head 在反向传播过程中使用 TAL 计算的学习信号自动调整其分类概率和定位预测, 对分类和定位的分布进行相应地调整。使用特征提取器来学习来自多个卷积层的任务交互特征堆栈。特征提取器使用具有激活函数的  $N$  个连续卷积层来计算任务交互特征, 如式(3)所示:

$$X_k^{inter} = \begin{cases} \delta(\text{conv}_k(X^{fpn})), k=1 \\ \delta(\text{conv}_k(X_{k-1}^{inter})), k>1 \end{cases} \quad \forall k \in \{1, 2, \dots, N\} \quad (3)$$

其中:  $X^{fpn}$  表示 FPN 的特征,  $\text{conv}_k$  表示  $k$  个卷积层和 ReLU 函数。

2.6 损失函数

在每个训练迭代中, 使用学生和教师之间的相对

预测信息。通过比较 Teacher 预测的边框与真实边框的回归距离, Student 预测的边框与真实边框的回归距离, 将具有较大不确定性的伪标签过滤, 减少质量较差的伪标签。使用负幂对数似然损失 (NPLL) 来作为回归损失。回归损失如式(4)所示:

$$L_{reg}^{sup} = \sum_i \eta_i \left( \sum \frac{(d_s - d_g)^2}{2\delta_s^2} \right) + \frac{1}{2} \log \delta_s^2 + 2 \log 2\pi \quad (4)$$

其中: 真值标签  $d_g$  来判断教师模型预测的好坏,  $d_s$  表示学生的预测边界距离,  $\eta_i$  表示预测框与真实框的 IOU 分数,  $\delta_s$  表示学生的预测不确定性。

明确考虑教师的定位不确定性  $\delta_t^i$  和学生的定位不确定性  $\delta_s^i$  来进行伪标签选择。通过选择教师的定位不确定性比学生低的有利实例, 无标签回归损失被定义为式(5)所示:

$$L_{reg}^{unsup} = \begin{cases} \sum_i \| \tilde{d}_t^i - \tilde{d}_s^i \|, \text{if } \delta_t^i + \sigma \leq \delta_s^i \\ 0, \text{otherwise} \end{cases} \quad (5)$$

其中:  $\tilde{d}_t^i$  是教师的回归预测,  $\tilde{d}_s^i$  是学生的回归预测,  $\sigma \geq 0$  为“教师”和“学生”的定位不确定度之差。无监督回归损失是在边界级别而不是盒级别计算的。

### 3 仿真实验

#### 3.1 实验数据集

本算法基于 Unbiased Teacher V2 进行改进,相比 STAC、Unbiased Teacher 等算法都有一定比例提升。采用 FCOS 无锚框模型,以 ResNet50 为主干,以 FPN+ASFF 为颈部,加入 T-head 作为头部。本实验采用 COCO 数据集和 VOC 数据集,遵循 Unbiased Teacher 的数据集划分策略,将数据集划分为 1%、5%、10% 的有标签数据,进行传统的目标检测,剩余数据作为无标注训练数据, val2017 作为验证数据集。对于 VOC 数据集, VOC07-trainval 作为有标注数据, VOC12-trainval 作为无标注训练数据, VOC07-test 作为验证数据集。

#### 3.2 实验环境

为避免实验环境在不同系统上的差异,本实验采用远程服务器进行,采用开源框架 Detectron2, Unbiased Teacher V2 开源框架进行开发。采用 8 块 3080Ti-12G 的 GPU,显存为 96G, CUDA 版本为 12.1, Detectron2 版本为 0.6, Pytorch 版本为 1.10。

本实验超参数设置初始学习率为 0.01 的 SGD 优化器,最大迭代次数为 180 000,图像批处理数为 16,其中 8 张图片为标注数据,剩余 8 张图片为未标注数据,指数移动平均率为 0.999 9,预训练次数为 30 000,使用无监督损失权  $\lambda_u = 3.0$ ,分类阈值  $\tau = 0.5$ ,采用 Unbiased Teacher 中使用的数据增强,并且沿用 Soft Teacher 的尺度抖动。

#### 3.3 不同半监督方法的性能比较

STAC、Unbiased Teacher、Unbiased Teacher V2 等不同半监督目标检测方法在 COCO 数据集下,随机抽取 1%、5%、10% 比例的数据,组成有标签数据集,其余作为无标签数据集,如表 2 所示。

表 2 实验结果对比 (COCO 数据集)

Table 2 Comparison of experimental results (COCO dataset)

	有标签数据集		
	1%	5%	10%
STAC	11.97	20.36	24.31
Unbiased Teacher	14.61	23.99	28.18
Instant-Teaching	18.05	26.75	30.04
Humble Teacher	16.96	27.70	31.62
Soft Teacher	18.8	28.10	31.10
Unbiased Teacher V2	22.71	30.08	32.61
本文方法	23.89	31.32	33.97

在 COCO 数据集下,本文方法相比原论文算法在 1% 的有标签数据集下精度提高了 1.18%,在 5% 的有标签数据集下精度提高了 1.24%,在 10% 的有标签数

据集下精度提高了 1.36%。再添加上 T-head 和 ASFF 后,精度都有一定的提升,图 4 是 Unbiased Teacher V2 和本文方法在 10% 有标签数据集下  $5 \times 10^4$  训练次数的平均精度性能表现。图 5 是这两种算法在  $9 \times 10^4$  训练次数中的伪标签数目比较。

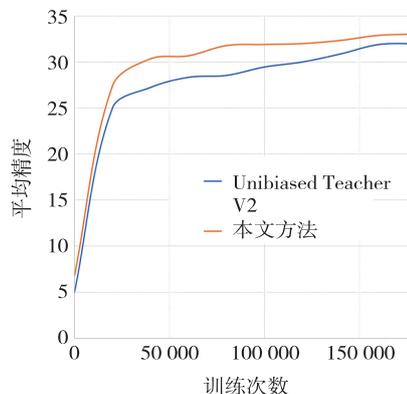


图 4 两种算法性能比较 (COCO 数据集)

Fig. 4 Comparison of performance between two algorithms (COCO dataset)

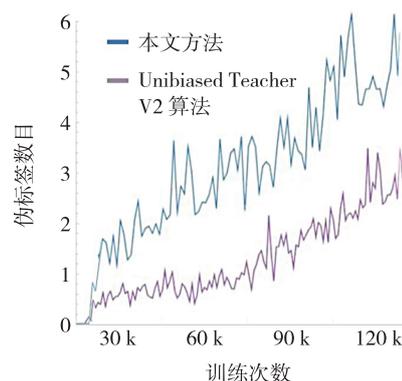


图 5 两种算法伪标签数目比较 (COCO 数据集)

Fig. 5 Comparison of the number of pseudo labels between two algorithms (COCO dataset)

表 3 展示了在 VOC 数据集下的性能比较,其中,  $AP_{50}$  指预测框与真值的 IOU 大于 0.5 的平均精度,  $AP_{50:95}$  指 IOU 从 0.5 到 0.95 的平均精度,测试结果如表 3 所示。

表 3 实验结果对比 (VOC 数据集)

Table 3 Comparison of experimental results (VOC dataset)

	$AP_{50}$	$AP_{50:95}$
STAC	77.45	44.64
Instant-Teacher	69.90	46.28
Humble Teacher	70.37	45.90
Soft Teacher	72.80	46.1
Unbiased Teacher	80.51	54.45
Unbiased Teacher V2	81.29	56.87
本文方法	82.67	57.45

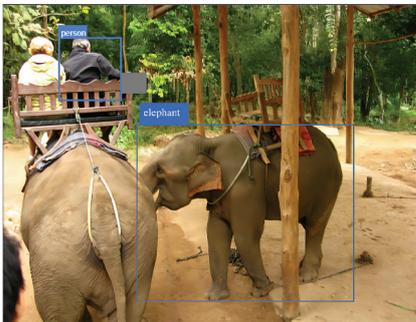
在 VOC 数据集上,有 20 个数据类,本文方法相比 Unbiased Teacher V2 模型,  $AP_{50}$  提高了 1.38%,  $AP_{50:95}$

提高了 0.58%。

在 VOC 数据集上,Unbiased Teacher V2 和本文算法可视化检测结果比较,如图 6 所示。



(a) Unbiased Teacher V2 图像可视化



(b) 本文方法图像可视化

图 6 图像可视化比较(VOC 数据集)

Fig. 6 Comparison of image visualization(VOC dataset)

### 3.4 消融实验对比

为了验证 T-head 和 ASFF 这两种方法的有效性,分别在 COCO 数据集和 VOC 数据集上进行消融实验。

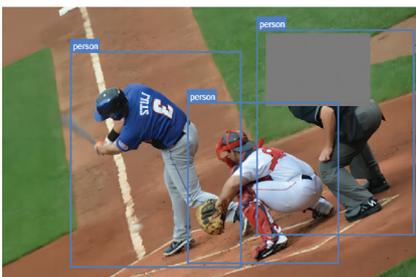
首先在 VOC 数据集进行 T-head 方法的消融实验对比,其实验数据如表 4 所示。

表 4 实验精度对比(VOC 数据集)

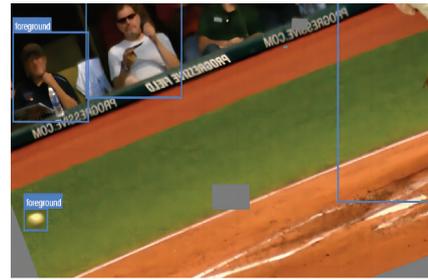
Table 4 Comparison of experimental accuracies(VOC dataset)

	AP <sub>50</sub>	AP <sub>50:95</sub>
Unbiased Teacher V2	81.29	56.87
Unbiased Teacher V2+T-head	81.56	57.23

在 VOC 数据集上,加入 T-head 后伪标签可视化结果如图 7 所示。



(a) Unbiased Teacher V2+T-head



(b) Unbiased Teacher V2

图 7 改进后模型产生的伪标签(20 k 训练次数)

Fig. 7 Pseudo labels generated in the improved model(20 k training times)

采用 VOC 数据集进行 ASFF 方法的消融实验对比,该数据集包含了 20 个数据类,其实验数据如表 5 所示。

表 5 ASFF 方法的实验精度对比(VOC 数据集)

Table 5 Comparison of experimental accuracies of ASFF method(VOC dataset)

	AP <sub>50</sub>	AP <sub>50:95</sub>
Unbiased Teacher V2	81.29	56.87
Unbiased Teacher V2+ASFF	81.47	57.33

随机抽取 1%、5%、10% 比例的数据,组成有标签数据集,其余作为无标签数据集,进行消融实验对比,其实验数据如表 6 所示。

表 6 改进后的精度对比(COCO 数据集)

Table 6 Comparison of improved accuracies(COCO dataset)

	有标签数据集		
	1%	5%	10%
Unbiased Teacher V2	22.71	30.08	32.61
Unbiased Teacher V2+T-head	23.12	30.54	32.92
Unbiased Teacher V2+ASFF	22.54	30.78	33.08

综上,针对 COCO 数据集和 VOC 数据集,通过改进 Unbiased Teacher V2 算法,精度在这两个数据集上都有一定的比例提升。实验数据可能在不同实验配置下会有一些的小比例波动。若本次实验能增加迭代次数,可能会得到更佳的实验效果。在图 6 和图 7 中,人物和背景出现一些灰色的方块,是在迭代过程中,由于刚开始迭代次数不充足,教师模型还未训练完全,数据预处理或标注过程中引入了噪声造成的,这些灰色方块不是目标或景的一部分。在实验环境部分提到,图像批处理大小为 16,它们是由于图像压缩、传输等因素导致的。

## 4 结论与展望

### 4.1 结论

通过改进的 Unbiased Teacher V2,在特征头中加入 T-head 实现特征对齐,实现了分类分支和定位分支的

对齐,增强了两个分支之间的交互性。在 FPN 中加入 ASFF 特征融合机制,在 COCO 数据集和 VOC 数据集上有了不同程度的提升,在一定程度上降低了伪标签的噪声,缓解了正负样本不均衡问题,并尝试降低无锚框检测器与 SSOD 之间的差距。

## 4.2 展望

通过消融实验进一步证明了全监督目标检测的某些方法可以应用到半监督目标检测中,进一步缓解了半监督模型伪标签不一致问题。今后也会尝试将更多全监督目标检测的一些方法应用到半监督目标检测中来提高算法精度,并且尝试应用到具体场景,来减少手工标注。

## 参考文献(References):

- [1] LEE D H. Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks [C]// Workshop on challenges in representation learning. ICML, 2013, 3(2): 896–904.
- [2] ABUDUWEILI A, LI X, Shi H, et al. Adaptive consistency regularization for semi-supervised transfer learning[C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Piscataway: IEEE Press, 2021: 6923–6932.
- [3] GIRSHICK R. Fast R-CNN [C]//Proceedings of the IEEE international conference on computer vision. Piscataway: IEEE Press, 2015: 1440–1448.
- [4] XU M, ZHANG Z, HU H, et al. End-to-end semi-supervised object detection with soft teacher[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 3060–3069.
- [5] DENG J, LI W, CHEN Y, et al. Unbiased mean teacher for cross-domain object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 4091–4101.
- [6] 吕君环,许柯,王东.基于指数移动平均知识蒸馏的神经网络低比特量化方法[J].模式识别与人工智能,2021,34(12):1143–1151.  
LYU Jun-huan, XU Ke, WANG Dong. Neural network low-bit quantization method based on exponential moving average knowledge distillation [J]. Pattern Recognition and Artificial Intelligence, 2021, 34(12): 1143–1151.
- [7] LUCAS T, WEINZAEPFEL P, ROGEZ G. Barely-supervised learning: Semi-supervised learning with very few labeled images[J]. Proceedings of the AAAI Conference on Artificial Intelligence. 2022, 36(2): 1881–1889.
- [8] 王睿,樊思杨,许婧文,等.采用 SVD 协同训练的半监督实例级目标检测[J].光学精密工程,2023,31(13):2000–2007.  
WANG Rui, FAN Si-yang, XU Jing-wen, et al. Semi-supervised instance-level target detection using SVD collaborative training[J]. Optical Precision Engineering, 2023, 31(13): 2000–2007.
- [9] LIU Y C, MA C Y, KIRA Z. Unbiased teacher v2: Semi-supervised object detection for anchor-free and anchor-based detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 9819–9828.
- [10] TIAN Z, SHEN C, CHEN H, et al. Fcos: Fully convolutional one-stage object detection[C]//Proceedings of the IEEE/CVF international conference on computer vision. Piscataway: IEEE Press, 2019: 9627–9636.
- [11] FENG C, ZHONG Y, GAO Y, et al. Tood: Task-aligned one-stage object detection[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 3490–3499.
- [12] QIU M, HUANG L, TANG B H. ASFF-YOLOv5: Multielement detection method for road traffic in UAV images based on multiscale feature fusion[J]. Remote Sensing, 2022, 14(14): 3498.
- [13] 陆健强,梁效,余超然,等.基于坐标注意力机制与高效边界框回归损失的线虫快速识别[J].农业工程学报,2022,38(22):123–132.  
LU Jian-qiang, LIANG Xiao, YU Chao-ran, et al. Rapid identification of nematodes based on coordinate attention mechanism and efficient bounding box regression loss [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(22): 123–132.
- [14] CAO X, CHEN H, LI Y, et al. Uncertainty aware temporal-ensembling model for semi-supervised abus mass segmentation[J]. IEEE transactions on medical imaging, 2020, 40(1): 431–443.
- [15] CUI W, LIU Y, LI Y, et al. Semi-supervised brain lesion segmentation with an adapted mean teacher model[C]// Lecture Notes in Computer Science, Cham: Springer International Publishing, 2019: 554–565.

责任编辑:李翠薇