

Data-Uncertainty Guided Multi-Phase Learning for Semi-Supervised Object Detection



合肥工业大学

张宇涛

2023.11.09

Wang Z, Li Y, Guo Y, et al. Data-uncertainty guided multi-phase learning for semi-supervised object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 4568-4577.

提 纲

01

研究介绍

02

理论方法

03

实验结果

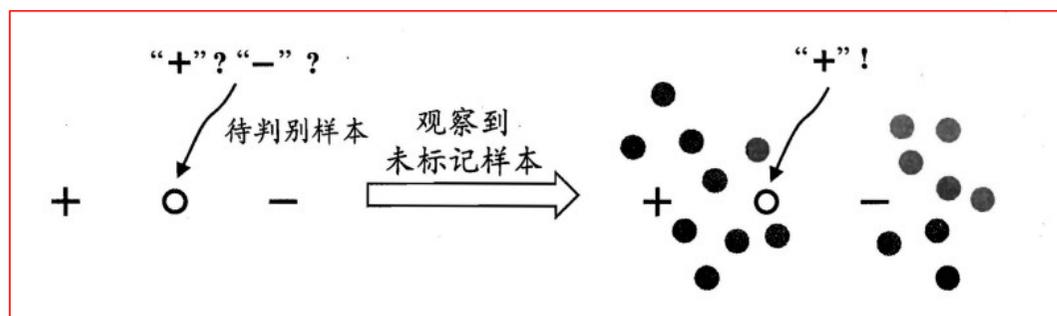
04

总结

— 研究介绍

➤ 研究背景

半监督学习 (semi-supervised learning) : 让学习器不依赖外界交互、自动地利用未标记样本来提升学习性能。



着色圆点标识未标记样本

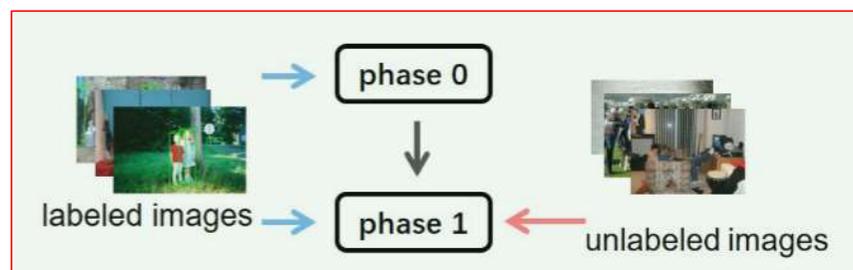
伪标记 (pseudo-label) : 对未标记数据进行预测, 并将预测结果作为标记数据。

— 研究介绍

➤ 研究背景

半监督目标检测 (semi-supervised object detection (SSOD))

在有少量已标记数据和大量无标记数据的情况下，通过利用已标记数据和无标记数据之间的关系，来提高目标检测性能的目标检测方法。



一般的SSOD方法:

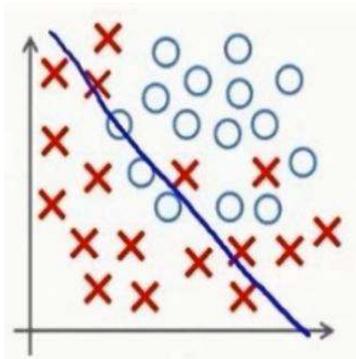
利用单阶段学习方案，通过在标记图像上预先训练的全监督模型，获得未标记图像的伪标记。这些伪标记被视为未标记数据的真实值，并与标记数据的注释集成以训练 SSOD 模型。

— 研究介绍

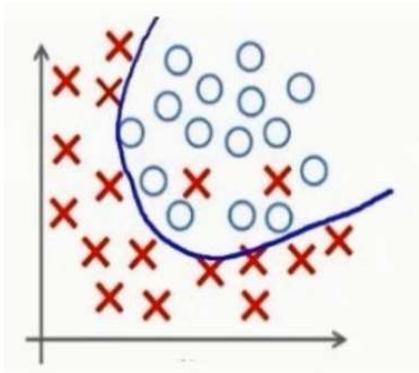
➤ 研究背景

单阶段SSOD方法的不足:

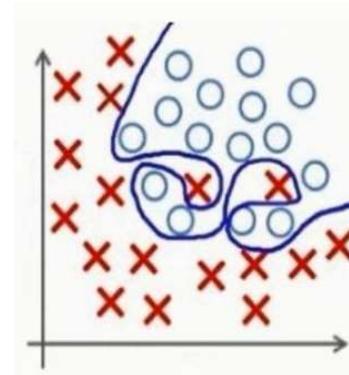
- 易受噪声标签的影响: 单阶段的SSOD方法直接将无标注数据预测为已标注数据, 因此容易受噪声标签的影响。
- 容易产生过拟合的问题。
- 对数据的依赖性较强, 需要大量的数据才能达到良好的性能。



欠拟合



好的拟合

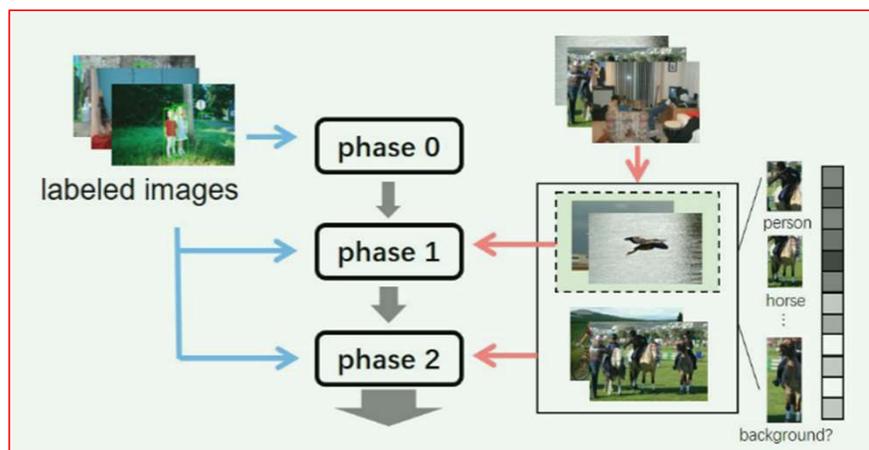


过拟合

二 理论方法

➤ SSOD多阶段学习

- 1) 对具有低不确定性的简单未标记图像进行 SSOD 训练。（生成的伪标记受噪声影响小）
- 2) 对具有高不确定性的复杂未标记图像进行 SSOD 训练。（利用多个模型，分别擅长不同难度级别的图像）

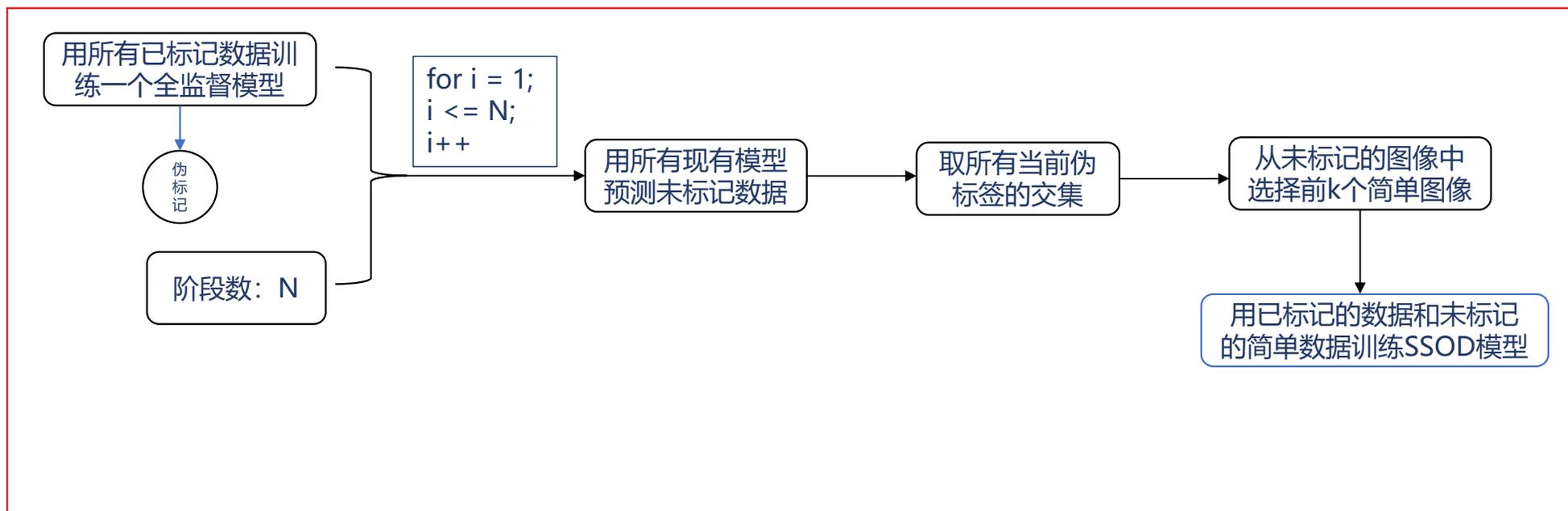


多阶段SSOD

- 引入不确定性来指导不同训练阶段的图像选择，通过数据的不确定性来进行SSOD的多阶段学习。
- 添加复杂图像来训练其他后续检测模型。最后，分析不同模型在不同复杂程度图像中的检测表现，充分利用不同模型的不同能力。

二 理论方法

➤ 多阶段 SSOD 学习的总体流程



- 综合之前阶段的所有模型来创建新的伪标签；
- 利用交集运算提高精度，降低伪标签的不确定性；
- 最后集成所有模型的测试结果，以生成最终结果。

二 理论方法

➤ 图像不确定性的选择

由于未标记的图像没有注释，因此在区分图像的不确定性时引入以下公式：

$$\{(bb_{mn}, s_{mn})\}_{m=1}^M$$

$$\bar{s}_m = \sum_{m=1}^M s_{mn}/M$$

其中，“ bb_{mn} ”表示边界框，“ s_{mn} ”表示置信度数值， M 表示边界框的数量。

s_m 小的图像被认为是困难的图像，并且在前几个阶段会被过滤掉。

这种基于图像不确定性的选择策略引导不同的检测模型选择更确定的图像，以减轻标签噪声过拟合的问题。

三 实验结果

➤ 实验

在PASCAL VOC 和 MSCOCO 上评估所提出的 SSOD 方法;

对于 VOC使用 VOC 2007 trainval (5,011 张图像) 作为标记数据, 使用 VOC 2012 trainval (11,540 张图像) 作为未标记数据, 然后在 VOC 2007 测试 (4,952 张图像) 上进行评估。

Model	Backbone	Method	L	Un	mAP
Faster RCNN	ResNet50	FS	VOC07	-	74.8
		Baseline	VOC07	VOC12	75.6
		DD [33]	VOC07	VOC12	76.0
		ours	VOC07	VOC12	78.6
		FS	VOC0712	-	81.2
SSD300	VGG16	FS	VOC07	-	70.2
		Baseline	VOC07	VOC12	71.8
		CSD [16]	VOC07	VOC12	72.3
		ISD [17]	VOC07	VOC12	73.3
		ours	VOC07	VOC12	74.5
		FS	VOC0712	-	77.2

(FS: 全监督模型)
(L: labeled data, Un: unlabeled data)
(CSD: 基于一致性的半监督学习)
(ISD: 基于插值的半监督学习)

对基于两阶段的Faster RCNN 和一阶段的SSD 的 SSOD 进行比较;

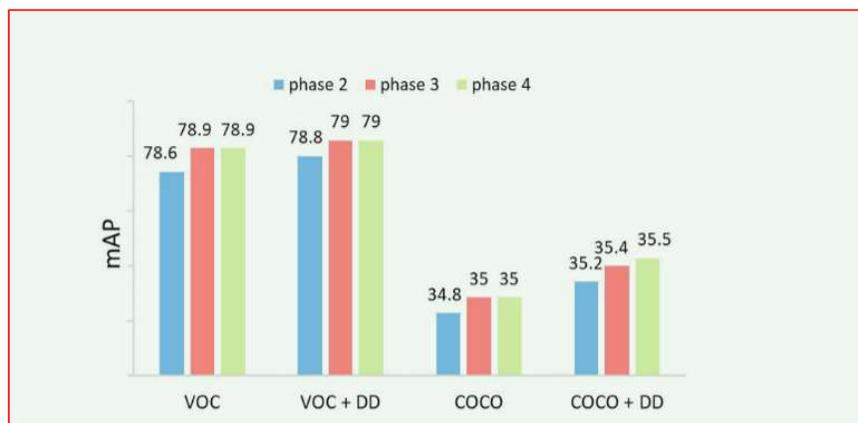
对于Faster RCNN, 基于数据不确定性的多阶段学习实现了78.6%的mAP, 与Baseline相比增加了3%;

对于SSD, 实现了74.5%的mAP;

四 结论

➤ 总结

最后将此方法扩展到更多的阶段。对于 VOC 数据集，性能从两阶段的78.6% 略微提高到三阶段的78.9%，而四阶段学习并没有更高的mAP。因此，生成两个半监督模型的两阶段学习已经能够充分描述大多数未标记信息。



该方法对挖掘未标记知识具有非凡的能力。而半监督目标检测是一个具有挑战性的问题，此后将进一步探索如何更有效地利用未标记数据。

Q&A



合肥工业大学