鲁棒性对抗性训练对小物体识别的提高

孙志华

辽宁何氏医学院　辽宁沈阳

【摘要】提出了一种基于对抗性训练的小物体识别方法，用于提高模型在复杂环境下的鲁棒性。该方法使用 YOLOv3 作为基础框架，冻结前几层并只更新后续层。通过真实图像和生成的复杂图像进行训练，使模型适应两种数据。采用细致的 Adam 优化器、较小的学习率和批大小进行训练。实验结果显示，该方法在小物体数据集上的 mAP 高于 YOLOv3，并在复杂测试集上具有高精度，表明对抗训练确实增强了模型的鲁棒性。然而，该方法的速度下降到 YOLOv3 的 0.7 倍，因为对抗图像较复杂，需要更长的前向传播时间。总之，对抗性训练可以显著提高小物体识别模型的鲁棒性，但也会带来速度下降和数据集依赖性增加的问题。需要进一步改进模型和训练策略，以在保持鲁棒性的同时尽量减少速度和数据集影响。综上所述，该研究提出了一种基于 YOLOv3 和对抗性训练的小物体识别方法，可以显著提高模型在复杂环境下的鲁棒性，但还需要进一步改进和优化。

【关键词】小物体识别；YOLO；对抗性训练；鲁棒性

【收稿日期】2023 年 4 月 6 日　【出刊日期】2023 年 5 月 23 日　【DOI】10.12208/j.aics.20230023

Robustness adversarial training to improve small object recognition

Zhihua Sun

Liaoning He university, Shenyang, Liaoning

【**Abstract**】A small object recognition method based on adversarial training is proposed to improve the robustness of the model in complex environments. The method uses YOLOv3 as the base framework, freezing the first few layers and updating only the subsequent layers. The model is trained through real images and generated complex images to adapt to both types of data. Training with a detailed Adam optimizer, small learning rate and batch size. Experimental results show that the proposed method has a higher mAP on small object datasets than YOLOv3, and has high accuracy on complex test sets, indicating that adversarial training does enhance the robustness of the model. However, the speed of this method drops to 0.7 times that of YOLOv3 because the counter image is more complex and requires a longer forward propagation time. In conclusion, adversarial training can significantly improve the robustness of small object recognition models, but it also brings the problem of reduced speed and increased dataset dependency. Further improvements to the model and training strategies are needed to minimize speed and data set impact while maintaining robustness. In summary, this study proposes a small object recognition method based on YOLOv3 and adversarial training, which can significantly improve the robustness of the model in complex environments, but it needs further improvement and optimization.

【**Keywords**】Small object recognition; YOLO; Adversarial training; Robustness

小物体识别是一个重要而具有挑战性的目标检测任务。小物体面积小，信息少，识别系统难以得到特征，导致精度低。特别是在复杂环境下，小物体易被遮挡或背景干扰，识别系统几乎无法正确检测，限制了小物体识别技术实用性。为提高小物体识别性能，现有方法主要从以下几个方面探索：

①更深入特征提取。使用更强大的卷积神经网络可以提取更丰富语义特征，为小物体分类判断提供更准确依据。

②更精细分类方法。使用 Softmax 分类、SVM

可以更好区分小物体，提高最终识别精度。

③更丰富训练数据。扩充数据集可以学习更全面与详细小物体特征，增强模型泛化能力。

④注意力机制。注意力模块使模型更聚焦小物体区域，避免背景干扰，提高小物体特征利用效率。

现有方法对小物体识别的改进有限，尤其在复杂环境下的识别精度不高，限制了其实用性。为此，提出了一种基于 YOLOv3 和对抗性训练的方法，以提高模型对这类图像的鲁棒性。该方法利用 YOLOv3 的快速和高精度优势作为基础框架，并结合对抗性训练思想，通过生成复杂图像来训练模型，提高其对这些图像的识别精度。该方法已在其他任务上取得显著提高，但在小物体识别方面应用较少。通过设计对抗数据生成器生成含小物体的复杂图像，将其与真实图像一起用于训练 YOLOv3 模型，以增强其鲁棒性。这种方法有望显著提高 YOLOv3 在复杂环境下小物体识别任务的适应性，但可能会降低速度，需要进一步改进。综上所述，小物体识别面临挑战，特别是在复杂图像中的识别精度不高，为此提出了一种基于 YOLOv3 和对抗性训练的方法，为小物体识别在复杂环境下的发展提供新思路。

1 资料与方法

1.1 YOLOv3 框架

YOLO 系列是实时目标检测方法，具有高速和高精度，适用于实际应用。YOLOv3 改进了 backbone 网络、预测网络和损失函数，提高了识别精度。损失函数包括类别、置信度、边界框回归、对象性和比例损失，使用交叉熵计算。通过最小化整个损失函数学习检测物体，产生精确的结果。因此，YOLOv3 适合对抗性训练方法，通过更新提高在复杂图像中的识别性能，增强鲁棒性。这将提升 YOLOv3 在小物体识别领域的实用性。

1.2 对抗数据生成器

通过设计一个对抗数据生成器，我们能够生成包含小物体的复杂图像。这些图像包括复杂的背景、不同光照条件下的小物体以及部分遮挡，增加了小物体识别的挑战。生成器采用 U-Net 结构，通过编码器和解码器提取和生成图像特征。训练生成器时，使用真实图像与标签作为输入，生成器产生具有相同布局和类别但具有变化的小物体和背景的图像。采用 L1 损失和语义分割损失来约束生成器。使用生成的对抗图像训练 YOLOv3 能够增强模型的泛化能力，使其在复杂数据上具有更高的识别精度和鲁棒性。这将在后续实验中得到验证。

1.3 实验结果与分析

我们在 3 个小物体数据集（Bottle、Coco-small、Tiny-vocs）上测试本方法，YOLOv3 的 mAP 分别提高 7.8%、4.3%与 3.1%。同时，我们构建一个复杂的测试集，含有许多复杂背景与光照变化的小物体图像，本方法在该集上的 mAP 达到了 63.7%，较 YOLOv3 的 57.4%有较大提高。这说明本方法确实增强了 YOLOv3 对复杂环境下小物体识别的能力，从而提高了其鲁棒性。我们也测试了本方法的速度，发现其大约是 YOLOv3 的 0.7 倍，这主要是因为对抗图像较为复杂，会带来额外的计算开销。我们对模型进行裁剪与量化，最终在 NVIDIATITANV 上实现了 51.2FPS，这已基本可以满足实时检测的要求，但仍需要进一步优化。

综上，本方法可以显著提高小物体识别模型 YOLOv3 的鲁棒性，使其可以在更加复杂环境下给出较高的识别精度。尽管还存在速度与数据集依赖性等问题，但我们相信随着技术的发展，这些问题可以得到很好的解决。

2 结论

通过本研究，我们证明了对抗性训练可以显著增强小物体识别模型的鲁棒性，使其可以在更加复杂的环境下也给出较高的识别精度。这为小物体识别技术的发展提供了一条新的思路，也使该技术在更加困难的数据集与应用场景中可以发挥更大的作用。但是，本方法也带来了一定的问题，主要体现在两个方面：

①速度下降。对抗图像的复杂性增加了计算开销，导致模型推理速度下降。尽管使用模型裁剪和量化可以部分恢复速度，但还需进一步优化。采用轻量级网络结构如 MobileNet、ShuffleNet 替代 Darknet-53 可在保证鲁棒性的前提下最大程度提高速度。此外，通过模型蒸馏或知识迁移等方法，生成更轻量的网络也是后续工作的探索方向。

②数据集依赖性。本方法需要大量真实图像与生成的对抗图像来训练模型，这使其变得更加依赖数据集。在数据集比较少的情况下，本方法的效果可能会有所下降。我们希望通过持续积累数据，不断优化训练策略，使模型可以在数据集不太丰富的情况下也可以达到较强的鲁棒性。此外，通过学习

统一的图像特征表示，或利用迁移学习与弱监督学习等思想，也可以在一定程度上减少对数据集的依赖，这将是我们后续工作的另一个重要方向。

综上，本研究证明对抗性训练可以显著提高小物体识别模型的鲁棒性，但也会带来速度与数据集依赖性的影响。通过持续优化与完善，这些问题必将得到较好的缓解。

3 附录

我们在 3 个小物体数据集上测试本方法，其中 Bottle 数据集包含 1167 张图像及 5 个类别（糖果瓶、啤酒瓶、搅拌器杯、牛奶瓶、可乐瓶）的小瓶子，Coco-small 包含 5854 张图像及 80 个物体类别，Tiny-vocs 包含 1578 张图像及 20 个类别的小物体。这 3 个数据集涵盖较为广泛的数据分布与物体类别，可以全面评价本方法的效果。

实验结果如表 1 所示。本方法在 3 个数据集的 mAP 分别达到了 88.4%、79.1%与 75.9%，较 YOLOv3 有 7.8%、4.3%与 3.1%的提高。这证明本方法可以有效增强小物体识别模型对标准数据集的适应性，从而提高最终的识别精度。我们也构建了一个复杂的测试集，包含 511 张图像，其中许多图像具有复杂的背景与光照效果，也包含部分遮挡的小物体，这增加了检测的难度。实验结果如表 2 所示，本方法的 mAP 达到 63.7%，较 YOLOv3 的 57.4%也有一定提高。这证明本方法确实增强了模型对复杂环境下小物体检测任务的适应性，这提高了模型的鲁棒性。

表 1　3 个小物体数据集上的识别精度比较（mAP%）

方法	Bottle	Coco-small	Tiny-vocs
YOLOv3	80.6	74.8	72.8
Propose	88.4	79.1	75.9

表 2　复杂测试集上的识别精度比较（mAP%）

方法	复杂测试集
YOLOv3	57.4
Propose	63.7

我们也测试了本方法的速度，发现其是 YOLOv3 速度的 0.7 倍，主要受对抗图像计算开销的影响。我们采取模型裁剪与量化等方法进行优化，最终实现了 51.2FPS，这已基本可以满足实时检测的要求。我们相信通过更加轻量级的网络结构，该速度还可以进一步提高。可以看出，本方法可以更加准确地检测图像中的小物体，这证明其较 YOLOv3 有更强的识别能力，而这主要源于对抗性训练提高的鲁棒性。

综上，详细的实验结果与示例图像证明，本方法可以显著提高小物体识别模型 YOLOv3 的识别精度与鲁棒性，使其可以很好地工作在复杂的环境下。这将有利于小物体识别技术在实际应用中的推广与落地。但是，本方法也存在速度下降的问题，这需要我们在后续工作中加以改进与优化。

参考文献

[1] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[2] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21-37.

[3] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.

[4] Jian S, Kaiming H, Shaoqing R, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision & Pattern Recognition. 2016: 770-778.

[5] Isola P, Zhu J Y, Zhou T, et al. Image-to-image translation with conditional adversarial networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1125-1134.

[6] Hoffman J, Tzeng E, Park T, et al. Cycada: Cycle-consistent adversarial domain adaptation[C]// International conference on machine learning. Pmlr, 2018: 1989-1998.