

[首页](#)[资讯](#)[深度资源](#)[产业视频](#)[GMIS峰会](#)[AI 商用搜索](#)[登录/注册](#)

谷歌开源TensorFlow Object Detection API物体识别系统

By [机器之心](#) 2017年6月16日 17:19

近日，谷歌在其开源博客上发表了一篇名为《Supercharge your Computer Vision models with the TensorFlow Object Detection API》的文章，通过 TensorFlow Object Detection API 将谷歌内部使用的物体识别系统（2016 年 10 月，该系统在 COCO 识别挑战中名列第一）开源给更大的社区，帮助打造更好的计算机视觉模型。机器之心对该文进行了编译，原文链接请见文末。

在谷歌，研究人员开发了最高水平的计算机视觉机器学习系统，不仅可以用于谷歌自身的产品和服务，还可以推广至整个研究社区。创造能够在同一张图片里定位和识别多种物体的机器学习模型一直是业内的核心挑战之一，谷歌宣称自己已投入大量时间训练和实验此类系统。



其中一个模型在示例图片中（来自 COCO 数据集）对象识别的效果

2016 年 10 月，谷歌内部的物体识别系统达到了业内最佳水平，在 COCO 识别挑战中名列第一。自那时起，该系统为大量文献提供了结果，并被一些谷歌的产品所采用，如 NestCam，同样的理念也被识别谷歌街景街道名称和门牌号的 Image Search 采用。

今天，我们很高兴通过 TensorFlow Object Detection API 把该系统开源给更大的研究社区。这个代码库是一个建立在 TensorFlow 顶部的开源框架，方便其构建、训练和部署目标检测模型。我们设计这一系统的目的是支持当前最佳的模型，同时允许快速探索和研究。我们的第一个版本包含：

- 一个可训练性检测模型的集合，包括：

- 带有 MobileNets 的 SSD (Single Shot Multibox Detector)
- 带有 Inception V2 的 SSD
- 带有 Resnet 101 的 R-FCN (Region-Based Fully Convolutional Networks)
- 带有 Resnet 101 的 Faster RCNN
- 带有 Inception Resnet v2 的 Faster RCNN
- 上述每一个模型的冻结权重 (在 COCO 数据集上训练) 可被用于开箱即用推理。
- 一个 Jupyter notebook 可通过我们的模型之一执行开箱即用的推理
- 借助谷歌云实现便捷的本地训练脚本以及分布式训练和评估管道

SSD 模型使用了轻量化的 MobileNet，这意味着它们可以轻而易举地在移动设备中实时使用。在赢得 2016 年 COCO 挑战的研究中，谷歌使用了 Fast RCNN 模型，它需要更多计算资源，但结果更为准确。

如需了解更多细节，请参阅谷歌发表在 CVPR 2017 上的论文：<https://arxiv.org/abs/1611.10012>

准备好了吗？

这些代码在谷歌的计算机视觉应用中非常有用，谷歌希望本次开源能够将这份便利带给所有人。谷歌同时也欢迎更多贡献者参与进来。现在，你可以下载代码，使用 Jupyter notebook 尝试在图片中识别物体，也可以开始在 Cloud ML 引擎中训练你自己的识别器了。

- 代码：https://github.com/tensorflow/models/tree/master/object_detection
- Jupyter notebook：https://github.com/tensorflow/models/blob/master/object_detection/object_detection_tutorial.ipynb
- Cloud ML：<https://cloud.google.com/blog/big-data/2017/06/training-an-object-detector-using-cloud-machine-learning-engine>

参考资料

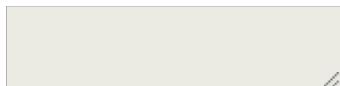
1. Speed/accuracy trade-offs for modern convolutional object detectors (<https://arxiv.org/abs/1611.10012>), Huang et al., CVPR 2017 (paper describing this framework)
2. Towards Accurate Multi-person Pose Estimation in the Wild (<https://arxiv.org/abs/1701.01779>), Papandreou et al., CVPR 2017
3. YouTube-BoundingBoxes: A Large High-Precision Human-Annotated Data Set for Object Detection in Video (<https://arxiv.org/abs/1702.00824>), Real et al., CVPR 2017 (see also our blog post (<https://research.googleblog.com/2017/02/advancing-research-on-video.html>))*
4. Beyond Skip Connections: Top-Down Modulation for Object Detection (<https://arxiv.org/abs/1612.06851>), Shrivastava et al., arXiv preprint arXiv:1612.06851, 2016

5. Spatially Adaptive Computation Time for Residual Networks (<https://arxiv.org/abs/1612.02297>), Figurnov et al., CVPR 2017
6. AVA: A Video Dataset of Spatio-temporally Localized Atomic Visual Actions (<https://arxiv.org/abs/1705.08421>), Gu et al., arXiv preprint arXiv:1705.08421, 2017
7. MobileNets: Efficient convolutional neural networks for mobile vision applications (<https://arxiv.org/abs/1704.04861>), Howard et al., arXiv preprint arXiv:1704.04861, 2017

原文地址：<http://opensource.googleblog.com/2017/06/supercharge-your-computer-vision-models.html>

声明：本文由机器之心编译出品，原文来自Google Open Source，转载请查看要求，机器之心对于违规侵权者保有法律追诉权。

[谷歌开源计算机视觉GitHub工程](#)



[提交评论](#)

登录后参与评论[去登录](#)



[关于我们寻求报道商务合作](#)

©2017版权所有 机器之心（北京）科技有限公司

京 ICP 备 12027496

全球人工智能信息服务

友情链接

[Synced Global](#)[机器之心](#) [Medium](#) [博客PaperWeekly](#)[网易智能动脉网](#)



联系电话：+86 010-57150141

联系邮箱：contact@jiqizhixin.com