

每日头条

[首页](#)[健康](#)[娱乐](#)[时尚](#)[游戏](#)[3C](#)[亲子](#)[文化](#)[历史](#)[动漫](#)[星座](#)[汽车](#)[家居](#)[情感](#)[f 分享引文](#)[太多新闻看不过来？搜索试试！](#)

谷歌通过TensorFlow Object Detection API开源其物体识别系统

2017-06-16 由 机器之心 发表于 科技

选自Google Open Source

作者：Jonathan Huang

机器之心编译

参与：黄小天、李泽南

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。

ADVERTISEMENT



其中一个模型在示例图片中（来自 COCO 数据集）识别对象的效果

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：

ADVERTISEMENT

- 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。

ADVERTISEMENT

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>

ADVERTISEMENT



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

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

相關文章



谷歌开源物体检测系统 **API**（附代码下载地址）

2017-06-16

雷锋网AI科技评论按：6.



开发**AI**前 先看看谷歌前机器学习大师的**9**条忠告吧

2016-05-05

Adam Berenzweig获得哥伦比亚大学博士学位后，在谷歌工作了超过十年的时间，建造了最早的谷歌音乐推荐系统，并参与了谷歌新闻、Goggles（Google的图像搜索应用）和实时搜索等项目工作。



李佳：人工智能的「**Another Badass Woman**」

2017-05-24

机器之心原创作者：微胖、吴攀李佳，谷歌云机器学习/人工智能研发负责人（Head of R&D, AI/ML, Senior Director at Cloud，Google）。

本科毕业于中国科大自动化系。新加坡南洋理工大学硕士。

李佳：人工智能的「**Another Badass Woman**」

2017-05-24



机器之心原创作者：微胖、吴攀李佳，谷歌云机器学习/人工智能研发负责人（Head of R&D, AI/ML, Senior Director at Cloud, Google）。本科毕业于中国科大自动化系。新加坡南洋理工大学硕士。



一文带你掌握谷歌的目标检测与识别API，附实际测试效果

2017-08-31

导读：深度学习做图像识别有很多不同的途径。谷歌最近发布了一个使用Tensorflow的物体识别API，让计算机视觉在各方面都更进一步。小编带大家一起去了解如何使用TF物体识别API。

深度 | 主流深度学习框架对比：看你最适合哪一款？

2016-09-13

选自deeplearning4j.



谷歌开源 TensorFlow 的图像文字说明功能

2016-09-23

在我们这一代人中，接近 100%的人都痴迷于 Instagram。遗憾的是，我早在 2015 年就离开了这个图片分享平台。原因其实很简单，我总在应该贴出哪些照片以及应该加哪些文字说明的事情上犹豫不决。



业界 | Google发布Open Images图像数据集，包含9百万标注图片

2016-10-01

选自Google Research机器之心编译参与：李亚洲最近，谷歌不断加大开源的力度。



Facebook 如何为计算机视觉技术寻求出路

2016-08-28

为提高计算机视觉能力，Facebook正为开发者社区越来越多地开源内部工具。计算机视觉隶属于人工智能（AI），可协助机器理解图像。计算机视觉不需要手工输入关键词或描述语等元数据，而是将图像分解并在逐个像素的基础上处理图像。

深度 | 从算法、硬件到研究社区，全面回顾2016年机器学习领域主要进展

2017-02-01

机器之心原创作者：微胖、吴攀人工智能圣杯虽然遥远，但这并不妨碍机器学习继续在2016年高歌猛进，其中，深度学习仍是最亮眼的明星。机器学习的重大进展离不开三个核心内容：算法（或软件）、硬件和数据。本文仅从算法（或软件）、硬件角度梳理2016年机器学习领域（主要是深度学习）主要进展。