Kaggle深度学习项目汇总

2018年7月24日

16:17

先挑选了一些看起来比较容易实现的，对硬件设备要求不是很高的来训练。

[Carvana Image Masking Challenge](https://www.kaggle.com/c/carvana-image-masking-challenge)

**问题目标**：自动识别图片中汽车的边界

**问题类型**：计算机视觉（CV）

**问题特点**：边界处理，模型提取。

有kernel,参赛队伍多（700+），可参考性高。

数据集大小：十万张图片。训练集不超过2G，测试集比较大十来个G。

[NOAA Fisheries Steller Sea Lion Population Count](https://www.kaggle.com/c/noaa-fisheries-steller-sea-lion-population-count)

**问题目标**：根据航拍图像内容，研究算法分析图片中有几个海狮

**问题类型**：计算机视觉（CV）；模式识别（Pattern Recognition）

**问题特点**：典型的计算机视觉和边界识别的综合问题。

有kernel，参赛队伍385，有一定的参考价值。

数据集接近100G，放弃吧。

[The Nature Conservancy Fisheries Monitoring](https://www.kaggle.com/c/the-nature-conservancy-fisheries-monitoring)

**问题目标**：根据视频内容，识别鱼的种类

**问题类型**： 计算机视觉（CV）；分类问题（Classification）；深度学习（DL）

**问题特点**：模式识别加上Deep Learning的内容。

有kernel，参赛队伍389，活跃度比海狮那个项目高。

数据集大小，不超过两个G，可以考虑。

[Dstl Satellite Imagery Feature Detection](https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection)

**问题目标**：从高空图像中准确辨别各种特征

**问题类型**：计算机视觉（CV）；深度学习（DL）

**问题特点**：根据卫星图像对地球上的物体进行识别。

有kernel，参赛队伍419，活跃度高。

数据集二十多G，比较大，不用服务器就只有放弃。

[State Farm Distracted Driver Detection](https://www.kaggle.com/c/state-farm-distracted-driver-detection)

**问题目标**：计算机视觉能否发现驾驶员分心？

**问题类型**：分类问题（Classification）；计算机视觉（CV）

**问题特点**：利用视频实时追踪驾驶员，判断驾驶员是否出现注意力不集中的现象。

有kernel，1440支队伍， 活跃度不高。

数据集大小4个G，可以考虑。

[Avito Duplicate Ads Detection](https://www.kaggle.com/c/avito-duplicate-ads-detection)

**问题目标**： 欺骗性、重复性的广告检测

**问题类型**： 分类问题（Classification）；计算机视觉（CV）

**问题特点**： 相似度分析。

有kernel，参赛队伍548，活跃度不高。

数据集也是四十多G，不用服务器可以放弃了。

[Right Whale Recognition](https://www.kaggle.com/c/noaa-right-whale-recognition)

**问题目标**：根据航拍照片，预测图片中是否有濒危动物脊美鲸(right whales)

**问题类型**：分类问题（Classification）；计算机视觉（CV）

**问题特点**：类似于物体检测的情况。

无kernel， 参赛队伍364，活跃度不高。

数据集大小8个G，还可以考虑。

[Galaxy Zoo - The Galaxy Challenge](https://www.kaggle.com/c/galaxy-zoo-the-galaxy-challenge)

**问题目标**： 对宇宙中的不同距离的星系的形态进行分类

**问题类型**： 分类问题（Classification）；计算机视觉（CV）

**问题特点**： 宇宙。

无kernel，参赛队伍326，活跃度低。

不超过两个G，还可以考虑。

[Multi-modal Gesture Recognition](https://www.kaggle.com/c/multi-modal-gesture-recognition)

**问题目标**： 多模态手势识别（2D/3D）

**问题类型**： 计算机视觉（CV）

**问题特点**： 卷积神经网络;RGB；降维。

无kernel，参赛队伍53，活跃度非常低。

数据集十个G以上，待定了。

[Challenges in Representation Learning: Facial Expression Recognition Challenge](https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge)

**问题目标**： 识别图像中人脸表情

**问题类型**： 计算机视觉（CV）；分类问题（Classification）

**问题特点**： 面部表情分类器；灰度图片。

无kernel，参赛队伍56，活跃度非常低。

好像没有足够的数据集。

[ICDAR2013 - Gender Prediction from Handwriting](https://www.kaggle.com/c/icdar2013-gender-prediction-from-handwriting)

**问题目标**： 预测手写笔记的作者是男还是女

**问题类型**： 计算机视觉（CV）；分类问题（Classification）

**问题特点**： 性别特征对字体形态的影响；神经网络

无kernel，参赛队伍190，活跃度非常低。

数据集大小合适，可以考虑。