

Deep learning for localization

姓名：张毅

学号：21721190

日期：2018-5

1 问题描述

在 CUBdataset 数据集上训练模型执行 localization 任务，框出小鸟位置。

数据集: CUBData (由助教提供)，按照要求, 随机选取 80% 作为训练数据, 其余 20% 作为测试数据。

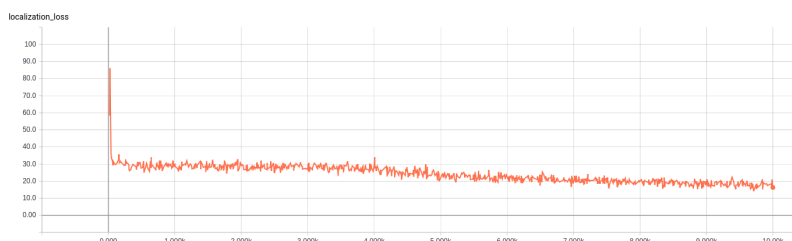
2 方法及原理

实验中采用了 vgg16 进行 fine-tuning (将在 ImageNet 上预训练好的 vgg16 最后一层替换为 4 个输出单元) 进行回归预测, 代码已经上传到 github 仓库 (yizhangzc/course), 本次实验代码放置在 localization 文件夹下, 运行方式及运行环境见 README.md。

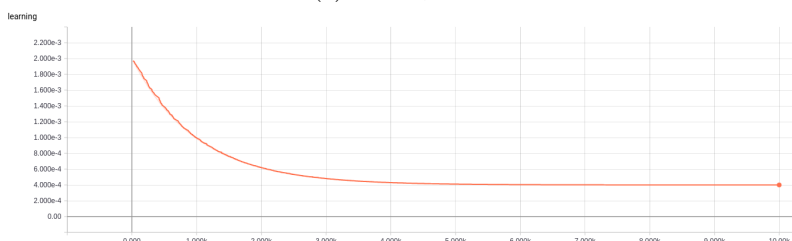
注: 课件示例使用了 ResNet18, 因为在网上没能找到训练好的 ResNet18 的 checkpoint 文件来进行 fine-tuning, 所以本次实验选择了 vgg16 进行。

3 实验结果

因为数据集较大, 训练很慢, 训练只迭代了 10000 次, 模型并未到达最佳效果, loss 仍然有下降趋势, 但是已经到达可接受范围。结果如下:



(a) loss 下降过程



(b) 学习速率下降过程

4 总结

(1) 损失函数为课件示例中给出的 SmoothL1 Loss (将对每个样本求和改为求均值), 训练采用了 Adam 优化器。(2) 实验中, 仍然采用了按指数规律下降的学习速率。(3) 使用 tensorboard 可以观测到 loss 的下降过程。(4) 图像输入模型前经过了预处理, resize 为 224x224, 并进行归一化。(5) 考虑到显存大小, 实验中 batch 大小取 50。