

猫狗大战 – 开题报告

项目背景

掌握解决具体计算机视觉任务的方法会帮助人类解决大规模系统的复杂问题。而近年来，卷积神经网络在图像分类、目标检测、图像语义分割等领域取得了一系列突破性的研究成果，其强大的特征学习与分类能力引起了广泛的关注，具有重要的分析与研究价值。

图像特征的提取与分类一直是计算机视觉领域的一个基础而重要的研究方向。卷积神经网络提供了一种端到端的学习模型，模型中的参数可以通过传统的梯度下降方法进行训练，经过训练的卷积神经网络能够学习到图像中的特征，并且完成对图像特征的提取和分类。

作为神经网络领域的一个重要研究分支，卷积神经网络的特点在于其每一层的特征都由上一层的局部区域通过共享权值的卷积核激励得到。这一特点使得卷积神经网络相比于其他神经网络方法更适合应用于图像特征的学习与表达。

问题描述

使用深度学习方法，用概率输出识别一张图片是猫还是狗

输入数据

数据集：[Dogs vs. Cats Redux: Kernels Edition](#)

训练 train 文件夹包含 25,000 张狗和猫的图像，根据每个图像的标签命名。

测试 test 文件夹包含 12,500 个图像，根据数字 ID 命名。

获取方式：从网站直接下载

选择此数据集的原因：样本文件给的，我也不知道为什么

解决办法

用卷积神经网络实现。具体方法是直接输入原始图像，用 CNN 网络对图片进行多次卷积层和池化层处理，在输出层给出两个节点并进行 softmax 计算得到两个类别各自的概率。

基准模型

使用基于 keras 的 resnet , Xception , VGG16 等网络模型去完成项目。因为这些模型在众多图像识别领域中表现出色，ResNet 最少可以达到 85%以上的准确率，而 Xception 最少可以达到 88%以上准确率。

评估指标

kaggle Public Leaderboard 前 10% , 即 131 位以前 , Log_loss 要小于 0.0167

项目设计

1. 数据集预处理
 - 对图像进行处理
 - 对数据进行分类
2. 模型建立
 - 建立 keras 模型
 - 具体怎么做，我不知道。。
3. 模型训练
 - 参数随机化
 - 用不同的优化器对模型进行调参，对比它们的性能
4. 模型评估
 - 用对数损失函数来评估模型
 - $\text{Log_loss} < 0.0167$
5. 数据可视化
 - 对模型优化过程中的数据进行可视化，便于直接观察和理论分析