**机器学习纳米学位开题报告**

**猫狗大战**

**刘珅珅**

**2019年3月29号**

1 背景

计算机视觉是让计算机能够感知和理解图像的意义，在现实生活中有广泛的应用。图像分类是计算机视觉的一个重要的研究方向，通过对不同类别图像的进行特征提取并构建模型来预测未知图像的类别，在很多实际问题上收到良好的效果，例如udacity课程中讲解的皮肤癌检测的例子，通过对各种类别的皮肤病图片的分类识别来判断患者是否患有皮肤癌，其准确率甚至超过了一些经验丰富的医生。kaggle推出dogs-vs-cats比赛，就是一个典型的图像分类问题，通过机器学习的方式来识别出图片是猫还是狗。选择这个项目作为毕业项目的原因是本人对计算机图像处理感兴趣，并且项目本身会用到深度学习和卷积神经网络的知识，这也是目前人工智能的热点。

2 问题陈述

该项目是一个监督学习的二分类问题，训练集通过图片名称已经打好了标签，区分是猫还是狗，通过机器学习构建模型并训练，然后预测测试集中的图片是猫还是狗。

3 数据集与输入

本项目的数据集为kaggle提供的数据集，训练集为25000张图片，其中12500张为猫类图片，12500张为狗类图片。测试集为12500张猫狗图片。考虑到交叉验证的情况，训练集中随机抽取5000张作为验证集，最终的数据集是训练集20000张图片，验证集5000张图片，测试集12500张图片。

4 解决方案

过去几年中，卷积神经网络在图像分类中有很好的效果，考虑到数据集较大并和ImageNet的图片集有很大的相似性以及Udacity课程中深度神经网络有关迁移学习的介绍，可以采用迁移学习的方式来解决问题。将问题分为两部分：特征提取和模型分类。特征提取是采用预训练模型提取实际图片的边缘特征、轮廓特征以及更高级的能区分猫狗的特征等。模型分类是利用之前提取的特征，输入到一个模型中进行分类训练和预测，并根据训练过程中的参数如准确率、损失大小等来判断是否满足要求来调整模型参数，最终得到符合要求的分类结果。

5 基准模型

本项目拟采用Keras中提供的5种预训练模型VGG16，VGG19，ResNet50，InceptionV3和Xception进行特征提取，然后构建一个统一的分类模型进行训练并预测，将预测结果提交到Kaggle上，选择效果最好的模型作为最终的模型。

6 评价指标

在kaggle竞赛中，该项目采用的评价指标为对数损失LogLoss，即对数损失。本项目采用与之相同的评价指标。毕业项目要求模型的LogLoss必须在kaggle pubic learnboard中排入前10%，即LogLoss<0.06127.

7 项目设计

1) 开通阿里云服务器，并在Linux系统安装Anaconda，Python3，Numpy，Pandas，Matplotlib，Tensorflow-gpu，Keras等。

2) 将训练集数据拆分为cat和dog两部分，并生成一个验证集。

3) 将图片转换为张量图像数据并进行预处理。

4) 采用预训练模型VGG16，VGG19，ResNet50，InceptionV3，Xception进行特征提取。

5) 构建分类模型并进行训练以及参数调优。

6) 对测试集进行预测，并将结果上传kaggle。如果效果不满足要求，则从4)或5)开始重复。