

# 计算机视觉项目开题报告

## 基于卷积神经网络的猫狗分类识别系统的设计与实现

### 项目背景

传统的图像识别多采用人工提取特征的方法，这种方法存在诸多缺点，例如，无法准确的提取特征，不能将图像良好的分类，每次的特征提取只能适用于一种图像，几乎没有针对性。因此急需一种方法能够从多样本学习中实现对图像的抽象，得到图像更本质的特征，避免人为提取特征的误差，进一步提高目标图像的识别率。并且，目标图像难免会出现模糊或不完整现象，目标的清晰度与完整性直接影响到目标图像的正确识别，如果对目标图像识别率提高，即使模糊或不完整的图像也容易做出较好的判断。

机器学习领域发生了很大的变化，特别是在深度学习和图像分析方面。利用卷积神经网络进行图像分类识别是深度学习关于图像处理的一个应用，卷积神经网络的优点是能够直接与图像像素进行卷积，从图像像素中提取图像特征，这种处理方式更加接近人类大脑视觉系统。另外，卷积神经网络的权值共享属性和池化层使网络需要训练的参数大大减小，简化了网络模型，提高了训练的效率。深度学习的优势在于它能够自动学习图像特征，在经过合理的训练后，模型具有极强的泛化能力。因此作为一种新的机器学习方法，深度学习的优势就格外明显。

### 问题描述

本项目需要使用TensorFlow框架搭建用于猫狗分类识别的卷积神经网络，通过实验分析，对网络进行调整。然后训练卷积神经网络，产生满足需要的模型。

### 输入数据

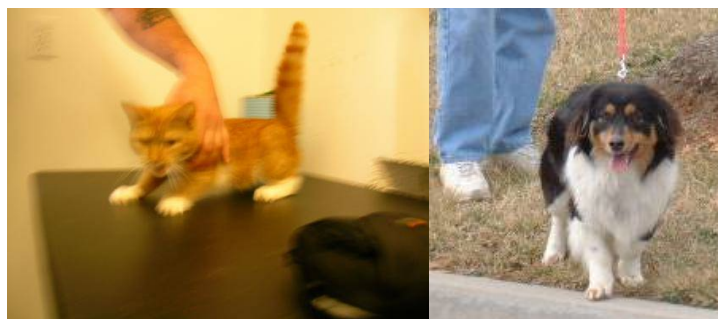


图1 Dogs vs. Cats数据集

本项目所用图片源于Kaggle中的Dogs vs. Cats。25,000张训练用图片，12,500张测试用图片，训练用图片中包含猫与狗的图片各12,500张。

由于图片是通过文件名来区分猫与狗的，而且图片大小不一，所以需要通过标签将训练用图片中的猫狗区分并调整成相同大小后封装，之后将封装好的图片集导入神经网络中训练，将测试用图片导入训练结束后生成的模型中。

## 解决方法

本项目的主要技术核心是搭建并训练神经网络与使用模型进行识别。所以，在本项目的设计过程中，首先需要做到收集图片数据，其次是搭建神经网络，使用流行的神经网络与调试网络超参数可以提高神经网络的训练效率与模型的准确率，之后是训练神经网络，将收集到的带标签图片预处理后导入神经网络进行训练，得到模型后测试模型，最后使用模型识别图片，使Loss 低于0.1。

## 评估标准

搭建一个合理的卷积神经网络并完成训练，生成的模型可以基本正确的识别照片中的猫与狗。

使用的Loss function 为：

$$L(\mathbf{w}) = \frac{1}{N} \sum_{n=1}^N H(p_n, q_n) = -\frac{1}{N} \sum_{n=1}^N \left[ y_n \log \hat{y}_n + (1 - y_n) \log(1 - \hat{y}_n) \right]$$

## 基准模型

本项目决定选用当前较为流行的卷积神经网络AlexNet，如图2所示，该网络将卷积神经网络的基本原理应用到了很深很宽的网络中的同时也使用到了很多新技术点，激活函数使用了ReLU来取代Sigmoid，使用最大池化层避免了模糊效果并提升了特征丰富性，训练时使用Dropout避免过拟合，提出了增强模型泛化能力的LRN层。

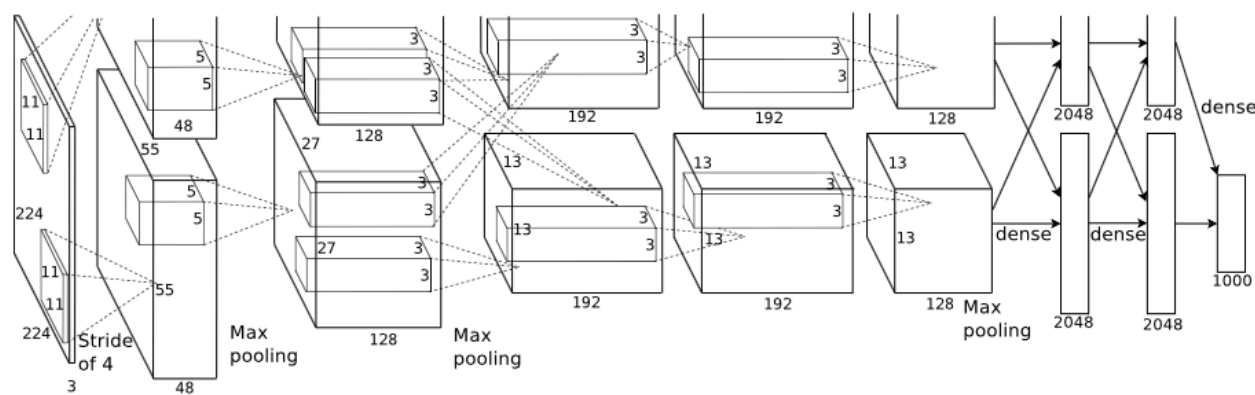


图2 AlexNet

不包括池化层和LRN层，整个AlexNet有8个需要训练的层，前五层为卷积层，后三层为全连接层。AlexNet最后一层做分类，LRN层在前两个卷积层后，ReLU函数应用在每一层的后面。

## 项目设计

首先，按照论文[2]中的描述，使用Python语言与TensorFlow框架搭建AlexNet。之后，将编写好的代码上传至AWS云服务器中使用，所需要的数据集直接通过服务器下载，整个训练过程均在服务器中实现。最后将训练结束后生成的模型从服务器中下载，编写后续用于识别的代码，使用OpenCV标示出图片中的猫或狗。

## 参考文献

- [1]Mitchell T M. 机器学习[M]. 机械工业出版社, 2003.
- [2].Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C] International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097–1105.
- [3].Géron A. Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems[J]. 2017.
- [4]Solem J E. Python 计算机视觉编程[J]. 2014.