

# 猫狗大战 – 开题报告

## 项目背景

神经网络的起源可追溯到 20 世纪 40 年代。1986 年，Rumelhart、Hinton 和 Williams 在《自然》杂志发表了著名的反向传播算法用于训练神经网络，从而引起巨大反响。

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。

2012 年，ImageNe 图像分类比赛的冠军是欣顿的研究小组，其采用的深度学习算法使深度学习在计算机视觉领域大放异彩。因为他们的准确率远超其之后的三组排名小组，这些采用传统的计算机视觉方法，准确率相差不到 1%。从而引起了深度学习的热潮。从此以后，每年的 ImageNet 图像分类比赛都是神经网络夺得冠军。比如 2012 年 AlexNet 使用 8 层神经网络，错误率 16.4%；2014 年 InceptionNet 使用 22 层神经网络，错误率 6.7%；2015 年 ResNet 使用 152 层神经网络，错误率 3.57%。

## 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。这是一个二分类问题，项目要求使用深度学习方法，通过训练模型，任意输入一张图片，模型能输出图片是猫或者是狗的概率。

## 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。

选择此数据集的原因：训练数据共有 25000 张图片，猫狗各占一半，分类数量相同，每张图片都带有类别标签。

获取方式：从网站直接下载。对彩色图片进行 resize，以便大小一致的输入到神经网络中。本项目采用 keras 的 ImageDataGenerator.flow\_from\_directory 的参数 target\_size 设置大小。

## 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。

kaggle Public Leaderboard 前 10%，即 131 位以前，我的最终结果要小于 0.06127

## 评估指标

采用对数损失函数来衡量

$$\text{logLoss} = -\frac{1}{n} \sum_{i=1}^n [y_i \log(y'_i) + (1 - y_i) \log(1 - y'_i)]$$

其中，n 是图片数量， $y'_i$  是模型预测为狗的概率， $y_i$  是类别标签，1 对应狗，0 对应猫

## 项目设计

### 1. 数据集预处理

从 kaggle 下载好图片，将猫和狗分类放在不同的文件夹，为 keras.ImageDataGenerator 准备好数据。

对图片进行 resize，保持输入图片信息大小一致。

对训练数据进行随机偏移、转动等图像变换处理，让训练数据多样化。

### 2. 模型建立

利用 keras 的 application 模块建立 keras 模型：

使用 ResNet50 等现有的去掉了全连接层训练模型；添加自己的全连接层到 ResNet50 网络。

### 3. 模型训练

导入预训练的网络权重；

冻结除了全连接层的所有层，获得 bottleneck 特征；

用不同的优化器对模型进行训练，对比它们的性能，选择最佳模型。

### 4. 模型评估

用对数损失函数来评估模型

### 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：

可视化原始数据；可视化模型训练过程的准确率曲线、损失函数曲线等

## 参考文献

- [1] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks
- [2] <http://keras-cn.readthedocs.io/en/latest/>
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition
- [4] Karen Simonyan, Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition
- [5] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich. Going Deeper with Convolutions
- [6] Diederik P. Kingma, Jimmy Ba. Adam: A Method for Stochastic Optimization

[7] Matthew D. Zeiler. ADADELTA: An Adaptive Learning Rate Method