

# 机器学习纳米学位—猫狗大战

## 开题报告

---

沈平岗

2019.02.18

### 项目背景

Dog vs. Cats (猫狗大战) 是 kaggle 在两年 2017 年举办的一场有趣的比赛。Kaggle 一共举行过两次猫狗大战图像分析比赛，第一次是在 2013 年，从那时候起，机器学习领域发生了很大的变化，特别是在深度学习领域和图像分析方面。如今 Dog vs. Cats 分类问题在也是一个热门话题。虽然现代的技术已经解决了这个问题。但是在数据集上使用新的技术，也是对未来的一种挑战。

项目参赛者需要构建一个模型，并对模型进行训练。达到输入一张图片，输出该图片的类别和对应的概率。其中概率越接近 1, 表示图片是 “狗” 的可能性越大，反之概率越接近 0，图片是 “猫” 的可能性越大。

### 问题描述

根据 kaggle 上提供的项目信息，Dog vs. Cats (猫狗大战) 实际是二分类问题，即给定一张图片，输出图像的类别。这在计算机视觉中，使用神经网络无疑是最佳选择。特别的，针对图像数据类型，需要用到卷积神经网络 (Convolutional Neural Networks)。在卷积神经网络中有多种优秀的网络模型框架，如 resnet、vgg、inception、xception 等。然而提高准确率，降低 loss 的一种方法就是使用迁移学习，即使用提供的卷积神经网络模型进行特征向量提取，训练。

### 数据集和输入

数据来自 kaggle: [Dogs vs. Cats Redux: Kernels Edition](#)

下载的 kaggle 数据集中解压后包含 3 个文件: train.zip、test.zip、sample\_submission.csv。

- train.zip 为训练数据集。包含了猫和狗的训练数据集合，总共 25,000 图片。其中猫和狗图片各有 12500 张。

- test.zip 为测试数据，没有标定图片是猫、还是狗。以数字命名，从 1 开始，总共 12,500 图片。

- sample\_submission.csv 为提交样本。需要将测试数据集的结果导入到样本中，然后提交给 kaggle、其中预测为狗则为 1、猫为 0。
- 提供的图片数据集中，部分图片存在模糊不清难以分清。部分图片标注错误。图片中不含有“猫”或“狗”。模糊不清、分类错误的图片占少数比例，约为 %1 左右。
- 提供的图片尺寸大小各不相同。有少部分图片尺寸相对比较大。

## 解决办法

项目采用卷积神经网络解决问题。主要通过迁移学习的方式，提前特征向量，构建神经网络，将提取的特征向量进行训练。以此提高准确率和降低 loss。在优秀的卷积神经网络模型中，使用 resnet50 以及其他模型进行特征提取，并对提取的特征向量进行保存，然后构建全连接网络模型，使用 sigmoid 激活函数输出类别概率。

## 基准模型

本项目的最低要求是 kaggle Public Leaderboard 前 10%。在 kaggle 上，总共有 1314 只队伍参加了比赛，所以需要最终的结果排在 131 位之前，131 位的得分是 0.06127，所以目标是模型预测结果要小于 0.06127。

## 评估指标

交叉熵（cross entropy）是深度学习中常用的一个概念，一般用来求目标与预测值之间的差距。是分类问题中广泛使用的一种损失函数。该项目是一个二分类问题，因此可以采用 logloss 损失函数作为评价指标。其公式如下：

$$LogLoss = -\frac{1}{n} \sum_{n=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

- n 数据集中的图像数
- $\hat{y}_i$  图像是狗的预测概率
- $y_i$  如果图像是狗，则为 1

## 项目设计

## 1. 数据预处理

- 从 kaggle 上下载好图片。
- 采用软连接的方式对图片类别进行区分。划分数据集为测试数据集和训练数据集。验证数据可在训练模型时用 fit 进行比例指定。
- 设置图片增强和做归一化处理。使用 keras 提供的 ImageDataGenerator 模型进行操作。训练数据集进行图片增强和归一化处理，测试数据集进行归一化处理。

## 2. 提取特征向量

- 模型使用 resnet50 获取特征向量。在 keras 中提供预训练模型。模型中有 1000 个类别，其中包含“猫”和“狗”。
- 保存特征向量，并在需要时候获取。

## 3. 搭建模型 & 训练模型

- 构建神经网络模型。使用 sigmoid 激活函数输出概率，优化器为 rmsprop、binary\_crossentropy 损失函数。
- 加载特征向量。
- 训练模型。将特征向量输入到神经网络模型，设定 epochs，采用 callback 方式并保存最佳模型，设置验证数据集为训练数据集的 20%比例。

## 4. 提高模型精度

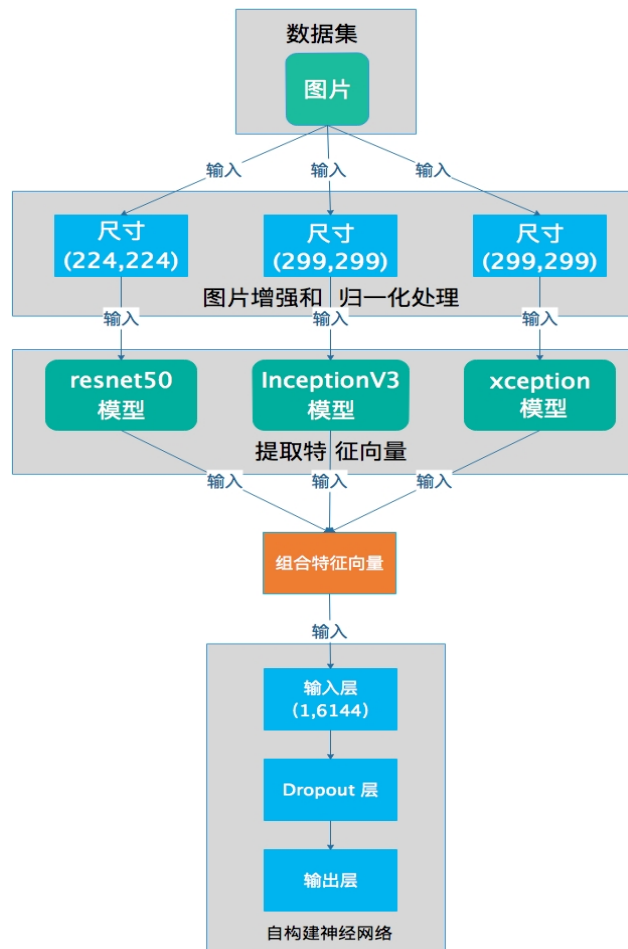
- 设法提高模型的准确率，降低损失。将使用 inceptionV3，xception 模型提取特征向量，加上之前 resnet50 提取的特征向量，一共三个特征向量。

（需要注意，inceptionV3 和 xception 模型要求图片的尺寸为 299 \* 299，所以在操作之前需要重新对图片进行预处理）

- 将 resnet50，inceptionV3，xception 三个模型提取的特征向量进行组合。重新构建网络模型进行训练。使用 sigmoid 激活函数输出概率，优化器为 rmsprop、binary\_crossentropy 损失函数。

- 训练新模型。将组合特征向量输入到神经网络模型，设定 epochs，采用 callback 方式并保存最佳模型，设置验证数据集为训练数据集的 20%比例。

整个执行流程如下图：



## 5. 模型评估

- 使用训练的模型对测试数据集进行预测。保存预测结果。
- 将预测结果上传到 kaggle 获得得分。

## 6. 可视化

- 进行数据探索并可视化。
- 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] Kaiming He & Xiangyu Zhang & Shaoqing Ren & Jian Sun. *Deep Residual Learning for Image Recognition* [C], cornell university, Submitted on 10 Dec 2015
- [2] *Convolutional neural network*[J], From Wikipedia, the free encyclopedia
- [3] 凌蓝风,*Dogs vs Cats For Udacity (异常值检验)*[J], 知乎,2018-02-27
- [4] 郭耀华, *深度学习-优化器算法详解*[J], cnblogs, 2018-03-10
- [5] Sandipan Dey , *Dogs vs. Cats: Image Classification with Deep Learning using TensorFlow in Python*[J], Data Science Central, 2017-8-14
- [6] Jeff Delaney,*CatdogNet - Keras Convnet Starter*[J],kaggle
- [7] Shivam Bansal,*CNN Architectures : VGG, ResNet, Inception + TL*[J],kaggle
- [8] Ashish Jagadish,*Dogs vs Cats: Keras Solution*[J],kaggle
- [9] MrMiaow, *Kaggle 猫狗大战准确率 Top 2%webapp 部署*[J],简书,2018.10.11
- [10] *Keras: 基于python 的深度学习*[M]
- [11] shizhengxin123, *利用 resnet 做 kaggle 猫狗大战图像识别, 秒上 98 准确率*[J], csdn, 2017-05-18