

# 猫狗大战：从图片中区分出猫或者狗

## 领域背景

计算机视觉在我们的日常生活中随处可见且变得越来越重要，如自动驾驶、OCR、图片分类等，这些应用的关键部分就是对于图像的侦测与分类。神经网络算法在图像的侦测与分类起到了巨大的作用，从最早的 LeNet、AlexNet 模型再到 VGGNet、ResNet 图像的识别准确率在不断的提升，最近的 Xception 模型更是在 ImageNet 的 Top-5 accuracy 中达到了 0.945。

本项目旨在依托前人创建的优秀模型，来构建一个可以对图片中猫狗分类的程序，项目中使用的数据集来自于 [kaggle dogs vs cats 比赛](#)。

## 问题阐述

从图片中侦测出猫狗并对其分类，可知此问题属于计算机视觉领域，且是识别出两类物体，因此又属于分类问题中的二分类问题。在选择的数据集中包含猫和够两种图片，每种图片都是被标记过得，图片要么是狗、要么是猫。我们的目标就是从给定的标记数据集中构建一个分类器，能够对输入的图片进行预测并分类标记。

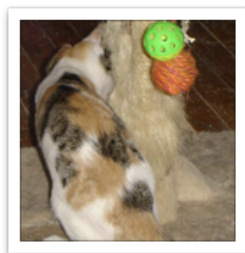
项目中使用开源计算机视觉库 OpenCV 来对图片进行处理；开源深度学习框架 TensorFlow 及 Keras 进行神经网络的构建。使用训练数据集对模型进行训练，训练好之后在测试数据集上进行测试。

## 数据集

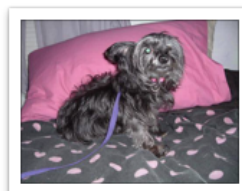
数据集分为两个部分：训练部分和测试部分。训练部分的数据集包含有 25000 张猫和狗的照片，其中猫和狗各有 12500 张图片，每张图片中含有猫或者狗，图片的标签即是文件的名字，如“cat.11.jpg”就表明这是一张猫的图片，图片的编号是 11（编号从 0 开始），如下图：



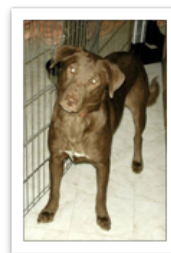
cat.12498.jpg



cat.12499.jpg

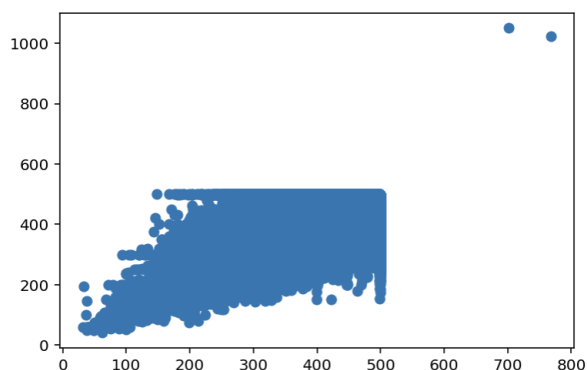


dog.0.jpg



dog.1.jpg

训练集中图片大小的分布如下：



测试部分的数据集包含 12500 张猫和狗的照片，这些没有被标记的图片，以数字编号作为文件名，这些数据用于模型对图片中含有狗的概率预测。

## 解决方案

在此项目中，主要使用卷积神经网络来对图片中的猫和狗进行识别。

## 基准模型

此处以 kaggle dogs vs cats 比赛中 leaderboard 成绩为标准，要求名次在前 10%，也就是以

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

公式计算的得分必须在 0.06 以内。其中  $n$  表示测试数据集图片的数量， $\hat{y}$  表示预测图片为狗的概率， $y_i$  表示如何图片是狗则为 1 是猫则为 0， $\log()$  表示以  $e$  为底数的自然对数。

## 评估标准

因为测试数据集并没有标签，因此使用准确率进行评估并不可行，所以使用 kaggle 官方的评估标准（即基准模型中提到的公式）。

## 项目设计

### 1. 数据处理

通过 opencv 将使用的数据集处理成可以供模型训练的数据集，如将图片的名字作为标签与图片对应起来，然后对图片进行裁剪、缩放、归一化等处理，之后将图片与标签的顺序打乱，让图片更具有随机性，最后将训练数据分割成训练数据集、验证集和测试数据集。

### 2. 模型选择

此处选择 xception[1] 模型

### 3.模型训练

将数据输入到模型中，训练的过程中不断调整参数，直到模型准确率达到标准

### 4.模型测试

将训练好的模型用在训练集上进行测试

## 引用

[1] <https://www.kaggle.com/c/dogs-vs-cats-redux-kernels-edition#evaluation>

[2] Francois Chollet . Xception: Deep Learning with Depthwise Separable Convolutions .