

机器学习纳米学位

沈龙飞

Dog VS Cats 开题报告

2018年2月15日



项目背景

选择的课题来自Kaggle上的一个竞赛题目，目标是训练一个模型，在给定的图像中分辨猫和狗。属于图像分类问题，是机器视觉中的一类问题。

视觉是人类感知外部世界的重要手段，而图像也是人类获取信息的重要途径，因此和视觉紧密相关的数字图像处理技术越来越多的受到关注，逐渐形成图像识别技术。

图像识别技术是利用计算机采集到的图像数据为基础，让机器模仿人类视觉，自动完成某些信息的处理功能，达到人类所具有的对视觉采集图像进行识别的能力，以代替人去完成图像分类及辨别的任务。

问题描述

Kaggle 中的Dog vs Cats 是从真实世界中采集的包含猫和狗的图片，图片本身存在很大的差异性，比如说拍摄设备不同造成的分辨率不同，拍摄角度照成的猫或者狗的主体不完整，猫或者狗的观察角度不同导致猫狗轮廓之间存在很大差异，真实场景中的猫和狗也处于不同的动作状态，猫狗本身的体积大小，让我们无法直接使用标准的轮廓去直接分辨图片中是否是猫或者狗。简而言之就是无法使用简单的规则直接判定是否是猫狗，而是需要数学模型对其中的猫狗特征进行提取和分类才能在不同场景下作出准确的判断。

输入数据

竞赛所需要的数据集可以在Kaggle的竞赛页面里面下载得到，一共包含两个zip文件，分别train.zip 和 test.zip，train.zip 是用来训练模型，test.zip 则是通过训练出来的模型进行预测，最后将预测结果提交到竞赛页面中，得出评分。

Dog vs Cats 数据集中的图片经过初步观察，训练集总数有25000张，其中猫有12500张，狗也是12500张。两者数量相等，在训练过程中则不需要针对图片数量给予不同的权重。

通过对下载的图片进行简单的预览，发现图片中的噪音较大，猫狗主体轮廓差异很大，同时轮廓中包含其他物体带来的异常情况。预计在训练过程中需要使用尽可能多的样本进行特征的提取。

由于图像质量的参差不齐，在输入图像时，需要对图片进行预处理，首先将图片resize为特定尺寸，在对图片进行中心裁剪，获取相同大小的输入图像。同时由于图片是彩色的，所以使用对应的RGB三通道转化成的三维矩阵作为模型的输入数据。

解决方法

在图像处理领域，卷积神经网络的训练效果非常好，初步选择使用卷积神经网络来训练对图片的识别。通过对图片使用滤波器进行卷积，然后使用最大池化提取有效特征。进行两次卷积和池化后，对结果进行评判，然后调整滤波器，和池化方法，以获得较好的结果，同时通过对比选择较好optimizer 降低训练时间。

基准模型

在本次竞赛中，由于需要获得kaggle比赛前10%的成绩，在kaggle竞赛页面中，截止目前总共参赛人员为1314名提交了成绩。前10%意味这需要分数超过第131名，在页面中可以看到第131名的得分为0.06127。

基准模型得分需要小于 0.06127.

评估指标

在kaggle上，官方提供了一个模型预测得分公式，通过对测试集的预测结果执行LogLoss 公式得出分数，以评估模型训练成效以达成对模型的性能评估。

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

公式中： n 表示测试集的数量

\hat{y}_i 表示图片预测为狗的概率

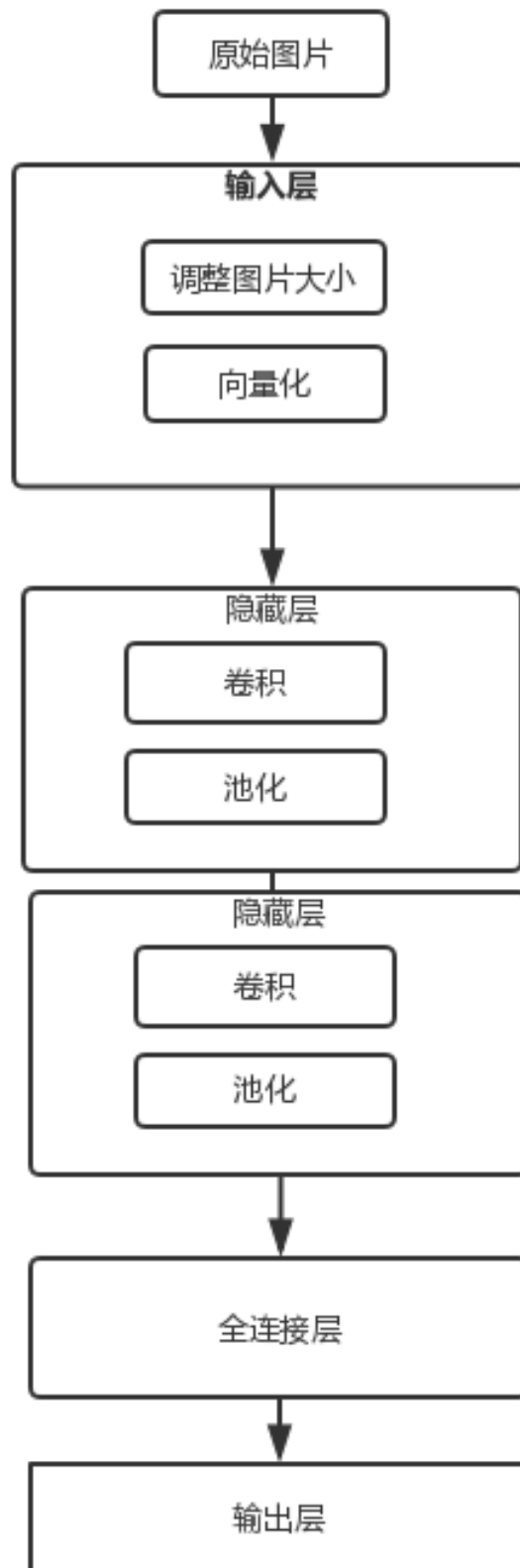
y_i 为1 表示为狗，为0 表示为猫

$\log()$ 表示以自然对数

设计大纲

本题主要使用卷积神经网络，在卷积神经网络中，主要关注点是卷积，池化，全连接同时要由于图像数据转化为矩阵之后可能会引起输入矩阵比预想的要巨大，在项目设计过程中要仔细考虑不同环节的影响。

神经网络的概念流程图如下：



如概念图中所示，在卷积神经网络中，会对结果产生影响的环节包括如下：

1. 数据预处理：

首先要对图片进行resize 并裁剪，图片裁剪用于减少输入图片大小，设置的图片大小不一致表示输入图像包含的信息数量也是不一致的，同时增大输入图像数量过程中也会额外增加一定的噪音数据进入。理论上输入数据越多训练效果应该更好。

2. 模型搭建

整个模型如图所示，隐藏层中包含两次卷积和池化，尝试使用keras搭建模型。

3.模型训练

依据输入数据在池化之后的规模，可能会选择不同的优化器进行对比训练效率和训练结果。使用adam 优化器进行训练。

4.模型调参

在卷积网络中，可供调整的参数包括：

- (1) 滤波器
- (2) 滤波器和池化的次数
- (3) Optimizer 选取，控制训练时长

5.模型评估

模型评估依旧采用竞赛提供的LogLoss 计算模型得分。

6.可视化

可视化部分使用matplotlib，编写辅助函数实现。

参考文献：

- [1] Alex Krizhevsky , Ilya Sutskever , Geoffrey E. Hinton . ImageNet Classification with Deep Convolutional Neural Networks
- [2] Jianxin Wu . Introduction to Convolutional Neural Networks
- [3] Diederik P. Kingma, Jimmy Lei Ba* . ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION