

机器学习纳米学位-开题报告

王烨

2018年8月30日

项目背景

猫狗大战 (Dog vs Cat) 是Kaggle举办的一个娱乐型的竞技比赛。在这个比赛中需要编写计算机算法来实现分类图像中是猫还是狗。Kaggle举办这个比赛的原因是，原来在很多服务器网站上为了防止机器人注册而使用了基于动物识别的验证码，这个验证是利用人能够很容易的区分猫狗，而机器不能，但是在人工智能发展之后，机器对猫狗的识别率已经大大提高了。本比赛是为了挑战计算机对猫狗的识别率。本项目是使用卷积神经网络 (Convolutional Neural Network, CNN)，来对图像进行识别。

项目选择的数据集是 Kaggle 竞赛提供的数据，训练集包括 12500 张被标记为猫的图片 和 12500 张被标记为狗的图片，测试集包括 12500 张未标记图片。对于每一张测试集中的图 像，模型需要预测出是狗图像的概率（1 代表狗，0 代表猫）。

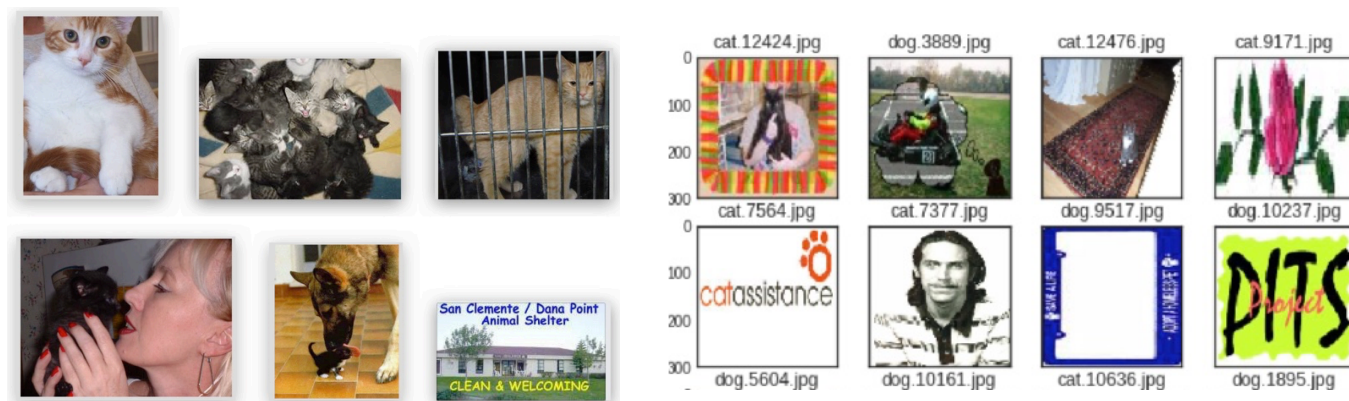
问题描述

1. 首先，这是一个分类问题。
2. 项目提供的数据是来源某收留流浪猫狗网站的真实的照片，背景复杂，分辨率不同，还有一些异常图片混在其中。

数据集分析

项目选择的数据集是 Kaggle 竞赛提供的数据，训练集包括 12500 张被标记为猫的图片 和 12500 张被标记为狗的图片，测试集包括 12500 张未标记图片。对于每一张测试集中的图 像，模型需要预测出是狗图像的概率（1 代表狗，0 代表猫）。

首先，简单观察一下训练数据有哪些类型，照片来自真实的场景，可以看到场景丰富。有一只猫，关在笼子后的猫，多只猫，人和猫，猫和狗 同时现出现的；还有少量的异常值。



解决方案

基于前面的数据探索与分析可知，猫狗识别问题其实是一个分类问题。输入一张图片得到这张图片是猫还是狗的概率。分类算法就我们知道的有 决策树、支持向量机、逻辑回归、贝叶斯、神经网络 等方法。基于经验，神经网络 尤其适合图像的识别。

基准模型

本次使用TensorFlow和 Keras ¹

Model	Top-1 Accuracy	Top-5 Accuracy	Parameters	Depth
Xception	0.79	0.945	22,910,480	126
VGG16	0.715	0.901	138,357,544	23
VGG19	0.727	0.91	143,667,240	26
ResNet50	0.759	0.929	25,636,712	168
InceptionV3	0.788	0.944	23,851,784	159
InceptionResNetV2	0.804	0.953	55,873,736	572
MobileNet	0.665	0.871	4,253,864	88
DenseNet121	0.745	0.918	8,062,504	121
DenseNet169	0.759	0.928	14,307,880	169
DenseNet201	0.77	0.933	20,242,984	201

选择基准模型为Xception 和 InceptionV3

评估指标

$$LogLoss = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad ^2$$

- n 是测试集中的图像数量
- \hat{y}_i 是图像是狗的预测概率
- y_i 图像如果是狗为1，如果是猫为 0
- $\log()$ 是以E为底的自然对数

项目设计

1. 下载数据集 [Kaggle Data](#)
2. 数据预处理
 - 2.1 将训练集分为两个目录
 - 2.2 利用预训练模型剔除异常值
3. 利用Keras的预训练模型提取特征向量并保存
4. 载入第3步中保存的特征向量
5. 构建编译模型
6. 训练模型
7. 预测测试集
8. Dropout参数优化

参考

¹ <https://keras.io/applications/#applications>

² <https://www.kaggle.com/c/dogs-vs-cats-redux-kernels-edition#evaluation>