

优达学城机器学习（进阶）毕业项目开题报告

Kaggle: Dog Breed Identification

1. 项目背景

本项目是 Kaggle 近期上线的一个 Playground 项目。训练数据来自于 ImageNet 关于狗品种的子集。通过这个项目可以让参与者了解图像识别算法，找到自己正在开发或者学习的算法中的问题。

本项目属于图像识别领域新出现的一个研究方向，细粒度视觉分类(Fine-Grained Visual Categorization, FGVC)，其目的是将某一大类事物的子类进行分类。比如，本项目中将狗这一大类事物，分出 80 个子类（80 个狗的品种）。

2. 问题描述

本项目的具体问题是将 10357 张狗的图片分成对应的 120 个品种。

3. 输入数据

本项目的输入数据包括三个部分：

- 1) Keras 预训练模型的权重，计划使用 Xception、VGG16、VGG19、ResNet50 和 InceptionV3 这 5 个模型。
- 2) 训练数据包括，10222 张已标记的彩色图片。
- 3) 测试数据包括，10357 张未标记的彩色图片，这些图片对应的分类的结果将作为输出，提交 Kaggle，验证最终算法的正确率。

4. 解决办法

针对本项目使用已经给出预训练权重的特点，主要将会使用 Xception、VGG16、VGG19、ResNet50 和 InceptionV3 这 5 个模型提取特征。然后，将这些特征输入一个多层神经网络训练，得到模型。

5. 基准模型

将使用 KNN 作为基准模型。将三维图像数据展开成一维数据，直接输入模型训练，每个图像输出对应 80 个分类的可能性。

KNN 代表统计学习方法，本项目中使用预训练权重的迁移学习做对比。

6. 评估指标

评价指标，选用 Kaggle 官方的 Multi Class Log Loss。

7. 设计大纲

本项目的解决方案将主要将会使用 Xception、VGG16、VGG19、ResNet50 和 InceptionV3 这 5 个模型提取特征。然后，将这些特征输入一个多层神经网络训练，得到模型。

引用：

- [1] <https://www.kaggle.com/c/dog-breed-identification>
- [2] <https://www.kaggle.com/gaborfodor/use-pretrained-keras-models-lb-0-3>
- [3] https://github.com/ypwhs/dogs_vs_cats
- [4] Xception: Deep Learning with Depthwise Separable Convolutions, , <https://arxiv.org/abs/1610.02357>
- [5] Very Deep Convolutional Networks for Large-Scale Image Recognition , <https://arxiv.org/abs/1409.1556>
- [6] Very Deep Convolutional Networks for Large-Scale Image Recognition , <https://arxiv.org/abs/1409.1556>
- [7] Deep Residual Learning for Image Recognition, <https://arxiv.org/abs/1512.03385>
- [8] Rethinking the Inception Architecture for Computer Vision, <https://arxiv.org/abs/1512.00567>
- [9] Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning , <https://arxiv.org/abs/1602.07261>
- [10] MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications , <https://arxiv.org/pdf/1704.04861.pdf>
- [11] <https://keras.io/>
- [12] <http://www.ulunwen.com/archives/111056>