图像分类与自监督

章剑鸿 2019101412

1. **数据集**

用python随机爬取谷歌图片50类，每类300张，合计15000张，共1.2G。每类按照 训练集/验证集/测试集 各划分成200/50/50张，图片大小为84\*84。

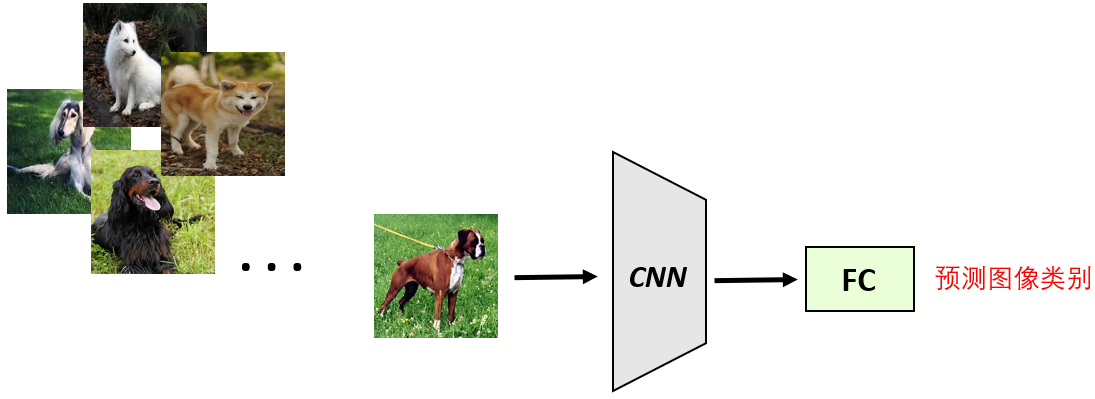
1. **模型**

采用的特征提取器是基础的四个block模型：conv2d(3, 64, 3)，conv2d(64, 64, 3)，conv2d(64, 64, 3)，conv2d(64, 512, 3)，连接一个最大池化层,maxpool2d(5)，其输出为512维的图像特征。对于分类任务，特征提取器后会连接一层全连接层映射到分类概率（图像分类概率或角度预测概率）。

1. **方法**

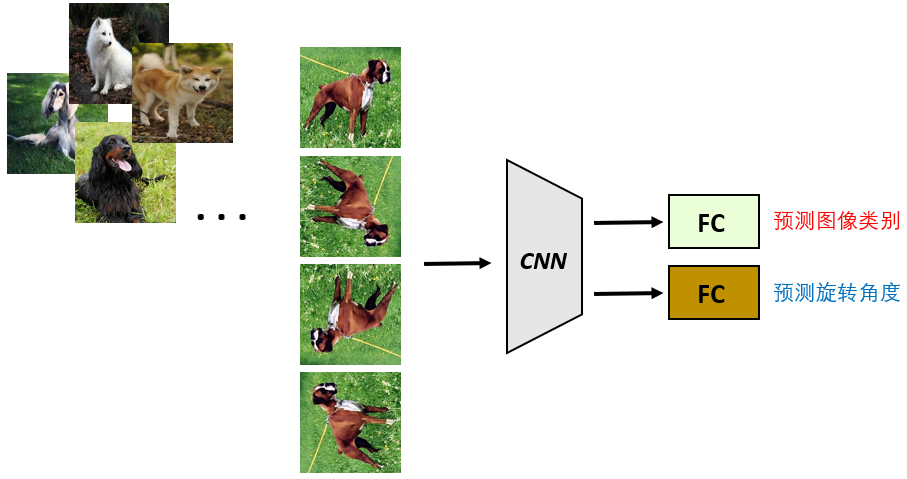
一共尝试了三种方法：一种原始方法，两种结合自监督的方法。

1. **原始方法**



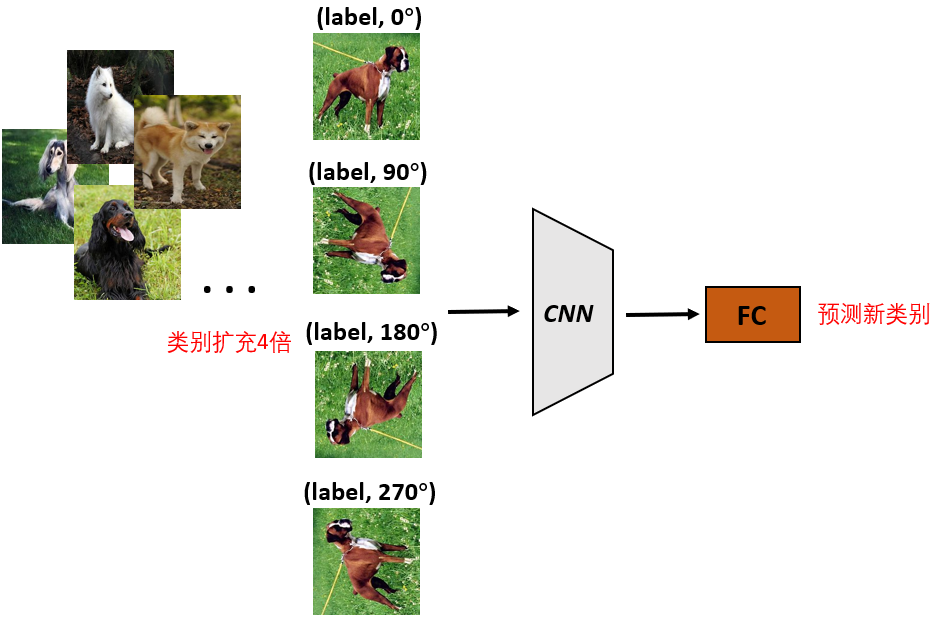
直接将图片输入CNN，再接FC预测图像类别，loss为单一的分类Cross\_Entropy Loss。

1. **自监督（多任务）**



先将图片做旋转操作（0，90，180或270度），再输入CNN，此时CNN后接两个FC，一个预测图像所属种类，一个预测图像旋转角度（4分类），以多任务方式训练。

1. **自监督（单任务）**



同将图片做旋转操作（0，90，180或270度），并与图片原始类别组合成新的四种类别，再输入CNN和FC来预测图像新的类别。

1. **结果与分析**

经实验，三种方法的准确率分别为45.26%，48.46%和50.02%。方法二通过旋转图片来引入自监督信息，改善模型分类效果；但多任务训练存在不稳定的情况，故方法三通过组合原始类别和旋转角度获得扩充的新类别，以单任务的方法训练模型，使结果得到进一步提升。