

天气图像分类算法报告

无人艇团队

团队成员：杨扬、沈纪元、祝嘉诚、董开元、戚文栋、蔡卓悦

指导老师：沈俊

1. 赛题分析

决赛题目为天气图像的分类问题，类别共有 3 类，分别是晴天、阴天和其他，其中其他天气，经过我们的观察，主要包括下雨、下雪、雷电、雾霾等天气。训练集包括 179584 张图片，测试集有 72778 张图片，数据量是比较大的。图 1 展示了部分训练集图像，可看出，图片整体大小不一，图片大致包括天空、地面、建筑、自然景物、人、交通工具等等，场景各异。

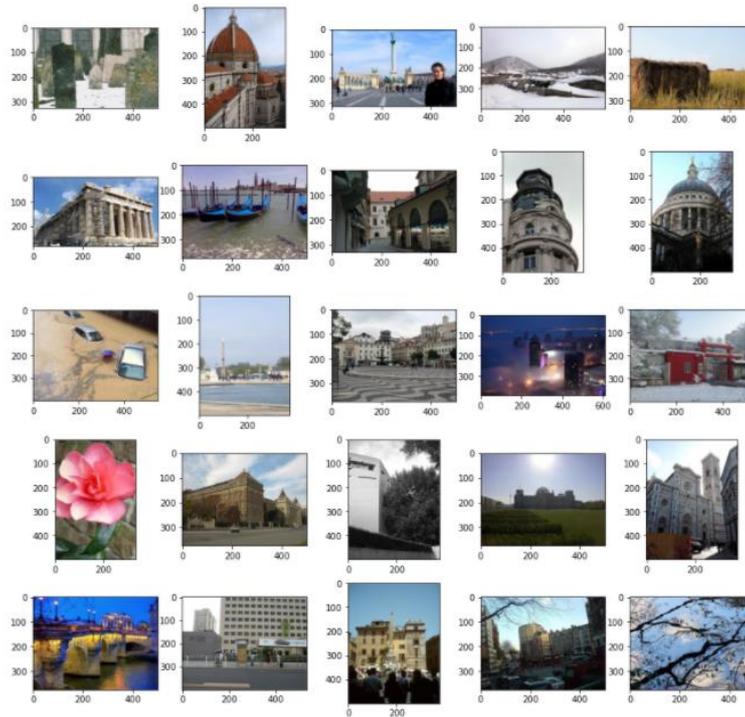


图 1 训练集图片

原始训练集包括四个部分，分别为 Part1、Part2、Part3、Part4，他们这样区分的原因是数据的来源不同，可能存在标注不一致等现象。如图 2 所示，我们统计了一下，Part1 包括晴天、阴天和雨天，分别有 56207 张、38614 张、74763 张，Part2-4 都只有晴天和阴天两个类别，总共有 6000、2000、2000 张图片，晴天阴天各占一半。考虑到数据来源不一致的情况，我们优先考虑使用 Part1 的数据，看情况再决定是否加入 Part2-4 的数据。总体上来说，晴天、阴天、其他这三类数据的数量还算是比较均衡的。

```
part1 sunny = 56207
part1 cloudy = 38614
part1 others = 74763
part2 sunny = 3000
part2 cloudy = 3000
part3 sunny = 1000
part3 cloudy = 1000
part4 sunny = 1000
part4 cloudy = 1000
```

图 2 训练集数据分布情况

天气不是一个具体的东西，可能和图片中的天空颜色、云朵、地面、光照、阴影、亮度等特征有关系。这些特征不一定都存在，有的图片里看不到天空，甚至只有一朵花、只有一栋建筑或者是只有一个人。卷积神经网络具有强大的特征提取能力，我们由此考虑主要使用卷积神经网络对天气图片进行分类。

我们认为该题目主要需要考虑以下方面：首先是图像大小不同如何处理，可以考虑直接缩放到固定大小，或者按比例缩放再随即裁剪、甚至是不做处理等等。其次就是数据增强，可以做的尝试包括水平/垂直翻转，随机旋转，亮度、饱和度、对比度的变化，色彩空间变化，还有一些高阶的图像变换等。然后就是模型方面，我们考虑一些目前比较主流的，在 ImageNet 上取得较高准确率的模型，也可以针对天气图像的特点对模型做一些改进。调参也是提升模型准确率的方法，包括学习率策略、优化器选择、正则化、损失函数选择等等。还有一个很重要的点在于训练集和验证集的划分，要避免数据泄露问题，并且要在验证集上较为精确的反映出模型的真实水平、泛化能力以及实用性等等。最后，可以通过模型集成的方式来提升模型的预测准确性，常见的方法有 bagging、boosting、stacking、blending 等。

2. 解题过程

这一部分，我们根据时间进程来梳理一下我们在解题过程中所做的一系列尝试。

在训练集的预处理上，我们首先计算了训练集的均值和方差，用于图像的标准化。然后，对比了两种图像缩放的方法。第一种是直接将图像 Resize 到指定尺寸(比如 256*256)。第二种是将短边缩放到某个长度(如 256)，将长边按比例缩放，然后在训练时随机裁剪到某个固定尺寸(比如 224*224)，在验证时进行中心裁剪[1]。通过对比实验，我们发现直接 Resize 要比随机裁剪更好，这可能是因为天气是全局特征，随即裁剪可能会丢失某些重要特征，比如天空、阴影等。直接 Resize 虽然改变了图像的比例，但是不影响全局天气的识别，而且在一定程度上也增强了模型的泛化能力。

在数据增强上，我们尝试了多种数据增强的方法，包括随机水平翻转、随机旋转、亮度变化、对比度变化、饱和度变化、色调变化，还尝试了比较新颖的随机数据增强[2]。但是整体上效果不太理想，尤其是亮度、对比度、饱和度等变换，这说明这几个特征对于天气的识别是非常重要的，因此不能随意修改这几个特征。而且，考虑到这次数据量本身就比较大，足够训练深度神经网络，我们决定只采用简单的随机水平翻转。

在模型的选择上，我们测试了 resnet、densenet、efficientnet、mobilenet、shufflenet、ViT 等模型，结果显示，efficientnet-b3 是分类准确率最高的，我们把它作为主力模型。但是，efficientnet 的训练速度较慢，可能要训练上一整天才能够拟合好，而 shufflenet、mobilenet 等轻量级的模型只要两个小时左右就能够训练到 80 左右，因此考虑在轻量级模型上测试各种方法的优劣，然后在主力模型上进行上分。

在优化器的选择上，我们使用了 Adam、带动量的 SGD 以及 Adam 的变种。Adam 训练比较快，最终效果也比较好，SGD 精度也很高，但是训练比较慢，而且容易陷入局部最优解，而无法达到整体最优。

在损失函数上，我们使用了分类问题常用的交叉熵损失函数。

在训练集和验证集的划分上，我们使用了多种不同比例的划分方式，将 Part1 划分为 15:2、16:1 等。在验证时，我们采用准确率、f1 值作为评价指标。我们先用 Part1 的数据进行训练和验证，在最终提交结果上分数不错，可见模型的准确率和泛化能力比较好。然后，再用该模型预测 Part2-4，发现得分非常低，只有 0.2 的准确率，因此我们认为可能是由于 Part1 和 Part2-4 数据来源不一致，标注方法不统一造成的。所以，果断舍弃了 Part2-4。

以上为在 10 月 22 日之前，我们得出最好成绩时所做的工作。

以下部分为我们在考虑如何进一步提高我们的分数时所做的工作。

如图 3 所示，我们观察了验证集中模型预测错误的图片，发现有一些比较容易混淆的图片，比如阴天和雾霾天，晴天和多云等。因此我们考虑手动提取一些特征，辅助神经网络进行判断。我们提取了图像的颜色直方图、暗通道特征、亮度、对比度等特征[3]，将它们与卷积层提出的特征进行拼接，再输入全连接层进行分类，实现了传统特征提取与深度学习的融合。

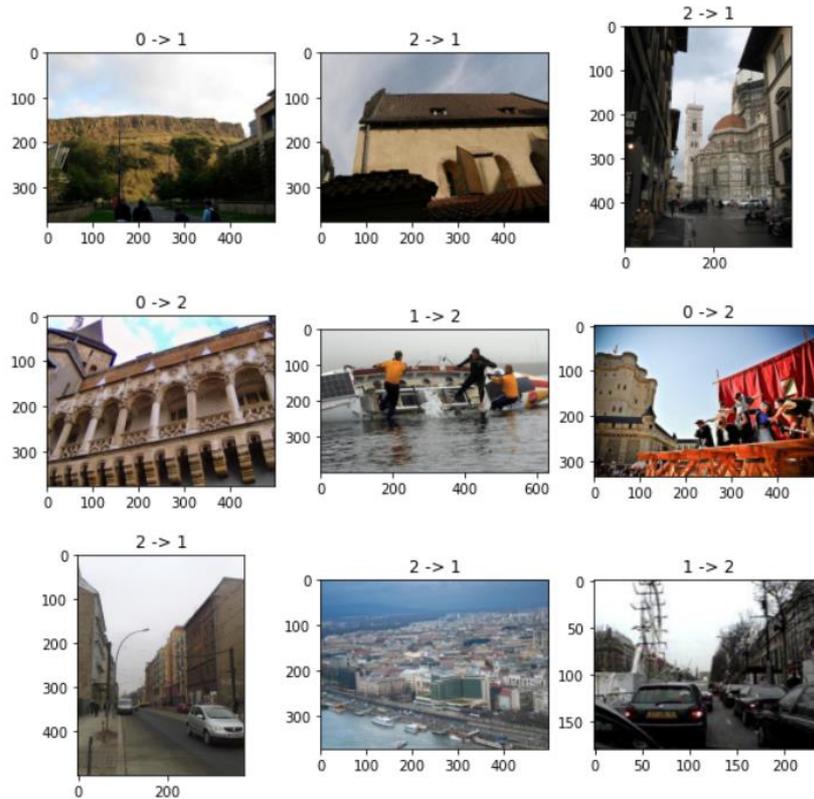


图 3 预测错误的图片示例(标签为 2，预测结果为 1，则记为 2->1)

除此之外，我们还使用了伪标签方法。将原始训练数据划分为训练集和验证集，然后先用训练集数据训练模型 A，得到测试集的预测结果，再将测试集的预测结果当作真实标签，与训练集结合，一起训练模型 B，采用原先划分的验证集来验证。

最后一点，我们将不同训练方法得到的结果进行集成，包括投票、加权投票等方法。

3.方案概览

我们得到最强单模型的方案如下:

- ① 计算训练集和验证集的均值和方差，用于数据标准化
- ② 将图片 Resize 到 300*300，进行随机水平翻转的数据增强
- ③ 将 Part1 中的数据按照 15:2 划分为训练集和验证集
- ④ 选取 efficientnet-b3 作为模型，载入了 ImageNet 的预训练权重，batch size 设置为 32，交叉熵作为损失函数，优化器使用 Adam，weight decay 设置为 0.0001，先以 0.001 的学习率训练 4 轮，然后降低 10 倍后训练 4 轮，再降低 10 倍后训练 2 轮，提交结果为 0.8322
- ⑤ 加入其他模型预测得到的伪标签，对 efficientnet-b3 再进行训练，先以 0.001 的学习率训练 4 轮，然后降低 10 倍后训练 4 轮，再降低 10 倍后训练 2 轮，提交结果为 0.8358

除了最强单模型之外，我们还生成了如下结果

结果 2: 使用模型为 shufflenetv2, 图片 256x256,也用了伪标签方法

结果 3: efficientnet-b3, 没有使用伪标签, 训练验证集划分与最强单模不同

结果 4: efficientnet-b0, 采用带动量的 SGD 优化器(其他用的都是 Adam), 没有用伪标签, 图片 224x224

结果 5: efficientnet-b3 采用伪标签, 训练验证集的划分方式与最强单模型不同

结果 6: efficientnet-b1, 图片大小 240x240, 训练验证集的划分方式也与之前不同

最后, 我们将以上 6 个结果进行投票集成, 一人一票, 取票数最高类别, 如果出现打平的情况, 则按照最高单模型的结果, 得到了最终提交结果 0.8381。

4.创新细节

- ① 提取了图像的颜色直方图、暗通道特征、亮度、对比度等特征, 将它们与卷积层提出的特征进行拼接, 再输入全连接层进行分类, 实现了传统特征提取与深度学习的融合。
- ② 采用了伪标签方法, 将模型的预测结果与训练集相结合来训练新的模型, 进一步提升了模型精度和泛化能力。

5.后续改进思考

从实用性的角度考虑, 我们后期需要进一步提升单模型预测的准确率。可能的途径是通过以下几个方面:

- ① 尝试高阶数据增强, 比如 mix up, 进一步提升模型的泛化能力
- ② 可以采用加权损失函数, 如 focal loss, 这可能会让模型对于某些类别具有偏好性
- ③ 采用混合精度训练, 加快训练速度
- ④ 选择更好的学习率调整策略, 比如 warm up、余弦退火调整、自适应调整等等, 从而让模型更好地收敛
- ⑤ 可以尝试改进模型的结构, 针对天气图像的特点进行优化

参考文献

- [1] A. Krizhevsky, I. Sutskever, and G. Hinton. Imagenet classification with deep convolutional neural networks[C]. Advances in neural information processing systems, 2012.
- [2] D. C. Ekin, Z. Barret, S. Jonathon, et al. RandAugment: Practical automated data augmentation with a reduced search space[C]. Computer Vision and Pattern Recognition, 2020.
- [3] 郭志强, 胡永武, 刘鹏. 基于特征融合的室外天气图像分类[J]. 计算机应用, 2020, 4: 1023-1029.

附录

团队成员：

杨扬 910022508@qq.com

沈纪元 815462629@qq.com

祝嘉诚 1398359669@qq.com

董开元 2331867197@qq.com

戚文栋 2066223131@qq.com

蔡卓悦 2318486955@qq.com

指导老师：

沈俊 jshen@t.shu.edu.cn