

CVDL(2020 Spring)第一次作业

一、作业任务：

请从以下两个任务中选择一个完成：

(A) 场景分类

本任务需要根据给定的数据训练并测试场景分类模型。数据集共包含50类的不同场景图像，其中训练集有超过9万张图像，测试集有5000张图像。每个场景包含500-3000张不等的图片。要求同学们建立场景分类模型，并对给定测试集的数据进行预测。

(B) 细粒度分类

本任务需要大家预测180种鸟类的种类。数据集共有24497张训练图像，900张验证集图像，900张测试集图像。在训练集内每类有100-200张图像不等；在测试集内，每类各有5张图像。要求同学们建立模型，在给定的测试集上进行预测。

二、参与方式

本次作业采用在线竞赛平台Kaggle，同学们通过以下链接参与竞赛（邀请制竞赛，只能通过以下链接进入），并按照要求下载数据和提交结果，详细要求见竞赛内页面说明

场景分类：

<https://www.kaggle.com/t/d254073775b04928a6478768ccd3081b>

细粒度分类：

<https://www.kaggle.com/t/f0614880cd3f40d9919dfecb808b004a>

本次竞赛持续一个月时间，截至日期为2020.5.23 23:59 (UTC+0)，北京时间为2020.5.24早8点，请务必尽早动手

三、实现要求

考虑到不是所有同学能够访问GPU资源，本作业提供两种方案，两种方案任选其一即可：

- CPU方案：实现一种传统框架的分类算法并完成该任务(如SIFT+BoW+SVM)，适合于没有GPU资源的同学
 - SIFT、SVM等算法不必自己实现，可以调用现成的包
 - 选择这一方案需要尝试不同的feature和分类算法，如SIFT和HOG, SVM和kernel SVM等，并且建议尝试调用包时改变具体的参数，不必尝试很多
- GPU方案：使用深度网络来完成该任务，适合于有GPU资源的同学
 - 没有GPU资源但同样想尝试这一方案的同学，可以使用Google的Colab，能够免费使用一块GPU/TPU，<https://colab.research.google.com/notebooks/welcome.ipynb>
 - 选择这一方案必须要尝试课上讲到的训练手段，包括数据增强、学习率调整、损失函数的选择、正则化等等，不必尝试很多组合，选择合适的即可

四、提交要求

- 参与竞赛并在线提交结果
- 并提交一份不超过3页的报告(pdf格式)，说明自己所使用的方法以及主要工作，特别要把以上要求尝试的部分写清楚，同时提交代码
- 请大家将最终材料打包发送到cvdl2020@163.com，命名格式为 学号_姓名_第一次作业.zip，提交的截止日期为5月23日晚23点59分(北京时间)
- 注意：在线评测的排名仅供参考，对于最终的成绩评定我们主要衡量报告质量、代码、算法设计等，在线排名并非主要，请同学们不要为了刷点采取一些不必要的手段

五、联系方法

有任何问题可以联系各位助教，邮箱或者微信都可以，几位助教邮箱如下：

姜博睿：jbr@pku.edu.cn

田贵宇：guiyutian@pku.edu.cn

薛犇：xueben@pku.edu.cn

张天远：1600012888@pku.edu.cn

周新哲：zhouxinzhe1023@pku.edu.cn

六、参考资料

pytorch tutorial: https://pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html

tensorflow tutorial: <https://www.tensorflow.org/tutorials/images/classification>

vlfeat开源library: <https://www.vlfeat.org/>

opencv开源library: <https://opencv.org/>

sklearn开源library: <https://scikit-learn.org/>