

《计算机视觉》期末大作业

作业要求：完成如下四道题，撰写实验报告，需提交可运行的源代码即实验测试结果图（建议用 python）。

实验报告需要包括问题的描述，求解过程，中间实验结果示例，结果展示和分析等。

作业包含数据集 data.zip 一份，数据集包含 imgs 和 gt 两个文件夹，gt 中的图片为 imgs 对应图像的前景区域标注，图片名称为 XXXX.png，XXXX 为图像 id。每位同学的实验测试图像为图像子集{id|id%100==学号后两位}，例如某同学的学号后两位为 01，则他实验报告中需要展示图像子集{1.png,101.png, 0201.png...0901.png}的处理结果，其余图像作为训练图像。

1. (30 分) 结合 “Lecture 6 Resizing “ 的 Seam Carving 算法，设计并实现前景保持的图像缩放，前景由 gt 文件夹中对应的标注给定。要求使用 “Forward Seam Removing” 机制，X, Y 方向均要进行压缩。压缩比例视图像内容自行决定（接近 1-前景区域面积/（2*图像面积）即可）。每一位同学从各自的测试子集中任选两张代表图，将每一步的 seam removing 的删除过程记录，做成 gif 动画格式提交，测试子集的其余图像展示压缩后的图像结果。
2. (20 分) 结合 “Lecture 7 Segmentation” 内容及参考文献[1]，实现基于 Graph-based image segmentation 方法（可以参考开源代码，建议自己实现），通过设定恰当的阈值将每张图分割为 50~70 个区域，同时修改算法要求任一分割区域的像素个数不能少于 50 个（即面积太小的区域需与周围相近区域合并）。结合 GT 中给定的前景 mask，将每一个分割区域标记为前景（区域 50%以上的像素在 GT 中标为 255）或背景（50%以上的像素被标为 0）。区域标记的意思为将该区域内所有像素置为 0 或 255。要求对测试图像子集生成相应处理图像的前景标注并计算生成的前景 mask 和 GT 前景 mask 的 IOU 比例。假设生成的前景区域为 R1，该图像的 GT 前景区域为 R2，则
$$IOU = \frac{R1 \cap R2}{R1 \cup R2}$$

[1] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation[J]. International journal of computer vision, 2004, 59(2): 167-181. <http://people.cs.uchicago.edu/~pff/papers/seg-ijcv.pdf>

3. (50 分) 从训练集中随机选择 200 张图用以训练，对每一张图提取归一化 RGB 颜色直方图（8*8*8=512 维），同时执行问题 2 对其进行图像分割，（分割为 50~70 个区域），对得到的每一个分割区域提取归一化 RGB 颜色直方图特征（维度为 8*8*8=512），将每一个区域的颜色对比度特征定义为区域颜色直方图和全图颜色直方图的拼接，因此区域颜色区域对比度特征的维度为 2*512=1024 维，采用 PCA 算法对特征进行降维取前 20 维。利用选择的 200 张图的所有区域（每个区域 20 维特征）构建 visual bag of words dictionary（参考 Lecture 12. Visual Bag of Words 内容），单词数（聚类数）设置为 50 个，visual word 的特征设置为聚

簇样本的平均特征，每个区域降维后颜色对比度特征（20 维）和各个 visual word 的特征算点积相似性得到 50 个相似性值形成 50 维。将得到的 50 维特征和前面的 20 维颜色对比度特征拼接得到每个区域的 70 维特征表示。根据问题 2，每个区域可以被标注为类别 1（前景：该区域 50%以上像素为前景）或 0（背景：该区域 50%以上像素为背景），选用任意分类算法（SVM, Softmax, 随机森林, KNN 等）进行学习得到分类模型。最后在测试集上对每一张图的每个区域进行测试（将图像分割为 50~70 个区域，对每个区域提取同样特征并分类），根据测试图像的 GT，分析测试集区域预测的准确率。

4. （20 分）

结合 “Lecture 10. Dimensionality Reduction” 中学习的方法，每一位同学从各自的测试子集中任选一张代表图，执行 PCA 压缩。先将图片尺寸缩放或裁减为 12 的倍数，以 12*12 patch 为单位执行 PCA 压缩，1) 展示 16 个最主要的特征向量的可视化图，2) 展示 144D, 60D, 16D 和 6D 的压缩结果。需要介绍算法流程和对应的结果展示。

总分 120 分，得分超过 100 分按 100 分计算。

评分占比：此次作业占总成绩的 50%，本次作业的评分细则（1）算法实现的准确性 50%；（2）实验报告的规范性 30%；（3）代码规范性 20%；

提交时间：2022 年 7 月 8 日 24:00 前，推迟一天成绩减 5 分（不足一天按一天计算），以有效作业提交的邮件时间戳为准。

提交方式：作业命名 “学号-姓名-计算机视觉期末作业.zip”，交到 TA 黄日聪（邮箱：cv_2022@163.com）处。

注意：不得抄袭，包括源代码 !!!