

1 目标任务

1. 将输入图片分割为 N 张子图片打乱输入，并将他们的正确排列作为输出，输出格式采用排列阵 P ，第 i 行第 j 列若为 1 则代表图片 i 应该排在位置 j ，每行每列最多有一个元素为 1。具体切分方式可以自行定义，但需要有相应的可视化效果说明自己的切分设计。
2. 设计一套针对图片拼接的神经网络，可以参考 DeepPermNet 的设计，即对每一个子图片过一个 CNN，将得到的特征进行拼接后过若干全连接网络后得到一个 $N \times N$ 的矩阵，通过 Sinkhorn 算法使其变成双随机矩阵 Q （每行每列的和固定为 1）作为神经网络的预测值。最后将预测值 Q （双随机矩阵）与真值 P （排列阵）作 Loss 后用于更新神经网络。
3. CNN 和 FC 网络等可以套用作业二中的代码，也可以自行设计新的网络。Sinkhorn 算法我们推荐使用 pygmtools 中的 sinkhorn（后面大作业可能也需要用到 pygmtools，可以提前熟悉一下）。
4. 在实现了图片拼接问题的网络以后，尝试能否把图片拼接作为一种预训练方式来更好的辅助作业二中的图片分类任务。