本次深度学习课程大作业基于“视觉”与“图”两大深度学习的重点研究方向，对于一个视觉问题中的应用（拼图问题），在参考已有论文的基础上，提出自己的解决方案。同学们需要在DDL前将大作业相关的文件（包括报告和代码）一起打包上传到Canvas上。（超过DDL十天以内，晚交会扣除一定的分数，**每**晚交一天扣一分，超过十天则不再收。）

目标：

在CIFAR-10数据集上，我们将在图像分类任务以外研究图片拼接这一任务，即，对于一张图片，我们将其切分为若干张子图片，例如水平竖直分为四份。接着，我们将这几张子图片打乱顺序作为输入，目标就是设计一个模型可以根据这些乱序的子图片还原出他们本来的排列位置。

相关参考论文：DeepPermNet: Visual Permutation Learning <https://arxiv.org/abs/1704.02729>

任务步骤：（这里只是推荐的实现步骤，如果有不同的设计也可以自行实现，最终我们将依据pdf报告中展示的完成质量综合考虑进行打分）

1. 修改作业二/作业三中的图片分类的代码框架，将输入图片分割为N张子图片，打乱它们的顺序作为输入，并将他们的正确排列作为label，输出格式采用排列阵P，其中第i行第j列为1则代表第i个子图片应该排在第j个位置。注意，每行与每列均最多有一个元素为1。具体切分方式可以自行定义。
2. 设计一个针对图片拼接的神经网络模型，可以参考DeepPermNet的设计，即将每一个子图片输入一个CNN进行特征提取，将得到的特征进行拼接后通过若干全连接网络得到一个N x N 的矩阵，使用Sinkhorn算法使其变成双随机矩阵Q（即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）作为损失函数的输入，计算Loss后用于更新神经网络。
3. 神经网络可以套用作业二/作业三中的代码，也可以自行设计新的网络。Sinkhorn算法推荐使用pygmtools中的Sinkhorn，这是我们实验室自行研发的开源平台。

本作业侧重探索性质，对精度不做要求，没有固定的解决方案，言之有理即可。最终要求提交一份五页pdf报告，在DDL前与代码一起上传到Canvas上。（报告限制pdf五页）