Homework——Self-supervised Learning by Predicting Absolute Location

Self-supervised learning

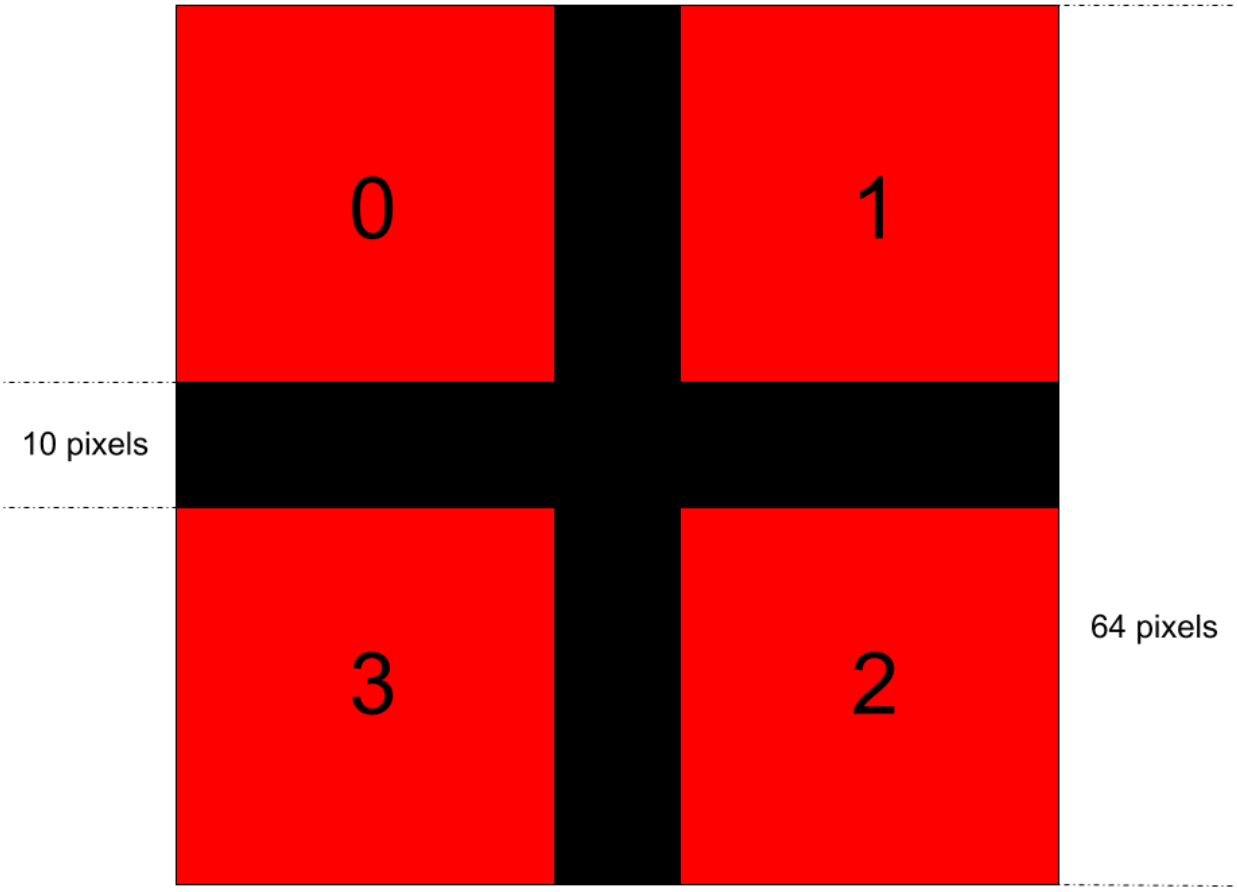
Self-supervised learning is one of the most important paradigms of unsupervised learning. It uses naturally existed signals for training and there is (almost) no human intervention. Here are some already existed self-supervised learning methods:

- Auto-encoders by bottlenecks
- Denoise from clean images with Gaussian noise
- Image Colorization
- Inpainting
- Predict whether an image is rotated

Predicting Absolute Location

In this homework, you need to predict the absolute location of an image patch. Specifically, given an image patch P cropped from an unknown image I, you need to predict the absolute location of P on I. You have four choices:

- top left –> 0
- top right –> 1
- bottom right –> 2
- bottom left –> 3



Dataset

All data is saved in DATA.zip. Unzip this file, there will be three folders:

- train: there are 5000 images each of which is a 64×64 color image. They are selected from 100 unknown classes and each class with 50 images. For example, the 19th image of 9th class is named as “009_019.JPEG”.
- test_crop_easy: there are 500 patches each of which is a 25×25 color image. They are cropped from unknown images. These unknown images are selected from the same classes of the training set.
- test_crop_hard: similar with test_crop_easy but selected from unseen classes.

Note the center of the cropped patch is at least 5 pixels away from the center of the whole image. See above figs and **utils.py** for reference.
random_crop in **utils.py** return a randomly cropped patch and its label, given an image array.
You need to crop training images by yourself.
Given a cropped patch, you need to assign a label $(0, 1, 2, 3)$ to it.

Submission

You need to submit two txt files, the one for *test_crop_easy*, the other for *test_crop_hard*.
They are named as *{student ID}_easy.txt* and *{student ID}_hard.txt*.
For example, 1502110189_easy.txt and 1502110189_hard.txt.
Each line in each file contains the name of an image file and its label, splitted by a comma (order doesn’t matter).
For example:

000_050.JPEG,0
000_051.JPEG,1
000_052.JPEG,2
000_053.JPEG,3

You can use any methods but the overall accuracy must be above 35%.