VRDL – HW4 Report

Model link: https://drive.google.com/file/d/158Lj8p-jYrZsJnTZMW1951QTMLJRc0EP/view?usp=sharing

1 GitHub link: https://github.com/andychan8877/2021_VRDL_HW4.git

2 Reference:

2.1 https://github.com/cszn/KAIR

2.2 https://arxiv.org/pdf/2108.10257.pdf

3 Introduction:

3.1 Model:

I use SwinIR from papers-with-codes, it's SOTA image resolution method based on transformers which show impressive performance on high-level vision tasks.

4 Methodology:

4.1 Setting:

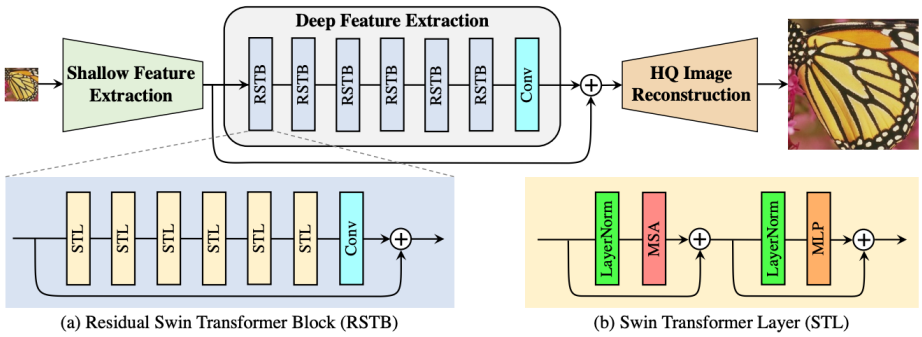
I noticed that the size of some of the training images is different, this may cause some problem in Dataloader, so I set the size of input images to 78. In addition, I set the batch size to 16 to deal with OOM problem, and set the scaling factor to 3.

4.2 Model architecture:

SwinIR is based on Swin transformer. Swin transformer layer(STL) consists of multi-head self-attention(MSA) of original transformer layer. But using shifted window mechanism to get the patches of images.

And multi-layer perceptron(MLP) which has two fully-connected layers with GELU to get feature information

The core module block Residual Swin Transformer Block(RSTB) consists of STLs and convolutional layer. By connecting RSTBs, SwinIR do deep feature extraction



(a) Residual Swin Transformer Block (RSTB) (b) Swin Transformer Layer (STL)

Figure 2: The architecture of the proposed SwinIR for image restoration.

5 Summary:

The opened code of SwinIR is very easy to use and it's very powerful for image resolution. Embodying the powerful functions and prospects of transformers for computer vision tasks. During this homework, I mainly spend most of my time training the model, in the end the PSNR was not getting more higher, I think I should do data-augmentation, maybe the PSNR can be more higher.

So, this is the last homework in VRDL, I successfully beat the baseline and see the LR-images turning into HR-images, it's very fulfilling. Visual Recognition is a really interesting field, I learned a lot this semester, hope that I will learning more in the future, too.