HW4 Image Super Resolution

Selected Topics in Visual Recognition using Deep Learning

GitHub link

https://github.com/samuelyutt/Selected-Topics-in-Visual-Recognition-using-Deep-Learning-course/tree/hw4/hw4-ImageSuperResolution

Utilities and reference

- EDSR-PyTorch: https://github.com/sanghyun-son/EDSR-PyTorch

- Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," 2nd NTIRE: New Trends in Image Restoration and Enhancement workshop and challenge on image super-resolution in conjunction with CVPR 2017.

Introduction

Image Super Resolution is a technique to enhance the resolution of an image. Most recent research on super-resolution problem relies a lot on deep neural networks, such as VDSR and SRResNet. The former architecture requires heavy computation of bicubic interpolated image as the input, and the latter simply applies the ResNet architecture with good performance. However, SRResNet results in suboptimal since ResNet was proposed to solve higher level computer vision problems.

Enhanced deep super-resolution network (EDSR) solves the above problems by first optimizing SRResNet architecture, analyzing, and removing unnecessary modules to simplify the network architecture. According to the paper, the authors claim that EDSR produces better results than many other SR methods, which shows it a powerful network.

Methodology

Data pre-process

Since the training images are only provided in high resolution format, I simply generate the input training by downscaling the given images with a factor of 3. To do so, I applied bicubic interpolation.

Model architecture

The EDSR structure is similar to SRResNet, but without having ReLU activation layers outside the residual blocks. Also, residual scaling layers are not applied since there are only 64 feature maps for each convolution layer.

The architecture of the proposed single-scale SR network (EDSR) are shown in figure 1.

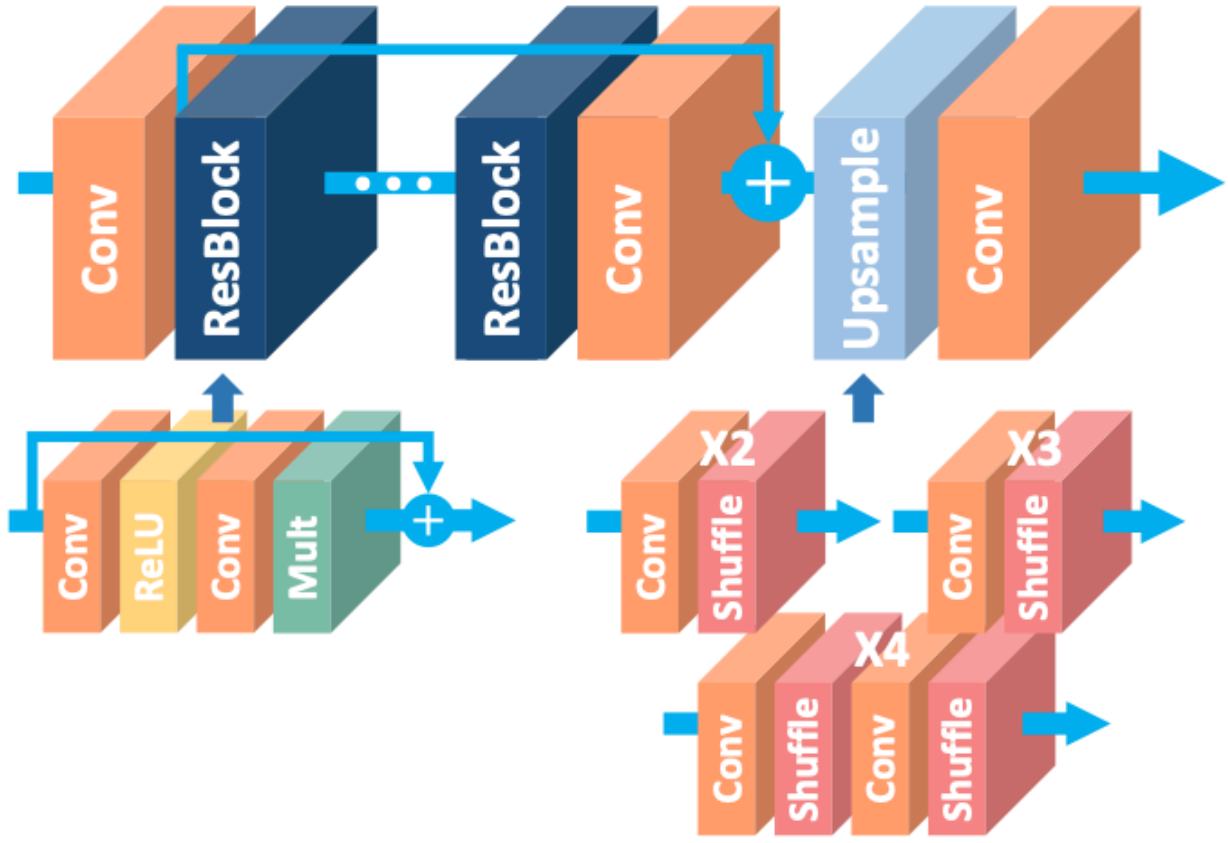


Figure 1

Hyperparameters

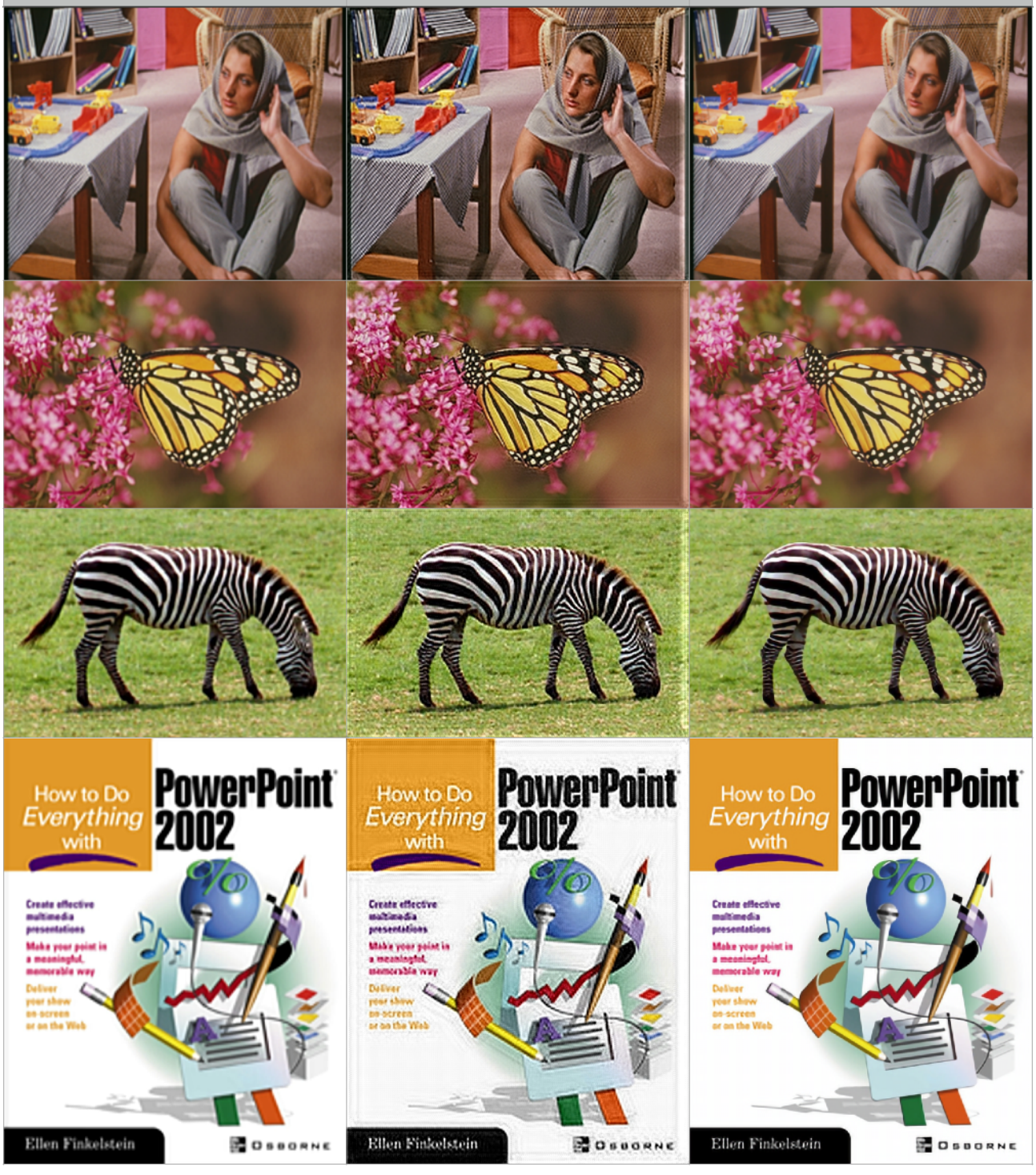
The learning rate is set to 1e-4 at the very beginning of the training, and the learning rate decay is set to 200 with a factor of 0.5.

The optimizer is ADAM, with an epsilon set to 1e−8 for numerical stability.

A checkpoint of weights are stored for each epoch.

Results

I have tried 3 methods from GitHub: VDSR, ISR, and EDSR.

VDSR	ISR	EDSR
		

Summary

In this homework, I have worked on the 3 aforementioned image super resolution method. VDSR was the most easiest to use and did not performs bad, though it did not reach the baseline. I then found ISR and trained it. The results of ISR is pretty bad. As we can see that the edges of the ISR results are quite strange. I believe that I might have done something wrong during the training. After training ISR, I did some paper research and found EDSR. EDSR is claimed to be improved by fixing several shortcomings of VDSR and SRResNet. Indeed, it outperforms the previous 2 methods and beats the baseline as well.

Training 3 networks is extremely time consuming, especially during finals. Therefore, I do not have sufficient time to search for optimal hyper parameters after training EDSR. Additionally, I have an idea to train with an upscale factor of 6. I can thus input a low resolution image and obtained a 6x high resolution output. To generate the 3x submission image, I can downscale the 6x image by training another network, or by simply applying bicubic interpolation with a factor of 2. I am curious about the quality of the output.

Due to the lack of time, I marked these as future work. Also, I will make sure to do the paper survey BEFORE I choose a method next time. Still, I learned a lot from this homework, such as training custom datasets on image resolution problem and doing research on the EDSR network architecture.