Cara Evangeline Chandra Mohan
190539421

Course Work 1: Image Super-resolution Using Deep Learning

1. Suppose the settings of a SRCNN as: f1=9, f2=3, f3=5, how many pixels of the low-resolution image are utilized to reconstruct a pixel of the high-resolution image with the SRCNN?

The estimation of a high resolution pixel utilizes the information of $(9+5-3)^2$ =121 pixels.

2. Why the deep convolutional model is superior to perform image super-resolution? Give one reason to explain it.

Reason 1:

- In general interpolation is used to upscale an image which gives very less visual quality, as the details (e.g. sharp edges) are often not preserved.
- More sophisticated methods exploit internal similarities of a given image or use datasets of low-resolution images and their high-resolution counterparts, learning a mapping between them.
- In external SR methods, like Sparse coding requires a dictionary to be found that will allow us to map low resolution images into an intermediate, sparse representation. In addition, the HR dictionary is learned, and will allow us to restore our estimate of a high resolution image. Such a pipeline usually involves several steps and not all of them can be optimized. Ideally we would like to have all of these steps combined in one big step with all of its parts being optimizable. This can be achieved by a deep neural network like deep CNN.

Reason 2:

- In Convolutional Neural network, instead of feeding the entire image as an array of numbers, the image is broken up into a number of tiles, based on the size of the filter. Here, each filter has it's own specific functionality and is very much useful in extracting texture and other properties.

Reason 3:

- Information exploited for reconstruction is comparatively larger than that used in any other methods (ie. 121 pixels incase of 9-3-5 filters for SRCNN)

Cara Evangeline Chandra Mohan
190539421

3. Please explain the physical meaning of peak signal-to-noise ratio (PSNR) in the context of image super-resolution. PS: place here the ground truth (GT) image, and the high-resolution images by SCRNN (HR-SRCNN) and bicubic interpolation (HR-BI) for reference. Also put the PSNR value below the high-resolution images.

Fig 1: Ground-truth image

Fig 2: Bicubic Interpolation

Fig 3: Super Resolution Image

<u>Result:</u>

```
Super Resolution - PSNR:  20.74699270372715
Bicubic/Blurred - PSNR:  20.497630173285614
(python3-tensorflow-gpu) bash-4.2$
```

- A metric commonly used to measure the quality of reconstruction of lossy compression is called Peak Signal to Noise Ratio (PSNR).
- It measures how much the distorted image deviates from the original high-quality image (Ground truth). Here, PSNR is the ratio of maximum possible pixel value of the image to maximum mean squared error (MSE) between the original image and its estimated version, expressed in logarithmic scale.
- The larger the PSNR values, the better the reconstruction, and therefore maximization of PSNR naturally leads to minimizing MSE as the objective function.