ECS795P Deep Learning and Computer Vision, 2021

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? (10% of CW1)

   Using the formula (f1 +(f2-1) + (f3-1))2 => (9 + 2 + 4)**2 = 225 pixels of the low-resolution image are applied to reconstruct a pixel of high-resolution image with SRCNN.

2. Why the deep convolutional model is superior to perform image super-resolution? Give one reason to explain it. (10%of CW1)

   Three operations: patch extraction, representation, non-linear mapping and reconstruction - are motivated by different intuitions, but they all lead to the same form as a convolutional layer. Another words, these operations are essential for solving the task that can be ideally reformulated into a CNN. In addition, its structure is intentionally designed with simplicity, and yet provides superior accuracy and faster speed compared with state-of-the-art example-based methods like Sparse coding.

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. (10% of CW1)

   PSNR is calculated using the Mean-Square-Error (MSE) of the pixels in a refurbishment image and the maximum possible pixel value (MAXI) in original image as follows:

$$\boldsymbol{PSNR} = 10 \cdot \log(\frac{MAX_I^2}{MSE})$$

GT
HR-BI (PSNR= 22.73)
HR-SRCNN (PSNR= 23.65)

HR Image

LR Image

SR Image