

# ECS795P Deep Learning and Computer Vision, 2022

---

## Course Work 1: Image Super-resolution Using Deep Learning

1. Suppose the settings of a SRCNN as:  $f_1=9$ ,  $f_2=3$ ,  $f_3=5$ , 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)

**Ans:** According to the formula:

$$r_0 = \sum_{l=1}^L \left( (k_l - 1) \prod_{i=1}^{l-1} s_i \right) + 1$$
$$((9-1 + 3-1 + 5-1) + 1) \times ((9-1 + 3-1 + 5-1) + 1) \Rightarrow 225 \text{ pixels}$$

225 pixels of the low-resolution image are utilized to reconstruct a pixel of the high-resolution image with the SRCNN.

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

**Ans:** The deep convolutional network gives us an end-to-end mapping between the low resolution and high-resolution images. In our convolutional neural network, the low-resolution dictionary, high-resolution dictionary, non-linear mapping, together with mean subtraction and averaging, are all involved in the filters to be optimized. So, our method optimizes an end-to-end mapping that consists of all operations. This optimized end-to-end mapping makes SRCNN faster than the traditional sparse coding methods. A typical and basic setting for SRCNN is  $f_1=9$ ,  $f_2=1$ ,  $f_3=5$ ,  $n_1=64$ , and  $n_2=32$ . Here, the estimation of a high-resolution pixel utilizes the information of  $(9+5-1)^2 = 169$  pixels. This information exploited for reconstruction is larger than that used in the existing sparse coding methods using  $(5+5-1)^2 = 81$  pixels. Clearly the SRCNN exploits more information for reconstruction. This is one of the prominent reasons why SRCNN gives superior performance.

3. Please explain the meaning of **peak signal-to-noise ratio (PSNR)** in the context of image super-resolution. PS: give the ground truth (GT) image, and the high-resolution images by SRCNN (HR-SRCNN) and interpolation (HR-Base) for reference. Also put the PSNR value below the high-resolution images. (10% of CW1)

**Ans:** The peak-signal-to-noise-ratio (PSNR) is a widely used metric for quantitatively evaluating image restoration quality and is at least partially related to

the perceptual quality. The proposal is that the higher the PSNR value is, the better degraded image has been reconstructed to match the original image and the better the reconstructive algorithm is. The main limitation of this metric is that it relies strictly on numeric comparison and does not actually consider any level of biological factors of the human vision system such as the structural similarity index. (SSIM).

GT



HR-Base (PSNR= 20.4976)



HR-SRCNN (PSNR= 22.9226)

