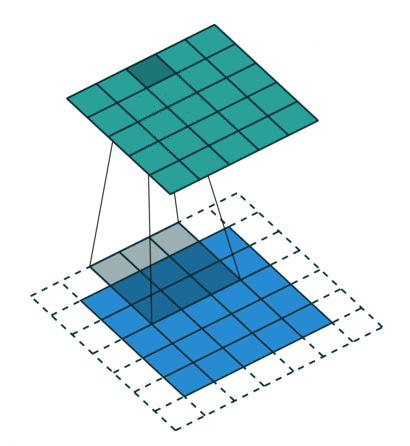
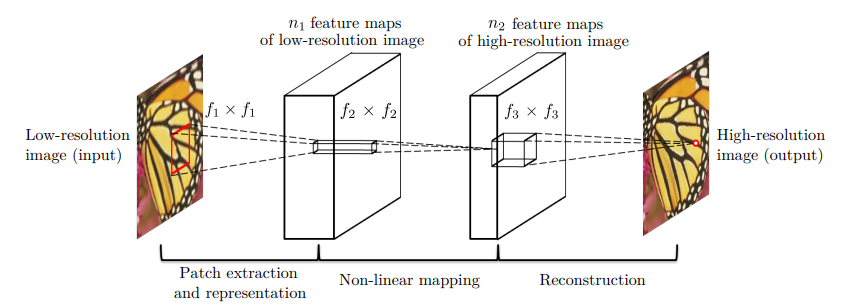
ECS795P Deep Learning and Computer Vision, 2022

**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)



Answer：225

As can be seen from the above figure, if f3=5, then a pixel in HR needs 5\*5 pixels in n2; If f2=3, then the 5\*5 pixels in n2 need (5+2)\*(5+2) pixels in n1; If f1=9, then (5+2)\*(5+2) pixels in n1 need (5+2+8)\*(5+2+8) pixels in LR, i.e.225 pixels.

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

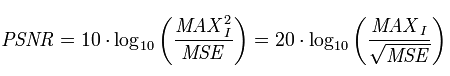
Answer:

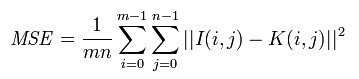
not all operations have been considered in the optimization in the sparse-coding-based SR methods. On the contrary, in convolutional neural network, the low-resolution dictionary, high-resolution dictionary, non-linear mapping, together with mean subtraction and averaging, are all involved in the filters to be optimized. So deep convolutional neural network optimizes an end-to-end mapping that consists of all operations.

For example, a typical and basic setting is f1 = 9, f2 = 1, f3 = 5, n1 = 64, and n2 = 32 (we evaluate more settings in the experiment section). On the whole, the estimation of a high resolution pixel utilizes the information of (9 + 5 − 1)^2 = 169 pixels. Clearly, the information exploited for reconstruction is comparatively larger than that used in existing external example-based approaches, e.g., using (5+5−1)^2 = 81 pixels. This is one of the reasons why the SRCNN gives superior performance.

1. 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 SCRNN (HR-SRCNN) and interpolation (HR-Base) for reference. Also put the PSNR value below the high-resolution images. (10% of CW1)

Peak signal-to-noise ratio (PSNR) is an objective criterion for evaluating images. It is the logarithm of the mean square error between the original image and the processed image relative to (2^n-1)^2 (the square of the maximum value of the signal, n is the number of bits per sample value), and its unit is dB.





where MSE is the mean squared error between the original image and the processed image. MAXI: Indicates the maximum value of the image color, 8-bit sampling points are expressed as 255. Peak refers to the maximum value of 255 in 8bits notation. MSE refers to MeanSquareError, I (superscript n) refers to the nth pixel value of the original image, and P (superscript n) refers to the nth pixel value of the processed image. So the larger the PSNR value, the less the distortion.

|  |
| --- |
| GT |
| HR-Base (PSNR= 20.49756351563886  ) |
| HR-SRCNN (PSNR= 22.923322214270737  ) |