ECS795P Deep Learning and Computer Vision, 2020

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)

Ans: $(f1 + f2 - 1 + f3 - 1)^2$

Therefore, $(9 + 2 + 4)^2 = 225$ pixels of LR image are used to construct a pixel of HR image.

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

Ans: SRCNN outperforms sparse-coding-based method. SRCNN learns an end-to-end mapping between the low resolution image (input) and the high resolution image (output). Unlike spare-coding method, SRCNN optimizes the end-to-end mapping making it faster than sparse-coding. Also, the sparse-coding solver is an iterative algorithm whereas, the SRCNN is fully feed-forward. The common setting for SRCNN is f1 = 9, f2 = 1, and f3 = 5 with n1 = 64 and n2 = 32 which results in the utilization of $(9 + 5 -1)^2 = 169$ pixels in the LR image for a pixel in HR image. This is a large value when compared to $(5 + 5 - 1)^2 = 81$ pixels for the sparse-coding method. Therefore, SRCNN exploits more information for reconstruction. This is one of the reasons for its superior performance.

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)

Ans: PSNR value or the peak signal to noise ratio is used to determine the similarity between two images. It is calculated as $PSNR=10 * log_{10}(\frac{L^2}{MSE})$, MSE is the mean squared error calculated at the pixel level. $MSE = \frac{1}{N} * \sum_{i=1}^{N}(\text{target-predicted})^2$ and L is the maximum possible value for a pixel (for an 8 bit RGB image L = 255). A high

PSNR means high similarity and vice versa. As PSNR is inversely proportional to MSE, therefore it is evident that PSNR only considers pixel difference. It does not represent perceptual quality. Rather, PSNR prefers blurry results without high frequency details as minimization of the errors averages the possible solutions.

GT 
HR-BI (PSNR=20.45) 
HR-SRCNN (PSNR=21.77)