ECS795P Deep Learning and Computer Vision, 2023 by:- Abhishek Kumar Singh (220982845)

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. This problem is the problem of finding the receptive field of our network with given filters f1,f2 and f3. We can find the number of pixels used to reconstruct the high resolution image using the formula $r_1 = s_2 * r_2 + (f2 - s_2), we\ have\ s = 1\ r_2 = 9, f2 = 3\ so\ we\ get\ r = 11\ similarly, we\ use\ this\ answer\ and\ calculate\ for\ the\ next\ layer, (11 * 1 + (5 - 1))^2 = (15)^2$. We find that 225 pixels of the low resolution image are utilized to reconstruct a pixel of high resolution image with 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. For the case of image super-resolution deep convolutional neural networks were found to be superior when compared to traditional approaches such as external and internal example-based approaches. This is because unlike the traditional approaches the convolutional neural network immaculately learns non-linear end to end mapping between the low-resolution images and high-resolution images, hierarchically which allows them to understand and learn the underlying image structure . This task is implicitly achieved by the network without the loss of finer details or blurring, unlike the traditional methods. DCNNs can capture a lot of abstract details and finer low-level features and hence results in a sharper, finer, and a high quality super-resolution images.

3. 1) In the context of image super-resolution, explain the definition (how to calculate) of peak signal-to-noise ratio (PSNR) and why can PSNR be applied to measure the quality of output images?

Ans. Peak Signal-to-noise ratio (PSNR) is a quantitative metric used for evaluating the quality of image super-resolution algorithms. PSNR is used to measure the difference between the output of the algorithm with its ground truth, by comparing the two images with the pixel values of two images and noise in the image. The PSNR is calculated by first calculating the mean squared error (MSE) for the ground truth and the output image. The formula for MSE in our case is:

$$\sum_{i=1}^{M} \sum_{j=1}^{N} \left[Y(i,j) - X(i,j)\right]^2$$

Where M,N are the number of pixels in the image/ rows and columns in the image.

After calculating the MSE, We can compute the PSNR using this equation:

$$\text{PSNR} = 10 * \log_{10} \frac{(\text{MAX})^2}{(\text{MSE})}$$

Where , (MAX) is the maximum possible pixel value of the image, in the case of an 8-bit grayscale it is 255.
PSNR can be applied to measure the quality of output images because it gives us how similar the output of the algorithm is to high resolution ground truth image. Hence, we can infer that higher the value of PSNR the closer the upscaled image is to the high-resolution ground truth image, that's why we can use to measure the quality of output images with respect to ground truth.
However that might not always be the case, as PSNR struggles with images with complex structures or textures, and they do not consider the perceptual factors of images, as can be seen from the huge disparity between Mean Opinion Scores(MOS) and PSNR values in certain cases.

2) Show the ground truth (GT) image, and the high-resolution images by interpolation (HR-Base) and SRCNN (HR-SRCNN). Also put the PSNR values below the high-resolution images. (10% of CW1)

GT(Original)	HR-Base (Interpolated)	HR-SRCNN
	 PSNR= 20.497	 PSNR= 22.922