ECS795P Deep Learning and Computer Vision, 2023

Course Work 1: Image Super-resolution Using Deep Learning

BY DAKSH RAMESH CHAWLA (ID:220718392)

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:

   The first convolutional layer in the SRCNN model has a receptive field size of f1=9, the second convolutional layer is f2=3, and the third convolutional layer is f3=5.

   In the first convolutional layer, a neuron's receptive field size is 9. A neuron in the second convolutional layer has a receptive field size of = f1 + (f2-1) * 1 = 11. A neuron in the third convolutional layer has a receptive field size of = 11 + (f3-1) * 1 = 15.

   Each pixel in the output high-resolution picture of the SRCNN model is a result of a 15x15 patch of the input low-resolution image. Using this SRCNN, 225 pixels from the low-resolution picture are used to reconstruct one pixel from the high-resolution image.

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

   ANSWER:

   As opposed to conventional techniques that might not be able to capture the underlying structure and complexity of the images, Deep Convolutional Neural Networks (DCNNs) have demonstrated improved performance in image super-resolution. DCNNs can learn highly non-linear and complex mappings from low-resolution to high-resolution images.

Deep neural networks can learn end-to-end mappings between low- and high-resolution pictures, which enables them to complete the whole super-resolution process in a single model without the need for additional stages or intermediary representations. By avoiding mistakes and artefacts that may result from intermediary processing processes, this end-to-end method increases the final image's overall correctness and quality.

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?

ANSWER:

Peak signal-to-noise ratio (PSNR) is a measure to evaluate the image quality of output images. The mean squared error between the high-resolution output image and the original high-resolution image, and the maximum pixel value of the original high-resolution image (denoted as P) is used to calculate the PSNR.

PSNR = 20 * log10(P/sqrt(MSE))

Since PSNR represents how similar the output image and the original image are to one another, it can be used to assess the output image quality in the context of image super-resolution. A single quantitative measure of image quality is provided by PSNR, which accounts for both distortion (represented by the MSE) and the maximum allowable signal amplitude (expressed by P).

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)

ANSWER:

GT

HR-Base (PSNR=20.50 DB)

HR-SRCNN (PSNR=22.92 DB)