ECS795P Deep Learning and Computer Vision, 2022

Bhanu Pratap Singh(210760006)
ec21279@qmul.ac.uk

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?

   For determining the number of pixels of the low-resolution image are utilized to reconstruct a pixel of the high-resolution image with SRCNN we are using the formula given below

   Input = (Output – 1) * S – 2P + F

   Where S is the Stride size
   P is the padding size
   F is the size if the filter used

   In this case we are back propagating from the output layer to the input layer to determine how many numbers of input image pixel is used. Here we are using the stride size =1 and as there is no information about the padding so padding size =0. Also, we have given f1=9, f2=3, f3=5.

   For the output
   Input = (Output – 1) * S – 2P + F
   Input = (1-1) *1-2*0 + 5
   Input =5
   We have taken one pixel from the output layer to determine how many inputs image pixel is used to make one pixel of pixel of output that why we have taken output =1 and we doing it for the output layer. So, we chose f3 filter size that is 5. Input = 5 implies for getting one pixel in the output channel 5 input pixel is intermediate image is used and further this input=5 is taken as output for further calculation
   Input = (Output – 1) * S – 2P + F
   Input = (5-1) *1-2*0 + 3
   Input =7
   Here we taken f2 filter size which is 3 and output=5 and getting the input=7 which implies for generating 5 pixel in intermediate output image 7 input image pixel is used and this input=7 is taken as output for further calculation.
   Input = (Output – 1) * S – 2P + F
   Input = (7-1) *1-2*0 + 9
   Input =15

Here we taken f1 filter size which is 3 and output=7 and getting the input=15 which implies for generating one pixel in output image using SRCNN 15 input image pixel is used from one channel but there are two channel in our input because the input is grey scaled.
=I*I
=15*15
=225
The total number of input image pixel is used is **225**

2. Why the deep convolutional neural network is superior to perform image super-resolution? Give one reason to explain it.

As compared to the traditional method we have chosen to use the deep convolutional neural network to perform image super resolution because unlike the traditional method which handle each component separately our method jointly optimizes all layers thus results in better overall reconstruction quality.

If we compare with the traditional method which uses external example-based SR method like sparse-coding method which can be viewed as kind of convolutional neural network with different non-linear mapping. But not all the operations have been considered in the optimization in the sparse coding-based SR method. On the contrary our convolutional neural network, the low-resolution dictionary, high-resolution dictionary, non-linear mapping, together with the mean subtraction and averaging, are involve in the filter to be optimized. So, our method optimizes end-to-end mapping that consist of all operations.

Deep convolutional neural network is superior to perform image super-resolution due to Its simplicity and robustness could be applied to low vision problems such as image de-blurring or simultaneously SR-denoising.

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

Peak signal-to-noise ratio (PSNR) refers to the ratio between the maximum possible value (power) of a signal and the power of distorting noise that affects the quality of its representation.
Better the PSNR values implies better the results or better the reconstruction of image, lower the value of PSNR implies that we are not getting the optimal reconstruction of image.
Here we are getting better PSNR for the SRCNN as compared to the HR image implies that the reconstruction quality is better in the SR image.

GT

HR-Base (PSNR=20.49)

HR-SRCNN (PSNR=22..92)