ECS795P Deep Learning and Computer Vision, 2018

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)

   The first layer uses a 9x9 patch to compute a feature, in the second layer 3 of previous patches which are overlapping (stride = 1) expand the range of information used for one feature presented to layer 3 to 11 x 11 (2 more on each side), the final layer takes 5 of such features of which the original patches also overlap in similar fashion, expanding the range to (11 + 4)= 15^2 = 225 , the total number of pixels involved in reconstructing 1 pixel is 225 pixels.

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

A deep convolution model achieves superior performance over sparse dictionary learning because it is at least as poweful as the sparse-coding method since it can be seen as a generalisation of it. The author of the paper argues throughout the paper how the first layer is a more general form of patch feature extraction, and the second layer leans an n-dimensional mapping and this also holds for the reconstruction filter which perform a more complex form of averaging on the local patches.

Not only can the model be considered analogeous to sparse coding but the total number of parameters learned is greater which yields a more precise model albeit longer to train.

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)

The PSNR is the ratio between the maximum power of the signal and the corrupting noise of the signal, which is computed as a Mean Sqaured error between the ground truth and the prediction. The meaning is a log scale metric which grows with the ratio between the total dynamic range possible and the amount of corruption. Hence a higher PSNR is exponentially better.

GT image:

HR-BI (PSNR=20.497630173285614)

HR-SRCNN (PSNR=21.77124830643141)