ECS795P Deep Learning and Computer Vision, 2018

Course Work 1: Image Super-Resolution Using Deep Learning

**Disclaimer:**

**This coursework has been submitted on multiple files:**

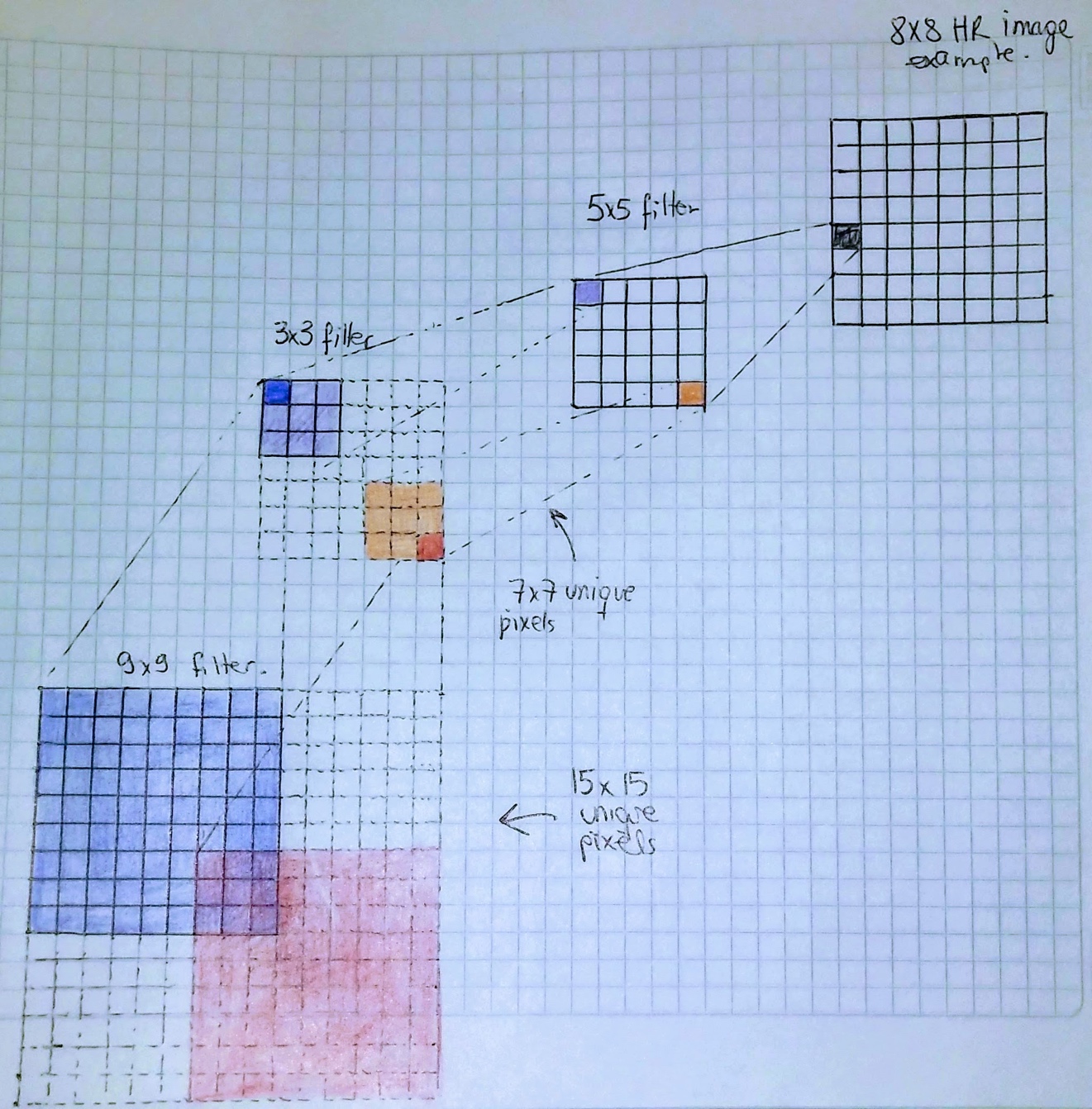
* **The current file answering the questions in the report.**
* **A set of Jupyter Notebooks that go over the exercises in the coursework guidance, as well as the completed code for the SRCNN to work. The reason for this was that it was easier to visualize the outputs of various things while running the scripts outside the Python file itself.**
* **A Python script file with complete code, but not some of the nicer things found in the Jupyter notebooks ☺**

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)

Assuming we have a stride of 1, the number of **unique** pixels used to reconstruct 1 pixel in the HR image will be:

(9 + 3 + 5 – 1 -1)^2 = 15x15 = 225

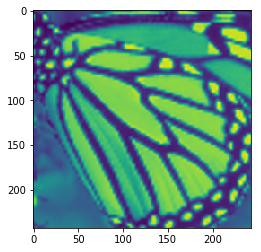
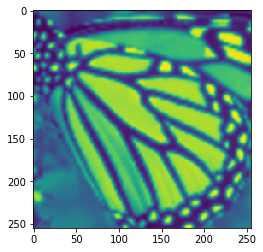
See below an illustration of the LR pixels used to estimate 1 HR pixel.



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

* More pixels are used compared to sparse-coding methods. Example in illustration in previous question, as mentioned in the paper, sparse-coding methods would use less LR pixels, with less neighborhood data used, the estimation should not be able to compete with a more superior CNN model.
* Inputs don’t have a lot of preprocessing involved. Indeed, as seen in the paper, the only pre-processing required is the bicubic interpolation of the LR image in order to match the size of the target HR image. Later in the FSRCNN model the same team actually removed this step, in order to train faster, hence no preprocessing whatsoever (just a deconvolution layer at the end to match HR size).
* All channels at the same time. CNNs can deal with all three channels at the same time, unlike other current (for the time) models.

1. 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 images above are original, LR after bicubic interpolation and LR after SRCNN.



PSNR is a measure of the quality of reconstruction of images (and other signals). This measures the ratio between the value of the signal and the power of distorting noise. Image enhancement or improving the visual quality of a digital image can be subjective. Saying that one method provides a better-quality image could vary from person to person. For this reason, it is necessary to establish quantitative/empirical measures to compare the effects of image enhancement algorithms on image quality.[1]

For further details of filters and images after each convolution please see the jupyter notebooks attached.

[1] <http://www.ni.com/white-paper/13306/en/>