Deep Learning and Computer Vision, 2022

**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.**

* **Receptive Field**

|  |  |  |
| --- | --- | --- |
| Convolution Layers  [Kernel-Padding-Stride] | Image Representation | Receptive Field |
| Conv1 [9-4-1] | Low Resolution Image  [255\*255] | 9 (initial) |
| Conv2 [3-1-1] | High Resolution Image  [255\*255] | 11 |
| Conv3 [5-2-1] | SRCNN High Resolution  [1,1,255,255] | 15 (final) |

Fig2. Table shows the receptive field representation for each convolution layer for the models f1=9, f2=3, f3=5

* **Definition:** Receptive field is defined as the size of the spatial units (number of pixels in this case) that works as the input to the feature vector (acts as the input to the next convolution layer).

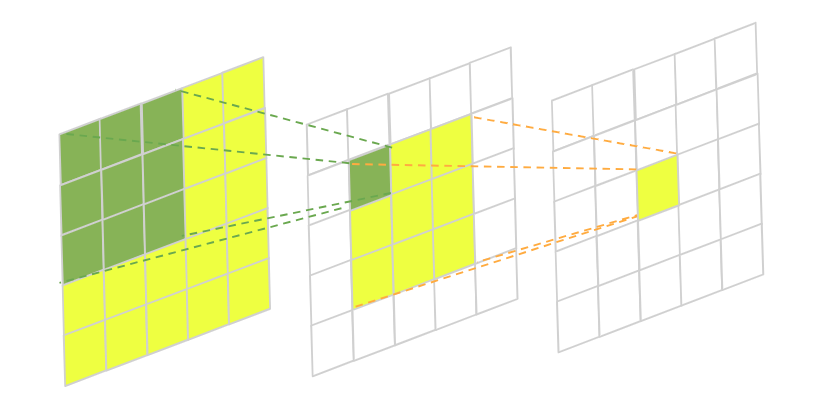
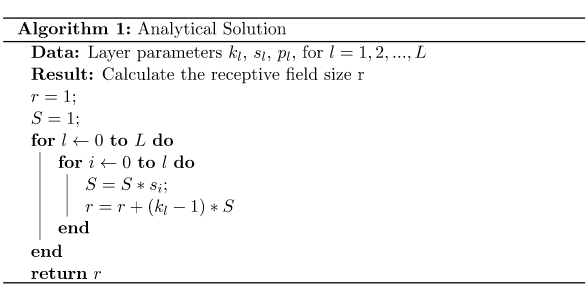
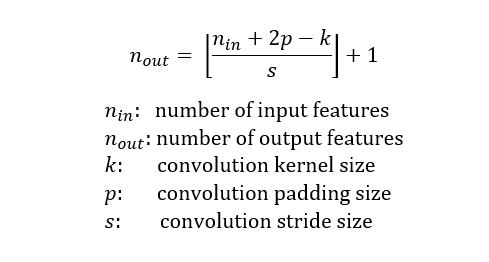
****

Fig1. Represents how the receptive field of the network reduces from layer 1 to layer 3. Layer 1 receptive field (green) and layer 2 receptive field (yellow).

* **Calculation of Receptive Field:**



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

* In the recent times CNN have been successfully applied to the computer vision fields, such as object detection, facial recognition, pedestrian detection and counting and other. Diverse range of factors are to be considered of central importance in this progress such as:
  + Training efficacy on GPUs,
  + Implementing Rectified Linear Unit which make convergence much faster while still presents good quality and
  + The easy access to an abundance of data (like ImageNet [9]) for training larger models. Our method also benefits from these progresses.
* The SRCNN methodology optimizes an end-to-end mapping and works at faster rate. Deep CNN is not only a quantitatively superior method but also denoise the system by removing the noisy patterns.

1. **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) is one of the most popular evaluation metrics that represents the reconstruction quality measurements of lossy transformation like image compression etc.
* In the context of super-resolution, PSNR is defined in terms of maximum pixel value defined by ‘S’ and the mean squared error (MSE) between images.
* Let’s assume the ground truth image I (x, y) with ‘N’ pixels and the reconstructed image I’ (x, y), the PSNR between I and I’ are defined as follows:

PSNR= (1)

|  |
| --- |
| GT |
| HR-Base (PSNR=20.497630181) |
| HR-SRCNN (PSNR=22.92269643) |