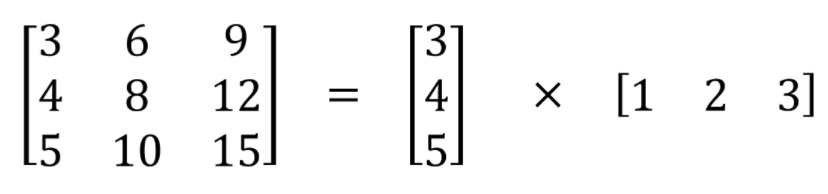
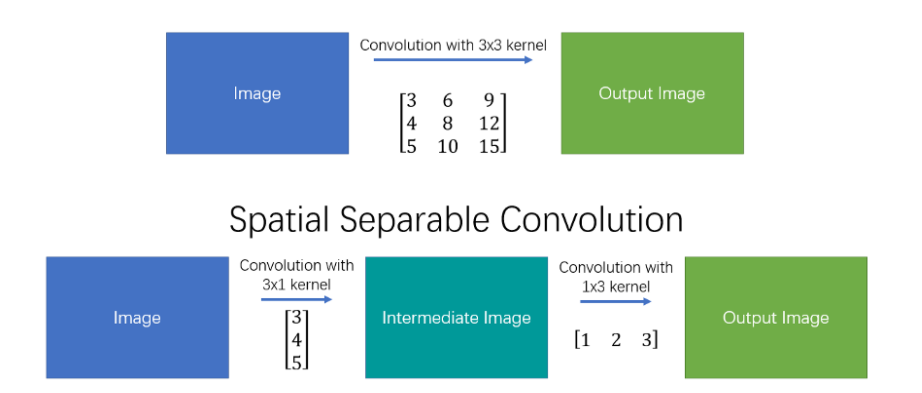
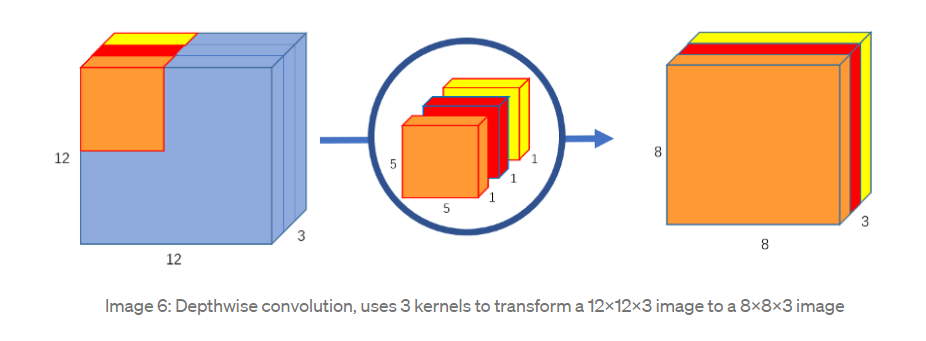
There are basically two types of separable convolutions – **spatial separable convolutions and depth-wise separable convolutions**. In spatial separable convolutions we basically breakdown the kernel into two dimensions and instead of applying it all at once we apply these two dimensions one by one. This reduces the number of multiplications that are to be performed as compared to the conventional convolutional but this method is not applicable for all the images.

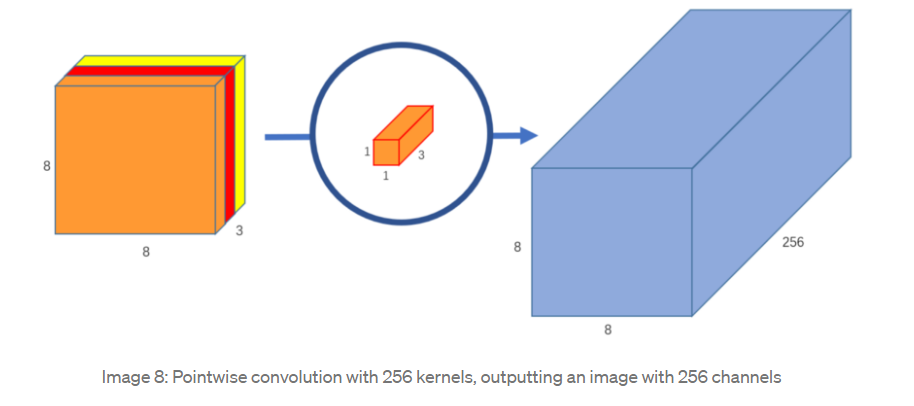




Depth-wise separable convolutions as the name suggests is applied on various layers. Each layer in the depth (number of channels) can be considered as a separate representation of the image in general. As the red channel can be thought of as the amount of ‘redness’ in the image and Blue channel as the amount of ‘blueness’ and so on for a RGB image. A depthwise separable convolution splits a kernel into 2 separate kernels that do two convolutions: the depthwise convolution and the pointwise convolution.



The first conversion is done by convoluting the image with 3 different 5x5x1 filters each of which are applied individually to the image generating 3 different images of dimensions 8x8x1 which are then stacked up to create an image of dimensions 8x8x3. After this we apply point wise convolutions which are nothing but 1x1x(depth of the image) outputted by the above convolution. These point filters, so to say, are applied in a fixed number which represents the number of output depth that is required.



The **main difference** is this: in the normal convolution, we are transforming the image 256 times. And every transformation uses up 5x5x3x8x8=4800 multiplications. In the separable convolution, we only really transform the image once — in the depthwise convolution. Then, we take the transformed image and simply elongate it to 256 channels. Without having to transform the image over and over again, we can save up on computational power.

**Spatial Pyramid Pooling (SPP)** – used for retaining the features from where it originally pooled by using either of the pooling methods (max pooling or average pooling)

**Skip connections** are used to communicate the information learned by lower level images to higher level images. Generally the symmetry helps in better performance.

**Image super resolution** is the process of recovering a High Resolution (HR) image from a given Low Resolution (LR) image. This is done by assuming that the noisy image is a result of the a noisy function that has been applied to the denoised (clear) image. The task is further proceeded by learning the mapping function between the low resolution input and the high resolution ouput.

Motivation

The paper combined both the existing methods for characterizing the rain streaks: simple super imposition and blending and then it was motivated by another paper which assumed that the load on the network could be alleviated by characterizing the rain streaks into different layers and assuming that each layer would consist of rain streaks of same size and density.

