DEEP LEARNING FOR IMAGE SEARCH

WEEK_2

BASIC CONCEPTS:

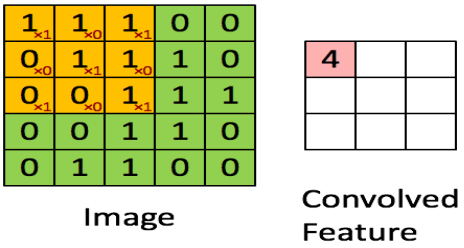
Deep learning is the implementation of neural networks with more than a single hidden layer of neurons .In deep learning, convolutional layers are exceptionally good at finding good features in images to the next layer to form a hierarchy of nonlinear features that grow in complexity (e.g. blobs, edges -> noses, eyes, cheeks -> faces).

A CNN consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalization layers.

PROCESS OF CNN:

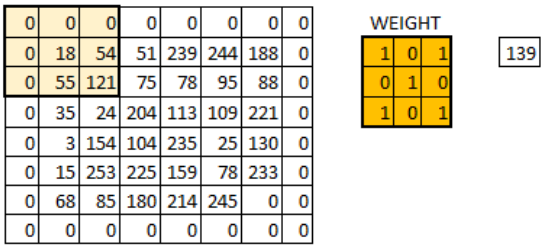
FILTERS:

A filter in a CNN is like a weight matrix with which we multiply a part of the input image to generate a convoluted output. Let's assume we have an image of size 28*28. We randomly assign a filter of size 3*3, which is then multiplied with different 3*3 sections of the image to form what is known as a convoluted output.



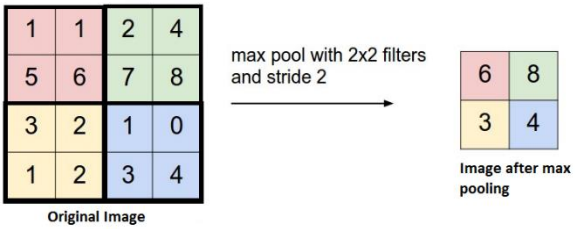
PADDING:

The image is padded with zeroes so that we get the image of same dimension.



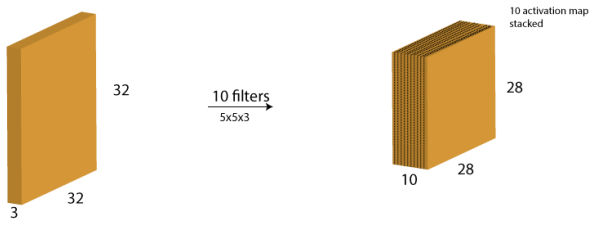
POOLING:

It is common to periodically introduce pooling layers in between the convolution layers. This is basically done to reduce a number of parameters and prevent over-fitting. The most common type of pooling is a pooling layer of filter size(2,2) using the MAX operation.



MULTIPLE FILTERS:

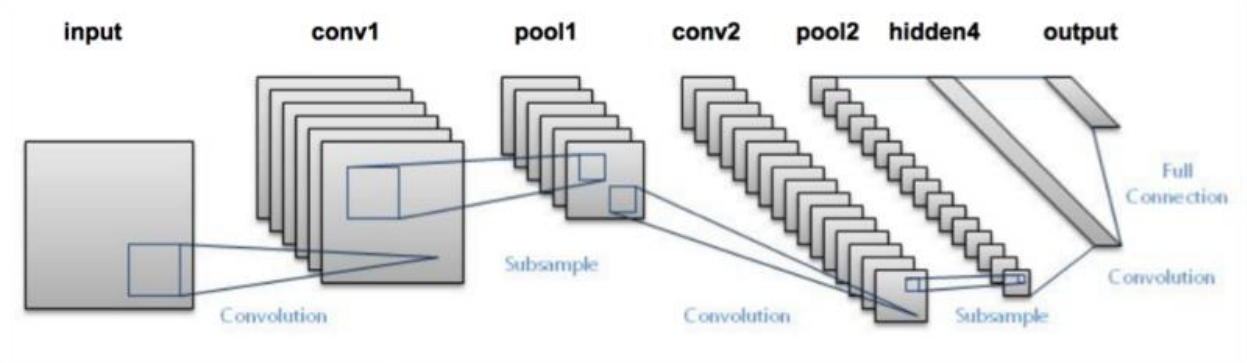
The weight extends to the entire depth of the input image. Therefore, convolution with a single weight matrix would result into a convolved output with a single depth dimension



The weight extends to the entire depth of the input image. Therefore, convolution with a single weight matrix would result into a convolved output with a single depth dimension.

CNN DIAGRAM:

CNN as you can now see is composed of various convolutional and pooling layers. Let's see how the network looks like.



The output is then generated through the output layer and is compared to the output layer for error generation. A loss function is defined in the fully connected output layer to compute the mean square loss. The gradient of error is then calculated.