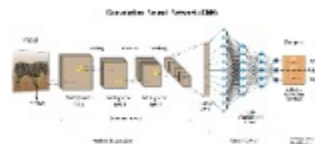


CNN (Convolutional Neural Network)

Type of feed forward neural network whose connectivity structure is inspired by the organization of the animal visual cortex.

It is a type of NN to evaluate visual information. These networks can handle a wide range of tasks images, videos, sounds and other media.



In CNN we have input layer, numerous hidden layers and millions of parameters, output layer like an ANN but it uses convolution and pooling processes to sub-sample the given input before applying an activation function.

All layers consist of hidden neurons that connect partially, with a completely connected layer at the end producing the output. The output shape is similar to the size of the input image.

Convolution is the process of combining two functions to produce the output of the other function. The input image is convoluted with the application of filters in CNNs, which results in a feature map.

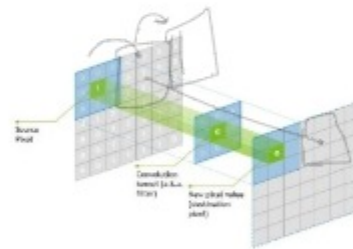
Feature map is the matrix in which one feature of the image has been captured by the convolution process.

Filters are weights and biases that are randomly generated vectors in the network. Instead of having individual weights and biases for each neuron, CNN uses the same weights and biases for all neurons. Many filters are created, each of which catches a different aspect from the input. Filters are also known as kernels.

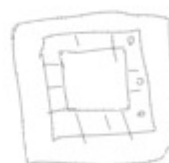
1. Convolutional layer is a fundamental part of CNN.
2. These layers involves input vectors, which is an image, filters (kernels or feature detectors), output vectors, often referred as feature maps.
3. The image as an input goes through abstraction into a feature map/activation map.
4. The process is called convolution, which enables detection of more complex features within the image.
5. ReLU is also applied within the layers to introduce non-linearity into a network.
6. After convolution pooling is applied to reduce the spatial dimensions of the feature maps, which leads to a more manageable output volume.

Feature map = input image x feature detector (filter)

A convolution is a grouping function in mathematics



In the above image we can see the convolution process in which the position of the kernel's centre element is above the source pixel. Then replace the source pixel with a weighted sum of itself and its neighboring pixels.

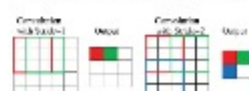


Padding and stride
They can be used to increase or decrease the dimensions of the input/output vectors.

The term describes how many pixels a CNN kernel adds to an image during processing.

If you set the padding 0 then every added pixel value will be zero. If we set the zero padding to one a one pixel border with a zero value will surround the image.

Stride determines how the filter convolves over the input matrix which means how many pixels shift. When stride is 1 the filter moves across one pixel at a time.



Pooling

It reduces the spatial representation of image by reducing the number of parameters and computations in the network.

The pooling layer treats each feature map separately.

There are two types of pooling max and average pooling.

Max pooling chooses the most significant element from the feature map. It is most commonly used. If there is no pooling the output has the same resolution as the input. It also prevents overfitting.