Convolutional Neural Networks

CNN’s are neural network that have the convolution operation instead of normal matrix multiplication in at least one of the layers.

**What is Convolution 🡪 Why use convolution 🡪 Pooling**

The best architecture for CNN’s are announced every week or so.

Text

Description automatically generated

Or A simpler way to write is just use an asterisk.

Text

Description automatically generated with medium confidence

X is the Input

W is the kernel

Output (s) is sometimes called the feature map

Discrete convolution: A picture containing text, watch, clock

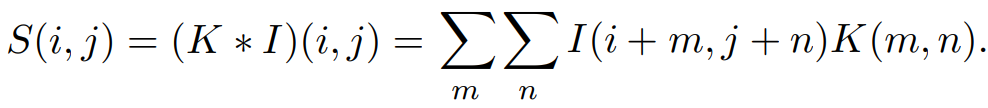
Description automatically generated

2-dimensional convolution also called as tensors:

Text, letter

Description automatically generated

Cross Correlation: Convolution used in a lot of machine learning libraries



Sparse Connection

One input can affect many outputs. Sparse connection is when the kernel is smaller than the input array.

Receptive Field

Output unit affected by input units. The receptive field of the units in the deeper layers of a convolutional network is larger than the receptive field of the units in the shallow layers. This effect increases if the network includes architectural features like strided convolution or pooling. This means that even though direct connections in a convolutional net are very sparse, units in the deeper layers can be indirectly connected to all or most of the input image

Convolution is more efficient than matrix multiplication.

Equivariance function is when the output changes in the same way like input.

Detector stage is called as the stage where the data is passe through the non linear activation function.

Invariance to translation means if there is a small change in the input, the values of the pooled output do not change. Use case: To determine if something is actually there than to find its exact position.

Pooling is used to decrease the computational power require to process the data by reducing the dimensionality of the images.

2 types of pooling: max pooling and average pooling.

Max pooling returns the max value from the portion of image covered by kernel.

Average pooling returns the average of all values from the portion of image covered by kernel.

Max pooling acts as noise suppressant.

Performance: Max pooling > Average pooling

Generally, the output of CNN has less dimensions than the input image. It is called Valid padding.

Same padding is applied to increase the dimensionality of output from the input image.

Or to keep the dimensions of output and input same.