# Weight Sharing in CNN, Fully Connected NN vs CNN

Ms. Khushi Upadhyay

Weight Sharing in CNN

Neural networks have become a popular method for solving complex problems and have revolutionized various industries. Two popular types of neural networks are Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). Despite their differences, both of them share a common concept called weight sharing.In this article we will explore weight sharing in CNN and understand how they use it to find patterns in the input.

Multiple Input Multiple Output (MIMO)
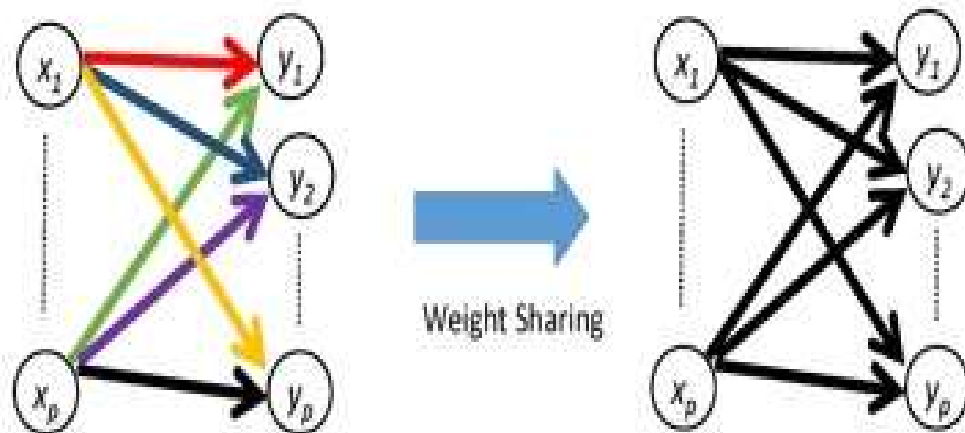
Weight Sharing

The different colours correspond to different weights, whereas the same colour illustrates weight sharing.
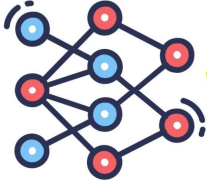
Weight sharing is not new and has proven very useful in adaptive learning, yet it has not been adopted in adaptive signal processing. On the other hand, this type of constraint on weights features prominently in neural networks, such as convolutional neural networks (CNNs) and feedforward neural networks. In practice, weight sharing constrains all 2D input samples ,X[i,j] , to be weighted by the *same* set of coefficients, F_{kl} to determine the output, Y[i,j]. The best known weight sharing approach in CNNs is the kernel image convolution operation, given by

$$\mathbf{Y}[i,j](n) = \sum_{k=1}^{K}\sum_{l=1}^{L}\mathbf{X}[i+k-1, j+l-1](n)\mathbf{F}_{kl}(n) \qquad \forall i,j$$

$$\text{subject to} KL \ll \text{card}(X(n))$$

where K and L are respectively the width and height of the 2D filter, $\mathbf{F}$, and card(X(n)) denotes the input size or the number of input samples at a time instant n. Typically, each iteration index n corresponds to one image presented to a CNN. Observe that

   i)   the kernel convolution captures the cross-correlation between the filter and the input samples;

   ii)   ii) KL≪card(X(n)) means that the total number of input samples exceeds significantly that of the weights at a iteration index n - thus implying weight sharing.

The physical interpretation of weight sharing is that the learning algorithm is forced to detect a common 'local' feature across all input samples . For example, the scanning of all image pixels with a kernel with the same set of coefficients at iteration index n enables the CNN to detect the commonality between similar patterns scattered throughout the image, as the same 'local' feature rather as several different features. Moreover, weight sharing leads to other advantages such as:

1.Lower computational cost as the number of free parameters is reduced.
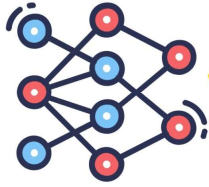
2.Fewer parameters mean the subsequent learning algorithms are easier to tune and faster converging.

3.Reduced risk of overfitting the data.

Fully Connected NN vs CNN

Fully connected neural network

- A fully connected neural network consists of a series of fully connected layers that connect every neuron in one layer to every neuron in the other layer.

- The major advantage of fully connected networks is that they are "structure agnostic" i.e. there are no special assumptions needed to be made about the input.

- While being structure agnostic makes fully connected networks very broadly applicable, such networks do tend to have weaker performance than special-purpose networks tuned to the structure of a problem space.
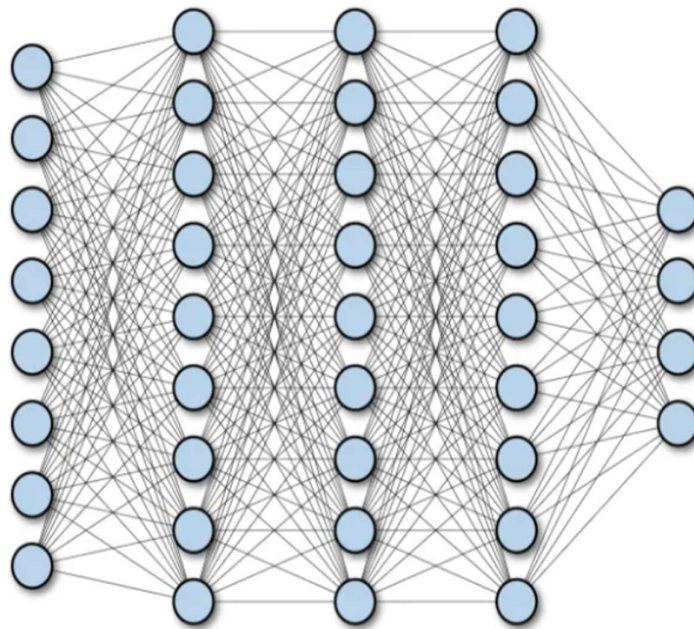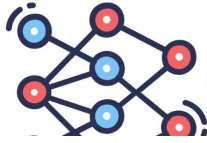
Convolutional Neural Network

- CNN architectures make the explicit assumption that the **inputs are images**, which allows encoding certain properties into the model architecture.

- A simple CNN is a sequence of layers, and every layer of a CNN transforms one volume of activations to another through a differentiable function. Three main types of layers are used to build CNN architecture: Convolutional Layer, Pooling Layer, and Fully-Connected Layer.