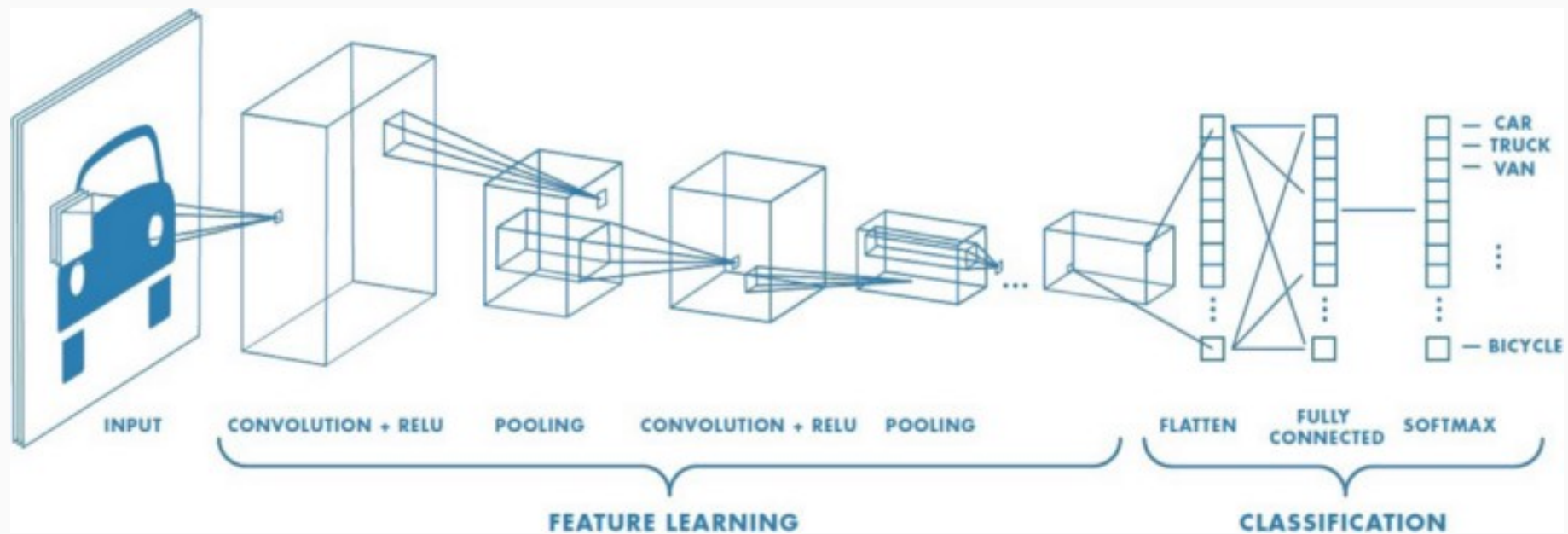


Convolutional Neural Networks ConvNet/CNN

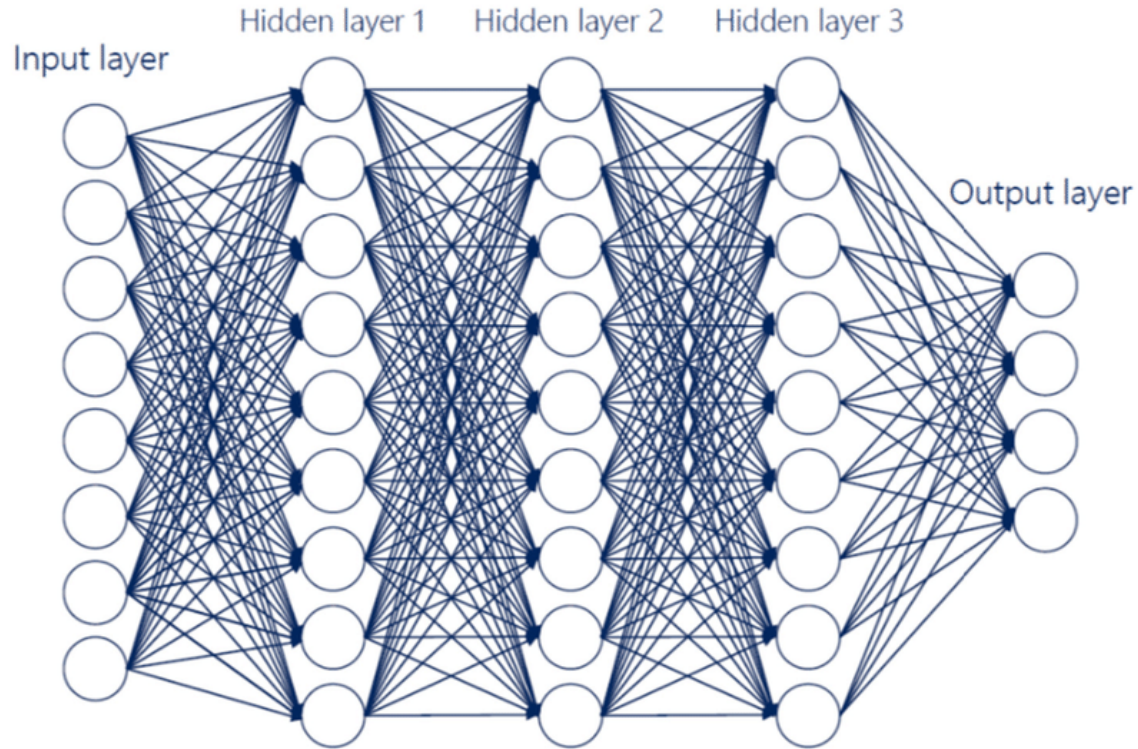
Motivation and Architecture



Definition

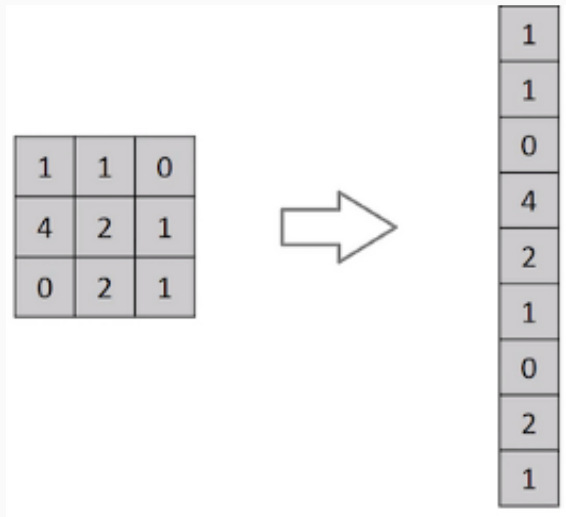
A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.

Feed Forward Neural Network - FFNN



CNN over FFNN

How to feed an image to FFNN?



CNN

Is able to successfully capture the **Spatial** and **Temporal** dependencies in an image through the application of relevant filters.

Spatial dependency - a pixel's value is influenced by nearby pixel's value in image.

Temporal dependency - dependencies of images taken from different moments of time. Correlations between the images are often used to monitor the dynamic changes of the object.

Efficiency

Things would get computationally intensive once the images reach dimensions, say 8K (7680×4320).

The role of the CNN is to reduce the images into a form which is easier to process, without losing features which are critical for getting a good prediction.

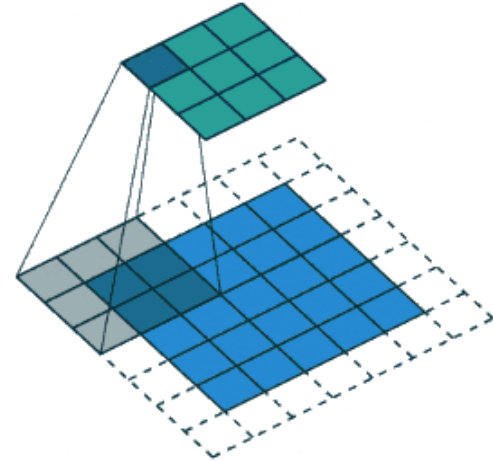
This is not only computationally efficient for learning but also is scalable to massive datasets.

Convolution

Is aimed to extract the high-level features such as edges, from the input image.

Conventionally, the first layer of CNN is responsible for capturing the Low-Level features such as edges, color, gradient orientation, etc.

With added layers, the architecture adapts to the High-Level features as well, giving us a network which has a full understanding of images.



Pooling

Is responsible for reducing the spatial size of the Convolved Feature (dimensionality reduction).

It is useful for extracting dominant features which are rotational and positional invariant, thus maintaining the process of effectively training of the model.

3.0	3.0	3.0
3.0	3.0	3.0
3.0	2.0	3.0

3	3	2	1	0
0	0	1	3	1
3	1	2	2	3
2	0	0	2	2
2	0	0	0	1

Flatten

Adding a Fully-Connected layer is a (usually) cheap way of learning non-linear combinations of the high-level features as represented by the output of the convolutional layer.

The Fully-Connected layer is learning a possibly non-linear function in that space to map the extracted features to a linear space.

Which then can be used for a specific problem say
Classification

Thank You