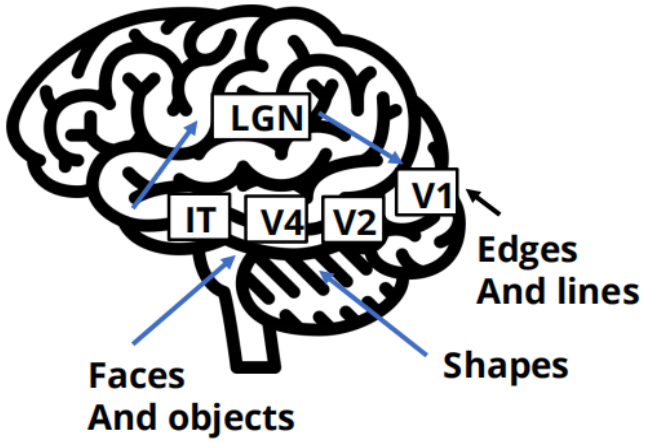
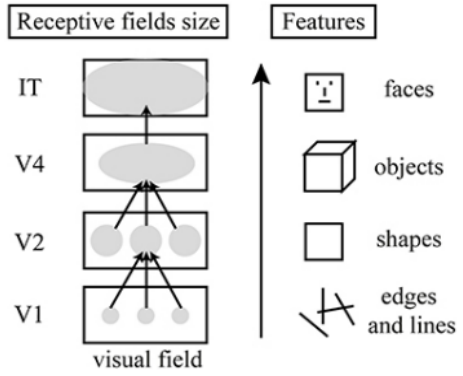
Convolutional Neural Net (CNN)

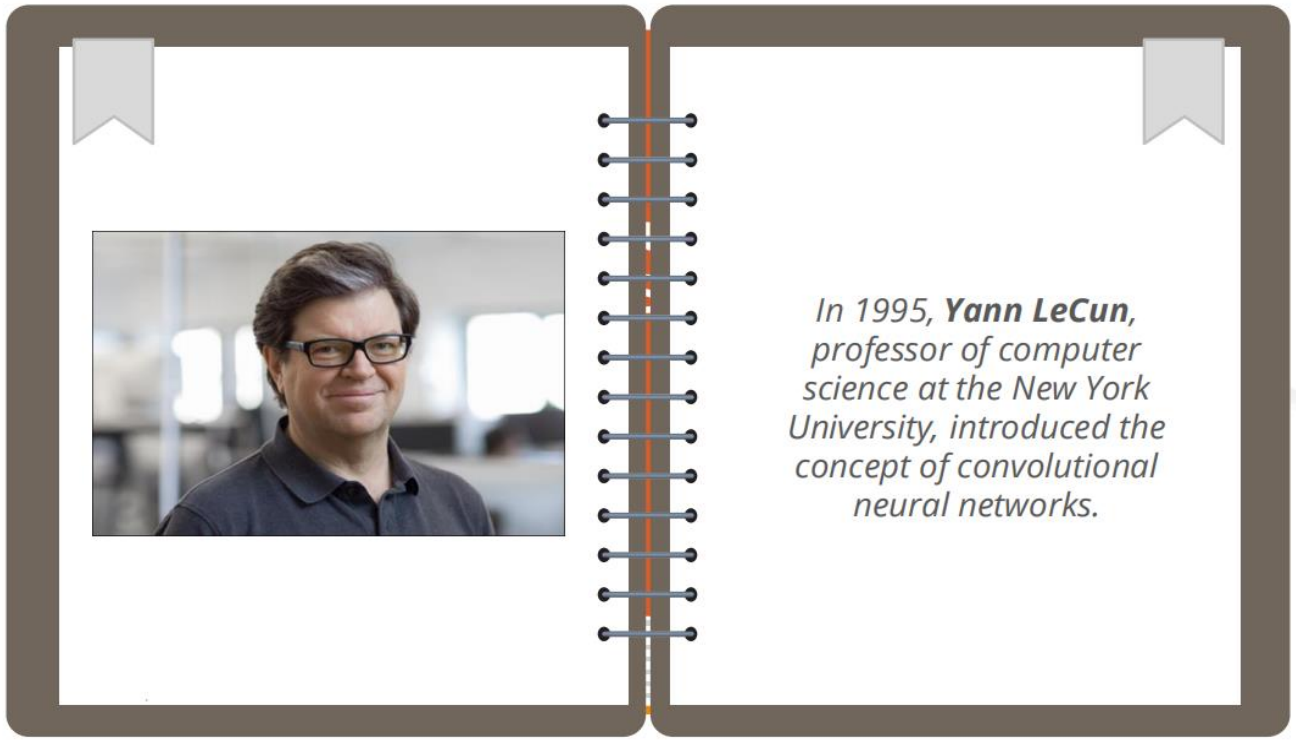
Human Visual and CNN





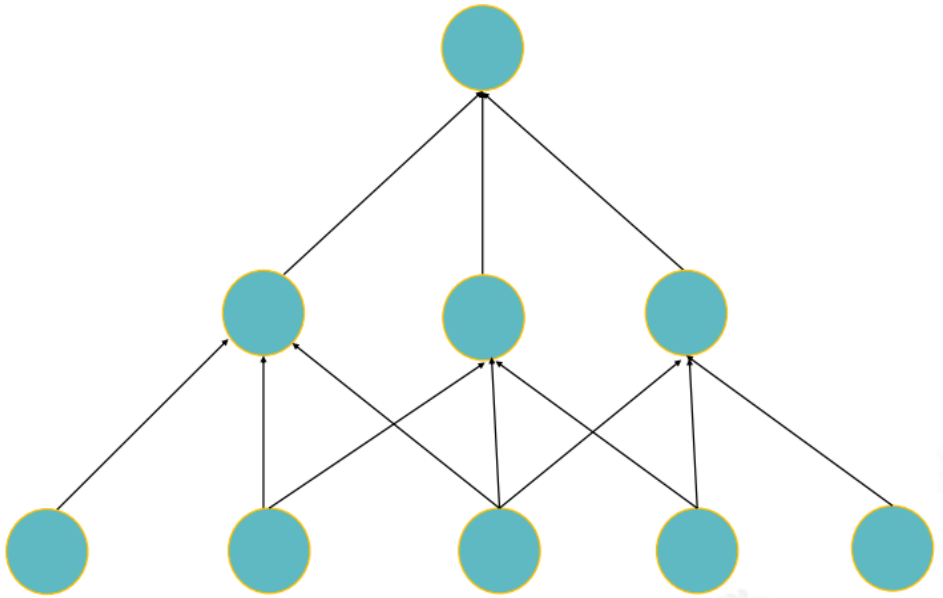
- The idea of CNNs was neurobiologically motivated by the findings of locally-sensitive and orientation-selective nerve cells in the visual cortex.
- Inventors of CNN designed a network structure that implicitly extracts relevant features.
- Convolutional Neural Networks are a special kind of multilayer neural networks.

History of CNN



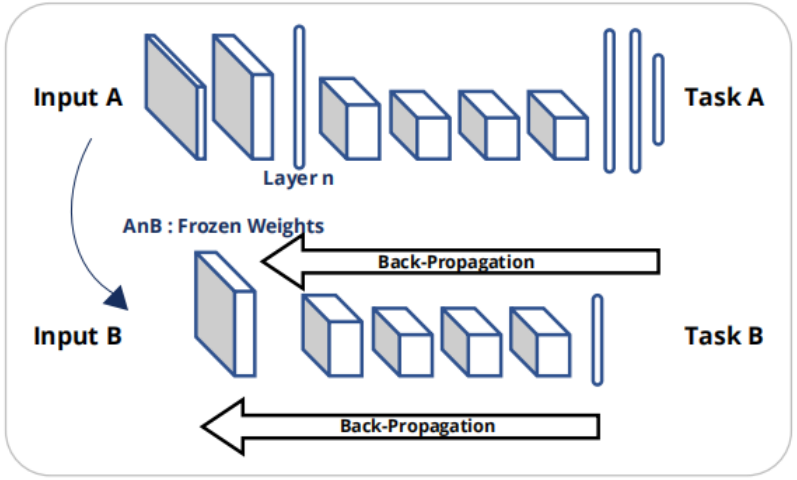
The Core Idea Behind CNN

- **Local Connections**
- Layering
- Spatial Invariance

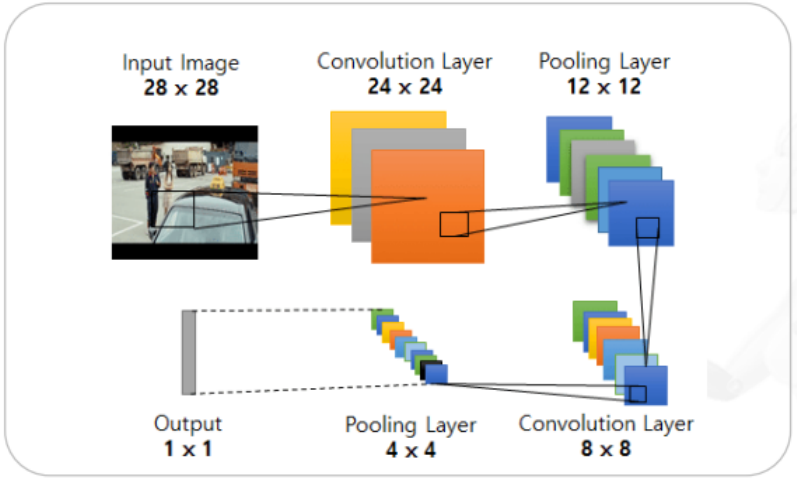


Represent how each set of neurons in a cluster is connected to each other, which in turn represents a set of features

CNN Applications



Transfer Learning and Fine Tuning



Feature Extraction

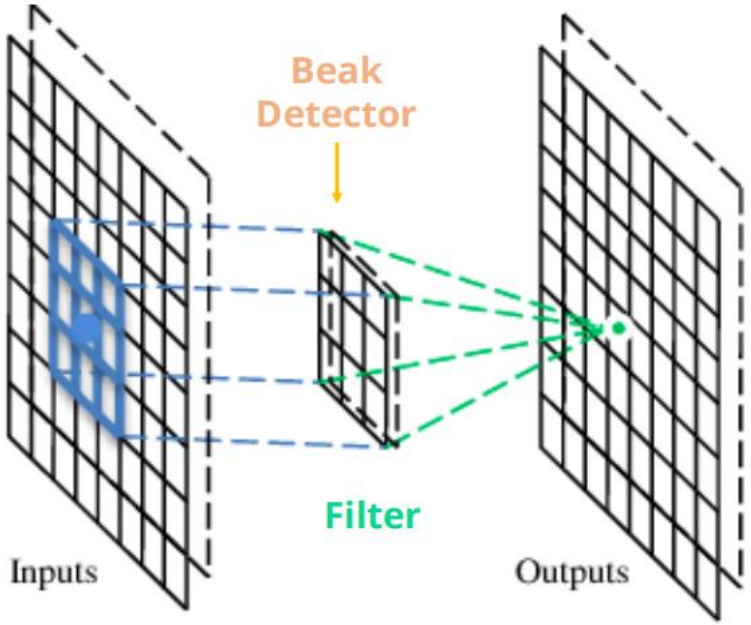
Learning an Image

CNN focuses on smaller and specific patterns than the whole image.

It's convenient and effective to represent a smaller region with fewer parameters, thereby reducing computational complexity.

The Convolutional Layer

A CNN is a neural network with convolutional layers (and other layers). A convolutional layer has several filters that perform the convolution operation.



The Convolution Operation

Consider a 6x6 image convolved with 3x3 filter(s) to give an output of size 4x4.

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

*

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

-1	1	-1
-1	1	-1
-1	1	-1

Filter 2

⋮ ⋮

Each filter detects a small pattern (3 x 3)

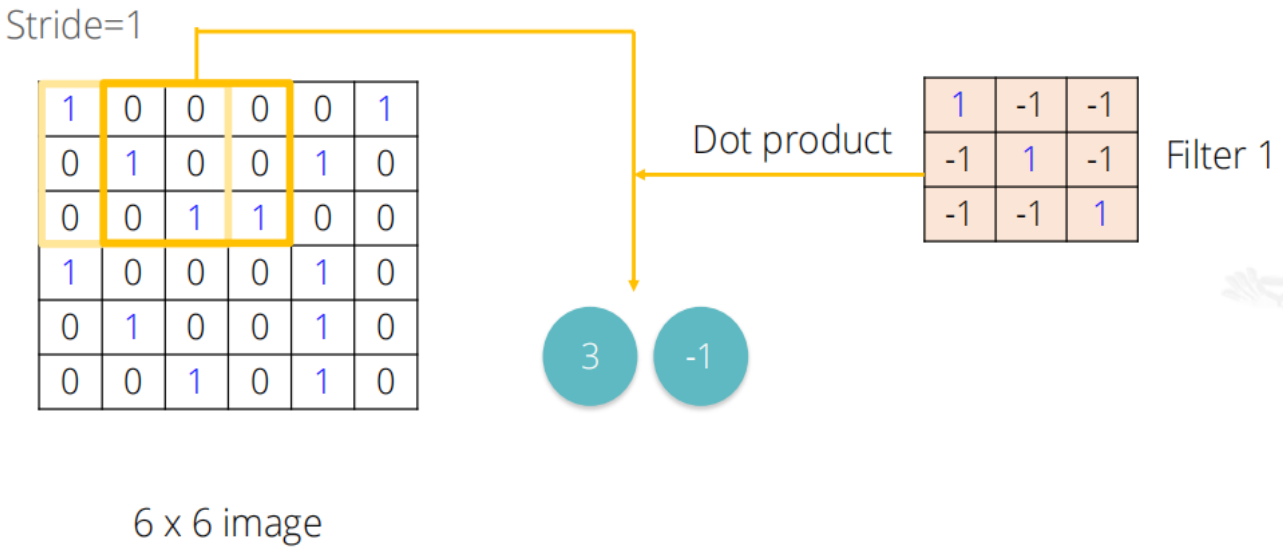
Note: Filters can be considered network parameters to be learned.

Straight Line Filters & Slanted Line Filters

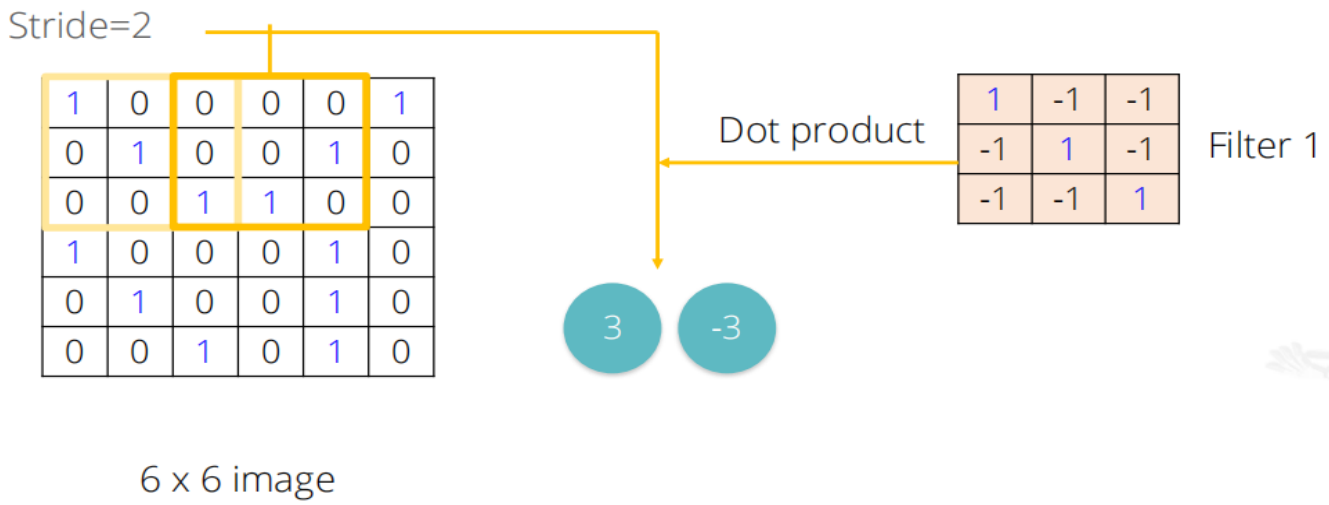
- **Straight Line Filters**
- **Vertical Line Filters:** Detect vertical edges or lines in an image. These filters respond strongly to vertical transitions from light to dark or dark to light.
- **Horizontal Line Filters:** Detect horizontal edges or lines in an image. These filters respond strongly to horizontal transitions from light to dark or dark to light.
- **Slanted Line Filters**
- **Slanted/Diagonal Line Filters:** Detect diagonal or slanted edges in an image. These filters respond to transitions that occur at an angle rather than strictly horizontal or vertical.
- **Purpose**
- **Edge Detection:** Both straight and slanted line filters help the CNN identify edges and boundaries within an image, which are crucial for understanding the structure and shapes within the scene.
- **Feature Extraction:** These filters enable the network to capture important features such as textures, contours, and shapes, contributing to the overall feature map used for classification or other tasks.

The Convolution Operation

Shift the filter around the input matrix (commonly known as stride) once a convolved output is achieved.

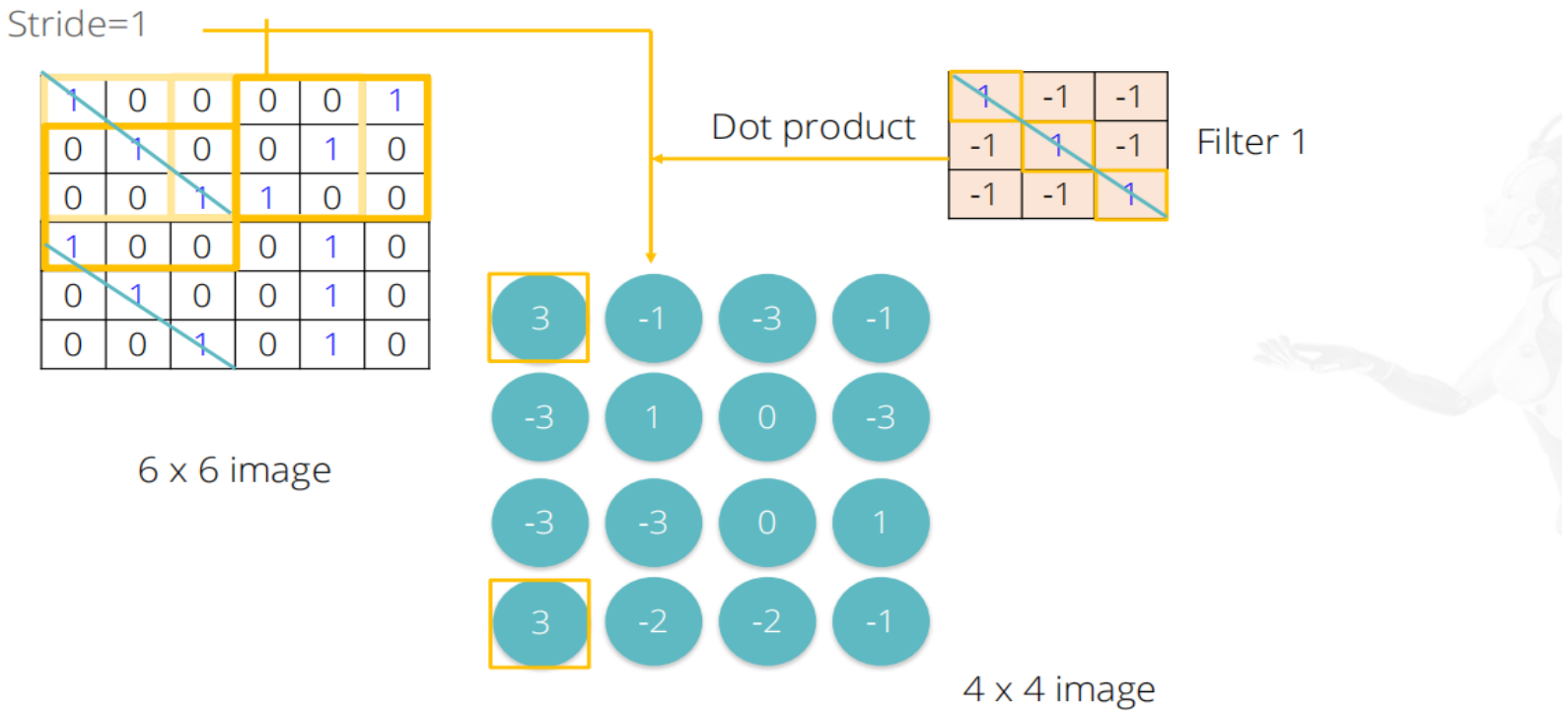


The Convolution Operation



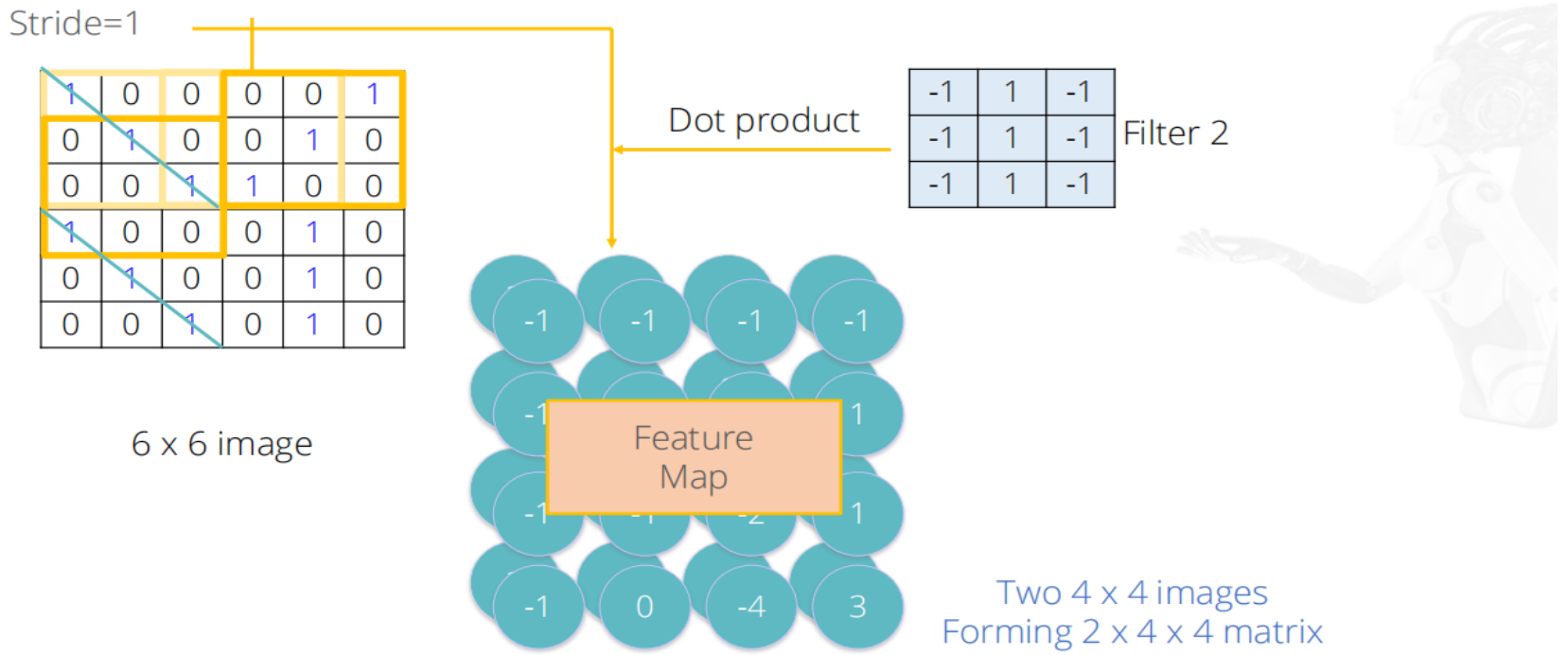
Note: If you change the stride size, the convolved output will vary (only outputting intense pixels).

The Convolution Operation



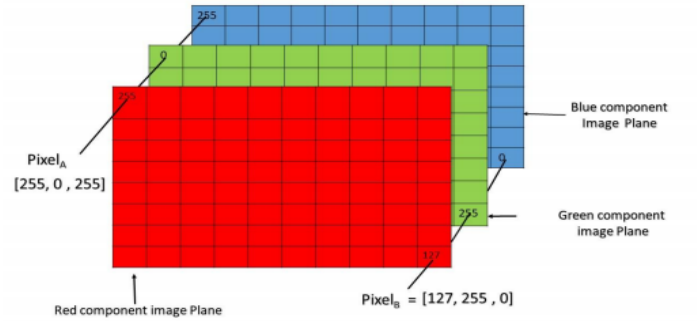
The Convolution Operation

The convolution operation gets repeated for each filter resulting in a feature map.

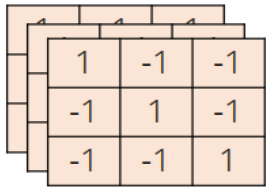


RGB Images

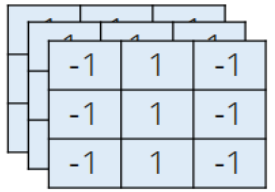
When RGB image is used as input to CNN, the depth of filter is always equal to the depth of image (3 in case of RGB).



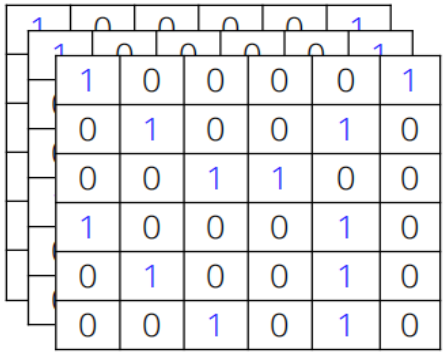
Pixel of an RGB image are formed from the corresponding pixel of the three component images



3-dimensional filter 1

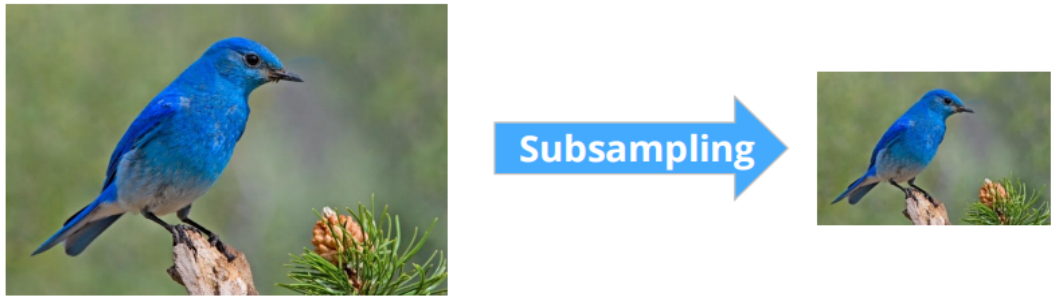


3-dimensional filter 2



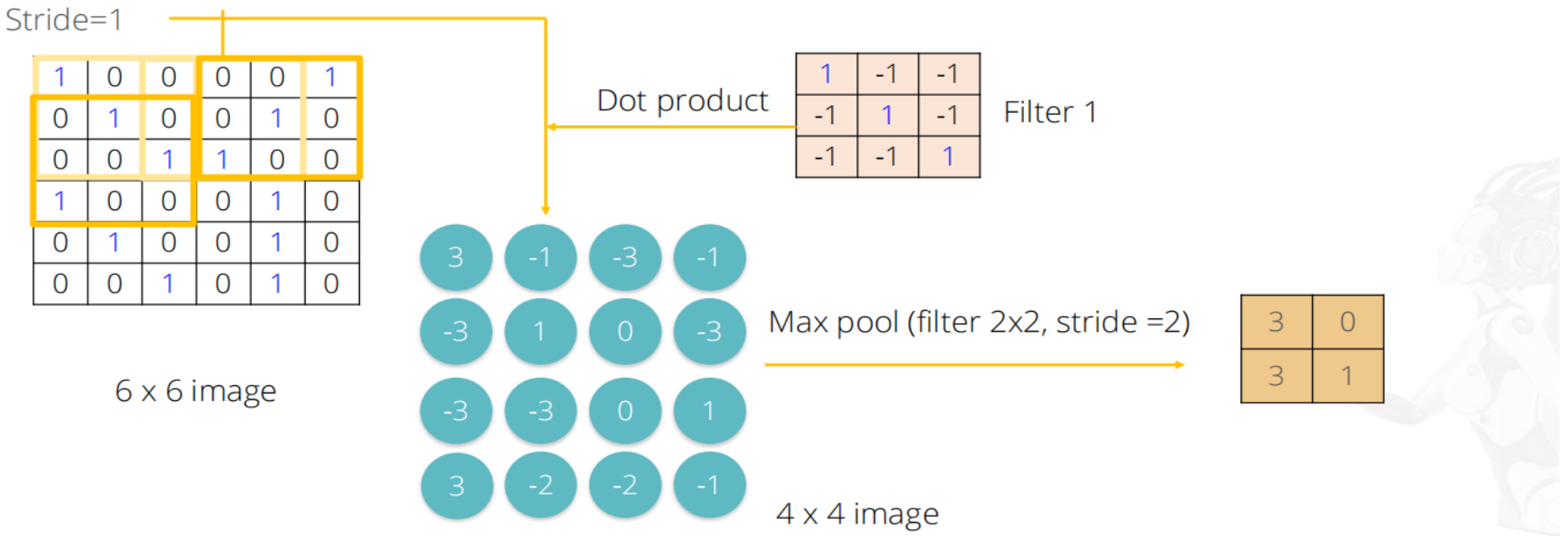
Pooling Layer

The pooling layer gradually reduces the spatial size of each matrix within the feature map such that the amount of parameters and computation is reduced in the network.



Note: The most commonly used pooling approach is max pooling.

Pooling Layer



Note: The most commonly used pooling approach is max pooling.

The CNN Architecture

The reduced image from these layers (convolution + pooling) is then passed through the activation function.