

# CNN

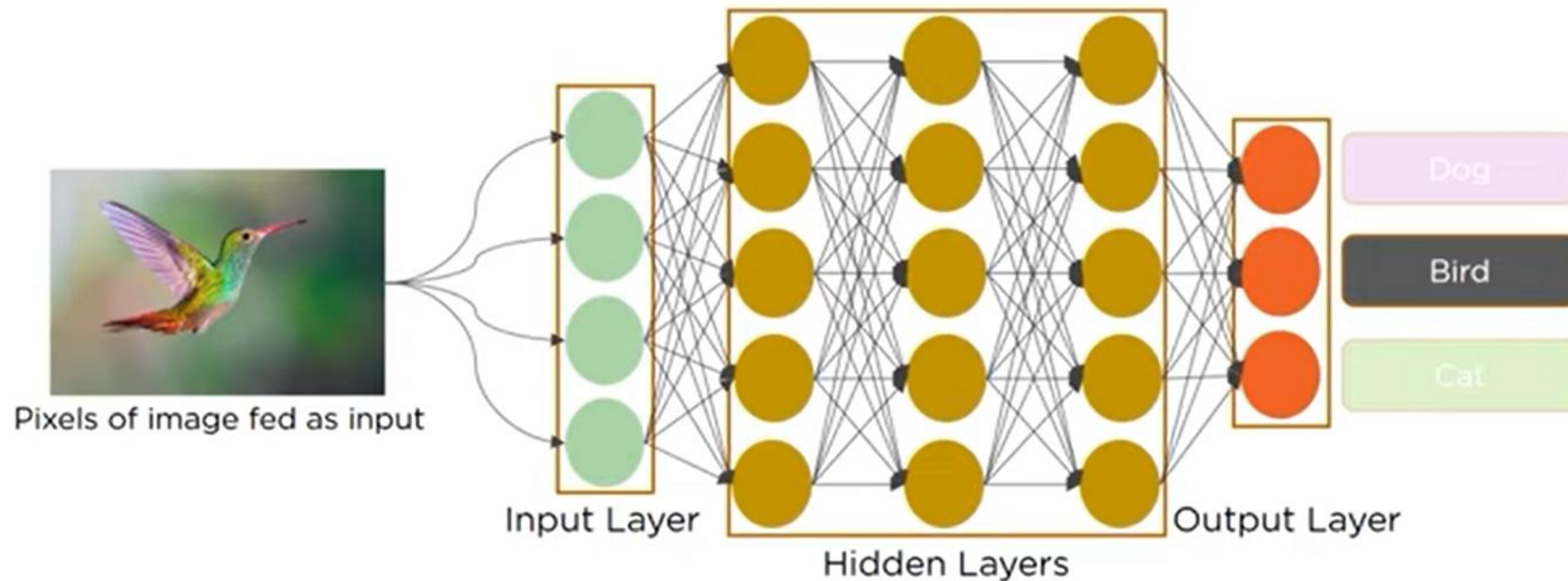
**Zahoor Tanoli (PhD)**

**CUI Attock**

# How image recognition works?

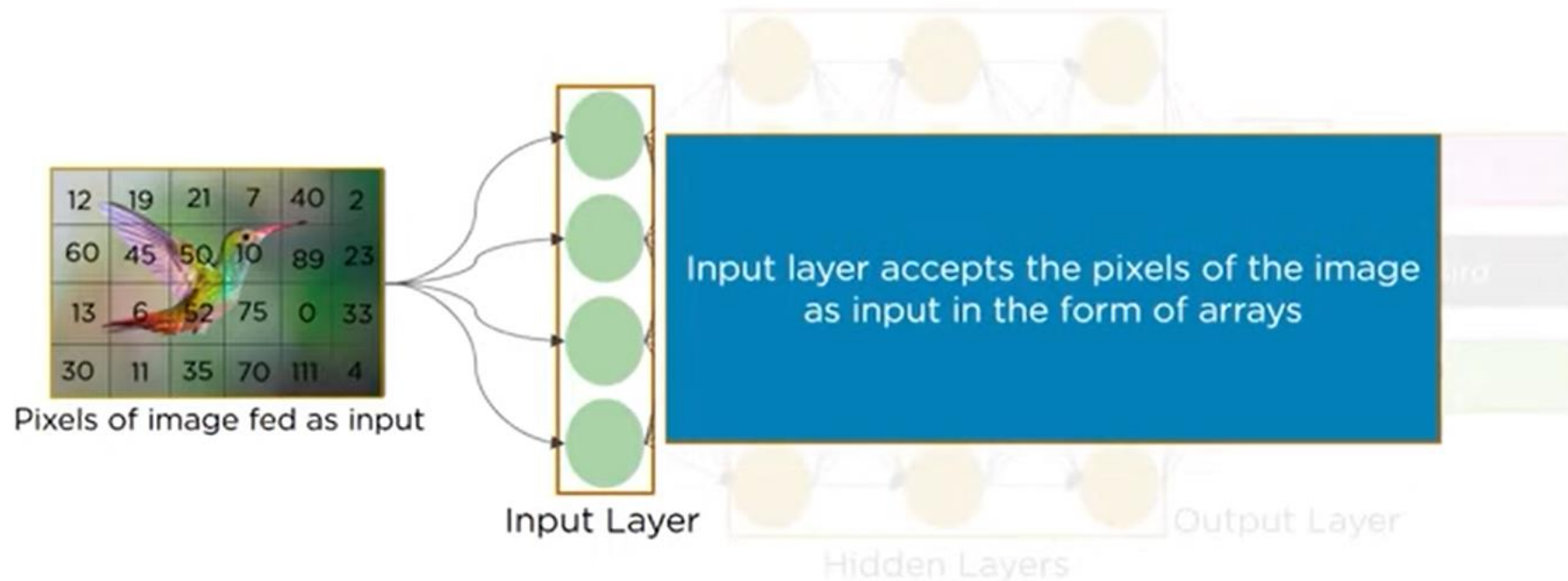
Do you know how Deep Learning recognizes the objects in an image?

It does it using a Convolution Neural Network



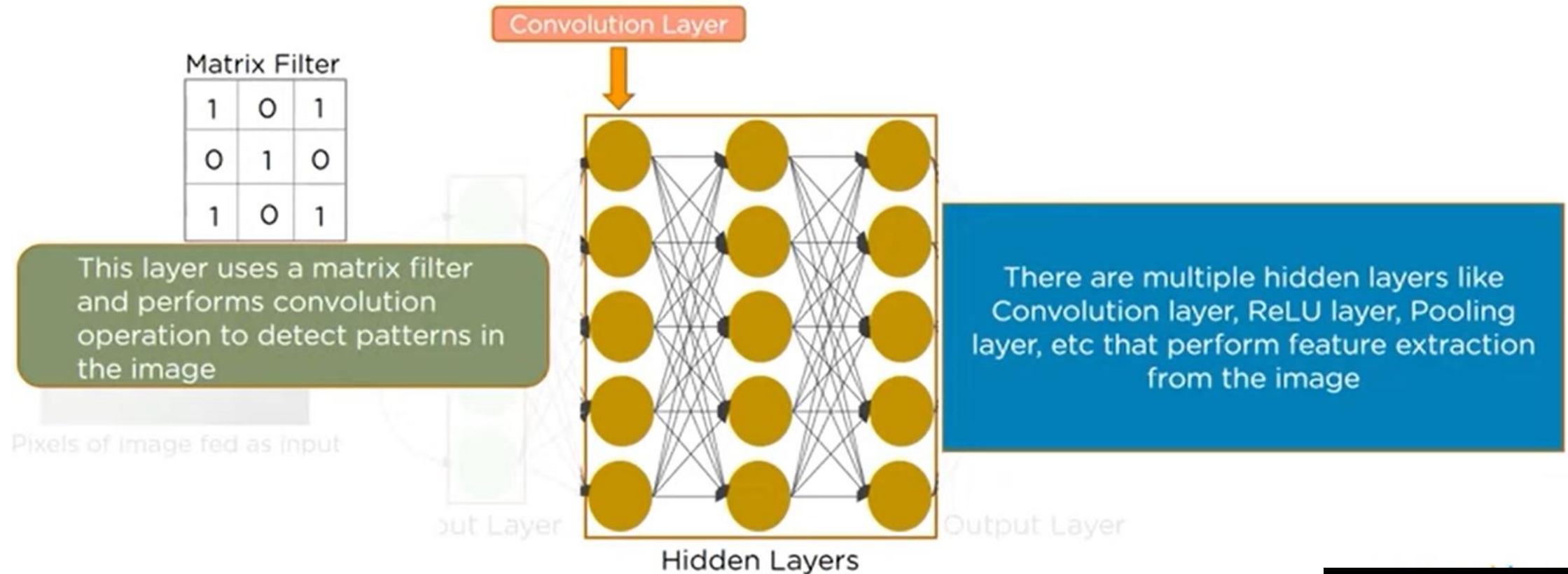
# How image recognition works?

Let's see how CNN identifies the image of a bird



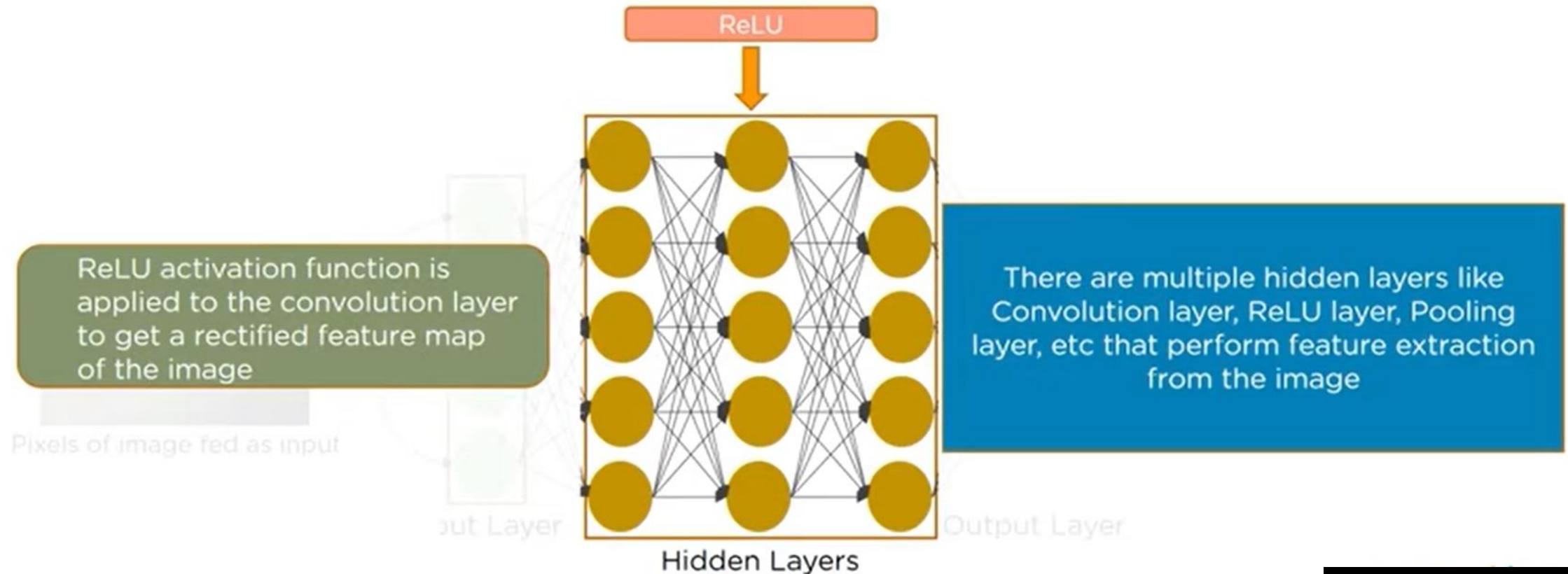
# How image recognition works?

Let's see how CNN identifies the image of a bird



# How image recognition works?

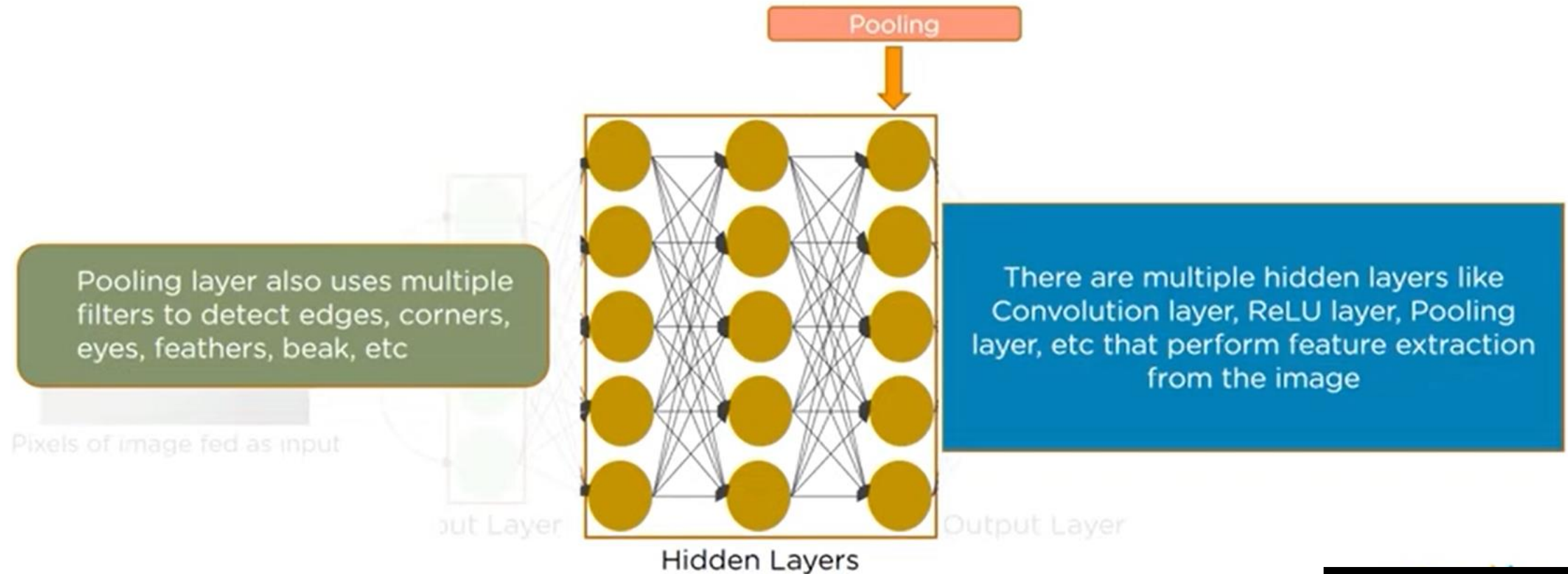
Let's see how CNN identifies the image of a bird





# How image recognition works?

Let's see how CNN identifies the image of a bird



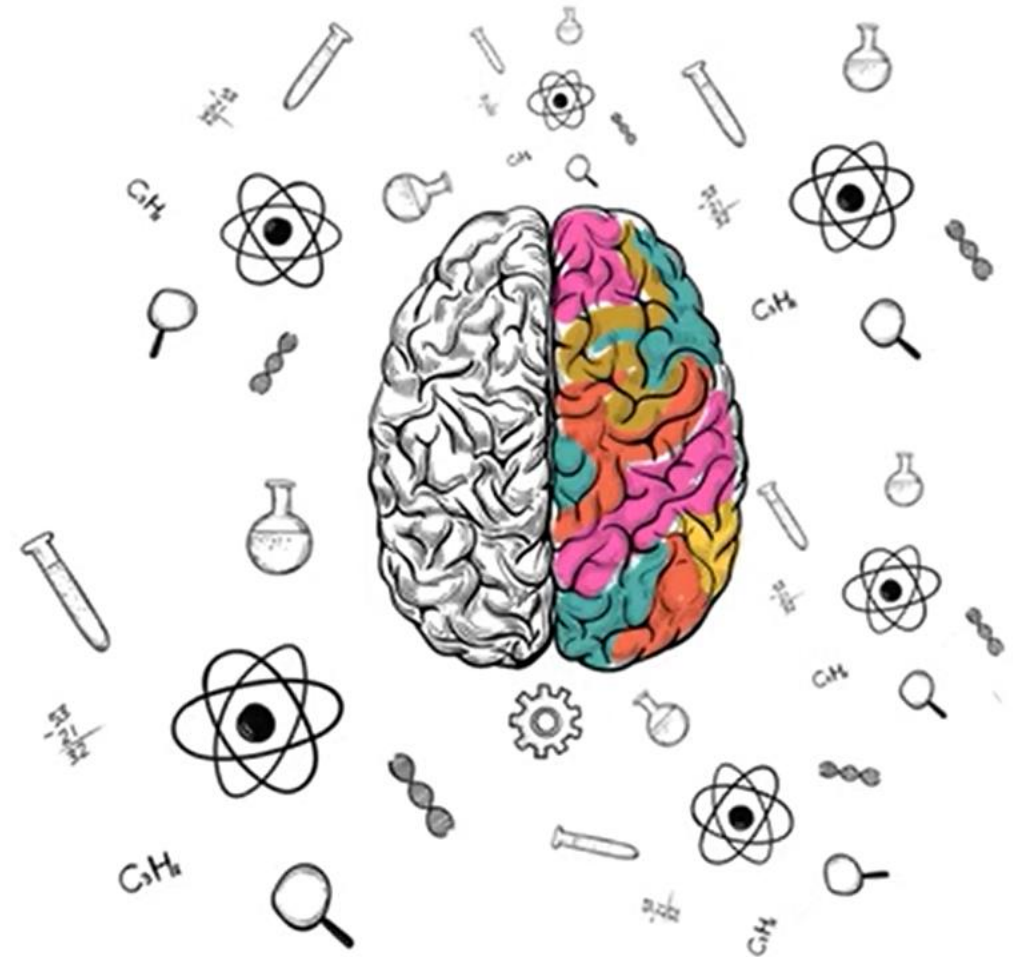
# How image recognition works?

Let's see how CNN identifies the image of a bird



# What's in it for you?

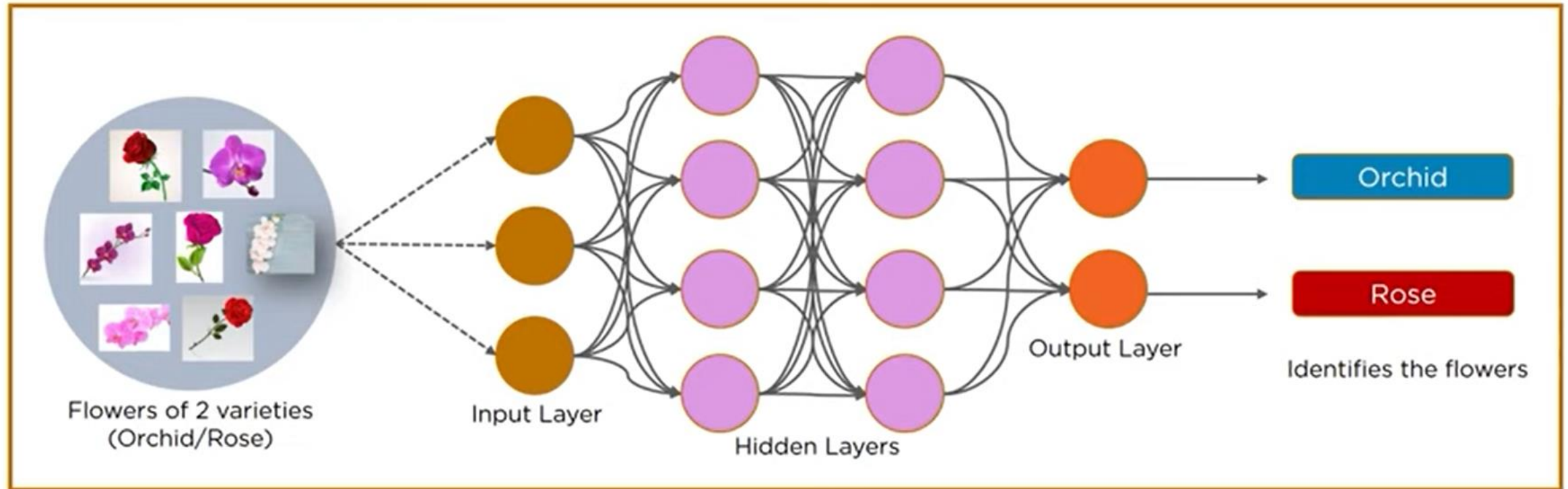
- ▶ Introduction to CNN
- ▶ What is Convolution neural network?
- ▶ How CNN recognizes images?
- ▶ Layers in convolution neural network
- ▶ Use case implementation using CNN





# What is a Convolution Neural Network?

CNN is a feed forward neural network that is generally used to analyze visual images by processing data with grid like topology. A CNN is also known as a “*ConvNet*”



# What is a Convolution Neural Network?

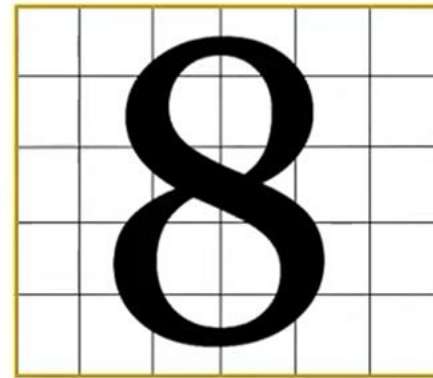
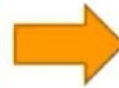
CNN is a feed forward neural network that is generally used to analyze visual images by processing data with grid like topology. A CNN is also known as a “*ConvNet*”

Convolution operation forms the basis of any Convolution Neural Network

In CNN, every image is represented in the form of arrays of pixel values



Real Image of the digit 8



Represented in the form of an array



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

Digit 8 represented in the form of pixels of 0's and 1's

# What is a Convolution Neural Network?

Let's understand the convolution operation using 2 matrices a and b of 1 dimension

a = [5, 3, 7, 5, 9, 7]  
b = [1, 2, 3]

Matrix a and b

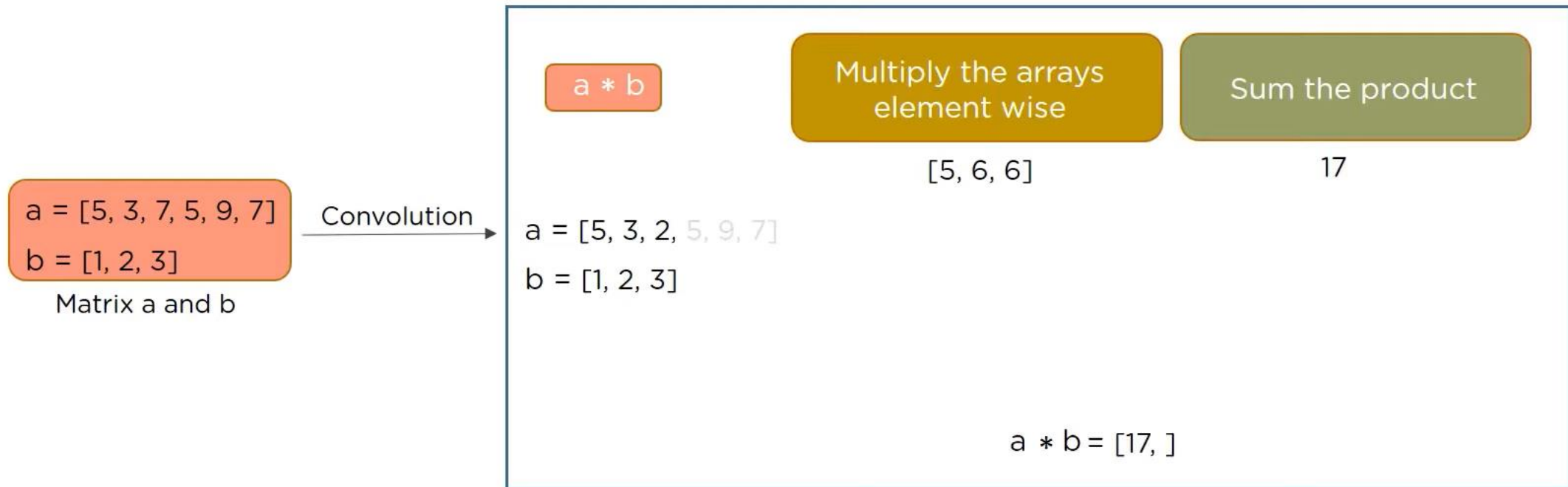
Convolution →

a \* b

a = [5, 3, 2, 5, 9, 7]  
b = [1, 2, 3]

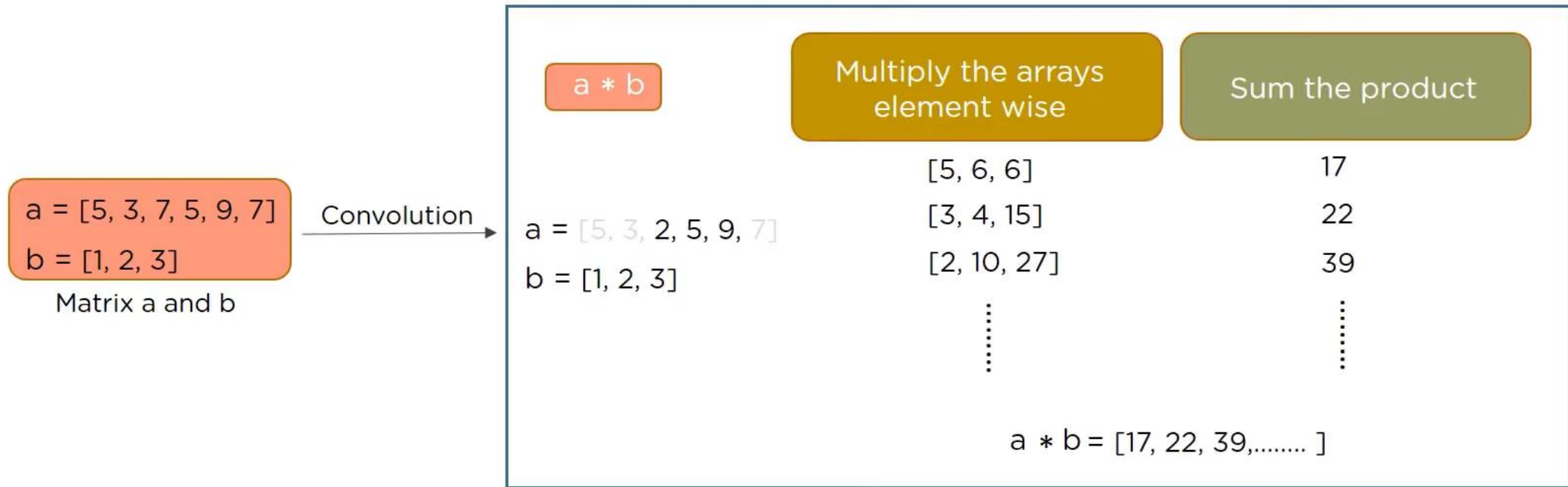
# What is a Convolution Neural Network?

Let's understand the convolution operation using 2 matrices a and b of 1 dimension



# What is a Convolution Neural Network?

Let's understand the convolution operation using 2 matrices a and b of 1 dimension





# How CNN recognizes images?

Consider the following 2 images:

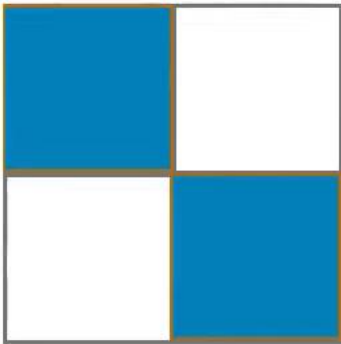
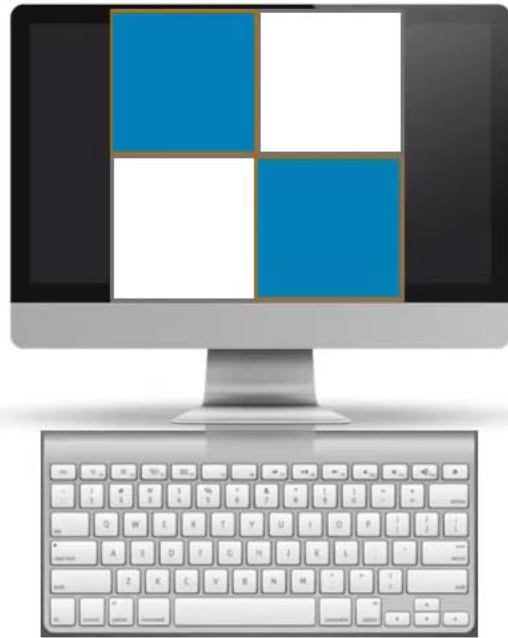


image for the symbol \



When you press \, the above image is processed

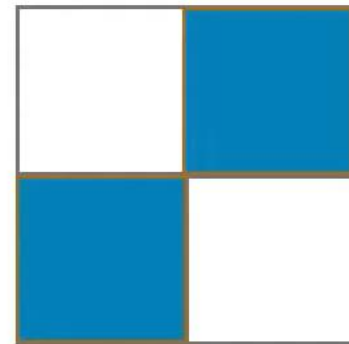
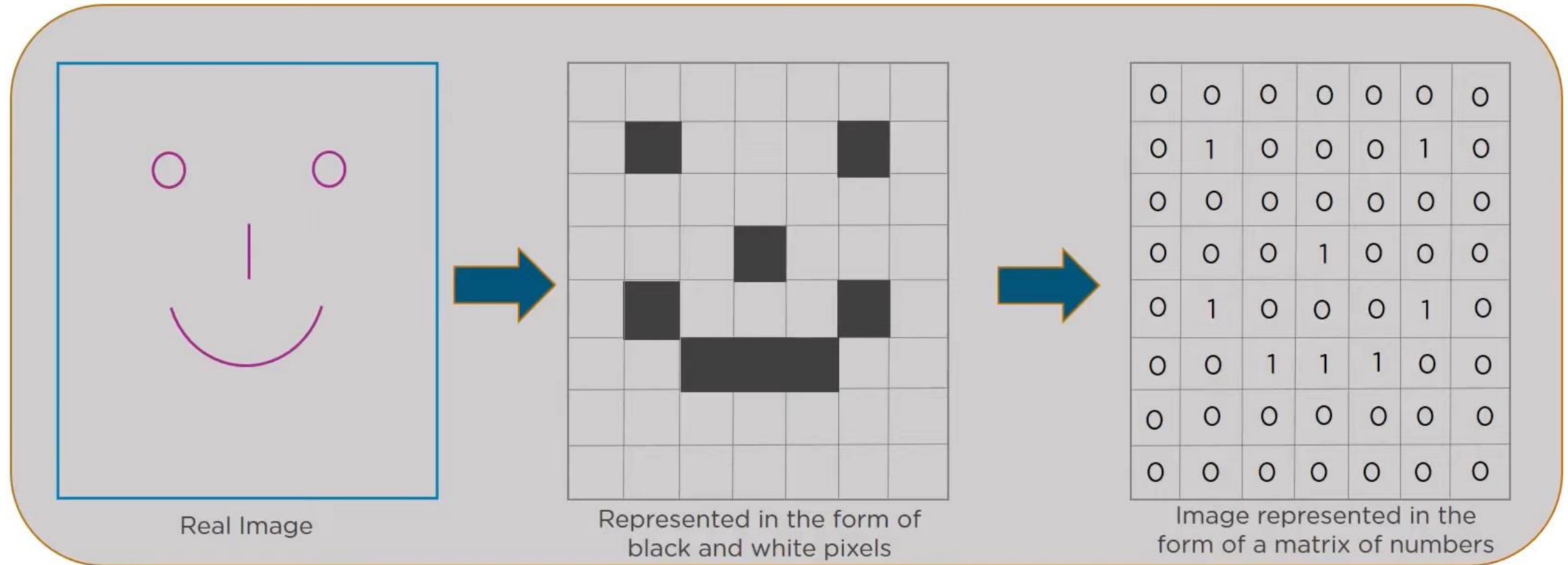
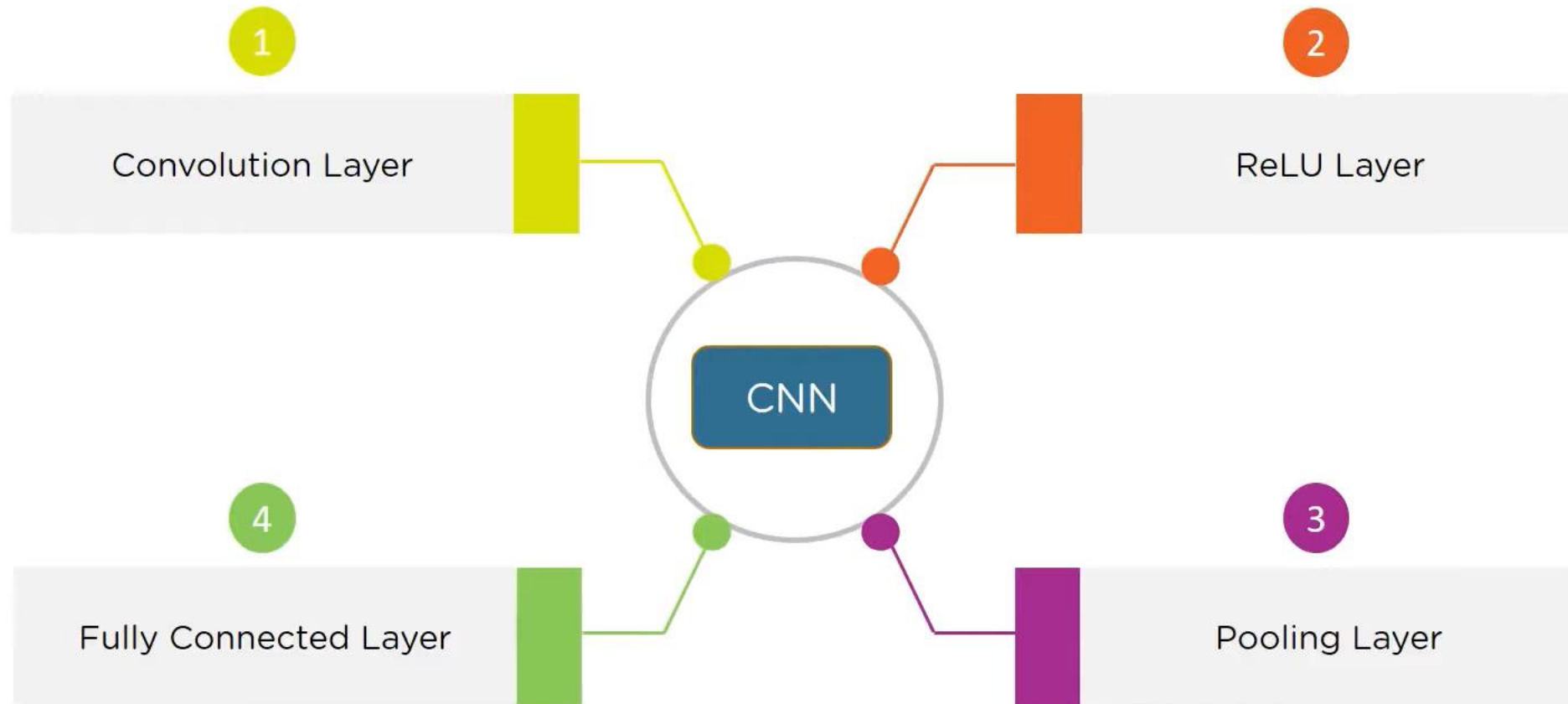


image for the symbol /

# How CNN recognizes images?



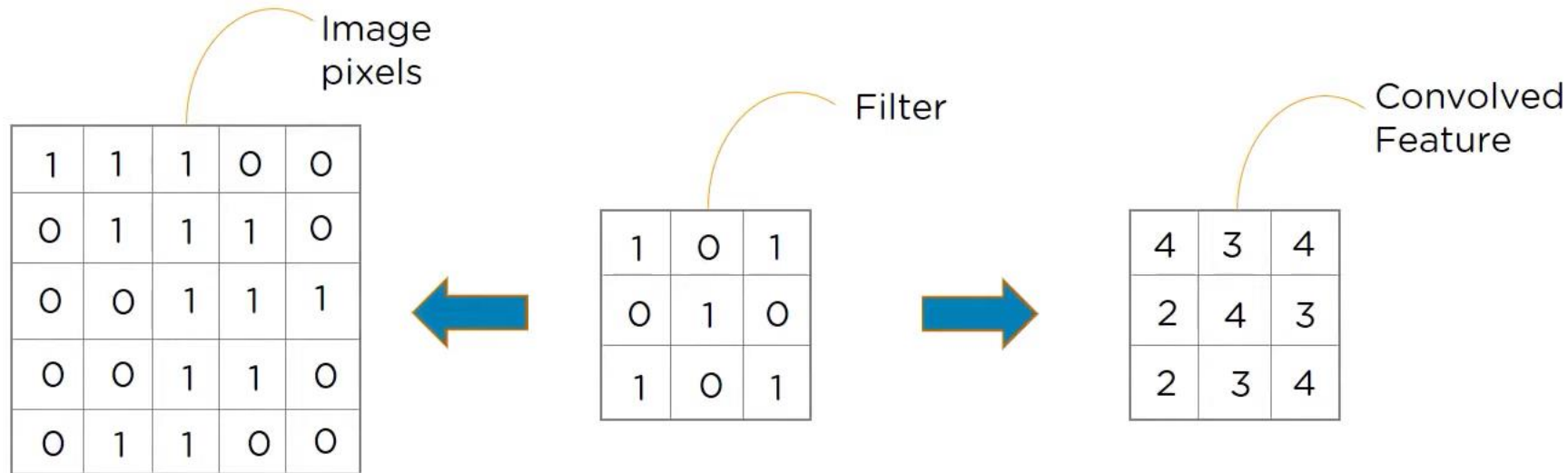
# Layers in Convolution Neural Network



# Convolution Layer

A Convolution Layer has a number of filters that perform convolution operation

Every image is considered as a matrix of pixel values.  
Consider the following 5\*5 image whose pixel values are only 0 and 1



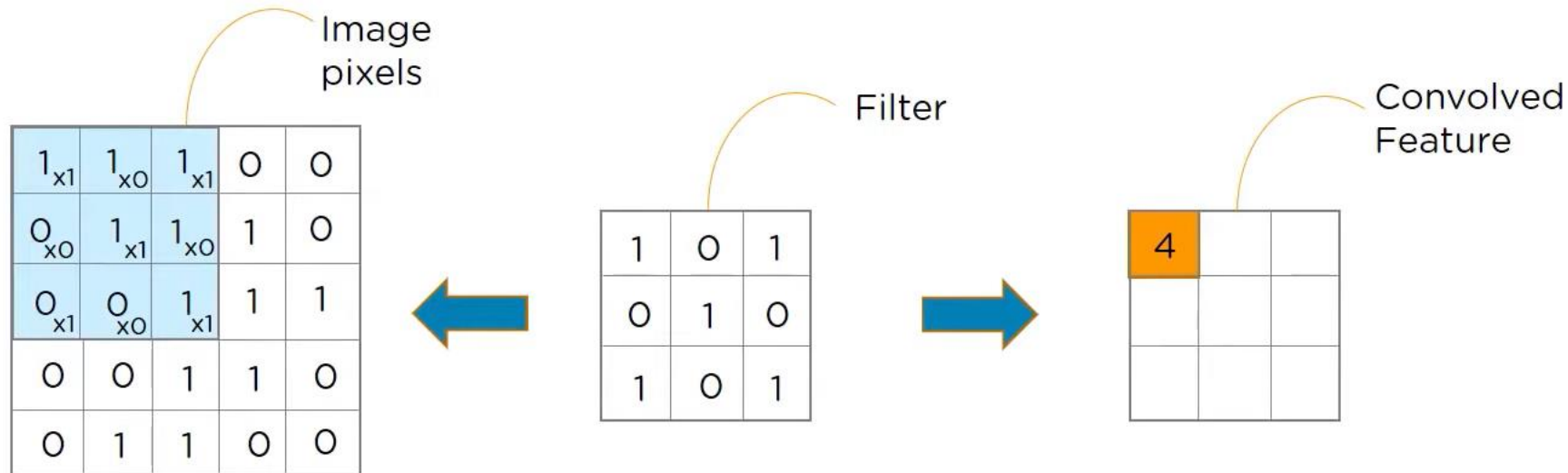
Sliding the filter matrix over the image and computing the dot product to detect patterns

# Convolution Layer

A Convolution Layer has a number of filters that perform convolution operation

Every image is considered as a matrix of pixel values.

Consider the following 5\*5 image whose pixel values are only 0 and 1



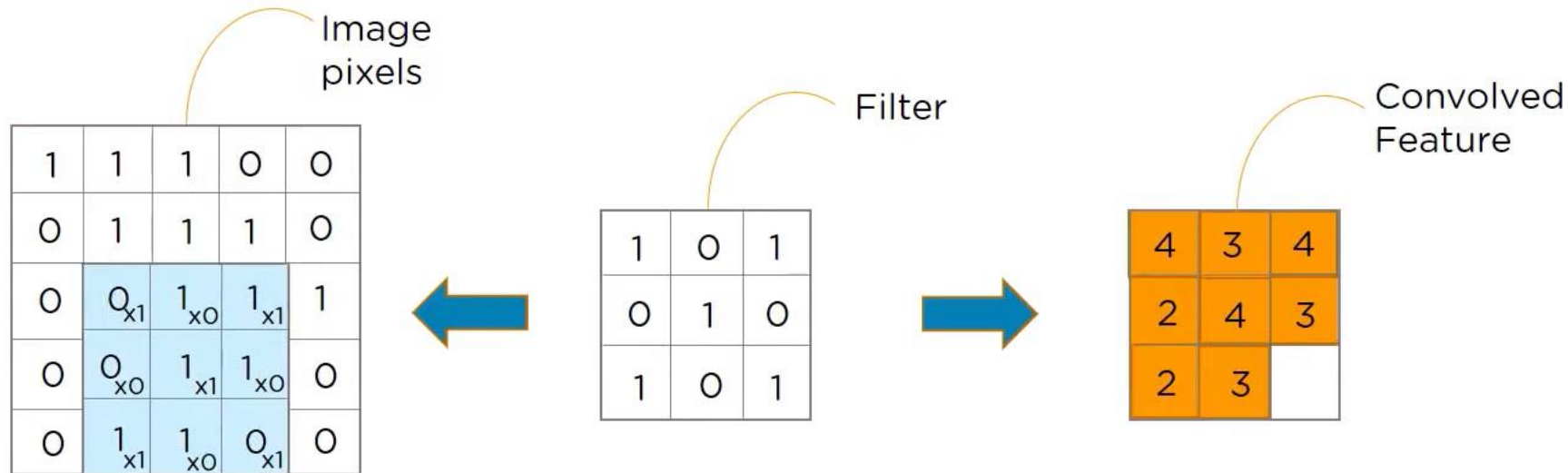
Sliding the filter matrix over the image and computing the dot product to detect patterns



# Convolution Layer

A Convolution Layer has a number of filters that perform convolution operation

Every image is considered as a matrix of pixel values.  
Consider the following 5\*5 image whose pixel values are only 0 and 1

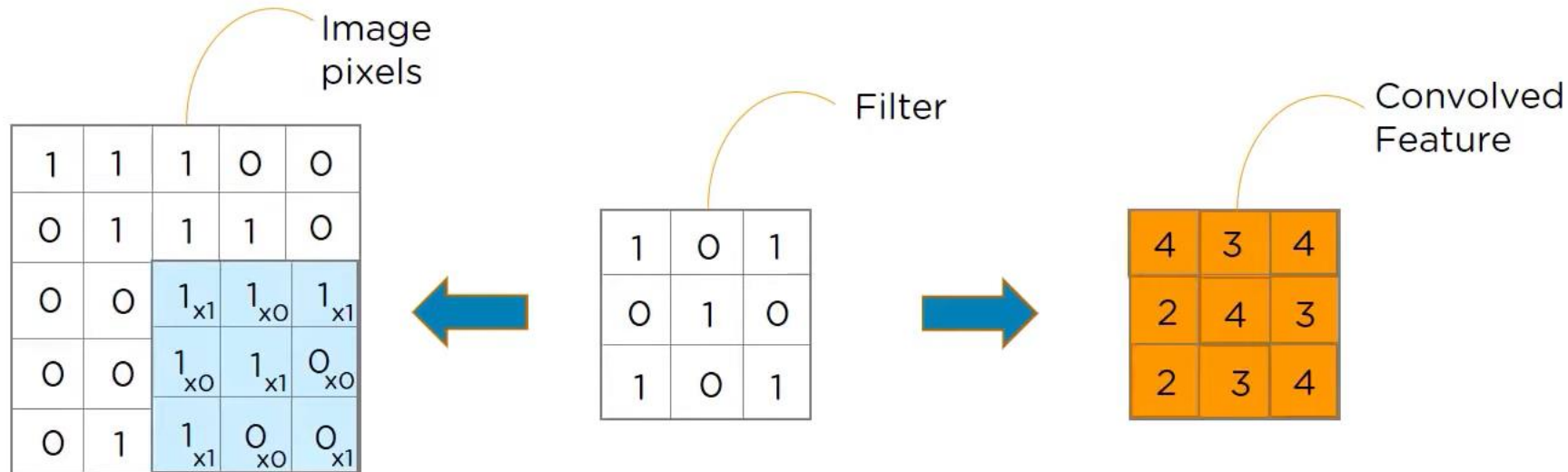


Sliding the filter matrix over the image and computing the dot product to detect patterns

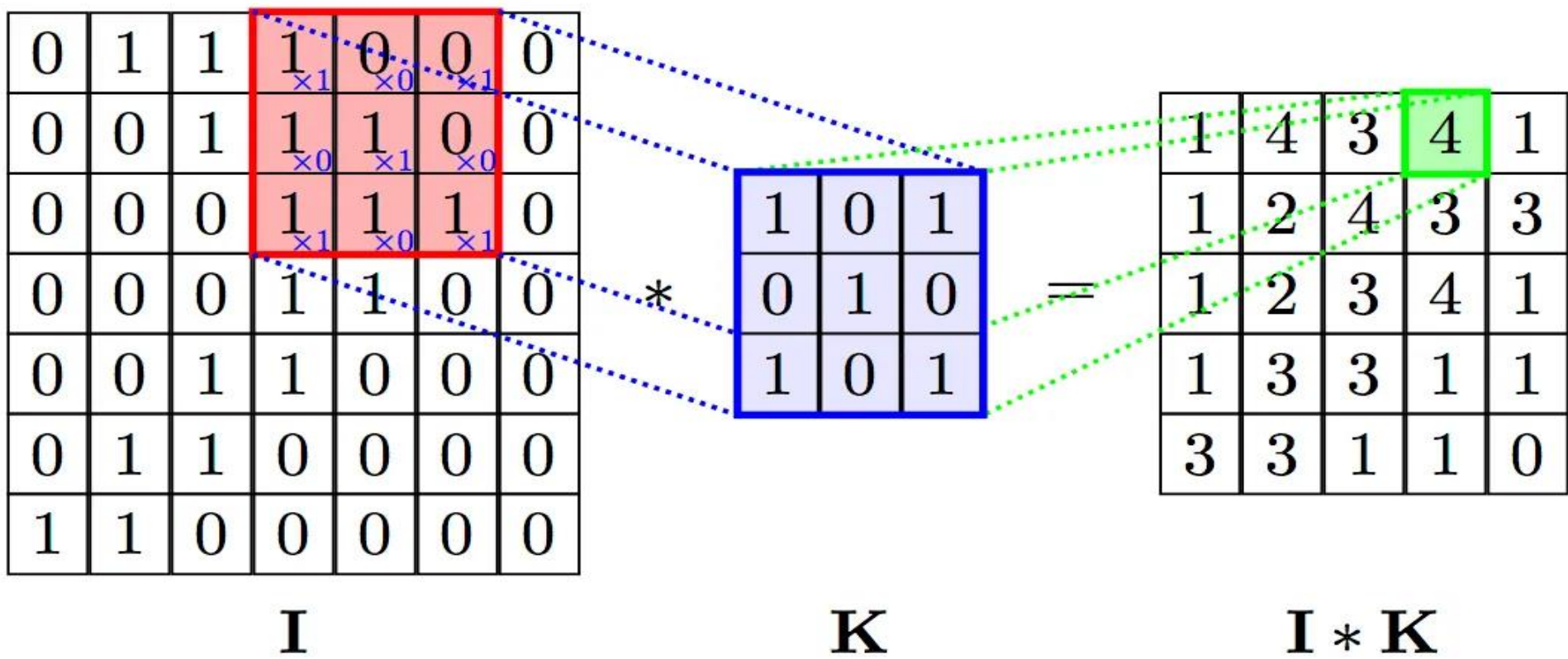
# Convolution Layer

A Convolution Layer has a number of filters that perform convolution operation

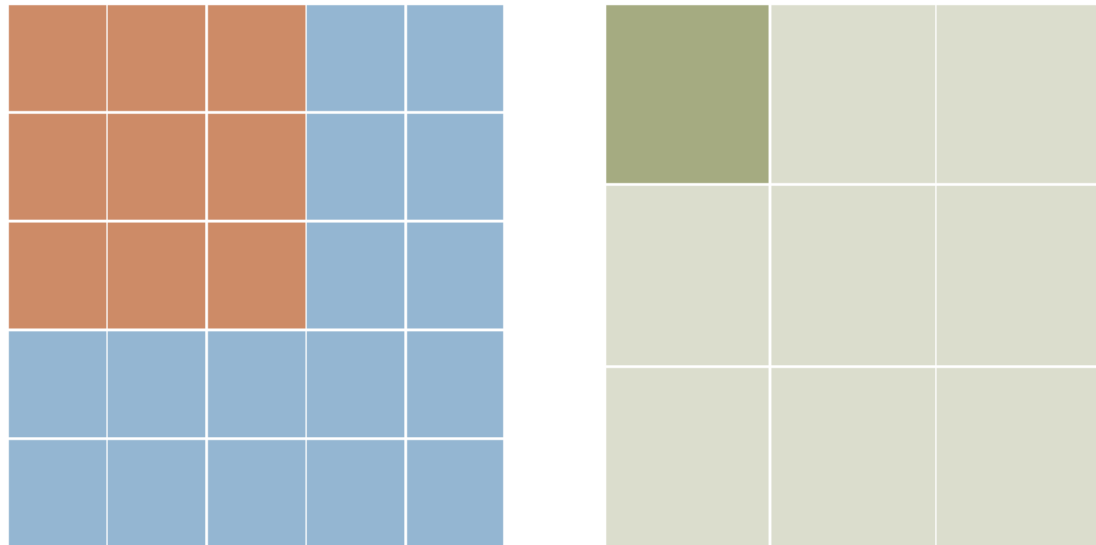
Every image is considered as a matrix of pixel values.  
Consider the following 5\*5 image whose pixel values are only 0 and 1



Sliding the filter matrix over the image and computing the dot product to detect patterns

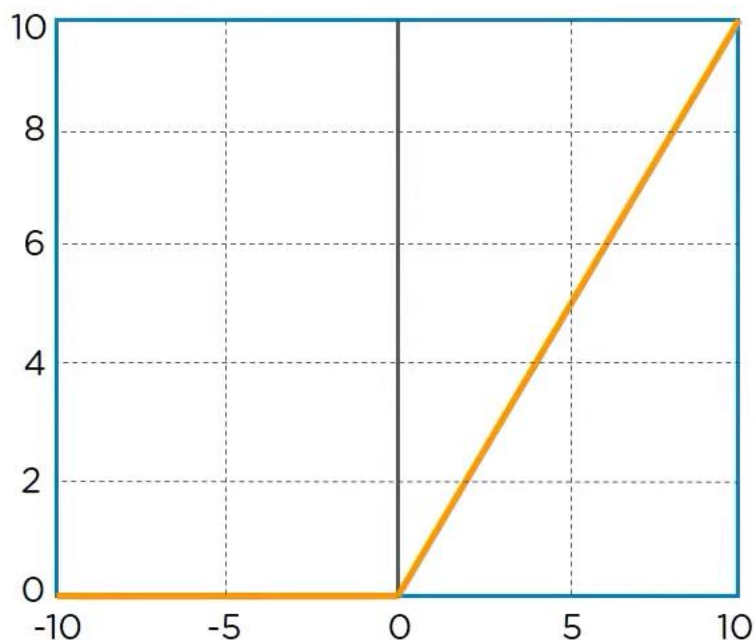


Type: conv - Stride: 1 Padding: 0

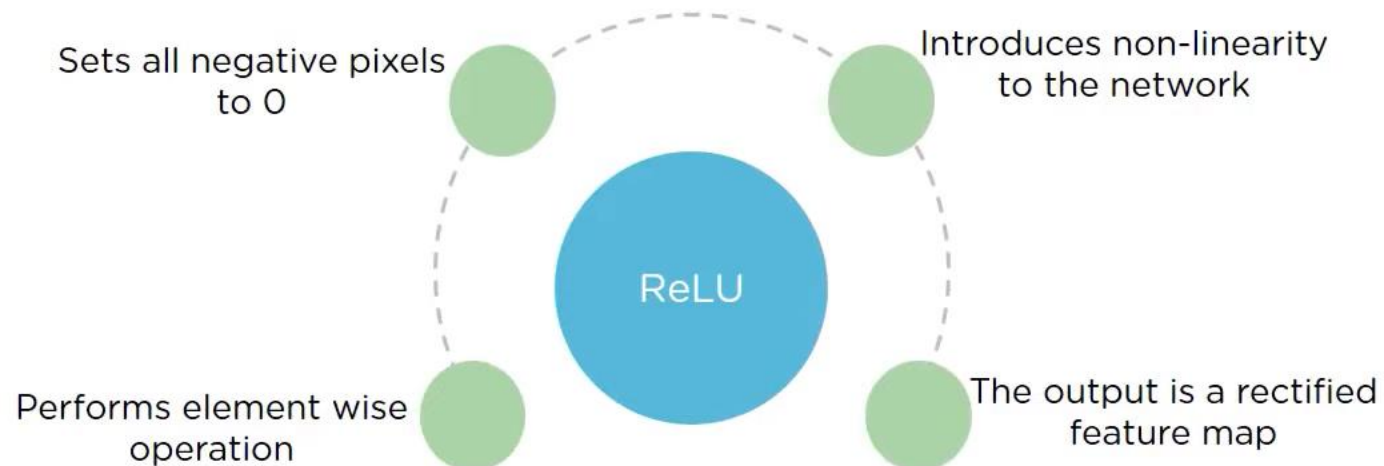


# ReLU Layer

Once the feature maps are extracted, the next step is to move them to a ReLU layer

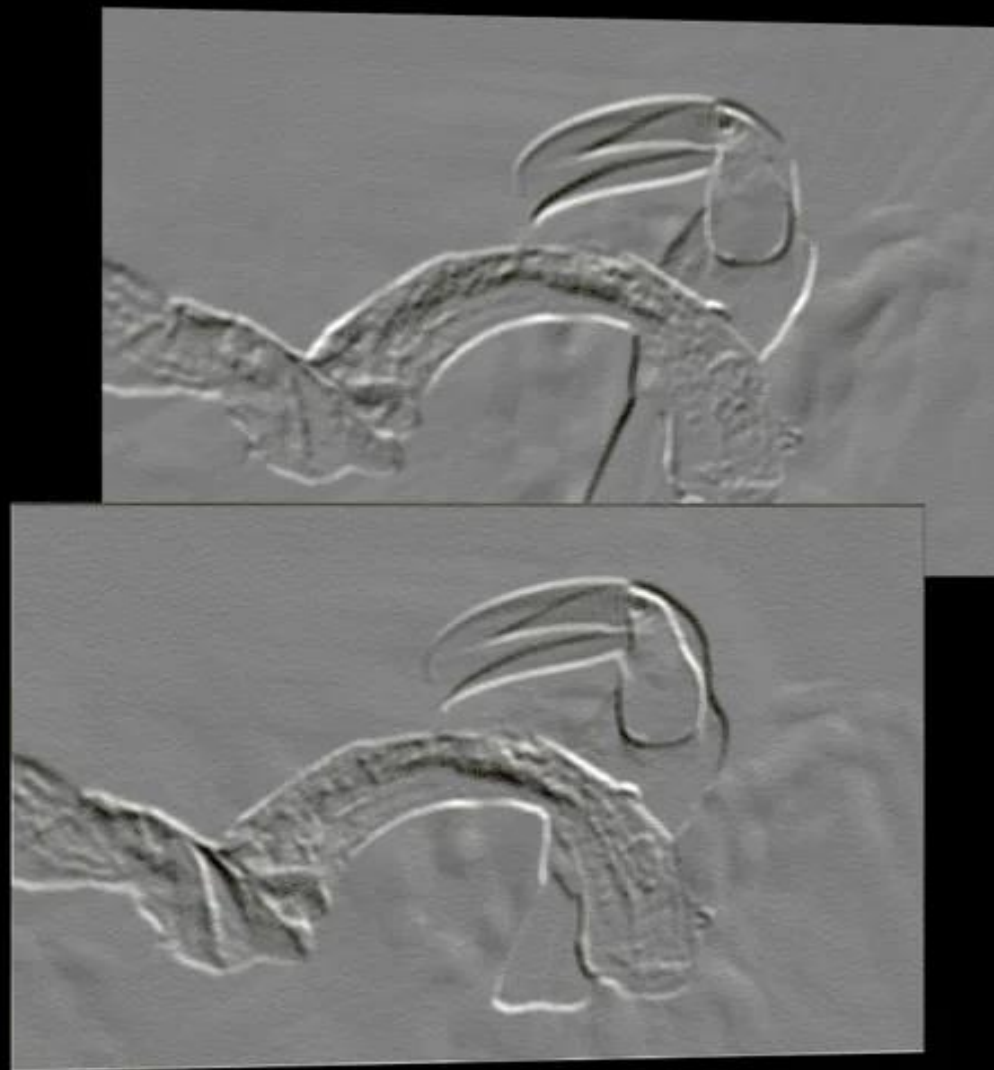


$$R(z) = \max(0, z)$$





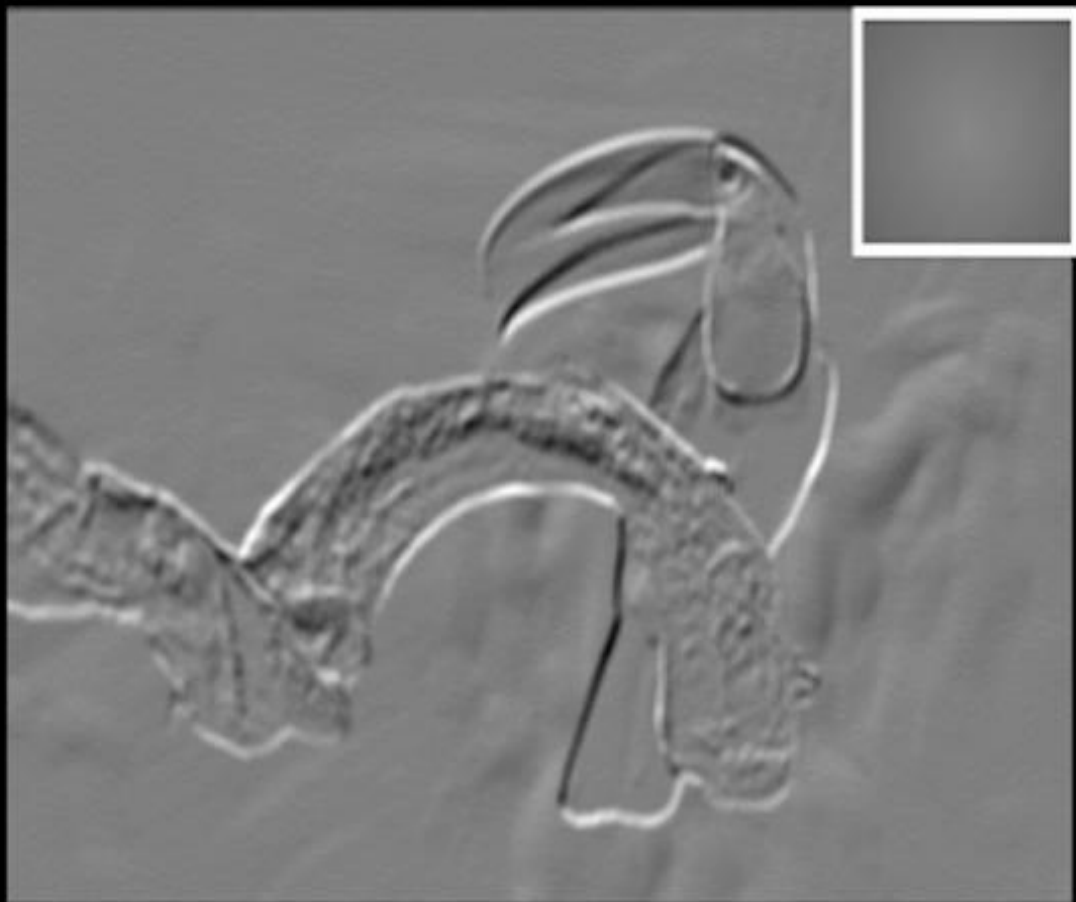
Input



Feature Map

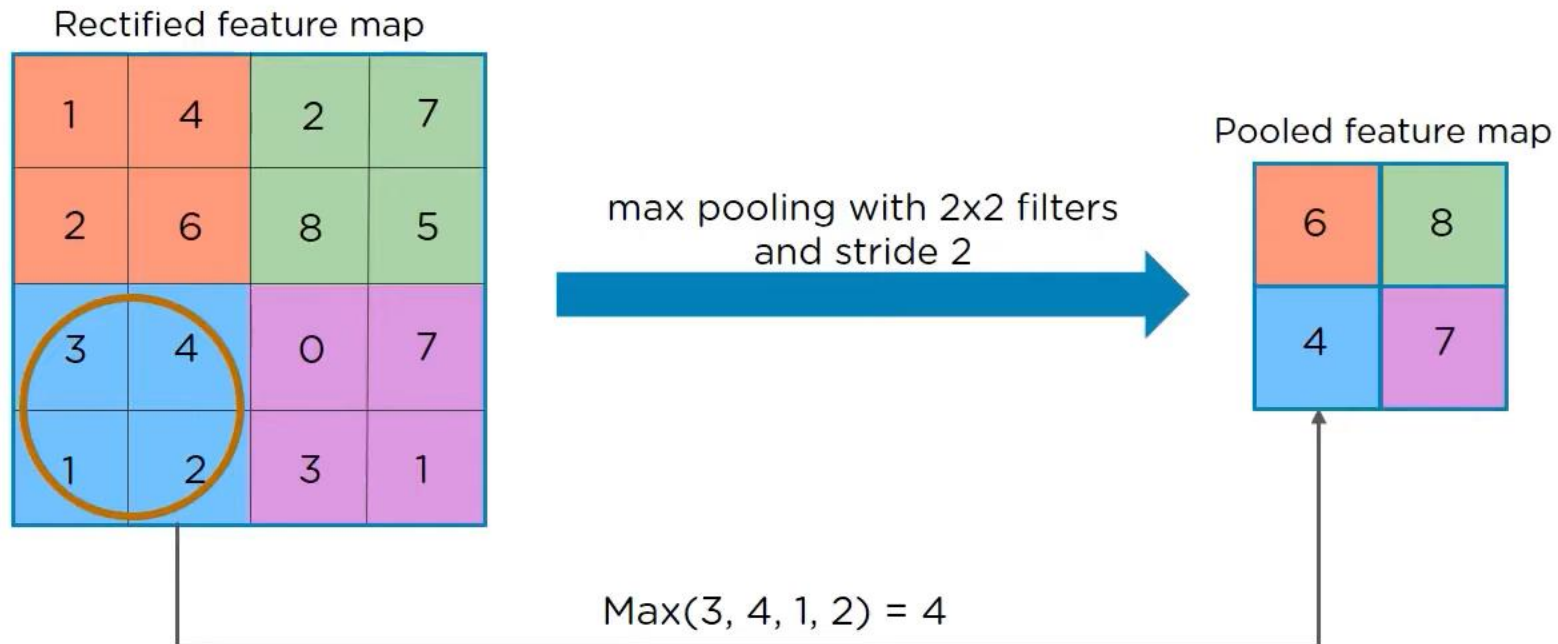


# Input Feature Map



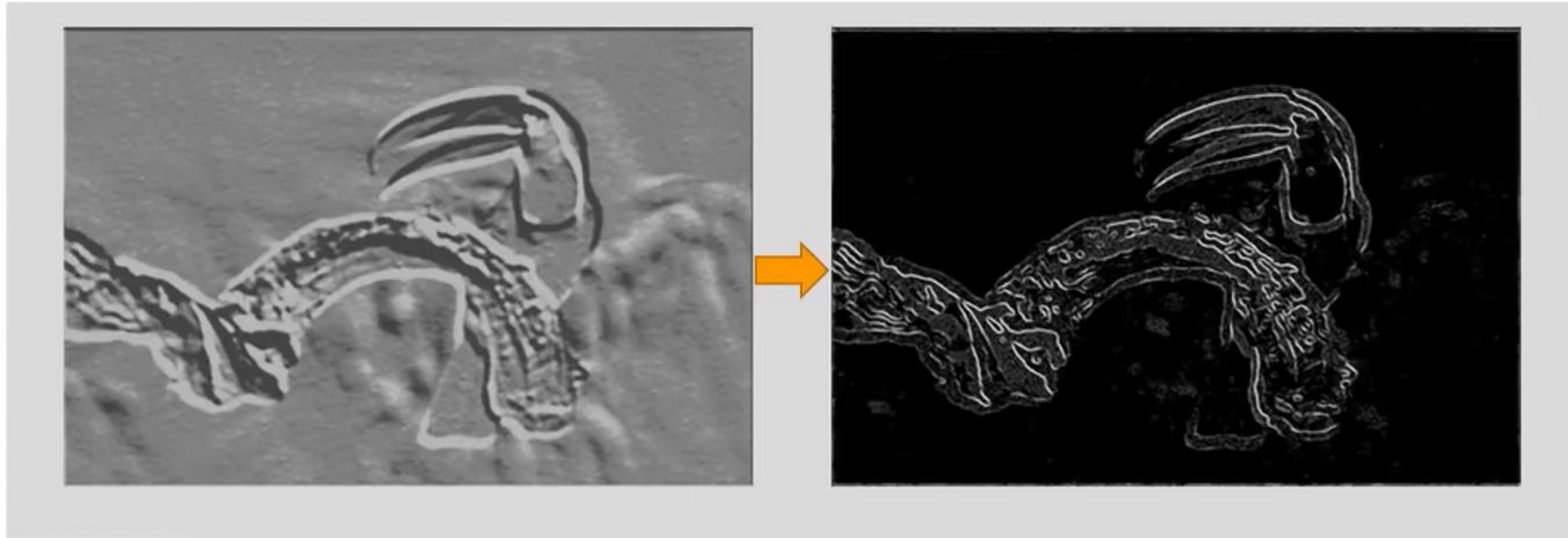
# Pooling Layer

The rectified feature map now goes through a pooling layer. Pooling is a down-sampling operation that reduces the dimensionality of the feature map.



# Pooling Layer

Pooling layer uses different filters to identify different parts of the image like edges, corners, body, feathers, eyes, beak, etc.



Identifies the edges, corners and other features of the bird

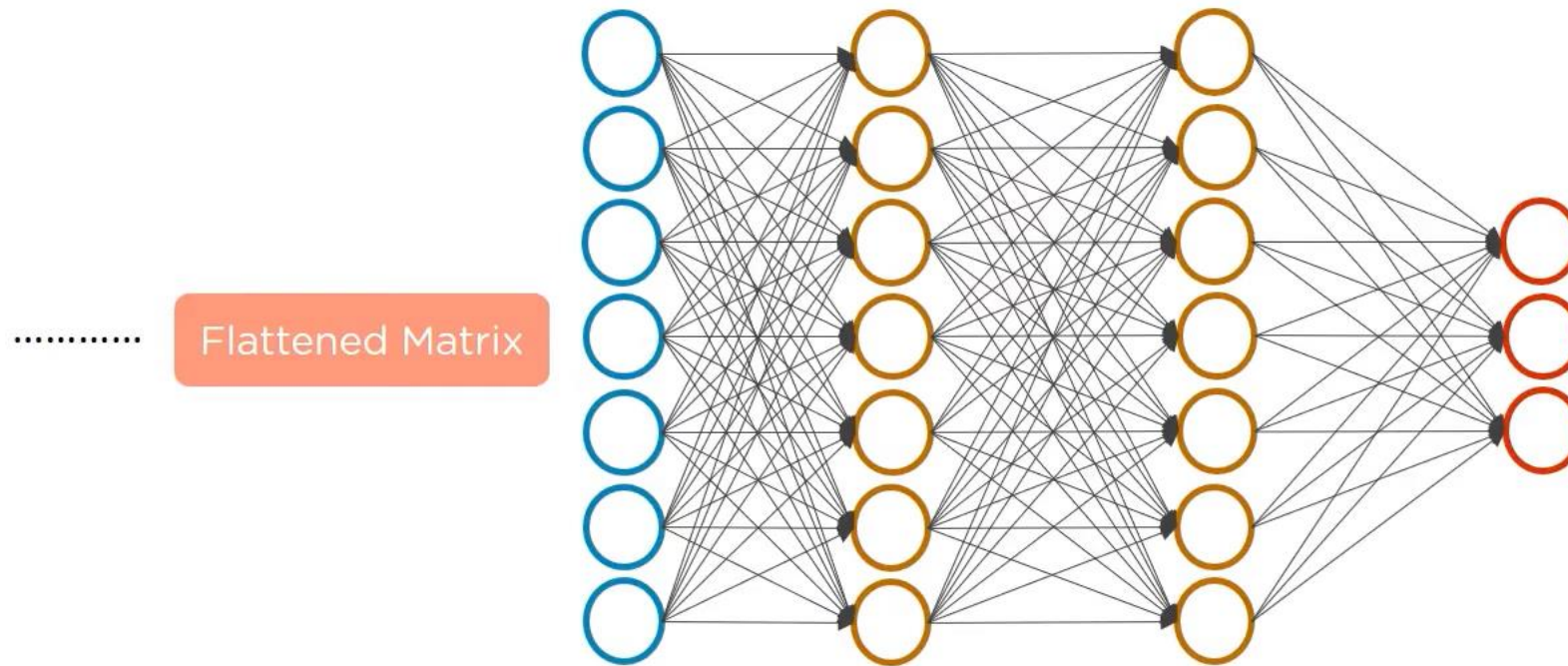
# Flattening

Flattening is the process of converting all the resultant 2 dimensional arrays from pooled feature map into a single long continuous linear vector.



# Fully Connected Layer

The Flattened matrix from the pooling layer is fed as input to the Fully Connected Layer to classify the image



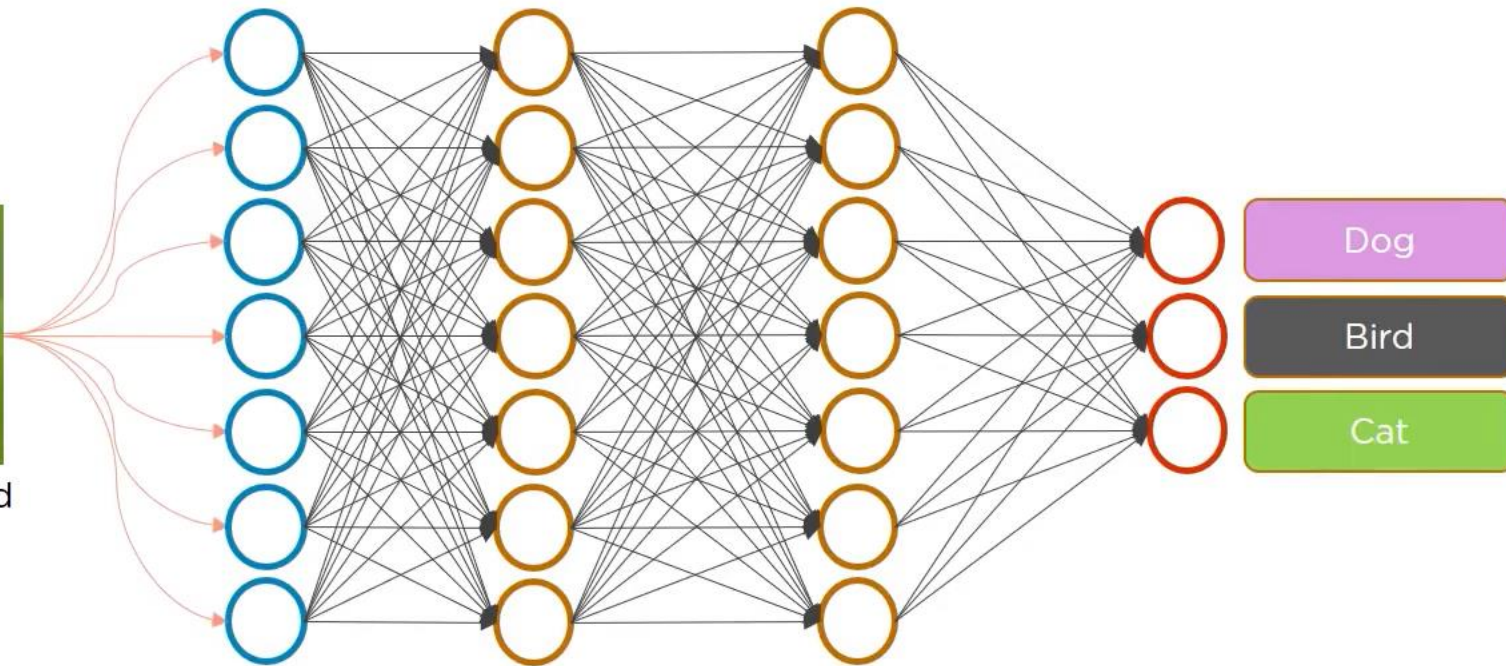


# Fully Connected Layer

The Flattened matrix from the pooling layer is fed as input to the Fully Connected Layer to classify the image



Pixels from the flattened matrix fed as input



# Fully Connected Layer

Lets see the entire process how CNN recognizes a bird

