Artificial Intelligence Masterclass

Convolutional Neural Network (CNN)

H.M. Samadhi Chathuranga Rathnayake

M.Sc in CS (SU), PG.Dip in SML (Othm), PG.Dip in HRM (LRN), B.Sc (Hons) in IS (UOC), B.Eng (Hons) in SE (LMU), P. Dip EP & SBO (ABE), Dip SE, Dip IT, Dip IT & E-Com, Dip B.Mgt, Dip HRM, Dip Eng

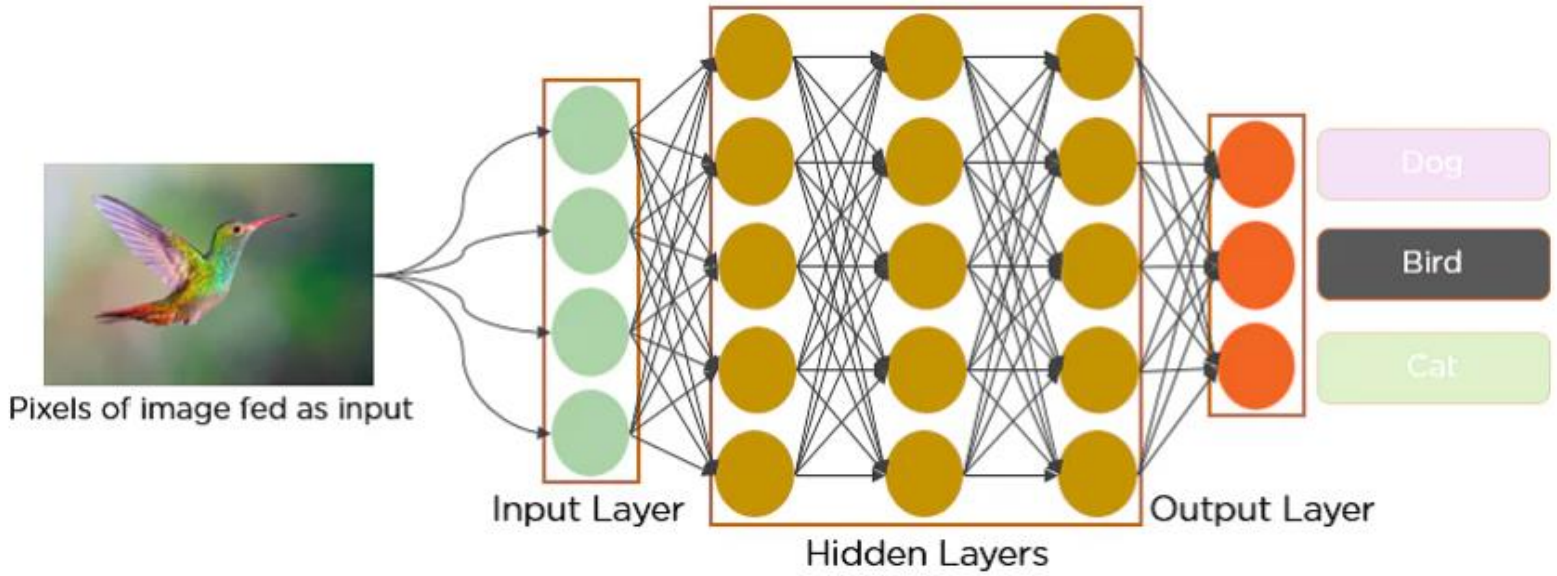
Introduction to CNN

- Convolutional neural networks were invented by Yann LeCun, who also serves as director of Facebook's AI Research Group.
- In 1988, he created the first convolutional neural network, known as LeNet.
- For character recognition tasks like reading zip codes and numbers, LeNet was employed.
- Have you ever wondered how object detection helps to construct self-driving cars, how facial recognition works on social media, or how disease identification is done using visual imaging in healthcare?
- All of this is made possible by convolutional neural networks (CNN).

Introduction to CNN

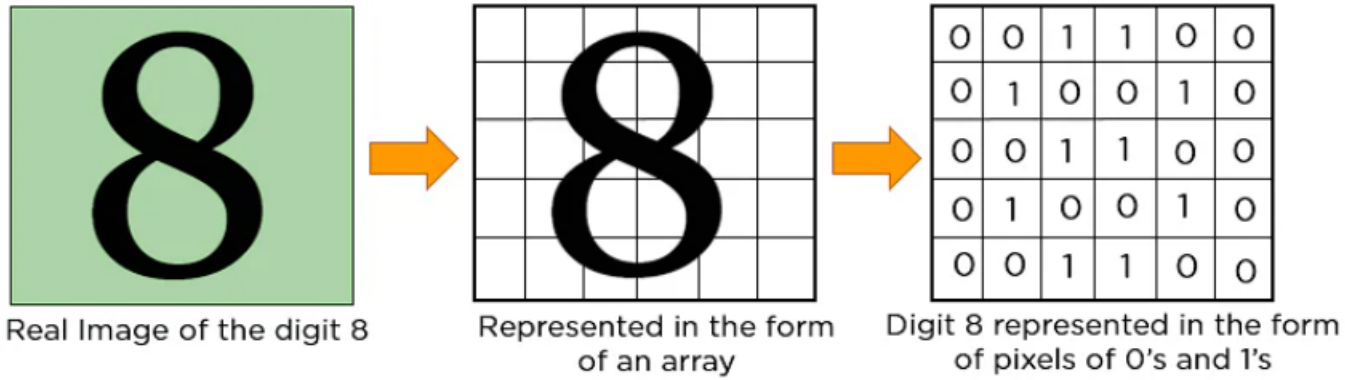
- Think there's an image of a bird, and you want to identify whether it's really a bird or some other object.
- The first thing you do is feed the pixels of the image in the form of arrays to the input layer of the neural network (multi-layer networks used to classify things).
- The hidden layers carry out feature extraction by performing different calculations and manipulations.
- There are multiple hidden layers like the convolution layer, the ReLU layer, and pooling layer, that perform feature extraction from the image.
- At the end, there's a fully connected layer that identifies the object in the image.

Introduction to CNN



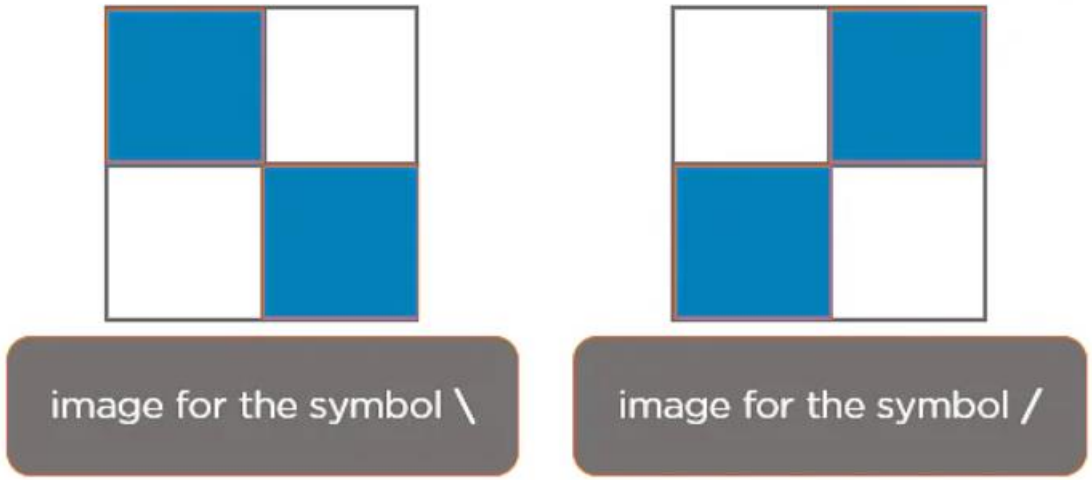
Introduction to CNN

- In CNN, every image is represented in the form of an array of pixel values.



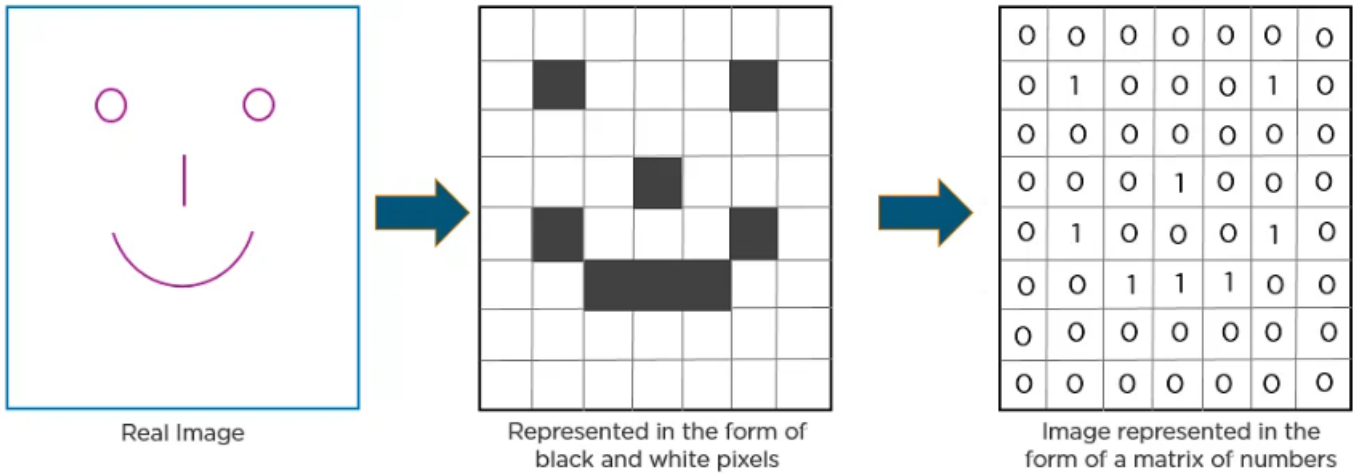
Introduction to CNN

- In CNN, every image is represented in the form of an array of pixel values.



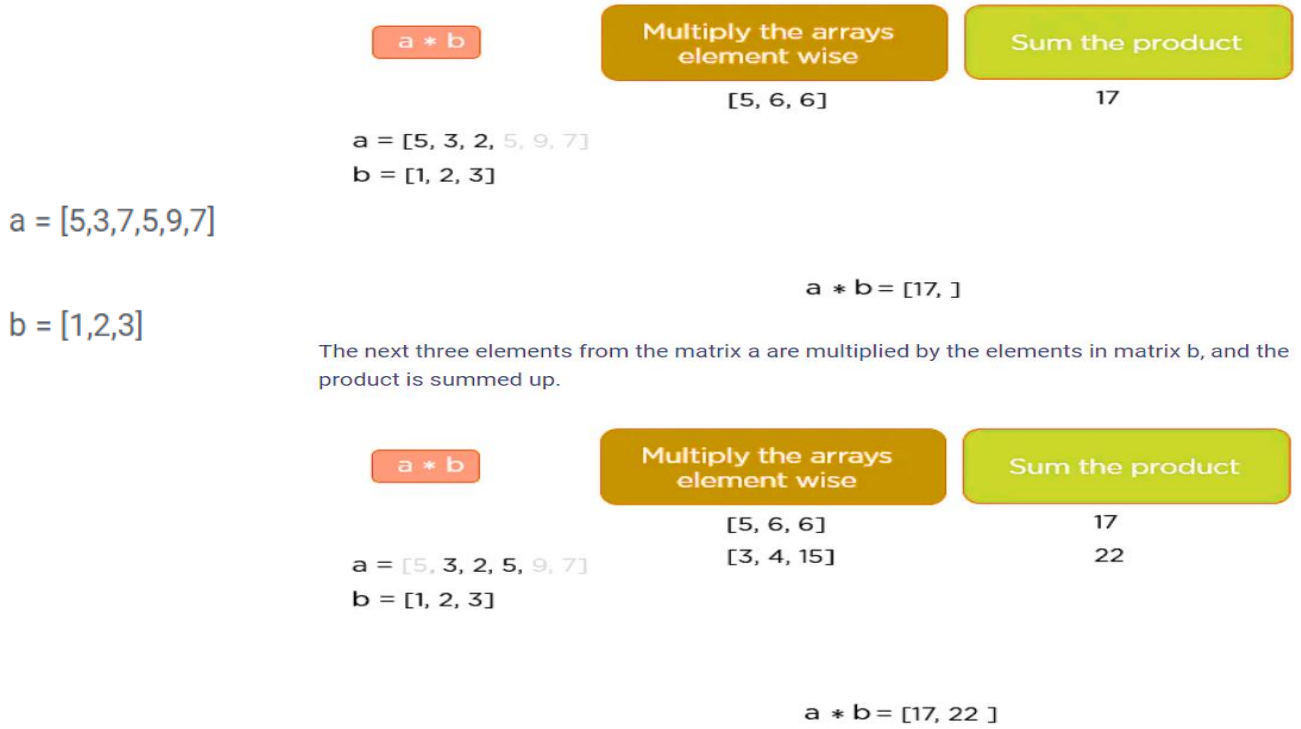
Introduction to CNN

- In CNN, every image is represented in the form of an array of pixel values.



Convolutional Operations

- Any convolutional neural network's foundation is the convolution process.
- Let's examine the convolution operation using two 1-dimensional matrices, a and b.



- This process continues until the convolution operation is complete.

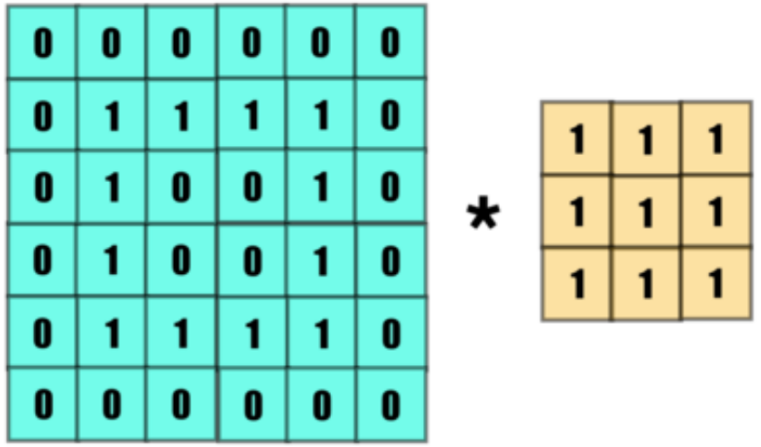
Layers of CNN

- A convolution neural network has multiple hidden layers that help in extracting information from an image. The four important layers in CNN are:

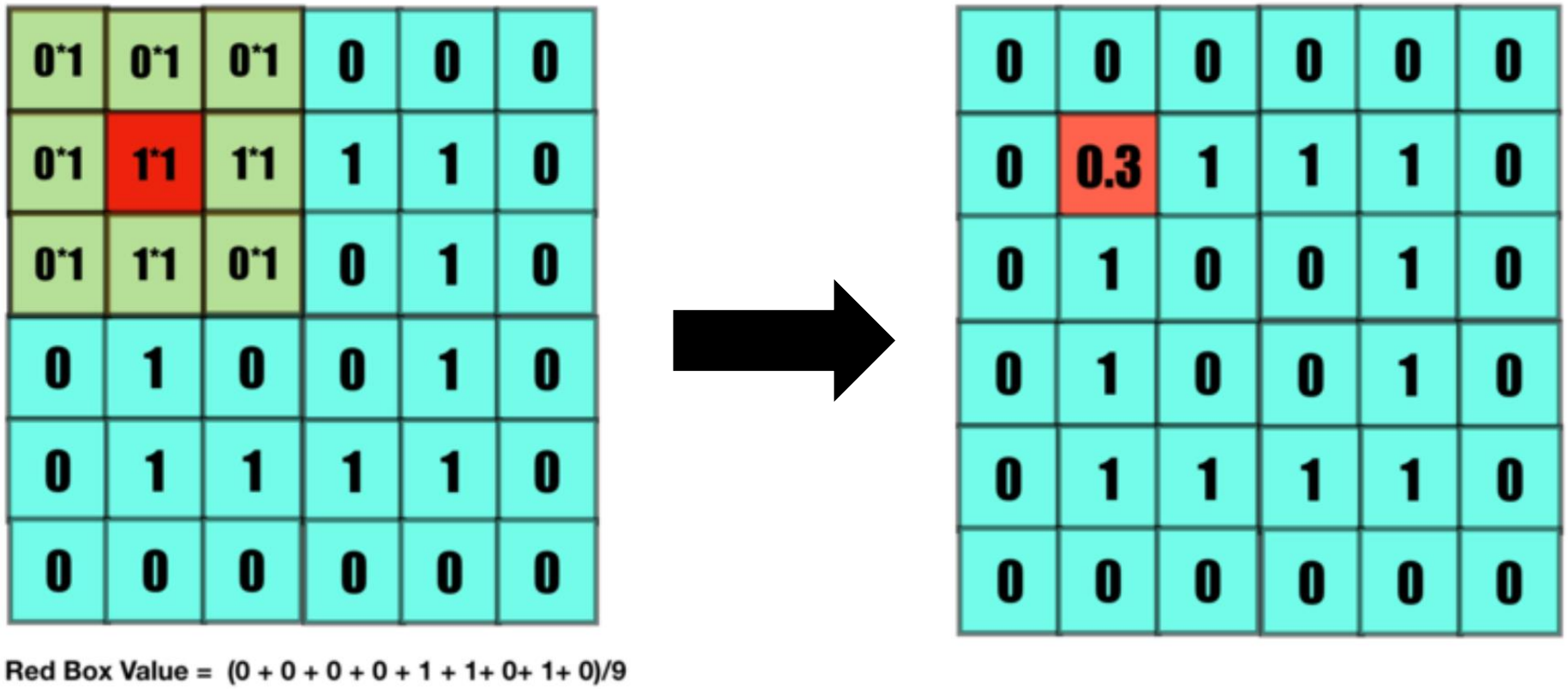
  - Convolution layer
  - ReLU (Rectified Linear Unit) layer
  - Pooling layer
  - Fully connected layer

Convolution Layer

- The process of removing useful elements from an image begins with this.
- Multiple filters work together to perform the convolution action in a convolution layer.
- Each image can be thought of as a matrix of pixel values.

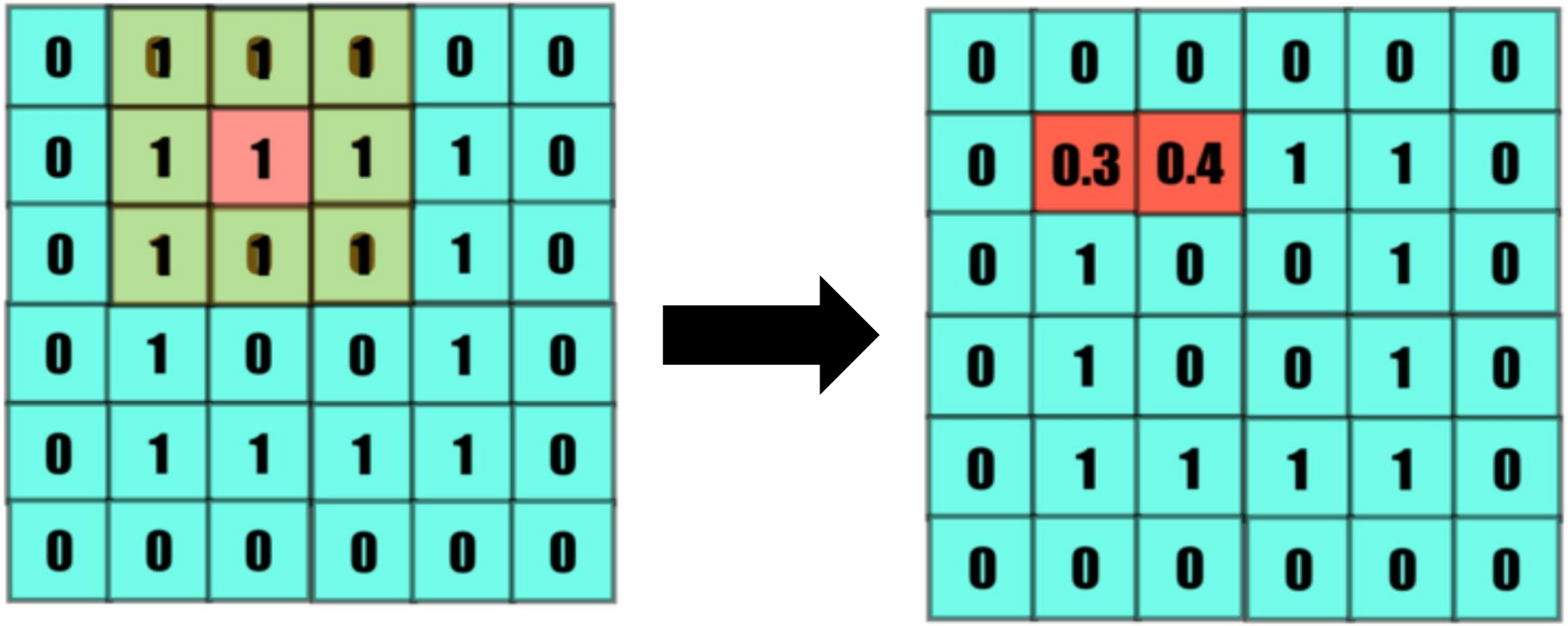


Convolution Layer



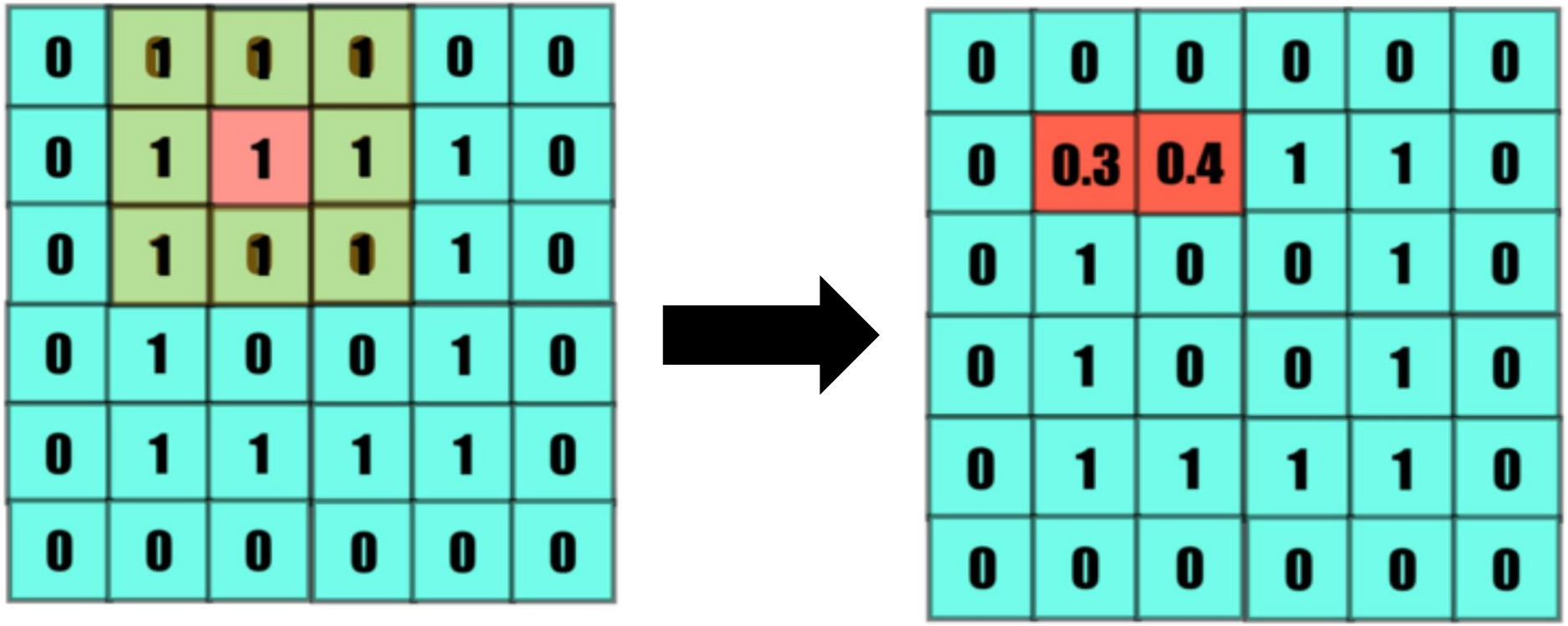
Red Box Value = (0 + 0 + 0 + 0 + 1 + 1+ 0+ 1+ 0)/9

Convolution Layer



Red Box Value = (0 + 0 + 0 + 1 + 1 + 1+ 1+ 0+ 0)/9

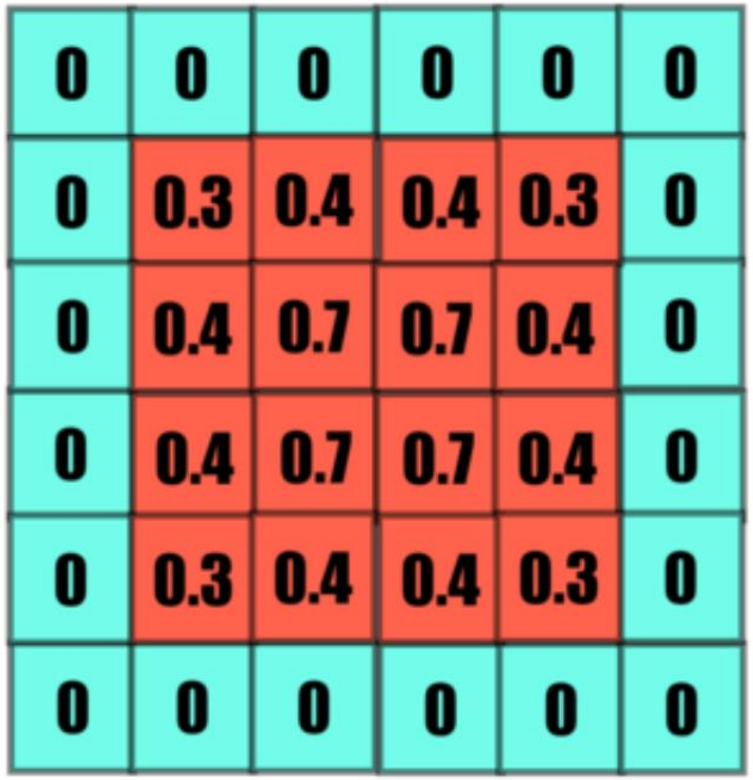
Convolution Layer



Red Box Value = (0 + 0 + 0 + 1 + 1 + 1+ 1+ 0+ 0)/9

Convolution Layer

- Get up to

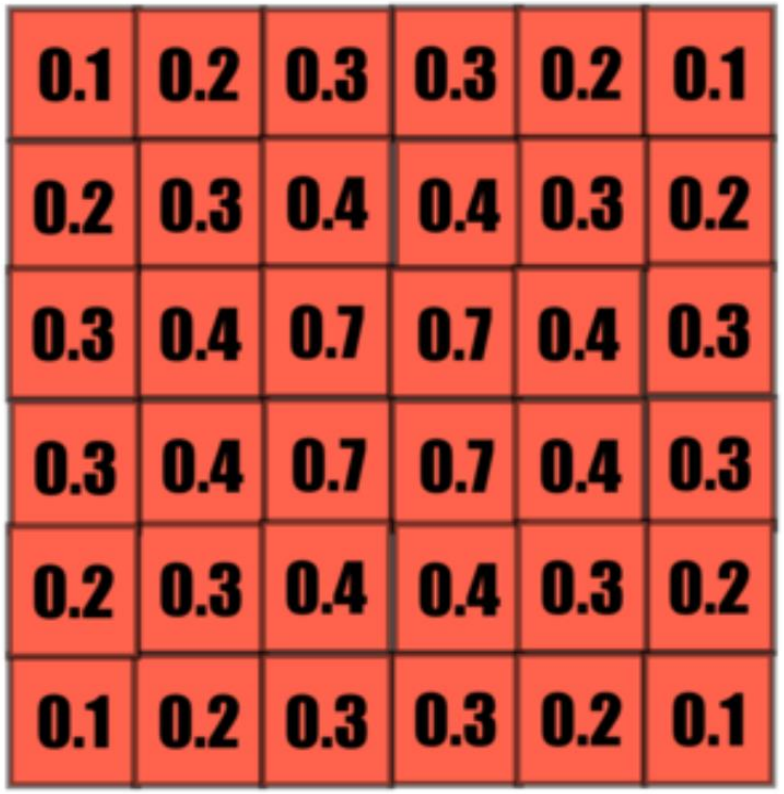


Convolution Layer

- **Padding** virtually extends the matrix to cater to border values as described in the image below. The pink layer isn't a part of the feature matrix, but helps in convolution.
- In below example padding is taken as 0.

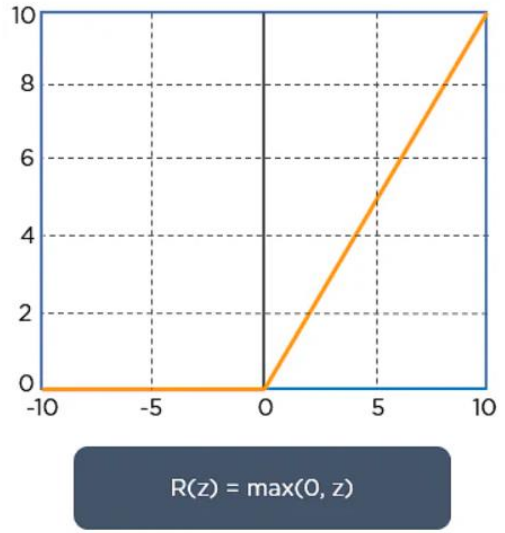
Convolution Layer

- Get up to

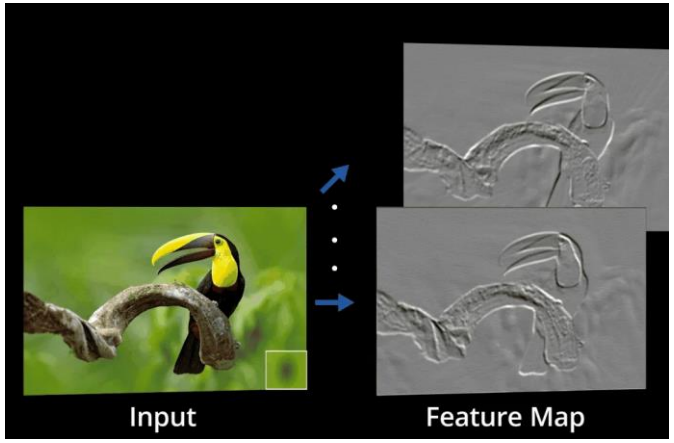


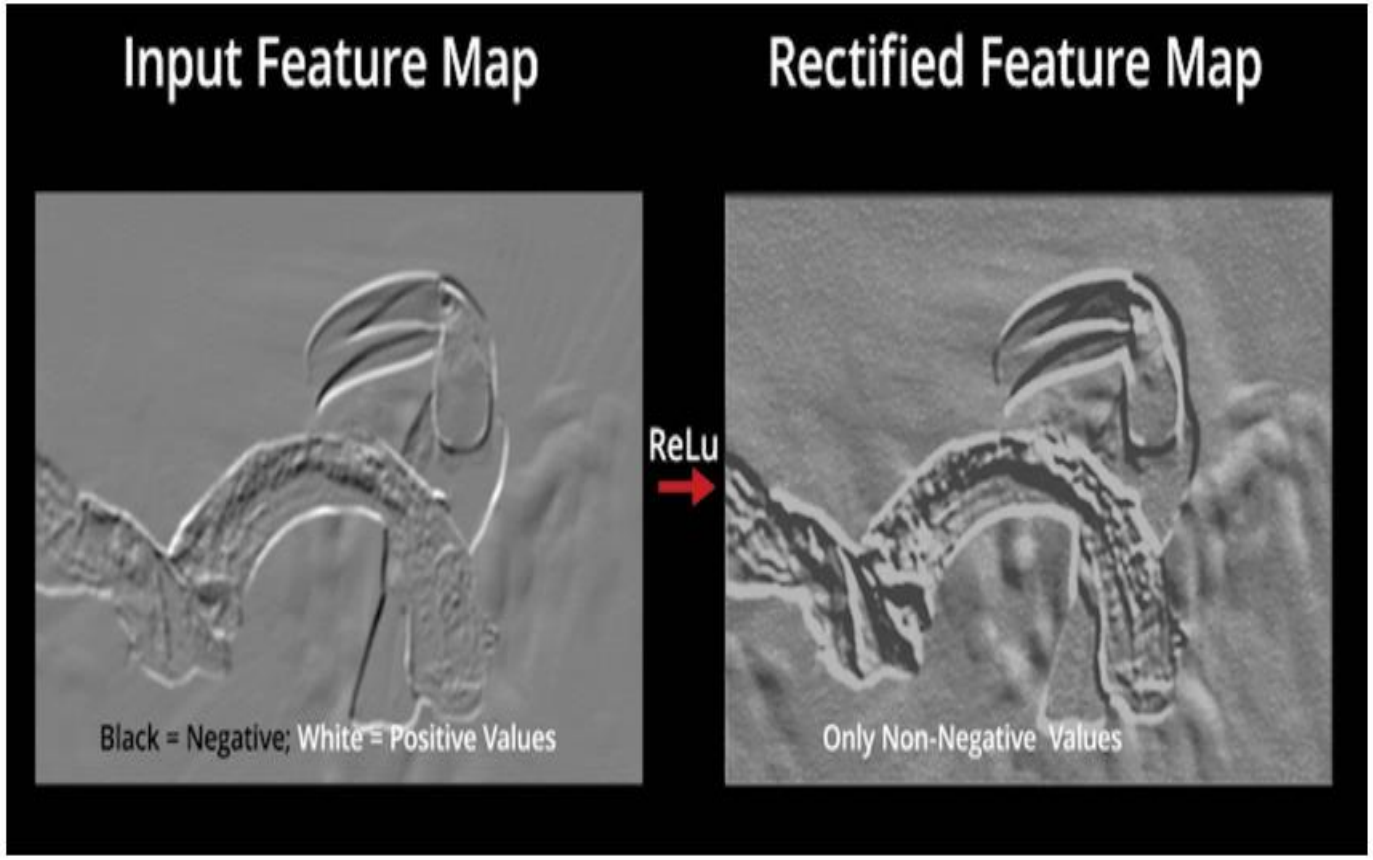
ReLU layer

- ReLU stands for the rectified linear unit.
- Once the feature maps are extracted, the next step is to move them to a ReLU layer.
- ReLU performs an element-wise operation and sets all the negative pixels to 0.
- It introduces non-linearity to the network, and the generated output is a rectified feature map.



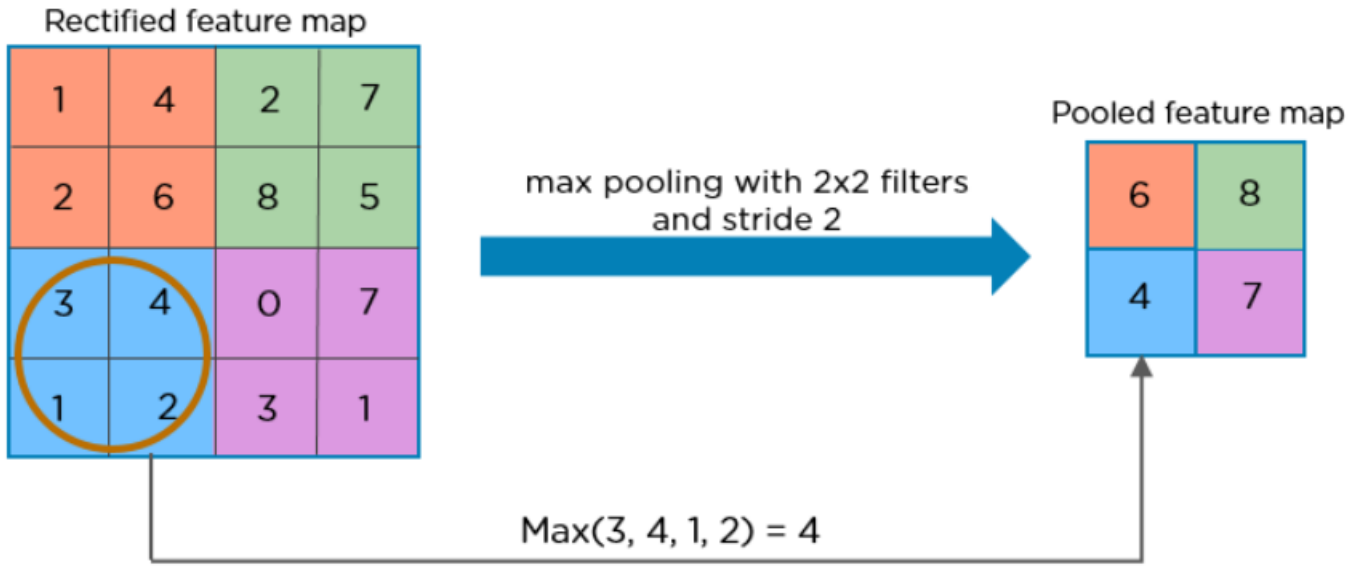
ReLU layer


Input
Feature Map


Input Feature Map
Rectified Feature Map
ReLu
Black = Negative; White = Positive Values
Only Non-Negative Values

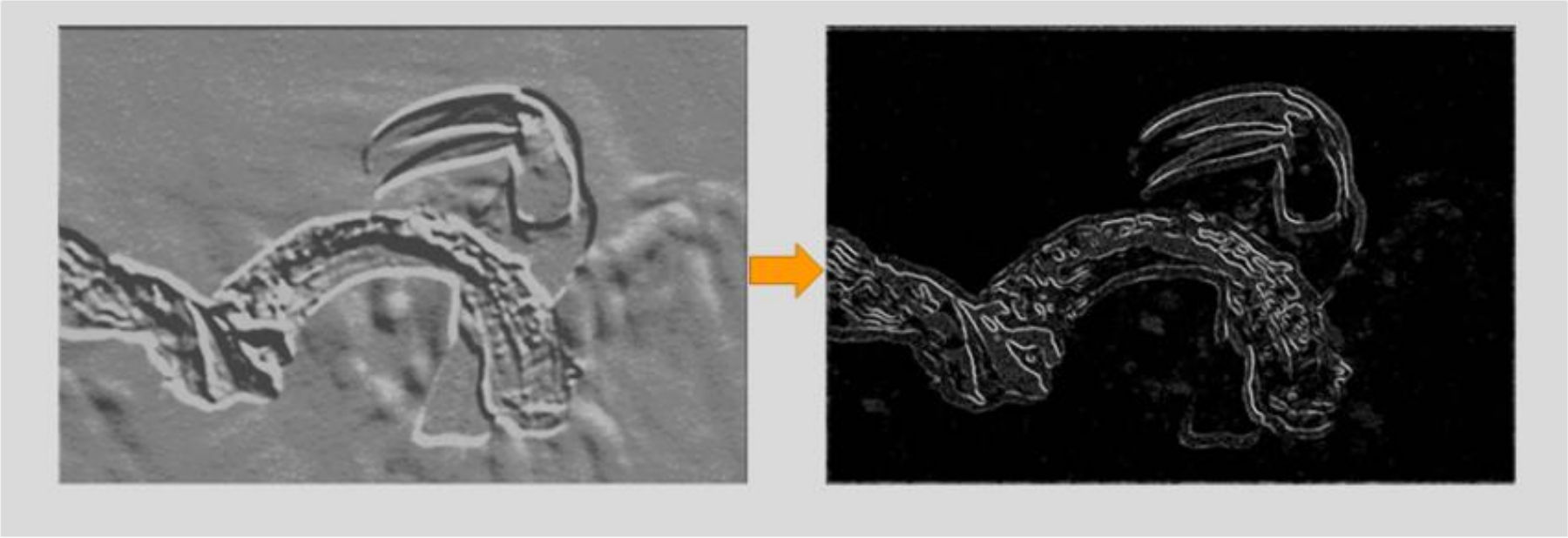
Pooling Layer

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



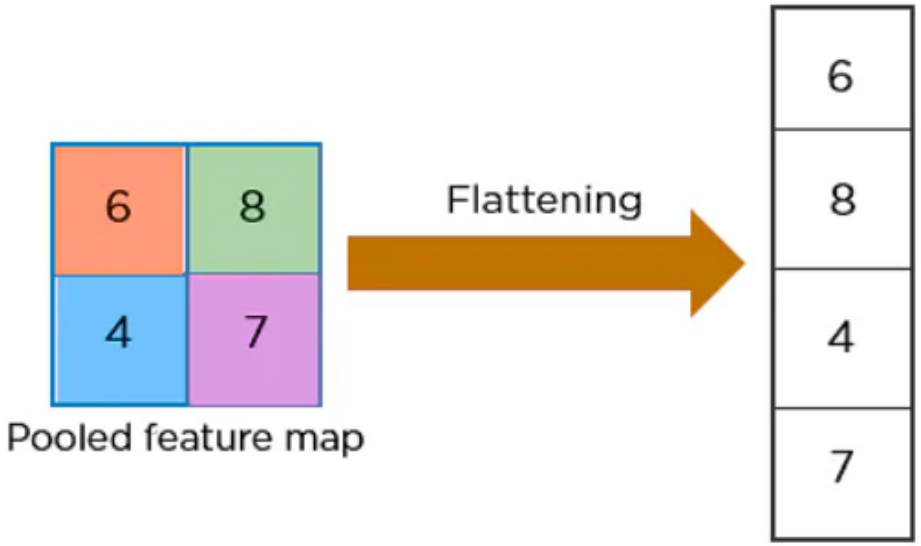
Pooling Layer

- To distinguish distinct portions of the image, such as edges, corners, bodies, feathers, eyes, and beak, the pooling layer employs a variety of filters.



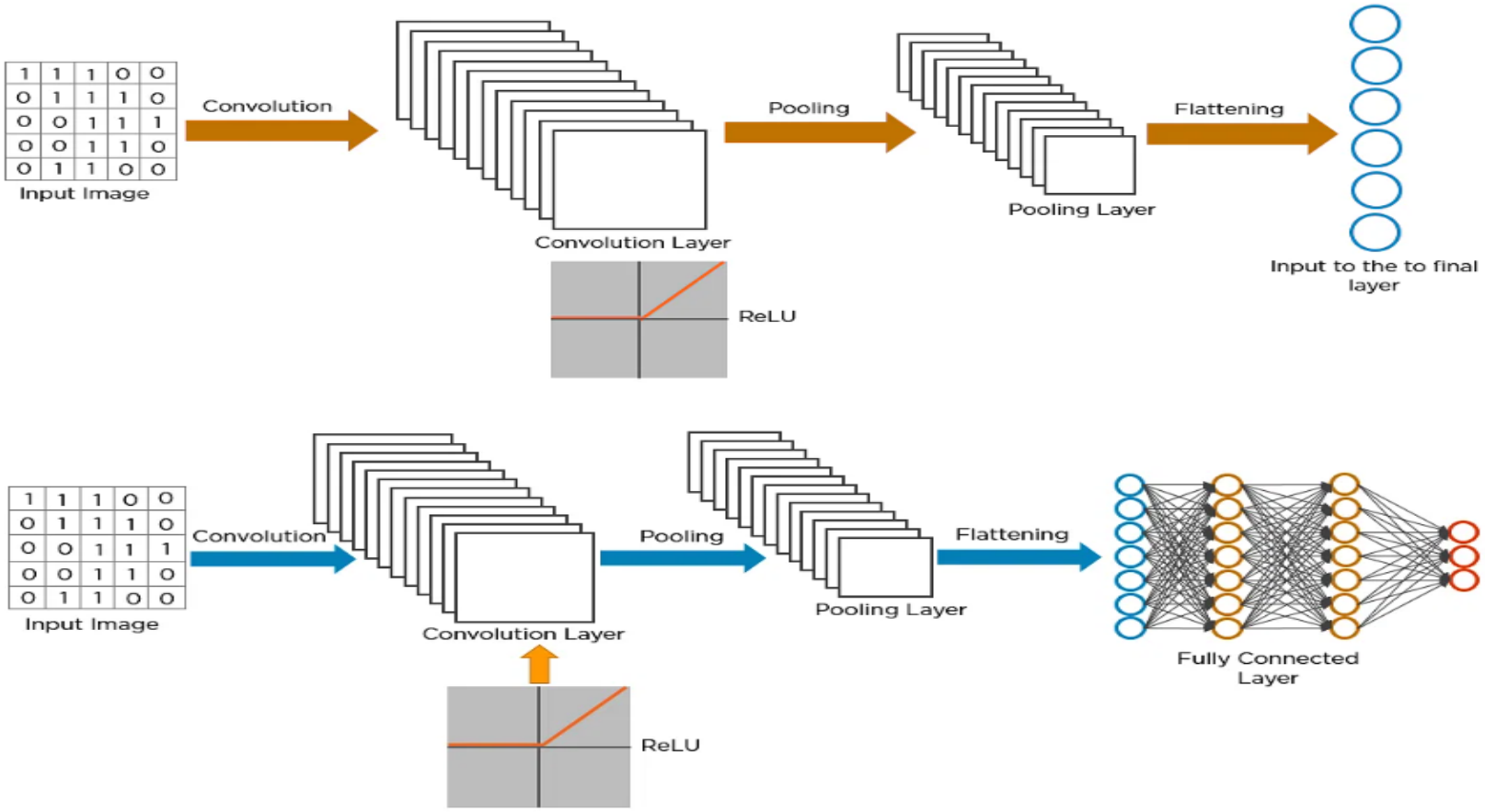
Flattening

- The next step in the process is called flattening.
- Flattening is used to convert all the resultant 2-Dimensional arrays from pooled feature maps into a single long continuous linear vector.



Fully Connected Layer

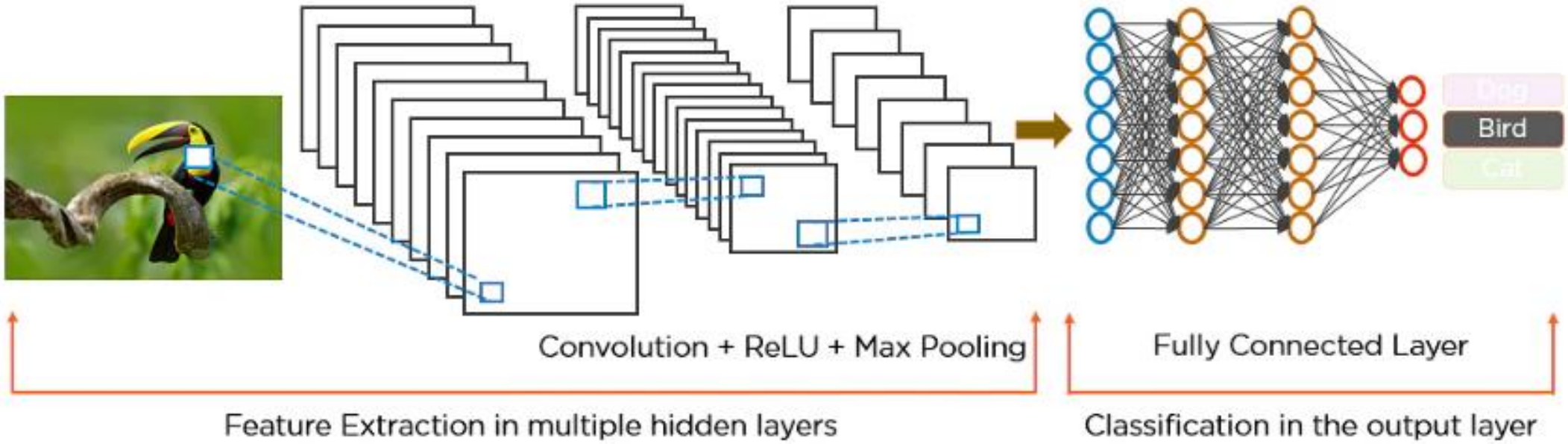
- The flattened matrix is fed as input to the fully connected layer to classify the image.



Summary

- The pixels from the image are fed to the convolutional layer that performs the convolution operation
- It results in a convolved map
- The convolved map is applied to a ReLU function to generate a rectified feature map
- The image is processed with multiple convolutions and ReLU layers for locating the features
- Different pooling layers with various filters are used to identify specific parts of the image
- The pooled feature map is flattened and fed to a fully connected layer to get the final output

Summary


Bird
Convolution + ReLU + Max Pooling
Fully Connected Layer
Feature Extraction in multiple hidden layers
Classification in the output layer