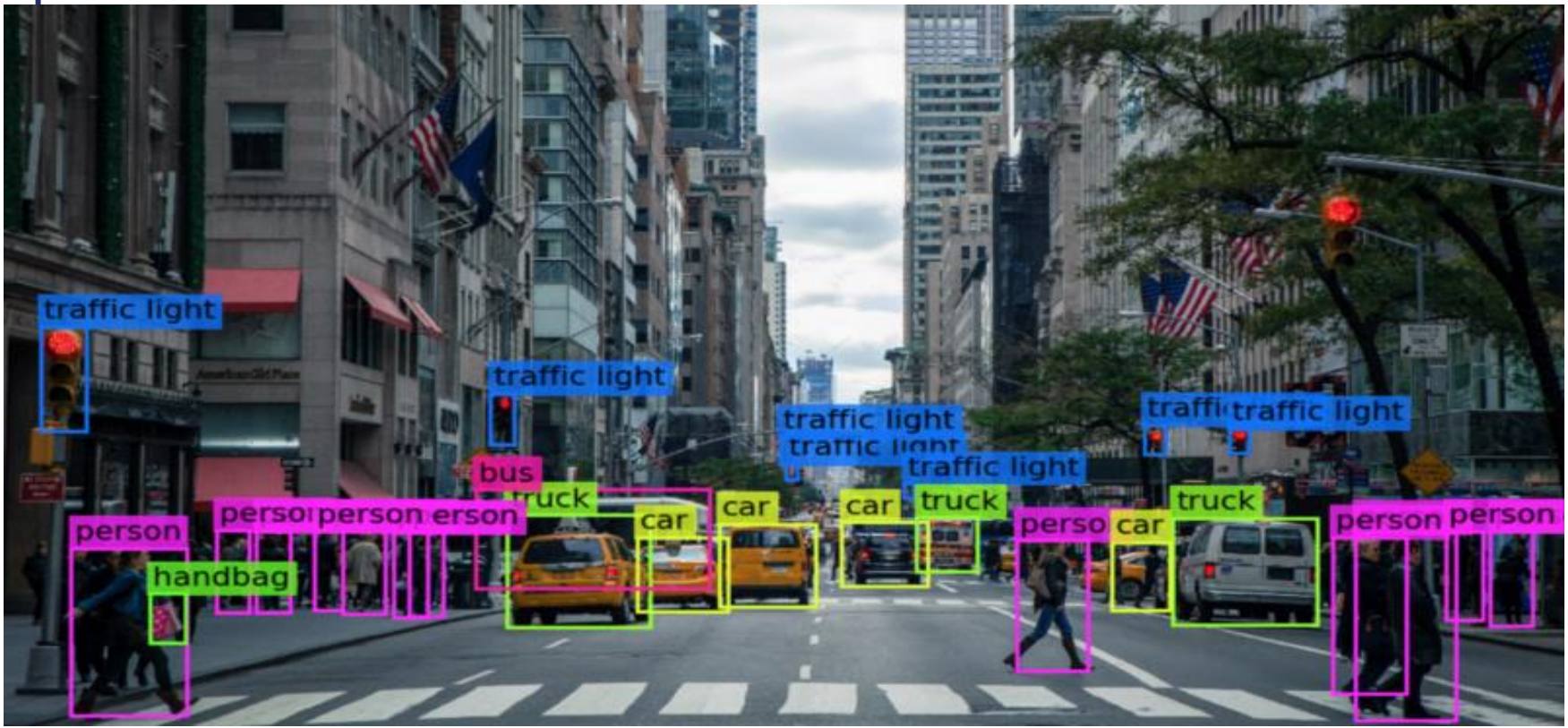
Unit 2
Computer Vision

Disclaimer

The content is curated from online/offline resources and used for educational purpose only

Computer Vision



Reference

Learning Objectives

You will learn in this lesson:

- What is Computer Vision?
- What is Image?
- Types of Image
- Working of Computer Vision
- Popular Python Libraries
- Basic operations on image
- Application of Computer Vision
- Convolutional Neural Network
- Transfer Learning

What is Computer Vision (CV)?

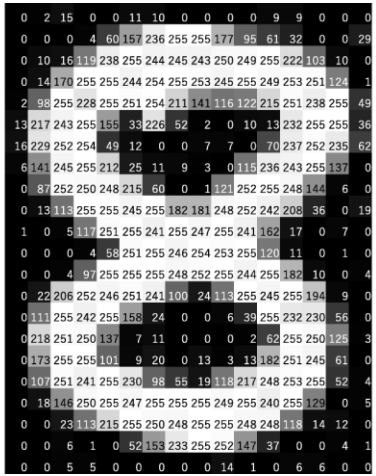
Computer vision (CV) is an artificial intelligence (AI) subcategory that focuses on developing and deploying digital systems that process, analyze, and interpret visual input.

The objective of computer vision is to allow computers to recognize an item or person in a digital image and take appropriate action.

Convolutional neural networks (CNNs) are used in computer vision to analyze visual input at the pixel level.

What is an Image?

Data in the form of matrix(Rows and Columns) consisting of Pixels



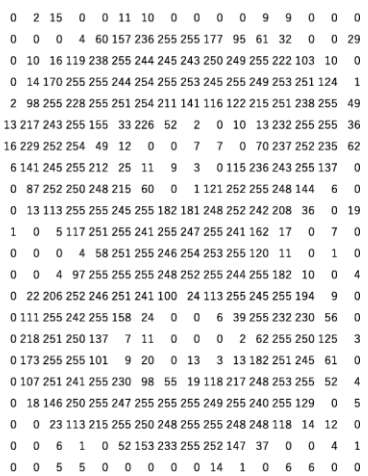


Image with Pixels

Click here

Reference link

Types of Images

(a)

Color Image

(b)

Grayscale Image

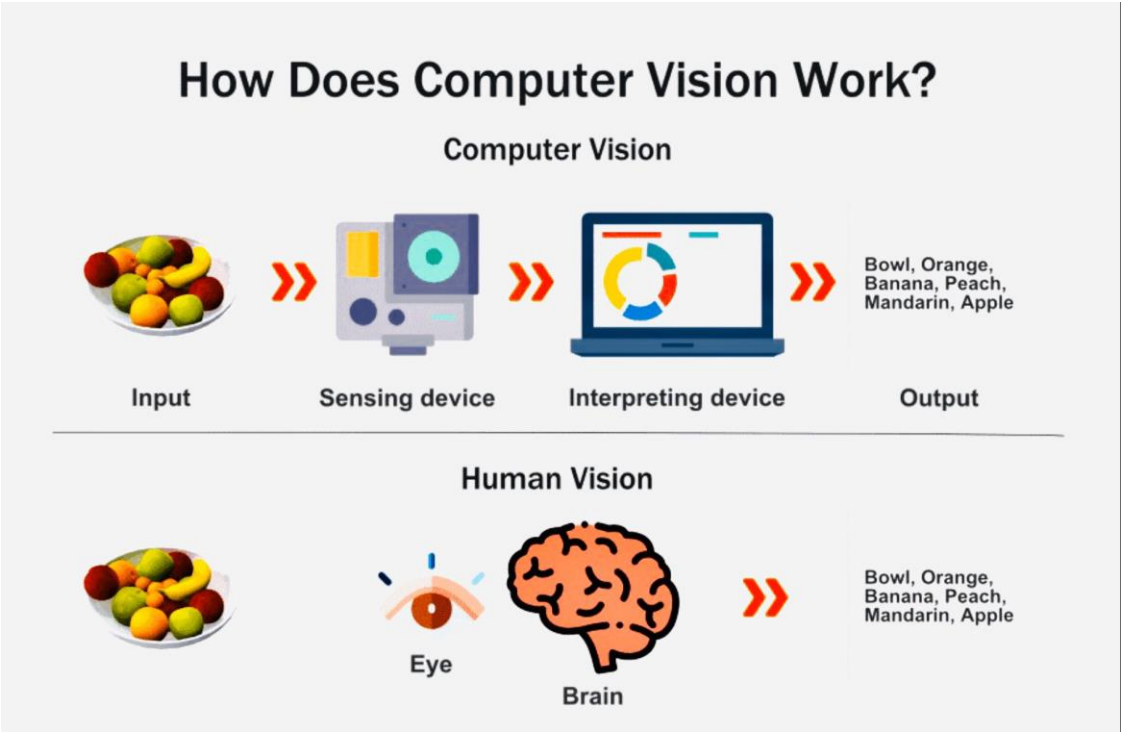
(c)

Binary Image

Click here

Reference link

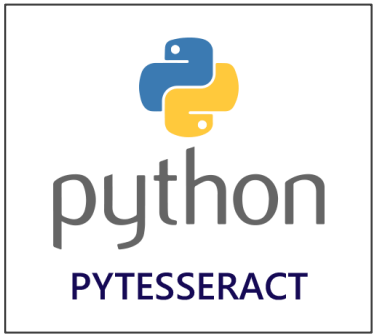
Working of Computer Vision



Click here

Reference link

Popular Python Libraries for CV



Basic Operations on Images using CV2

- Modify pixel values by gaining access to them.
- Image attributes may be accessed.
- Choosing an Image Region (ROI)
- Image Splitting and Merging

Click here

Reference link

Functions for accessing Image using OpenCV

Importing OpenCV library: **import cv2**

Loading an image : **img=cv2.imread(“cat.jpg”)**

Splitting and Merging Image Channels: When necessary, an image's B, G, and R channels can be split into their component planes. The different channels may then be merged to generate a BGR picture once more.

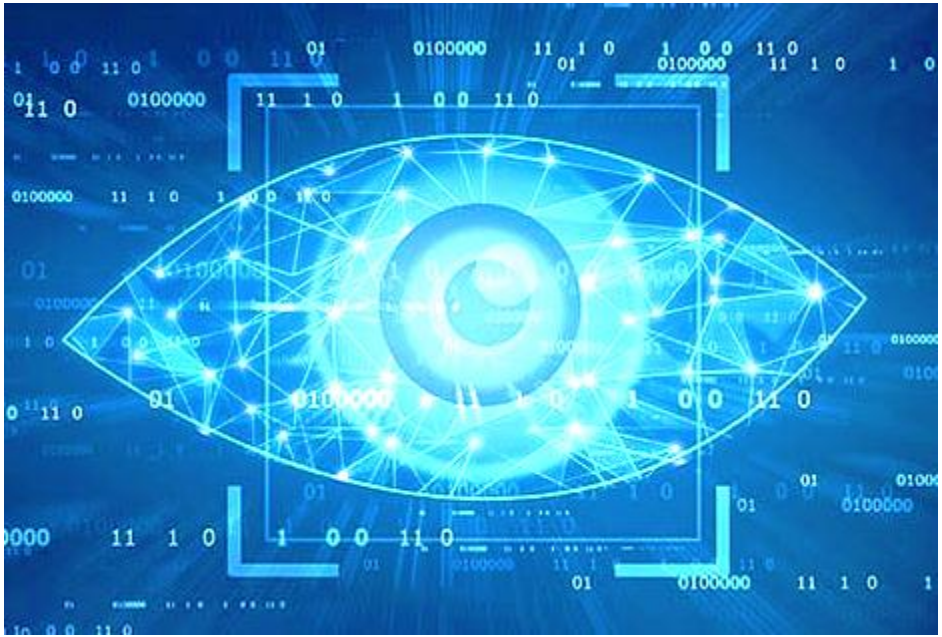
```
b,g,r=cv2.split(img)
img=cv2.merge((b,g,r))
```

Making Borders for Images (Padding): You may use the **cv2.copyMakeBorder()** method to build a border around an image, similar to a photo frame. However, it has additional uses for convolution operations, zero padding, and so forth.

Applications of Computer Vision(CV)

Computer vision may be used for a variety of purposes, including:

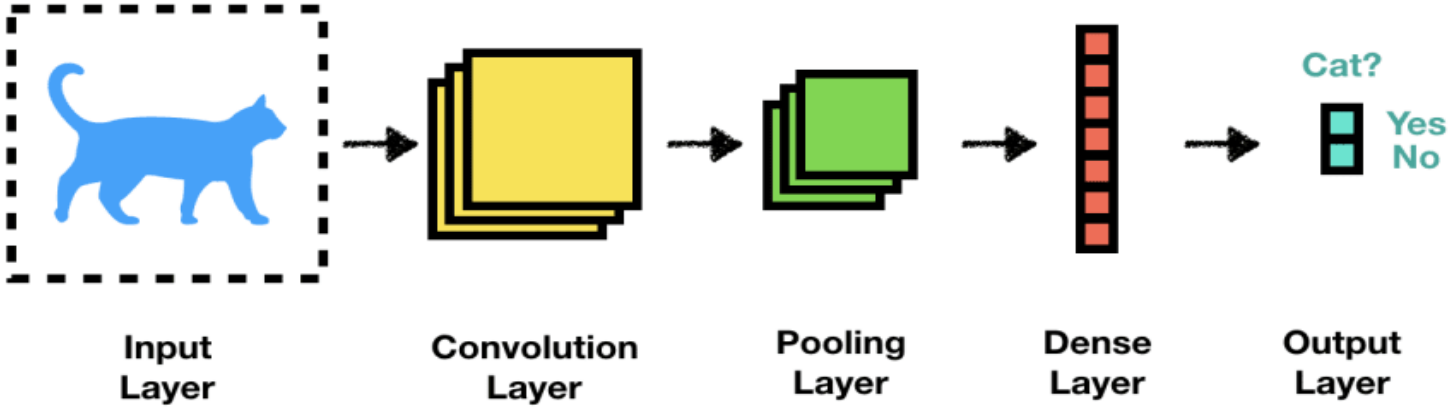
- CV plays a crucial part in both facial and iris recognition in biometric access management.
- CV lets industrial robots and self-driving automobiles to avoid accidents and travel safely.
- CV may be used in conjunction with other forms of artificial intelligence programming to automate the examination of X-rays and MRIs in digital diagnostics.
- CV allows mixed reality programs to know where a virtual item should be positioned using augmented reality.



Lab 1 – Perform Image Filters using Cv2

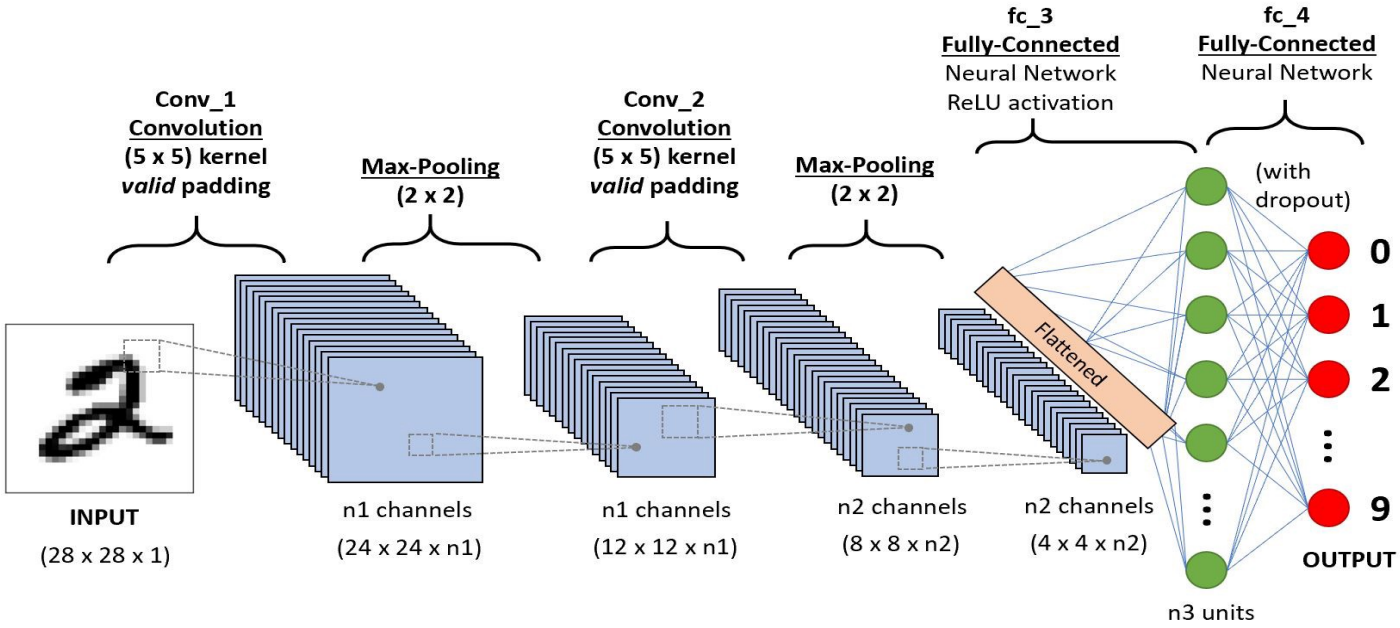
Convolutional Neural Network

- CNNs have stacked layered architecture of several Convolution and Pooling Layers



Reference link

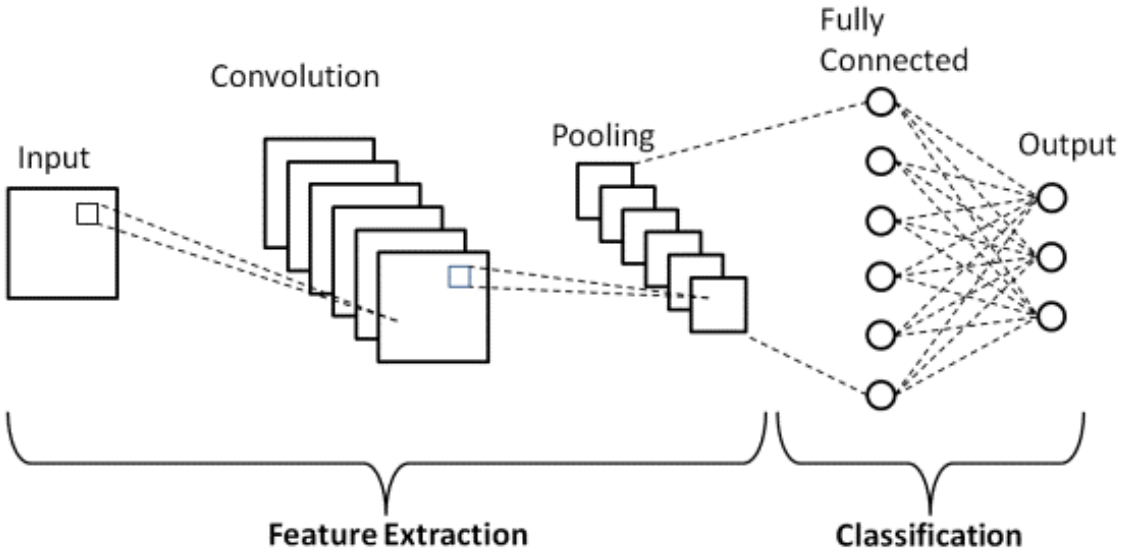
Convolutional Neural Network



Reference link

Understanding CNNs

- Convolutional Layers.
- Pooling Layers.
- Fully-Connected Layers.
- Padding

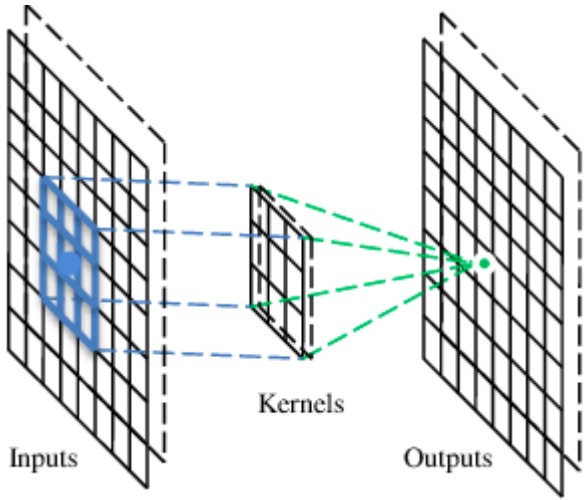


Click here

Reference link

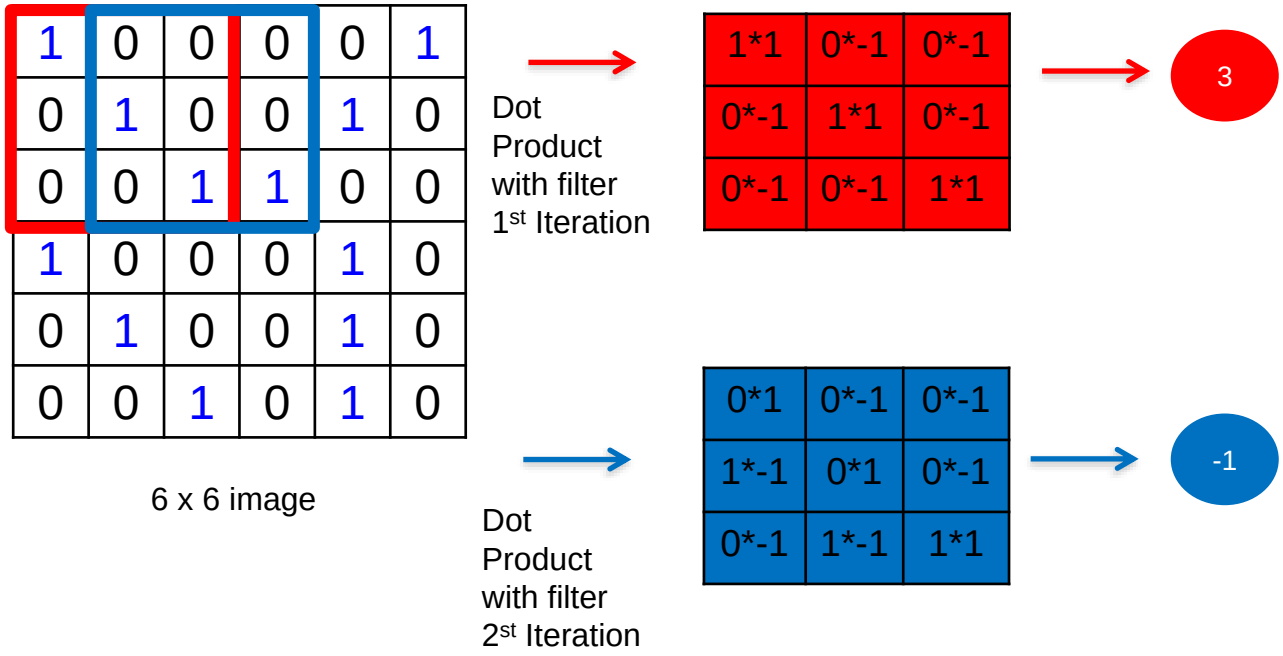
Convolutional Layer

- A CNN is a neural network with some convolutional layers (and some other layers).
- A convolutional layer has a number of filters that does convolutional operation.
- The convolutional layers are the key component of a CNN,
- Filters are applied to the input image to extract features such as edges, textures, and shapes.

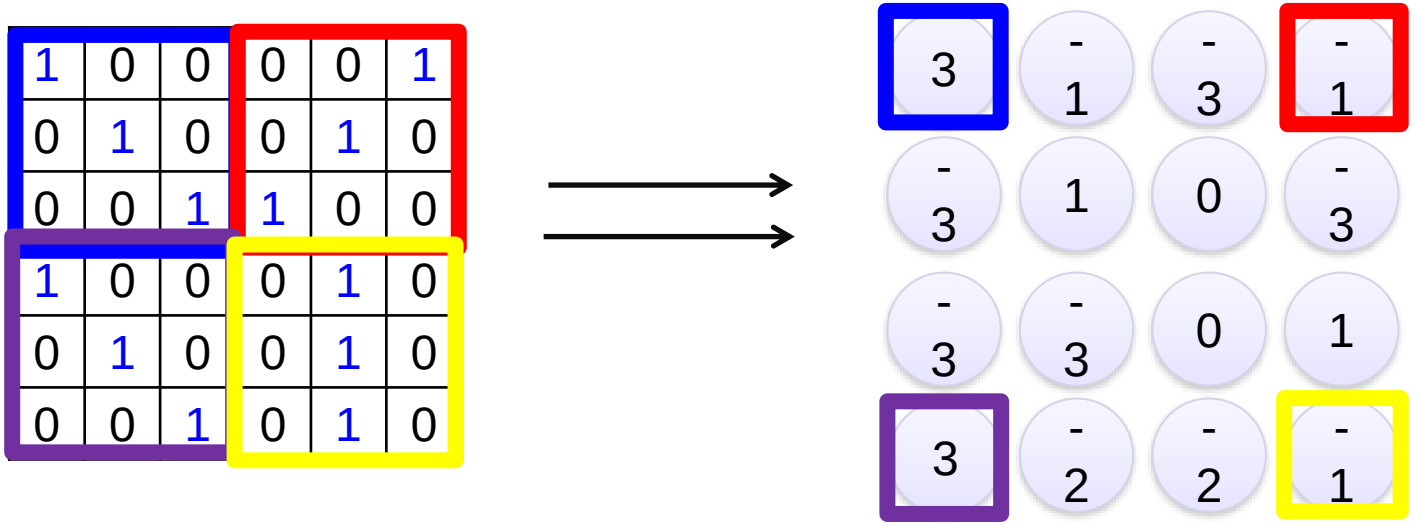


Reference link

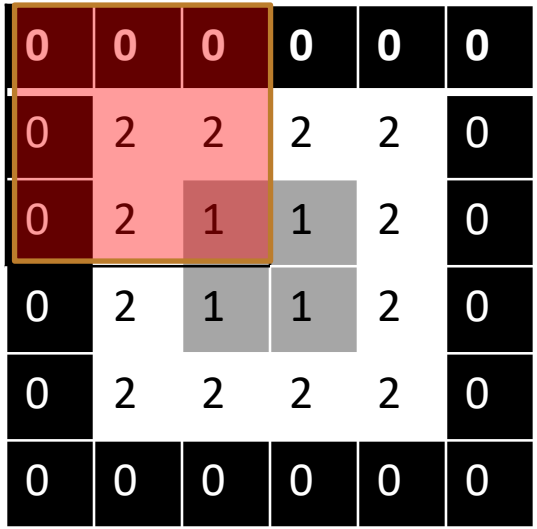
Convolutional Process



Convolutional Process



Filter



0	0	0
0	-1	1
0	1	-1

3*3 FILTER

0X0	0X0	0X0
0X0	2X-1	2X1
0X0	2X1	1X-1

Multiply by Filter Weights

0	0	0
0	-2	2
0	2	-1

SUM THE RESULT = 1

Stride Distance = (1,1)

Note: The resolution will decrease because we are taking 9 input values and one output value

Convolution Layers

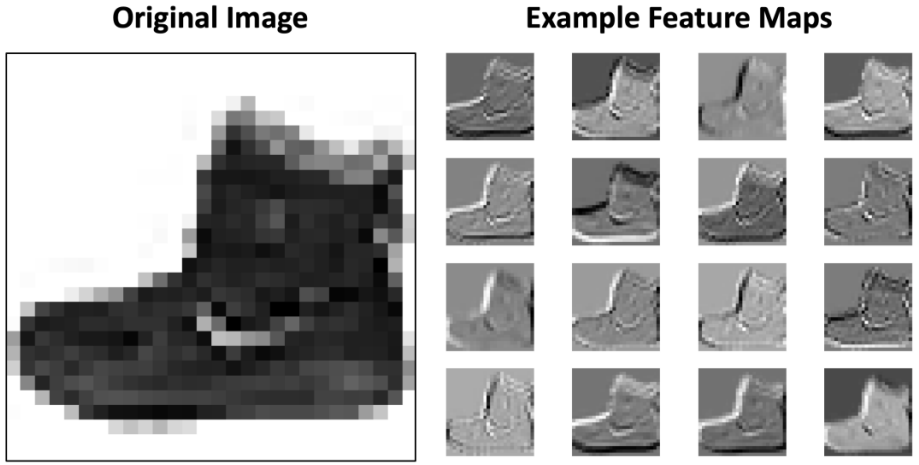
- Conv1D is used for input signals which are similar to the voice. By employing them you can find patterns across the signal. For instance, you have a voice signal and you have a convolutional layer. Each convolution traverses the voice to find meaningful patterns by employing a cost function.

- Conv2D is used for images. This use case is very popular. The convolution method used for this layer is so called convolution over volume. This means you have a two-dimensional image which contains multiple channels, RGB as an example. In this case, each convolutional filter should be a three-dimensional filter to be convolved, cross-correlated actually, with the image to find appropriate patterns across the image.

- Conv3D is usually used for videos where you have a frame for each time span. These layers usually have more parameters to be learnt than the previous layers. The reason we call them 3D is that other than images for each frame, there is another axis called time containing discrete values, and each of them corresponds to a particular frame.

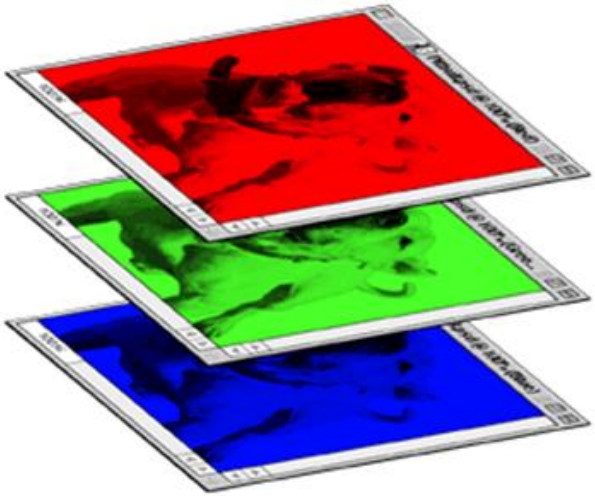
Convolutional Layers

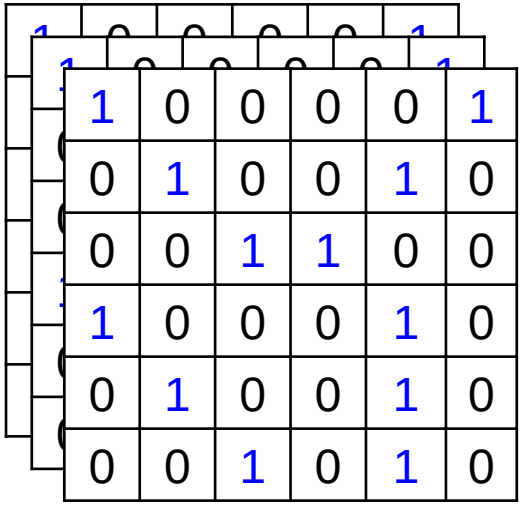
- Perform FEATURE EXTRACTION on Source Images by applying filters(eg. 16)
- Understand various pattern with regards to Image Structures (Textures, Edges, Corners and Patterns)
- They extract Feature Maps

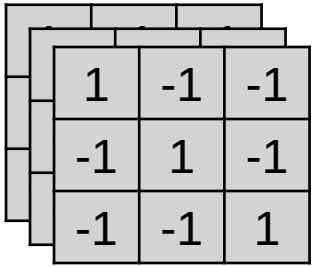


Color image: RGB 3 channels

Color image

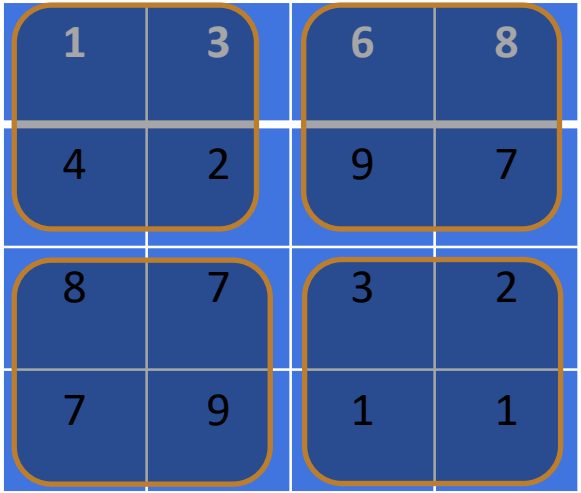






Filter

Max Pooling



Downsamples Feature maps from Conv Layers

4	9
9	3

Window 2x2 Stride : 2

Pooling Layers help in reducing Dimensionality after convolutions(compression)

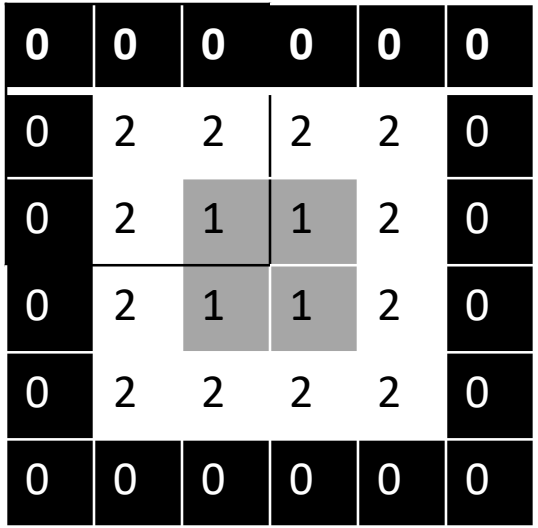
Padding

- Padding is a term relevant to convolutional neural networks as it refers to the amount of pixels added to an image when it is being processed by the kernel of a CNN. For example, if the padding in a CNN is set to zero, then every pixel value that is added will be of value zero.

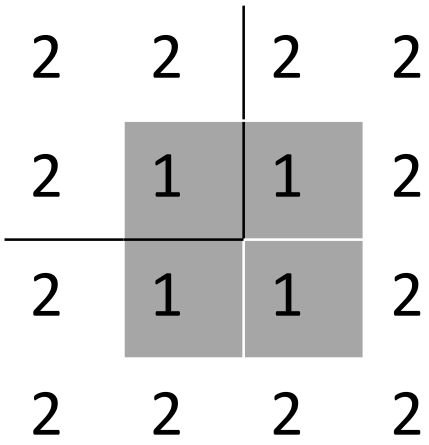
- **Same Padding** - In this type of padding, the padding layers append zero values in the outer frame of the images or data so the filter we are using can cover the edge of the matrix and make the inference with them too.

- **Valid Padding** - This type of padding can be considered as no padding

Padding



Same Padding



Valid Padding

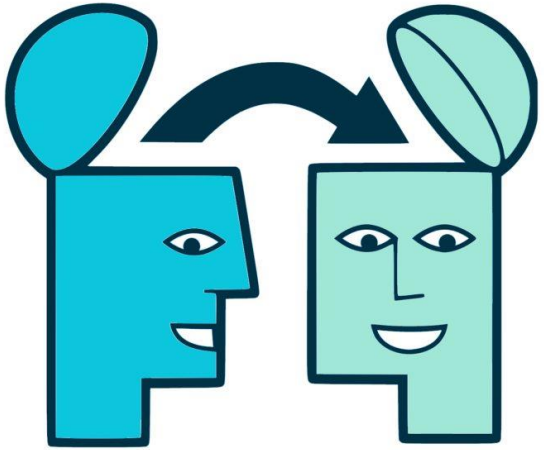
Lab 2 - Cat-Dog Classification using CNN

Transfer Learning

- In transfer learning we first train a base network on a base dataset and task, and then we repurpose the learned features, or transfer them, to a second target network to be trained on a target dataset and task.

- This process will tend to work if the features are general, that is, suitable to both base and target tasks, instead of being specific to the base task.

- In practice, very few people train an entire Convolutional Network from scratch because it is relatively rare to have a dataset of sufficient size.



Lab 3 – Object Identification Using Transfer Learning resnet50

Summary

- Computer vision (CV) is an artificial intelligence (AI) subcategory that focuses on developing and deploying digital systems that process, analyze, and interpret visual input.
- The objective of computer vision is to allow computers to recognize an item or person in a digital image and take appropriate action.
- Convolutional neural networks (CNNs) are used in computer vision to analyze visual input at the pixel level.
- Transfer learning involves the approach in which knowledge learned in one or more source tasks is transferred and used to improve the learning of a related target task.

Quiz

1) Computer vision is concerned with modelling and replicating human vision using computer software and hardware.

A. TRUE
B. FALSE
C. Can be true or false
D. Can not say

A) TRUE

Quiz

2) Which of the following is an Applications of Computer Vision?

A. Robotics
B. Medicine
C. Security
D. All of the above

D. All of the above

Quiz

3) _____in which we give input to our model.
A. Input layer
B. Output layer
C. Hidden layer
D. None

A. Input layer

Quiz

4) What is the default stride distance of Max Pooling

A) 3*3
B) 1*1
C) 0*0
D) 2*2

D). 2*2

Quiz

5) What is the default stride distance of Convolution Process

A) 3*3
B) 1*1
C) 0*0
D) 2*2

D). 1*1

Reference

- https://setosa.io/ev/image-kernels/
- https://towardsdatascience.com/understand-transposed-convolutions-and-build-your-own-transposed-convolution-layer-from-scratch-4f5d97b2967
- https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning-algorithm
- https://medium.com/voice-tech-podcast/text-classification-using-cnn-9ade8155dfb9

Thank you...!