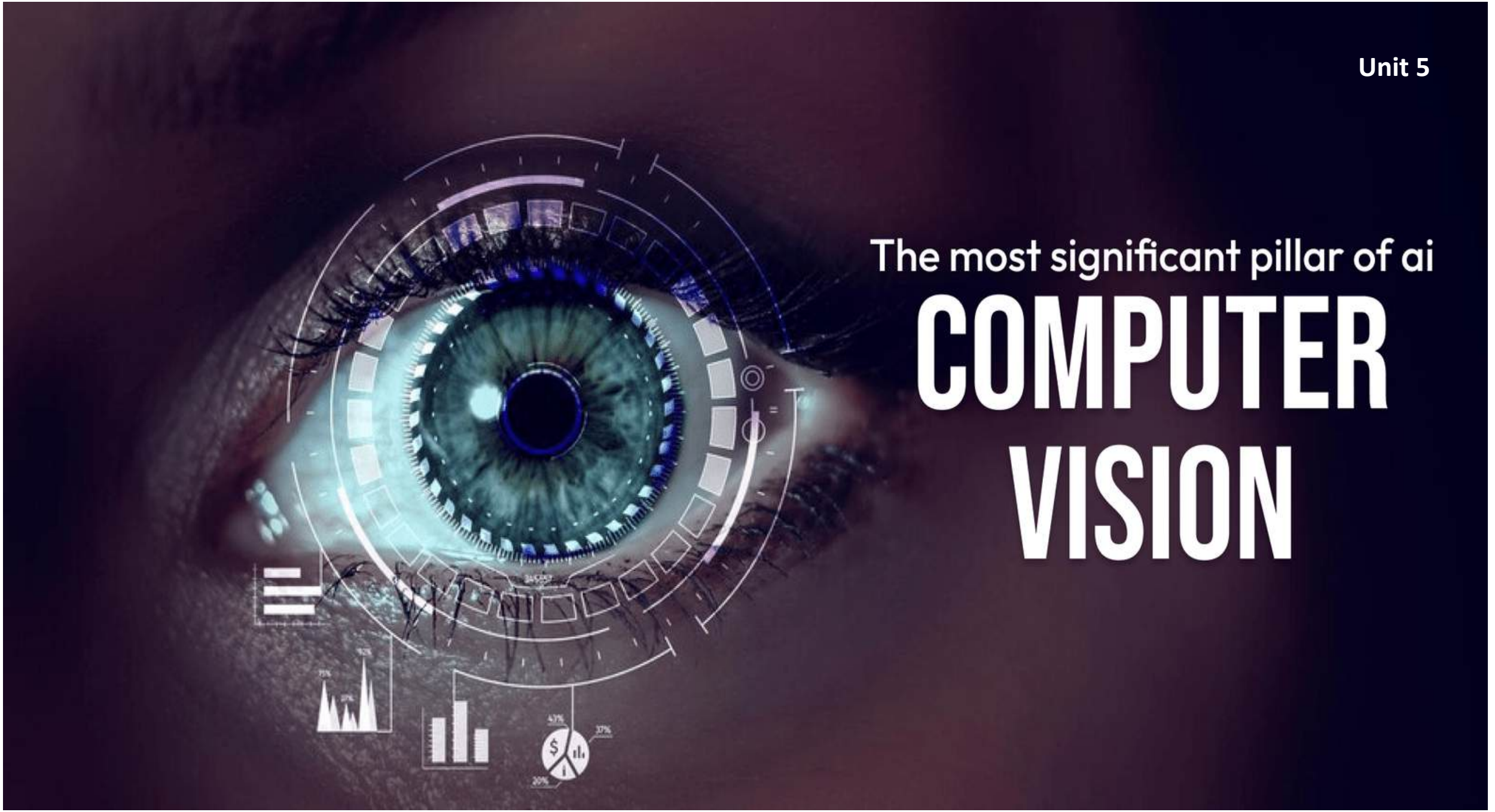
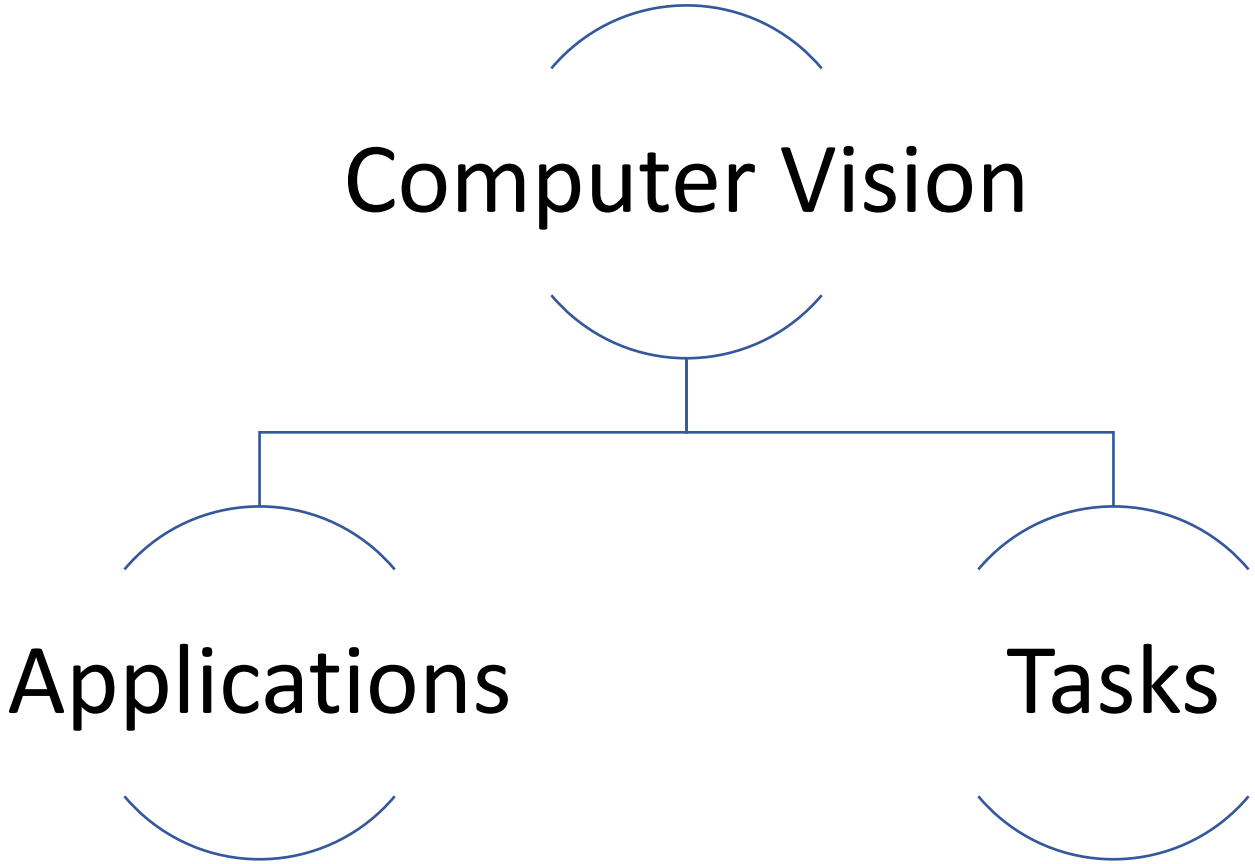
Unit 5



Introduction to

Computer Vision - CV

- Computer Vision is a branch of Artificial Intelligence (AI) that enables computers and systems to extract useful information from digital photos, videos, and other visual inputs and to execute actions or make recommendations based on that information.


Computer Vision
Applications
Tasks

Applications of Computer Vision

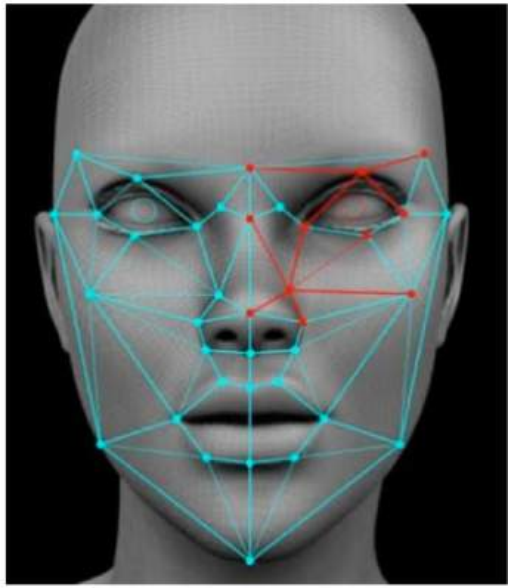
- The concept of computer vision was first introduced in the 1970s. Computer Vision technology advanced enough to make applications available to everyone at ease today. However, in recent years the world witnessed a significant leap in technology that has put computer vision on the priority list of many industries. Let us look at some of them:

1.Facial Recognition

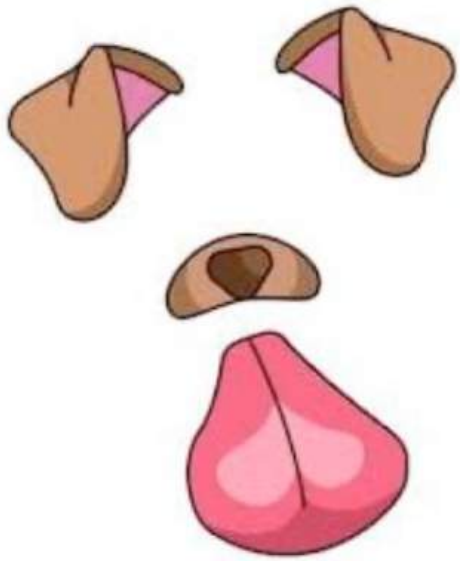
Facial recognition technology has many uses, including:

- **Security**: Facial recognition can be used to verify a person's identity, such as when checking in to a hotel, unlocking a phone, or boarding an airplane. It can also be used to identify potential shoplifters or deny entry to problem gamblers.
- **Law enforcement**: Facial recognition can be used to compare mugshots to databases to identify suspects.
- **Healthcare**: Facial recognition can be used to help healthcare providers access patient records and detect diseases.
- **Banking**: Facial recognition can be used to authenticate users of banking apps.
- **Education**: Facial recognition can be used to analyze students' facial micro expressions to help identify emotions like boredom, confusion, or frustration. This can help professors adjust their curriculum to better reflect student preferences.

Facial recognition compares the characteristics of one face with another or with a group of others. Comparing one face with another is called 1:1 verification, while comparing one face with a group is called identification.



2.Face Filters



Many of the functionality in today's apps, including Instagram and Snapchat, rely on computer vision. One of them is the usage of facial filters. The computer or algorithm may recognize a person's facial dynamics through the camera and apply the chosen facial filter.

3.Google's Search by Image

The majority of data that is searched for using Google's search engine is textual information, but it also has the intriguing option of returning search results via an Image. This makes use of computer vision since it examines numerous attributes of the input image while also comparing them to those in the database of images to provide the search result.

4.Computer Vision in Retail

One of the industries with the quickest growth is retail, which is also utilizing computer vision to improve the user experience. Retailers can analyze navigational routes, find walking patterns, and track customer movements through stores using computer vision techniques.

5.Self-Driving Cars

Computer Vision is the fundamental technology behind developing autonomous vehicles. Most leading car manufacturing in the world are reaping the benefits of investing in artificial intelligence for developing on-road version of hands-free technology

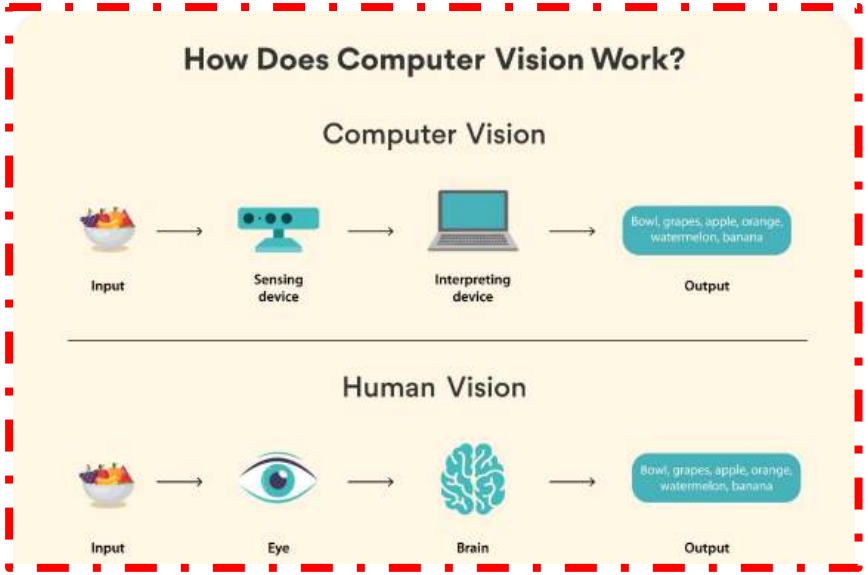
6.Medical Imaging

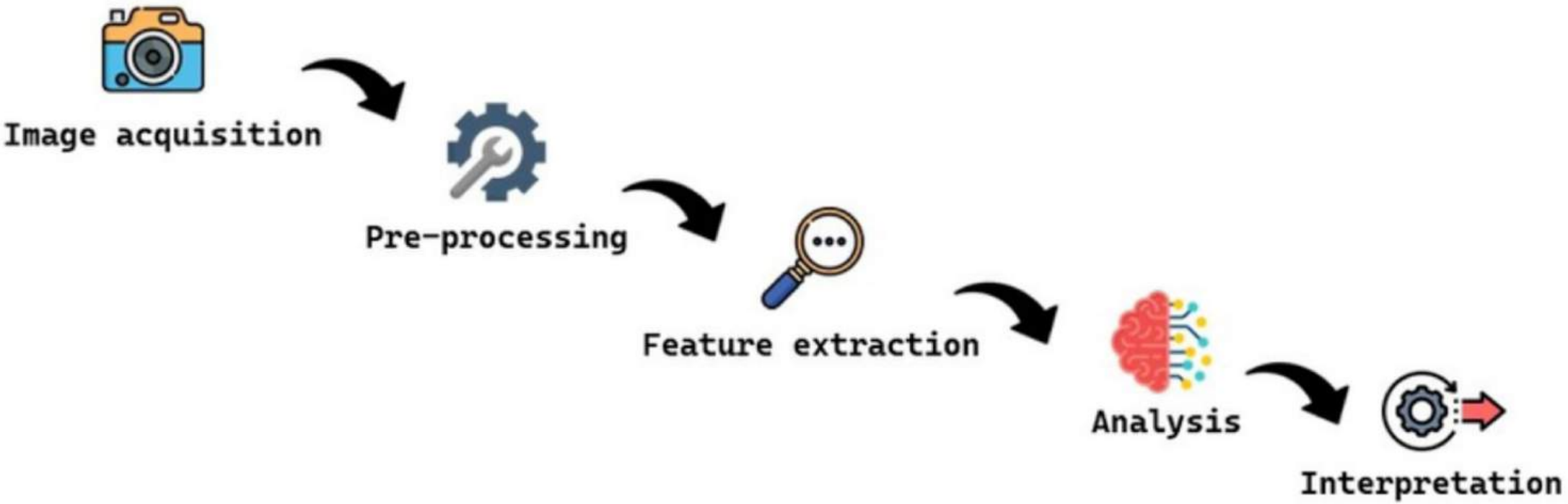
A reliable resource for doctors over the past few decades has been computer-supported medical imaging software. It doesn't just produce an analyze images; it also works as a doctor's helper to aid in interpretation. The software is used to interpret and transform 2D scan photos into interactive 3D models that give medical professionals a thorough insight of a patient's heath.

7.Google Translate App

To read signs written in a foreign language, all you have to do is point the camera on your phone at the text, and the Google Translate software will immediately translate them into the language of your choice. This is a useful application that makes use of Computer Vision, utilizing optical character recognition to view the image and Augmented reality to overlay an accurate translation

Computer Vision - Process



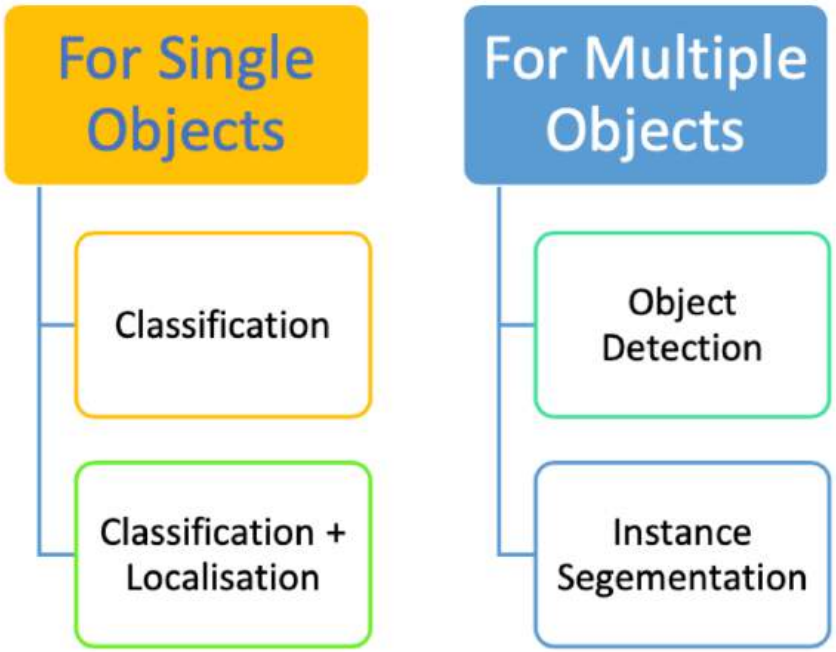


Tasks

The many computer vision applications are based on a variety of tasks that are carried out to extract specific information from the input image that may be utilized for prediction or serves as the foundation for additional analysis. A computer vision application performs the following tasks.

1.Classification

Image classification problem is the tasks of assigning an input image one label from a fixed set of categories. This is one of the core problems in CV that, despite its simplicity, has a large variety of practical applications.



2.Classification + Localization

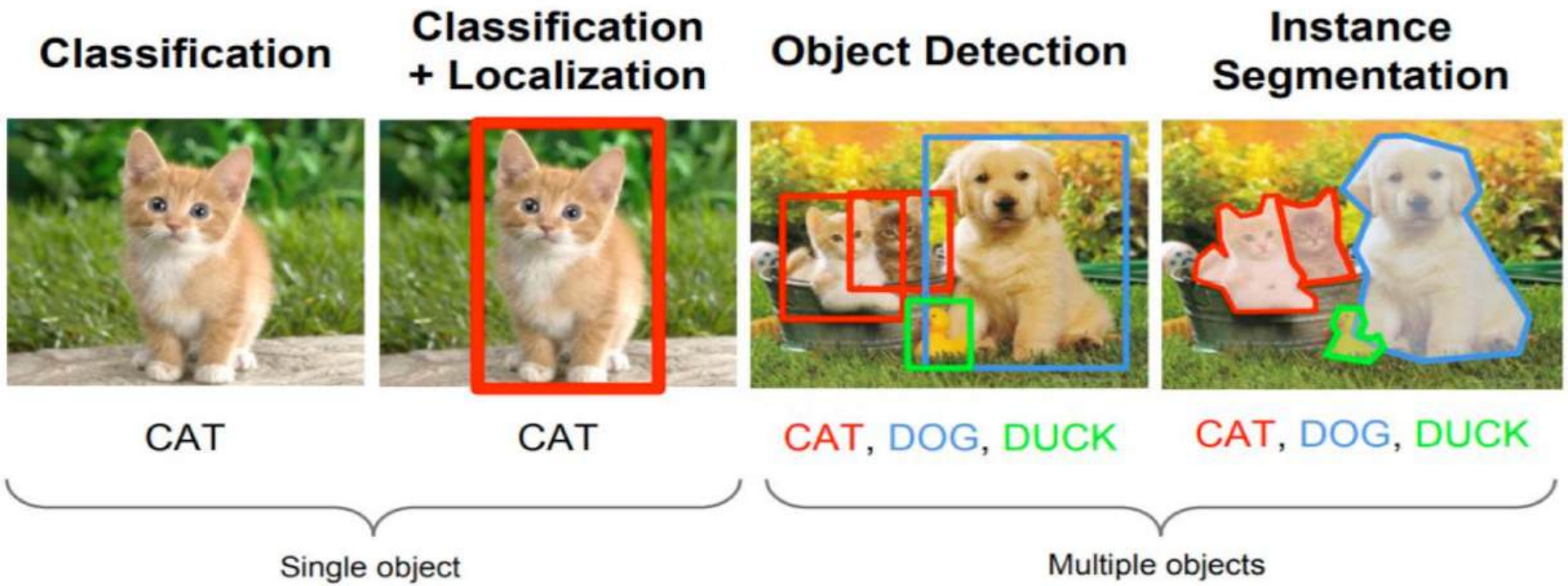
This is the task which involves both processes of identifying what object is present in the image and at the same time identifying at what location that object is present in that image. It is used only for single objects.

3.Object Detection

Object detection is the process of finding instances of real-world objects such as faces, bicycles, and buildings in images or videos. Object detection algorithms typically use extracted features and learning algorithms to recognize instances of an object category. It is commonly used in applications such as image retrieval and automated vehicle parking systems.

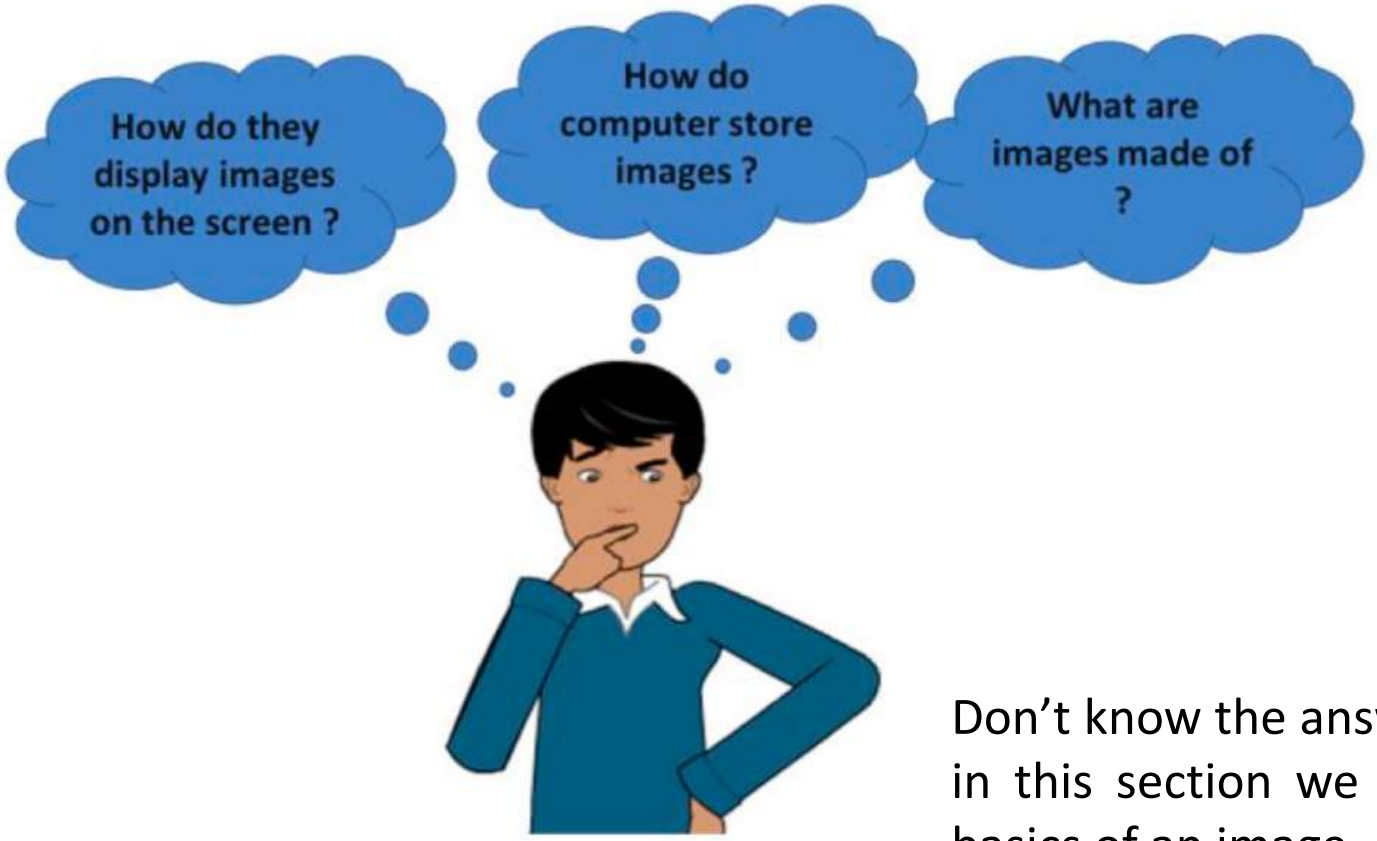
4.Instance Segmentation

Instance Segmentation is the process of detecting instances of the objects, giving them a category and then giving each pixel a label on the basis of that. A segmentation algorithm takes an image as input and outputs a collection of regions (or segments).



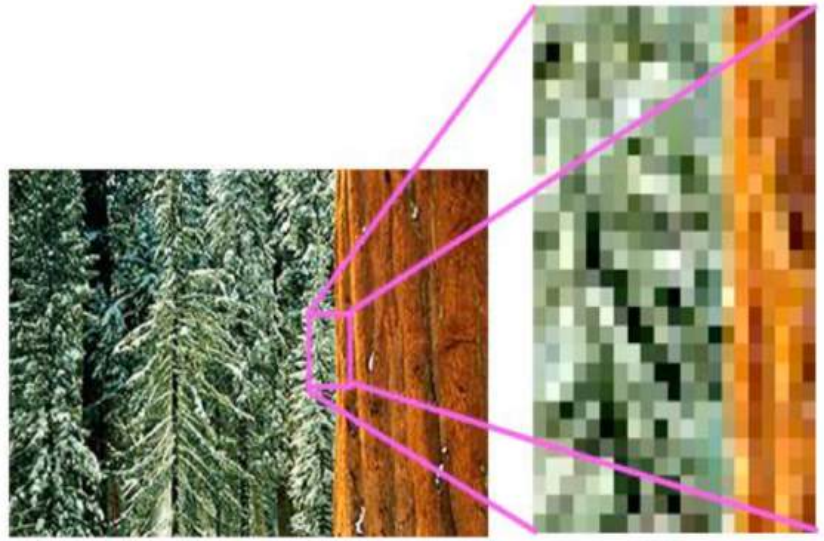
Basics of Images

We all see a lot of images around us and use them daily either through our mobile phones or computer system. But do we ask some basic questions to ourselves while we use them on such a regular basis.



Don’t know the answer yet? Don’t worry, in this section we will study about the basics of an image

Basics of Pixels



The word “pixel” means a picture element. Every photograph, in digital form, is made up of pixels. They are the smallest unit of information that make up a picture. Usually round or square, they are typically arranged in a 2-dimensional grid.

Resolution

Resolution refers to the number of pixels in an image. It can be expressed as width by height (e.g., 1280×1024), or as a single number (e.g., 5 megapixels, where one megapixel equals one million pixels). For example, a 1280×1024 monitor has about 1.31 megapixels (1280 x 1024 = 1,310,720 pixels).

Pixel value

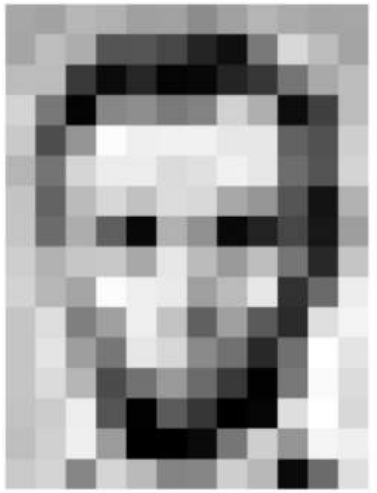
Each pixel in an image has a value that defines its brightness and color. In the common 8-bit format, this value ranges from 0 to 255, where 0 represents black and 255 represents white. This range comes from the binary system, where each pixel uses 1 byte (8 bits), and each bit can be either 0 or 1, allowing for 256 possible values (0 to 255).

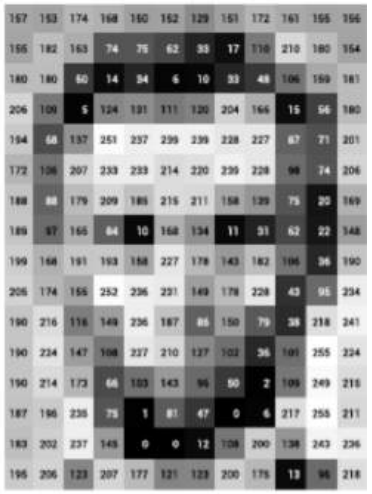
Number of bits	Different patterns	No. of patterns	No. of patterns
1	0 1	2^1	2
2	00 01 10 11	2^2	4
3	000 001 010 100 011 101 110 111	2^3	8

$2^8 = 256$

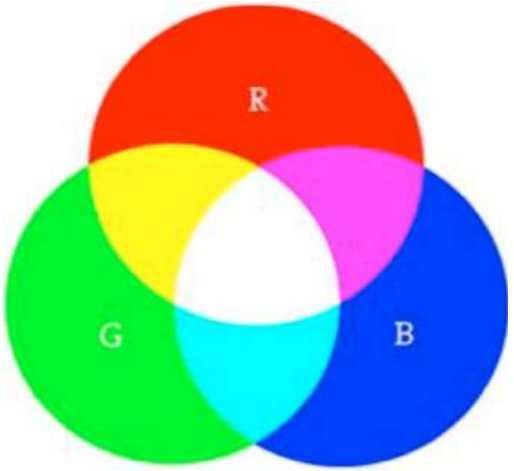
Grayscale images

Grayscale images display varying shades of gray without color, ranging from black (pixel value 0) to white (pixel value 255). Intermediate shades are created by equal brightness of the three primary colors. Each pixel in a grayscale image is 1 byte, arranged in a 2D array. The image size is defined by its height and width





RGB Images



All the images that we see around are colored images. These images are made up of three primary colors Red, Green and Blue. All the colors that are present can be made by combining different intensities of red, green and blue.

Let us experience!

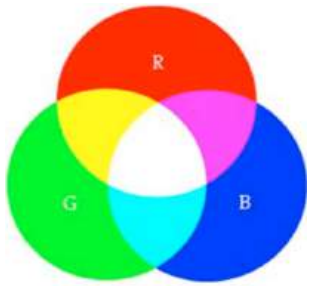
Go to this online link https://www.w3schools.com/colors/colors_rgb.asp.

Image Features

In Computer Vision and image processing, a feature is a piece of information which is relevant for solving the computation tasks related to a certain application. Features may be specific structures in the image such as points, edges or objects.

Activity #1

RGB Images



1) What is the output colour when you put R=G=B=255 ?

2) What is the output colour when you put R=G=B=0 ?

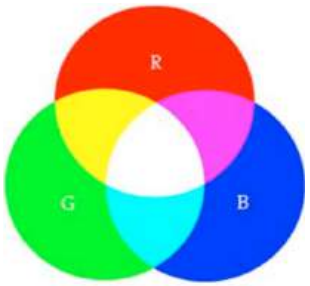
3) How does the colour vary when you put either of the three as 0 and then keep on varying the other two?

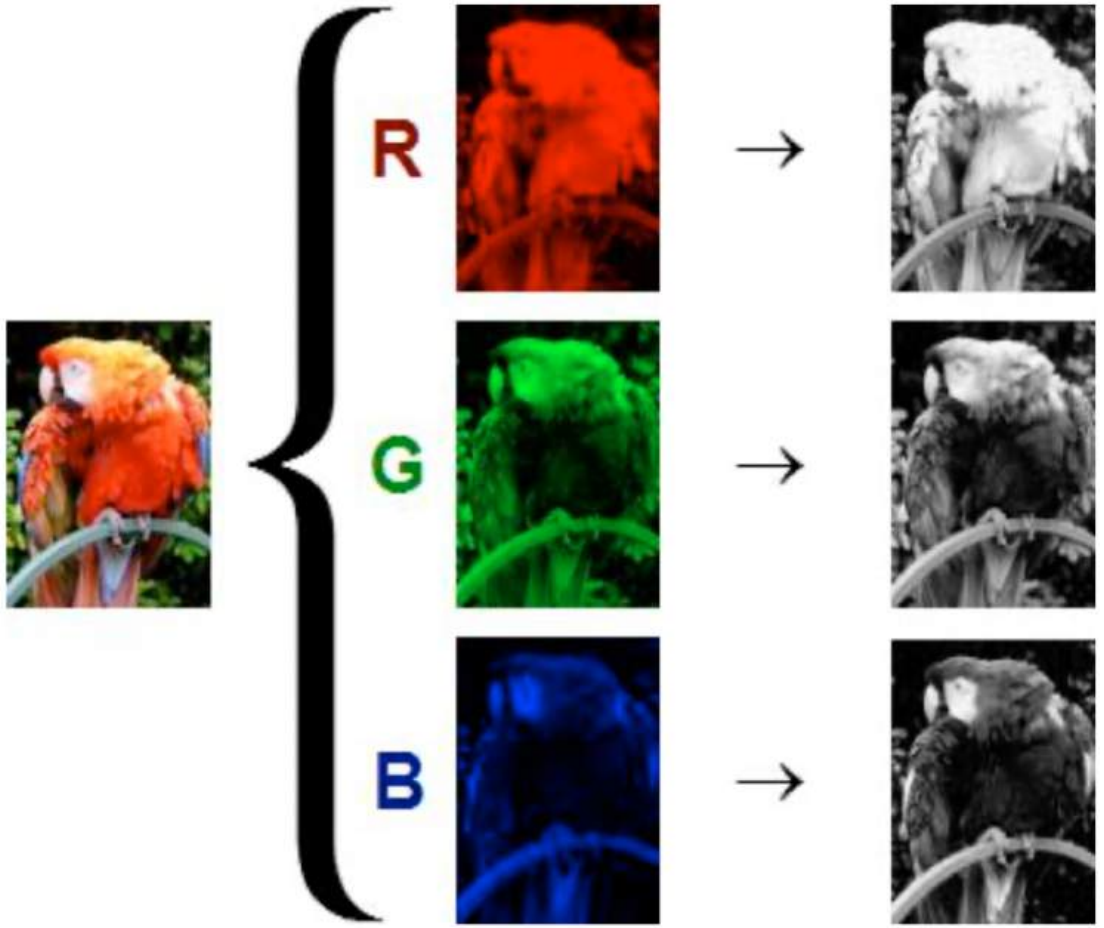
4) How does the output colour change when all the three colours are varied in same proportion ?

5) What is the RGB value of your favourite colour from the colour palette?

Activity # 2

RGB Images





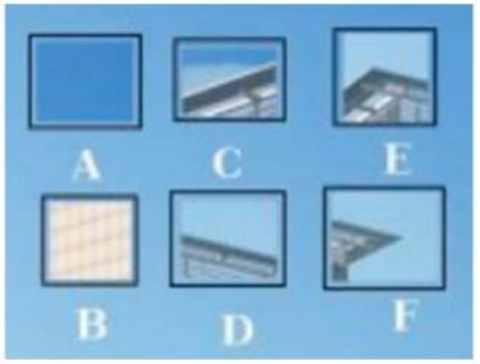
Task :

Go to the following link www.piskelapp.com and create your own pixel art. Try and make a GIF using the online app for your own pixel art.

Image Features

In computer vision and image processing, a **feature** is a piece of information which is relevant for solving the computational task related to a certain application. Features may be specific structures in the image such as points, edges or objects.

For example: Imagine that your security camera is capturing an image. At the top of the image we are given six small patches of images. Our task is to find the exact location of those image patches in the image. Take a pencil and mark the exact location of those patches in the image.



Introduction to

OpenCV

OpenCV (Open Source Computer Vision Library) is a tool used to extract image features for processing. It supports image and video analysis, enabling tasks like object, face, and handwriting recognition.

Convolution and Kernel are computational techniques used in image processing, machine vision, and signal processing:

- ✓ **Convolution**: A technique that involves splitting an image into small pieces and extracting features from each piece.
- ✓ **Kernel**: A 2D structure that defines how a filter alters the pixel values in an image. It's also known as a convolution matrix or mask.

Summary:

1.Convolution is a common tool used for image editing.
2.It is an element wise multiplication of an image and a kernel to get the desired output.
3.In computer vision application, it is used in Convolutional Neural Network (CNN) to extract image features.

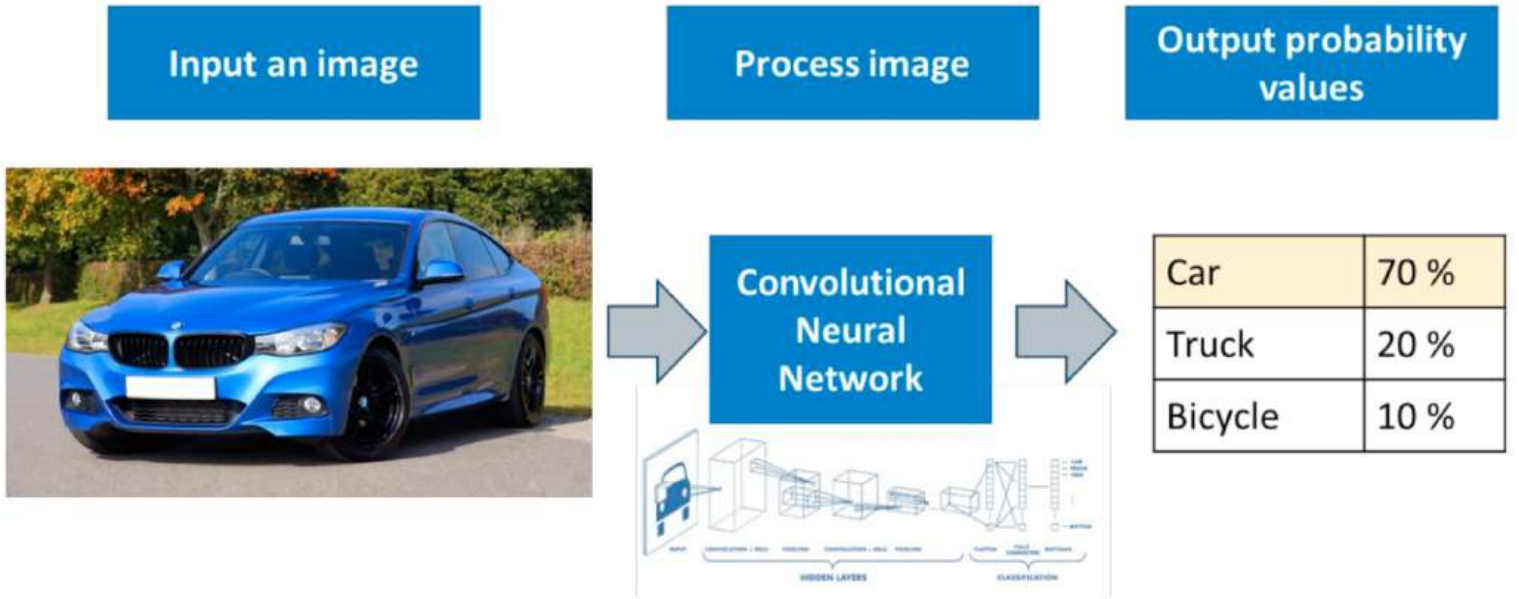
Introduction to

Convolution Neural Networks (CNN)

What is a Convolutional Neural Network?

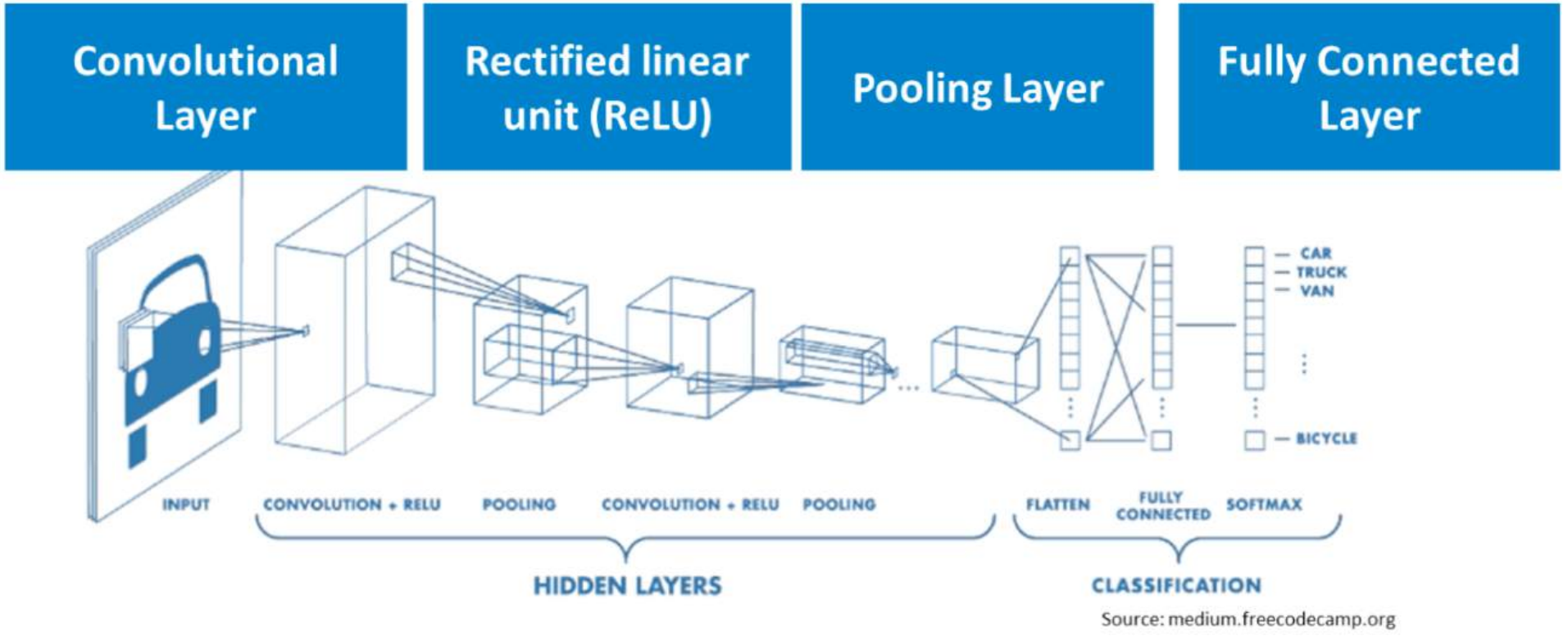
A **Convolutional Neural Network (CNN)** is a *Deep Learning algorithm* which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other.

The process of deploying a CNN is as follows:

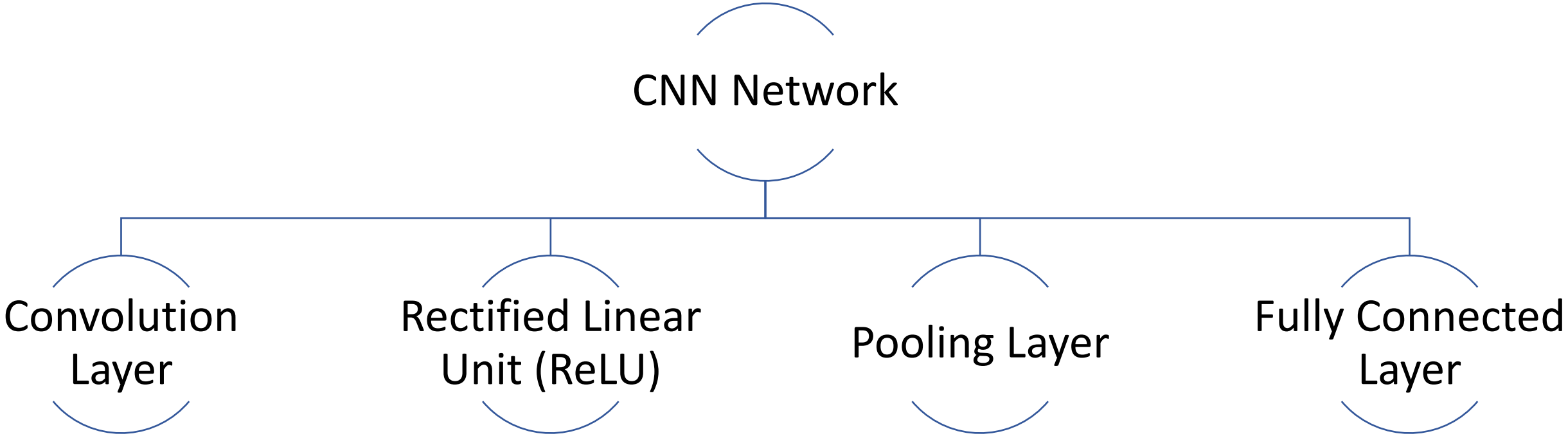


In the above diagram, we give an input image, which is then processed through a CNN and then gives prediction on the basis of the label given in the particular dataset.

The different layers of a Convolutional Neural Network (CNN) is as follows:



A Convolutional Neural Network consists of the following layers:



Convolution Layer

The first layer of a CNN is the **Convolutional Layer**, which **extract the high-level features such as edges, and also low-level features such as color, gradient orientation** from the input image. As more layers are added, the network learns higher-level features, giving it a deeper understanding of the image. The layer uses multiple **kernels** to produce **feature maps** (also called **activation maps**), which help in reducing image size and focusing on relevant features. For example, recognizing facial features like eyes and nose may be enough for identifying a person, rather than the entire face.

Rectified Linear Layer

After we get the feature map, it is then passed onto the ReLU layer. This layer simply gets rid of all the negative numbers in the feature map and let the positive number stay as it is.

Pooling Layer

Pooling layer is responsible for reducing the spatial size of the Convolved feature while still retaining the important features. There are two types of pooling which can be performed on an Image.

1. **Max Pooling**: Selects the max value from a small region of the feature map, keeping the most prominent feature.
2. **Average Pooling**: Computes the average value of the same small region, summarizing the features by their average.

Fully Connected Layer

The objective of a fully connected layer is to take the results of the convolution / pooling process and use them to classify the image into a label. Example: if the image is of a cat, features representing things like whiskers or fur should have high probabilities for the lalel “cat”.

Unit 6 - NLP

Important Questions

Computer vision is concerned with modeling and replicating human vision using computer software and hardware.

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

Explanation: True, Computer vision is concerned with modeling and replicating human vision using computer software and hardware.

Computer vision is a discipline that studies how to reconstruct, interrupt and understand a 3d scene from its _____.

A.1d images **B. 2d images** C. 3d images D. 4d images

Explanation: Computer vision is a discipline that studies how to reconstruct, interrupt and understand a 3d scene from its 2d images.

3. Computer vision is divided into _____ basic categories

A. 1 B. 2 **C. 3** D. 4

Explanation: Computer vision is divided into three basic categories : ***Low-level vision***, ***Intermediate-level vision*** and ***High-level vision***.

Which vision includes object recognition and 3D scene Interpretation?
A. Low-level vision B. ***Intermediate-level vision*** C. High-level vision D. All of the above

Explanation: Intermediate-level vision: includes object recognition and 3D scene Interpretation

The input and output of image processing are?
A. signal and image B. signal only C. **image only** D. depends on input

Explanation: Image processing studies image to image transformation. The input and output of image processing are both images.

Which of the following studies various techniques to classify patterns?
A. Image Processing B. Photogrammetry C. Image Recognition D. **Pattern Recognition**

Explanation: Pattern Recognition: it studies various techniques to classify patterns

VERY SHORT ANSWER QUESTIONS (2 MARKS)

What is Resolution ?

Ans : The resolution of an image is occasionally referred to as the number of pixels. The resolution of an image is occasionally referred to as the number of pixels.

2. What is RGB ?

Ans : Every image we encounter is a colored image. Three main colors—Red, Green, and Blue— make up these graphics.

SHORT ANSWER QUESTIONS (3 MARKS)

1. What is Object Detection ?

Ans : Finding occurrences of real-world items like faces, bicycles, and buildings in pictures or movies is a process known as object detection. To identify occurrences of a certain object category, object identification algorithms frequently employ extracted features and learning techniques.

2. What is instance segmentation ?

Ans : The process of identifying instances of the items, categorising them, and then assigning each pixel a label based on that is known as instance segmentation.

LONG ANSWER QUESTIONS (4 MARKS)

Differentiate the terms Facial Recognition and Face Filters.

Computer vision is essential to the advancement of the home in the era of smart cities and smart homes. The most crucial application of computer vision is facial recognition in security. Either visitor identification or visitor log upkeep is possible.

Face Filters : Many of the functionality in today's apps, including Instagram and Snapchat, rely on computer vision. One of them is the usage of facial filters. The computer or algorithm may recognize a person's facial dynamics through the camera and apply the chosen facial filter.


GOOD
LUCK
IN YOUR
EXAM