

**SUB-TOPIC- Introduction of Computer Vision**

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.

**DETAILED EXPLANATION : Applications of Computer Vision**

In the 1970s, computer vision as a concept was first introduced. Everyone was excited by the new uses for computer vision. However, a considerable technological advance in recent years has elevated computer vision to the top of many companies' priority lists. Let's examine a few of them:

**Facial Recognition**

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 recognise a person's facial dynamics through the camera and apply the chosen facial filter.

**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.

**Computer Vision in Retail**

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

**Self-Driving Cars**

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

**Medical Imaging**

A reliable resource for doctors over the past few decades has been computer-supported medical imaging software. It doesn't just produce and analyse 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 health.

**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 very immediately translate them into the language of your choice. This is a useful application that makes use of Computer Vision, utilising optical character recognition to view the image and augmented reality to overlay an accurate translation.

**Computer Vision 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 utilised for prediction or serves as the foundation for additional analysis. A computer vision application performs the following tasks:

**Classification**

Image Classification problem is the task 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.

### **Classification + Localisation**

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.

### **Object Detection**

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. Applications like image retrieval and automatic car parking systems frequently employ it.

### **Instance Segmentation**

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. An image is sent into a segmentation algorithm, which produces a list 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.

### **Basics of Pixels**

A picture element is referred to as a “pixel.” In digital form, pixels make up each and every image. They are the tiniest piece of data that go into a picture. They are normally structured in a 2-dimensional grid and are either circular or square.

### **Resolution**

The resolution of an image is occasionally referred to as the number of pixels. One approach is to define resolution as the width divided by the height when the phrase is used to describe the number of pixels, for example, a monitor resolution of 1280×1024. Accordingly, there are 1280 pixels from side to side and 1024 pixels from top to bottom.

### **Pixel value**

Each of the pixels that make up an image that is stored on a computer has a pixel value that specifies its brightness and/or intended colour. The byte image, which stores this number as an 8-bit integer with a possible range of values from 0 to 255, is the most popular pixel format. Zero is typically used to represent no colour or black, and 255 is used to represent full colour or white.

### **Grayscale Images**

Grayscale images are images which have a range of shades of gray without apparent colour. The darkest possible shade is black, which is the total absence of colour or zero value of pixel. The lightest possible shade is white, which is the total presence of colour or 255 value of a pixel. Intermediate shades of gray are represented by equal brightness levels of the three primary colours.

### **RGB Images**

Every image we encounter is a coloured image. Three main colors—Red, Green, and Blue—make up these graphics. Red, green, and blue can be combined in various intensities to create all the colours that are visible.

### **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.

### **Introduction to OpenCV**

OpenCV or Open Source Computer Vision Library is that tool which helps a computer extract these features from the images. It is used for all kinds of images and video processing and

analysis. It is capable of processing images and videos to identify objects, faces, or even handwriting.

### **What is a Kernel?**

A Kernel is a matrix, which is slid across the image and multiplied with the input such that the output is enhanced in a certain desirable manner. Each kernel has a different value for different kind of effects that we want to apply to an image.

### **Convolution Neural Networks (CNN)**

A Convolutional Neural Network (CNN) is a Deep Learning algorithm that can take in an image as input, assign importance (learnable weights and biases) to various elements and objects in the image, and be able to distinguish between them.

A convolutional neural network consists of the following layers:

- 1) Convolution Layer
- 2) Rectified linear Unit (ReLU)
- 3) Pooling Layer
- 4) Fully Connected Layer

#### **Convolution Layer**

The Convolution Operation's goal is to take the input image's high-level characteristics, such edges, and extract them. There is no requirement that CNN use just one Convolutional Layer. There are several kernels that are used to produce several features. The output of this layer is called the feature map. A feature map is also called the activation map. We can use these terms interchangeably. There's several uses we derive from the feature map: We reduce the image size so that it can be processed more efficiently. We only focus on the features of the image that can help us in processing the image further.

#### **Rectified Linear Unit Function**

The next layer in the Convolution Neural Network is the Rectified Linear Unit function or the ReLU 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 lets the positive number stay as it is.

#### **Pooling Layer**

Similar to the Convolutional Layer, the 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 : Max Pooling returns the maximum value from the portion of the image covered by the Kernel.
- 2) Average Pooling: Max Pooling returns the maximum value from the portion of the image covered by the Kernel.

#### **Fully Connected Layer**

The final layer in the CNN is the Fully Connected Layer (FCP). 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.

### **MCQ'S ( Computer Vision )**

1. 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

Ans : A

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

2. 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

Ans : B

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 Ans : C

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

4. 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 Ans B Explanation: Intermediate-level vision: includes object recognition and 3D scene Interpretation

5. The input and output of image processing are?

A. signal and image B. signal only C. image only D. depends on input Ans : C

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

6. Which of the following studies various techniques to classify patterns?

A. Image Processing B. Photogrammetry C. Image Recognition D. Pattern Recognition Ans : D

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

#### VERY SHORT ANSWER QUESTIONS (2 MARKS )

1. 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 coloured 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)

1. **Differentiate the terms Facial Recognition and Face Filters.**

Ans : 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 recognise a person's facial dynamics through the camera and apply the chosen facial filter.