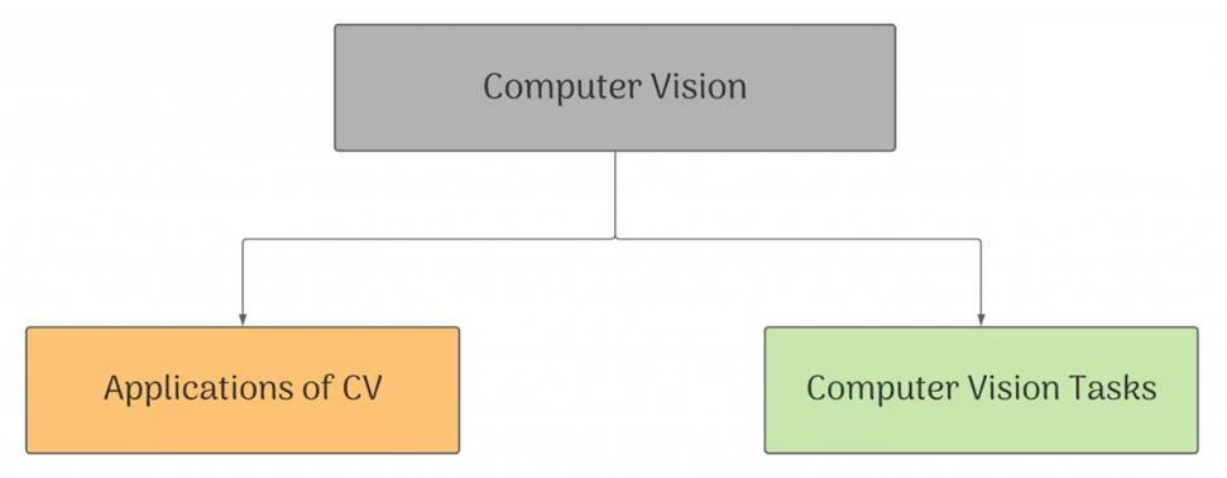
UNIT-5: COMPUTER VISION

Introduction to Computer Vision

Computer vision is a branch in the **Domain of AI** that enables computers to analyze meaningful information from images, videos, and other visual inputs.

Computer vision is the same as the human eye, it enables us see-through images or visual data, process and analyses them on the basis of algorithms and methods in order to analyse actual phenomena with images.



Applications of Computer Vision

This decade and the upcoming one can witness a significant leap in technology that has put computer vision on the priority list. Some common uses of Computer Vision are:

Facial recognition

The most frequently used technology is smartphones. It is a technology to remember and verify a person, object, etc from the visuals from the given pre-defined data. Such kinds of mechanics are often used for security and safety purposes.

For eg: Face security lock-in devices and traffic cameras are some examples using facial recognition.

Facial filters

Modern days social media apps like Snapchat and Instagram use such kinds of technology that extract facial landmarks and process them using AI to get the best result.

Goggle lens

To search data, Google uses Computer vision for capturing and analysing different features of the input image to the database of images and then gives us the search.

Automotive

The machinery in industries is now using Computer vision. Automated cars are equipped with sensors and software which can detect the 360 degrees of movements determine the location, detect objects and establish the depth or dimensions of the virtual world.

For eg: Companies like Tesla are now interested in developing self-driving cars.

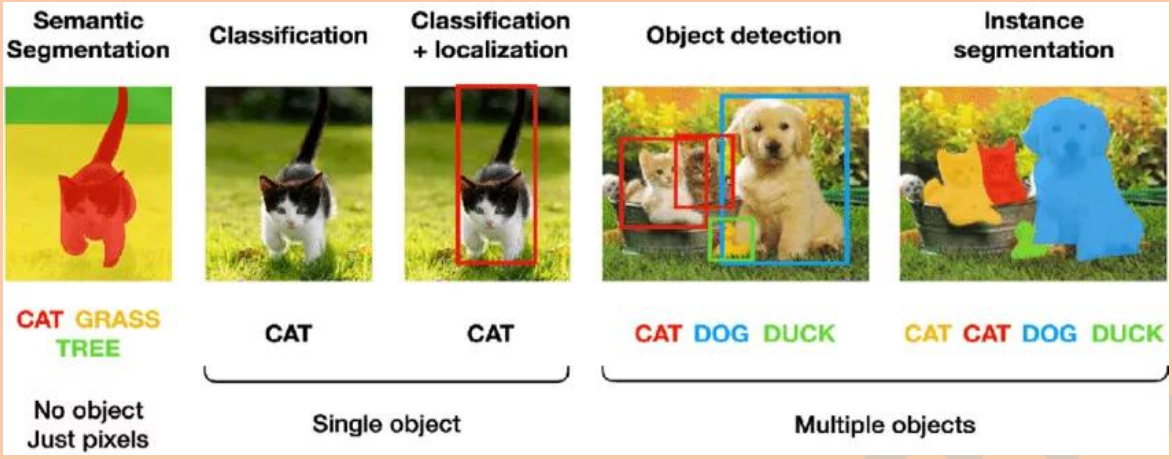
Medical Imaging

For the last decades, computer vision medical imaging application has been a trustworthy help for physicians and doctors. It creates and analyses images and helps doctors with their interpretation.

The application is used to read and convert 2D scan images into interactive 3D models.

Computer Vision Tasks

The Application of the computer is performed by certain tasks on the data or input provided by the user so it can process and analyse the situation and predict the outcome.



Single object	Multiple object
Image Classification: - Image Classification is the task of identifying an object in the input image and label from a predefined category	**Object detection:** - Object detection tasks extract features from the input and use learned formulas to recognize instances of an object category.
Classification + Localization: - As the name suggests, the task identifies the object and locates it in the input image.	**Instance segmentation**: - Instance segmentation assigns a label to each pixel of the image. It is used for tasks such as counting the number of objects

Basics of Images

The word "pixel" means a picture element.

Pixels

- Pixels are the fundamental element of a photograph.
- They are the smallest unit of information that make up a picture.
- They are typically arranged in a 2-dimensional grid.
- In general term, the more pixels you have, the more closely the image resembles the original.

Resolution

- The number of pixels covered in an image is sometimes called the resolution
- Term for area covered by the pixels in conventionally known as resolution.
- For eg :1080 x 720 pixels is a resolution giving numbers of pixels in width and height of that picture.
- A megapixel is a million pixels.

Pixel value

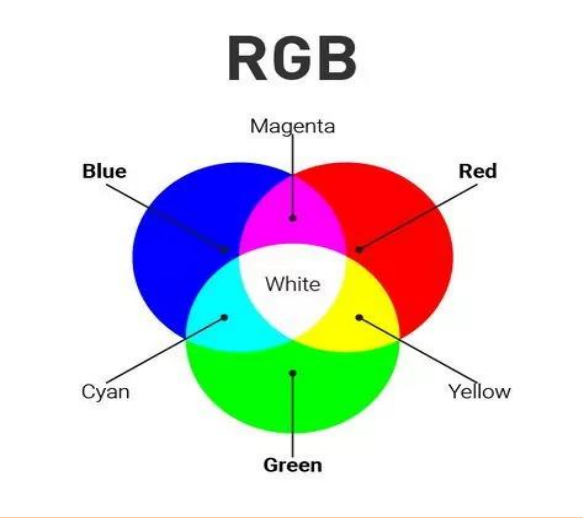
- Pixel value represent the brightness of the pixel.
- The range of a pixel value in 0-255(2^8-1)
- where 0 is taken as Black or no colour and 255 is taken as white

Why do pixel values have numbers?

Computer systems only work in the form of ones and zeros or binary systems. Each bit in a computer system can have either a zero or a one. Each pixel uses 1 byte of an image each bit can have two possible values which tells us that the 8 bits can have 255 possibilities of values that start from 0 and ends at 255.

Grayscale Images

- Grayscale images are images which have a range of shades of gray without apparent colour.
- The lightest shade is white total presence of colour or 255 and darkest colour is black at 0.
- Intermediate shades of gray have equal brightness levels of the three primary colours RBG.
- The computers store the images we see in the form of these numbers.



RBG colours

- All the coloured images are made up of three primary colours Red, Green and Blue.
- All the other colour are formed by using these primary colours at different proportions.
- Computer stores RGB Images in three different channels called the R channel, G channel and the B channel.

Image Features

- A **feature** is a description of an image.
- Features are the specific structures in the image such as points, edges or objects.
- Other examples of features are related to tasks of CV motion in image sequences, or to shapes defined in terms of curves or boundaries between different image regions.

Open CV or Open Source Computer Vision Library is that tool that helps a computer to extract these features from the images. It is capable of processing images and videos to **identify objects, faces, or even handwriting.**

Questions on Computer Vision

One Mark Questions	
Q.No	**Question**
1.	**What is the primary goal of Computer Vision?** A) To enhance image quality B) To interpret and make decisions based on visual data C) To store images efficiently D) To increase resolution of images
2.	**Which of the following is an example of an application of Computer Vision?** A) Word processing B) Speech recognition C) Facial recognition D) Data encryption
3.	**What does a pixel represent in an image?** A) A segment of a video B) The smallest unit of an image C) The color depth of an image D) The brightness of an image

4.	**Which of the following tasks involves identifying and locating objects within an image?** A) Image compression B) Feature extraction C) Object detection D) Image enhancement
5.	**A _____ is a technology based on computer vision that identifies, verifies, or matches a digital image of a human face against a database of stored facial images.**
6.	**What does segmentation in Computer Vision refer to?** A) Enhancing image details B) Dividing an image into parts or regions C) Reducing image size D) Increasing image resolution
7.	**A grayscale image represents intensity values ranging from 0 to _____.**
8.	**What is feature extraction in Computer Vision?** A) Reducing image noise B) Identifying and describing relevant characteristics from an image C) Increasing image contrast D) Storing image data
9.	**What is the main function of the Google Translate App when interpreting foreign language signs?** A) To provide dictionary definitions B) To teach grammar rules C) To translate text into your preferred language almost instantly D) To convert voice to text
10.	**What does the pixel value represent in a grayscale image?** A) The color B) The intensity C) The contrast D) The brightness
11.	**Which of the following is a common use of Computer Vision in medical imaging?** A) Audio transcription B) Image segmentation C) Video streaming D) Data encryption
12.	**Which of the following is a primary color in the RGB color model?** A) Yellow B) Cyan C) Green D) Magenta

13.	**What is a common application of Computer Vision in security systems?** A) Document editing B) Video streaming C) Facial recognition D) Web browsing
14.	**Which format is typically used to store a color image in digital form?** A) Grayscale B) Binary C) RGB D) Indexed
15.	**_____ is the core technology behind the development of autonomous vehicles**
16.	**_____allows you to point your phone's camera at the words and tell you what it means in your preferred language almost instantly**.
17.	**True/False** A higher resolution in an image implies less detail.
	Assertion Reasoning Questions
18.	**Assertion (A):** Computer vision is a field of artificial intelligence that enables computers to interpret and make decisions based on visual data from the world. **Reasoning (R):** Computer vision uses algorithms to process and analyse images and videos, enabling tasks like object detection and facial recognition. A. Both A and R are true, and R is the correct explanation for A. B. Both A and R are true, but R is not the correct explanation for A. C. A is true, but R is false. D. A is false, but R is true. E. Both A and R are false.
19.	**Assertion (A):** Image classification is the process of categorizing and labeling groups of pixels or vectors within an image based on specific rules. **Reasoning (R):** Image classification is a crucial step in medical imaging, allowing for the diagnosis of diseases from X-rays or MRI scans. A. Both A and R are true, and R is the correct explanation for A. B. Both A and R are true, but R is not the correct explanation for A. C. A is true, but R is false. D. A is false, but R is true. E. Both A and R are false.

20	**Assertion (A):** Computer vision can be used in automated quality inspection in manufacturing industries. **Reasoning (R):** Automated quality inspection systems use computer vision to identify defects or irregularities in products on a production line. A. Both A and R are true, and R is the correct explanation for A. B. Both A and R are true, but R is not the correct explanation for A. C. A is true, but R is false. D. A is false, but R is true. E. Both A and R are false.
	Subjective Questions-2 marks
1.	Explain the term "Computer Vision" and its primary goal.
2.	What is the significance of Computer Vision in AI?
3.	Name two applications of Computer Vision in the healthcare industry.
4.	Describe an application of Computer Vision in the automotive industry.
5.	What is the role of feature extraction in Computer Vision?
6.	Explain the concept of image segmentation.
7.	Differentiate between object detection and image classification.
8.	What is the significance of edge detection in Computer Vision tasks?
9.	Define pixel and explain its importance in digital images
10.	What is resolution, and how does it affect image quality?
11.	Explain the difference between grayscale and RGB images
12.	How is pixel value represented in a grayscale image?
13.	Describe the role of color channels in an RGB image.
14.	What are the advantages of using high-resolution images in Computer Vision tasks?
15.	Explain how pixel density affects the visual quality of an image.
	Subjective Questions-4 marks
1.	Explain two different real-world applications of Computer Vision and how they benefit society.
2.	Explain the concepts of pixel value, resolution, and color channels in digital images, and how they collectively affect image quality.
3.	Discuss the evolution of Computer Vision and its impact on modern technology.
4.	Describe the concept of feature extraction and its importance in different Computer Vision tasks.
5.	Compare and contrast object detection, image classification, and image segmentation in Computer Vision.
	Case Study/Application-Based Questions on Computer Vision- 5 marks

1.	**Case Study: Autonomous Vehicles** An autonomous vehicle relies on Computer Vision to navigate and make driving decisions. Explain how Computer Vision helps in tasks such as lane detection, pedestrian recognition, and traffic sign recognition. Additionally, discuss the potential challenges that the vehicle might face in adverse weather conditions and how these can be mitigated
2.	**Case Study: Medical Imaging** In the healthcare industry, Computer Vision is utilized for analysing medical images such as MRI and CT scans. Describe how Computer Vision can assist in detecting abnormalities such as tumours, and discuss the advantages of using this technology in early diagnosis and treatment planning.
3.	**Case Study: Surveillance Systems** Consider a smart surveillance system that employs Computer Vision for security purposes. Explain how object detection and facial recognition are used in this system to enhance security. What ethical considerations should be considered when implementing such a system?
4.	**Case Study: Digital Image Restoration** A company specializes in restoring old, damaged photographs using Computer Vision techniques. Explain how understanding pixel values, resolution, and color channels is crucial in this process. How does Computer Vision enhance the quality of these restored images?
5.	**Case Study: Retail Industry** In the retail industry, Computer Vision is used to improve customer experience and store management. Describe how Computer Vision can be applied for tasks such as inventory management, customer behaviour analysis, and automated checkout. What are the benefits and challenges associated with implementing these technologies in retail stores?

Answers

Q.No	One Mark Answers
1.	B) To interpret and make decisions based on visual data
2.	C) Facial recognition
3.	B) The smallest unit of an image
4.	C) Object detection
5.	facial recognition system
6.	B) Dividing an image into parts or regions

7.	255
8.	B) Identifying and describing relevant characteristics from an image
9.	C)To translate text into your preferred language almost instantly
10.	B) The intensity
11.	B) Image segmentation
12.	C) Green
13.	C) Facial recognition
14.	C) RGB
15.	Computer vision
16.	Google Translate app
17.	FALSE
18.	Both A and R are true, and R is the correct explanation for A.
19.	B) Both A and R are true, but R is not the correct explanation for A
20.	Both A and R are true, and R is the correct explanation for A.
	Subjective Questions-2 marks
1.	Computer Vision is a field of Artificial Intelligence that enables computers to interpret and make decisions based on visual data, with the primary goal of automating tasks that the human visual system can do.
2.	Computer Vision is significant in AI because it allows machines to understand and interpret visual information, leading to automation in areas like image recognition, object detection, and scene understanding, which are crucial for applications such as autonomous driving and medical imaging.
3.	Two applications of Computer Vision in healthcare are medical imaging analysis (e.g., detecting tumors in MRI scans) and surgical assistance systems (e.g., guiding robotic surgery with real-time imaging).
4.	In the automotive industry, Computer Vision is used in autonomous vehicles for tasks such as lane detection, pedestrian recognition, and traffic sign recognition, enhancing safety and enabling self-driving capabilities
5.	Feature extraction involves identifying and describing relevant characteristics from an image, which can then be used for tasks such as object recognition, classification, and tracking.
6.	Image segmentation is the process of dividing an image into multiple regions or segments, each representing a different part of the image. This helps in isolating objects and understanding the structure of the scene.
7.	Object detection identifies and locates objects within an image, providing bounding boxes for each object, whereas image classification assigns a label to the entire image based on the objects it contains without providing their locations.
8.	Edge detection is significant because it helps in identifying the boundaries and structure within an image, which is essential for object detection, segmentation,

	and recognition tasks.
9.	A pixel is the smallest unit of a digital image, representing a single point in the image with a specific color or intensity. Pixels are important because they collectively form the entire image, determining its resolution and detail.
10.	Resolution refers to the number of pixels in an image, usually measured in pixels per inch (PPI). Higher resolution means more pixels and greater detail, resulting in better image quality.
11.	Grayscale images consist of shades of gray, ranging from black to white, with each pixel representing an intensity value. RGB images use three color channels (Red, Green, Blue), where each pixel is a combination of these three colors, allowing for a wide range of colors in the image.
12.	In a grayscale image, the pixel value is represented by an intensity level ranging from 0 to 255, where 0 represents black, 255 represents white, and values in between represent different shades of gray.
13.	In an RGB image, each pixel has three color channels (Red, Green, Blue). The intensity of each channel determines the final color of the pixel. By combining different intensities of these three channels, a wide range of colors can be represented.
14.	High-resolution images provide more detail and clarity, which can improve the accuracy of Computer Vision tasks such as object detection, recognition, and segmentation, as they allow for better feature extraction and analysis.
15.	Pixel density, measured in pixels per inch (PPI), affects the sharpness and clarity of an image. Higher pixel density means more pixels are packed into a given area, resulting in a crisper and more detailed image, which is particularly important for high-quality displays and prints.
	Subjective Questions-4 marks Answers
1.	**Healthcare:** Computer Vision is used in medical imaging to detect anomalies such as tumours, improving early diagnosis and treatment outcomes. It also assists in robotic surgeries, providing precise guidance and enhancing surgical accuracy. **Security:** In security systems, Computer Vision is used for facial recognition to identify individuals in surveillance footage, aiding in crime prevention and investigation. It enhances public safety by monitoring public spaces and alerting authorities to suspicious activities.

2.	**Pixel Value:** In digital images, pixel value represents the intensity or color information of a pixel. In grayscale images, it ranges from 0 (black) to 255 (white). In RGB images, it is defined by the intensities of red, green, and blue channels. Resolution: Resolution refers to the number of pixels in an image, typically measured in pixels per inch (PPI). Higher resolution means more pixels and greater detail, enhancing image clarity and quality. **Color Channels:** In RGB images, each pixel is composed of three-color channels (red, green, blue). The combination of these channels at varying intensities produces a wide range of colors. High-quality images require accurate representation of these color channels. **Collective Impact:** High pixel values, resolution, and well-defined color channels contribute to a detailed, sharp, and color-rich image. Lower values or resolution can result in blurred, pixelated, or distorted images, reducing visual quality and effectiveness in Computer Vision tasks.
3.	Computer Vision has evolved from basic image processing techniques to advanced AI-driven algorithms that can understand and interpret complex visual data. This evolution has had a significant impact on modern technology, enabling advancements in areas such as autonomous driving, facial recognition, medical diagnostics, and augmented reality. As a result, many industries have seen improved efficiency, safety, and innovation through the integration of Computer Vision technologies.
4.	Feature extraction involves identifying and isolating significant information from an image, such as edges, textures, and shapes. This process is crucial for various Computer Vision tasks: **Object Recognition:** Features help in identifying objects within an image by matching extracted features with known patterns. **Image Classification:** Features are used to classify images into categories based on their content. **Tracking:** Extracted features allow for tracking objects across frames in video analysis, crucial for surveillance and motion detection. **Augmented Reality:** Features are used to overlay virtual objects accurately onto real-world scenes

5.	**Object Detection:** This task involves identifying and locating objects within an image, providing bounding boxes around detected objects. It focuses on detecting multiple objects and their positions. **Image Classification:** This task assigns a single label to an entire image based on its content. It does not provide the locations of objects, only categorizes the image as a whole. **Image Segmentation:** This task divides an image into segments, each representing a different object or region. It provides pixel-level classification, offering detailed information about the structure and boundaries within the image.
	Case Study/Application-Based Questions on Computer Vision- 5 marks- Answers
1.	Computer Vision helps in lane detection by using cameras to identify lane markings on the road, ensuring the vehicle stays within its lane. Pedestrian recognition involves detecting and tracking pedestrians to avoid collisions. Traffic sign recognition uses image processing to identify and interpret traffic signs, allowing the vehicle to respond accordingly. Challenges in adverse weather conditions include reduced visibility and accuracy. These can be mitigated by using additional sensors such as radar and LIDAR, as well as implementing advanced algorithms to enhance image processing in poor visibility.
2.	Computer Vision algorithms can analyze medical images to detect abnormalities like tumors by identifying unusual patterns and shapes that indicate the presence of disease. The advantages include faster and more accurate diagnosis, early detection of diseases, and improved treatment planning. This technology reduces the workload on medical professionals and increases the chances of successful treatment by identifying issues at an early stage.
3.	Object detection is used to identify and monitor objects within the surveillance area, alerting security personnel to any suspicious activity. Facial recognition identifies individuals by comparing captured images with a database of known faces, enhancing security by recognizing potential threats. Ethical considerations include privacy concerns, potential biases in recognition algorithms, and the need for transparency and accountability in how the data is used and stored.
4.	Understanding pixel values helps in identifying the intensity and color information of each pixel, which is essential for correcting damaged areas. Resolution knowledge is important for maintaining image detail during restoration. Color channels are used to accurately restore the colors in RGB images. Computer Vision enhances quality by using algorithms to fill in missing parts, correct color imbalances, and sharpen details, resulting in a restored image that closely resembles the original.

5.	Computer Vision can track inventory levels in real-time by analyzing shelf images, ensuring timely restocking. It can analyze customer behavior by monitoring movement patterns and product interactions, helping in optimizing store layout and marketing strategies. Automated checkout systems use image recognition to identify products and streamline the payment process. Benefits include increased efficiency, reduced labor costs, and improved customer satisfaction. Challenges include the high cost of implementation, potential technical issues, and ensuring data privacy and security.