Computer Vision

Computer Vision

1. Introduction

- **Computer Vision:** a field of AI that enables machines to interpret, analyse and derive meaning from digital images and video.
- It aims to replicate aspects of human visual perception so computers can recognise objects, scenes and activities.
- Combines image processing, pattern recognition, and machine learning.

2. Core Tasks

- **Image classification:** assigning a label to an entire image.
- **Object detection:** locating and labelling multiple objects with bounding boxes.
- **Segmentation:** partitioning an image into regions or pixel-level masks.
- **Tracking:** following objects across video frames over time.

3. Key Techniques

- **Convolutional Neural Networks (CNNs):** deep models that learn hierarchical spatial features from pixels.
- Traditional methods include edge detection, feature descriptors (SIFT, HOG) and thresholding.
- Data augmentation and transfer learning improve performance on limited datasets.

4. Applications

- Facial recognition, medical image diagnosis, and quality inspection in manufacturing.
- Autonomous vehicles use vision for lane and pedestrian detection.
- Augmented reality, document OCR, and agricultural crop monitoring.

5. Challenges

- Sensitivity to lighting, occlusion, viewpoint and background clutter.
- Need for large labelled datasets and high computational resources.
- Bias in training data and privacy concerns in surveillance use.