



Computer Vision



What is Computer Vision?

Computer vision is the science and technology of machines that see.

Concerned with the theory for building artificial systems that obtain information from images.

The image data can take many forms, such as a video sequence, depth images, views from multiple cameras, or multi-dimensional data from a medical scanner

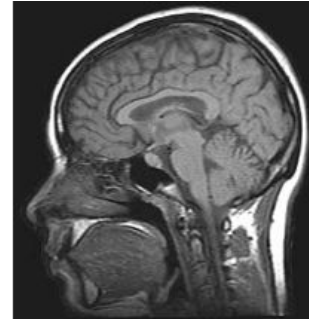


Image Processing Human vs Computer

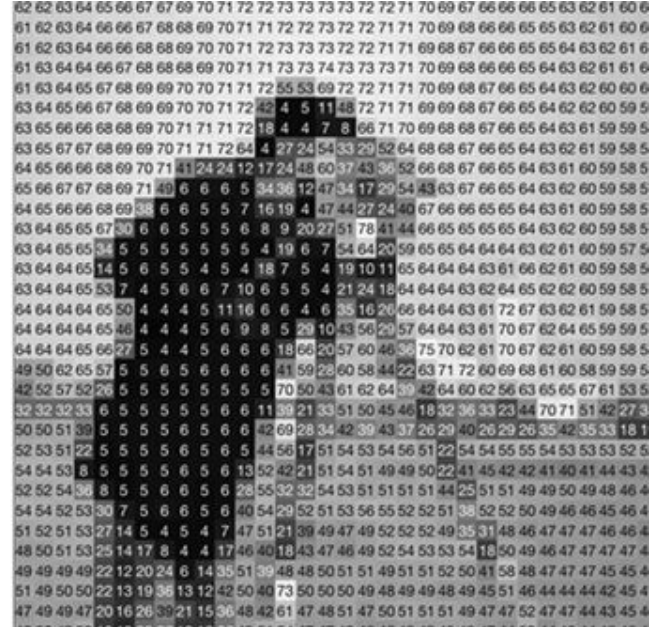




History of computer vision

- 1963 - Computers were able to interpret the tridimensionality of a scene from a picture.
- 1974 - Optical character recognition (OCR) was introduced to help interpret texts printed in any typeface.
- 1980 - Dr. Kunihiro Fukushima proposed Neocognitron, a hierarchical multilayered neural network.
- 2000 - 2001 - Studies on object recognition increased, helping in the development of the first real-time face recognition application.
- 2010 - ImageNet data were made available containing millions of tagged images across various object classes.
- 2014 - COCO has also been developed to offer a dataset.

What Human vs Computer see?

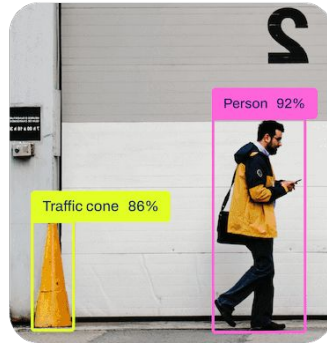


Major Classification of Computer Vision

Classify



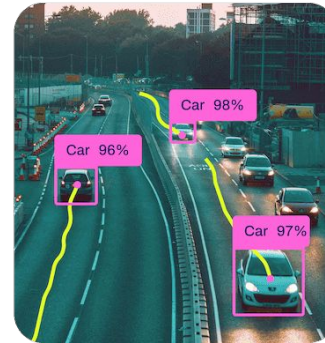
Detect



Segment



Track



Pose

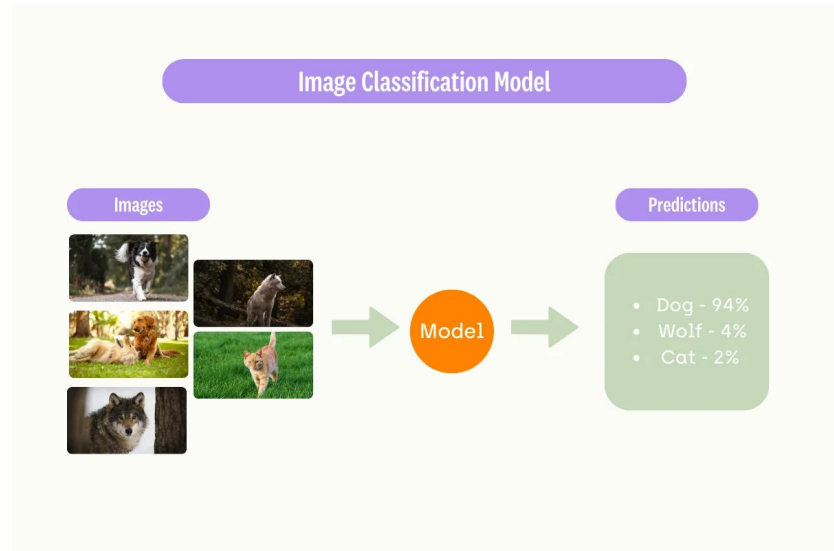


Classification

Classification is a technique that involves identifying and categorizing objects within an image or video into predefined classes based on their features and characteristics using machine learning algorithms.

Use Case:

- Retail and Inventory Management
- Healthcare and Medical Imaging



Detection

Object detection is a technique that involves identifying and locating objects within an image or video, providing both the classification of objects and their spatial coordinates using bounding boxes.

Use Case:

- Surveillance and Security
- Augmented Reality (AR) Applications
- Industrial Automation



Segmentation

Computer segmentation involves partitioning a digital image into multiple segments to simplify and change its representation for easier analysis.

Use Case:

- **Autonomous Vehicles**
- **Medical Imaging**



Tracking

Object tracking is a technique that involves monitoring the movement and position of an object over time across a sequence of images or video frames.

Use Case:

- Sports Analytics
- Traffic Monitoring



Pose

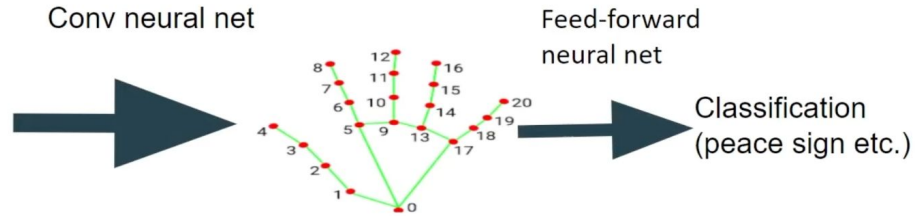
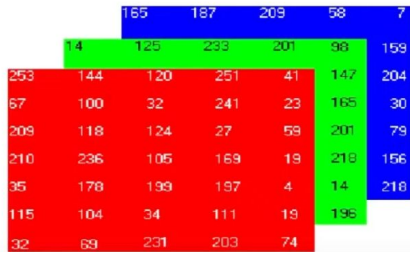
Pose estimation is a technique that involves detecting and predicting the spatial positions of key points on a human body or object, enabling the analysis of posture, movement, and orientation.

Use Case:

- Fitness and Rehabilitation
- Human-Computer Interaction (HCI)



How Computer understands?





Gesture recognition

- Gesture recognition is a technology that allows computers to understand human body language.
- It is a way for computers to understand human body language that allows them to interpret human gestures via mathematical algorithms.
- It allows humans to interface with the machine (HM) and interact naturally without any mechanical devices.

Example

A smart TV recognizing the hand movement to control the channels or to select a option from the menu.



Code Overview
