

# Computer Vision

## Introduction

Course Instructor:  
Dr. Suman Kumar Maji

# What is Computer Vision?



- Computer vision is the science of giving computers the ability to see and understand images and videos.
- It focuses on replicating parts of the human visual system by acquiring, processing, analyzing and understanding digital images.

# Historical Background

A brief history of Computer Vision's development through time -

- **Early Beginnings (1960s):** The early stages of computer vision started in the 1960s when researchers began finding ways to let machines understand pictures. Algorithm ideas remained fairly based on geometric shapes and pattern recognition initially.
- **Feature Detection and Recognition (1970s):** Work continued in the area of edge detection and feature extraction techniques. Researchers designed algorithms to identify and recognize simple shapes, laying the groundwork for more advanced analyses.

## Continued...

**Machine Learning Era (1980s-1990s):** This era marked a shift in emphasis towards machine learning. For the first time, statistical models were adopted to enhance object recognition. More robust algorithms were developed to handle real variations in images.

**Neural networks:** Neural networks gained popularity, particularly with the rise of deep learning in the last decade. The depth of Convolutional Neural Networks (CNNs) became one of the powerful tools for image classification and object detection, enhancing precision and efficiency.

**Current Trends (2010s-Present):** Today, computer vision is a leading area in AI research and applications, impacting various industries. Current trends are driven by hardware development (e.g., GPUs) and the availability of large datasets for model training. Key applications include real-time video analysis, facial recognition, and autonomous systems.

# Applications of Computer Vision

Computer vision has a wide range of applications in many different fields, some of which include -

- **Healthcare:** It is used for the diagnosis of diseases through the study of medical images.



**malignant**



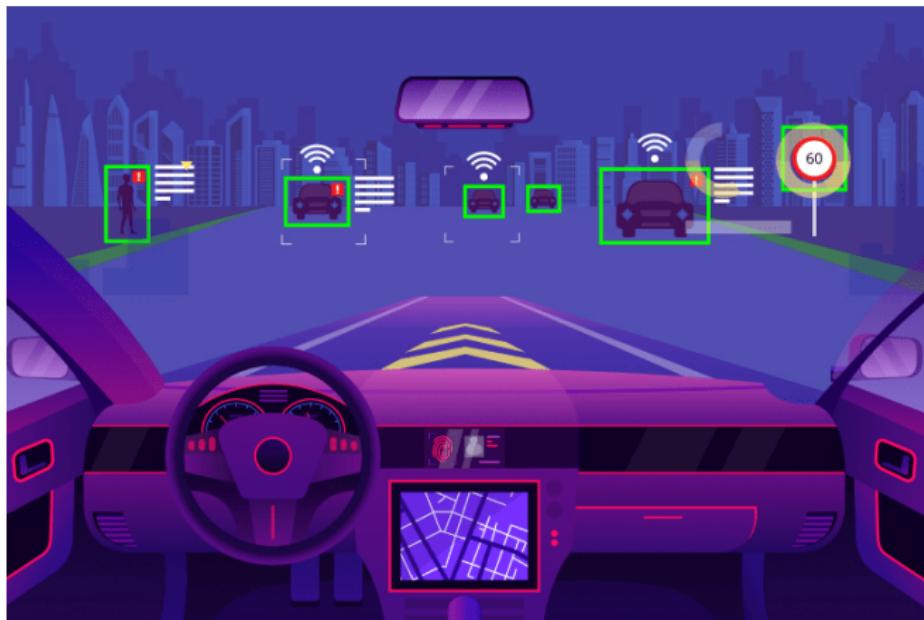
**benign**



Example of malignant/benign classification with Computer Vision

continued...

**Autonomous Vehicles:** Helping self-driving cars to recognize objects, pedestrians, and road signs.



Continued...

**Security and Surveillance:** Monitoring video feeds to detect suspicious activities.



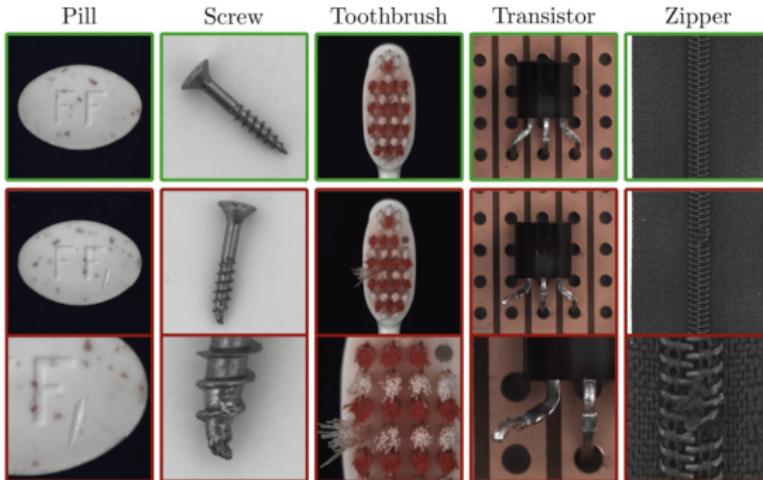
# Continued...

**Retail:** Visual search and automated checkout systems to improve customer experience.



# Continued...

**Manufacturing:** Product inspection in assembly lines to ensure quality control.



Examples of manufactured products without (first two rows) and with (last row) defects. Image originally appears in [MVTec AD – A Comprehensive Real-World Dataset for Unsupervised Anomaly Detection](#), and is courtesy of MVTech.

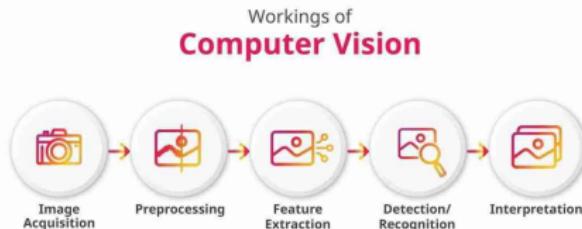
# Human Vision v/s Computer vision



62 62 63 64 65 66 67 67 69 70 71 72 72 73 73 73 73 72 72 71 70 69 67 66 66 66 65 63 62 61 60 6  
61 62 63 64 66 66 67 68 68 69 70 71 72 72 73 73 73 72 72 71 71 70 69 68 66 66 65 65 63 62 61 60 6  
61 62 63 64 66 66 68 68 69 70 70 71 72 73 73 73 72 72 71 71 69 68 67 66 66 65 65 64 63 62 61 6  
61 63 64 64 66 67 68 68 69 70 71 71 73 73 74 73 73 73 71 70 69 68 66 66 65 64 63 62 61 61 6  
61 63 64 65 67 68 69 69 70 71 71 72 55 53 69 72 71 71 70 69 68 67 66 65 64 63 62 60 6  
63 64 65 66 67 68 69 69 70 71 72 42 4 5 11 48 72 71 71 69 69 68 67 66 65 64 62 62 60 59 5  
63 65 66 66 68 68 69 70 71 71 72 18 4 4 7 8 66 71 70 69 68 68 67 66 65 64 63 61 59 59 5  
63 65 67 67 68 68 69 70 71 71 72 64 4 27 24 54 33 20 52 64 68 68 67 66 65 64 63 62 61 59 58 5  
64 65 66 66 68 69 70 71 21 24 24 12 17 24 48 60 17 43 30 52 66 68 67 66 65 64 63 61 60 59 58 5  
65 66 67 67 68 68 69 71 49 6 6 6 5 34 35 12 47 34 17 29 54 43 63 67 66 65 64 63 62 60 59 58 5  
64 65 66 66 68 69 50 6 6 5 5 7 16 19 4 47 44 27 24 40 67 66 66 65 65 64 63 61 60 59 58 5  
63 64 65 65 67 50 6 6 5 5 6 8 9 20 27 51 78 41 44 66 65 65 64 63 62 60 59 58 5  
63 64 65 65 5 5 5 5 5 5 4 19 6 7 54 64 20 59 65 65 64 64 63 62 61 60 59 57 5  
63 64 64 65 14 5 6 5 5 4 5 4 18 7 5 4 19 10 11 05 64 64 64 63 61 66 62 61 60 59 58 5  
63 64 64 65 53 7 4 5 6 6 7 10 6 5 5 4 21 24 18 64 64 64 63 62 64 65 62 62 60 59 58 5  
64 64 64 65 50 4 4 4 5 11 16 6 6 4 6 35 16 20 66 64 64 63 61 72 67 63 62 61 59 58 5  
64 64 64 65 48 6 4 4 5 6 9 8 5 29 10 43 50 29 57 64 64 63 61 70 67 62 64 65 59 59 5  
64 64 64 65 66 27 5 4 4 5 6 8 6 18 60 20 57 60 46 46 75 70 62 61 70 67 62 61 60 59 58 5  
49 50 62 66 57 5 5 6 6 6 41 59 22 60 58 44 22 63 71 72 60 69 68 61 60 58 59 59 5  
42 57 52 26 5 5 5 5 5 5 5 5 7 50 43 61 62 64 39 42 64 60 62 56 63 65 67 61 53 5  
22 32 32 33 6 5 5 5 5 5 6 6 11 39 21 33 51 50 45 46 18 32 36 33 23 44 70 71 51 42 27 3  
50 50 51 39 5 5 5 5 8 5 6 6 42 69 28 34 42 39 43 37 26 29 40 26 29 26 35 42 35 33 18 1  
52 53 51 22 5 5 5 5 6 6 5 44 56 17 51 54 53 54 56 51 22 54 55 55 54 53 53 52 5  
54 54 53 8 5 5 5 5 6 6 13 52 42 21 51 54 51 49 49 50 22 41 45 42 42 41 40 41 44 43 4  
52 52 54 54 8 5 5 6 6 6 26 55 32 32 54 53 51 51 51 51 44 25 51 51 49 49 49 50 49 48 46 4  
54 54 52 53 30 7 5 6 6 5 6 40 54 29 52 51 53 56 55 52 52 51 35 52 52 50 49 46 46 45 46 4  
51 52 51 53 27 14 5 4 5 4 7 47 51 21 39 49 47 49 52 52 52 49 35 31 48 46 47 47 47 46 46 4  
48 50 51 53 25 14 17 8 4 4 17 46 40 18 43 47 46 49 52 54 53 53 54 18 50 49 46 47 47 47 47 4  
49 49 49 22 12 20 24 6 14 35 51 39 48 48 50 51 51 49 51 51 52 50 41 58 48 47 47 45 45 4  
51 49 50 50 22 13 19 38 13 12 42 50 40 73 50 50 50 49 48 49 49 48 49 45 51 46 44 44 44 42 45 4  
47 49 49 47 20 16 26 39 21 15 30 48 42 61 47 48 51 47 50 51 51 51 49 47 47 52 47 47 44 43 45 4

what we see v/s what computers see

# How Does Computer Vision Work?



**Image Acquisition:** Capturing images or videos using cameras or sensors.

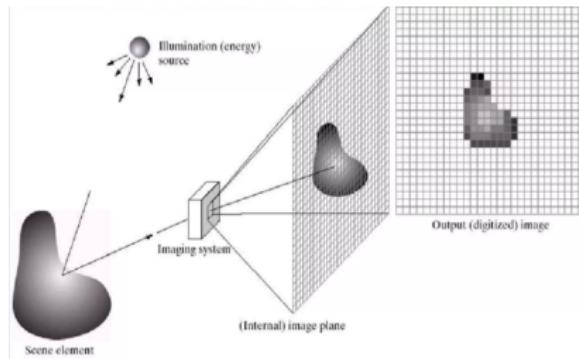
**Preprocessing:** Enhancing the quality of images, such as adjusting brightness or removing noise.

**Feature Extraction:** It means to find such important part of the image which is helpful in representation like edge, texture or shape.

**Object Detection and Recognition:** To find and classify the object/part of objects in image.

**Understanding and Interpretation:** Making decisions or predictions based on the recognized objects.

# Image Acquisition



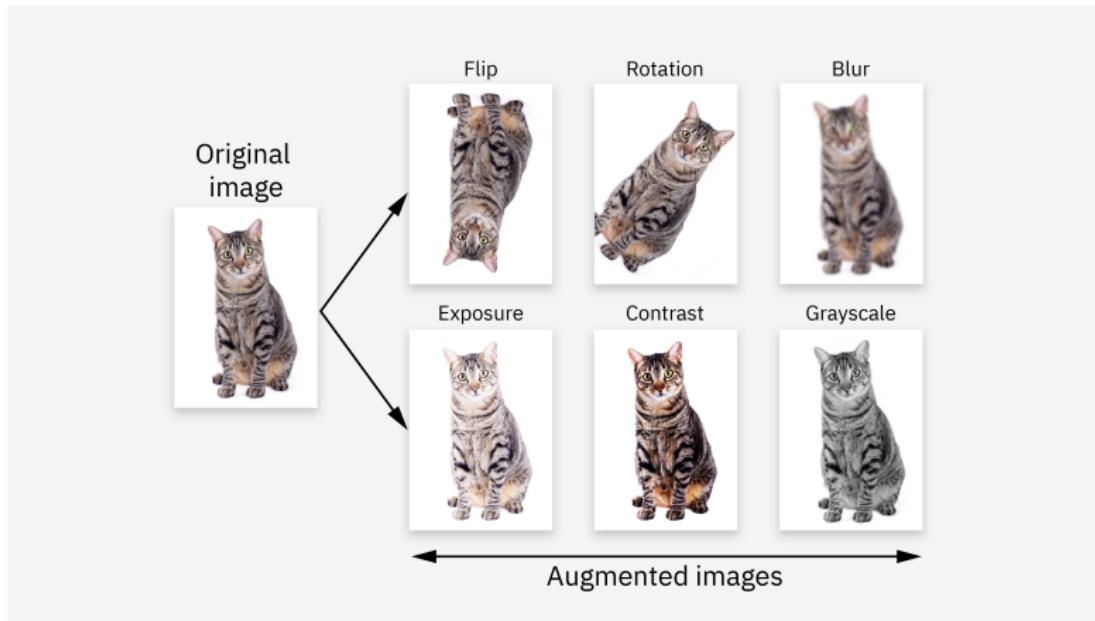
- In computer vision, the initial stage is image acquisition.
- Image acquisition simply involves capturing images using cameras, smartphones, or other specialized sensors.
- The quality of images captured in this phase has great influence on the following parts of computer vision.

# Preprocessing

Preprocessing refers to the transformations applied to our data before feeding it to the algorithm.

- In Python, we generally perform preprocessing using the scikit-learn library. The library provides the class `sklearn.preprocessing` for preprocessing data.
- ① **Noise Reduction:** Elimination of undesirable random fluctuations in the image.
- ② **Contrast Enhancement:** Adjusting the brightness and contrast to highlight important features.
- ③ **Image Resizing:** Changing the image size to match the requirements of the analysis algorithms.

continued...



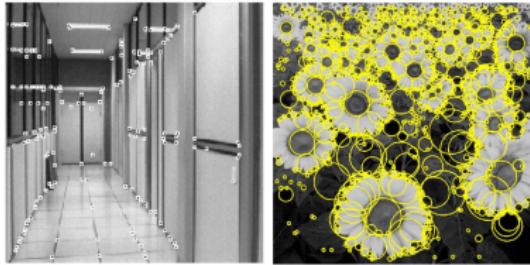
## Continued...

- Many modern object detection algorithms depend upon resizing the images to a predetermined size as part of their pre-processing step.
- We will start with extracting two features from an image: the lines that are present in the image and the color pixels.
- To do this, we need to understand what a line is and what a pixel is.
- The first thing you need to understand here is that whenever we talk about identifying or extracting something from an image, it basically means that we are trying to derive some meaningful information (features) about that thing (like height, weight, coordinates).

# Feature Extraction

Feature extraction is all about finding those parts of an image which will help us solve the problem at hand. Some common features are

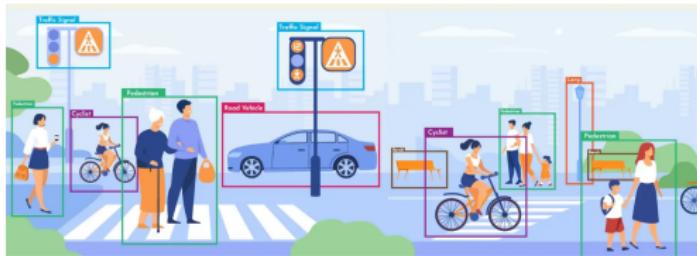
## Feature extraction: Corners and blobs



- **Edges:** Lines where there is a sharp change in color or intensity.
- **Corners:** Locations in the image where two edges meet.
- **Textures:** Repeating patterns of local intensity variations.

# Object Detection and Recognition

Object detection involves finding objects within an image and drawing bounding boxes around them. Object recognition goes a step further by identifying what those objects are.

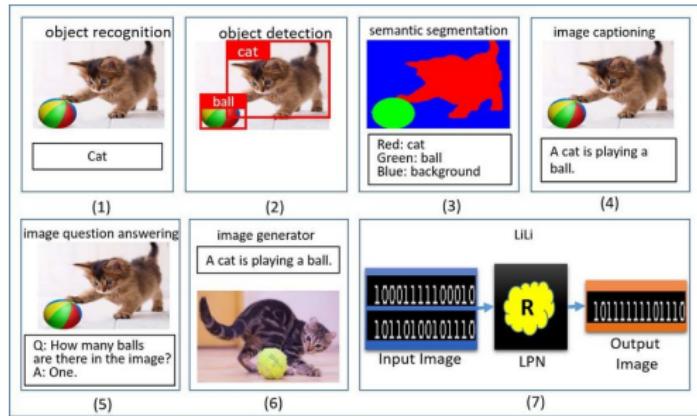


Some object detection techniques used are as shown below-

- **Template Matching:** The parts of the image are compared with some predefined templates.
- **Machine Learning:** Some algorithms learn on a big set of labeled images.
- **Deep Learning:** Leveraging neural networks for automatically learning features and classifying objects.

# Understanding and Interpretation

The final step is to interpret the recognized objects and make decisions.



This could involve -

- **Counting Objects:** Determining the number of certain types of objects in an image.
- **Tracking Movement:** Following the movement of objects across a series of frames in a video.
- **Scene Understanding:** Analyzing the overall context of the image to understand what is happening.

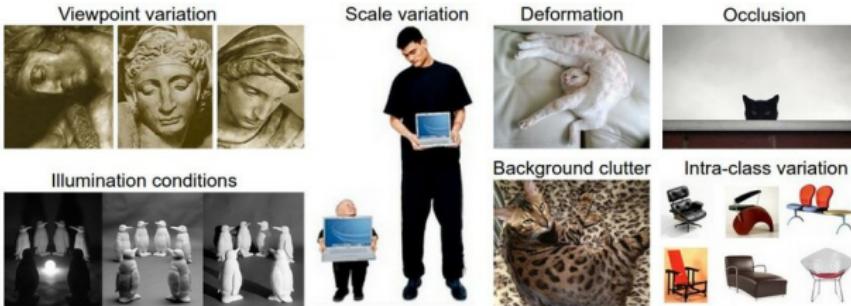
# Techniques in Computer Vision

There are various techniques that are used in computer vision to analyze and interpret visual data

- **Machine Learning:** Training algorithms on large datasets to recognize patterns and make predictions.
- **Deep Learning:** Using neural networks, especially Convolutional Neural Networks (CNNs), to process and analyze visual data.
- **Image Filtering:** Enhancing images by applying filters to remove noise and highlight important features.
- **Edge Detection:** Identifying the boundaries of objects within an image.

# Challenges in Computer Vision

Computer Vision faces several challenges that can hinder accurate image analysis and interpretation



- **Variability in Images:** Images can vary in quality, lighting, angle, and background, making it difficult to analyze.
- **Occlusion:** Objects in an image may be partially hidden by other objects, complicating detection and recognition.
- **Real-time Processing:** Analyzing images and making decisions in real-time requires significant computational resources.