

Chapter 1: Introduction to Computer Vision

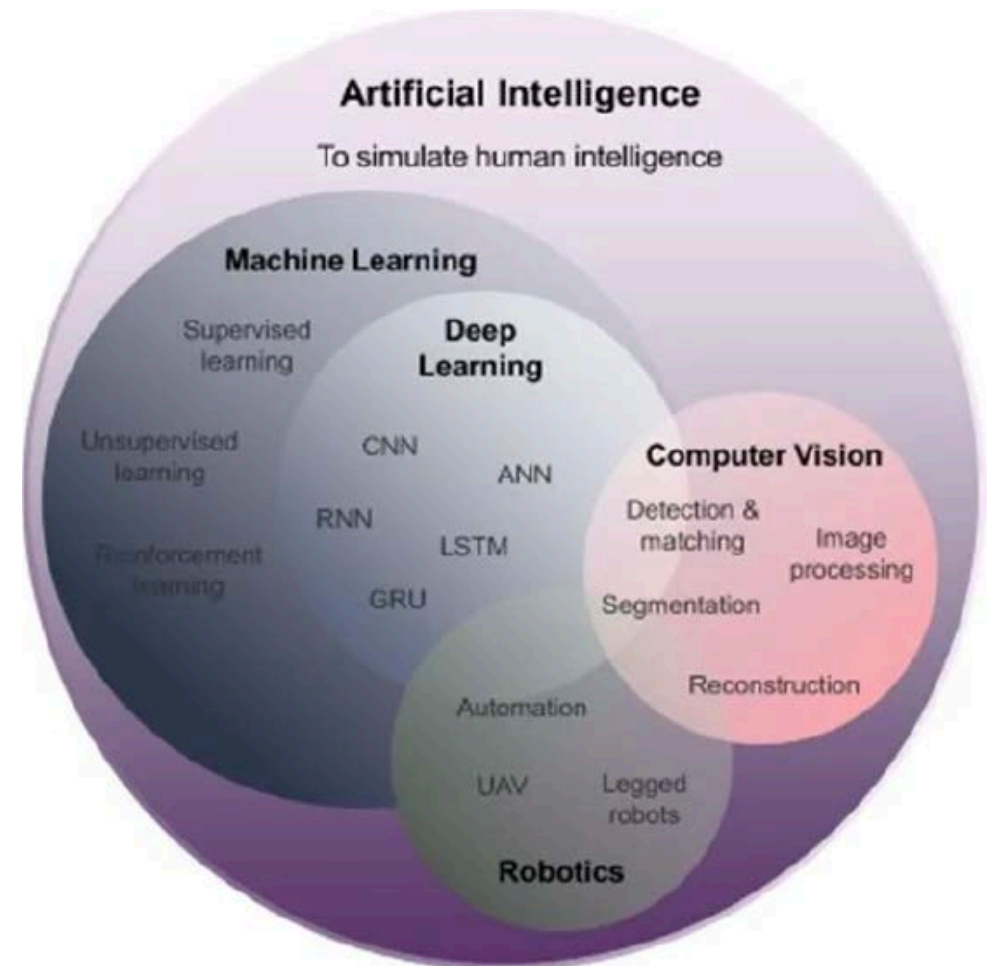
Faculty of Digital Engineering
Institute of Digital Technology
2026

Agenda

1. Introduction to Computer Vision
2. Why use Computer Vision
3. How Computer Vision Works
4. Application of Computer Vision
5. Challenges in Computer Vision
6. Current Research in Computer Vision

What is Computer Vision

- A field of Computer Science that focuses on analyzing and processing visual images or videos intelligently like humans.
- It is part of AI that focuses on developing algorithms and models that enable computers to interpret and understand visual information from the world.
- It is providing computer ability to **see** and **understand** image at pixel level through training and validation.

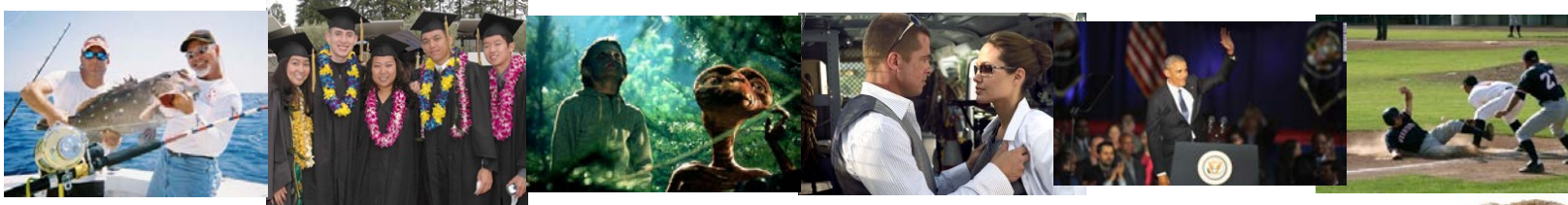


A brief story about Computer Vision

- 1940s – first programmable machine.
- 1950s – Turing test, AI birth.
- 1960s – birth of Computer Vision, the summer vision project at MIT.
- 1970s and 80s – part of AI – understanding human vision and emulating human perception.
- 1990s – depart from AI, geometric approach.
- Today – various mathematical methods (statistics, differential equations, optimization), applications (security, robotics, graphics).

Why use Computer Vision?

- Billions of images/videos captured per day
- Huge number of potential applications



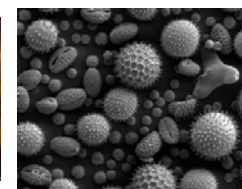
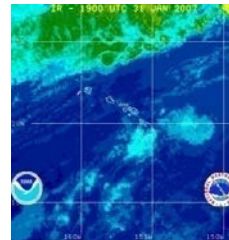
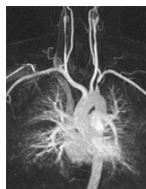
flickr



Google Photos

YouTube

TikTok



Every image tells a story



- Goal of computer vision: perceive the “story” behind the picture
- Compute properties of the world
 - 3D shape
 - Names of people or objects
 - What happened?

Can computers match human perception?



0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

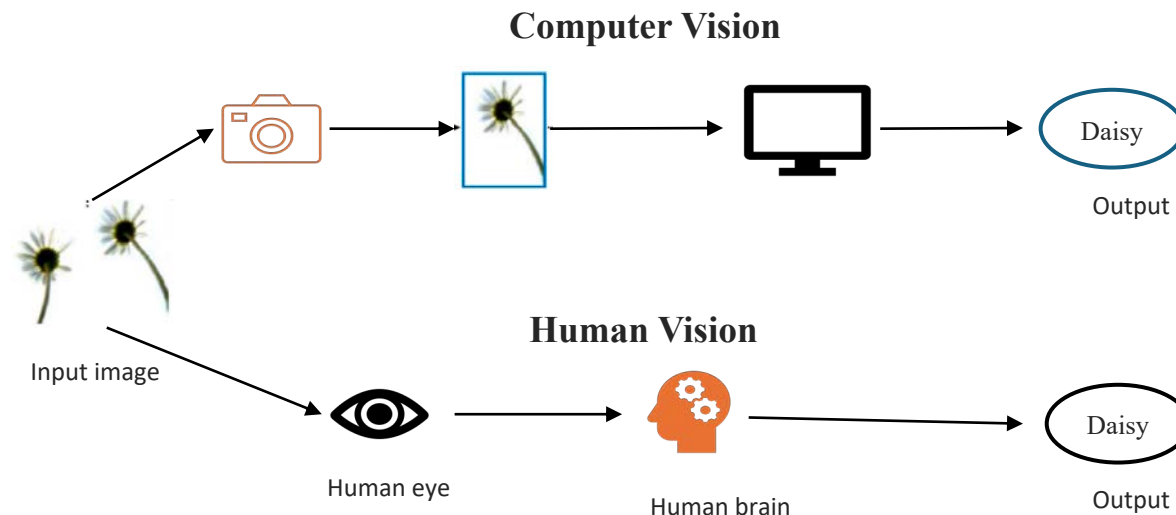
Can computers match human perception?



- Yes and no (mainly no)
 - computers can be better at “easy” things
 - humans are better at “hard” things
- But huge progress
 - accelerating in the last five years due to deep learning
 - what is considered “hard” keeps changing

How Computer Vision works?

- Computer vision emulates how the human brain recognizes objects visually.
- The brain decodes objects by recognizing patterns, a key concept in computer vision.
- Computer vision algorithms process large datasets to identify and label patterns.
 - For example, analyzing millions of flower images helps create a 'flower' model.
 - This enables computers to reliably identify flowers in images.

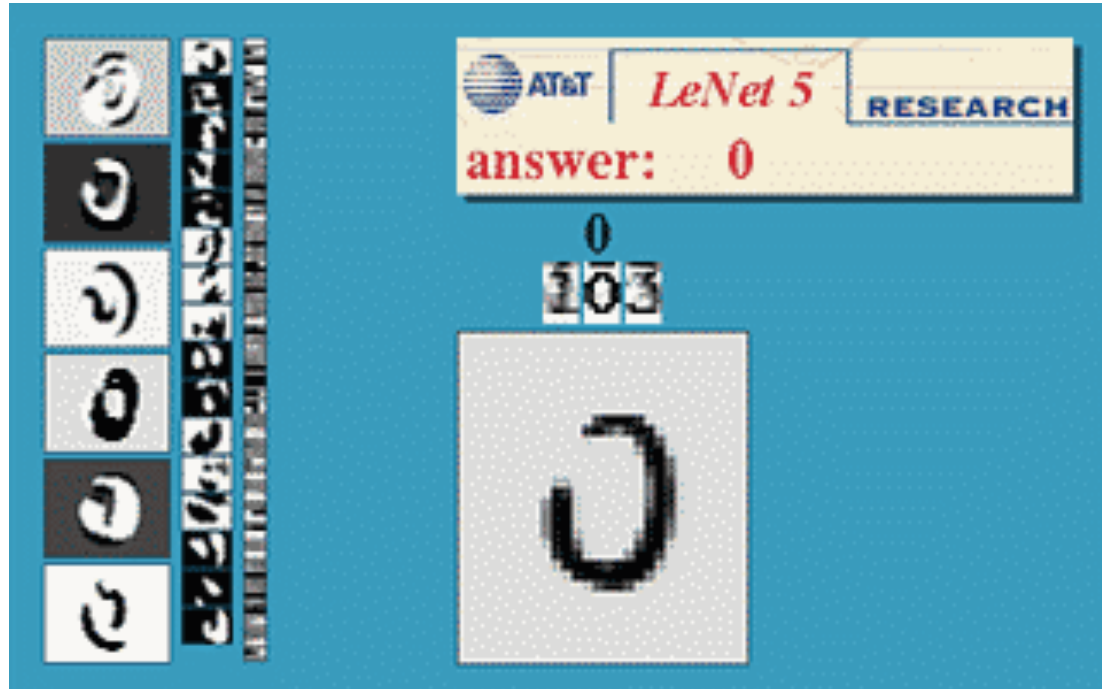


Applications of Computer Vision

- Computer vision has applications in all industries and sectors.
 - Automotive
 - Food and Beverage
 - Manufacturing
 - Pharmaceutical
 - Supply Chain
 - Energy and Utilities
 - Hospitality
 - Life Science
 - Human Resource
 - Insurance.



Optical Character Recognition (OCR)



Digit recognition, AT&T labs (1990's)
<http://yann.lecun.com/exdb/lenet/>



Automatic check processing

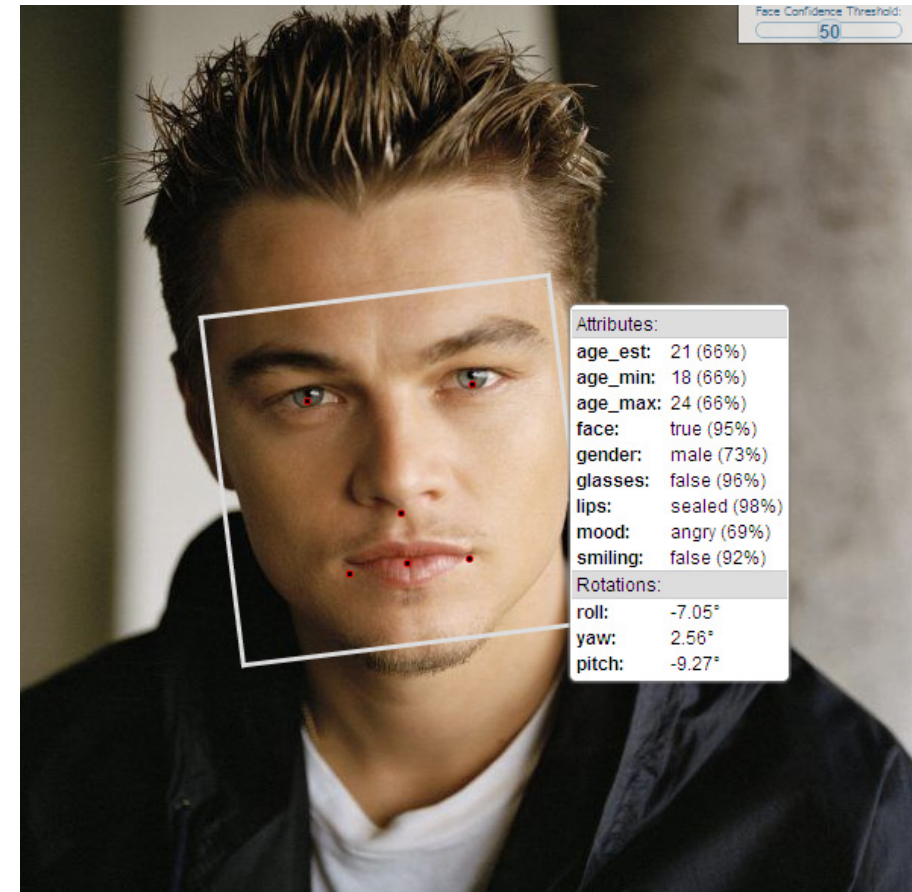


License plate readers
[http://en.wikipedia.org/wiki/Automatic number plate recognition](http://en.wikipedia.org/wiki/Automatic_number_plate_recognition)

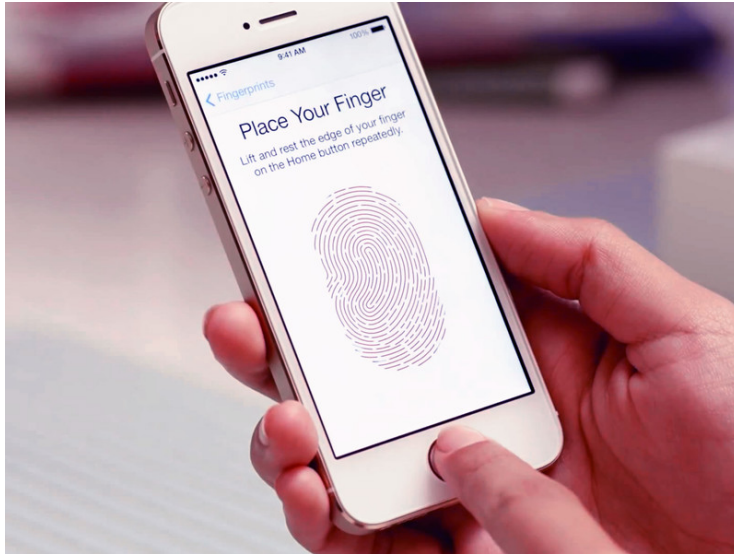


Sudoku grabber
<http://sudokugrab.blogspot.com/>

Face detection and recognition



Authentication without a password



Fingerprint scanners on many new smartphones and other devices



Face unlock on Apple iPhone X
See also <http://www.sensiblevision.com/>

Smart cars – auto driving cars



The screenshot displays the Mobileye website with a top navigation bar for 'manufacturer products' and 'consumer products'. The main banner features a car with three camera fields of view: 'rear looking camera', 'forward looking camera', and 'side looking camera', under the heading 'Our Vision. Your Safety.'.

Below the banner are three product highlights:

- EyeQ Vision on a Chip**: Accompanied by an image of the EyeQ chip and a '> read more' link.
- Vision Applications**: Describes 'Road, Vehicle, Pedestrian Protection and more' with an image of a pedestrian and a '> read more' link.
- AWS Advance Warning System**: Shows a car icon with a '0.8' warning on a screen and a '> read more' link.

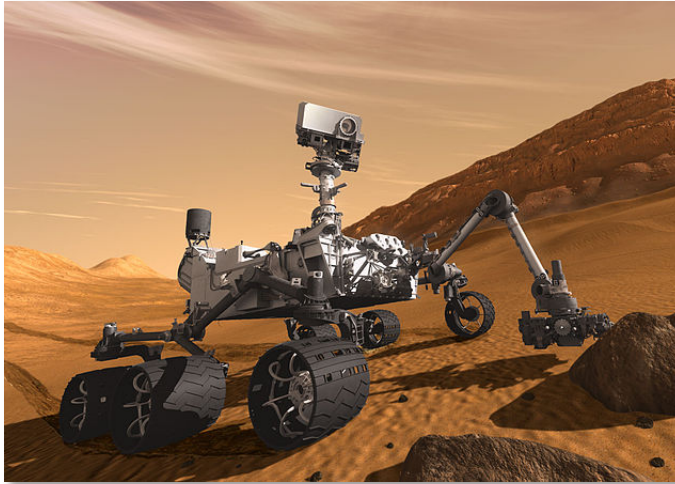
On the right side, there are two sections:

- News**: Contains two headlines: 'Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System' and 'Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end', with an '> all news' link.
- Events**: Lists 'Mobileye at Equip Auto, Paris, France' and 'Mobileye at SEMA, Las Vegas, NV', with a '> read more' link.

<https://www.mobileye.com/>

Safety features in many cars

Robotics



NASA's Mars Curiosity Rover

[https://en.wikipedia.org/wiki/Curiosity_\(rover\)](https://en.wikipedia.org/wiki/Curiosity_(rover))



Amazon Picking Challenge

<http://www.robocup2016.org/en/events/amazon-picking-challenge/>

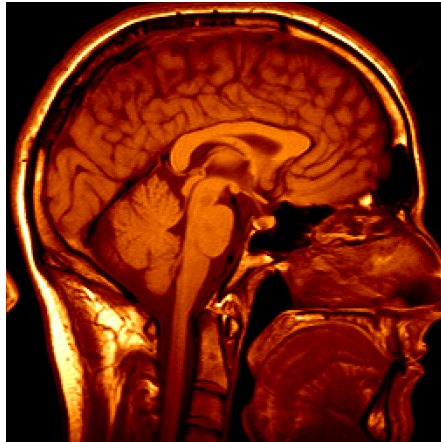


Amazon Prime Air

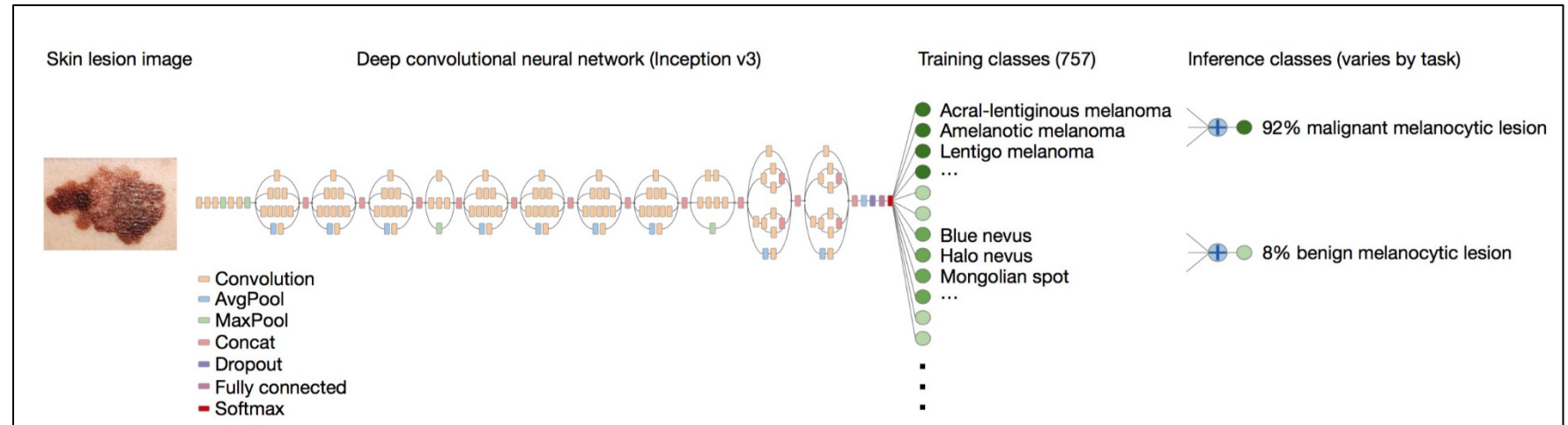


Amazon Scout

Medical imaging



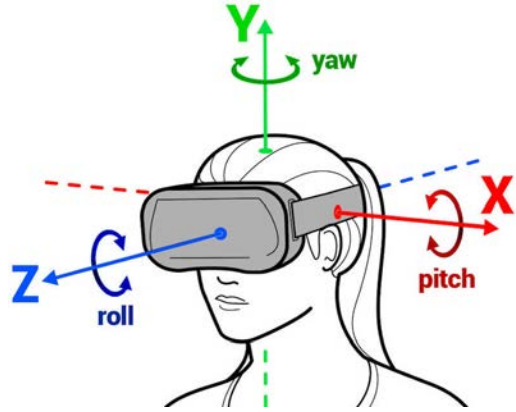
3D imaging
(MRI, CT)



Skin cancer classification with deep learning

<https://cs.stanford.edu/people/esteva/nature/>

Virtual & Augmented Reality



6DoF head tracking



Hand & body tracking



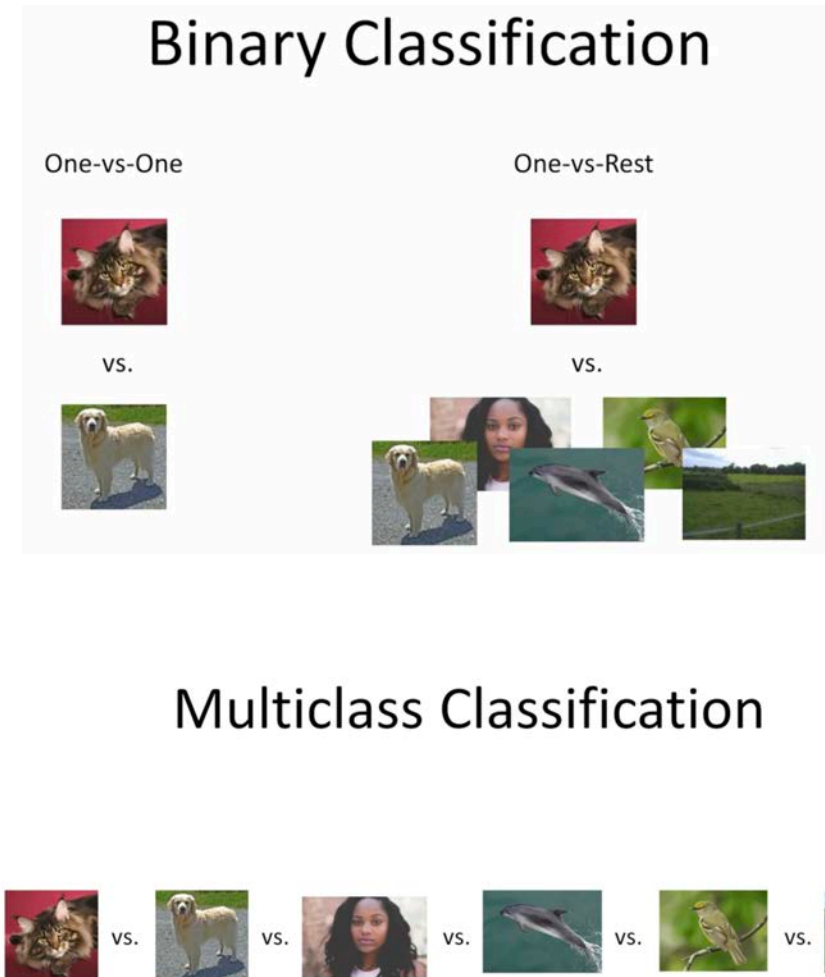
3D scene understanding



3D-360 video capture

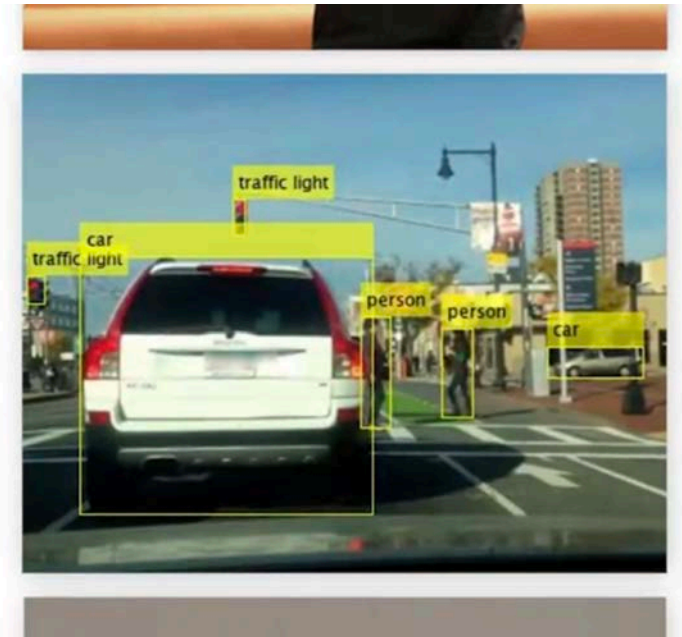
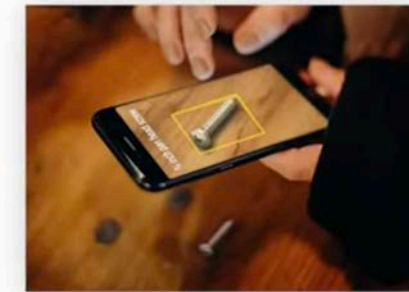
Computer Vision in object classification

- A process of predicting the probability of the presence of a specific object class in the input images.
- **Binary Classification:** Distinguishes between two ground truth labels
- **Multiclass Classification:** Differentiates among more than two labels



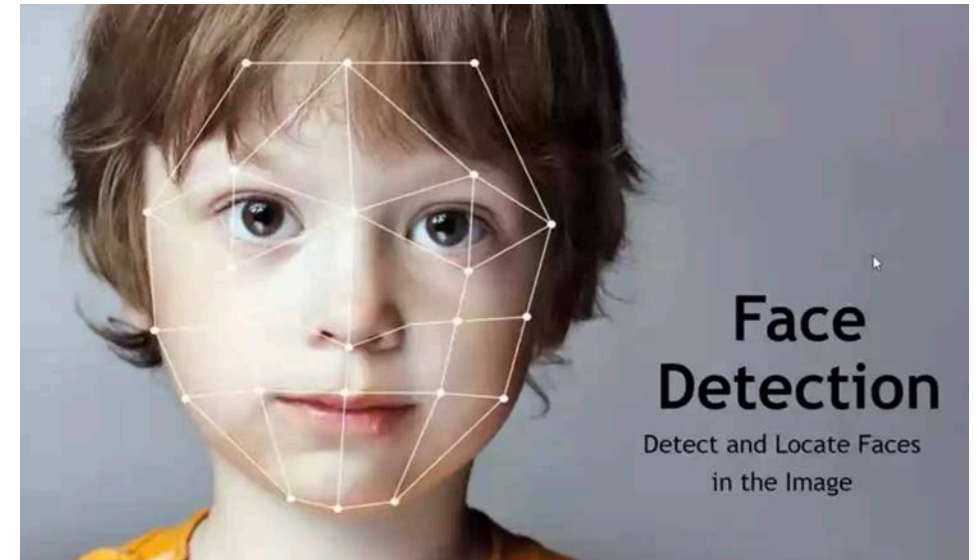
Computer Vision in object detection

- Object detection involves identifying the location of objects within an image and localizing them by drawing bounding boxes around the objects.
- Applications of object detection span various fields, including:
 - **Healthcare:** Detecting lesions.
 - **Industry/Robotics:** Identifying features of objects in industrial processes.
 - **Autonomous Vehicles:** Detecting traffic lights, obstacles (cars, people, etc.), and roads.
 - **Surveillance:** Monitoring environments for specific objects or activities.



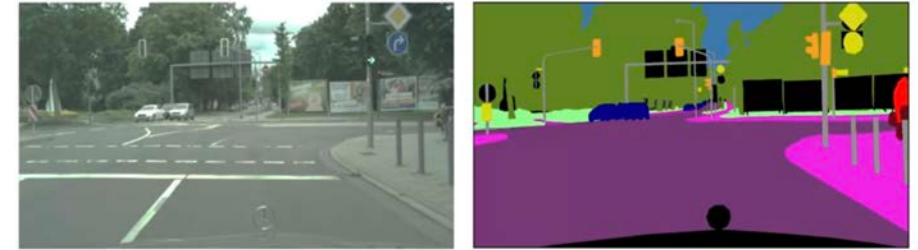
Computer Vision in facial recognition

- **Detect and Locate Faces:** Identifies faces in an image and determines their positions. Applies in security and surveillance.
- **Extract Unique Facial Features:** Analyzes key facial traits and converts them into numerical data. Used in smartphone unlocking.
- **Match to Stored Database:** Compares extracted features with a database to verify or identify individuals. Implemented in retail and marketing analysis, attendance systems, etc.

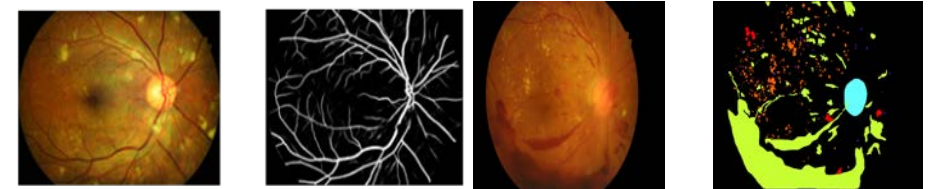


Computer Vision in segmentation

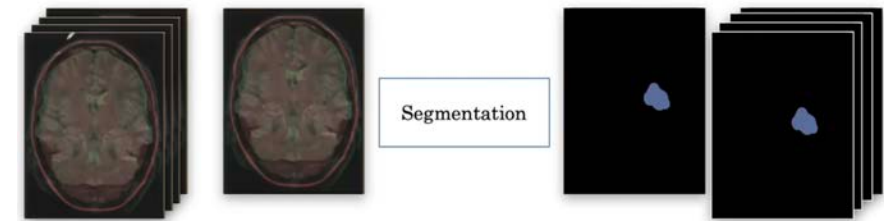
- Segmentation refers to the process of dividing an image into meaningful parts or regions to simplify its analysis.
 - to identify and classify every pixel in an image as belonging to a specific object or region of interest.
- There are two types of segmentation:
 - semantic segmentation
 - instance segmentation
- Application of segmentation in healthcare, autonomous vehicles, agriculture, surveillance, and image editing.



Semantic segmentation from Cityscapes



Blood vessel and lesion detection from Fundus image.



Lesion detection form Magnetic Resonance Imaging (MRI) images.

Challenges in Computer Vision

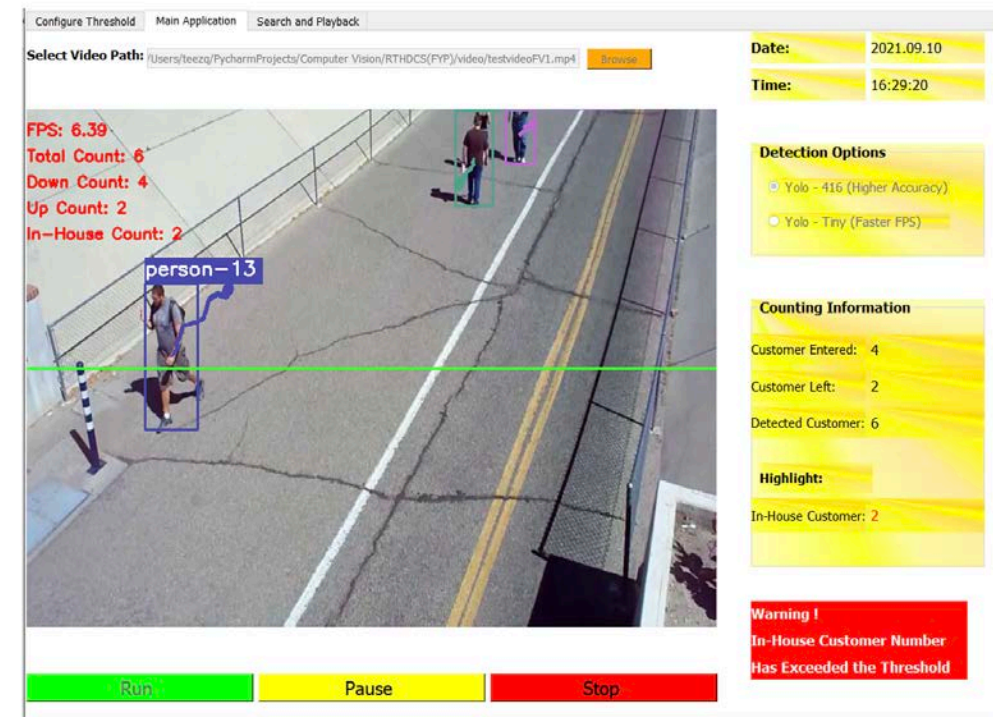
- Implementation Challenges
 - Lack of interdisciplinary knowledge
 - Data privacy protection
 - Cost
- Technical Limitations
 - Lack of Training Data
 - Non-transparency of Deep Learning Models
 - The Integration with Other Data Sources

Current Research in Computer Vision

Real-time human detection

- In 2023, Mokayed H. et al. proposed a real-time human detection system for detecting and counting people in shopping malls.
- Main Methods Used:
 - YOLOv3: For detecting and classifying human objects.
 - DeepSORT Tracking Algorithm: For tracking each detected human object.
- Dataset Used: COCO Dataset.
- Results: The proposed system demonstrated an accurate human detection performance with an accuracy of 91.07%.

Figure 2
Proposed system main application tab



Mokayed, H., Quan, T. Z., Alkhaled, L., & Sivakumar, V. (2023). Real-time human detection and counting system using deep learning computer vision techniques. In *Artificial Intelligence and Applications* (Vol. 1, No. 4, pp. 221-229).

Current Research in Computer Vision

Autonomous Driving Systems

- In 2024, Tan K. et al. discussed autonomous driving systems that integrate advanced computer vision with AI algorithms. This system is designed for application in autonomous vehicles, enabling them to perceive and understand their surroundings, as well as achieve intelligent decision-making and control.
- Three Main Tasks:
 - **Object Detection:** Identifies and locates various objects in the environment.
 - **Image Segmentation:** Classifies image pixels into target classes (e.g., ground, car, person, tree, etc.).
 - **Distance Estimation:** Calculates the distance to objects in the surroundings.

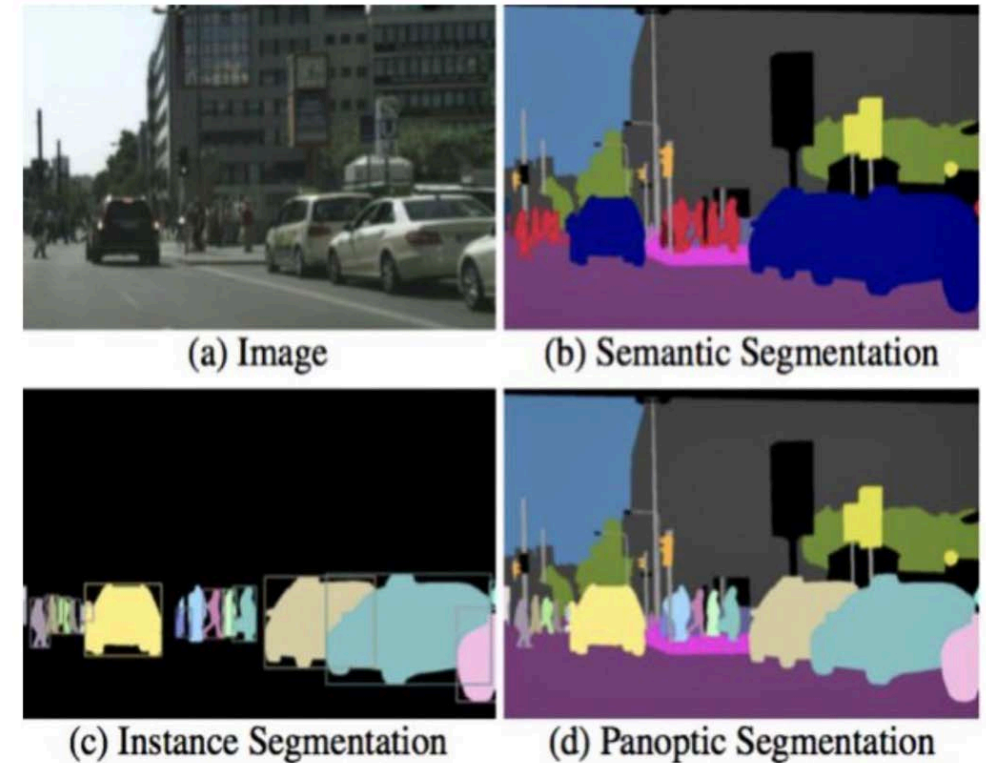


Figure 4: Driving section image segmentation program

Tan, K., Wu, J., Zhou, H., Wang, Y., & Chen, J. (2024). Integrating Advanced Computer Vision and AI Algorithms for Autonomous Driving Systems. *Journal of Theory and Practice of Engineering Science*, 4(01), 41-48.

