Machine Learning

Computer Vision

Salymbekov University

Miss Aliia Beishenalieva

aliya.beiwenalieva@gmail.com

What is the computer vision?

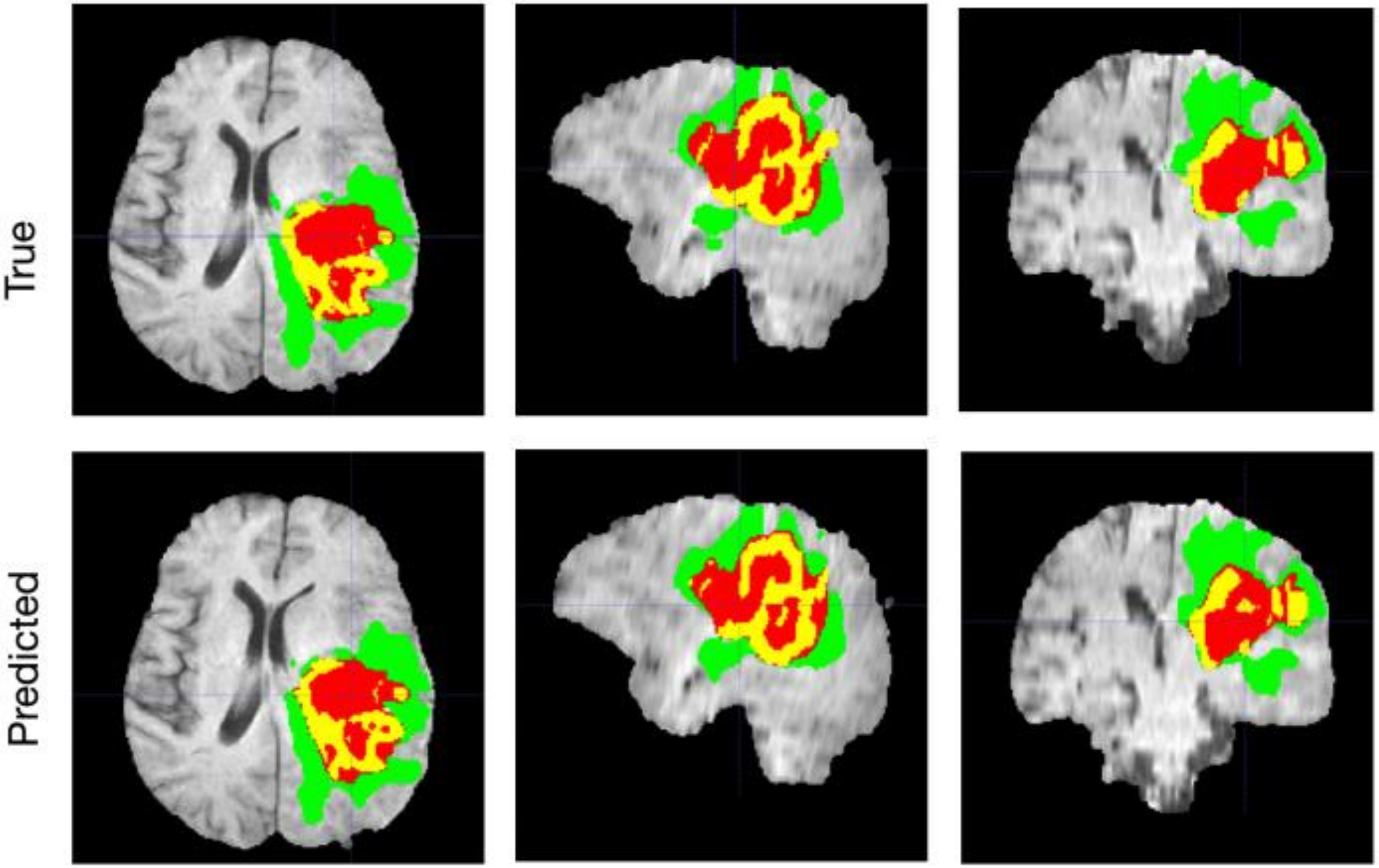
Computer Vision (CV) is a field of artificial intelligence that enables computers to interpret and make decisions based on visual data (images or videos). In essence, it seeks to teach machines to "see" the world as humans do.

Goals: The primary goal of computer vision is to develop algorithms and systems that can understand the visual world. This includes detecting objects, recognizing patterns, and interpreting scene structure.

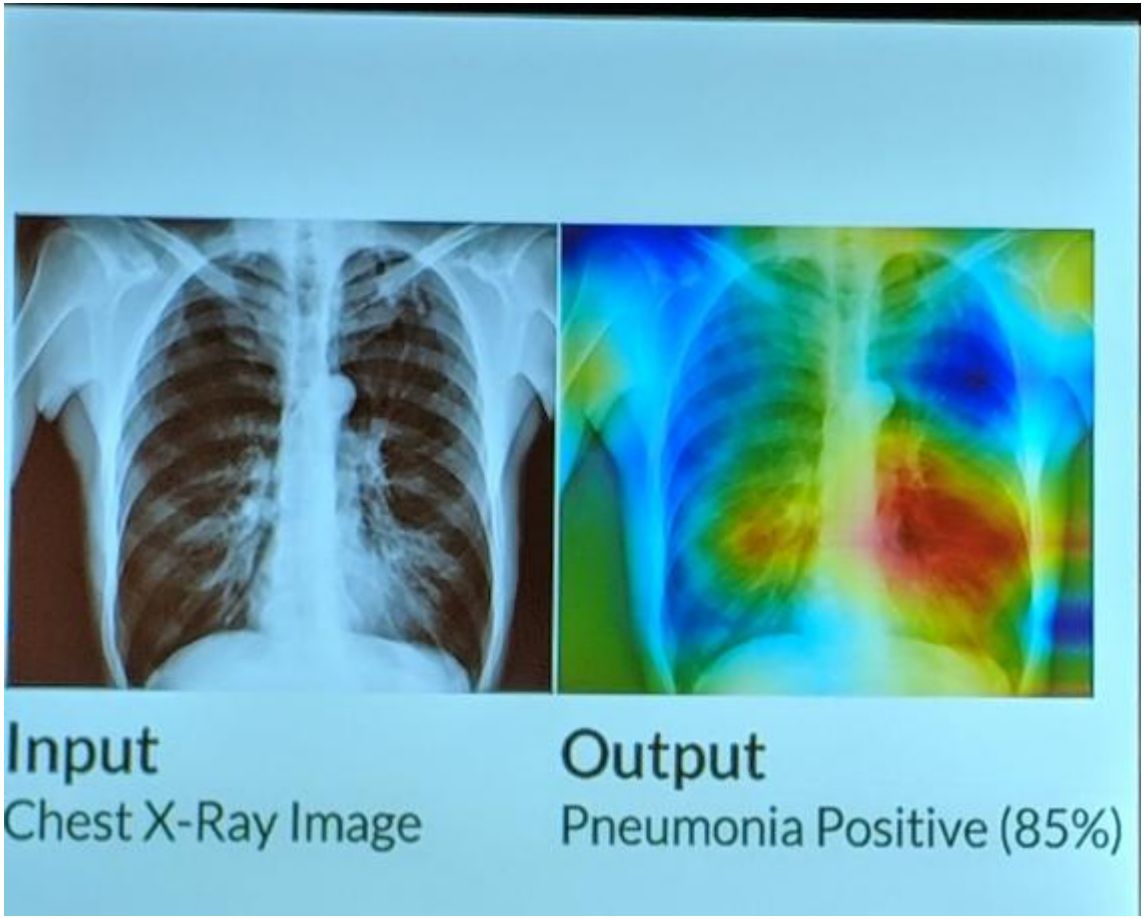
Applications

- **Healthcare**: Analyzing medical images for disease diagnosis.
- **Autonomous Vehicles**: Understanding the surrounding environment to make driving decisions.
- **Surveillance**: Real-time monitoring of spaces for security.
- **Retail**: Automated checkout systems that recognize products.

Healthcare



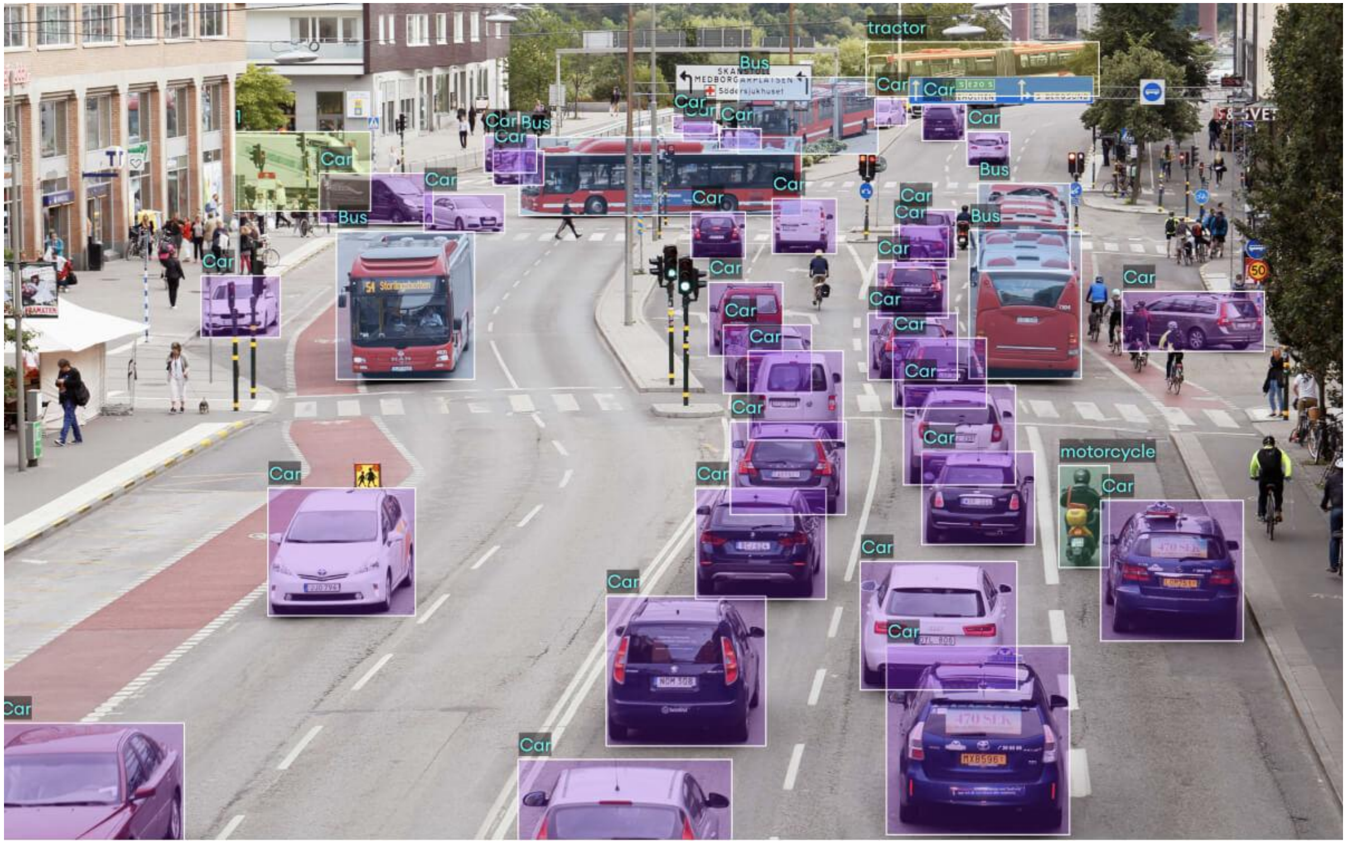
Healthcare



Input
Chest X-Ray Image

Output
Pneumonia Positive (85%)

Autonomous Vehicles



Surveillance

Surveillance

Retail

Brief History of Computer Vision

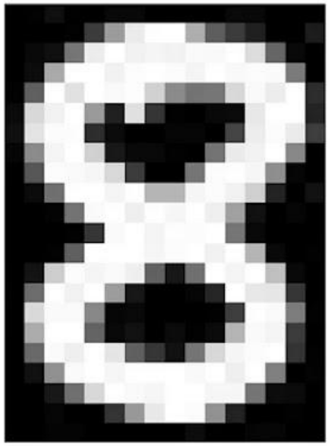
Early Beginnings (1960s-1980s): The field started with basic image processing and pattern recognition. Simple tasks included edge detection and identifying geometric shapes in images.

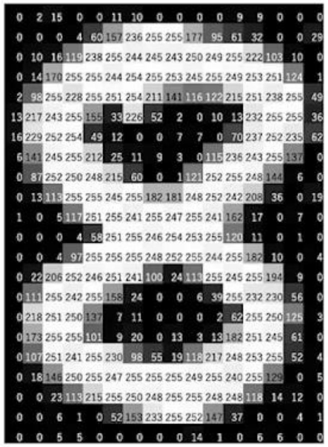
- **Rise of Machine Learning (1990s-2000s)**: As machine learning advanced, so did computer vision. Algorithms could classify simple images using methods like Support Vector Machines and K-nearest neighbors.

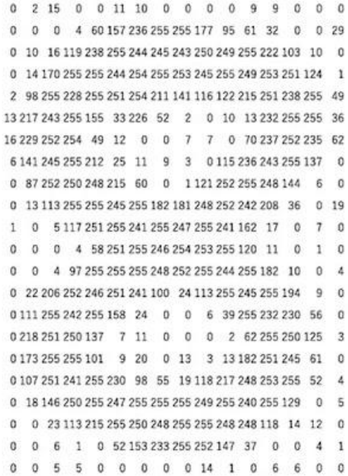
- **Deep Learning Era (2010s-Present)**: With neural networks, computer vision saw explosive growth. Convolutional Neural Networks (CNNs) allowed for high-accuracy image classification, object detection, and even generating new images.

What is an image in CV?

Introduction to Image Data

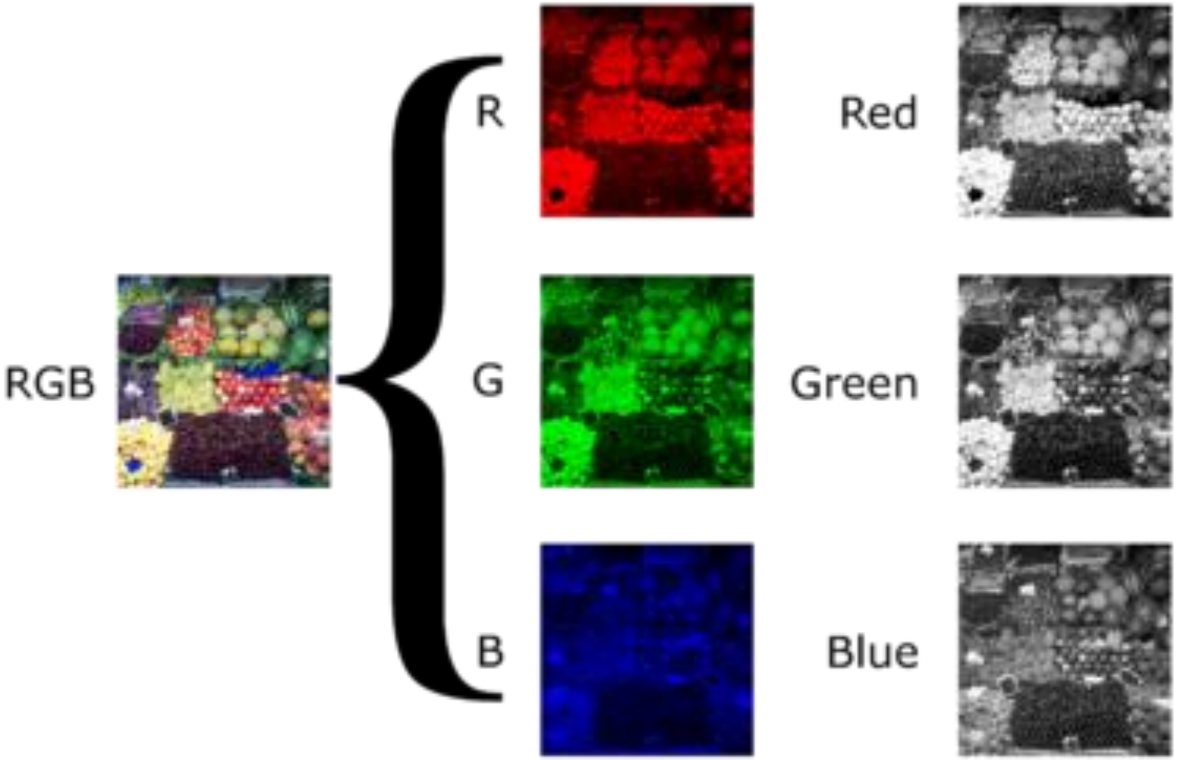






0	2	15	0	0	11	10	0	0	0	0	9	9	0	0	0
0	0	0	4	60	157	236	255	255	177	95	61	32	0	0	29
0	10	16	119	238	255	244	245	243	250	249	255	222	103	10	0
0	14	170	255	255	244	254	255	253	245	255	249	253	251	124	1
2	98	255	228	255	251	254	211	141	116	122	215	251	238	255	49
13	217	243	255	155	33	226	52	2	0	10	13	232	255	255	36
16	229	252	254	49	12	0	0	7	7	0	70	237	252	235	62
6	141	245	255	212	25	11	9	3	0	115	236	243	255	137	0
0	87	252	250	248	215	60	0	1	121	252	255	248	144	6	0
0	13	113	255	255	245	255	182	181	248	252	242	208	36	0	19
1	0	5	117	251	255	241	255	247	255	241	162	17	0	7	0
0	0	0	4	58	251	255	246	254	253	255	120	11	0	1	0
0	0	4	97	255	255	255	248	252	255	244	255	182	10	0	4
0	22	206	252	246	251	241	100	24	113	255	245	255	194	9	0
0	111	255	242	255	158	24	0	0	6	39	255	232	230	56	0
0	218	251	250	137	7	11	0	0	0	2	62	255	250	125	3
0	173	255	255	101	9	20	0	13	3	13	182	251	245	61	0
0	107	251	241	255	230	98	55	19	118	217	248	253	255	52	4
0	18	146	250	255	247	255	255	255	249	255	240	255	129	0	5
0	0	23	113	215	255	250	248	255	255	248	248	118	14	12	0
0	0	6	1	0	52	153	233	255	252	147	37	0	0	4	1
0	0	5	5	0	0	0	0	0	14	1	0	6	6	0	0

What is an image in CV?



Practice

1. Find 3 cases for CV in healthcare, surveillance, and retail
2. Briefly explain them

Machine Learning

How can computer see an image?

Salymbekov University

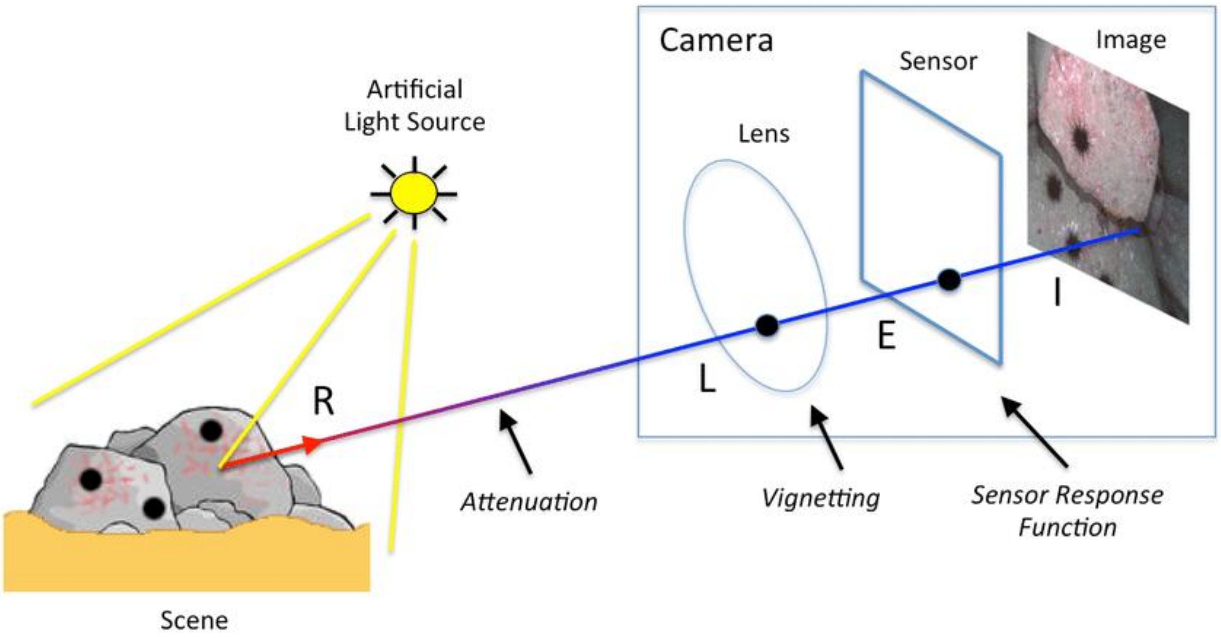
Miss Aliia Beishenalieva

aliya.beiwenalieva@gmail.com

Computer Vision

If AI enables computers to think, computer vision enables them to see, observe and understand.

How is an image formed?

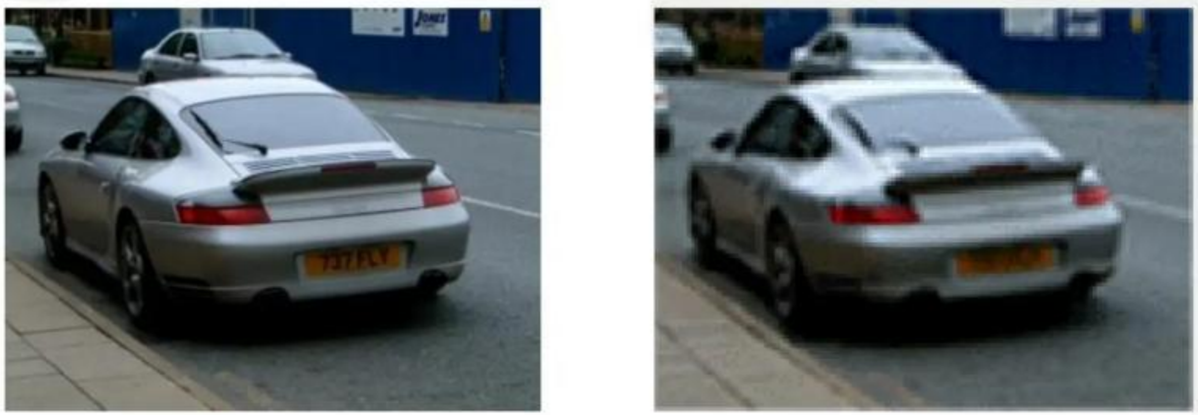


Images are formed by capturing light reflected or emitted from objects in the world around us. When light hits an object, it is reflected back and captured by a camera, which records the intensity and color information of the light. The camera's sensor converts the light into digital data, which is stored as an image file.

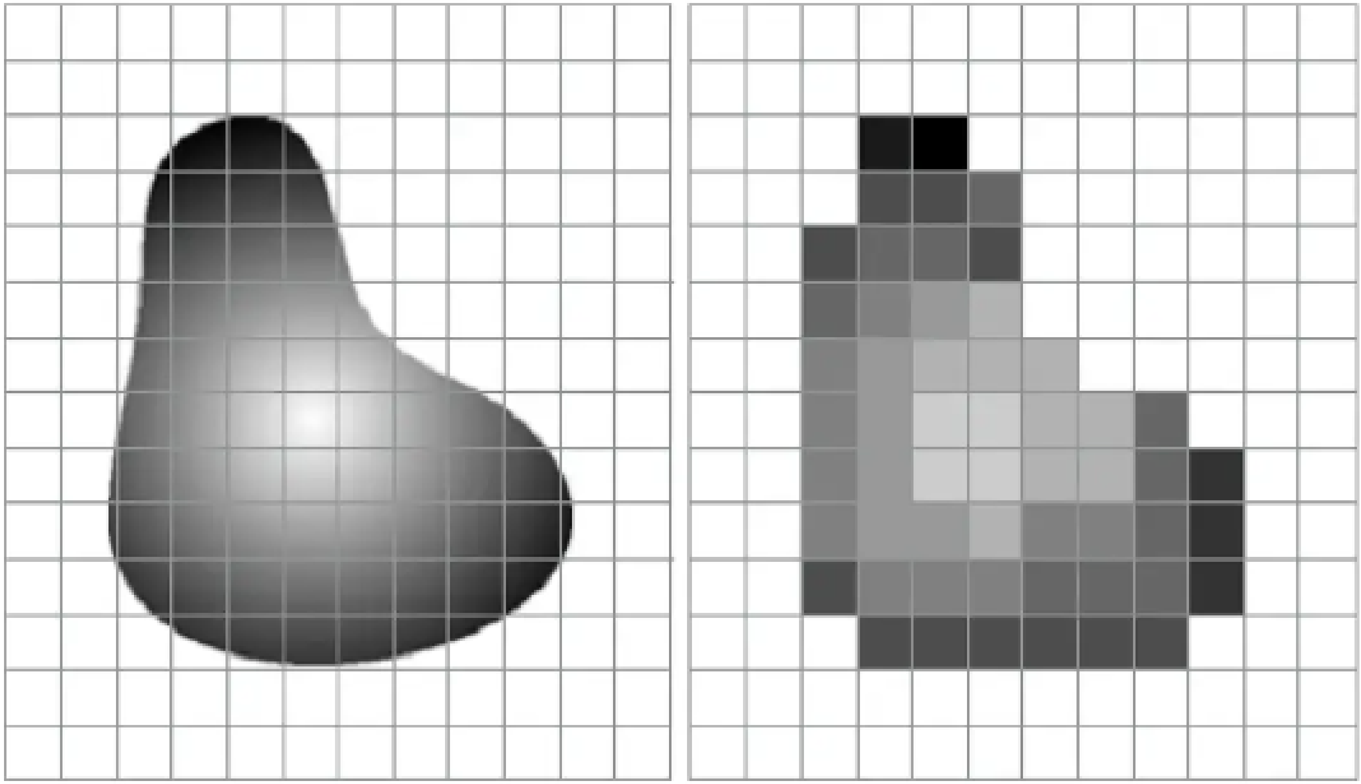
How can computer see an image?

The resolution of the image is determined by the number of pixels in the image, with each pixel representing a tiny unit of the image. The higher the number of pixels, the greater the detail and resolution of the image.

Resolution: How Much is Enough?



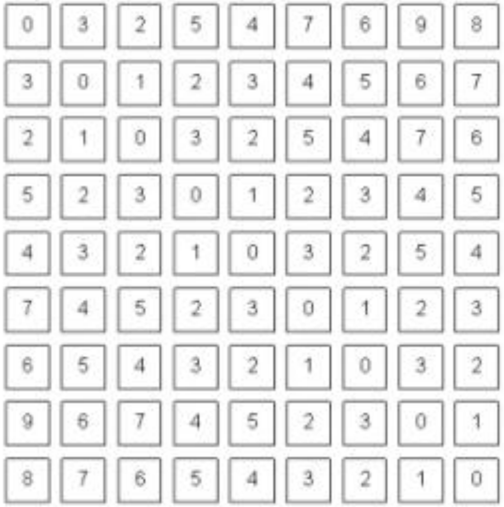
How can computer see an image?



How can computer see an image?

For a computer to "see" images, it needs to process them as data. Unlike humans, computers don’t actually interpret images in the way our eyes and brains do. Instead, they break down images into a format that they can analyze mathematically.

What we see



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

What a computer sees

How can computer see an image?

The range from **0 to 255** is fundamental in digital imaging, as it represents the intensity or brightness level of each pixel in an image. This range is due to how computers store color information in **8-bit** format, which is the standard for many digital images.

❑ **Understanding the 0-255 Range: 8-Bit Depth**

In an **8-bit system**, each pixel value is stored as a binary number using 8 bits (or binary digits).

With 8 bits, there are 2^8 =256 possible values, which means we can represent intensity levels from **0** to **255**.

This range gives a good balance between color detail and efficient storage. More bits per pixel (like 16-bit or 32-bit) allow for more detail but require more memory.

How can computer see an image?

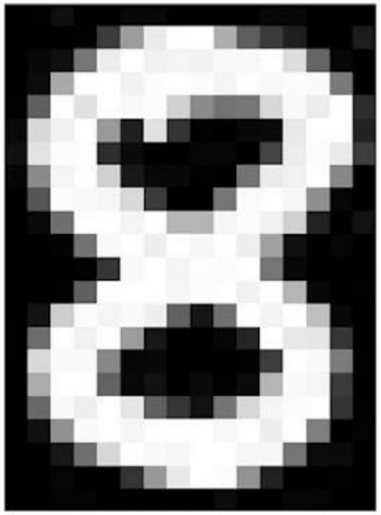
❑ **0-255 in Grayscale Images**

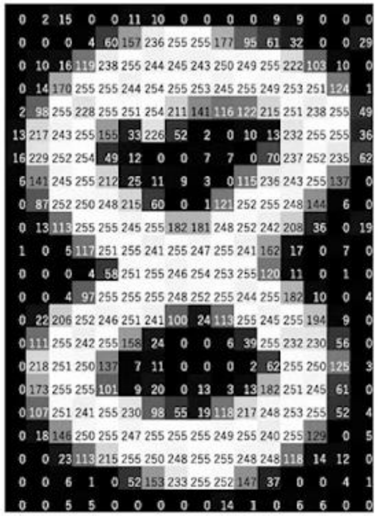
In grayscale images, each pixel has a single intensity value from **0** to **255**:

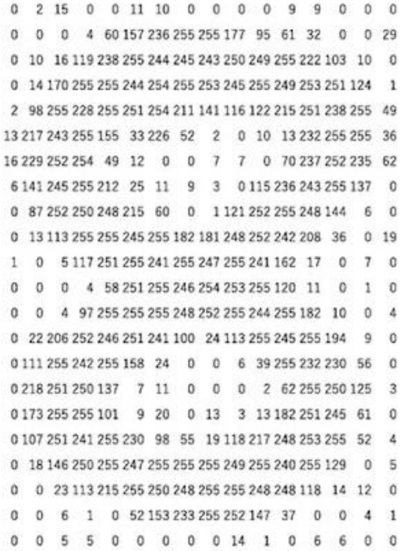
- **0** represents pure black (no brightness).
- **255** represents pure white (maximum brightness).
- Values in between represent shades of gray. For example, 127 would be a medium gray.

This grayscale scale allows an image to show varying levels of brightness, providing depth to the image without color.

How can computer see an image?







0	2	15	0	0	11	10	0	0	0	0	9	9	0	0	0
0	0	0	4	60	157	236	255	255	177	95	61	32	0	0	29
0	10	16	119	238	255	244	245	243	250	249	255	222	103	10	0
0	14	170	255	255	244	254	255	253	245	255	249	253	251	124	1
2	98	255	228	255	251	254	211	141	116	122	215	251	238	255	49
13	217	243	255	155	33	226	52	2	0	10	13	232	255	255	36
16	229	252	254	49	12	0	0	7	7	0	70	237	252	235	62
6	141	245	255	212	25	11	9	3	0	115	236	243	255	137	0
0	87	252	250	248	215	60	0	1	121	252	255	248	144	6	0
0	13	113	255	255	245	255	182	181	248	252	242	208	36	0	19
1	0	5	117	251	255	241	255	247	255	241	162	17	0	7	0
0	0	0	4	58	251	255	246	254	253	255	120	11	0	1	0
0	0	4	97	255	255	255	248	252	255	244	255	182	10	0	4
0	22	206	252	246	251	241	100	24	113	255	245	255	194	9	0
0	111	255	242	255	158	24	0	0	6	39	255	232	230	56	0
0	218	251	250	137	7	11	0	0	0	2	62	255	250	125	3
0	173	255	255	101	9	20	0	13	3	13	182	251	245	61	0
0	107	251	241	255	230	98	55	19	118	217	248	253	255	52	4
0	18	146	250	255	247	255	255	255	249	255	240	255	129	0	5
0	0	23	113	215	255	250	248	255	255	248	248	118	14	12	0
0	0	6	1	0	52	153	233	255	252	147	37	0	0	4	1
0	0	5	5	0	0	0	0	0	14	1	0	6	6	0	0

How can computer see an image?

❑ **0-255 in Color Images (RGB)**

In **color images**, each pixel has three color channels: Red, Green, and Blue (RGB). Each channel also has values from **0 to 255**.

By combining the intensity levels of these three channels, we can create different colors.

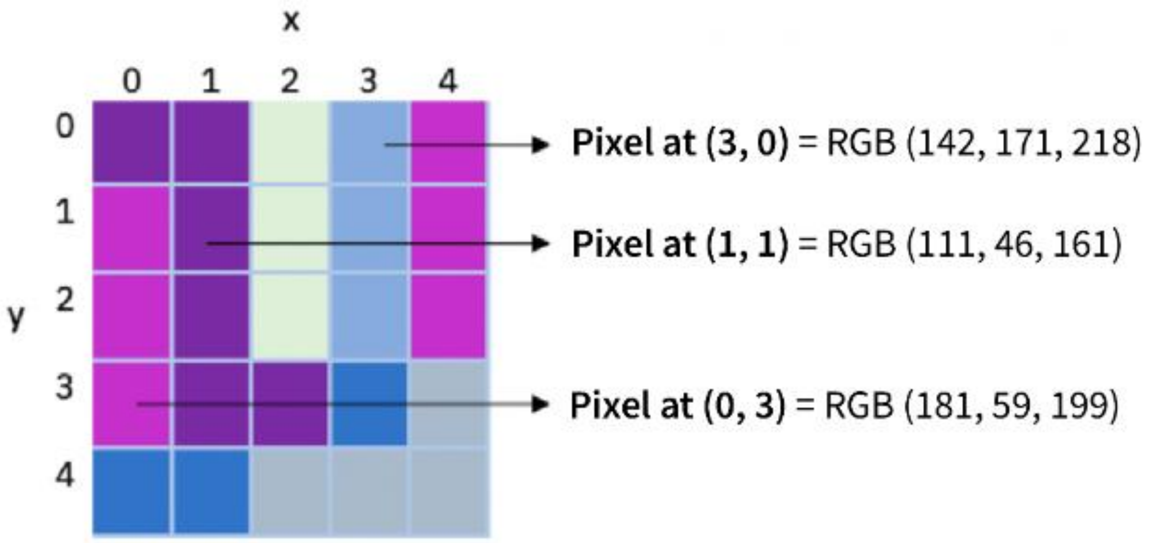
Here’s how the values for each channel translate:

- **(0, 0, 0)** = Black (no light in any channel)
- **(255, 255, 255)** = White (maximum light in each channel)
- **(255, 0, 0)** = Pure Red (maximum red, no green or blue)
- **(0, 255, 0)** = Pure Green (maximum green, no red or blue)
- **(0, 0, 255)** = Pure Blue (maximum blue, no red or green)
- **(255, 255, 0)** = Yellow (mixing red and green at full intensity)

By varying these values, you get a wide range of colors. For example, a shade of purple might be represented by (128, 0, 128), which gives a mix of red and blue at half intensity.

How can computer see an image?

For color images, each pixel is represented by a combination of values for the three primary colors: red, green, and blue (RGB). The pixel values for each color channel range from 0 to 255, so a single pixel in an RGB image is represented by a three-dimensional vector of values (R, G, B).


x
0 1 2 3 4
y
0 1 2 3 4
Pixel at (3, 0) = RGB (142, 171, 218)
Pixel at (1, 1) = RGB (111, 46, 161)
Pixel at (0, 3) = RGB (181, 59, 199)

How can computer see an image?

Image processing basically includes the following three steps:

1. Importing the image

2. Analysing and manipulating the image

3. Output in which result can be altered image or report that is based on image analysis

How can computer see an image?

How does a computer read an image?

Computers don’t “see” images in the way humans do. Instead, they interpret images as arrays of numerical values. The basic process of how a computer reads and processes an image are:

1.Pixel Values: An image is made up of pixels, which are the smallest units of information in an image. Each pixel has a value that represents its color and intensity. In the case of an RGB image, there are three values for each pixel corresponding to the Red, Green, and Blue channels.

2.Digital Representation: The RGB values are usually represented as integers ranging from 0 to 255. 0 represents the absence of color (black), and 255 represents the maximum intensity of that color (full brightness).

3.Image Matrix: The computer reads the image as a matrix of numbers, where each element in the matrix corresponds to the pixel value at that location. For a color image, there are typically three matrices, one for each RGB channel.

4.Image Processing: Image processing algorithms are applied to manipulate these numerical representations. Common operations include resizing, cropping, filtering, and more.

How the pixel information is used in CV?

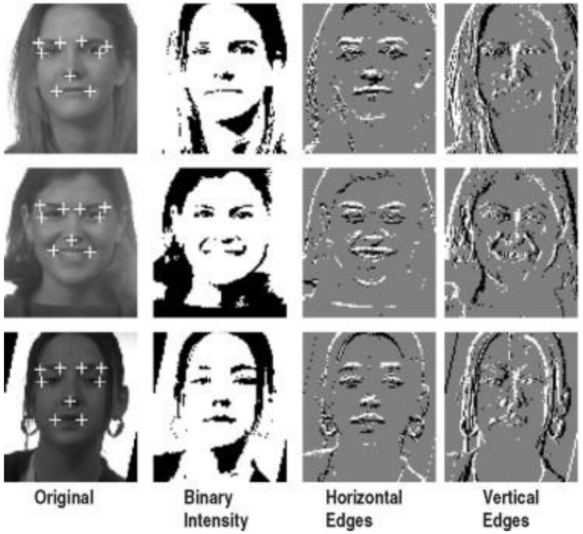
In computer vision (CV), the pixel values from 0 to 255, which represent intensity or color information, are essential because they serve as the raw data that models and algorithms analyze to "understand" and interpret images.

How the pixel information is used in CV?

Feature Detection

- **What it is**: Finding important parts in an image, like edges or corners, to help identify objects.

- **Example**: Edge detection can help find the boundaries of a car in a photo. This is useful for tasks like face detection or recognizing shapes.



How the pixel information is used in CV?

Image Preprocessing

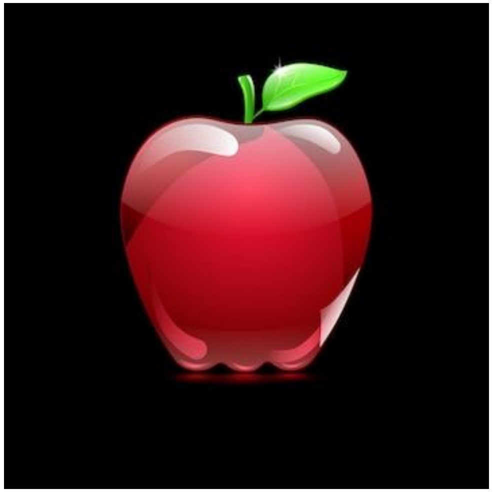
- ❑ **What it is**: Cleaning and preparing an image to make it easier for a computer to understand.

- ❑ **Example**: If an image is blurry or noisy, techniques like smoothing or adjusting brightness can help highlight important features, like a person's face in a crowded scene.

How the pixel information is used in CV?

Color Analysis

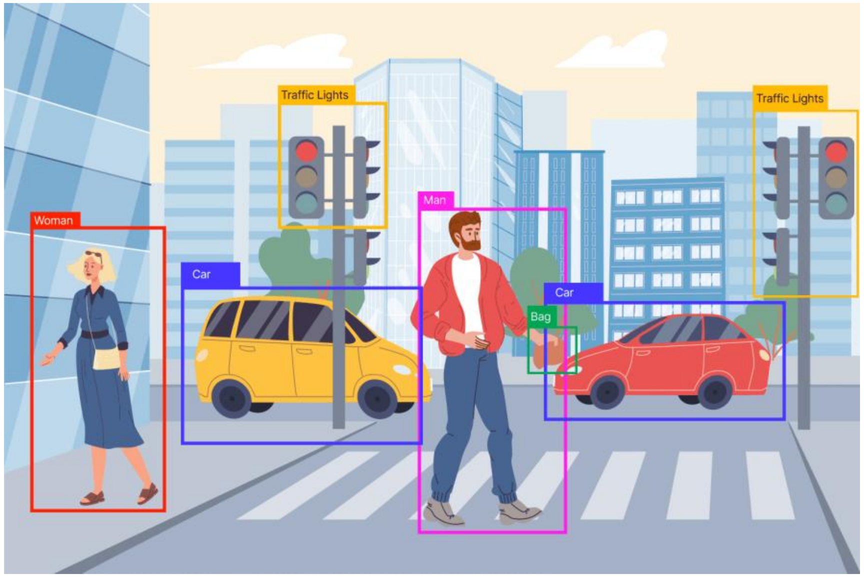
- ❑ **What it is**: Using colors to recognize objects or regions in an image.
- ❑ **Example**: In a photo of fruit, a computer can look at the red color of an apple to recognize it as an apple, even if the background changes.



How the pixel information is used in CV?

Object Detection

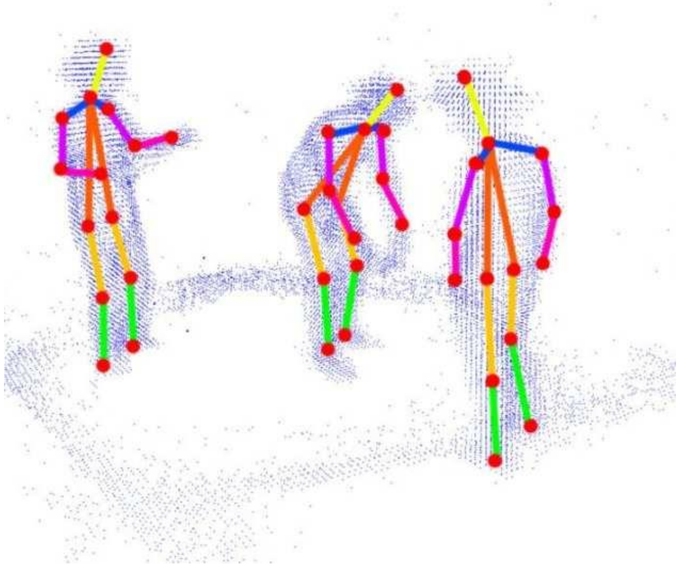
- ❑ **What it is**: Identifying and locating objects in an image.
- ❑ **Example**: In a street photo, detecting cars, pedestrians, and traffic lights. This is important for self-driving cars to understand their environment.



How the pixel information is used in CV?

Optical Flow (Motion Detection)

- ❑ **What it is**: Tracking movement by watching how pixels change over time in a video.
- ❑ **Example**: Detecting someone running in a video by seeing how their position changes from frame to frame.

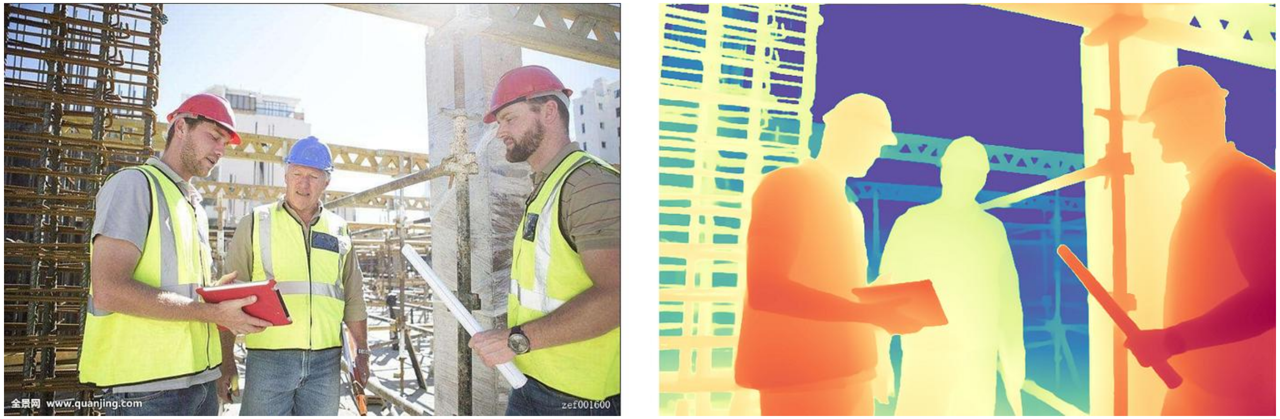


How the pixel information is used in CV?

Depth Estimation

❑ **What it is**: Measuring how far away objects are in an image.

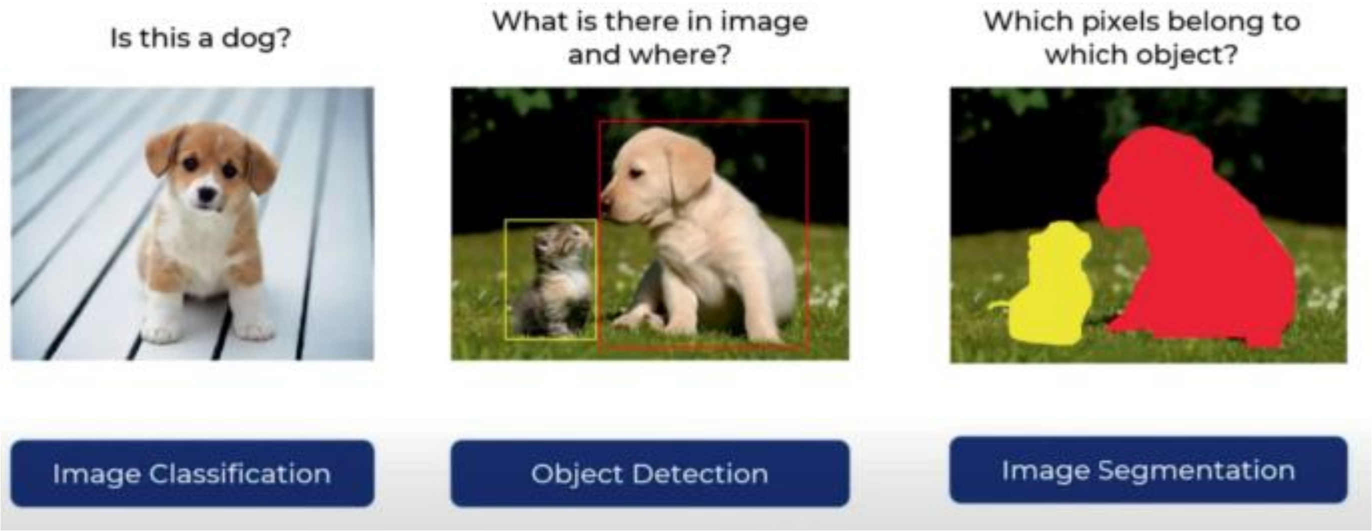
❑ **Example**: In 3D maps or when driving a car, understanding how far away obstacles are by analyzing pixel data.



How the pixel information is used in CV?

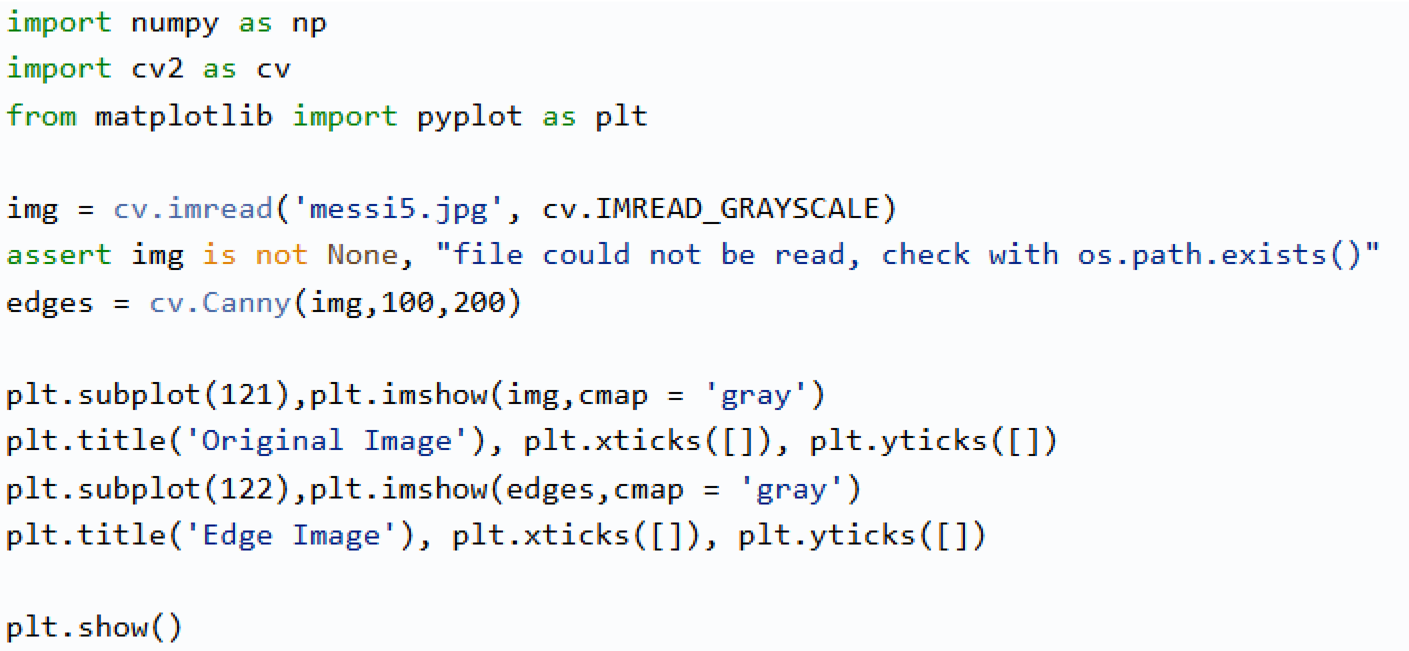
Image Classification

- ❑ **What it is**: Teaching a computer to recognize what objects are in an image.
- ❑ **Example**: Telling the difference between a cat and a dog in a photo by analyzing the patterns in the image's pixels.


Is this a dog?
What is there in image and where?
Which pixels belong to which object?
Image Classification
Object Detection
Image Segmentation

How the pixel information is used in CV?

- ❑ **Canny Edge Detection**



```python
import numpy as np
import cv2 as cv
from matplotlib import pyplot as plt

img = cv.imread('messi5.jpg', cv.IMREAD_GRAYSCALE)
assert img is not None, "file could not be read, check with os.path.exists()"
edges = cv.Canny(img,100,200)

plt.subplot(121),plt.imshow(img,cmap = 'gray')
plt.title('Original Image'), plt.xticks([]), plt.yticks([])
plt.subplot(122),plt.imshow(edges,cmap = 'gray')
plt.title('Edge Image'), plt.xticks([]), plt.yticks([])

plt.show()
```

https://docs.opencv.org/4.x/da/d22/tutorial_py_canny.html

Thank you for your attention

Computer Vision

Salymbekov University
Miss Aliia Beishenalieva
aliya.beiwenalieva@gmail.com