



Odisha Corporate Foundation

# Computer Vision

AI MACHINE LEARNING SUMMER SCHOOL

# ABHIJEET PARIDA



## EDUCATION

### **B. Tech. Mechanical Engineering**

Amrita School of Engineering  
Bengaluru, India

Major - Hypersonic Flows and  
Gas Dynamics

### **M. Sc. Computational Science and Engineering**

Technical University of Munich  
Munich, Germany

Major - Deep Learning for  
Medical Imaging

### **Research Interests-**

Deep Learning for Medical  
Imaging, Few Shot Learning,  
UAD, Machine Learning,  
Semantic Segmentation, Deep  
Learning for Fluid Flows.

## WORK EXPERIENCE

### **Data Scientist**

deepc GmbH  
(05/2019- 05/2022)



Development of the  
orchestration of AI  
models and design of  
workflows for processing  
of radiological image  
data.

### **R&D DevOps Eng**

Children's National  
(06/2022-)



Development of Children  
specific Imaging DL  
Models for Surgical  
Innovation and Precision  
Medicine with GWU and  
CHLA.

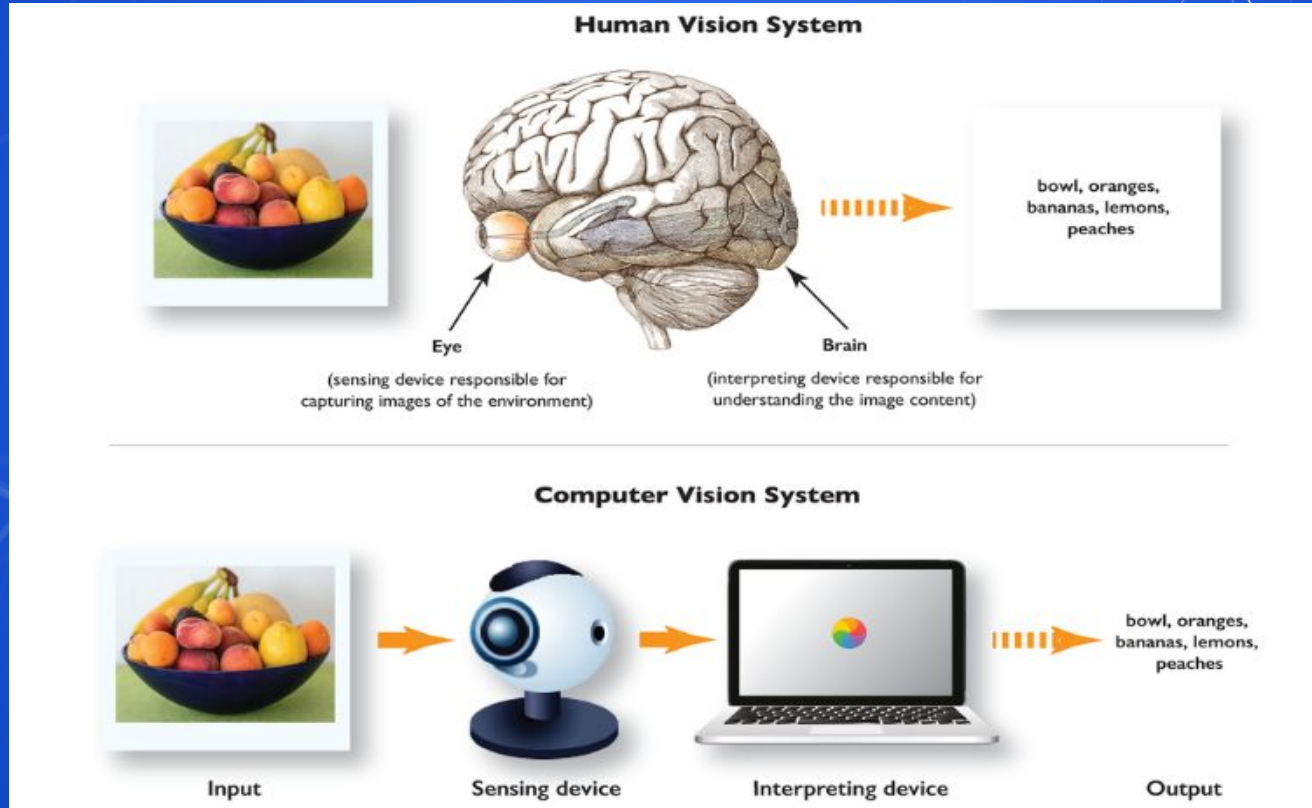
## OPEN MODEL

- Neural Network-Based Segmentation of MS Lesions in FLAIR Images ([DockerHub](#))
- Anatomy Based GM, WM and CSF Segmentation ([DockerHub](#))

# CONTENT

- What is Computer Vision?
- Why Computer Vision is Hard?
- What is an Image?
- What is a Video?
- What are major Image Processing Techniques?
- Demo of a task
- Outlook
- Outro

# What is Computer Vision?



# What is Computer Vision?: CV Tasks

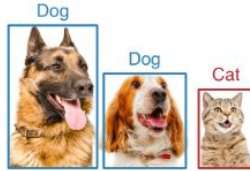
## Image Classification:

recognize an object in an image.



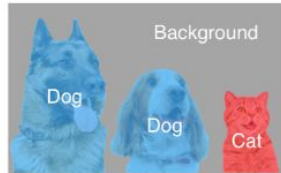
## Object Detection:

detect multiple objects with their bounding boxes in an image.



## Semantic Segmentation:

associate each pixel of an image with a categorical label.



## Instance Segmentation:

associate each pixel of an image with an instance label.





# What is Computer Vision?: CV Tasks

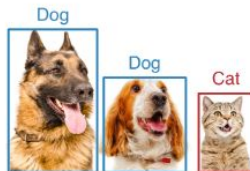
## Image Classification:

recognize an object in an image.



## Object Detection:

detect multiple objects with their bounding boxes in an image.



## Semantic Segmentation:

associate each pixel of an image with a categorical label.



## Instance Segmentation:

associate each pixel of an image with an instance label.



Image Classification



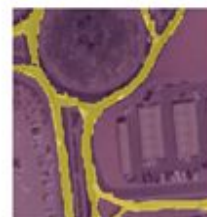
Object Detection



Semantic Segmentation



Instance Segmentation



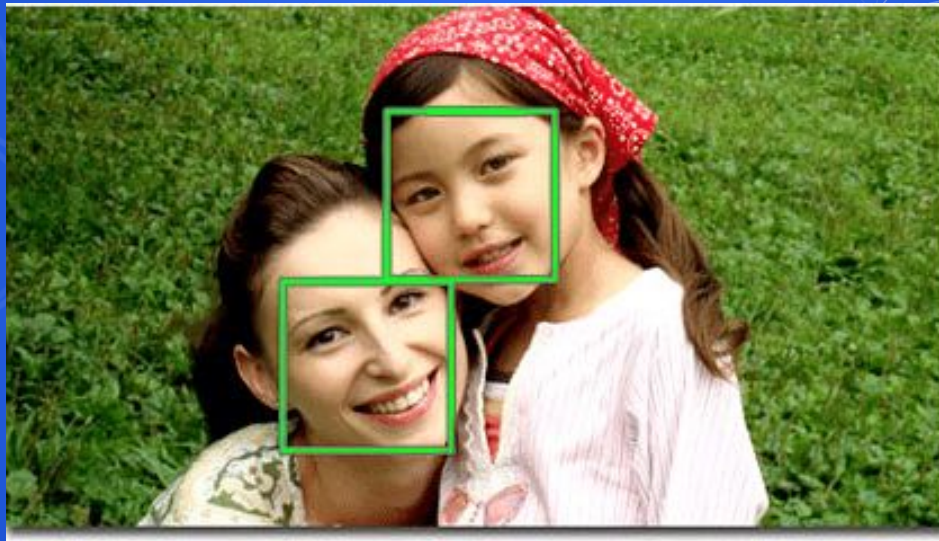
# What is Computer Vision?: Applications

## Optical Character Recognition



# What is Computer Vision?: Applications

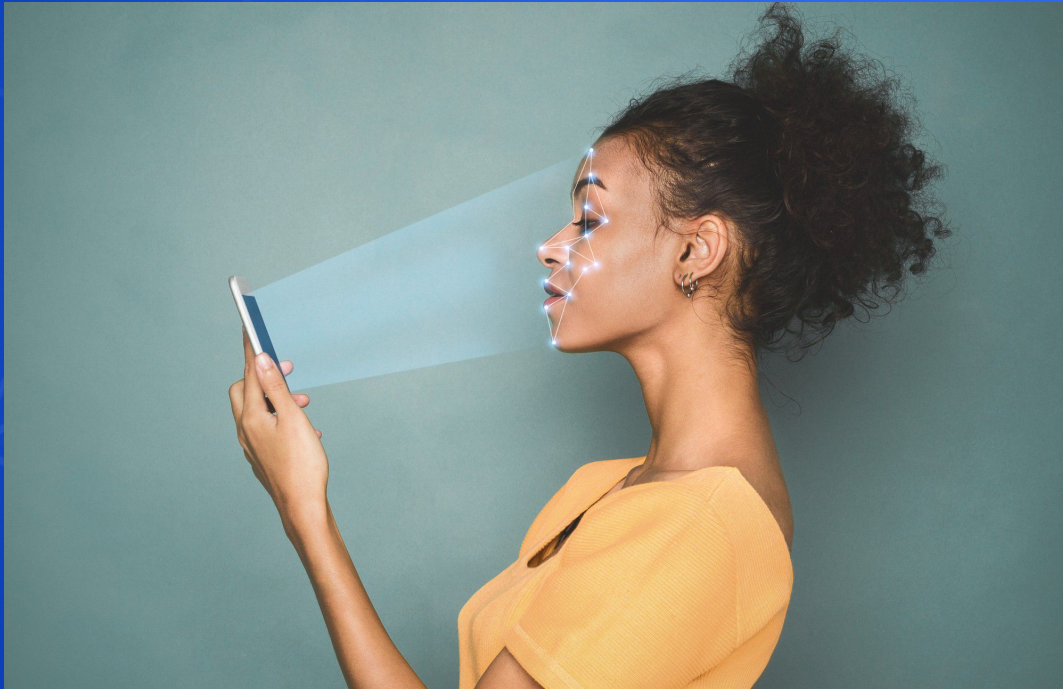
Face Detection/ Sony Smile Shutter





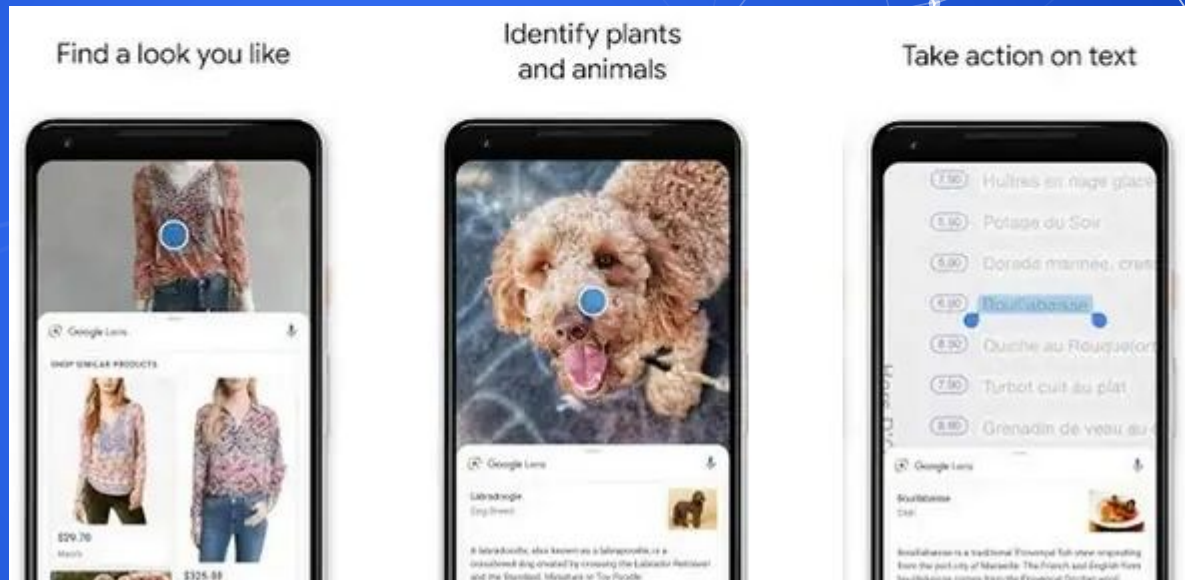
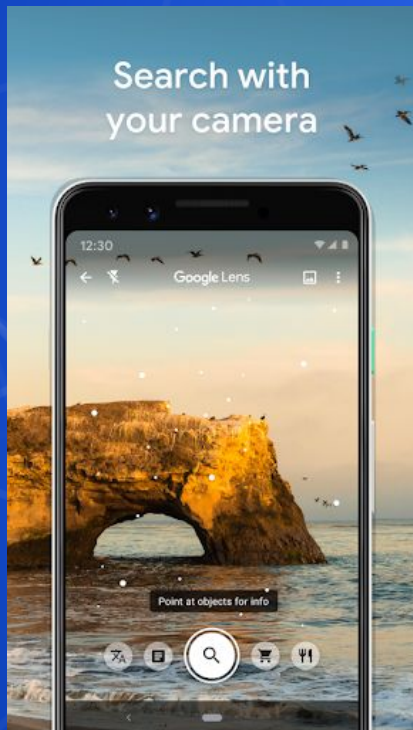
# What is Computer Vision?: Applications

Facial Recognition/ ID Card Scanning



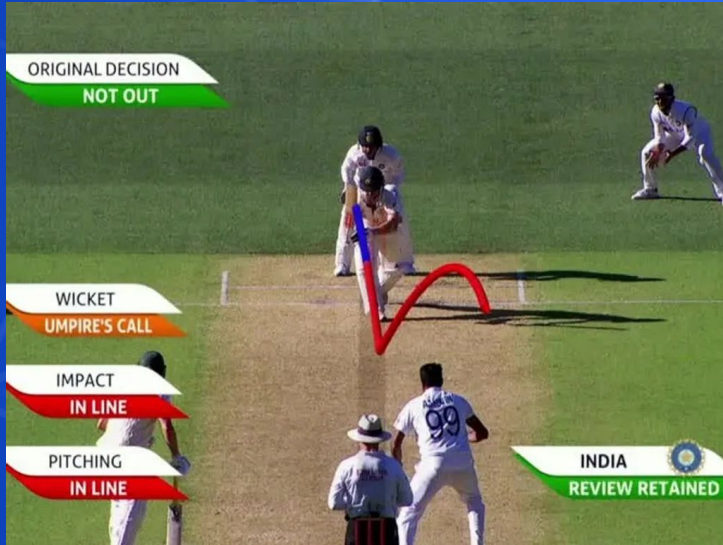
# What is Computer Vision?: Applications

## Google Lens



# What is Computer Vision?: Applications

## Sports Ball Tracking





# What is Computer Vision?: Applications

Medical Imaging

# Why Computer Vision is Hard?

Google

konark sun temple



Konark Sun Temple - Wikipedia  
en.wikipedia.org



Konark Sun Temple - World History ...  
worldhistory.org



Konark Sun Temple - Wikipedia  
en.wikipedia.org



Konark Sun Temple  
tripadvisor.com



The Legend of the Konark Sun Temple ...  
ancient-origins.net



Konark Sun Temple in Odisha: Essential ...  
tripsavvy.com



Added by British for stability, sand ...  
indianexpress.com



Legend of the Konark Sun Temple ...  
newindianexpress.com



13 Captivating Konark Sun Temple Facts ...  
kidadi.com



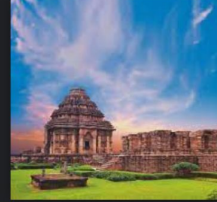
Awe-inspiring Konark Sun Temple ...  
navrangindia.blogspot.com



Konark Sun Temple is a ...  
in.pinterest.com



Konark Sun Temple  
tripadvisor.com

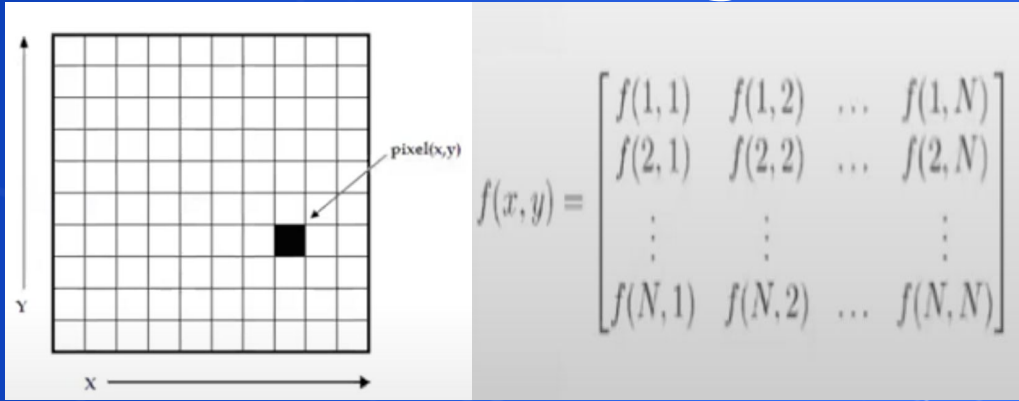


Konark: Chariot of the sun - G...  
artsandculture.google.com

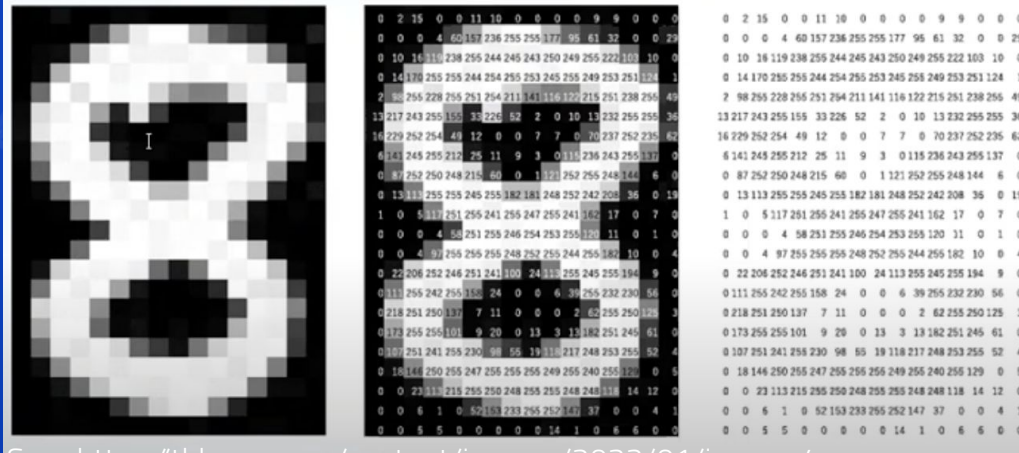
- Viewpoint
- Lighting
- Scale
- Camera Limitation
- Occlusion
- Obfuscation



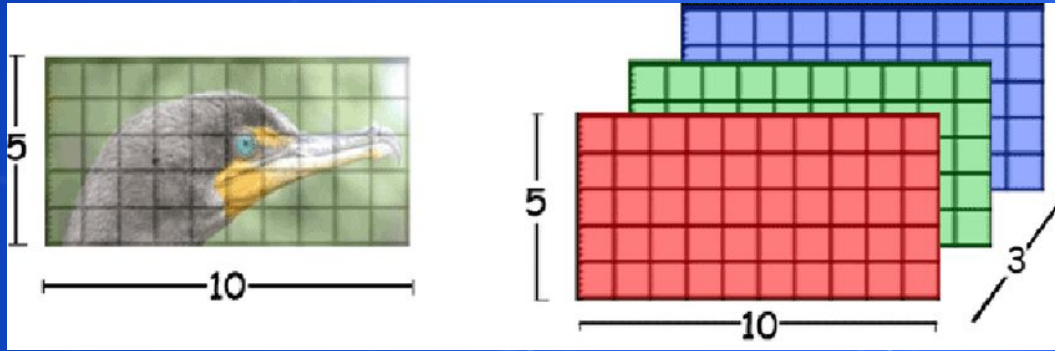
# What is an Image?: Grayscale



- Image is a combination of Pixels
- Pixels are spatial coordinate with a value
- The value is a number with range  $[0, 255]$
- Computer representation is a matrix

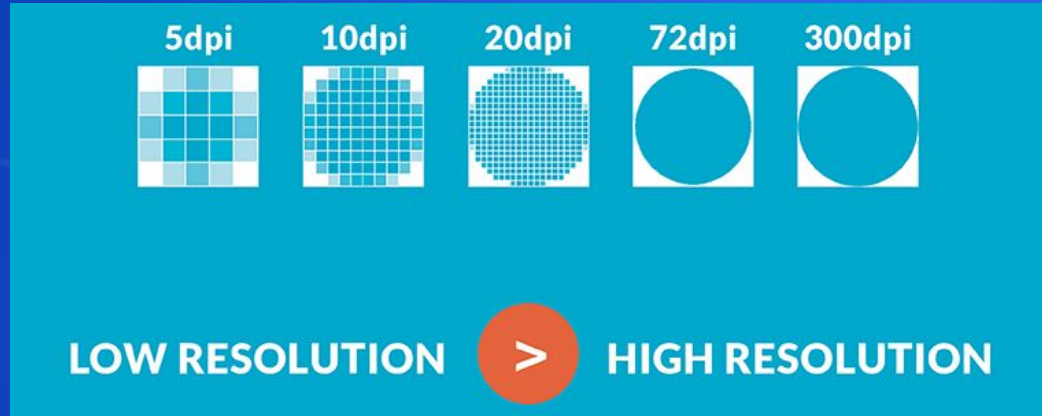


# What is an Image?: RGBA Image

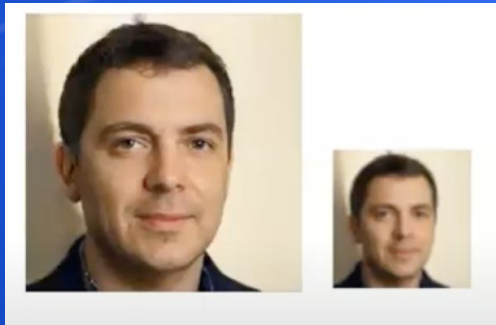


- Stacked multiple Channel to be interpreted as colors
- Pixel value signify intensity of the color
- Each Pixel of channel has range  $[0, 255]$
- All possible colors  $256^3 = 1,67,77,216$
- PNGs can have A value to signify transparency

# What is an Image?: Image Resolution



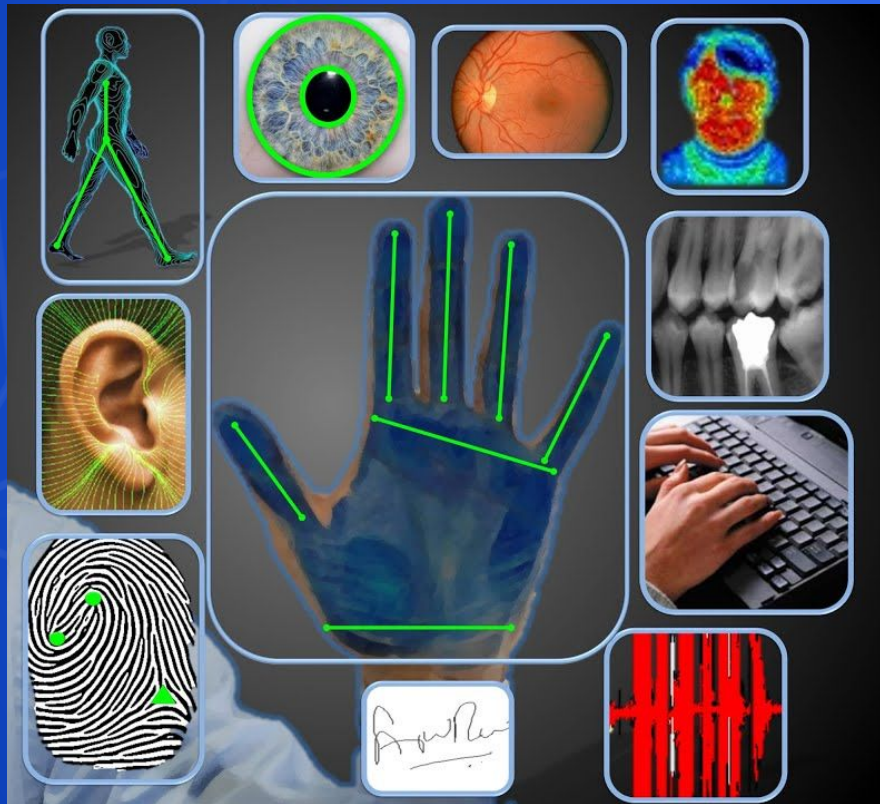
- No of Pixels per inch of screen
- Called PPI/DPI
- Higher resolution; higher PPI; more details
- Change of Resolution leads to loss of details



# What is a Video?



# Image Processing



- The mathematical notations equivalence of images are matrices
- The process modifying the values of the image matrix using algorithms
- They enhance and/or analyse the content
- DIY photoshop



# Image Processing: Cropping

The background of the slide is a solid blue color. Overlaid on this background is a complex, abstract geometric pattern. This pattern consists of numerous white dots of varying sizes, which are connected by thin, white lines. These lines and dots form a network of interconnected polygons and triangles, creating a sense of depth and structure. The overall effect is a modern, tech-oriented aesthetic.

# Image Processing: Resizing

The background of the slide is a solid blue color. Overlaid on this background is a complex, abstract geometric pattern. This pattern consists of numerous white dots (nodes) connected by thin white lines (edges). The connections form a variety of polygonal shapes, including triangles, quadrilaterals, and larger, more complex polygons. Some of these shapes are interconnected, creating a network-like structure, while others are isolated. The overall effect is a modern, technical, and digital aesthetic.

# Image Processing: Flipping/Rotation

The background is a solid blue gradient. Overlaid on this are several abstract geometric patterns. These consist of thin white lines connecting small white dots. The dots are scattered across the frame, and the lines form a network of interconnected polygons and triangles, creating a complex, web-like structure that resembles a molecular model or a network diagram. The overall aesthetic is clean, modern, and technical.

# Image Processing: Brightness/Contrast

The background of the slide is a solid blue gradient. Overlaid on this are several abstract geometric patterns. These consist of thin white lines connecting small white dots, forming a network of interconnected triangles and polygons. The lines and dots are more densely packed in some areas, particularly towards the top and right, and more sparse in others, creating a sense of depth and complexity. The overall effect is a modern, technological aesthetic.

# Image Processing: Edge Detection

The background of the slide is a solid blue color. Overlaid on this background is a complex, abstract geometric pattern. This pattern consists of numerous white dots of varying sizes, which are interconnected by thin, white lines. These lines and dots form a network of irregular polygons and geometric shapes, creating a sense of depth and connectivity. The overall effect is a modern, technological aesthetic that complements the technical nature of the topic.



# Image Processing: Dilation/Erosion

The background of the slide is a solid blue color. Overlaid on this background is a complex, abstract geometric pattern. This pattern consists of numerous white dots of varying sizes, which are interconnected by thin, white lines. These lines and dots form a network of irregular polygons and geometric shapes, creating a sense of depth and connectivity. The overall effect is a modern, tech-oriented aesthetic.

# Image Processing: Connected Components



# Image Processing: HDR

The background is a solid blue gradient. Overlaid on this are several abstract geometric patterns. These consist of thin white lines connecting small white dots. The dots are scattered across the frame, and the lines form a network of interconnected polygons, some of which are more complex and multi-faceted than others. The overall effect is a modern, technical, and digital aesthetic.

# Demo: Traffic Management System

The background of the slide is a solid blue color. Overlaid on this background is a complex, abstract pattern of thin white lines and small white dots. These lines and dots form a network-like structure, with some lines connecting dots to form polygons of various shapes and sizes. The pattern is more dense in the upper right and lower right areas, and more sparse in the lower left area.

# Outlook:CNNs



# Outlook: 3D Computer Vision



# Outro: Learning Goals

- What is Computer Vision?
- Why Computer Vision is Hard?
- What is an Image?
- What is a Video?
- What are major Image Processing Techniques?
- How multiple tasks can come together for an important application