



Odisha Corporate Foundation

AI MACHINE LEARNING SUMMER SCHOOL

Computer Vision

By Abhijeet Parida

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.

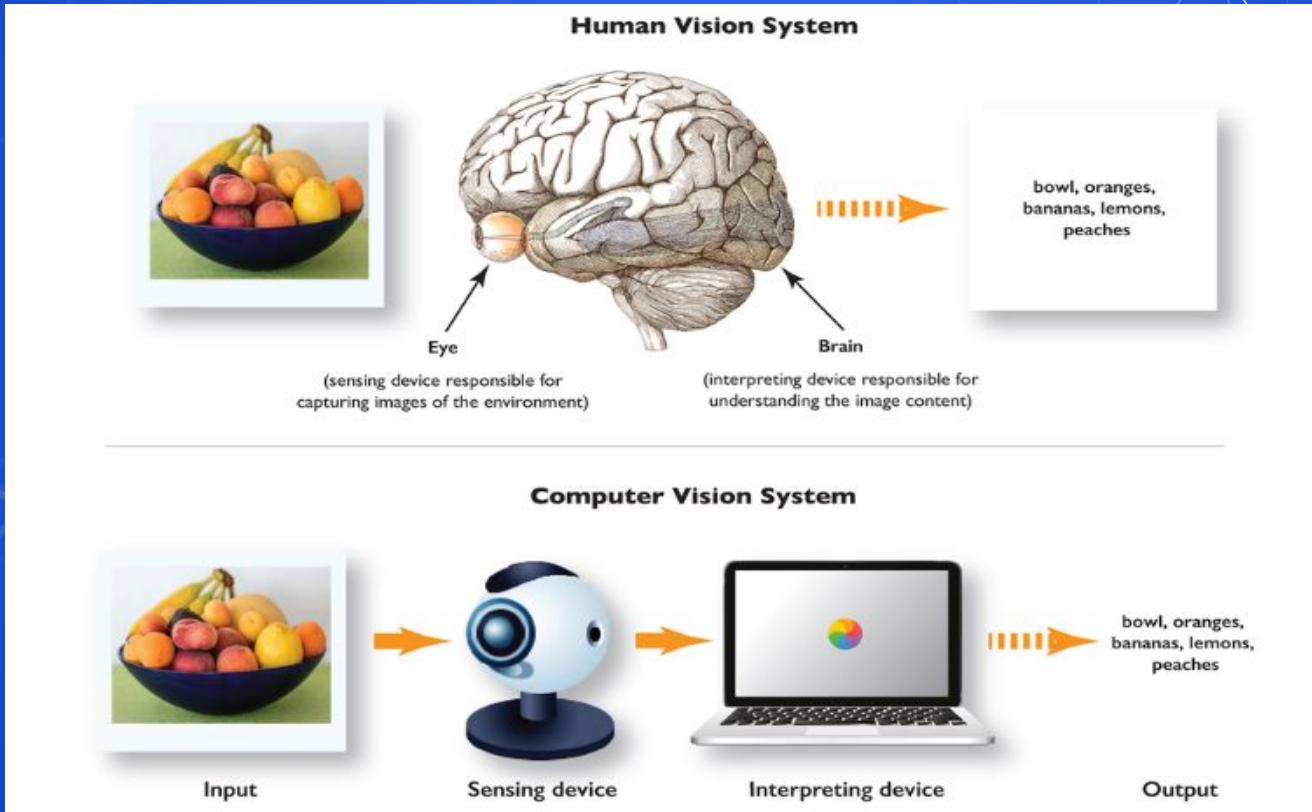
OPEN MODEL

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

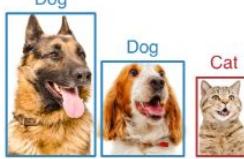
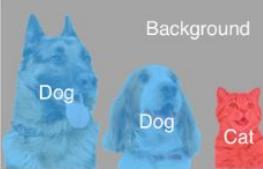
CONTENT

- 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

| | |
|---|--|
| <p><u>Image Classification:</u> recognize an object in an image.</p> |  |
| <p><u>Object Detection:</u> detect multiple objects with their bounding boxes in an image.</p> |  |
| <p><u>Semantic Segmentation:</u> associate each pixel of an image with a categorical label.</p> |  |
| <p><u>Instance Segmentation:</u> associate each pixel of an image with an instance label.</p> |  |

What is Computer Vision?: CV Tasks

Image Classification:

recognize an object in an image.

Dog



Object Detection:

detect multiple objects with their bounding boxes in an image.

Dog



Semantic Segmentation:

associate each pixel of an image with a categorical label.

Background



Instance Segmentation:

associate each pixel of an image with an instance label.

Background



Image Classification



Object Detection



Semantic Segmentation

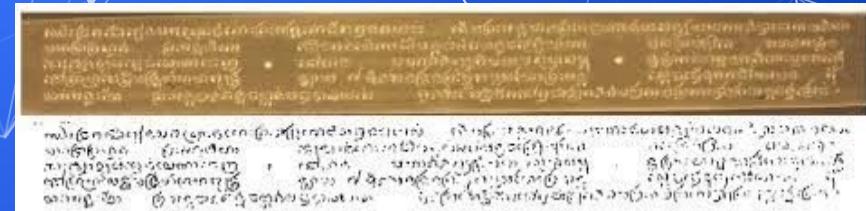
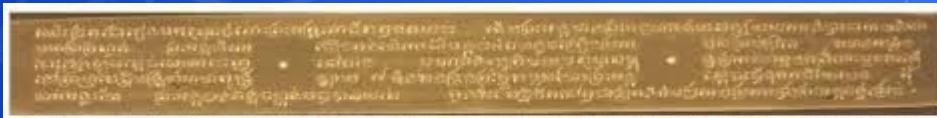


Instance Segmentation



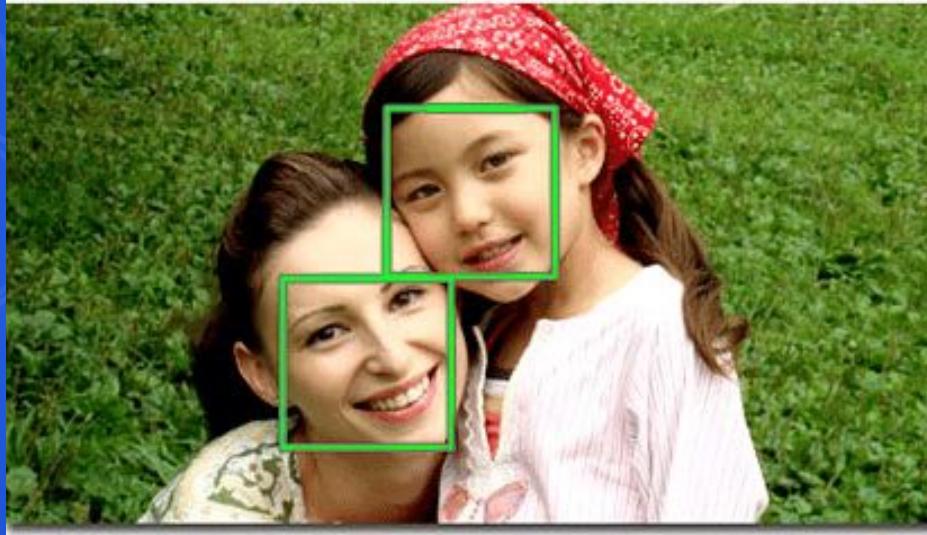
What is Computer Vision?: Applications

Optical Character Recognition



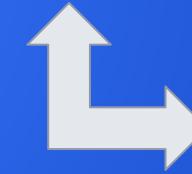
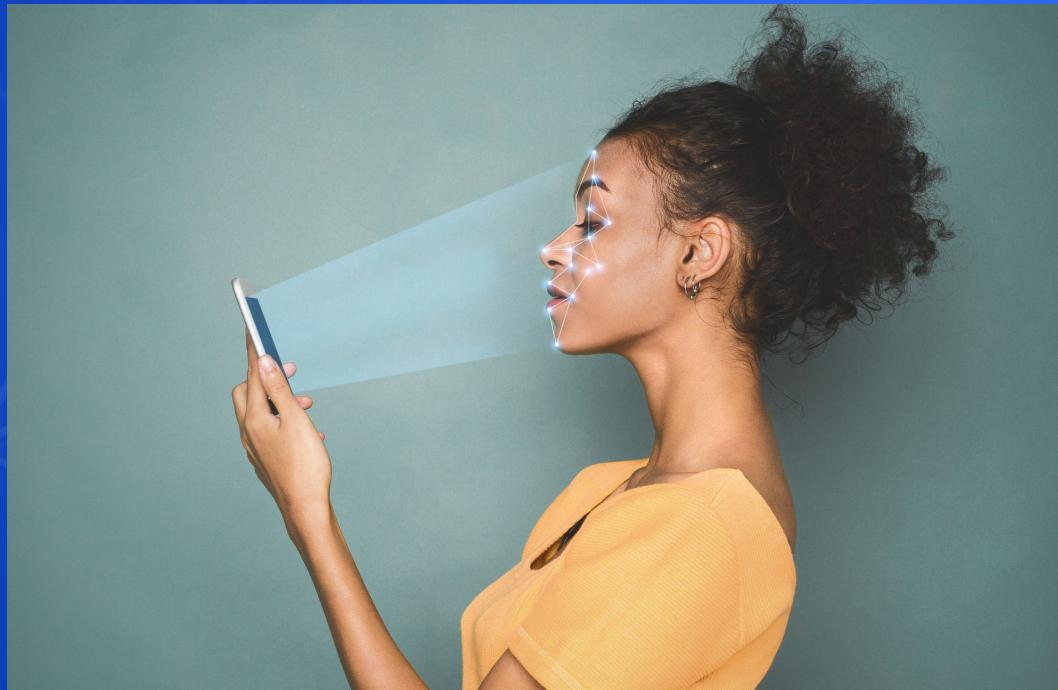
What is Computer Vision?: Applications

Face Detection/ Sony Smile Shutter



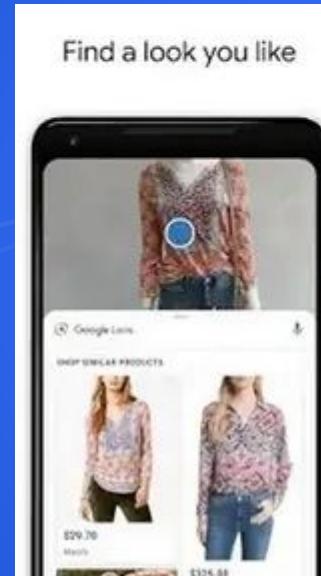
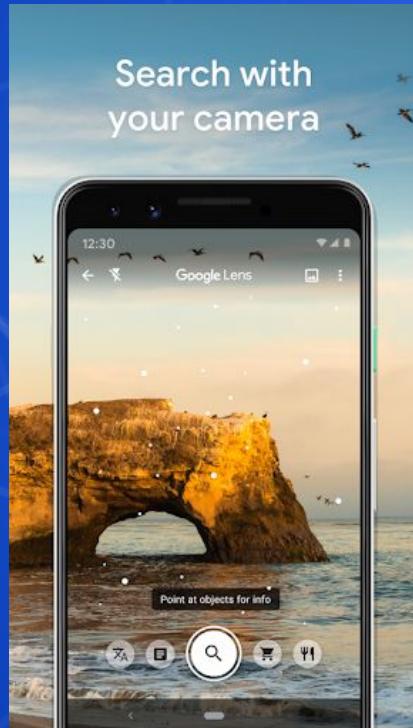
What is Computer Vision?: Applications

Facial Recognition/ ID Card Scanning



What is Computer Vision?: Applications

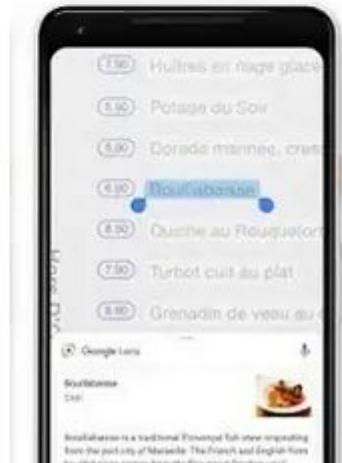
Google Lens



Identify plants
and animals

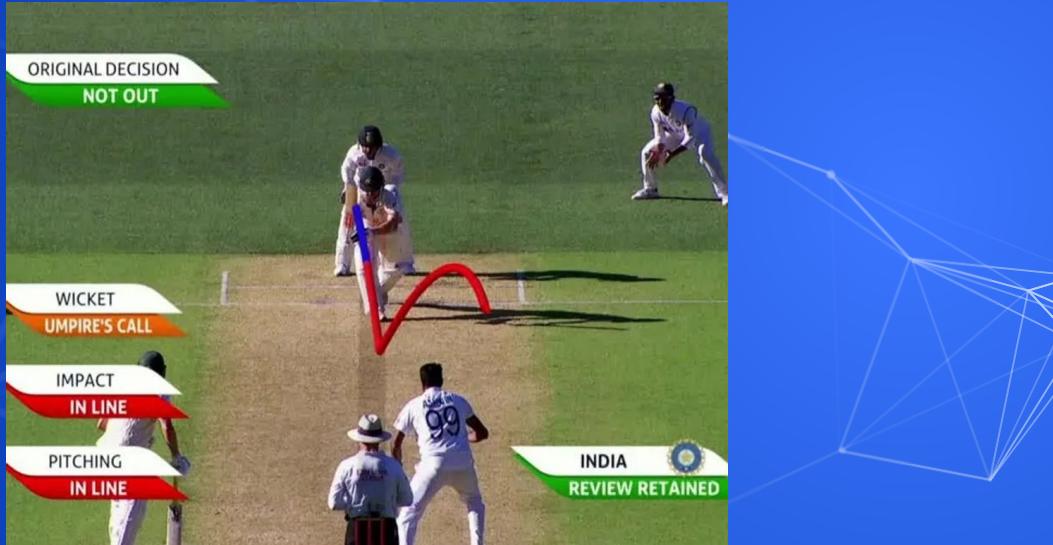


Take action on text



What is Computer Vision?: Applications

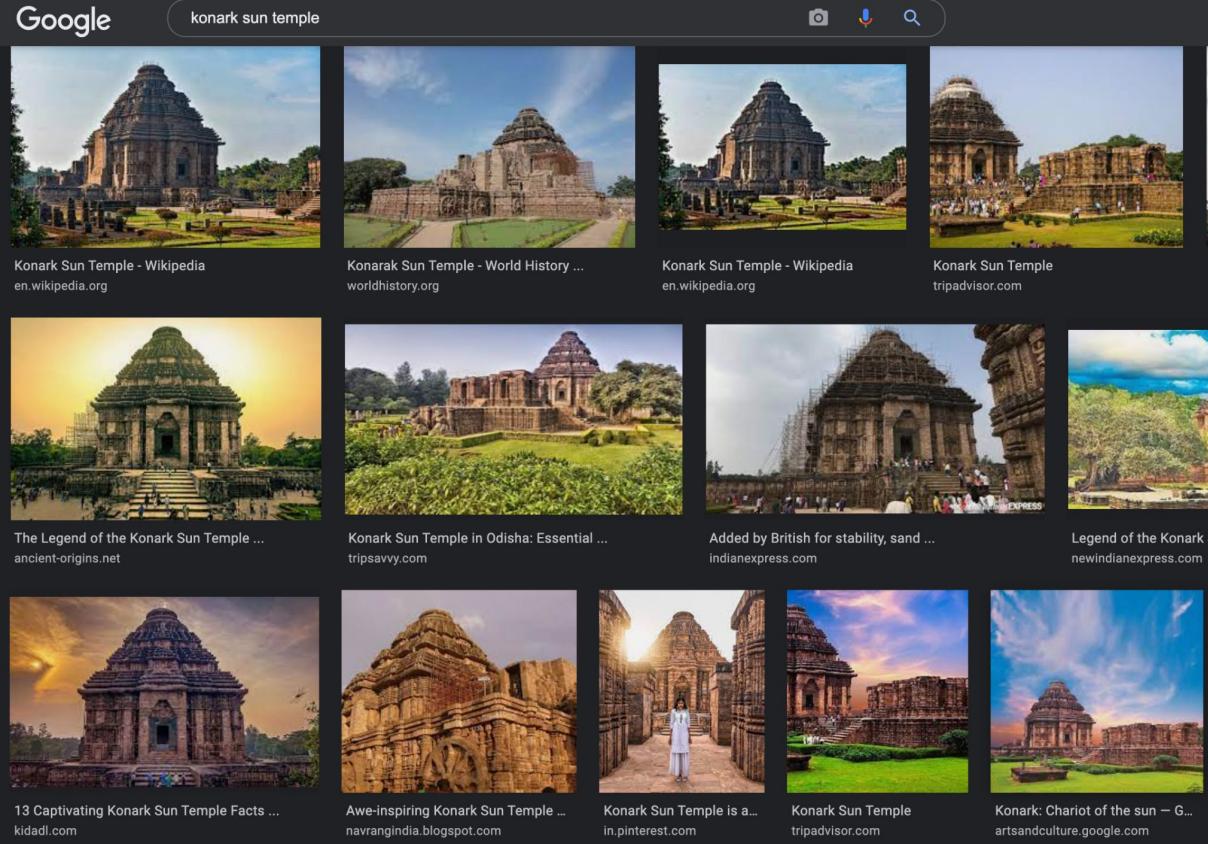
Sports Ball Tracking



What is Computer Vision?: Applications

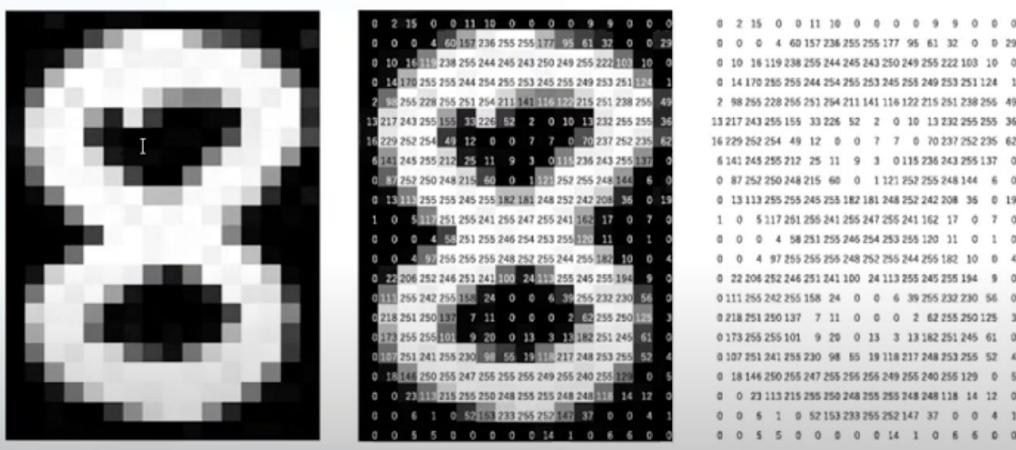
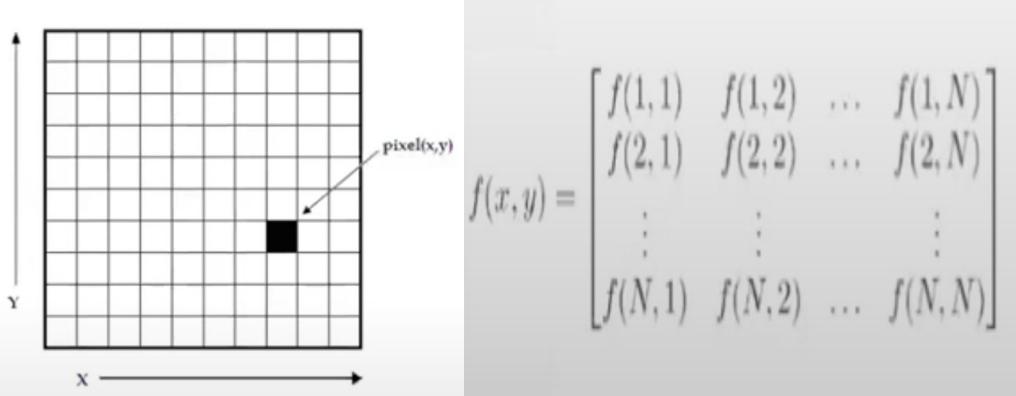
Medical Imaging

Why Computer Vision is Hard?



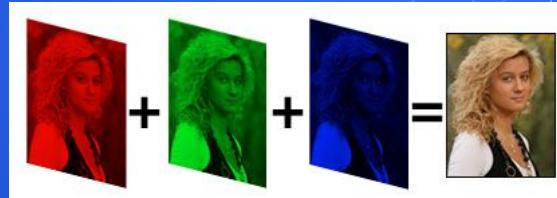
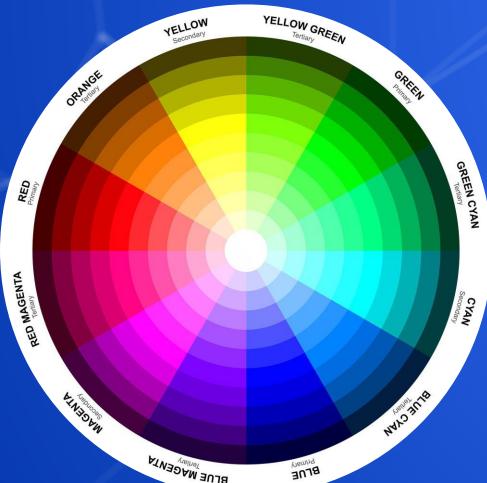
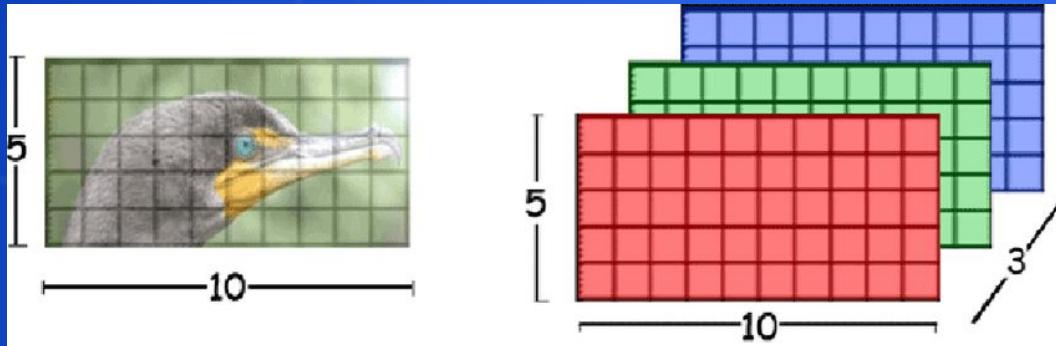
- Viewpoint
- Lighting
- Scale
- Camera Limitation
- Occlusion
- Obfuscation

What is an Image?: Grayscale



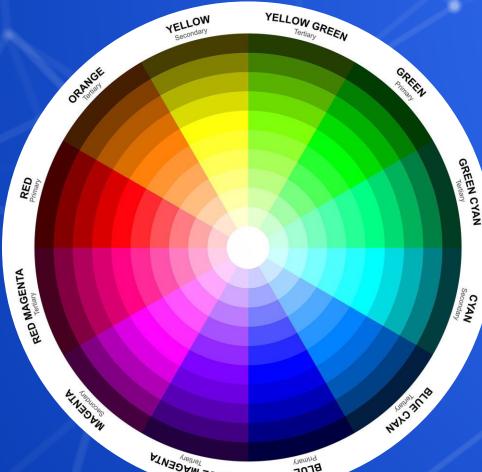
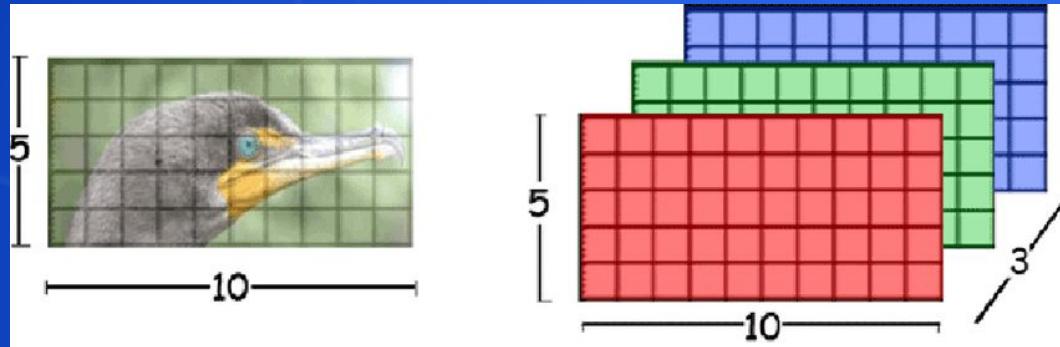
- Image is a combination of Pixels
- Pixels are spatial coordinate with a value
- The value is a number with range [0, 255]
- Computer representation is a matrix

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?: 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



LOW RESOLUTION



HIGH 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?



- Combination of picture frames stacked together
- FPS is the measure of captured frames
- Min 15 fps required to perceive motion
- More the frames slow motion effect

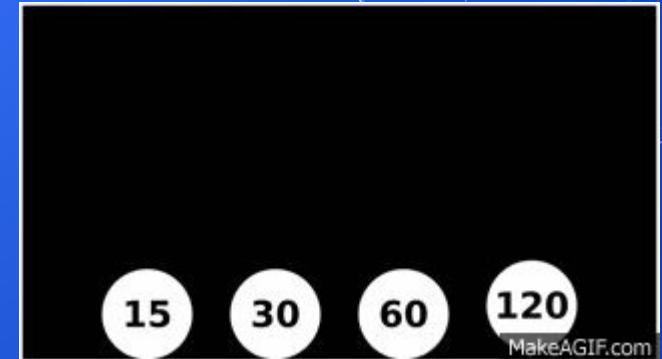
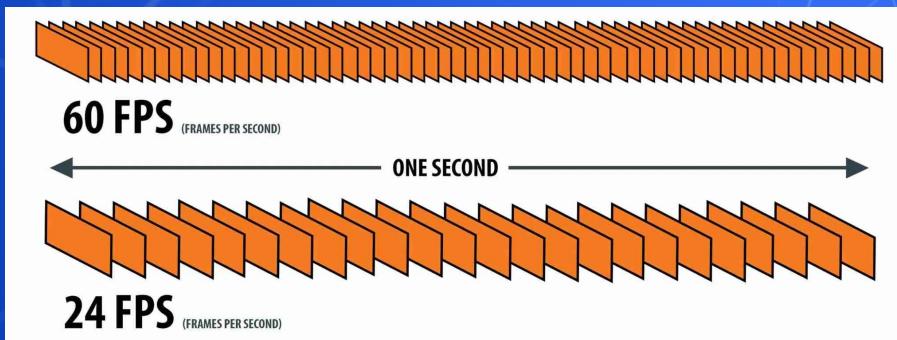
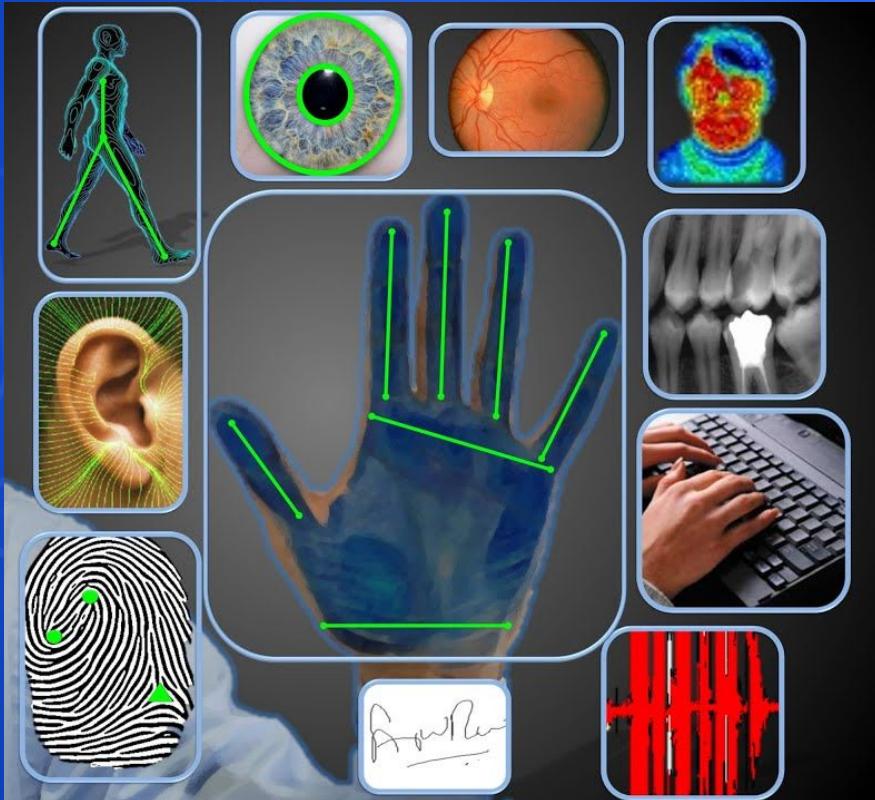


Image Processing



- 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



| | | | | | | | | | | | | | | | | |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|
| 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 | 96 | 61 | 32 | 0 | 0 | 29 | |
| 0 | 10 | 18 | 119 | 238 | 259 | 244 | 245 | 243 | 250 | 249 | 255 | 222 | 103 | 10 | 0 | |
| 0 | 14 | 170 | 255 | 255 | 244 | 254 | 255 | 253 | 245 | 258 | 249 | 253 | 251 | 124 | 1 | |
| 7 | 98 | 256 | 228 | 255 | 251 | 254 | 211 | 141 | 116 | 122 | 215 | 251 | 238 | 256 | 49 | |
| 13 | 217 | 243 | 255 | 155 | 33 | 228 | 52 | 2 | 0 | 10 | 13 | 232 | 255 | 255 | 36 | |
| 16 | 229 | 252 | 254 | 49 | 12 | 0 | 0 | 7 | 7 | 0 | 70 | 237 | 252 | 236 | 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 | 145 | 6 | 0 | |
| 0 | 0 | 13 | 113 | 255 | 255 | 245 | 255 | 182 | 181 | 248 | 252 | 242 | 208 | 36 | 0 | 19 |
| 1 | 0 | 5 | 117 | 251 | 255 | 241 | 265 | 247 | 255 | 241 | 162 | 17 | 0 | 7 | 0 | |
| 0 | 0 | 0 | 4 | 58 | 231 | 255 | 246 | 254 | 253 | 255 | 120 | 11 | 0 | 1 | 0 | |
| 0 | 0 | 0 | 4 | 97 | 255 | 255 | 255 | 248 | 252 | 255 | 244 | 255 | 182 | 10 | 0 | 4 |
| 0 | 0 | 22 | 206 | 252 | 246 | 251 | 241 | 100 | 24 | 113 | 255 | 245 | 255 | 194 | 9 | 0 |
| 0 | 0 | 111 | 255 | 242 | 255 | 158 | 24 | 0 | 0 | 6 | 39 | 255 | 232 | 230 | 56 | 0 |
| 0 | 0 | 218 | 251 | 250 | 137 | 7 | 11 | 0 | 0 | 2 | 62 | 255 | 250 | 125 | 3 | |
| 0 | 0 | 173 | 255 | 255 | 101 | 9 | 20 | 0 | 13 | 3 | 13 | 182 | 251 | 245 | 61 | |
| 0 | 0 | 107 | 251 | 241 | 258 | 230 | 98 | 55 | 19 | 118 | 217 | 248 | 253 | 255 | 52 | |
| 0 | 0 | 18 | 144 | 250 | 255 | 247 | 255 | 255 | 249 | 245 | 256 | 129 | 0 | 5 | 0 | |
| 0 | 0 | 0 | 23 | 113 | 215 | 255 | 250 | 248 | 255 | 248 | 248 | 118 | 14 | 12 | 0 | |
| 0 | 0 | 0 | 6 | 1 | 0 | 52 | 153 | 233 | 255 | 252 | 147 | 37 | 0 | 0 | 3 | |
| 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 1 | 0 | 6 | 6 | 0 | |

- It is the most widely used technique
- removal of unwanted outer areas
- Selection of a particular region within the image matrix for all the given channels

Image Processing: Cropping



| | | | | | | | | | | | | | | |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0 | 2 | 15 | 0 | 0 | 11 | 10 | 0 | 0 | 0 | 9 | 9 | 0 | 0 | 0 |
| 0 | 0 | 0 | 4 | 60 | 157 | 236 | 255 | 255 | 177 | 95 | 61 | 32 | 0 | 0 |
| 0 | 10 | 18 | 119 | 238 | 259 | 244 | 245 | 243 | 250 | 249 | 255 | 222 | 103 | 10 |
| 0 | 14 | 170 | 255 | 255 | 244 | 254 | 255 | 253 | 245 | 258 | 249 | 253 | 251 | 124 |
| 7 | 98 | 256 | 228 | 255 | 251 | 254 | 211 | 141 | 116 | 122 | 215 | 251 | 238 | 49 |
| 13 | 217 | 243 | 255 | 155 | 33 | 228 | 52 | 2 | 0 | 10 | 13 | 232 | 255 | 36 |
| 16 | 229 | 252 | 254 | 49 | 12 | 0 | 0 | 7 | 7 | 0 | 70 | 237 | 252 | 62 |
| 0 | 141 | 245 | 255 | 212 | 25 | 11 | 9 | 3 | 0 | 115 | 236 | 243 | 255 | 137 |
| 0 | 87 | 252 | 250 | 115 | 15 | 0 | 0 | 1 | 1 | 55 | 248 | 144 | 6 | 0 |
| 0 | 0 | 13 | 113 | 255 | 55 | 245 | 255 | 182 | 181 | 248 | 252 | 42 | 208 | 36 |
| 1 | 0 | 5 | 117 | 81 | 255 | 241 | 265 | 247 | 255 | 241 | 62 | 17 | 0 | 7 |
| 0 | 0 | 0 | 4 | 58 | 231 | 255 | 246 | 254 | 253 | 245 | 20 | 31 | 0 | 1 |
| 0 | 0 | 0 | 4 | 97 | 55 | 255 | 255 | 248 | 252 | 255 | 244 | 255 | 182 | 10 |
| 0 | 22 | 206 | 252 | 46 | 251 | 242 | 100 | 24 | 113 | 255 | 45 | 255 | 194 | 9 |
| 0 | 111 | 255 | 242 | 255 | 158 | 24 | 0 | 0 | 6 | 39 | 255 | 232 | 230 | 56 |
| 0 | 218 | 251 | 250 | 137 | 7 | 11 | 0 | 0 | 0 | 2 | 62 | 255 | 250 | 125 |
| 0 | 173 | 255 | 255 | 101 | 9 | 20 | 0 | 13 | 3 | 13 | 182 | 251 | 245 | 61 |
| 0 | 107 | 251 | 241 | 258 | 230 | 98 | 55 | 19 | 118 | 217 | 248 | 253 | 255 | 52 |
| 0 | 18 | 144 | 250 | 255 | 247 | 255 | 255 | 256 | 249 | 255 | 240 | 255 | 129 | 0 |
| 0 | 0 | 23 | 113 | 215 | 255 | 250 | 248 | 255 | 255 | 248 | 248 | 118 | 14 | 12 |
| 0 | 0 | 6 | 1 | 0 | 52 | 153 | 233 | 255 | 252 | 147 | 37 | 0 | 0 | 3 |
| 0 | 0 | 5 | 5 | 0 | 0 | 0 | 0 | 0 | 14 | 1 | 0 | 6 | 6 | 0 |

```
cropped_img = img[324:330,162:170]
```

- It is the most widely used technique
- removal of unwanted outer areas
- Selection of a particular region within the image matrix for all the given channels

Image Processing: Cropping



| | | | | | | | | | | | | | |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0 | 2 | 15 | 0 | 0 | 11 | 10 | 0 | 0 | 0 | 0 | 9 | 0 | 0 |
| 0 | 0 | 4 | 60 | 157 | 236 | 255 | 255 | 177 | 95 | 61 | 32 | 0 | 0 |
| 0 | 10 | 16 | 119 | 238 | 255 | 244 | 245 | 243 | 250 | 249 | 255 | 222 | 103 |
| 0 | 14 | 170 | 255 | 255 | 244 | 254 | 255 | 253 | 245 | 255 | 249 | 253 | 121 |
| 0 | 98 | 265 | 228 | 256 | 251 | 254 | 211 | 141 | 118 | 122 | 215 | 213 | 238 |
| 3 | 127 | 243 | 255 | 155 | 132 | 226 | 52 | 0 | 10 | 13 | 232 | 255 | 250 |
| 6 | 229 | 252 | 256 | 49 | 12 | 0 | 0 | 7 | 7 | 0 | 70 | 237 | 252 |
| 6 | 141 | 245 | 255 | 212 | 25 | 11 | 9 | 3 | 0 | 115 | 236 | 243 | 157 |
| 0 | 87 | 252 | 250 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 248 | 144 |
| 0 | 13 | 113 | 255 | 255 | 245 | 255 | 182 | 181 | 248 | 252 | 242 | 208 | 36 |
| 1 | 0 | 5 | 117 | 31 | 258 | 241 | 255 | 247 | 255 | 241 | 162 | 17 | 0 |
| 0 | 0 | 0 | 4 | 0 | 251 | 255 | 246 | 254 | 253 | 250 | 21 | 11 | 0 |
| 0 | 0 | 4 | 97 | 35 | 251 | 255 | 248 | 252 | 255 | 244 | 162 | 10 | 4 |
| 0 | 22 | 206 | 252 | 255 | 251 | 251 | 100 | 34 | 113 | 255 | 245 | 255 | 184 |
| 0 | 111 | 255 | 242 | 255 | 158 | 24 | 0 | 6 | 39 | 255 | 232 | 230 | 10 |
| 0 | 282 | 251 | 250 | 137 | 7 | 11 | 0 | 0 | 0 | 2 | 62 | 255 | 250 |
| 0 | 173 | 255 | 255 | 101 | 9 | 20 | 0 | 13 | 3 | 13 | 182 | 251 | 246 |
| 0 | 137 | 251 | 241 | 258 | 230 | 98 | 55 | 19 | 118 | 217 | 248 | 253 | 252 |
| 0 | 18 | 144 | 250 | 255 | 247 | 255 | 256 | 249 | 255 | 246 | 250 | 129 | 0 |
| 0 | 0 | 23 | 113 | 215 | 255 | 250 | 248 | 255 | 248 | 248 | 114 | 12 | 4 |
| 0 | 0 | 6 | 1 | 0 | 52 | 153 | 233 | 255 | 252 | 147 | 37 | 0 | 4 |
| 0 | 0 | 5 | 5 | 0 | 0 | 0 | 0 | 14 | 1 | 0 | 6 | 6 | 0 |

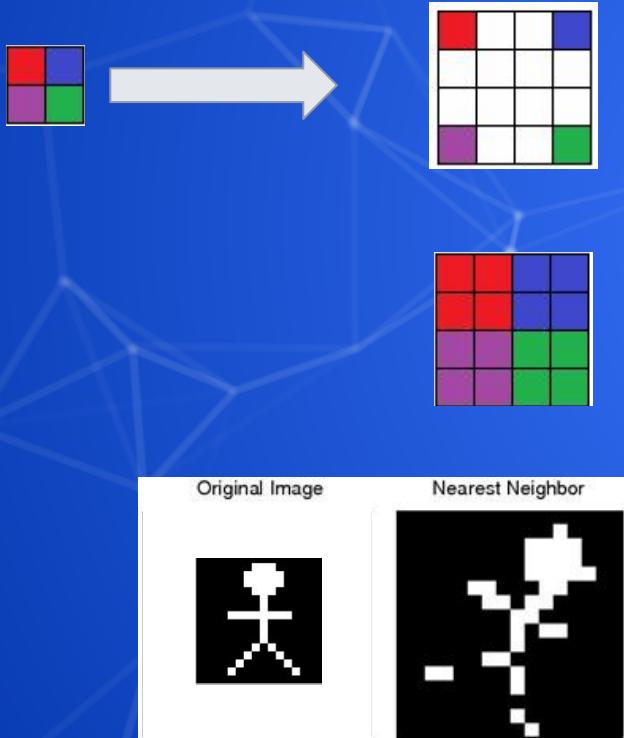


```
cropped_img = img[324:330,162:170]
```

6 255 245 255 182 181 248 25
7 261 255 241 255 247 255 24
4 58 251 255 246 254 253 25
7 255 255 255 248 252 255 24
2 246 251 241 100 24 113 25

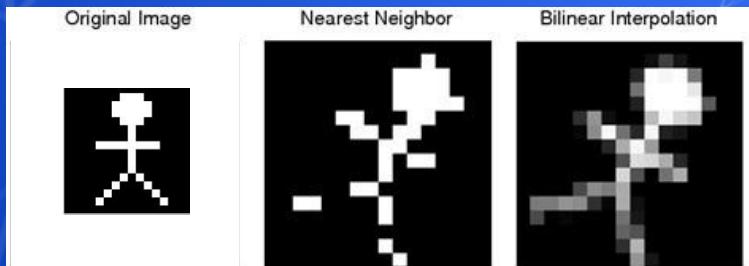
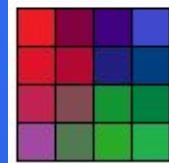
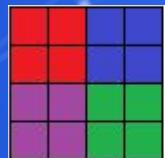
- It is the most widely used technique
 - removal of unwanted outer areas
 - Selection of a particular region within the image matrix for all the given channels

Image Processing: Resizing



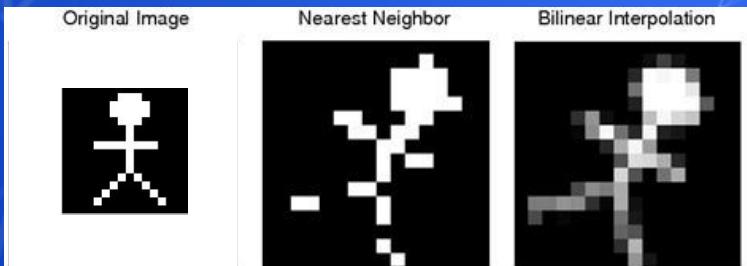
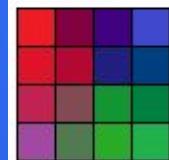
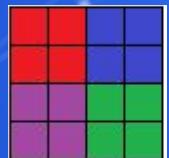
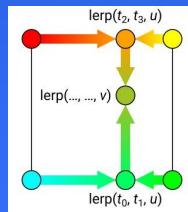
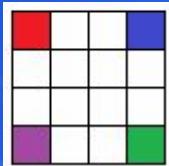
- The process of increasing or decreasing the size of an image
- Leads to loss of information
- Requires interpolation algorithms to compensate loss
- Nearest Neighbor, Bilinear
- Different from Padding

Image Processing: Resizing



- The process of increasing or decreasing the size of an image
- Leads to loss of information
- Requires interpolation algorithms to compensate loss
- Nearest Neighbor, Bilinear
- Different from Padding

Image Processing: Resizing



- The process of increasing or decreasing the size of an image
- Leads to loss of information
- Requires interpolation algorithms to compensate loss
- Nearest Neighbor, Bilinear
- Different from Padding

Image Processing: Flipping/Rotation

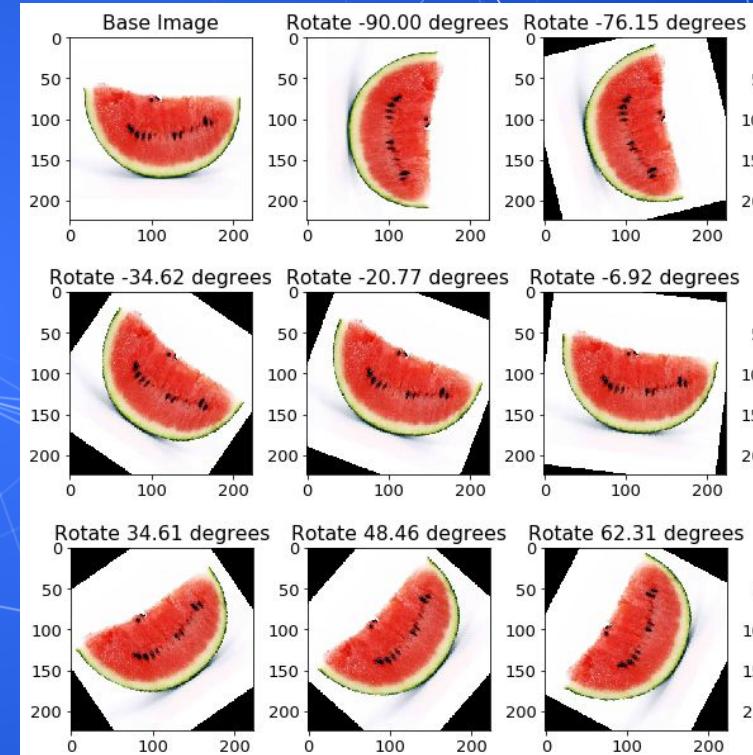
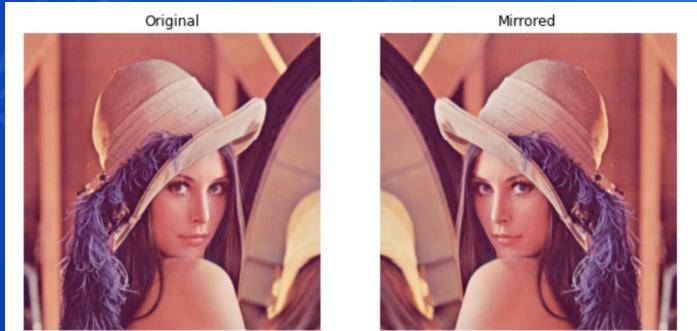
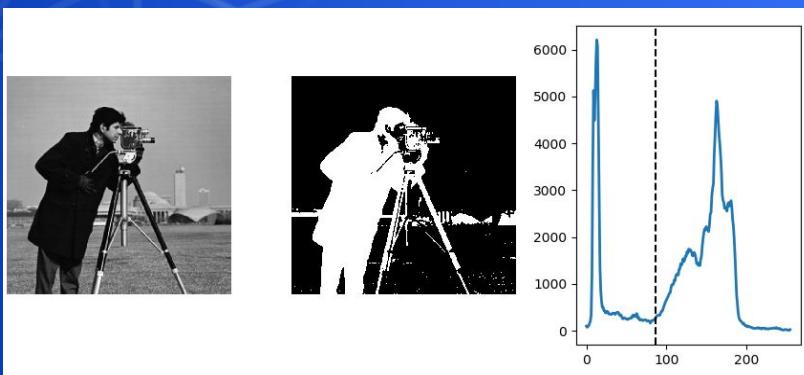
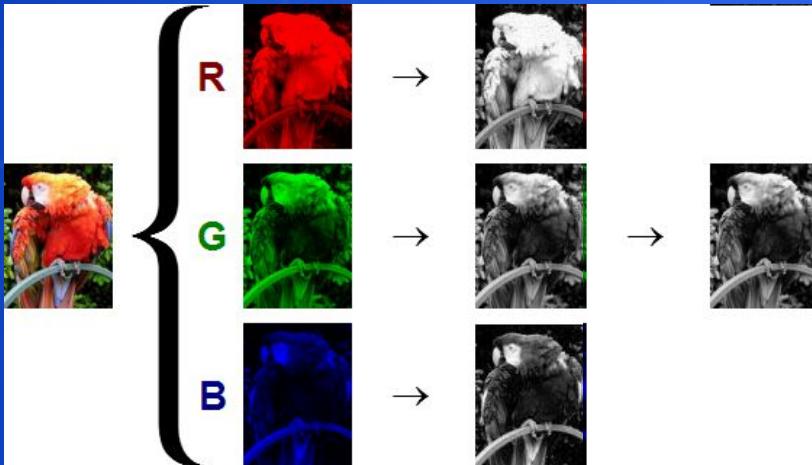
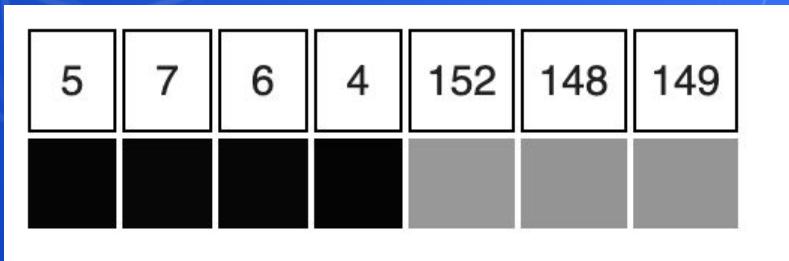
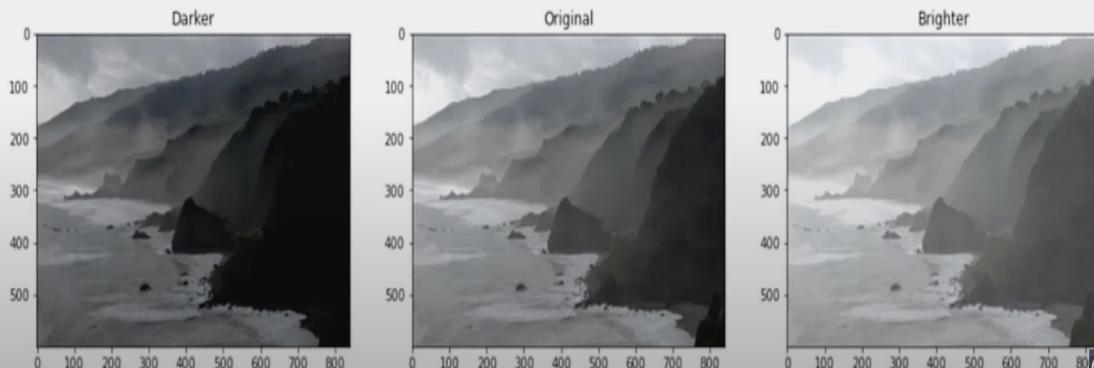


Image Processing: Color2Gray/Threshold



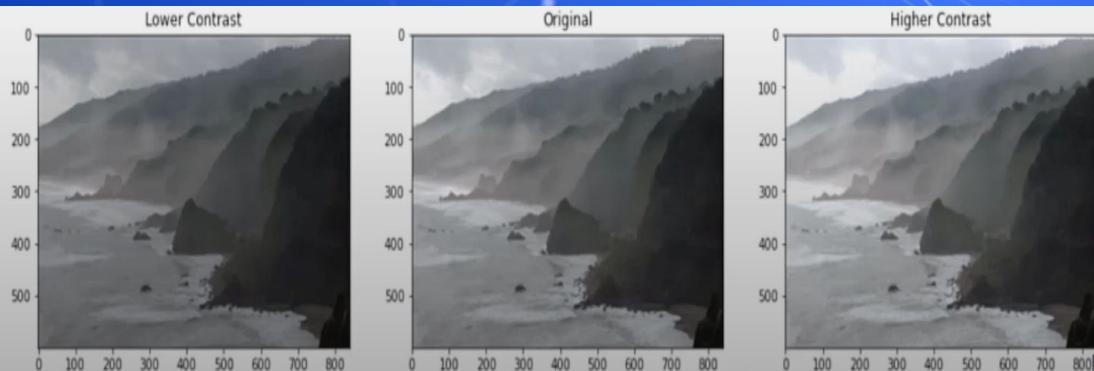
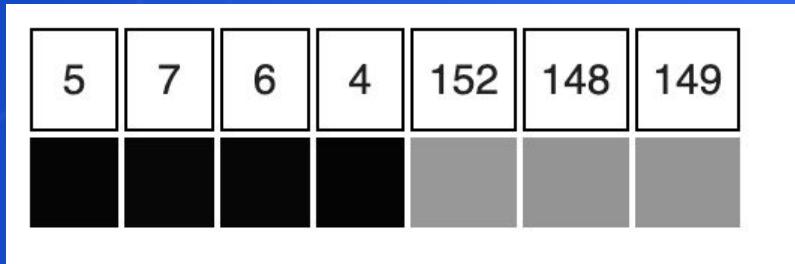
- Most of the common tasks don't require color
- Reduce image size by grayscaling
- Weighted sum is the most common technique
- Thresholding is converting gray images to binary
- Simplest way to remove background

Image Processing: Brightness/Contrast



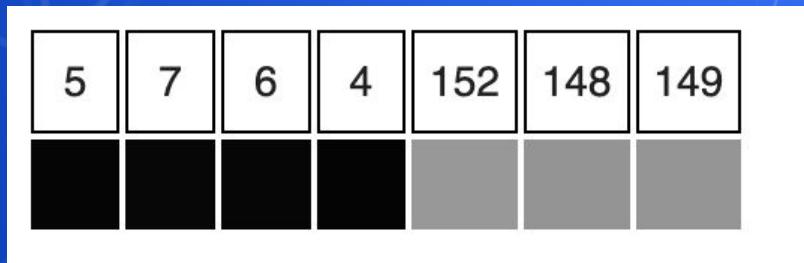
- Most easiest image enhancement technique
- Brightness- how much light comes from a object addition/subtraction of a fixed value
- Contrast- differentiate bright and dark bodies; define boundary
- mult/div of a fixed value

Image Processing: Brightness/Contrast



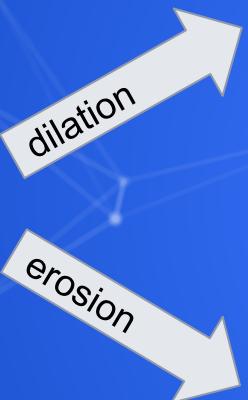
- Most easiest image enhancement technique
- Brightness- how much light comes from a object
- addition/subtraction of a fixed value
- Contrast- differentiate bright and dark bodies; define boundary
- mult/div of a fixed value

Image Processing: Edge Detection



- Most basic feature identification technique
- identifying edges, curves where image brightness changes sharply
Mathematically called discontinuity(abrupt change)
- Identifiable with very high gradient value

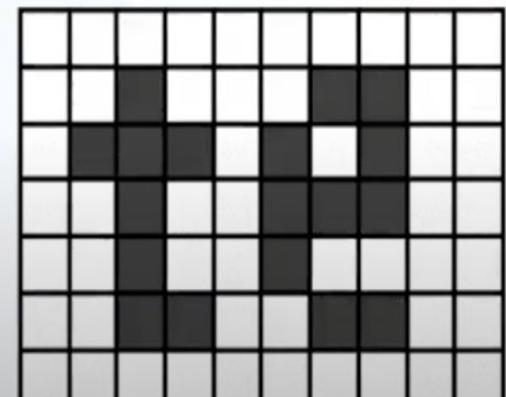
Image Processing: Dilation/Erosion



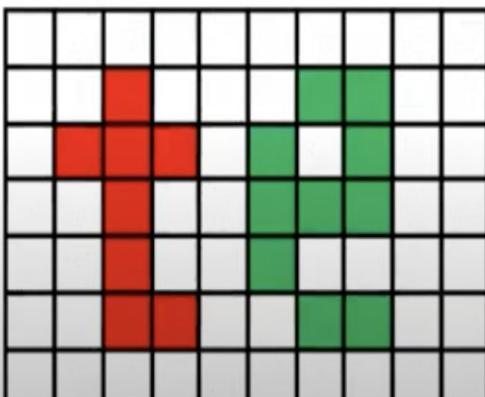
- Morphological operations to reduce noise
- Dilation - increase the size or thickness of the foreground object
Helps connect the disjoint region
- Erosion- decrease the thickness of the object
- Helps disconnect joined region

Image Processing: Connected Components

- Technique to identify pixel groups by neighbourhood connectivity.
- Detect and select objects with irregular shapes and sizes



a an input grayscale image



b the output: two connected-components

Image Processing: HDR

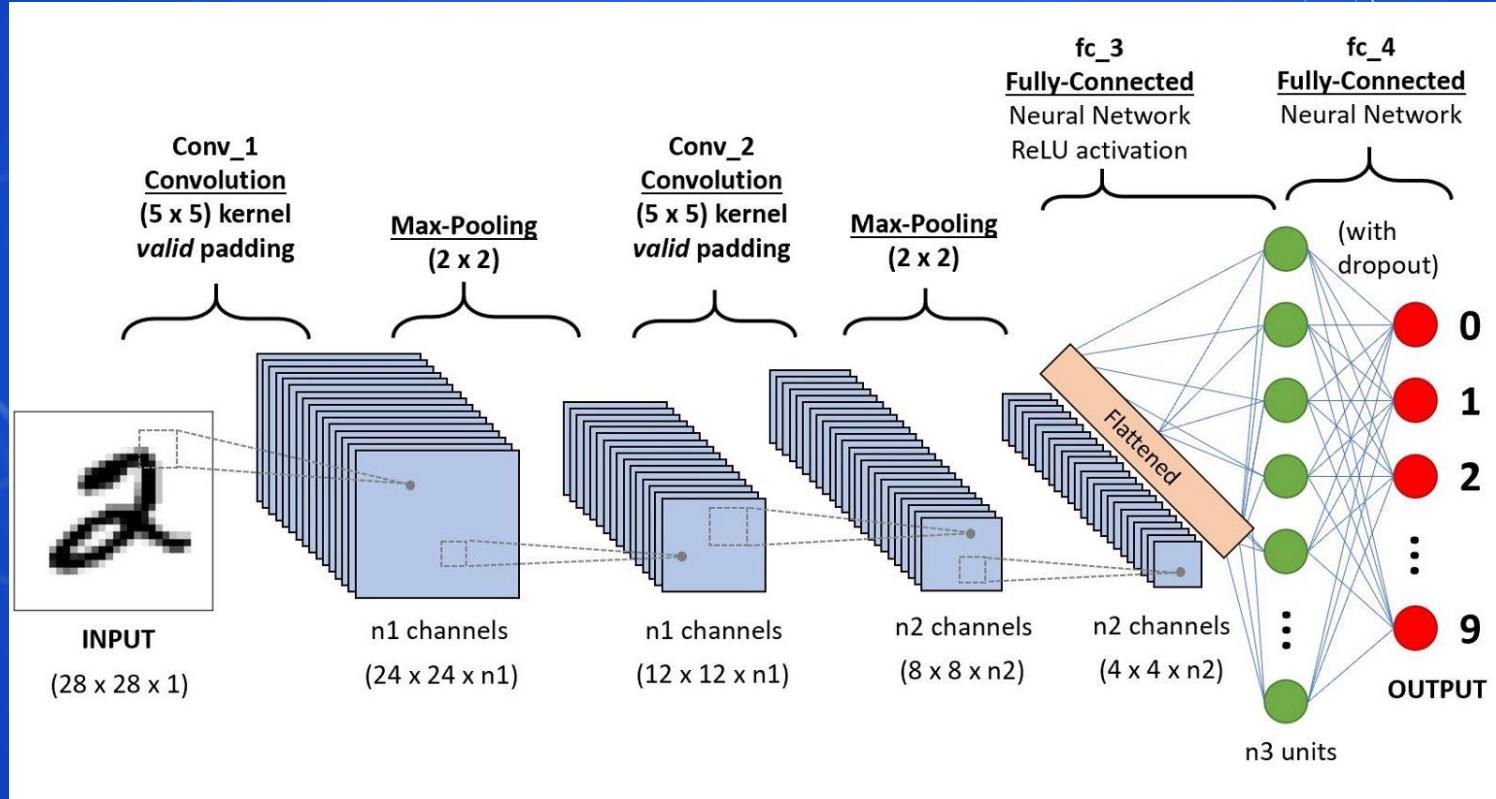


- lot of variation in light levels within a scene or an image
- capturing multiple frames of the same scene but with different exposures
- Gamma correction is the simplest combination technique

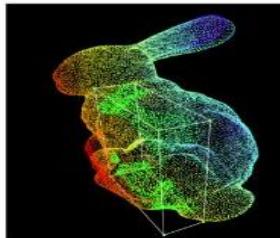
Demo: Traffic Management System

<https://github.com/odisha-ml/SummerSchool2022>

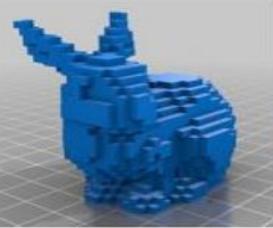
Outlook: CNNs



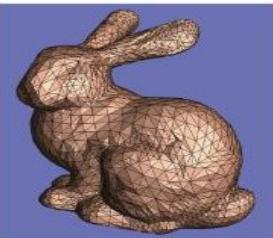
Outlook: 3D Computer Vision



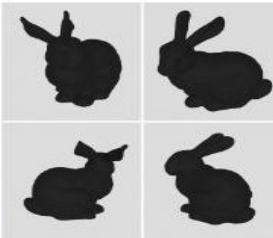
Pointclouds



Voxel grids



Meshes



Multiple views

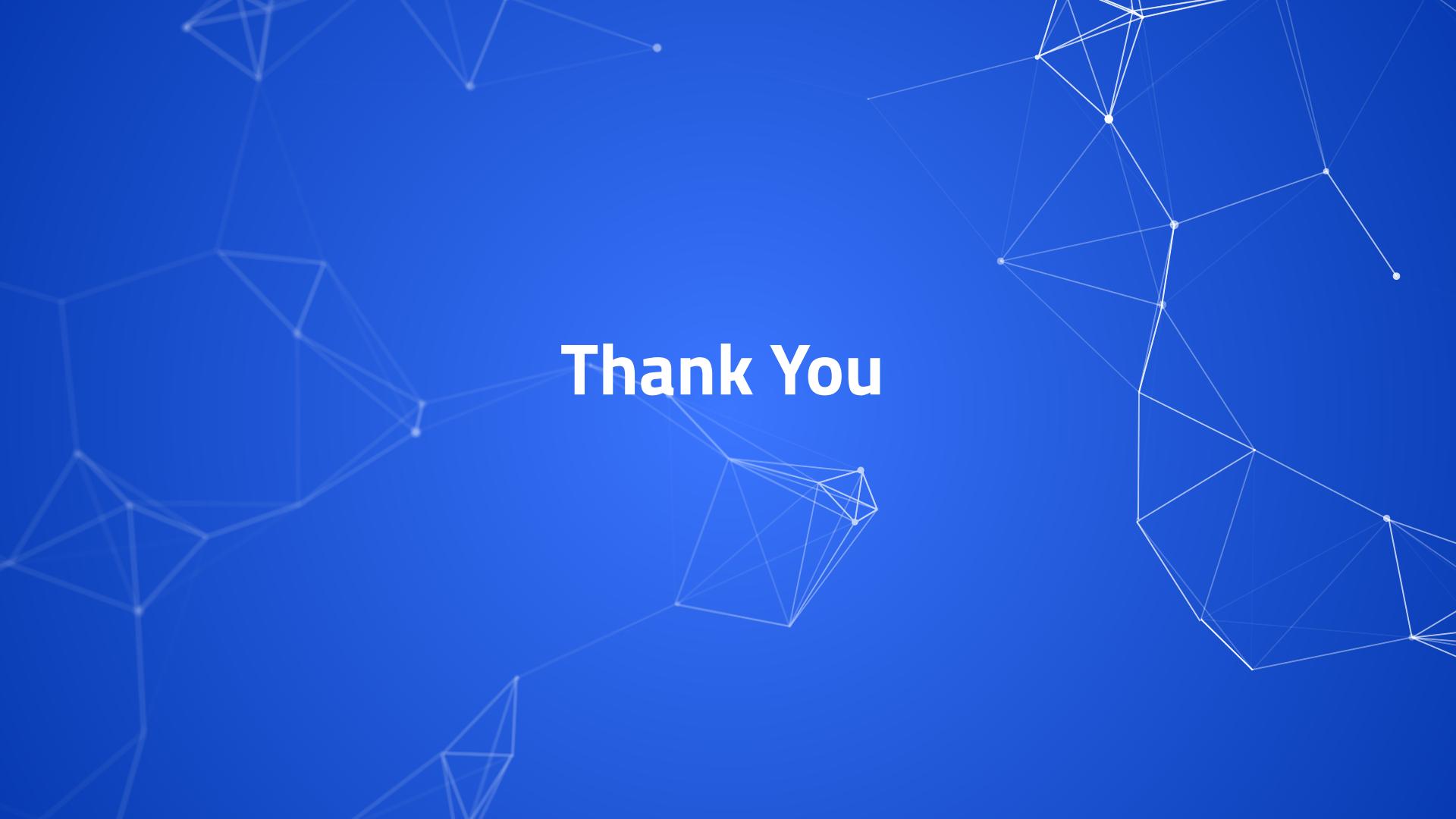


Depth maps



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



Thank You