

AI Workshop – Images and Machine Learning

Vijay Kalivarapu | Sr. AI Engineer | Pella Corporation

June 6, 2025



Who am I

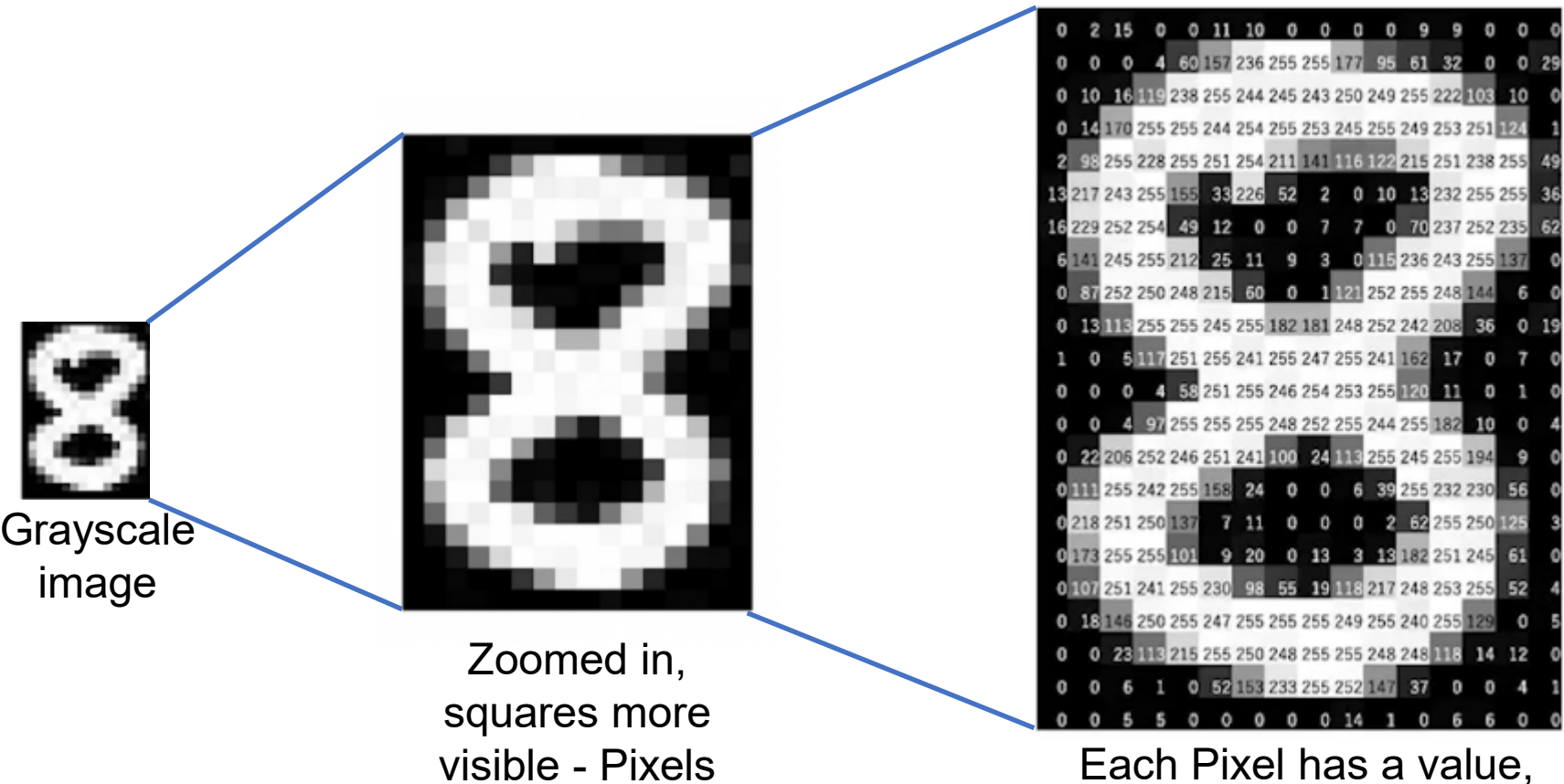
- Ph.D. in Mechanical Engineering, Human-Computer Interaction | Iowa State University
 - Sr. AI Engr., Pella Corporation
 - Teaching Professor, ISU
 - 3D Visualization, VR/AR, Design Optimization, Simulations, CV/ML
- 



Overview

- Image Representations
 - Convolving images
 - Machine Learning (ML)
 - Tasks in Computer Vision and ML
 - ML with Roboflow
 - STEM career discussion
- 

Image Representations - Grayscale



```
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
```

Pixel values are stored in the computer as a matrix - Channel

Image Representations - Grayscale

- Pixel intensity scale 0 – 255
 - 0 is black, 255 is white
 - $2^8 = 256$ colors
 - 8-bit or bit-depth = 8

Image Representations - Colors

- Colors generated from three primary colors Red, Green, Blue
 - 3 Channels



Colour Image

=



Red

+



Green

+



Blue

- Stored in three matrices, one for each primary color

Image Representations – RGB Colors

- 3-Channel colors stored in three matrices, one for each primary color R, G, and B
- 8-bit for each of R, G, B results in $8 + 8 + 8 = 24$ -bit color (e.g., .jpg)
- Each channel has 0 – 255 pixel intensity values
 - For R channel, 0 is no red and 255 is all red
- $256 \times 256 \times 256 = 16,777,216$ million colors!
- Computer reads color images as $N \times M \times 3$, i.e., width x height x # channels

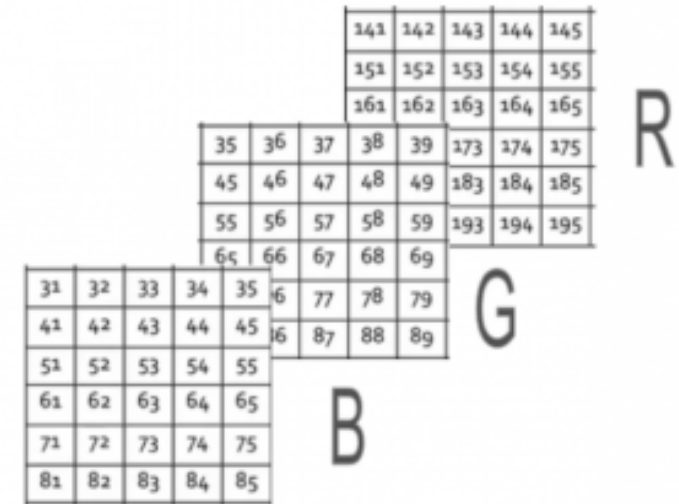


Image Representations – RGBA colors

- 4-Channel colors stored in four matrices, one for each primary color R, G, B and A
 - A is alpha channel
 - A goes from 0 – 255
 - $A = 0 \Rightarrow$ fully transparent
 - $A = 255 \Rightarrow$ fully opaque
- 8-bit for each of R, G, B, A results in $8 + 8 + 8 + 8 = 32$ -bit color (e.g., png, webp, tiff, tga)

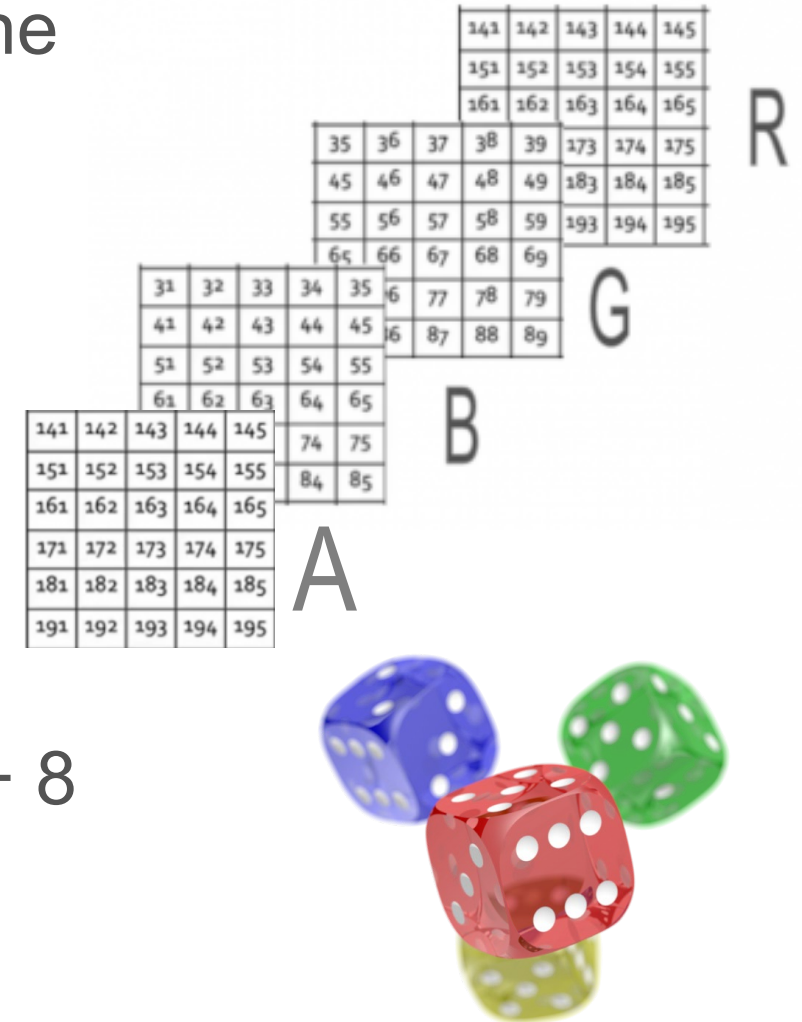


Image Representations

- Check bit-depth in Windows Explorer
 - Grayscale (Convert png to grayscale in GIMP)
 - RGB (Convert png to jpg in Paint)
- RGBA



Grayscale (8-bit)



RGB (24-bit)



RGBA (32-bit)

Image Convolutions

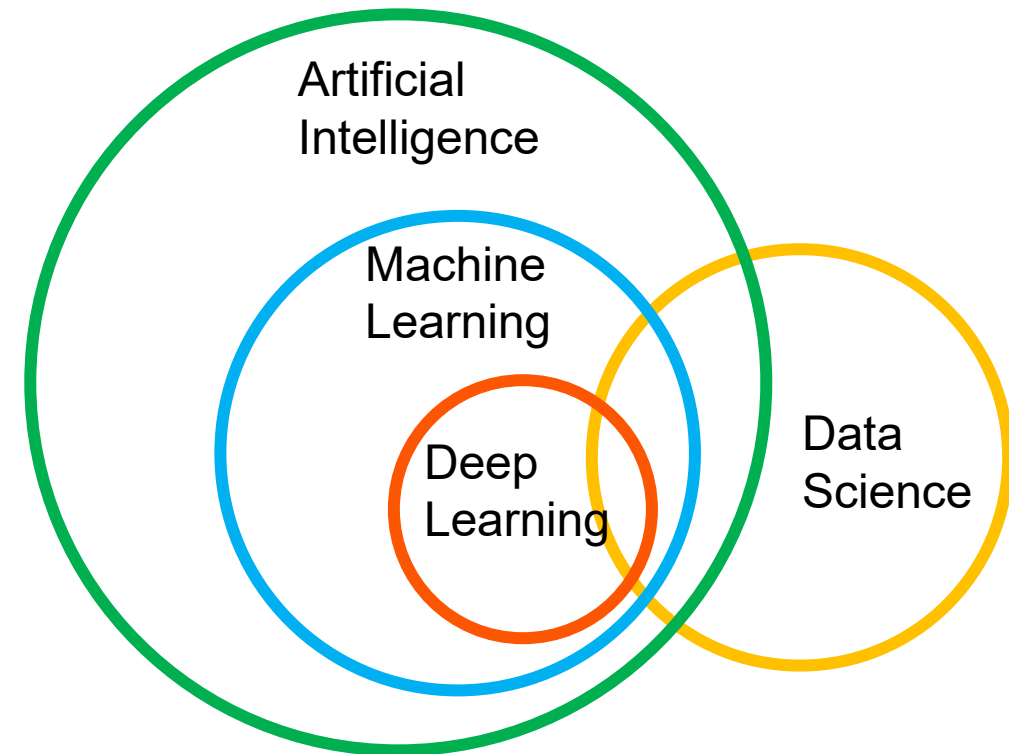
- Process of applying changes to images
 - Blurring, sharpening, edge detections, embossing, etc.
- Multiply a stride of pixels with a 3x3 matrix (kernel) to each color channel
 - [https://en.wikipedia.org/wiki/Kernel_\(image_processing\)](https://en.wikipedia.org/wiki/Kernel_(image_processing))
 - Fixed kernel for a specific purpose
- How does it work?
 - <https://setosa.io/ev/image-kernels/>
- Why do we need edge detections?

Machine Learning

- What is Machine Learning?
 - Teaching computers to learn by example
- How do computers 'see' images?
 - Using pixel values as numerical data
- Predictions in images
 - Classifying objects (e.g., is this a cat or a dog?)
 - Detecting objects (e.g., identify cars, vegetation, people, etc.)
- General pipeline
 - Input >> Extract Features >> Learning Model >> Predictions

Machine Learning Vs Artificial Intelligence

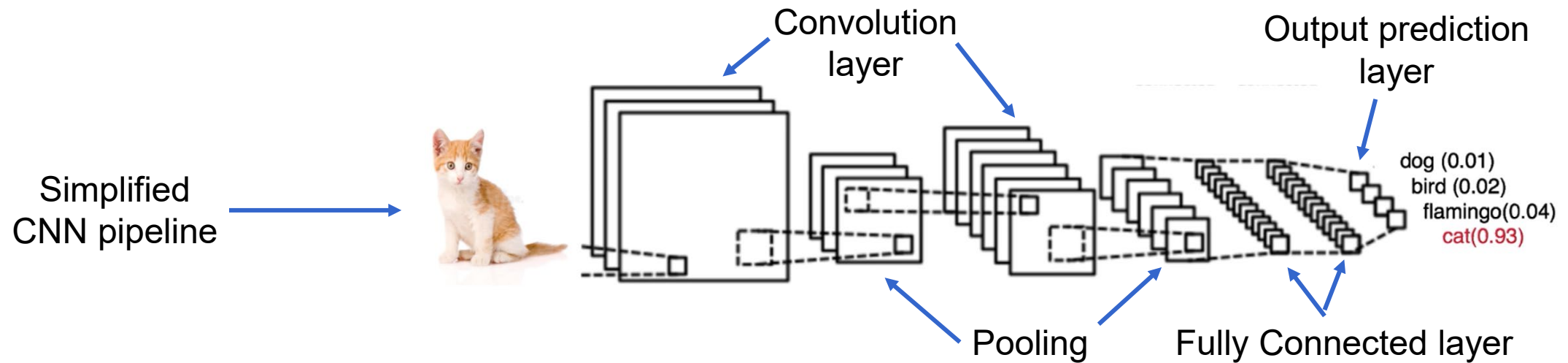
- AI can sense, reason, act, and adapt
 - Self-driving cars, virtual assistants, etc.
- ML is a subset of AI that learns from data without explicit programming
 - Fraud detection, customer segmentation, image recognition



- ML is like teaching a student step-by-step how to solve a problem
- DL is like letting the student explore many examples of math problems and figure out their own strategy to solve them

Machine Learning

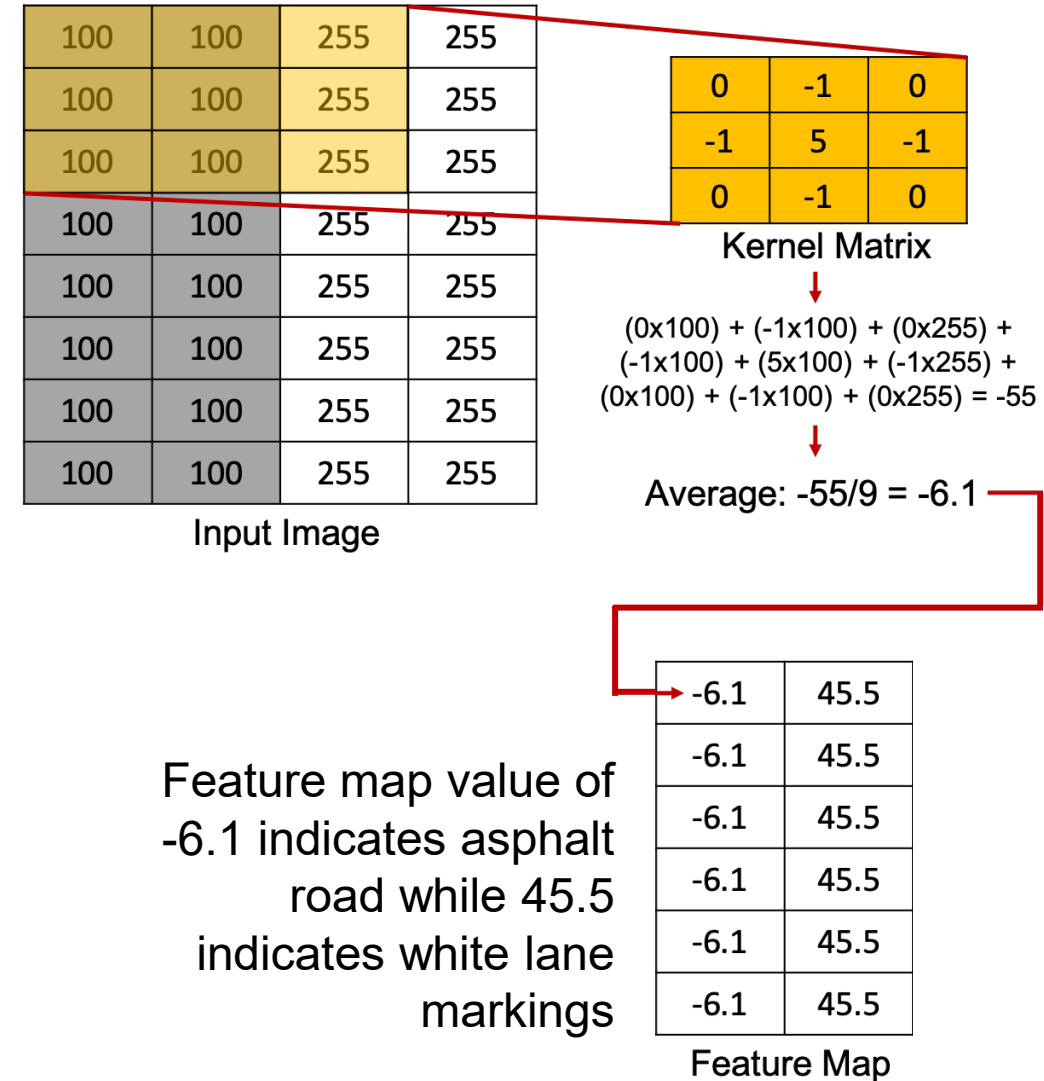
- Deep Learning with Convolutional Neural Networks
 - Highly effective for image recognition and classification



- Feature extraction through convolutions

Machine Learning – Feature Extraction

- Convolution layer for feature extraction
 - A kernel Matrix processes each input image pixel, starts randomly but automatically learns
 - Feature map computed via element wise matrix multiplication
- Pooling layer reduces the dimensionality of each feature
- Several convolutions + pooling layers extracts numerical information from an input image



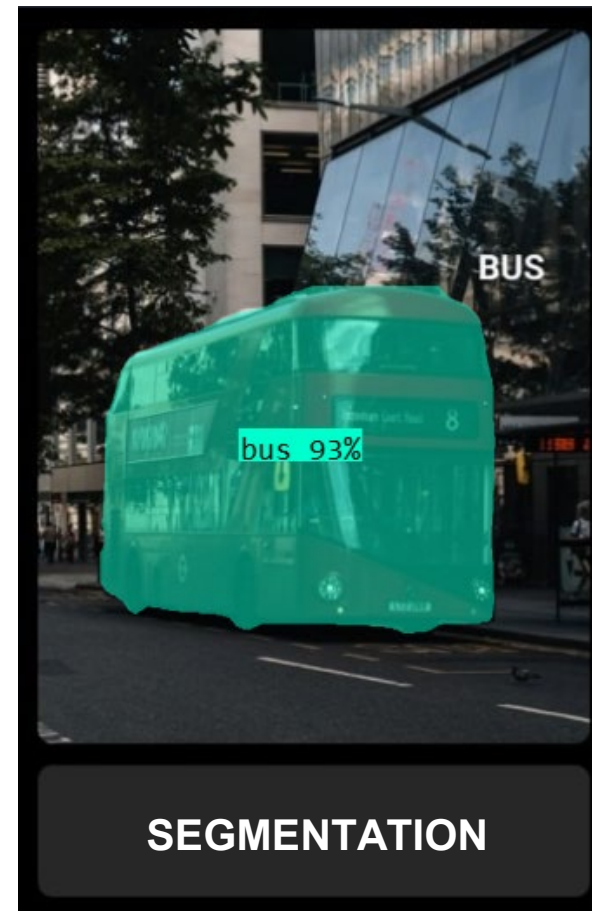
Machine Learning – Tasks



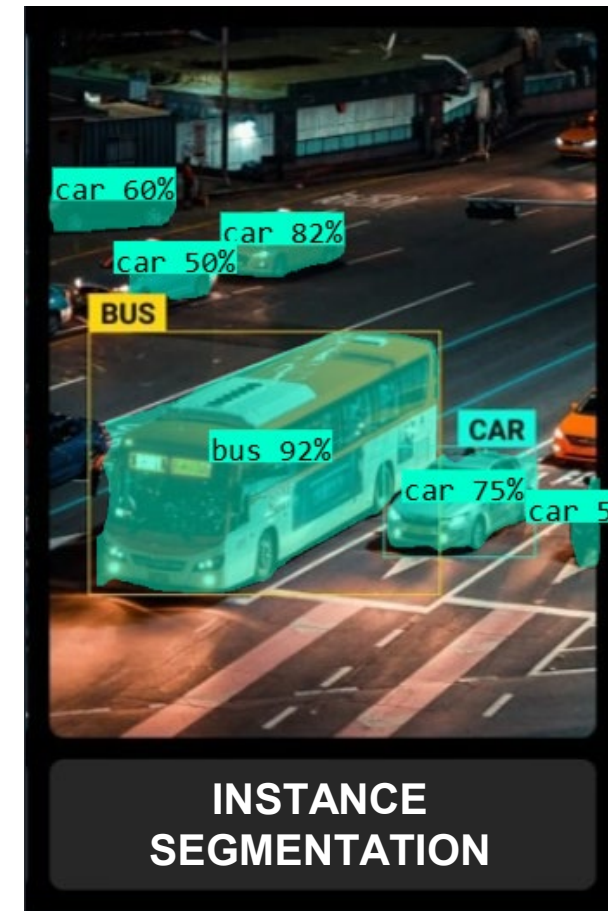
Categorizing images



Identifying and locating objects



Localization + Tight boundary around objects



Individually locate and identify different objects

Machine Learning – Segmentation

- Semantic: Links each pixel in the image to a class/label
- Instance: Identifies and separates each class in an image
- Panoptic: Combines semantic and instance



(a) Image



(b) Semantic segmentation



(c) Instance segmentation



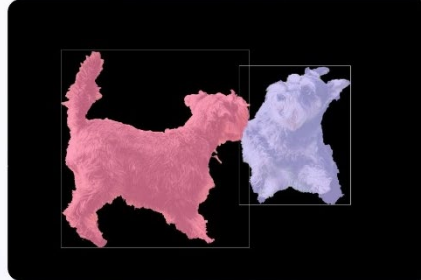
(d) Panoptic segmentation



(a) Image



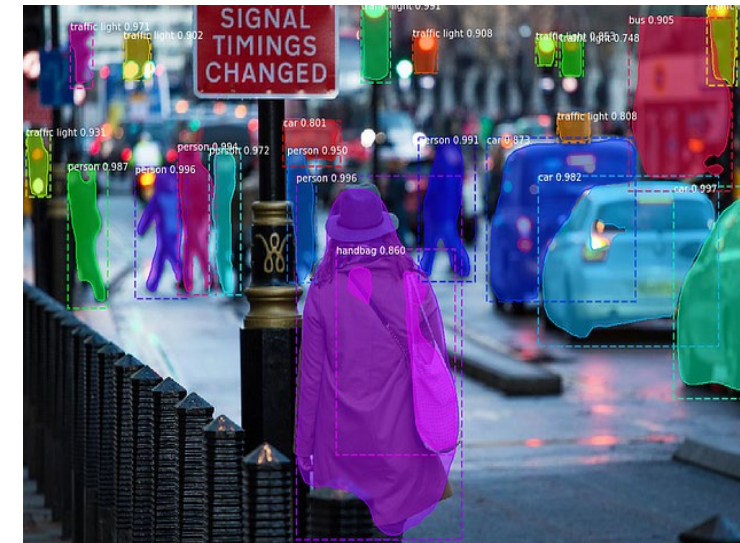
(b) Semantic Segmentation



(c) Instance Segmentation



(d) Panoptic Segmentation



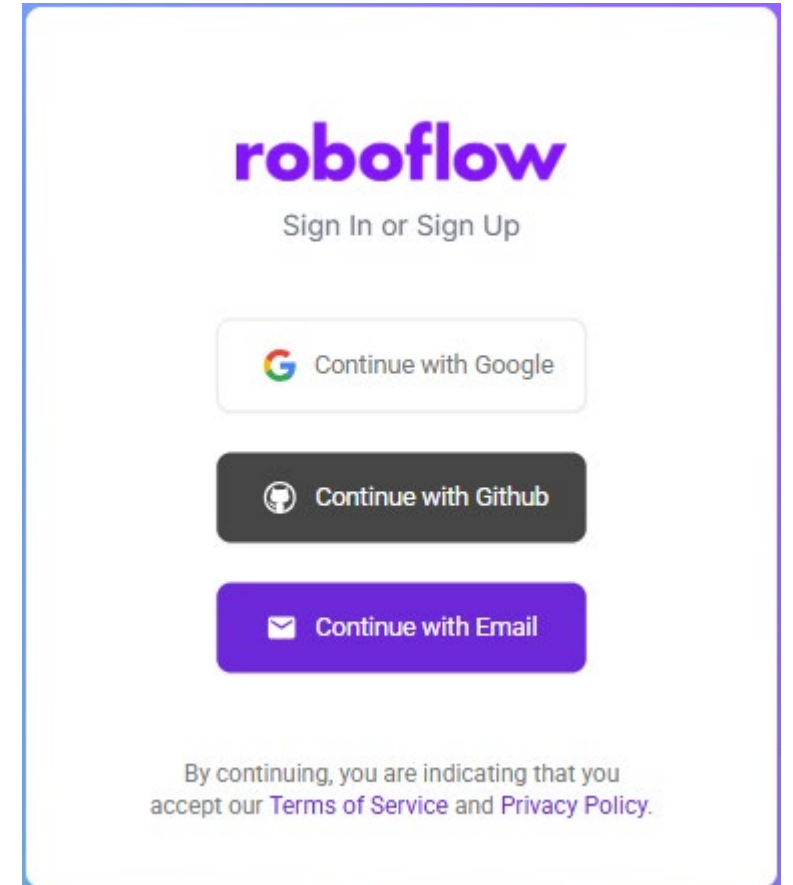
Instance segmentation

Instance Segmentation - Complexities

- Requires pixel-perfect training datasets
 - Poor labeling >> unreliable training >> inaccurate results
- Overlapping objects confuse models (e.g., multiple cars in a traffic scene)
- Computationally very intense (e.g., Mask R-CNN) for training and inference
- Models trained on certain datasets cannot be generalized to other domains
 - Retraining or fine-tuning will be required (time and cost)

Machine Learning with Roboflow

- Create account on Roboflow's website
 - <https://app.roboflow.com>



Machine Learning with Roboflow

- Roboflow Universe | Try
 - Classification project
 - Red and Green apples (<https://universe.roboflow.com/apples-1lrlt/red-and-green-apples-single-classification>)
 - Object detection project
 - Cats and dogs (<https://universe.roboflow.com/project/cat-and-dog-ekgyb>)
 - Road signs in real-time (<https://universe.roboflow.com/fsevval/traffic-sign-guy19>)
 - Instance segmentation
 - Train detection (<https://universe.roboflow.com/roboflow-universe-projects/train-detection-4ud7c>)

Machine Learning with Roboflow

- Let's create an ML model
 - Download the dataset from here:
 - <https://universe.roboflow.com/kidsplayground/2d-geometry-figures>
 - Try in Browser >> Dataset (left side) >> Download Dataset Button >> Download dataset (zip file) >> COCO format >> Download zip to computer
 - Google search for 'simple 2d shapes'
 - Download 4-5 of them
 - You will use them for labeling to build your ML model