CV / VLM

Unit 1: Introduction to Computer Vision (CV)

1.1.1

Basics of Computer Vision

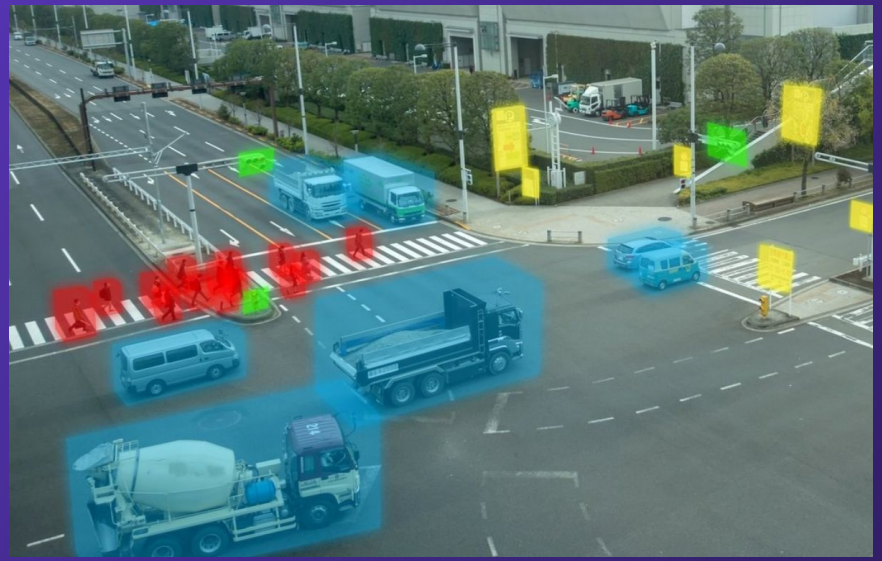
What is computer vision?

TIL-AI
TODAY I LEARNED-AI

What is Computer Vision?

- A branch of AI that teaches computers to see and understand visual data in a way that mimics human vision

- Utilizes algorithms to analyse and interpret visual inputs

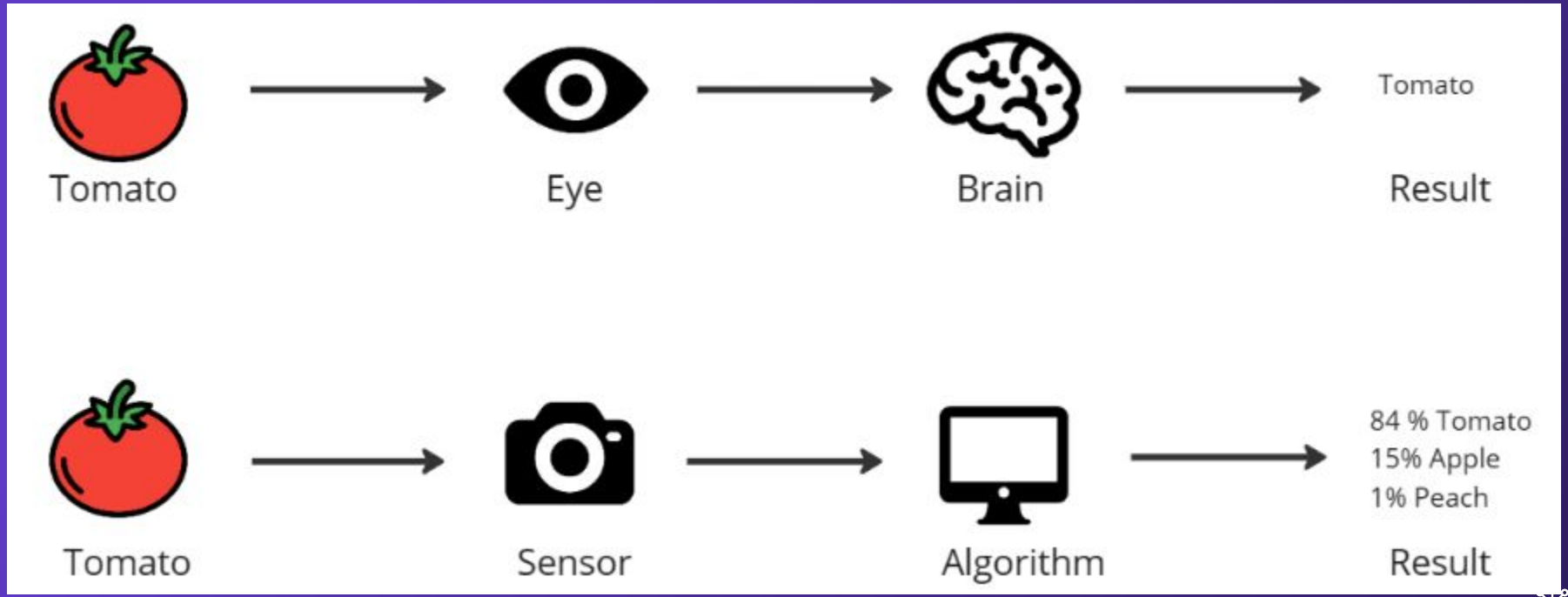


Adapted from viso.ai

TIL-AI
TODAY I LEARNED AI

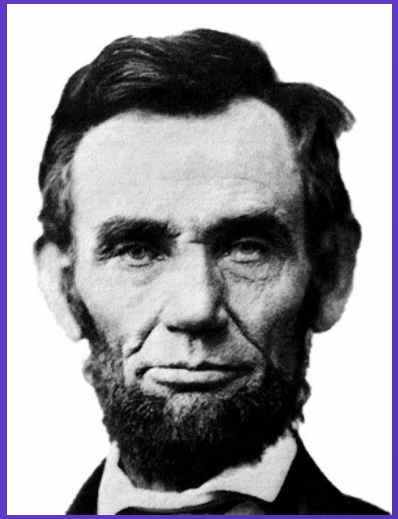
Analog Comparison

Human Vision vs Computer Vision



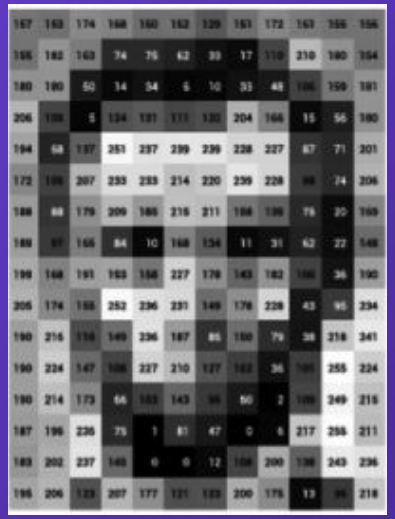
What Computers See - Pixel Values

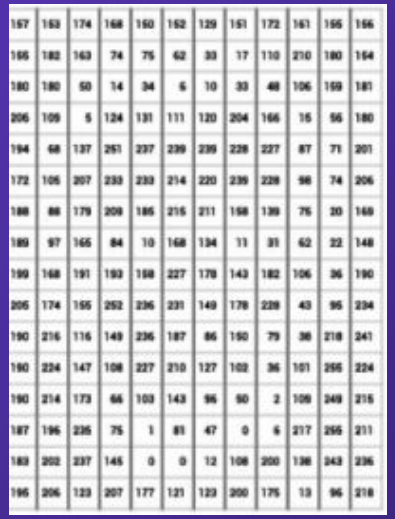
What we see



Grayscale
(Black and White)

What computers see





Array of Pixel values

Colored Images could be coded in RGB (Red, Green, Blue), HSL (Hue, Saturation, Lightness), etc, which would be channels (depth) of pixel value arrays.

TIL-AI
TODAY I LEARNED-AI

Motivation for CV

- Making computers see like us, allowing them to take vision as an input and respond accordingly.

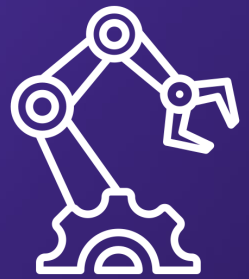
F()= G

Automation

Allowing machine to aid in assessing rules and recognising objects, standardizing and formalising vision inputs.

Interpretation

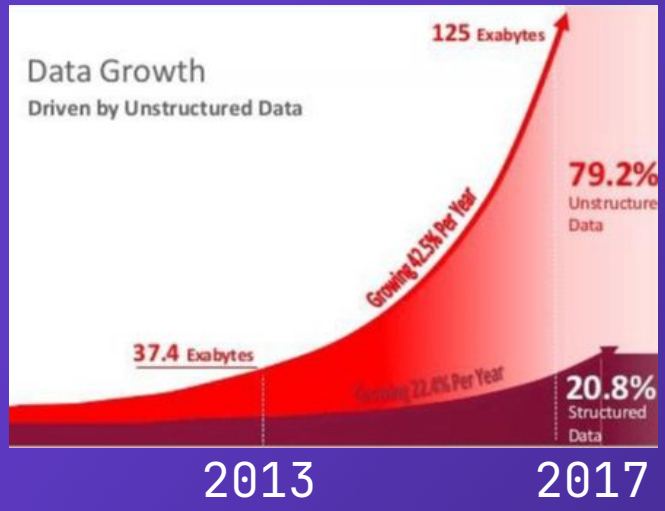
Allowing machine to take in vision as inputs, aid in classification, localizing and other vision tasks.



TIL-AI
TODAY I LEARNED AI

Motivation for CV

- Unlocking the power of images and videos: The world is full of data, and the majority of them are unstructured



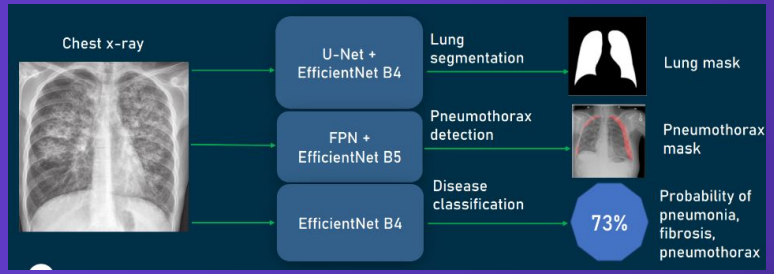
Unstructured Data
Video, Sound, Images

Structured Data
Tabular Data, Logs, Forms

Player	Minutes	Points	Rebounds	Assists
A	41	20	6	5
B	30	29	7	6
C	22	7	7	2
D	26	3	3	9
E	20	19	8	0
F	9	6	14	14
G	14	22	8	3
I	22	36	0	9
J	34	8	1	3

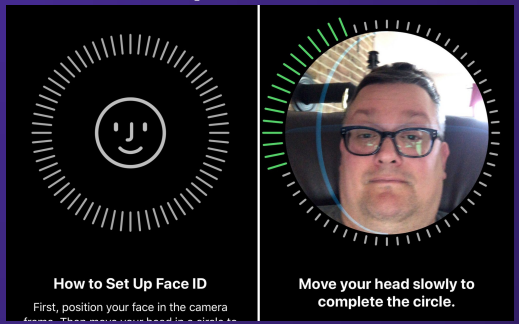
Motivation for CV
Use Cases

Lung Disease Diagnosis
Predictions of Disease



Car Navigation
Auto-pilot, Recognising objects

Phone Security
FaceID



TIL-AI
TODAY I LEARNED-AI

History of CV Techniques and Tools

19X0s

- 1960s: Camera-Computer Connection & Edge Detection
- 1970s: Feature Extraction & Object Recognition
- 1980s: Scale-Space Analysis & Shape Inference
- 1990s: Camera Calibration, 3D Reconstruction & Segmentation
- 1990s: OCR - Machines Read Text

1980s

CNN (Convolutional neural network) is invented.
Feature Extractions
Shape Inference

2009

ImageNet
Large Public Annotated Dataset for object classification

2012

AlexNet
Showed how deep neural networks - Convolutional neural network (CNN) can be used for image classification tasks.

2016

YOLO
Speed (45FPS)
Detection accuracy
Good generalization
Open-source

Present

Widespread Applications

- Self- Driving Cars,
- Facial Recognitions
- Medical Imaging
- Industry 4.0

Curated Data (Ingredients)

Architectual (Cookbook)

YOLO: You Only Look Once