CV / VLM

Unit 1: Introduction to Computer
Vision (CV)



1.1.1

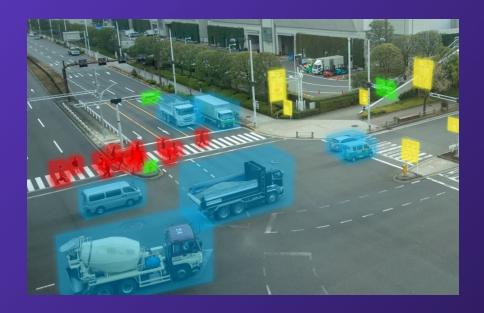
Basics of Computer Vision

What is computer vision?



What is Computer Vision?

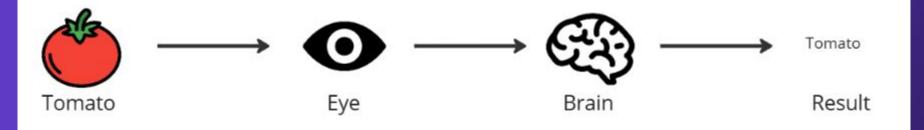
- A branch of AI that teaches computers to see and understand visual data in a way that mimics human vision
- Utilizes algorithms to analyse and interpret visual inputs

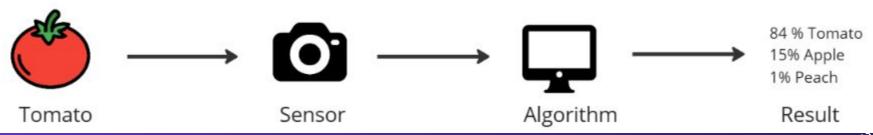


Adapted from <u>viso.ai</u>



Analog Comparison Human Vision vs Computer Vision

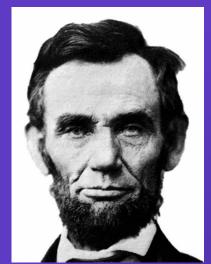






What Computers See - Pixel Values

What we see



Grayscale
(Black and White)

What computers see

-	_	_	_	_	_	_	_	_	-	_	_
157	163	174	160	160	162	129	151	172	167	356	15
156	182	163						110	210	180	35
180	180		34	54		10	33	49	106	150	18
206	1 = 1	5	134	181	100	130	204	164	15	56	100
194	*	187	251	287	299	239	228	227			20
172	100	207	233	283	214	220	259	228	-		20
188		179	209	185	216	211	184	136			760
188	**	166	м		168	134		31	62		Œ
198	168	191	165	156	227	178	143	102	110	26	19
205	174	188	252	236	291	149	176	228	43	16	23
190	216	114	149	236	187	*	150		38	218	34
190	224	147	108	227	210	127	162	36		255	22
190	214	173	66	163	143		86		100	249	21
187	196	236							217	255	21
188	202	237	141				144	200	126	243	29
196	204	122	207	177	121	123	200	175	13	-	211

157	163	174	168	150	162	129	151	172	161	155	156
166	182	163	74	75	62	33	.17	110	210	180	164
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	16	56	180
194	68	137	261	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	56	74	204
188	**	179	209	185	216	211	158	139	76	20	166
189	97	166	84	10	168	194	33	21	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
206	174	166	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	36	218	241
190	224	147	108	227	210	127	102	ж	101	255	224
190	214	173	66	103	143	95	50	2	109	249	216
187	196	236	75	1	#1	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	136	243	236
196	206	123	207	177	121	123	200	175	13	96	216

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.



Motivation for CV

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

Automation

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

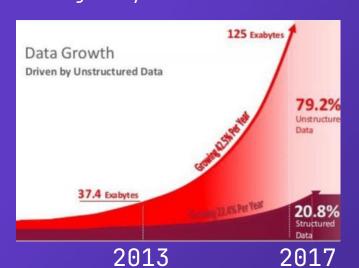
Interpretation

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

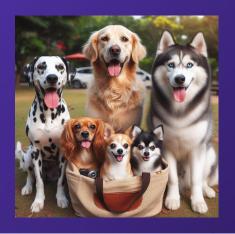


Motivation for CV

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

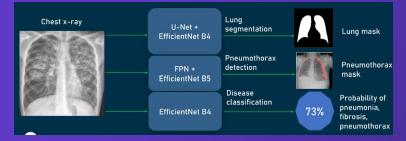


Structured Data
Tabular Data, Logs,
Forms

Player	Minutes	Points	Rebounds	Assists
Α	41	20	6	5
В	30	29	7	6
С	22	7	7	2
D	26	3	3	9
E	20	19	8	0
F	9	6	14	14
G	14	22	8	3
1	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





History of CV Techniques and Tools

19X0s

1980s

2009

2012

2016

Present

- · 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

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

ImageNet

Large Public Annotated Dataset for object classification

Curated Data (Ingredients)



AlexNet

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

YOLO

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

Widespread **Applications**

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

Architectual (Cookbook)

YOLO: You Only Look Once

