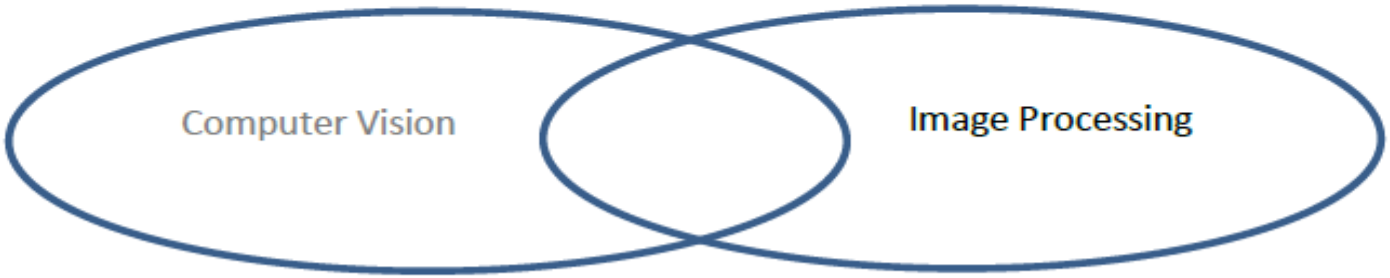
Lec. 1
Introduction to Computer Vision

Assist. Prof. Dr. Saad Albawi

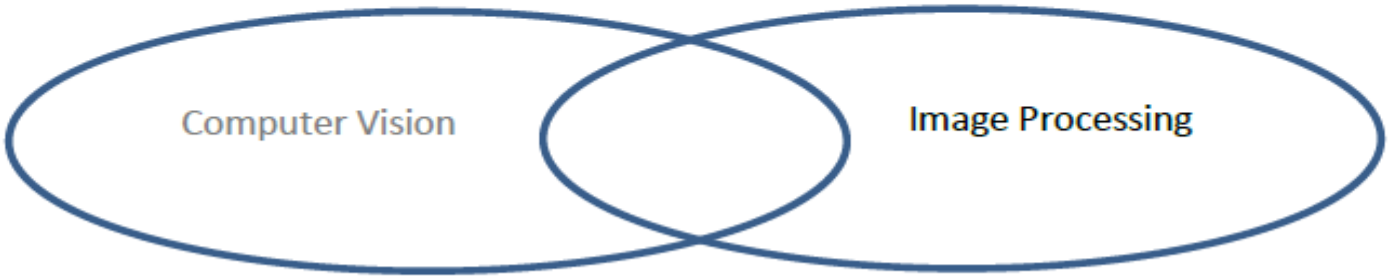
Computer Imaging



Computer vision has been around since the 1960s. Recent developments:

- Increasing availability of cheap, powerful cameras (e.g. digital cameras, webcams) and other sensors.
- Increasing availability of massive amounts of image and multimedia content on the web (e.g. face databases, streaming video or image-based communication).
- Increasing availability of cheap, powerful computers (processor speed and memory capacity).
- Introduction of techniques from machine learning and statistics (complex, data-driven models and algorithms).

Computer Imaging



Computer vision has been around since the 1960s. Recent developments:

- Increasing availability of cheap, powerful cameras (e.g. digital cameras, webcams) and other sensors.
- Increasing availability of massive amounts of image and multimedia content on the web (e.g. face databases, streaming video or image-based communication).
- Increasing availability of cheap, powerful computers (processor speed and memory capacity).
- Introduction of techniques from machine learning and statistics (complex, data-driven models and algorithms).

What is Computer Vision?

The goal of computer vision is to develop algorithms that allow computer to “see”.

Also called

- Image Understanding
- Image Analysis
- Machine Vision

Is a computer imaging where the application does not involve a human being in the visual loop. In other word, the image are examined and acted upon by a computer. Although people are involved in the devvelopment of the system, the final application requires a computer to use the visual information directly. One of the major topics within the filed of computer vision is image analysis.

What is Computer Vision?

- Make computers understand images and videos.

- What kind of scene?
- Where are the cars?
- How far is the building?

What is (computer) vision?

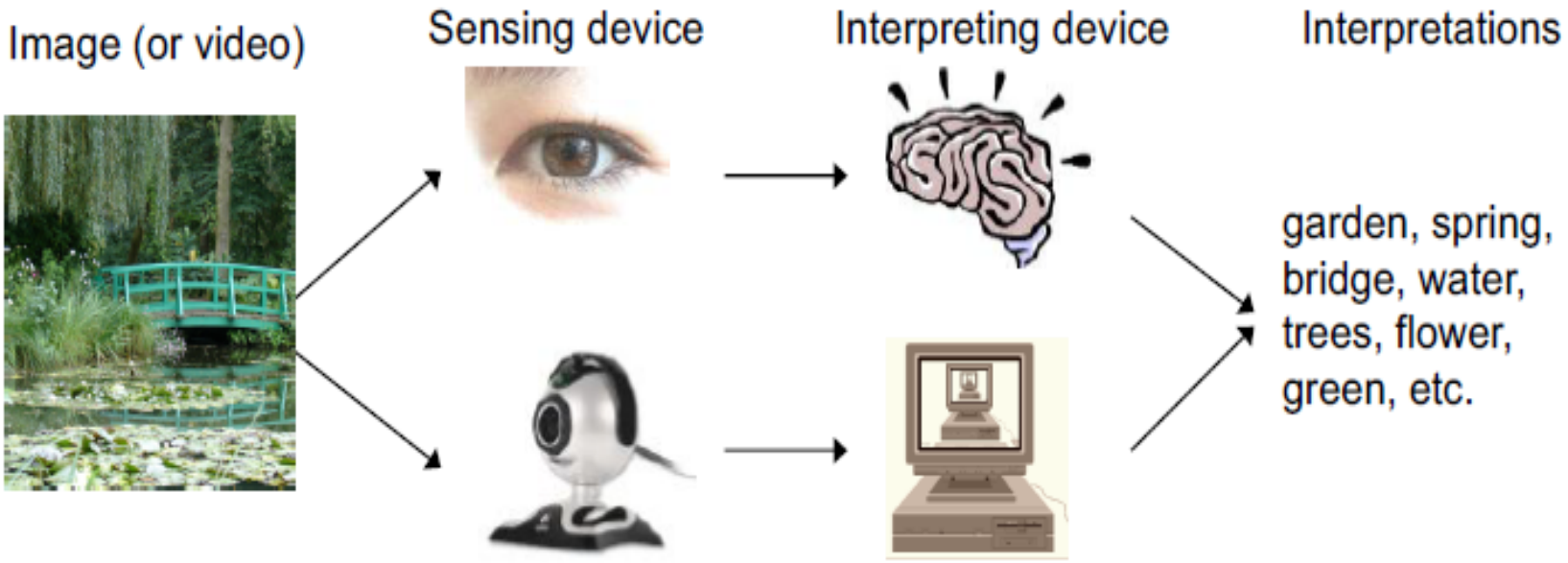


Image analysis involve the examination of the image data to facilitate solving a vision problem. The image analysis process involves two other topics: Feature extraction and pattern classification.

Features extraction is the process of acquiring higher level image information, such as shape or color information

Pattern classification is the act of taking this higher level information objects within the image.

Computer vision systems are used in many and various types of enviroments from manufacturing plants to hospital surgical suites to the surface of Mars. For example. In manufacturing systems, computer vision is often used for quality control. There, the computer vision system will scan manufactured items for defects and provide control signals to robotic manipulater to remove defective parts automatically.

General visual perception is hard

Digital Image

232	182	163	151	151	148	148	143	145	139	143	136	139	136	134	132	129	130	126	124	115	116	115	104	109	102	100	101
244	218	160	149	145	147	145	143	139	142	140	139	134	134	130	131	125	120	120	116	110	110	107	100	100	97	95	97
246	233	196	145	145	146	141	141	137	134	140	133	133	125	131	125	114	121	116	116	109	101	95	101	97	87	89	91
248	242	222	161	142	140	145	137	138	135	129	127	127	122	124	118	116	113	102	110	99	102	98	94	91	88	91	90
252	246	234	192	143	139	136	134	133	129	131	127	124	121	117	114	111	105	108	95	101	102	86	88	91	84	84	99
252	249	242	215	151	137	134	134	129	126	126	121	120	116	113	111	108	104	99	94	102	93	89	96	79	87	92	112
252	248	242	227	169	134	135	124	122	120	125	121	116	115	105	112	102	99	92	98	93	88	89	74	87	65	97	111
253	246	244	236	192	134	125	123	119	120	118	116	112	107	110	95	104	94	89	96	84	86	79	77	65	79	105	119
252	250	246	238	210	144	126	118	120	115	116	116	98	105	103	102	96	93	91	82	80	79	75	70	82	81	108	119
250	251	247	239	219	161	127	117	117	109	105	107	100	104	99	100	98	79	98	70	75	80	72	65	86	83	113	124
252	249	247	241	226	177	122	120	116	106	108	110	91	103	93	99	89	88	79	80	72	74	76	65	84	87	109	123
249	249	247	241	231	191	131	116	110	109	106	98	95	90	102	83	78	86	85	80	70	75	69	81	79	96	122	121
248	249	247	244	238	206	133	121	101	103	94	97	91	87	87	83	83	82	77	78	81	61	73	65	86	99	118	120
247	250	248	244	237	215	149	115	102	105	91	94	80	91	79	83	81	70	71	75	74	71	78	74	76	108	117	119
250	247	246	243	239	218	159	108	100	87	100	88	92	83	85	77	81	63	80	70	63	73	70	78	81	110	120	116
248	245	244	241	239	224	170	113	103	94	89	86	84	83	74	81	68	78	76	66	66	70	73	65	92	108	115	123
248	244	244	242	237	226	179	123	98	94	84	74	88	77	71	76	71	78	68	67	63	72	72	75	94	109	115	124
247	244	245	241	238	221	183	123	95	87	89	73	77	79	71	65	78	56	69	66	62	61	70	69	90	113	118	118
247	246	244	242	236	219	185	120	100	84	82	79	66	67	76	72	69	55	61	56	65	57	70	78	95	106	119	116
246	245	244	241	231	216	190	126	91	86	77	77	72	71	76	60	69	60	57	52	66	55	62	75	87	110	110	117
245	244	244	237	231	221	189	133	97	83	70	73	62	59	77	44	65	66	60	70	51	43	67	75	95	107	116	111
244	244	241	237	230	222	188	133	90	83	77	77	59	78	60	67	62	61	66	72	62	51	62	71	96	105	115	108
242	242	237	236	232	219	187	126	83	79	70	64	58	66	63	67	54	65	51	65	58	54	62	73	77	92	107	94
241	241	238	236	229	216	186	125	85	77	70	66	64	53	63	55	56	53	67	39	52	25	23	9	11	51	66	77
241	239	237	237	228	214	185	127	92	83	64	66	69	62	61	65	32	42	12	7	6	15	65	123	146	160	167	172
240	239	237	236	225	208	178	123	89	67	72	67	49	54	27	10	7	23	103	142	162	167	169	168	171	172	172	178
238	236	236	229	221	203	174	125	77	82	55	33	23	9	79	135	163	173	174	175	174	170	171	167	167	172	169	173
235	235	231	228	215	198	165	122	84	43	14	57	132	166	176	175	179	177	176	178	178	173	169	172	167	168	171	162
231	231	227	223	210	191	163	110	44	95	159	174	175	179	178	180	183	180	179	177	175	175	174	173	169	168	171	156
230	226	225	220	202	187	169	151	175	180	182	177	182	182	183	184	184	184	181	182	181	178	182	179	172	161	160	155
223	224	220	213	198	191	185	186	182	182	178	179	184	185	191	189	189	192	188	192	193	194	192	187	179	161	153	147
220	219	213	203	191	182	181	177	176	173	175	180	182	184	192	192	193	195	200	203	203	206	205	202	192	164	150	151
212	209	200	188	177	173	174	171	169	165	173	176	180	187	191	192	195	195	201	203	207	210	208	212	201	177	147	143

The goal of computer vision

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What we see

What a computer sees

Related fields

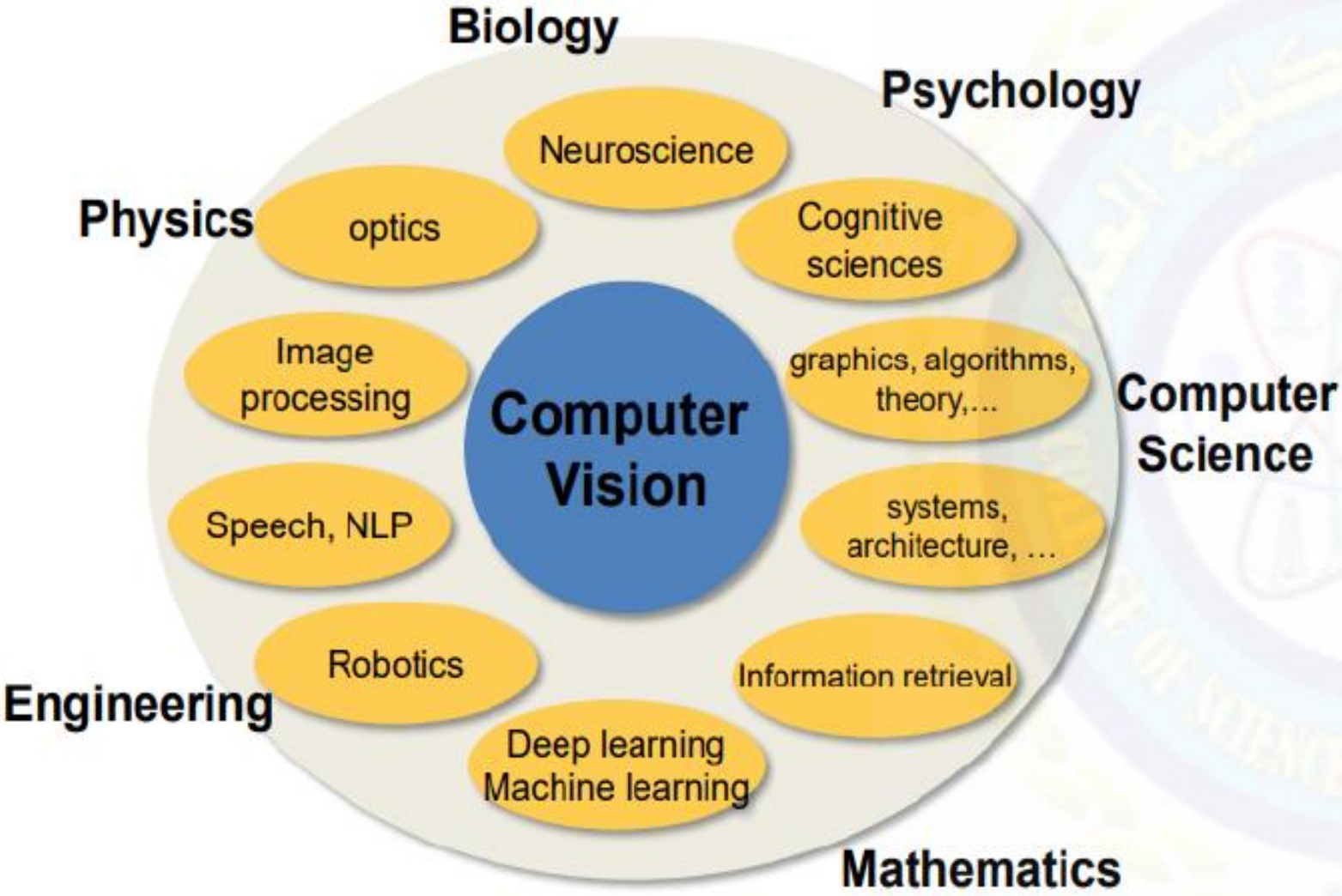
- Image Processing
- Pattern Recognition
- Photogrammetry
- Computer graphics

Our Time

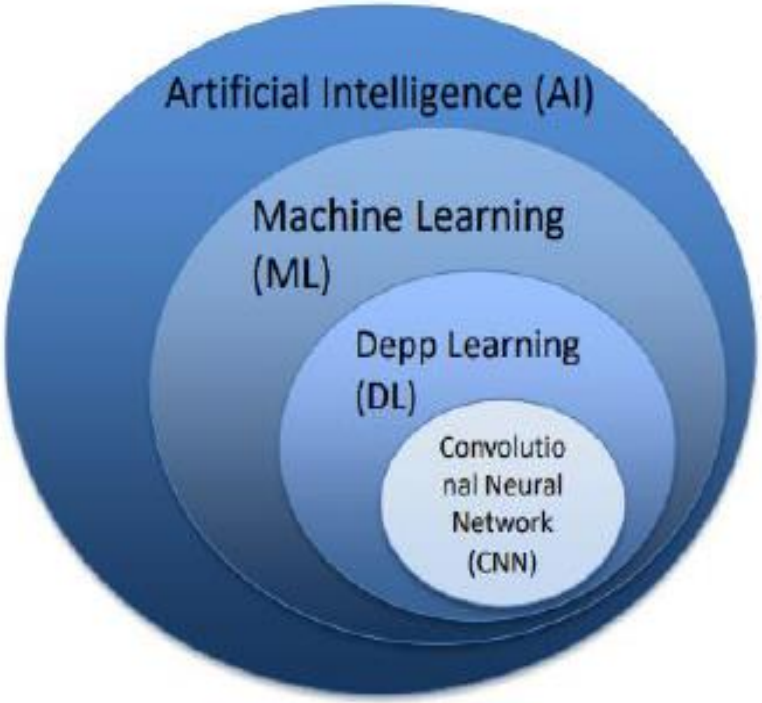
It is a good time to do computer vision now, because:

- Powerful computers
- Inexpensive cameras
- Algorithm improvements
- Understanding of vision systems

What is Computer Vision?



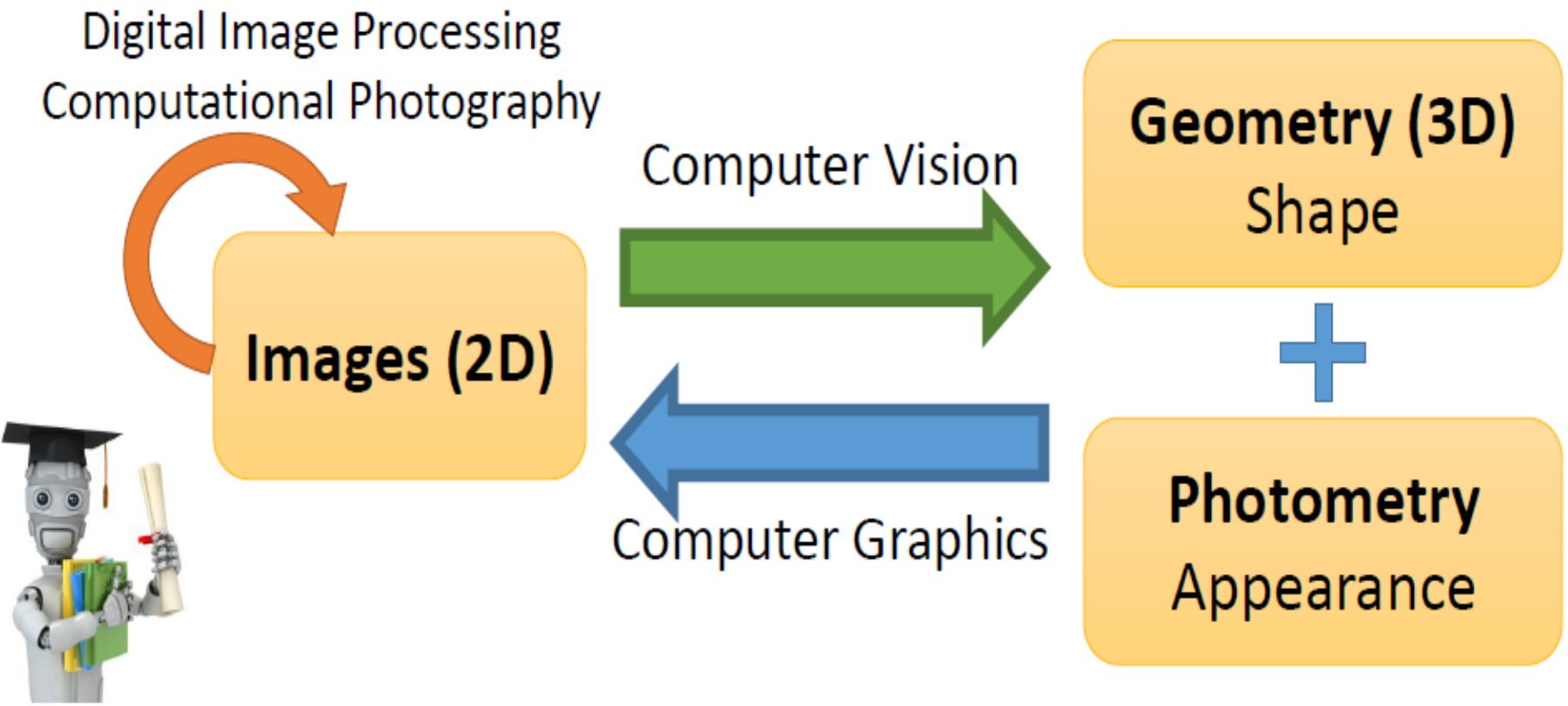
What is Computer Vision?



Computer Vision

- Object detection
- Object classification
- Scene understanding
- Semantic scene segmentation
- 3D reconstruction
- Object tracking
- Human pose estimation
- Activity recognition
- VQA
- ....

Computer Vision and Nearby Fields



Machine learning:

Vision = Machine learning applied to visual data

Where did we come from?

- Edges, segmentation, and perception

D. Lowe. *IJCV*, 1992

Low Level Topics in Computer Vision:

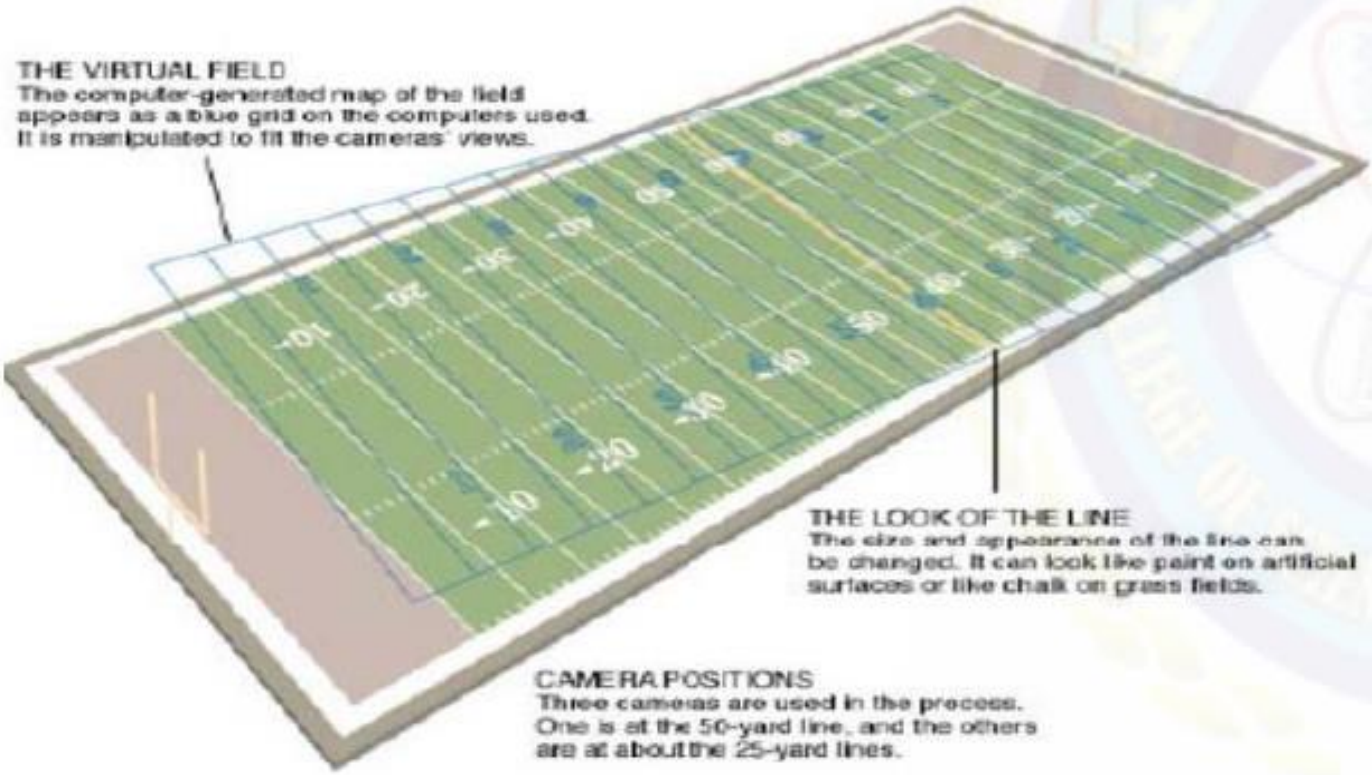
- Low level operations.
  - Image enhancement.
  - Feature detection.
  - Region segmentation.

Low Level Topics in Computer Vision:

- Low level operations.
  - Photo Stitching.

Low Level Topics in Computer Vision:

- Low level operations.
  - 3D shape reconstruction, stereopsis, structure from motion



Low Level Topics in Computer Vision:

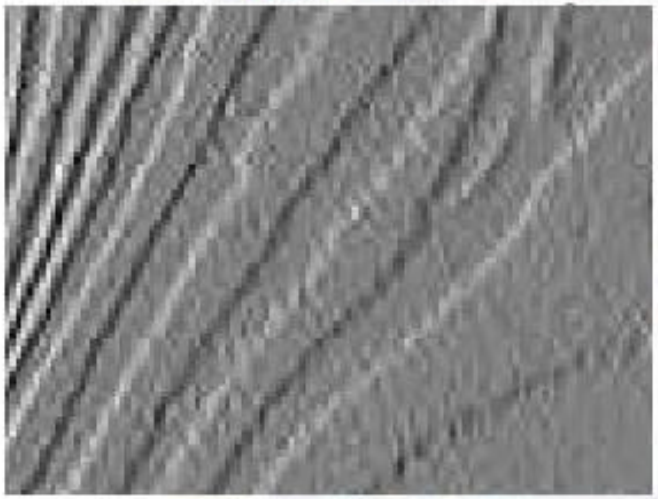
- Detecting sharp edges by convolution.

Convolve this image

With this kernel

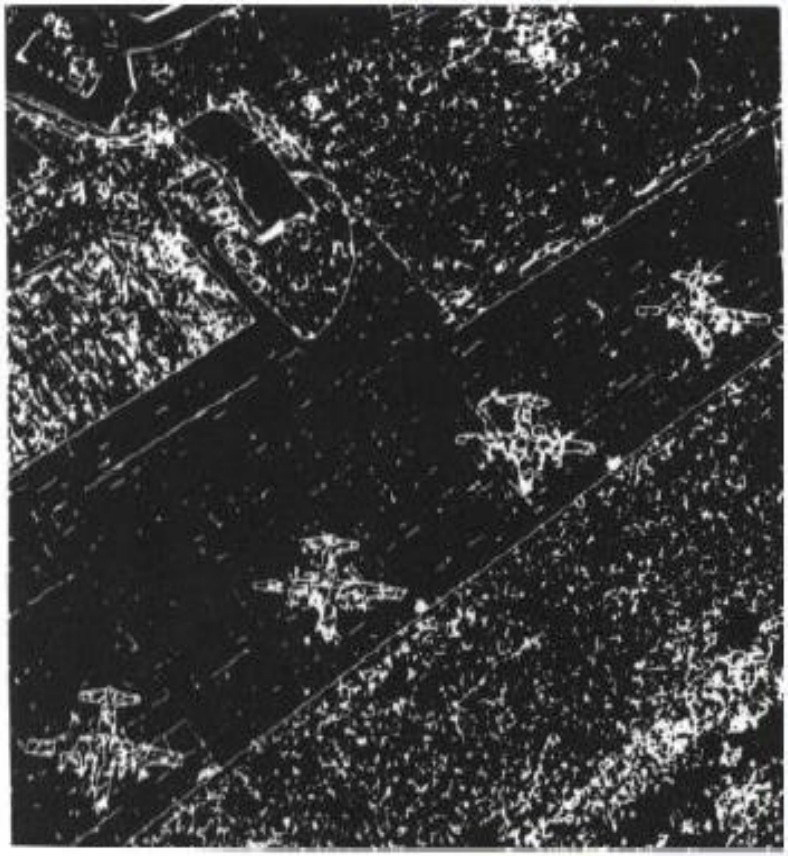
→

To get this



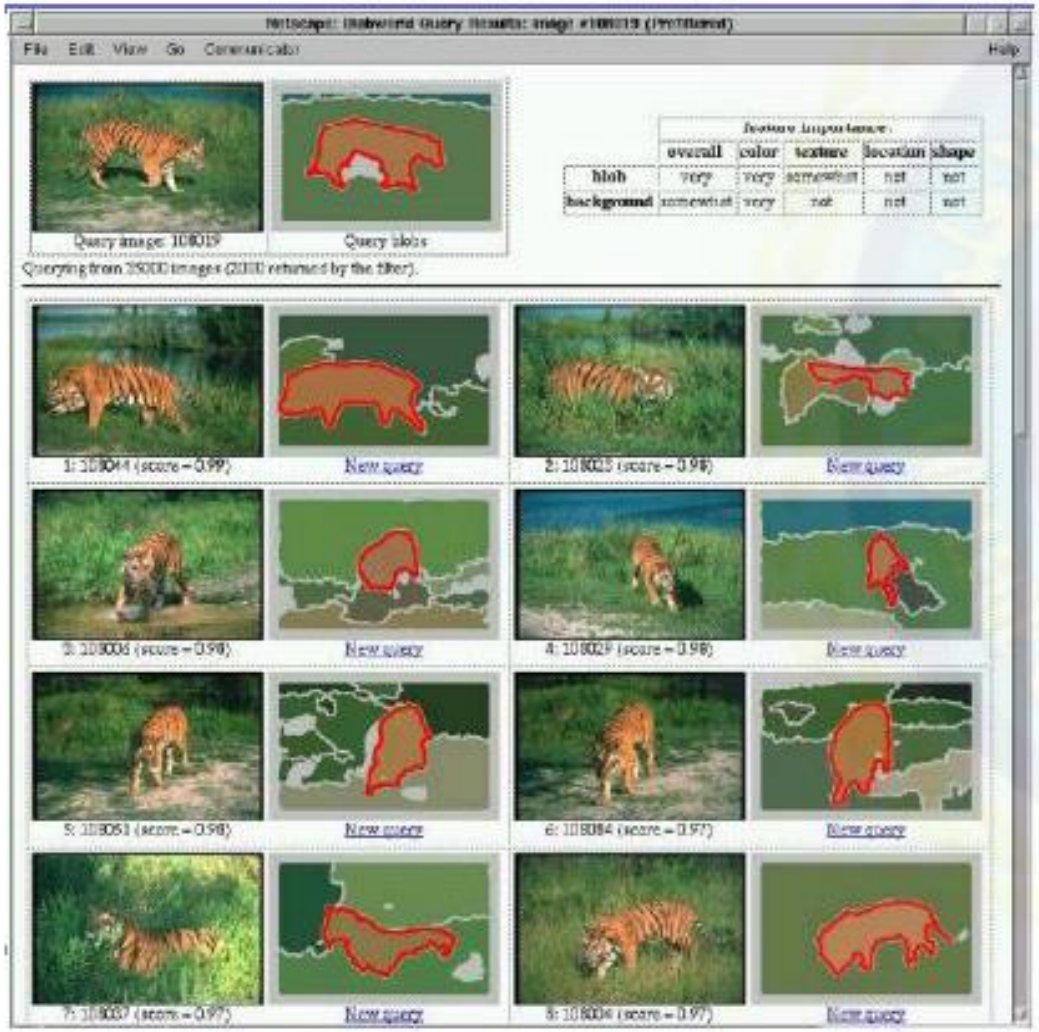
Low Level Topics in Computer Vision:

- Edge Detection Results:



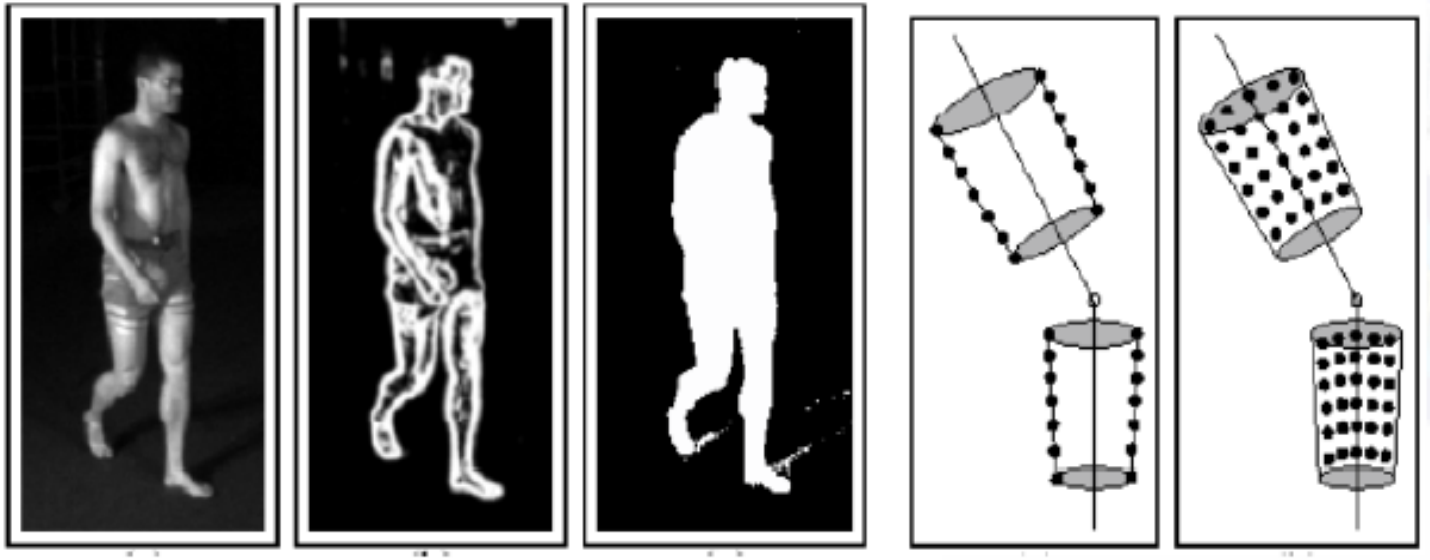
Mid Level Topic in Computer Vision:

- Segmentation:



Mid Level Topic in Computer Vision:

- Tracking:
  - Use a model to predict next position and refine using next image.
  - Model
    - Simple dynamic models (second order dynamics).
    - kinematic models.
    - etc.
  - Face tracking and eye tracking now work rather well.



Mid Level Topic in Computer Vision:

- Finding coherent structure so as to break the image or movie into big units.
  - Segmentation:
    - Breaking images and videos into useful pieces.
    - E.g., finding video sequences that correspond to one shot.
    - E.g., finding image components that are coherent in internal appearance.
  - Tracking:
    - Keeping track of a moving object through a long sequence of views

High Level Topic in Computer Vision:

- Recognition of people, places, events.

BBC NEWS

UK version | International version | About the versions

News Front Page
World
UK
England
Northern Ireland
Scotland
Wales
Business
Market Data
Your Money
E-Commerce
Economy
Companies
Politics
Health
Education

Last Updated: Monday, 6 February 2006, 14:29 GMT

E-mail this to a friend | Printable version

Face-hunting cameras boost Nikon

Japanese camera maker Nikon has tripled its profits on the back of strong sales of digital cameras that automatically focus on human faces.

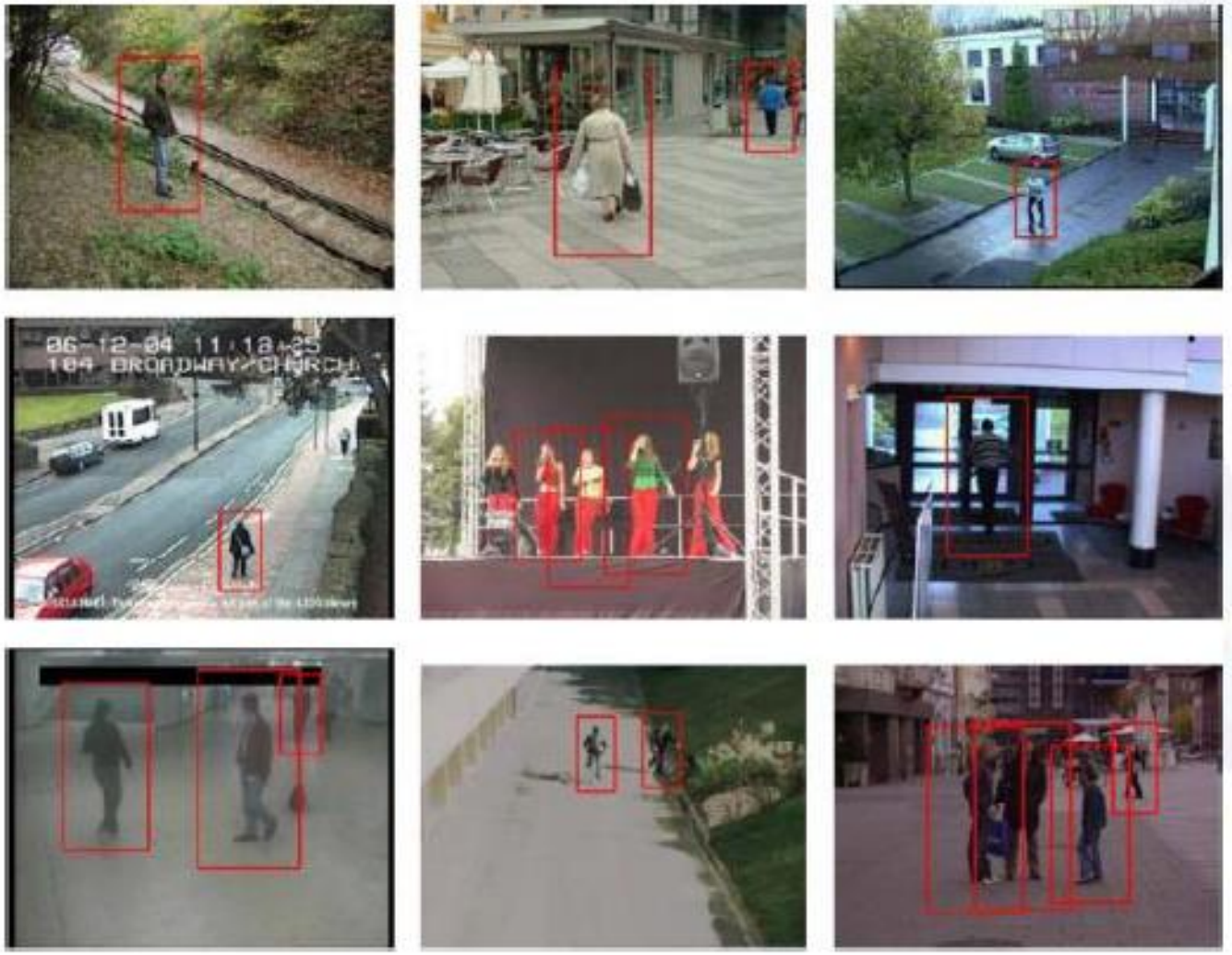
Face recognition cameras like the Coolpix L1 are popular

High Level Topic in Computer Vision:

- Face Detection.

High Level Topic in Computer Vision:

- Human Detection.



High Level Topic in Computer Vision:

- Object Recognition and Detection.
  - Object recognition:
    - Recognizing the object category.
  - Object Detection:
    - Detecting whether certain type of object presents type of object presents and where it locates.
  - Database and Challenges.
    - Pascal 2009, 2008.
    - Caltech 101, 256.

High Level Topic in Computer Vision:

- Object Recognition:

Caltech 101 images

Fei-Fei et al. 2004

Visual Object Classes Challenge 2009 (VOC2009)

PASCAL2
Pattern Analysis, Statistical Modelling and Computational Learning

[click on an image to see the annotation]

Everingham et al. 2006-2012

High Level Topic in Computer Vision:

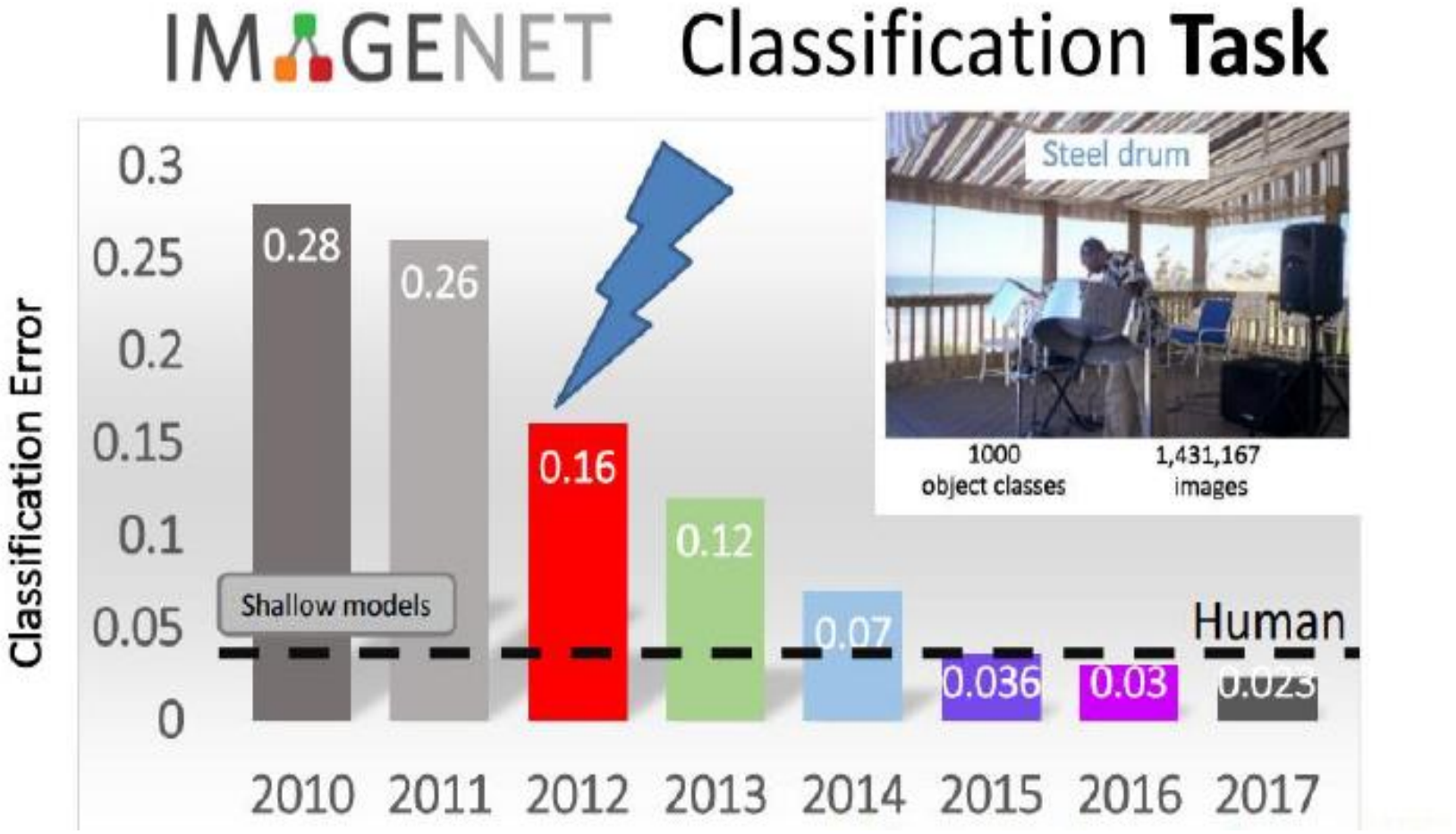
IMAGENET

22,000 categories ⋮ 15,000,000 images

J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li & L. Fei-Fei. *CVPR*, 2009.

High Level Topic in Computer Vision:

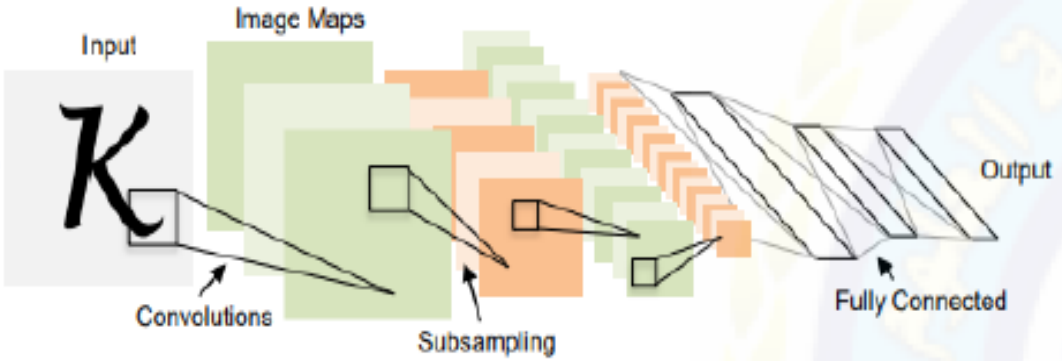
- Object Recognition:



High Level Topic in Computer Vision:

- Object Recognition:

1998
LeCun et al.



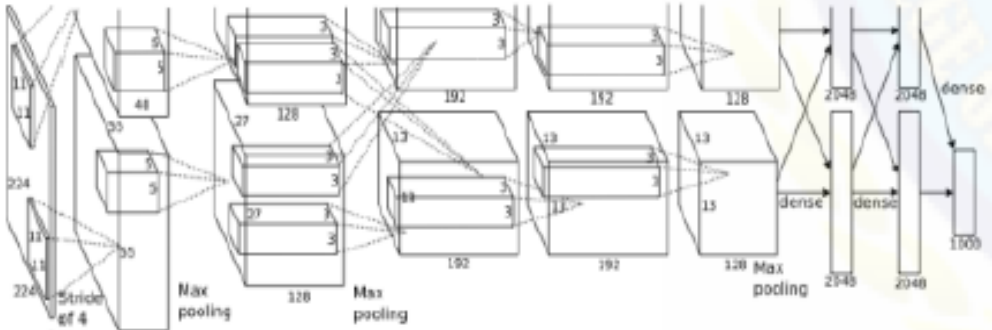
# of transistors

10^6

# of pixels used in training

10^7 NIST

2012
Krizhevsky et al.



# of transistors GPUs

# of pixels used in training

High Level Topic in Computer Vision:

- Object Recognition:

IMAGENET **Large Scale Visual Recognition Challenge**

Year 2010

NEC-UIUC

Dense descriptor grid: HOG, LBP

Coding: local coordinate, super vector

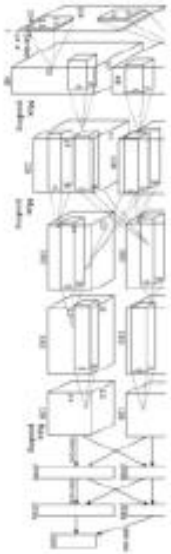
Pooling, SPM

Linear SVM

[Lin CVPR 2011]

Year 2012

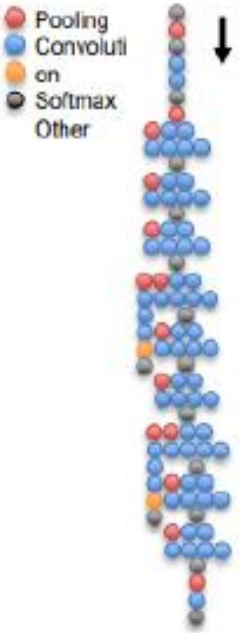
SuperVision



[Krizhevsky NIPS 2012]

Year 2014

GoogLeNet



[Szegedy arxiv 2014]

VGG

[Simonyan arxiv 2014]

Year 2015

MSRA

[He ICCV 2015]

High Level Topic in Computer Vision:

- Image Captioning: Richer Descriptions

Prior Work

Image Captioning → *A man riding a horse drawn carriage down a street*

Our Recent Work

Dense Captioning → JKF, CVPR 2016

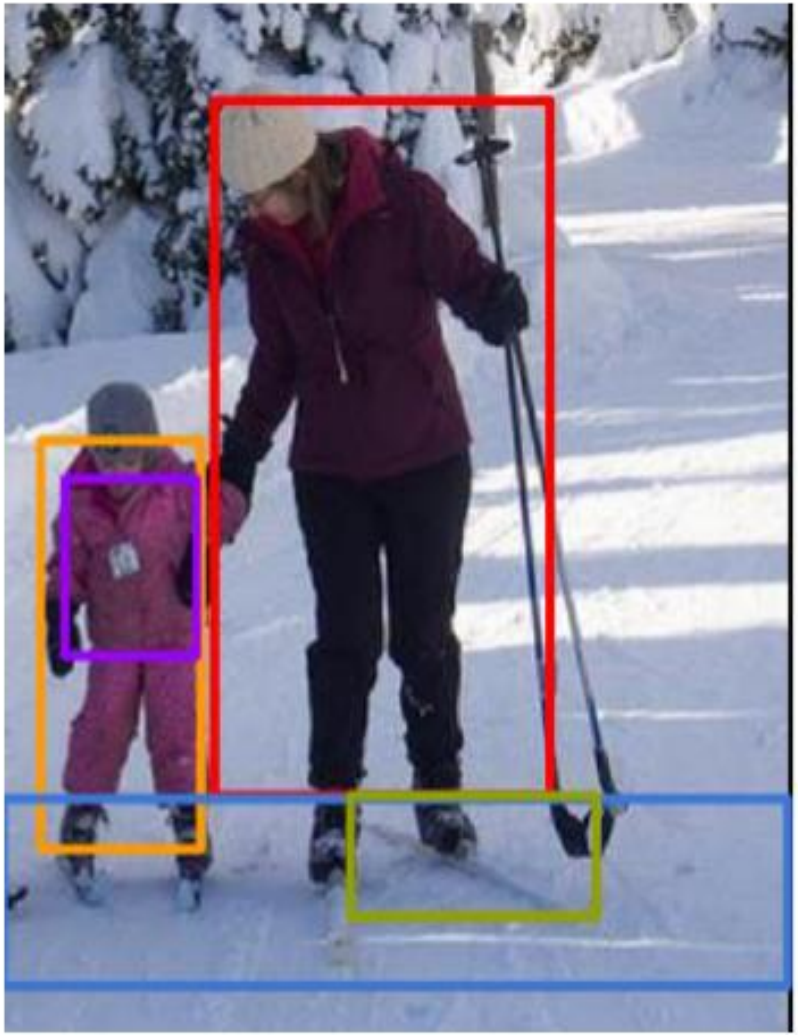
Horse pulling a cart. A wheel on a cart. A window on a building. A horse in a picture. A large white umbrella. A woman sitting on a bench. Man sitting on a motorcycle. A cart with a cart.

Paragraph Captioning → KJKF, CVPR 2017

A man is riding a carriage on a street. Two people are sitting on top of the horses. The carriage is made of wood. The carriage is black. The carriage has a white stripe down the side. The building in the background is a tan color.

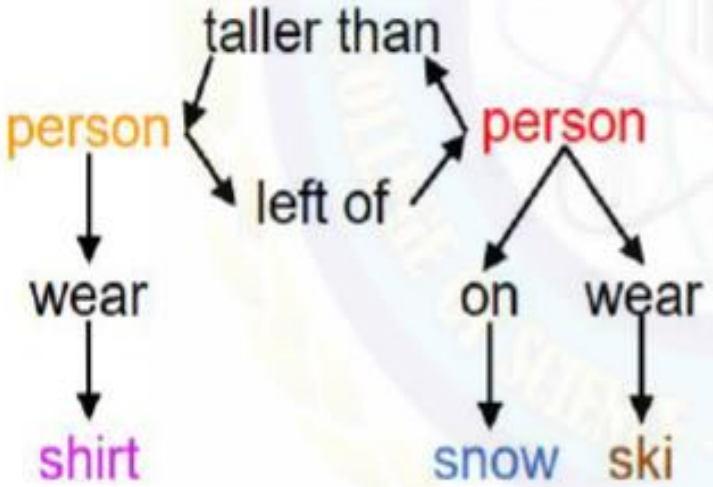
High Level Topic in Computer Vision:

- Image Captioning: Richer Descriptions



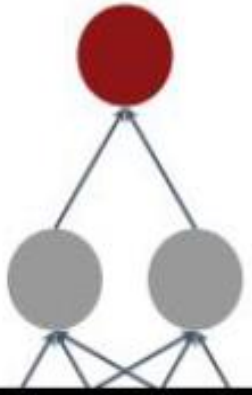
Results:

spatial, comparative, asymmetrical, verb, prepositional

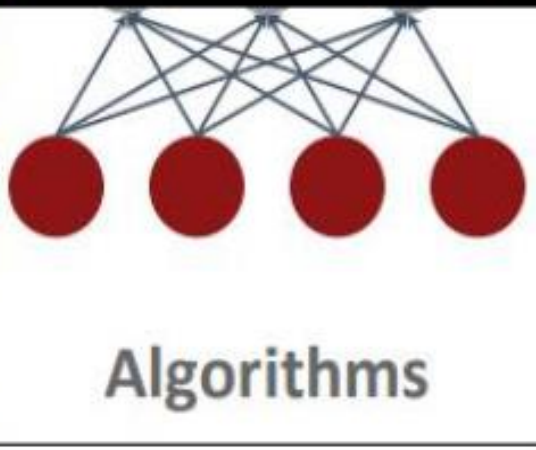


Krishna*, Lu*, Bernstein, Fei-Fei, *ECCV* 2016

High Level Topic in Computer Vision:



The Deep Learning Revolution



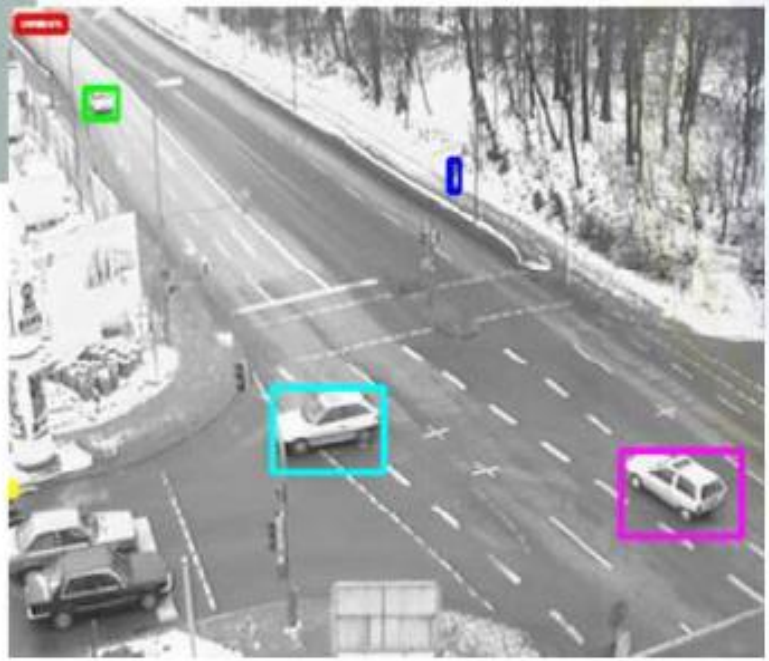
High Level Topic in Computer Vision:

- Action Recognition and Activity Anal
  - Action recognition:
    - Recognizing the action of a subject.
  - Activity Analysis:
    - Composition of several actions.
    - Multiple people involved.
  - Database and Challenge.
    - Trecvid 2009, 2008.
    - KTH, Weizmann

High Level Topic in Computer Vision:

- What is Video Surveillance?

Warn me if there is a mugging in the grove



High Level Topic in Computer Vision:

- Background Subtraction, Human Detection and Tracking in a Video and Tracking in a Video.

High Level Topic in Computer Vision:

- Real Surveillance video for an accident.

High Level Topic in Computer Vision:

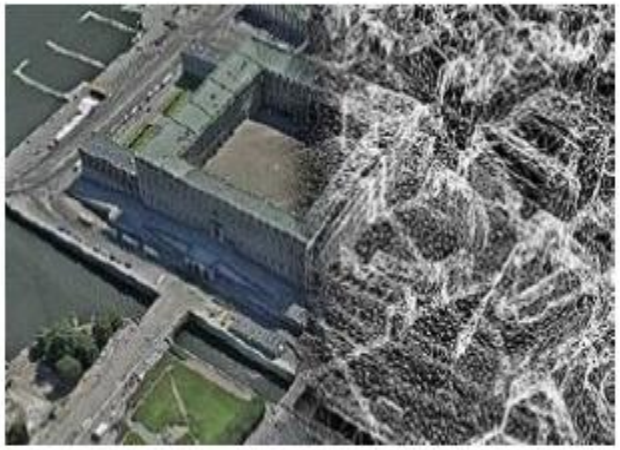
- Face Detection
  - Almost all digital cameras detect faces.
  - Snapchat face filters.

High Level Topic in Computer Vision:

- Smile Detection:

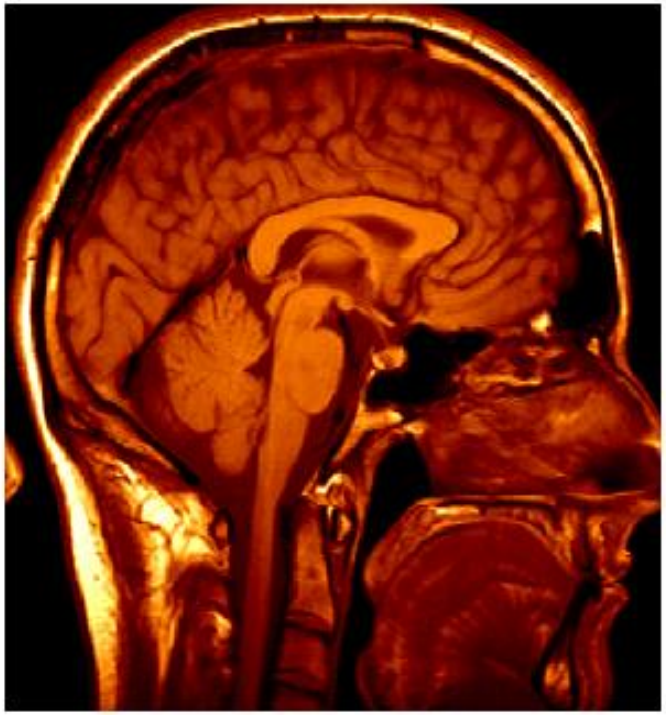
High Level Topic in Computer Vision:

- 3D from Images



High Level Topic in Computer Vision:

- Medical Imaging:



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

High Level Topic in Computer Vision:

- AutoCars - Uber bought CMU's lab:

High Level Topic in Computer Vision:

- Mobile robots

NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover

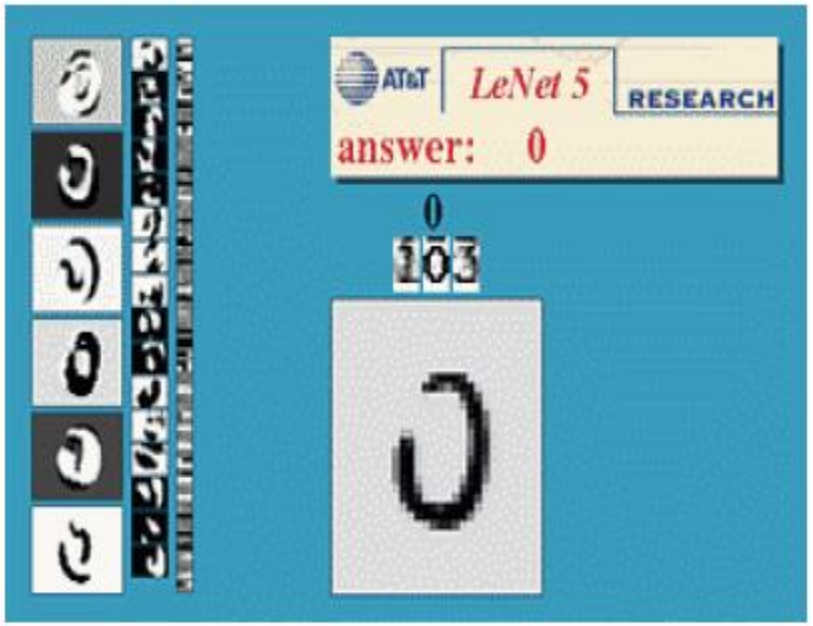
http://www.robocup.org/

amazon Prime Air

2020's: autonomous vehicles

Optical Character Recognition (OCR)

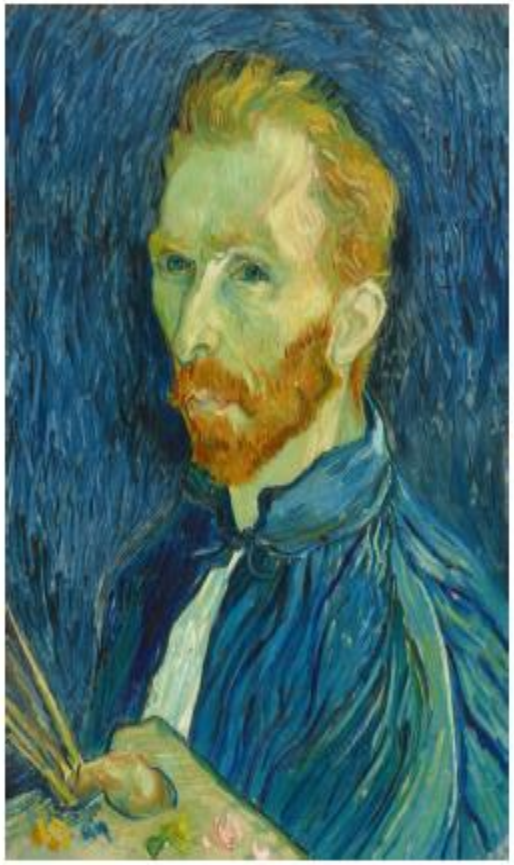
- Technology to convert scanned docs to text
  - If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Style transfer



Source image (**Style**)

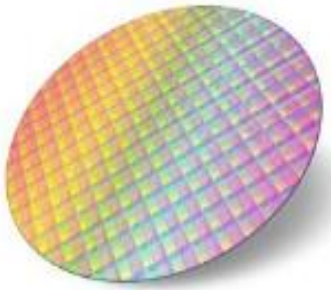
Target image (**Content**)

Output (deepart)

Panoramic Mosaics

Machine vision

Automated visual inspection



Deep Face