Advanced systems of computer vision

1. Introduction and basic concepts

Computer vision

- **Computer vision** is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to understand and automate tasks that the human visual system can do.

Computer vision

- Defectoscopy in manufacturing
- Faces, gestures and human recognition
- Medical recognition
- Barcodes recognition
- Traffic recognition
- Optical characters recognition
- Fingerprint recognition, iris recognition

Image and computer vision

- "It's better to see something once, than to hear about in a thousand times." - Asian Proverb.
- 'A hundred listenings do not [equal|compare with] one look'. - Japan.
- "One picture is worth a thousand [ten thousand] words."
- “To see is to believe”.

Computer vision tasks

- Acquiring,
- processing,
- analyzing,
- understanding digital images,
- extraction of high-dimensional data in order to produce numerical or symbolic information.

How the computer see

HUMAN SEES BEAUTY BUT COMPUTER SEES NUMBERS

WHAT IS THIS?

=

[137	137	137	137	137	140	134	133	129	130	131	131	131	131	132	131	129	130
131	127	132	128	133	129	130	129	126	131	132	130	131	132	130	129	136	134
133	131	133	140	140	137	134	133	137	137	141	137	139	137	136	136	134	136
133	132	134	131	134	137	133	137	135	140	143	140	144	148	144	150	145	153
149	147	152	146	146	143	137	143	135	131	137	134	124	124	117	118	116	110
101	102	94	86	85	80	70	64	63	55	55	60	62	59	52	52	51	56
62	57	60	64	53	67	57	62	58	53	69	65	55	54	57	58	57	68
58	64	60	60	62	61	61	66	67	65	68	72	65	67	63	71	67	71
66	67	64	65	65	75	73	67	61	69	61	60	66	62	60	69	64	64
57	62	66	59	64	67	62	60	61	63	70	65	65	67	68	67	65	63
68	64	65	61	65	69	63	69	72	75	74	69	75	78	76	76	80	73
74	75	68	73	73	83	69	76	77	79	75	75	73	78	74	71	72	74
78	73	77	73	69	76	66	69	73	65	63	68	74	70	73	66	74	70
72	66	69	65	66	68	69	66	66	60	73	69	63	64	66	65	72	70
68	64	65	65	59	70	64	65	66	59	60	65	61	64	65	70	65	65
73	67	66	63	65	71	76	71	67	67	68	68	67	67	71	65	71	66
65	56	53	63	56	54	58	51	46	57	52	48	51	50	42	48	41	49
56	50	52	44	49	49	49	46	52	49	52	53	56	53	61	61	56	44
50	45	48	46	42	47	40	49	42	51	55	47	45	38	49	39	42	44
48	35	46	41	40	40	34	22	27	32	28	22	28	30	30	25	25	33
31	32	25	20	43	32	24	29	17	39	26	28	23	20	18	16	22	19
20	22	22	26	27	21	17	15	23	23	19	21	24	30	30	30	39	48
69	88	102	112	110	116	124	123	125	119	121	118	112	108	107	109	113	104
81	67	47	36	26	14	12	13	9	9	5	10	14	10	8	12	9	6
8	6	7	6	7	11	4	5	7	6	9	8	7	7	8	10	7	8
14	13	8	13	21	20	18	14	15	13	16	12	12	11	14	9	9	8
14	16	15	17	25	21	18	32	45	63	70	63	57	59	62	56	47	43
39	29	28	23	16	20	19	20	28	24	25	38	27	31	29	24	25	25
25	24	26	31	23	18	22	22]										

How the computer see

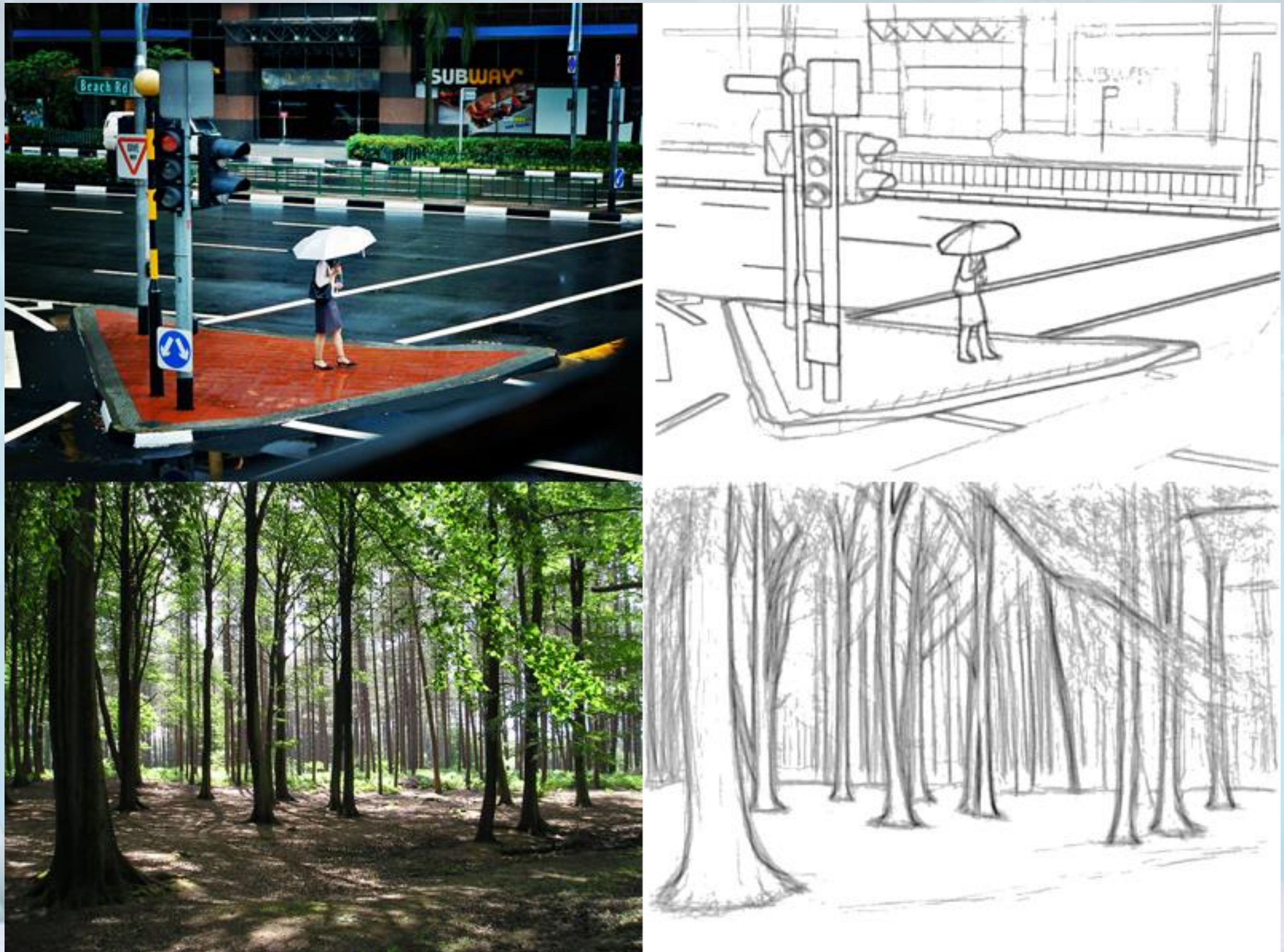
Computer vision vs human vision

What we see

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

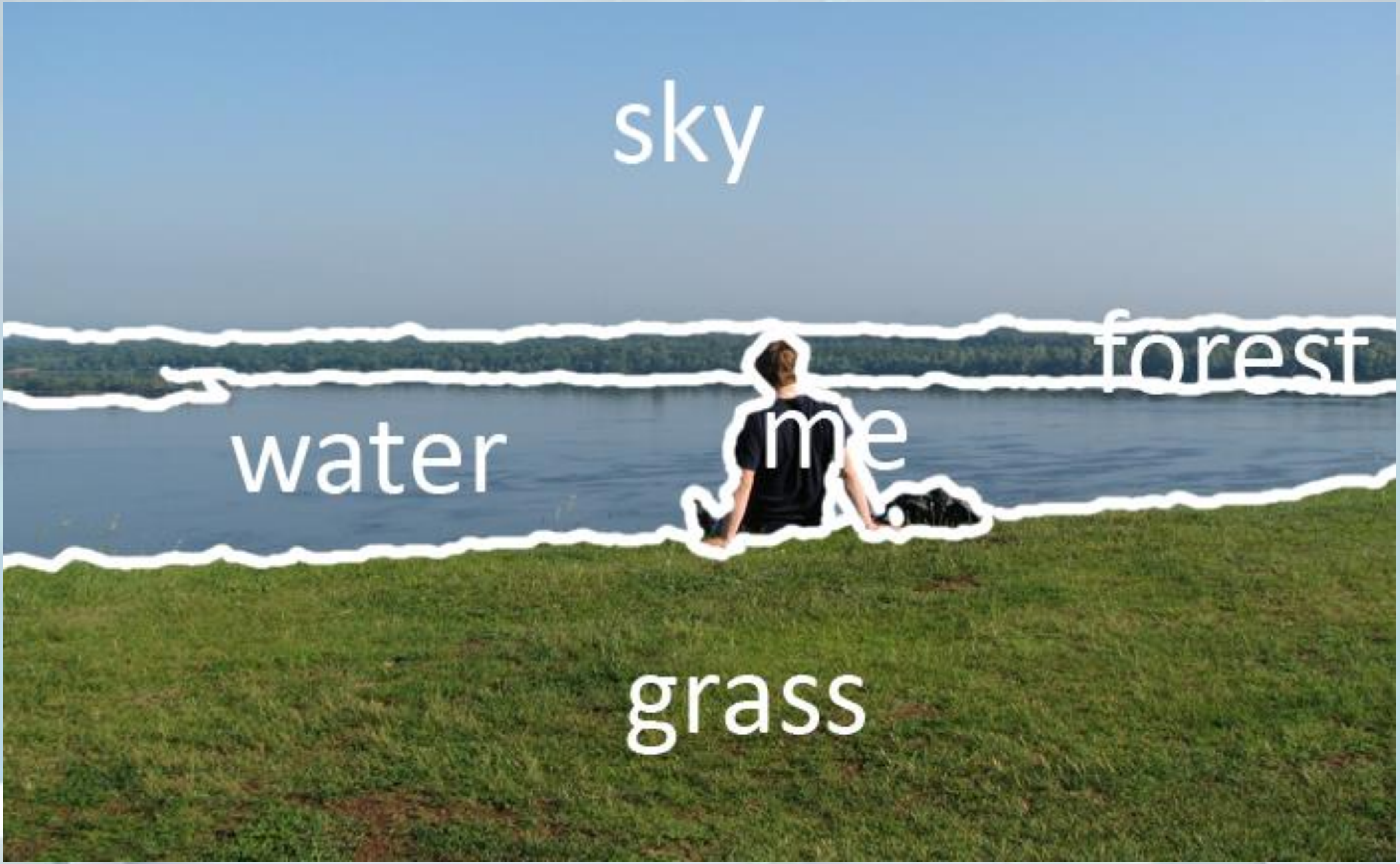
What a computer sees

Scene classification

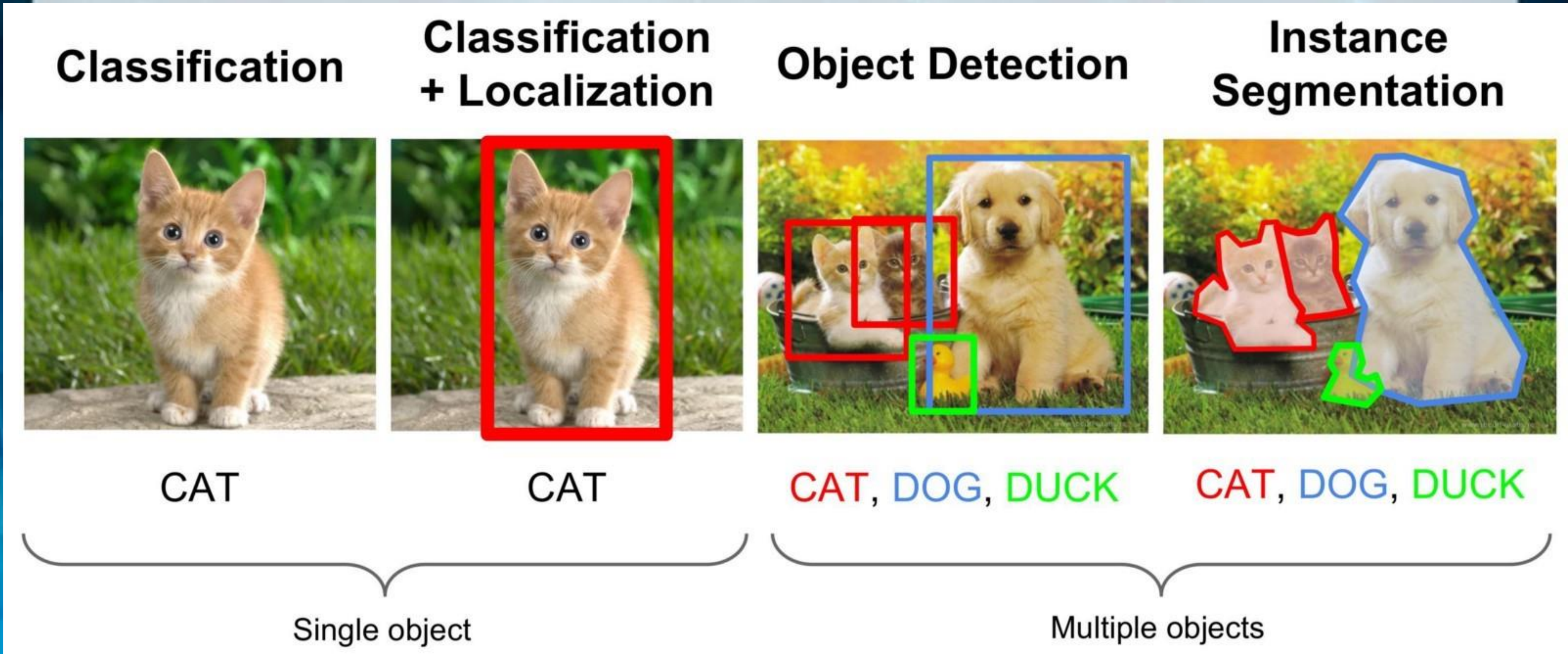


Scene classification

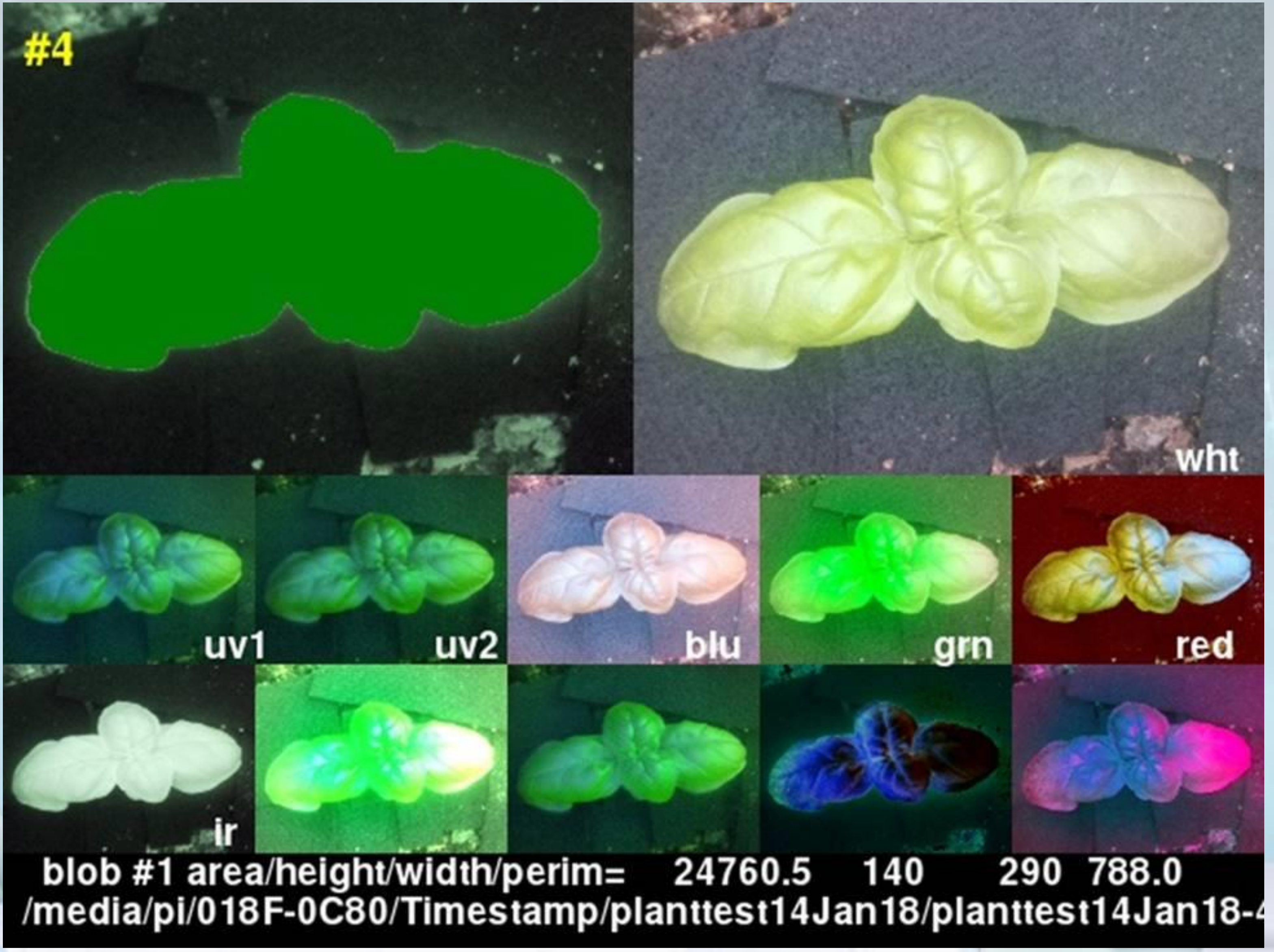
Object segmentation



Classification, localization, detection, segmentation



Measurements



Image

- Computer vision
- Signal
- Digital image
- Video
- Acquiring, processing, analyzing, recognition image
- Artificial system

Application domain

Recognition:
- Digital images and video
- Speech
- Biosignals
- Sounds
- Electromagnetic signals

Images processing

- Images recognition
- Machine vision
- Images restoration

Measurements

- Spectral analysis
- Speed and position measurements
- Noise measurement

Audio

- Speech recognition
- Speech synthesis

Military

- Communication
- Radars
- guidance system, automatic flies

Biomedicine

- Diagnosis
- Patients observations
- Electroencephalograms (EEG)
- Electrocardiography (EKG)
- X-Rays, Medical ultrasound

Consumer goals

- Mobile phones
- Systems of communication
- Digital cameras
- Quality analysis

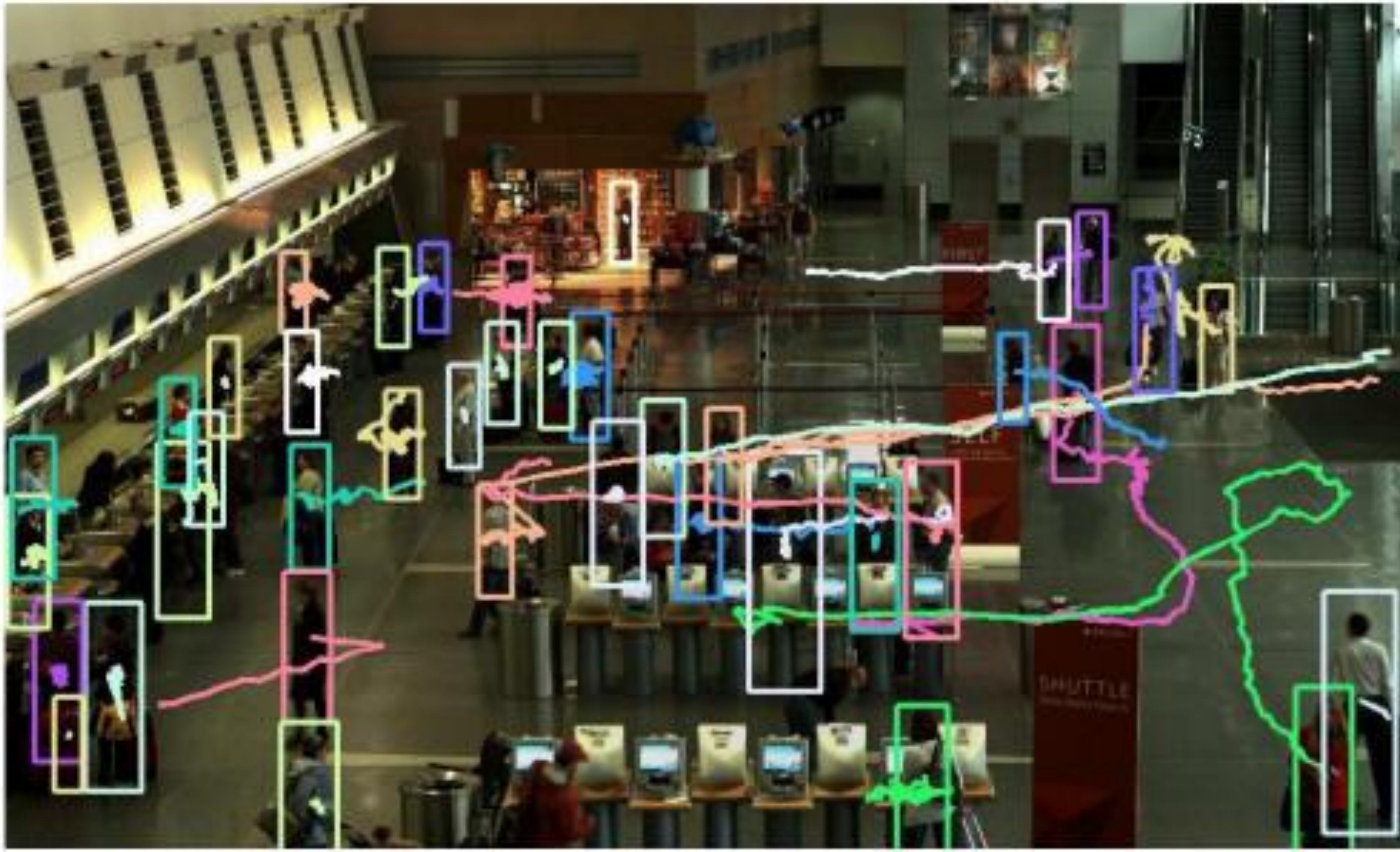
Robots

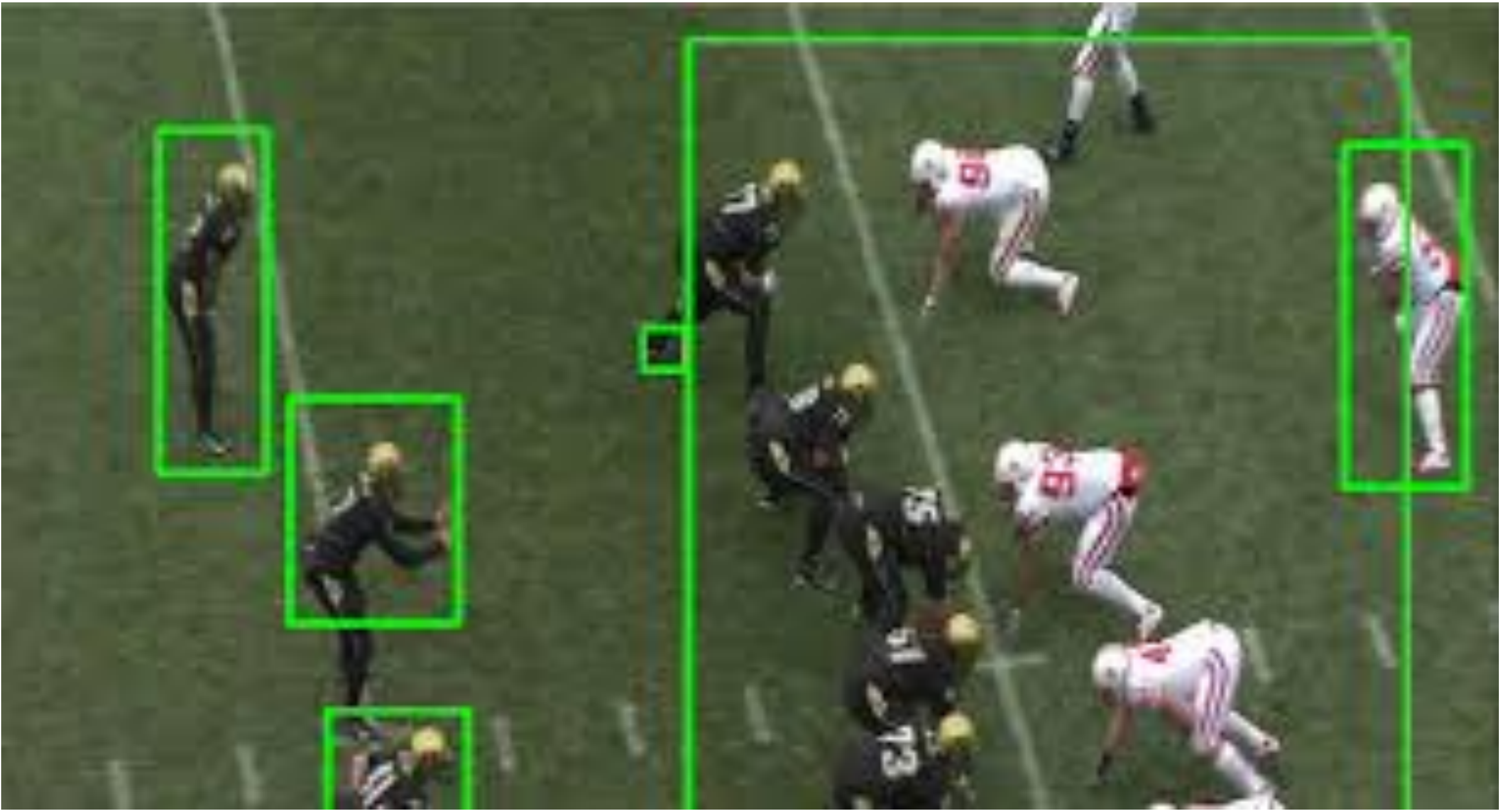
- Automotive
- Autonomous robotic systems
- Production
- Mobile platforms

Neurobiology

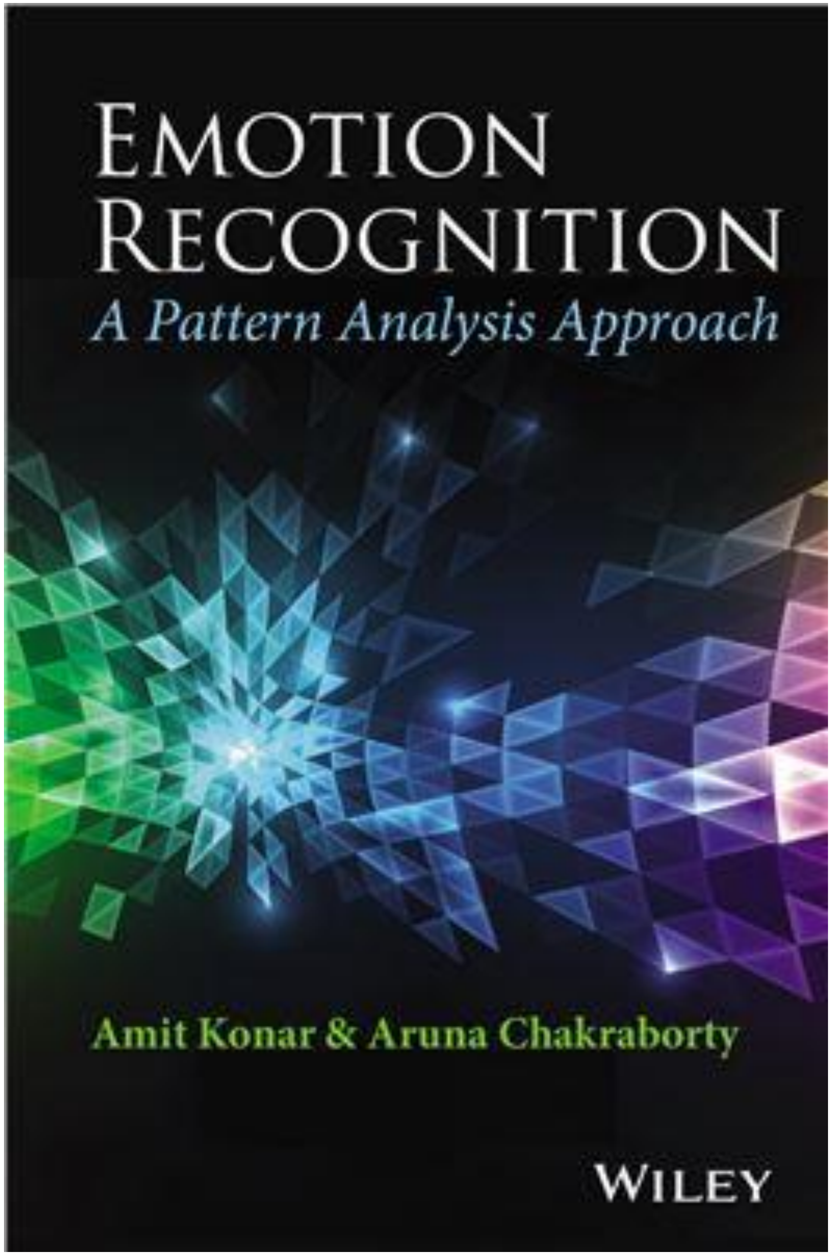
- Study of human and animal vision
- Systems Technology and Human-Computer Interaction, Neurocomputer interface

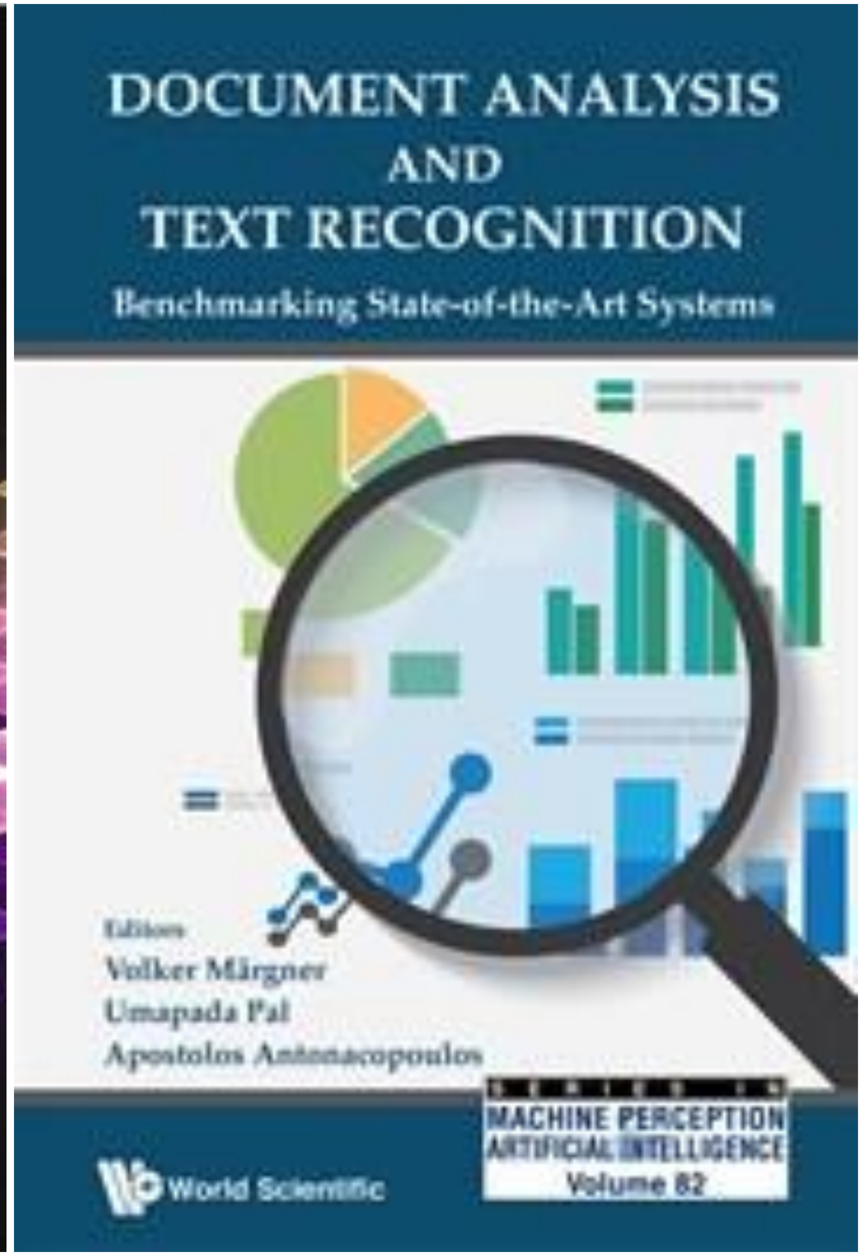

sky
building
STOP
sky
WEST ST
sky
tree
tree
building
person
building
tree
tree
tree
building
STOP
car
tree
tree
tree
car
car
car
car
tree
car
car
car
car
road
car
car
car
road
car
car
building
car
road
road
road
road
grass
person

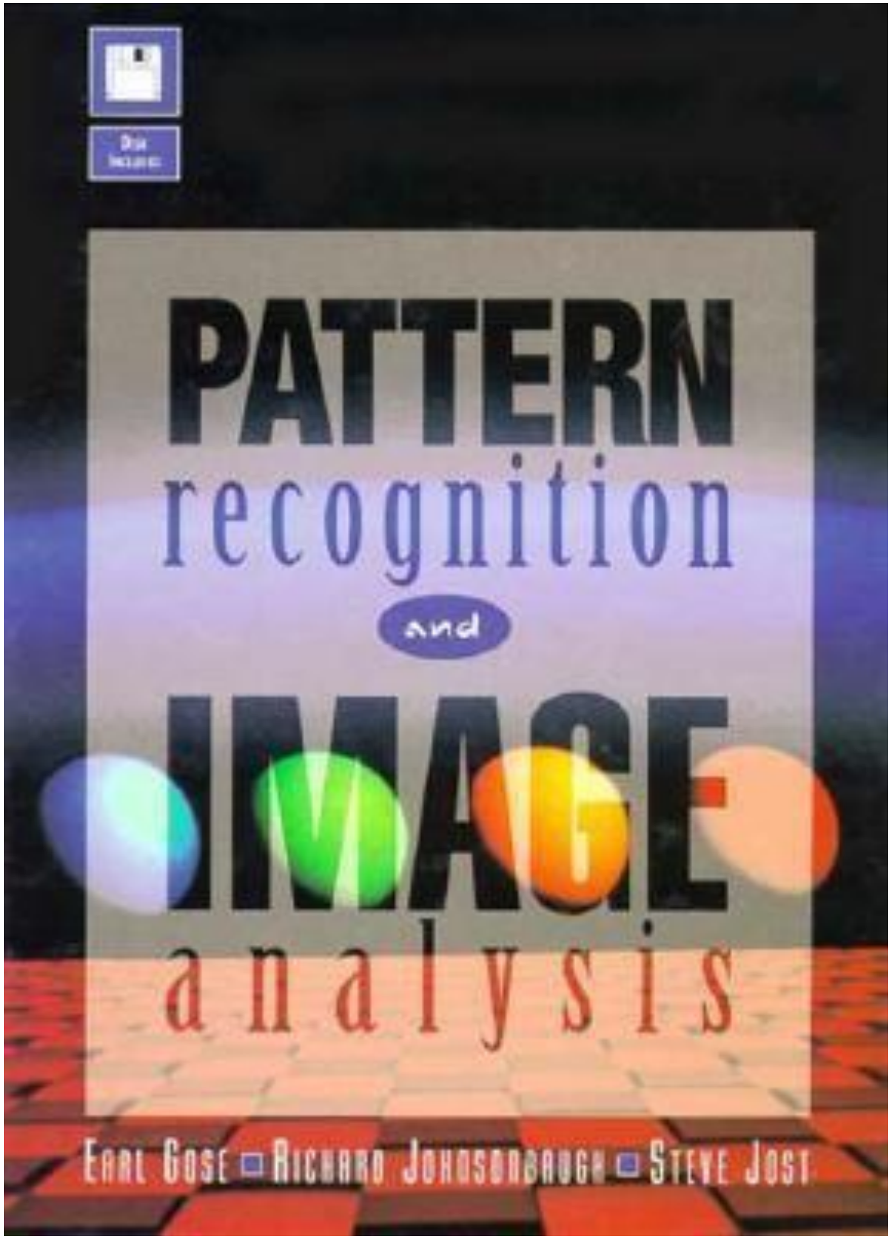

SHUTTLE

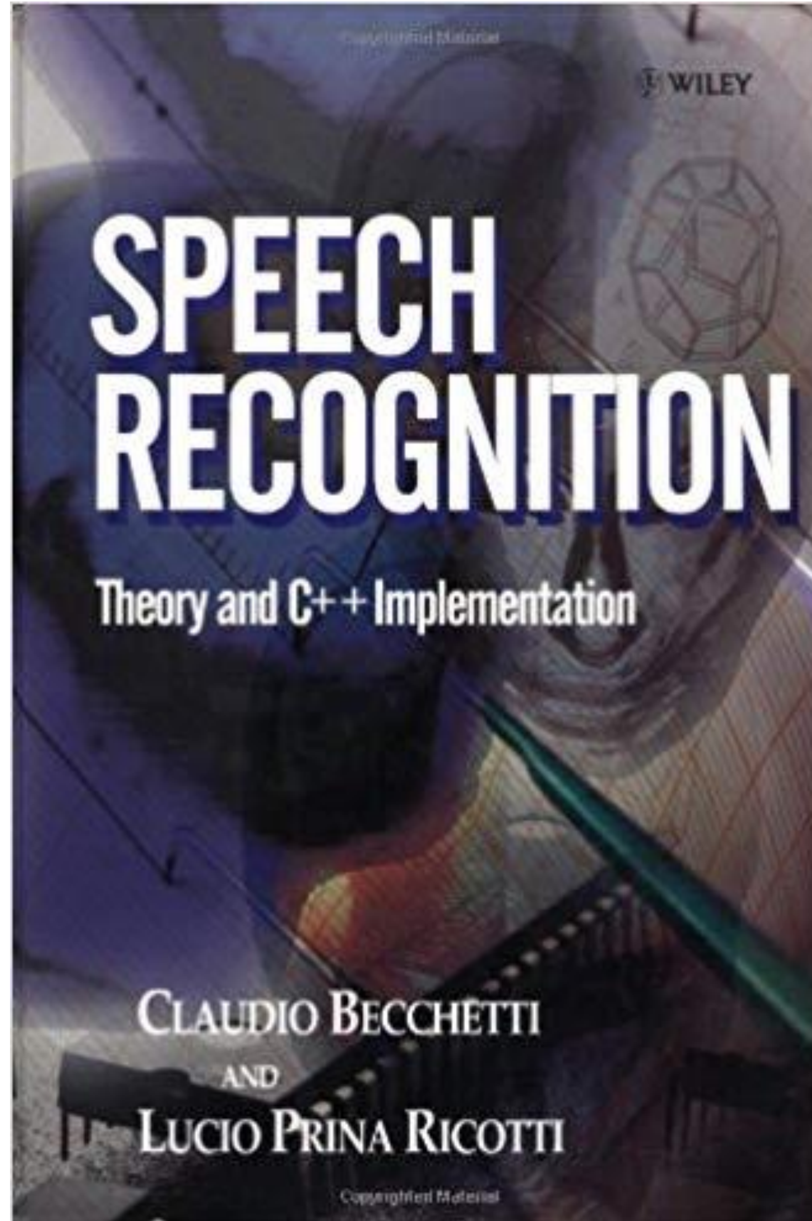


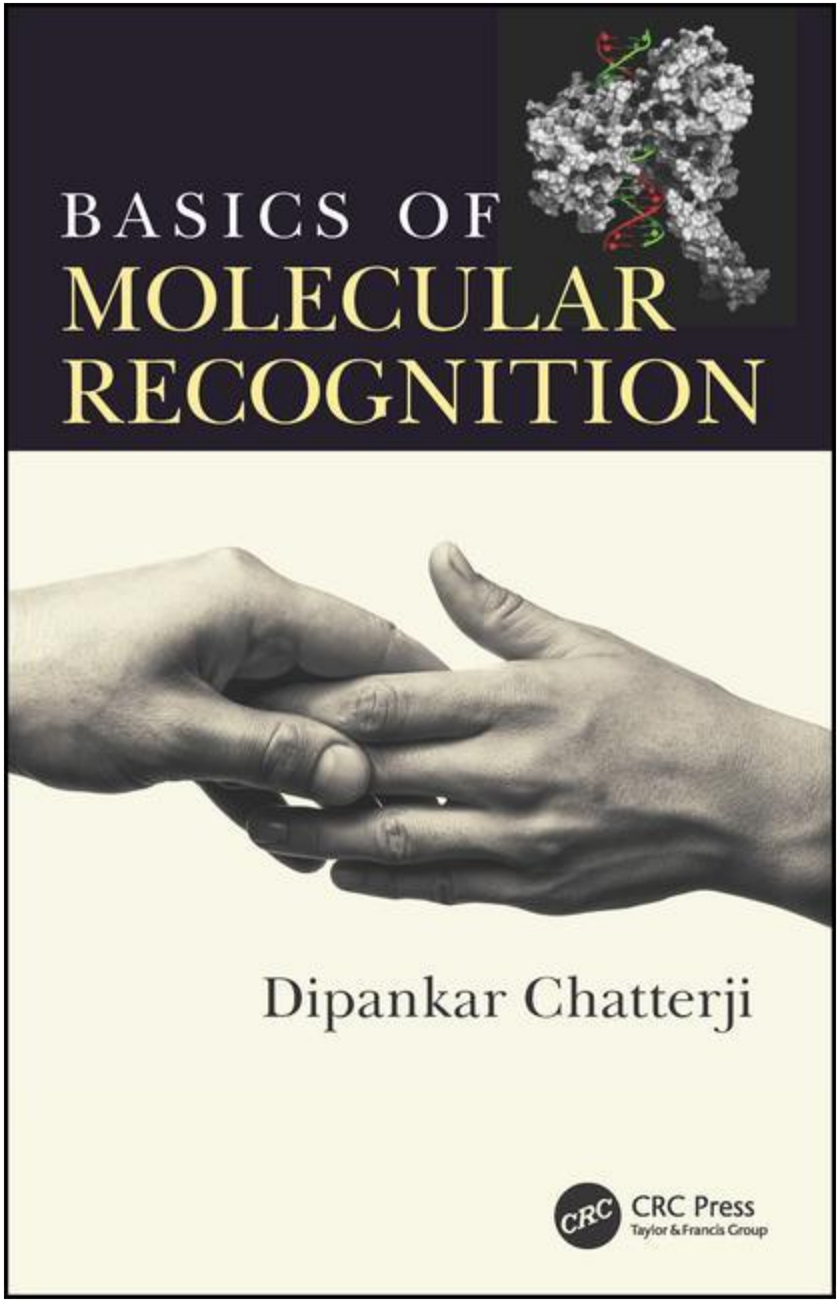

Happy
Surprise
Angry
Sad

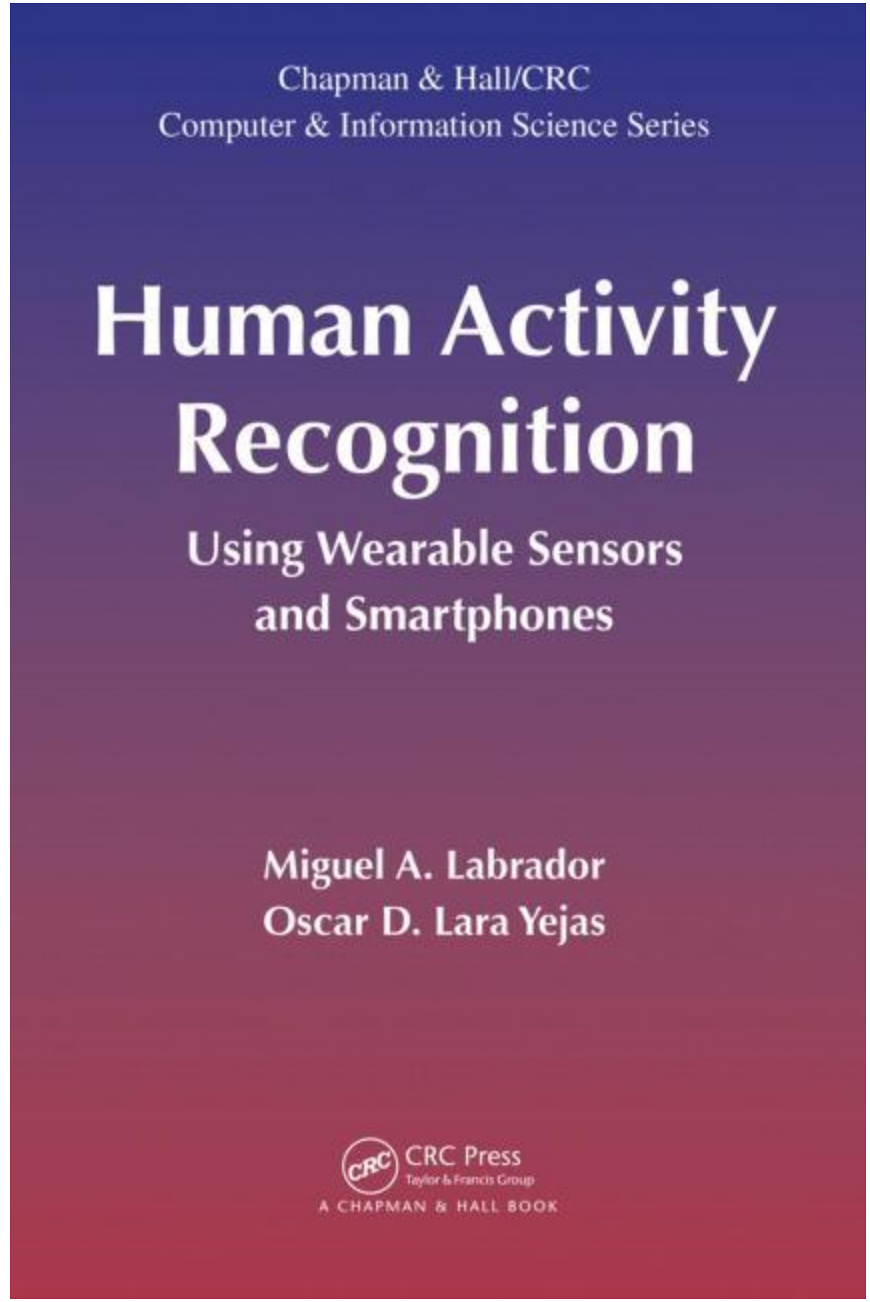

EMOTION
RECOGNITION
A Pattern Analysis Approach
Amit Konar & Aruna Chakraborty
WILEY

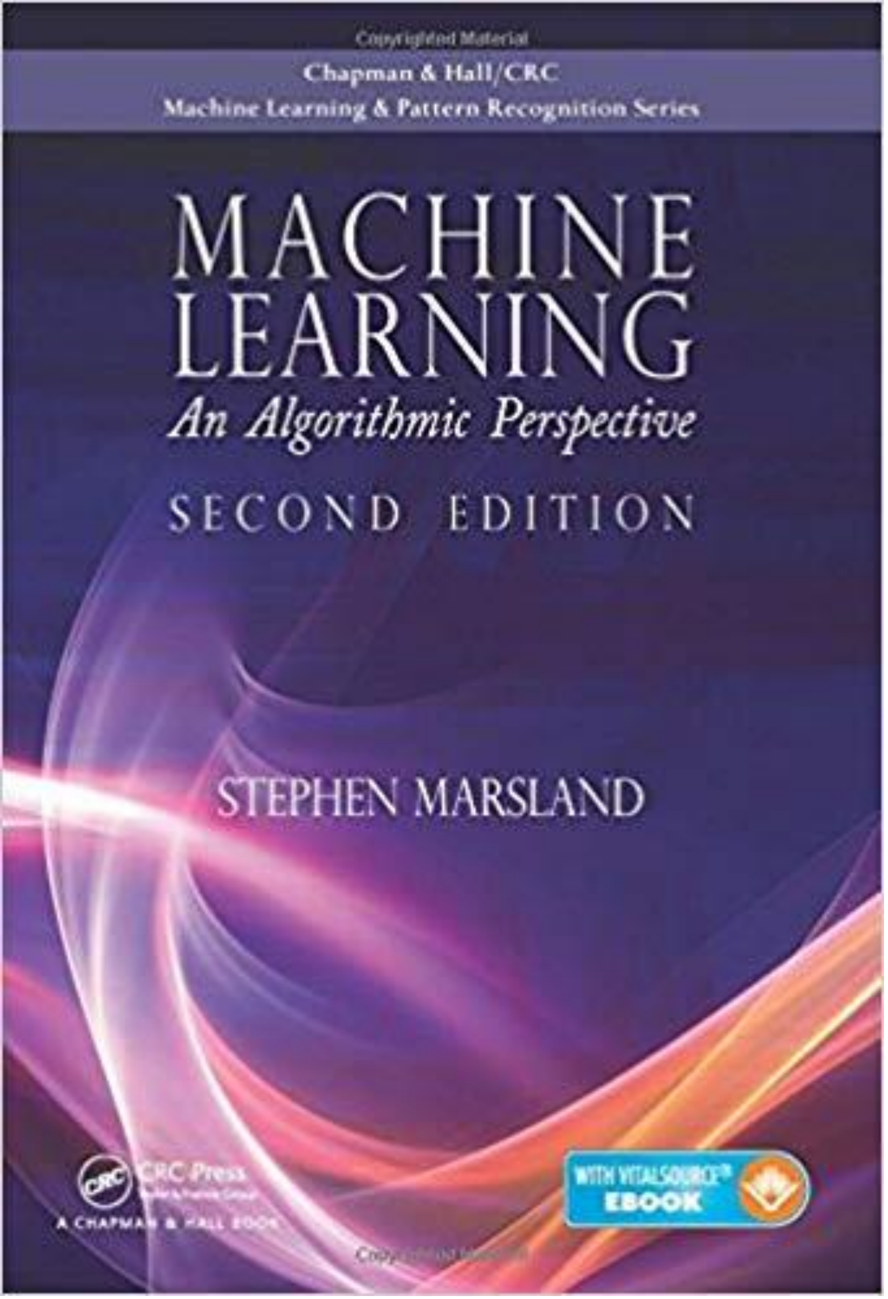

DOCUMENT ANALYSIS
AND
TEXT RECOGNITION
Benchmarking State-of-the-Art Systems
Editors
Volker Märgner
Umapada Pal
Apostolos Antonacopoulos
SERIES IN
MACHINE PERCEPTION
ARTIFICIAL INTELLIGENCE
Volume 82
World Scientific

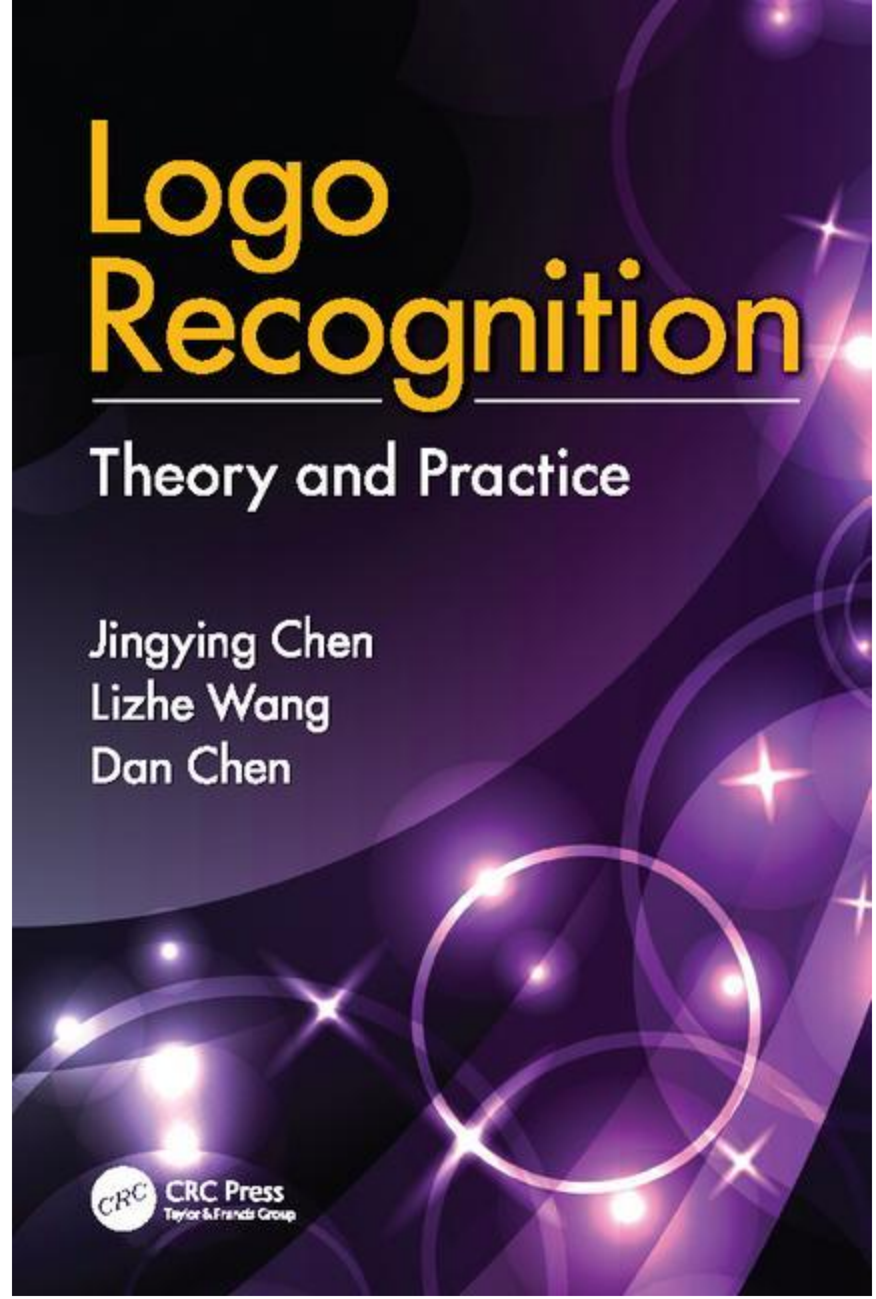

Disk Included
PATTERN
recognition
and
IMAGE
analysis
Earl Gose
Richard Johnsonbaugh
Steve Jost


Copyrighted Material
WILEY
SPEECH
RECOGNITION
Theory and C++ Implementation
Claudio Becchetti
and
Lucio Prina Ricotti
Copyrighted Material

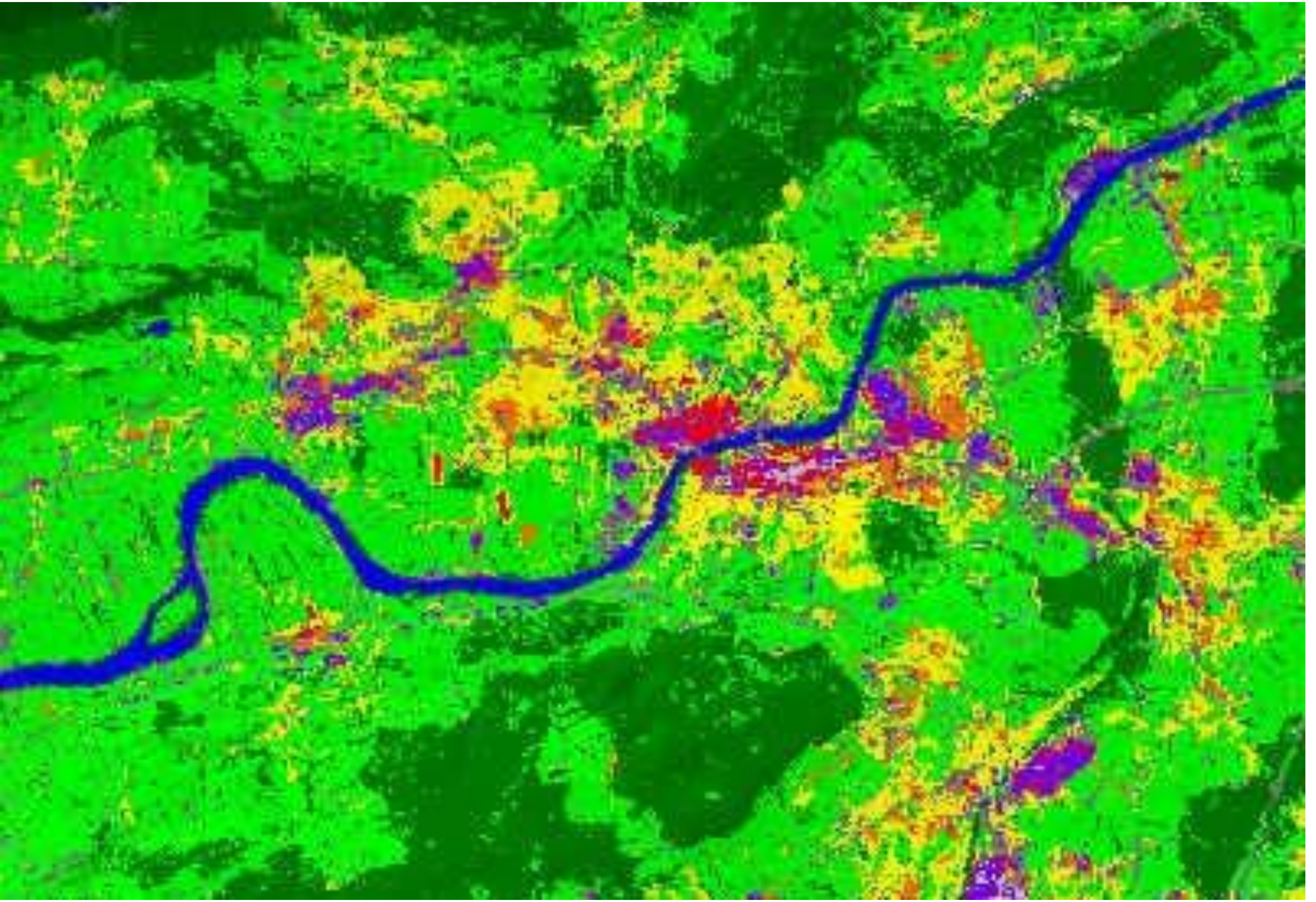

BASICS OF
MOLECULAR
RECOGNITION
Dipankar Chatterji
CRC Press
Taylor & Francis Group

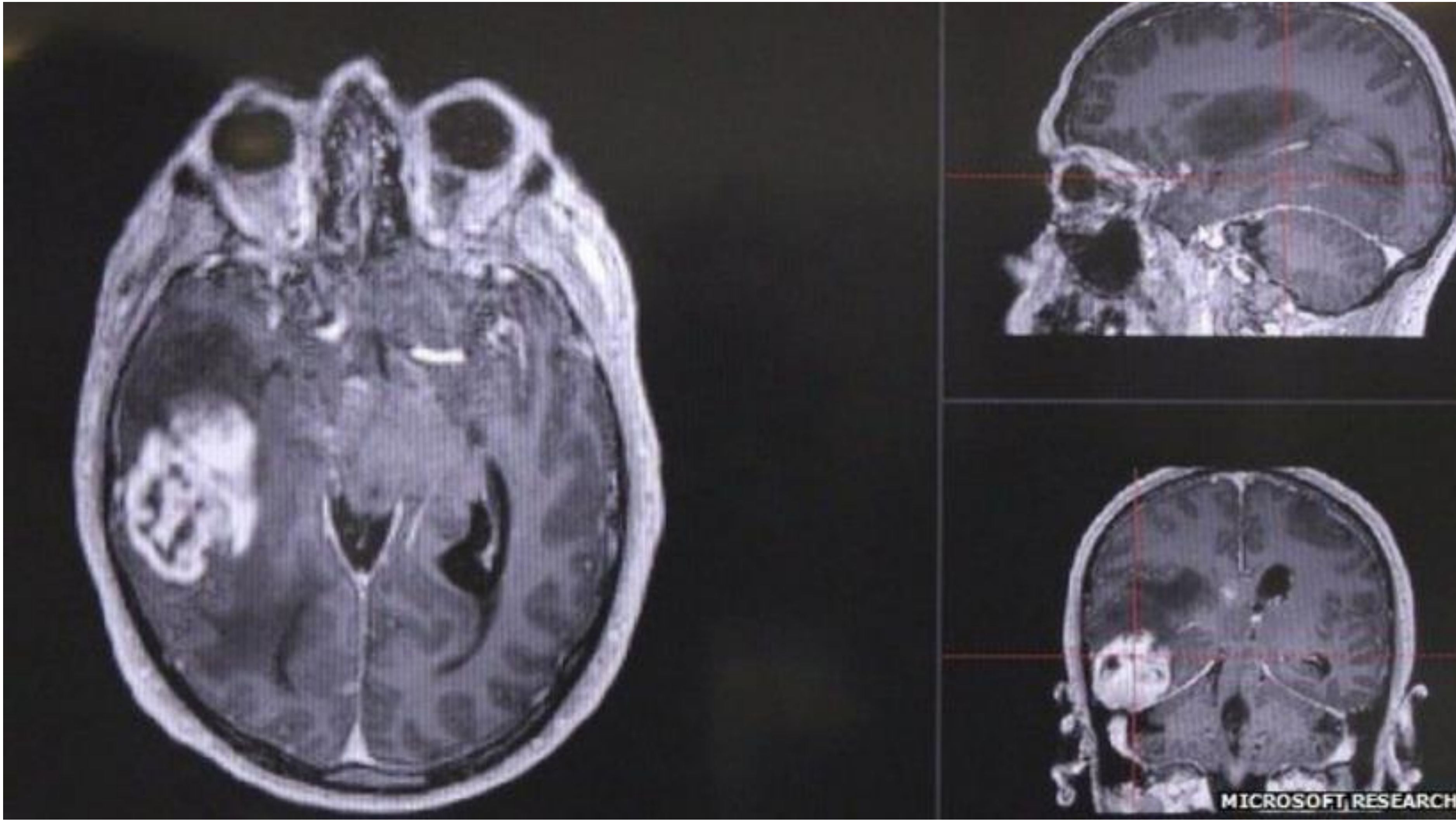

Chapman & Hall/CRC
Computer & Information Science Series
Human Activity
Recognition
Using Wearable Sensors
and Smartphones
Miguel A. Labrador
Oscar D. Lara Yejas
CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK


Copyrighted Material
Chapman & Hall/CRC
Machine Learning & Pattern Recognition Series
MACHINE LEARNING
An Algorithmic Perspective
SECOND EDITION
STEPHEN MARSLAND
CRC Press
WITH VITALSOURCE® EBOOK

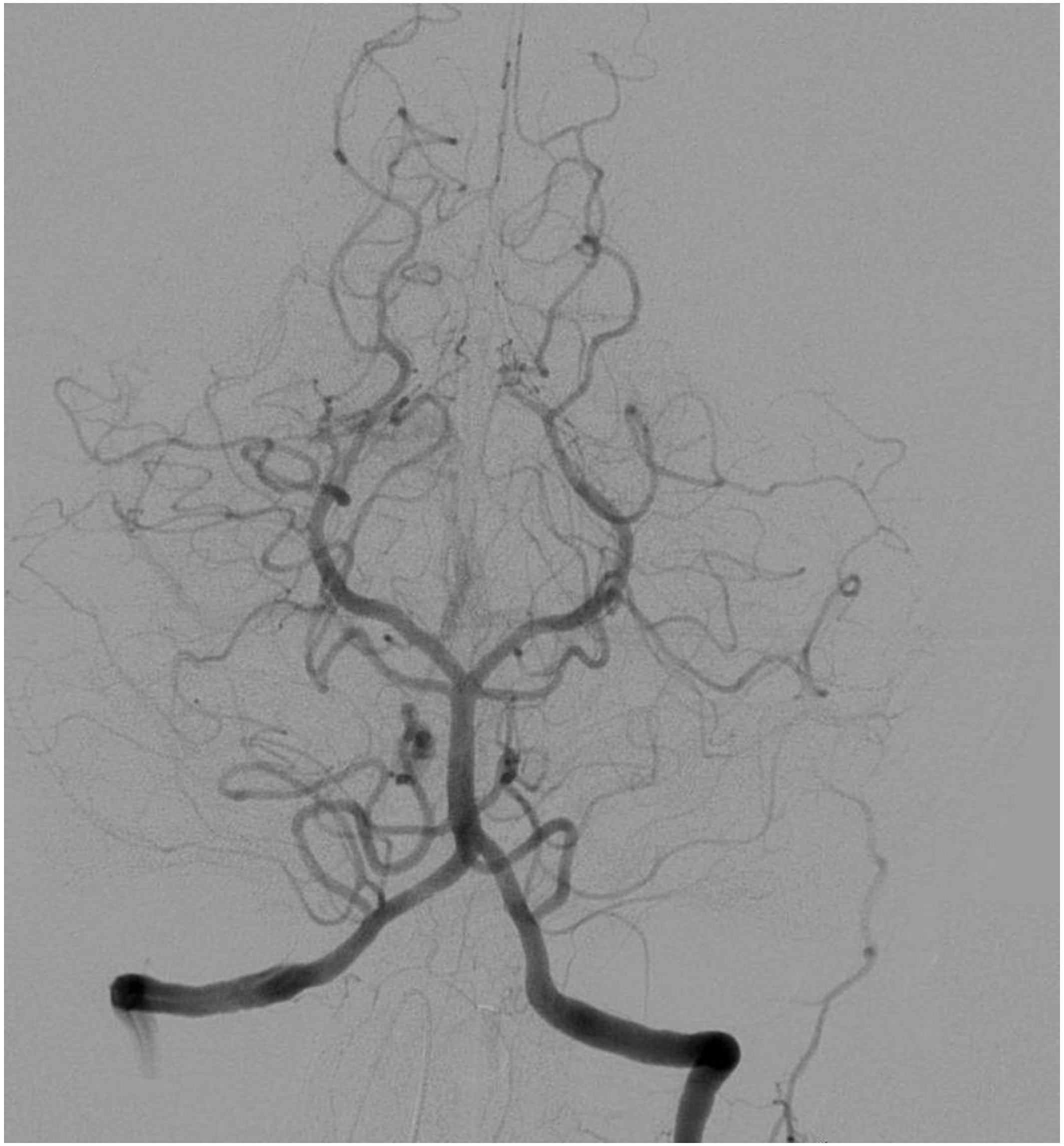

Logo Recognition
Theory and Practice
Jingying Chen
Lizhe Wang
Dan Chen
CRC Press
Taylor & Francis Group


Customer 27
Pick Up
Cereal Box
Cart
Orange $ 9
Eggs $ 5
Chicken $ 20
Milk $ 10
Subtotal $ 44





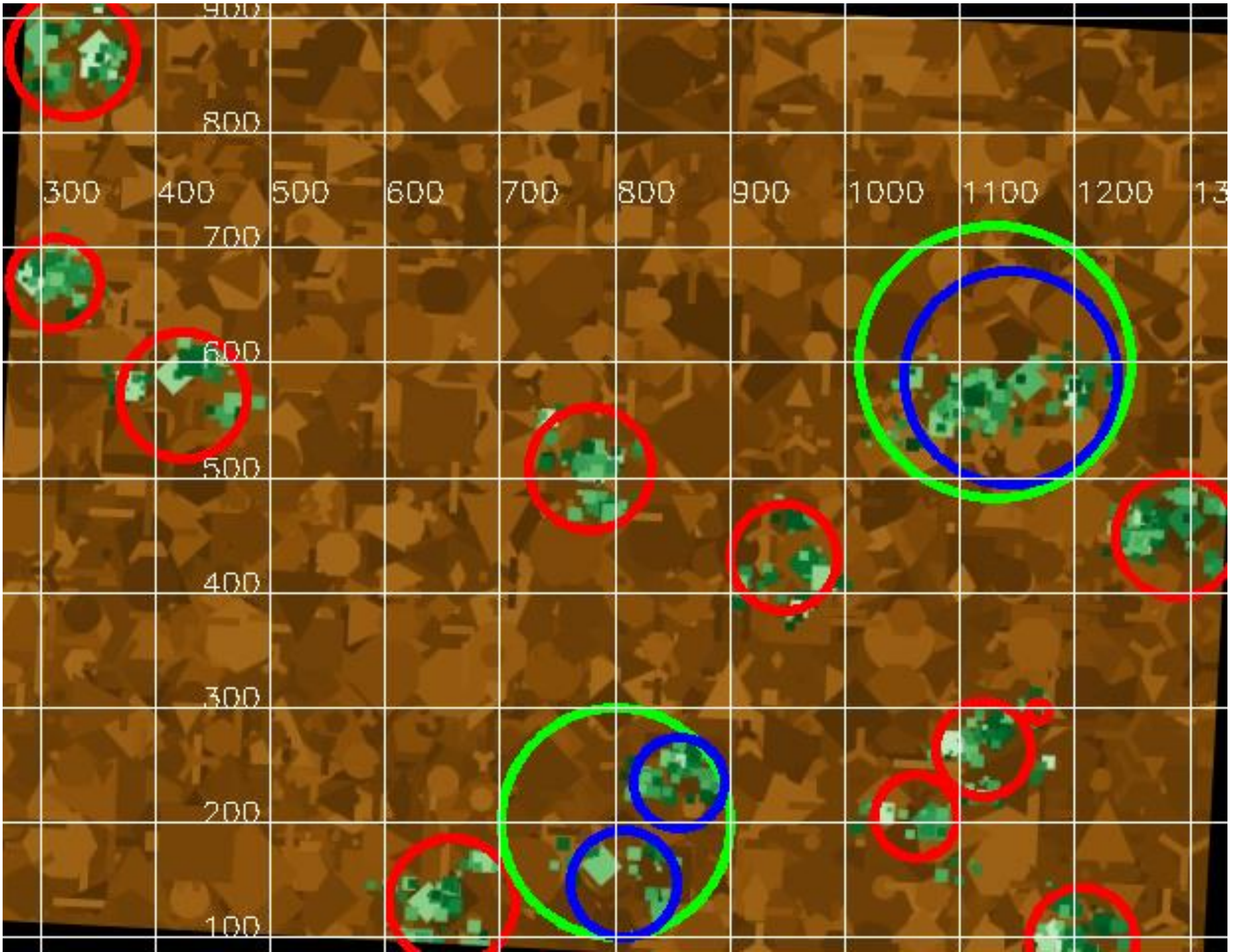
Tomography

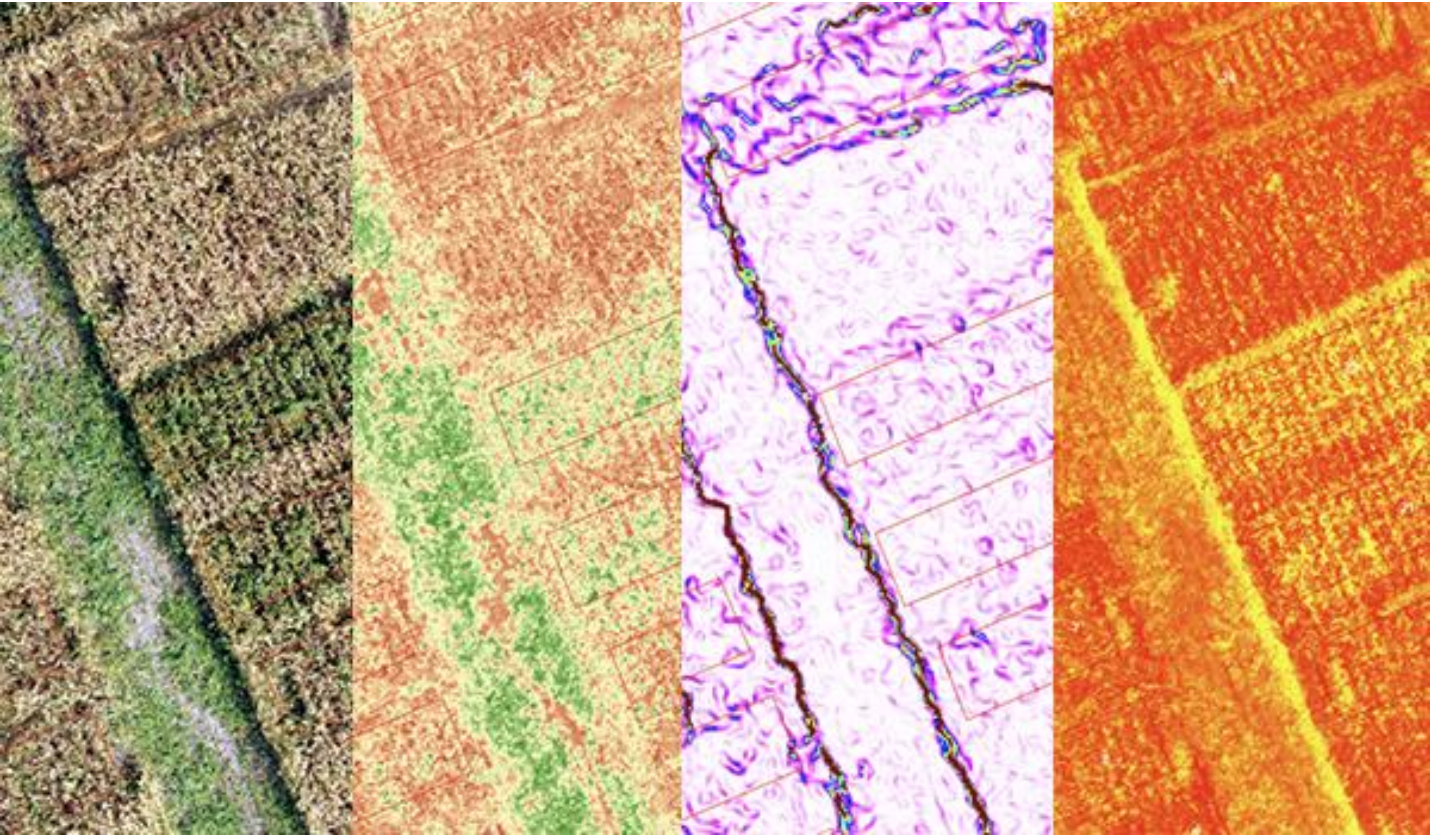


Angiogram


LP 53569
LP 53569
LP 53569
0 1 2 3 4 5 6 7 8 9

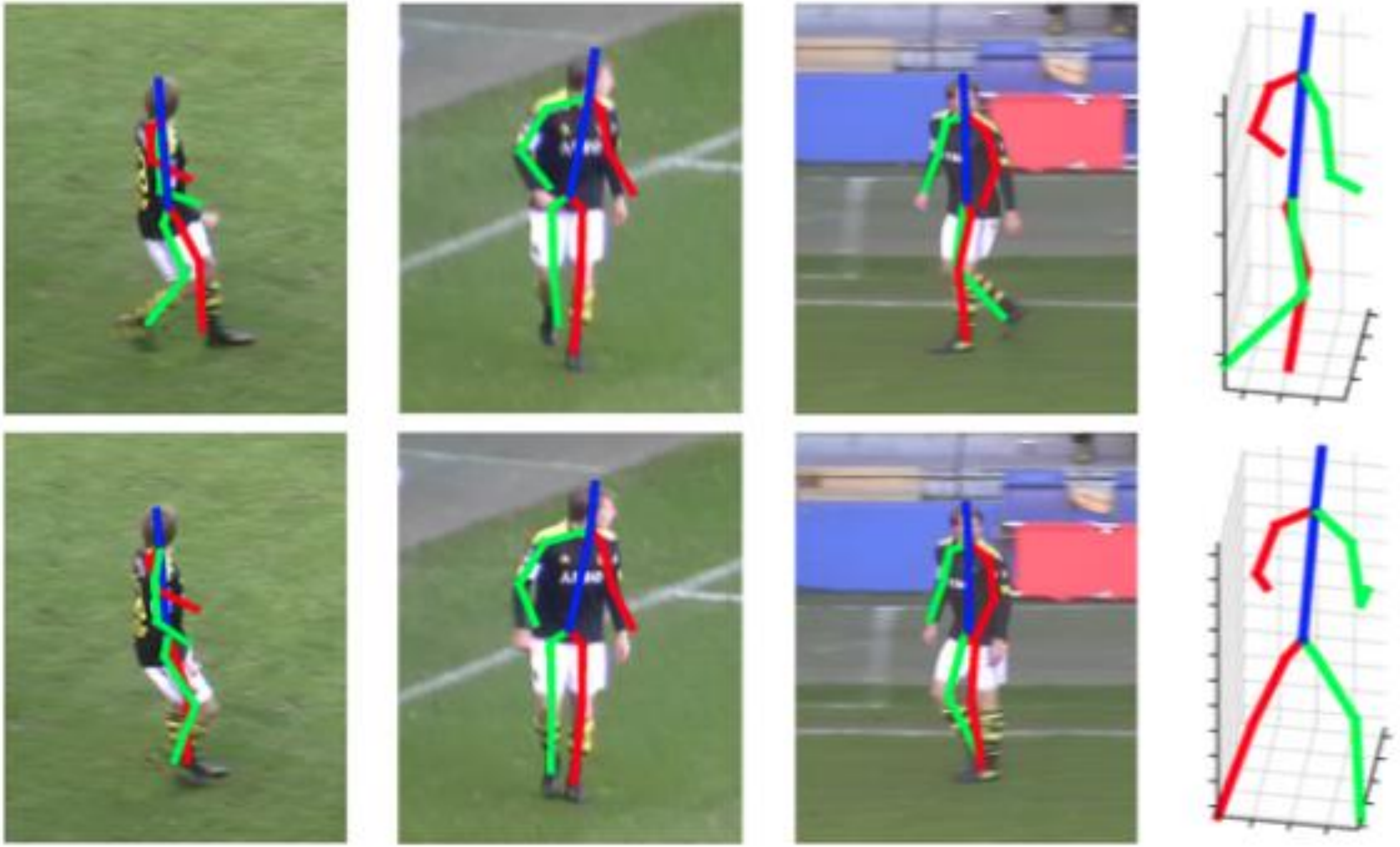

ABBEY
ROAD NW8
CITY OF WESTMINSTER


900
800
300
400
500
600
700
800
900
1000
1100
1200
700
600
500
400
300
200
100



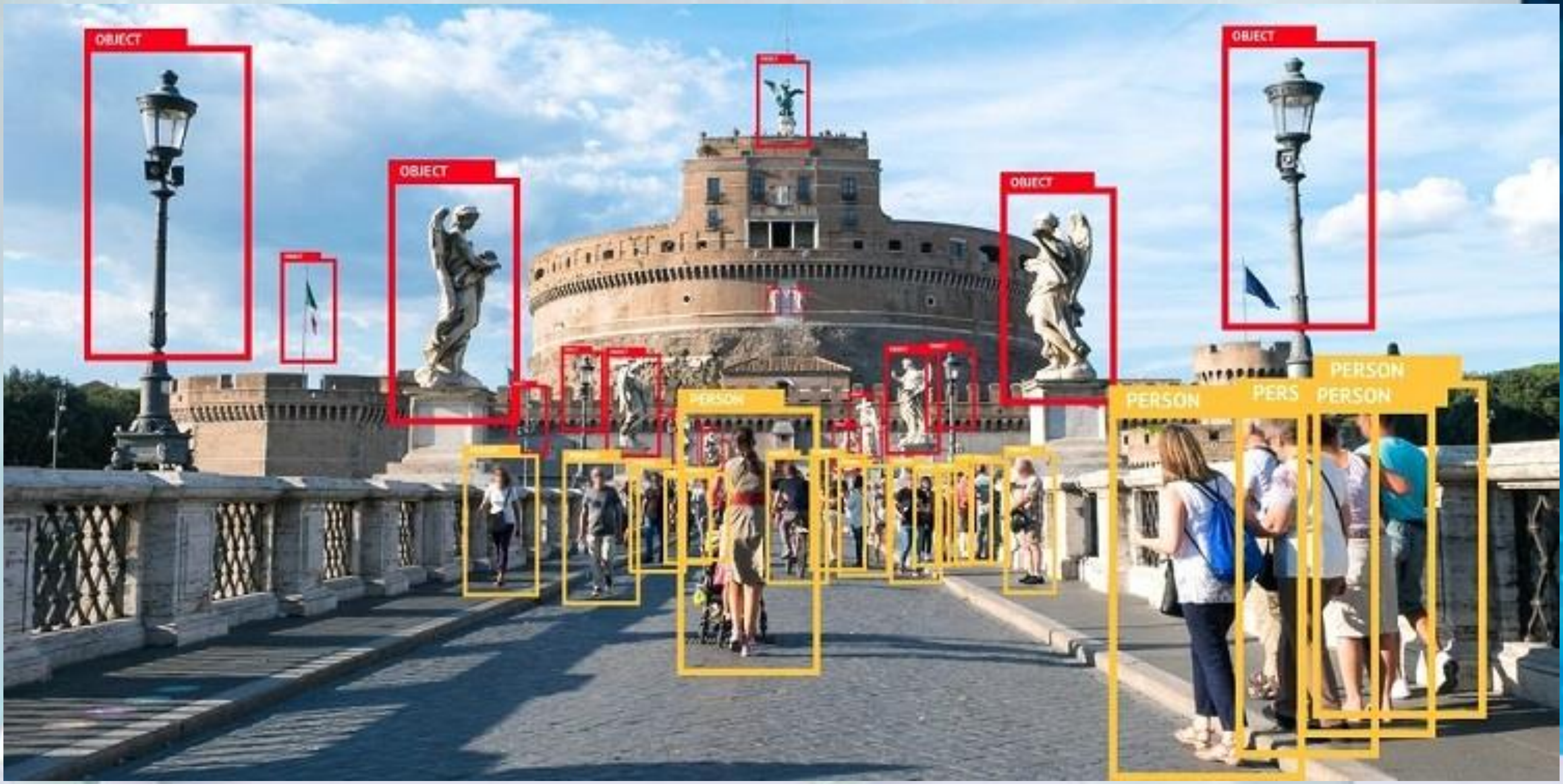
Recognition problems

Recognition problems - location

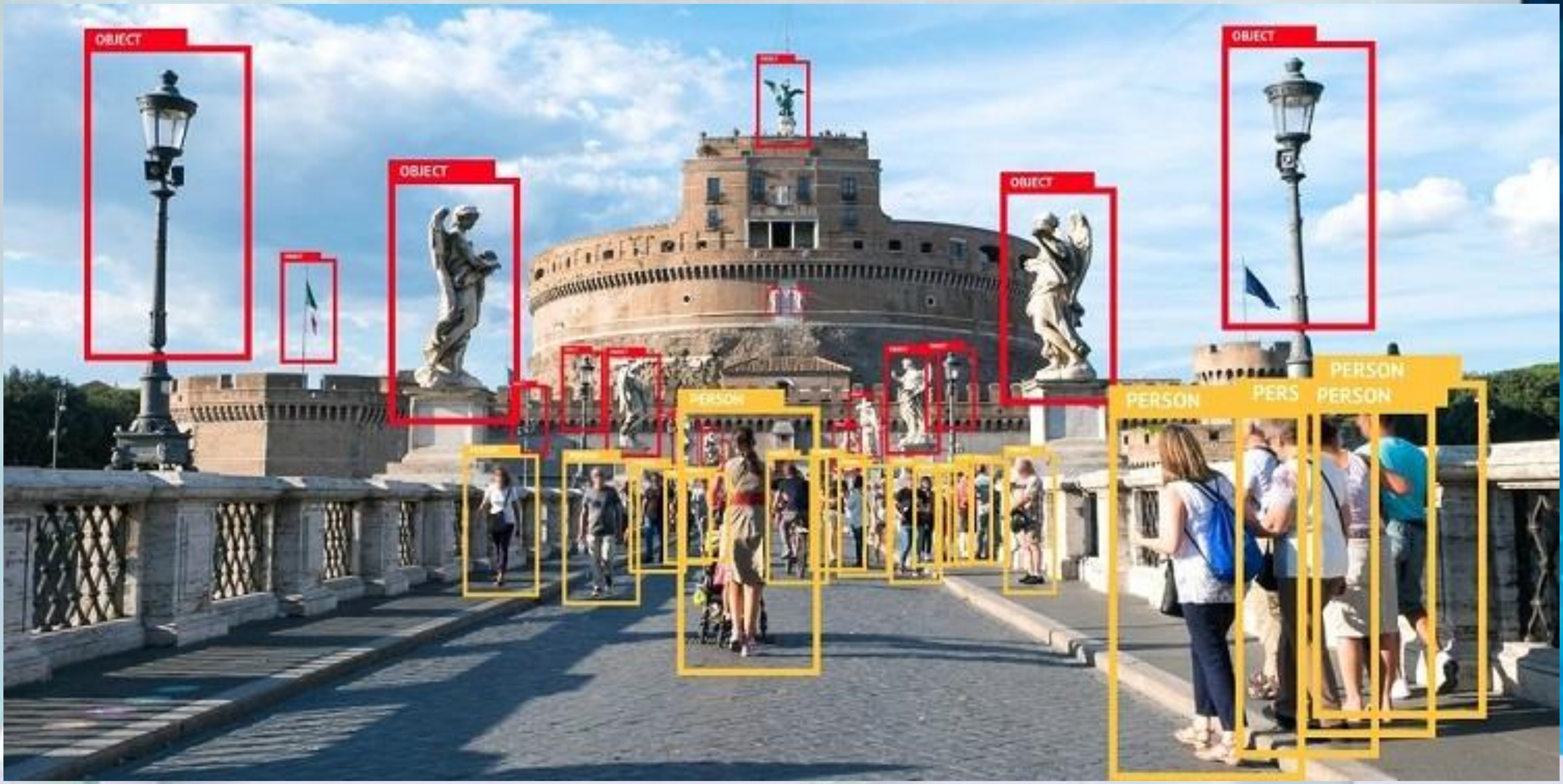


Recognition problems– color and illumination

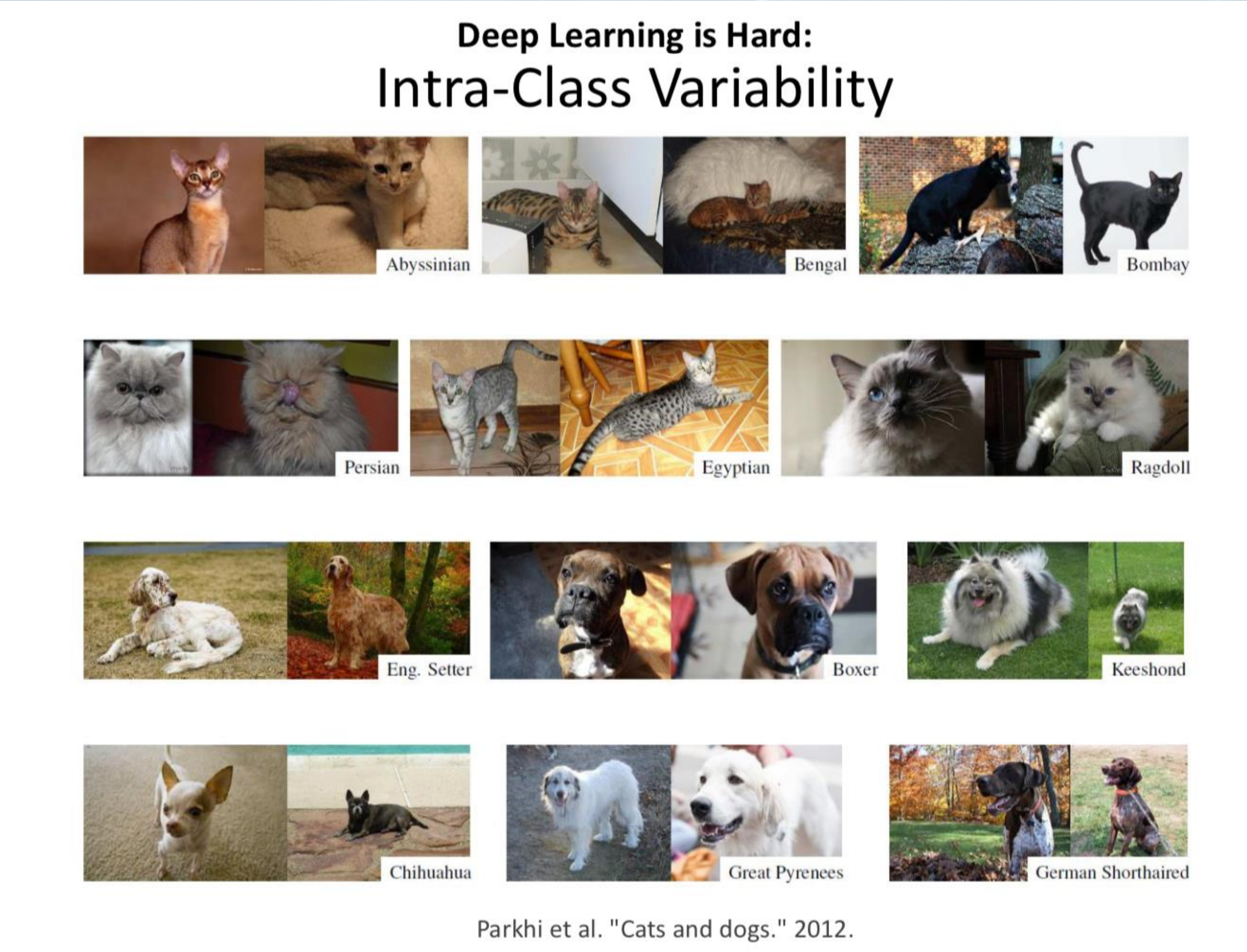
illumination– deformation and overlapping



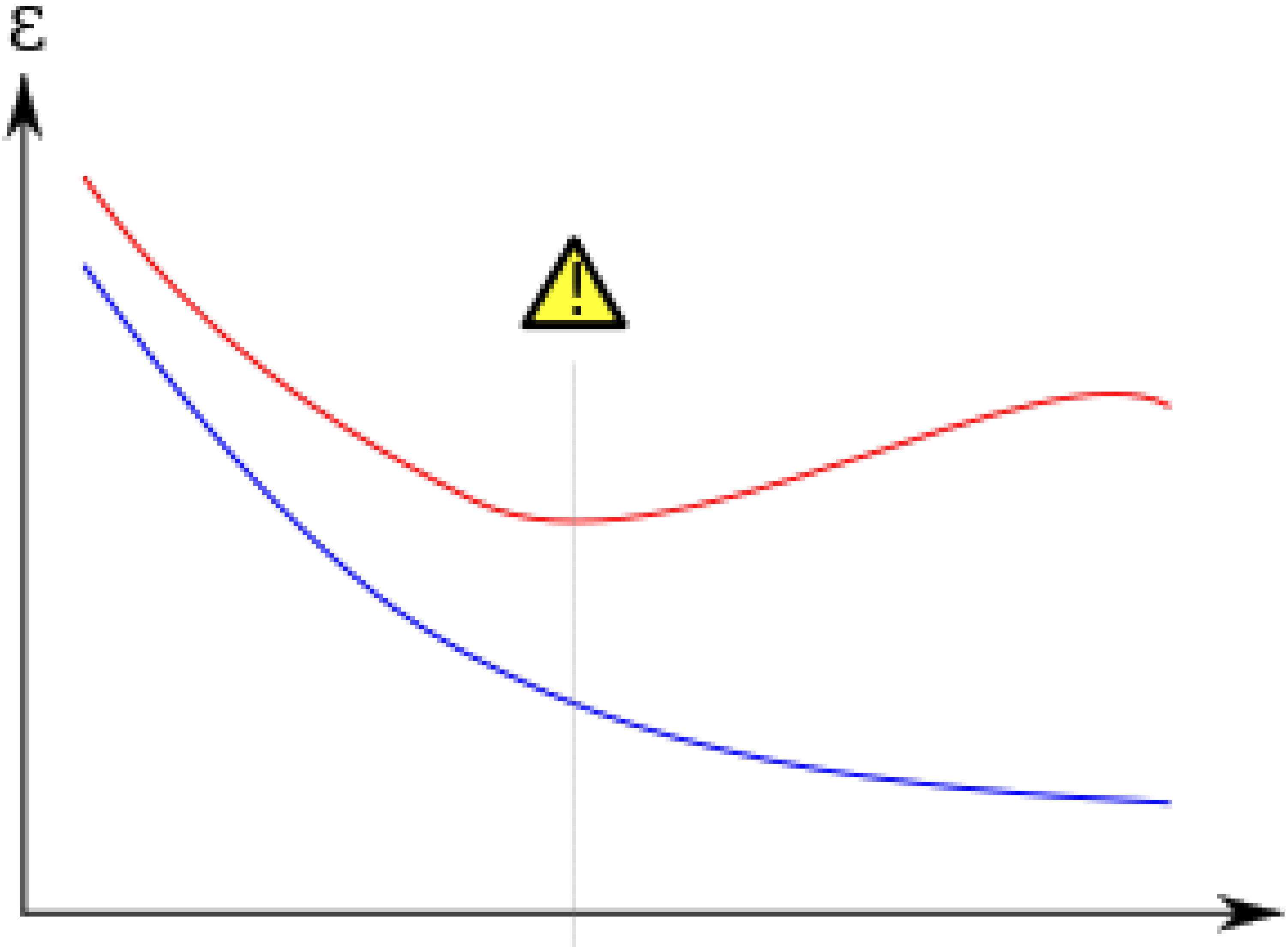
Recognition problems– moving



Recognition problems – intra-class variations



Neural network - retraining and overtraining



Three attributes that we need to describe

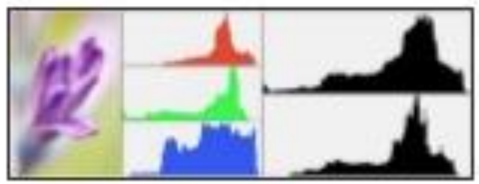
How is it done with "classic" Computer Vision?

Shape

Edge Detectors
Image Moment

Color

Color Histograms



Texture

HOG / HOF / SIFT
Fourier / Wavelet

Table 1.1: *Function modules of human and machine vision*

Task	Human vision	Machine vision
Visualization	Passive, mainly by reflection of light from opaque surfaces	Passive and active (controlled illumination) using electromagnetic, particulate, and acoustic radiation
Image formation	Refractive optical system	Various systems
Control of irradiance	Muscle-controlled pupil	Motorized apertures, filter wheels, tunable filters
Focusing	Muscle-controlled change of focal length	Autofocus systems based on various principles of distance measurements
Irradiance resolution	Logarithmic sensitivity	Linear sensitivity, quantization between 8- and 16-bits; logarithmic sensitivity
Tracking	Highly mobile eyeball	Scanner and robot-mounted cameras
Processing and analysis	Hierarchically organized massively parallel processing	Serial processing still dominant; parallel processing not in general use

- The three major criteria for the performance of computer vision algorithms are: Successful solution of task. Any practitioner gives this a top priority. But also the designer of an algorithm should define precisely for which task it is suitable and what the limits are.
- Accuracy. This includes an analysis of the statistical and systematic errors under carefully defined conditions (such as given signal-to-noise ratio (SNR), etc.).
- Speed. Again this is an important criterion for the applicability of an algorithm.

Visual perception of scenes depends on appropriate **illumination** to visualize objects. The **human visual system** is limited to a very narrow portion of the spectrum of electromagnetic radiation, called **light**. In some cases natural sources, such as solar radiation, moonlight, lightning flashes, or bioluminescence, provide sufficient ambient light to navigate our environment. Because humankind was mainly restricted to daylight, one of the first attempts was to invent an artificial light source—fire (not only as a food preparation method).

Computer vision is not dependent upon visual radiation, fire, or glowing objects to illuminate scenes. As soon as imaging detector systems became available other types of radiation were used to probe scenes and objects of interest. Recent developments in imaging sensors cover almost the **whole electromagnetic spectrum** from x-rays to radiowaves. In standard computer vision applications illumination is frequently taken as given and optimized to illuminate objects evenly with high contrast. Such setups are appropriate for object identification and geometric measurements. Radiation, however, can also be used to visualize quantitatively physical properties of objects by analyzing their interaction with radiation.

Electromagnetic radiation consists of electromagnetic waves carrying energy and propagating through space. Electrical and magnetic fields are alternating with a temporal frequency v and a spatial wavelength λ. The metric units of v and λ are cycles per second (s^{-1}), and meter (m), respectively. The unit 1 s^{-1} is also called one hertz (1 Hz). Wavelength and frequency of waves are related by the speed of light c:

$$c = v\lambda$$

The speed of light depends on the medium through which the electromagnetic wave is propagating. In vacuum, the speed of light has the value 2.9979×10^8 m s^{-1}, which is one of the fundamental physical constants and constitutes the maximum possible speed of any object. The speed of light decreases as it penetrates matter, with slowdown being dependent upon the electromagnetic properties of the medium.