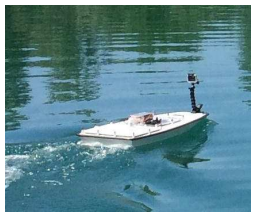
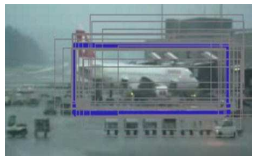
Machine Learning
Master in Artificial Intelligence and Robotics
Sapienza University of Rome







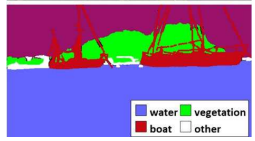
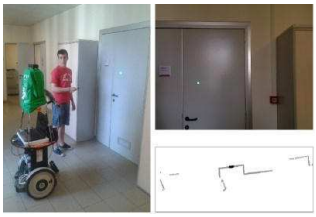
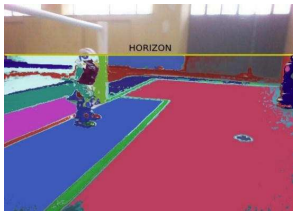


Image Classification and Detection

Domenico Daniele Bloisi

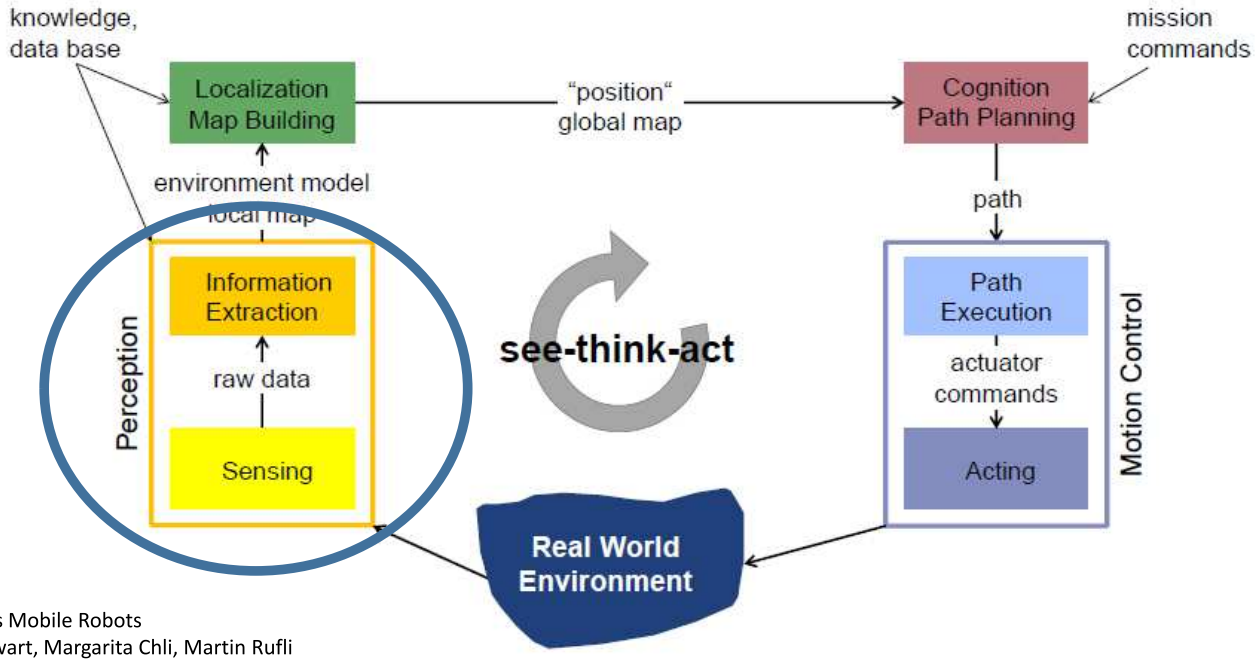
November 2018

About me

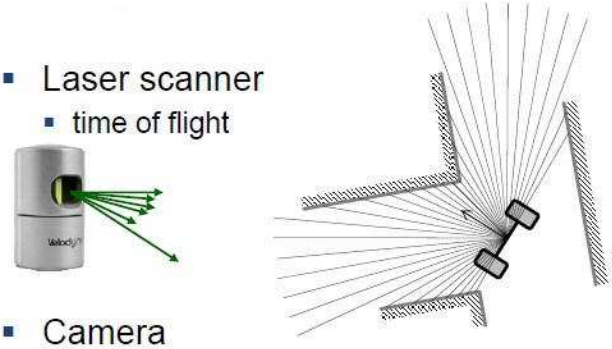
- Assistant professor with the Department of Computer Science at University of Verona
  profs.scienze.univr.it/~bloisi
- Team manager SPQR Robot Soccer Team
  www.diag.uniroma1.it/~bloisi
  spqr.diag.uniroma1.it
- Research interests: intelligent surveillance, multi-sensor data fusion, image processing, robotic vision, steganography

See-Think-Act Cycle

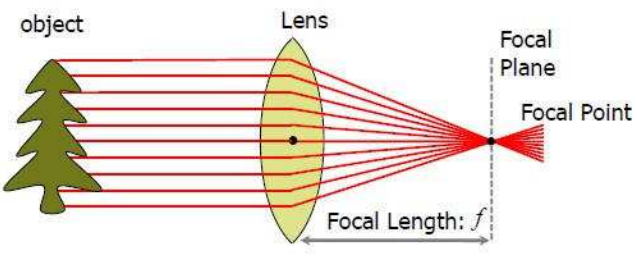


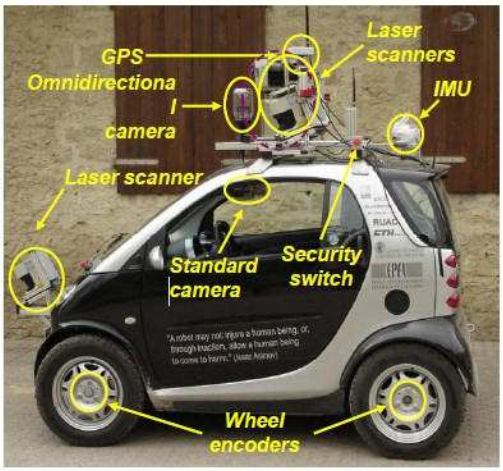
Sensors

- Laser scanner
  - time of flight

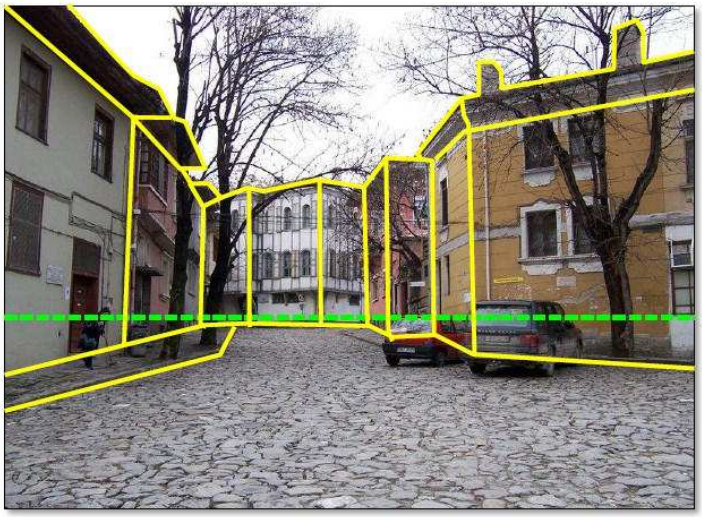


- Camera





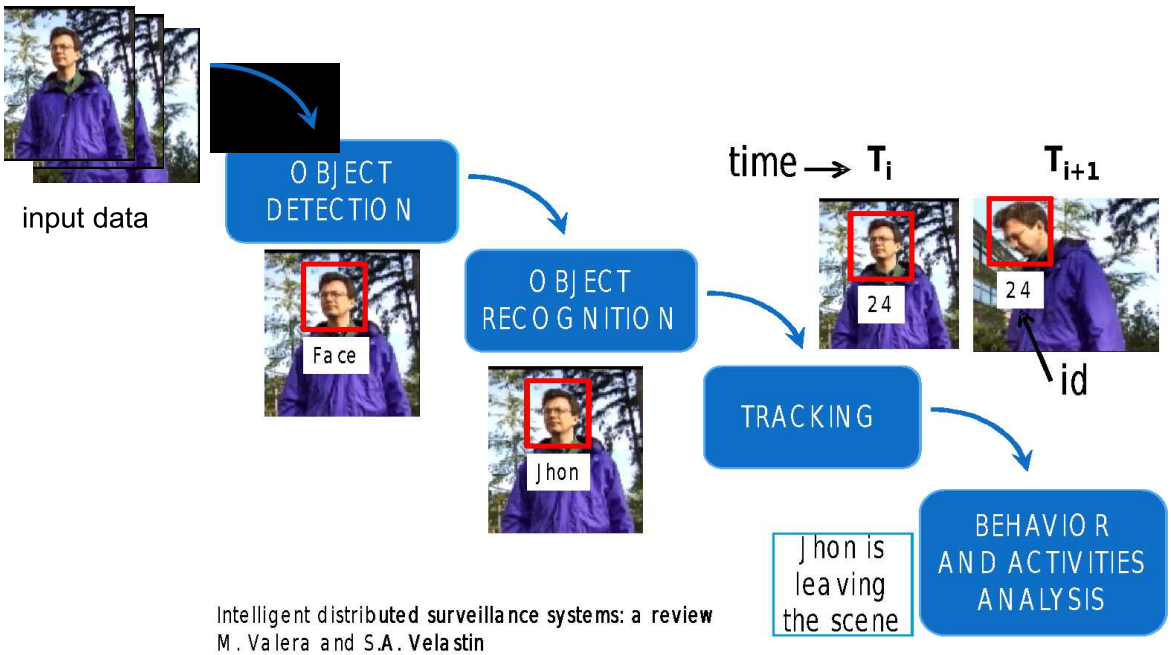
Geometrical and Semantic data



Geometrical info

Semantic info

Flow of processing visual data

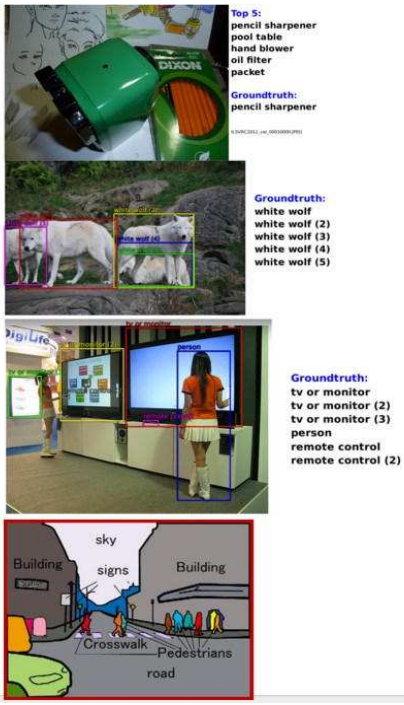


Intelligent distributed surveillance systems: a review
M. Valera and S.A. Velastin

Detecting objects in images

Classification, Localization, Detection, and Segmentation

- classification
- localization
- detection
- segmentation



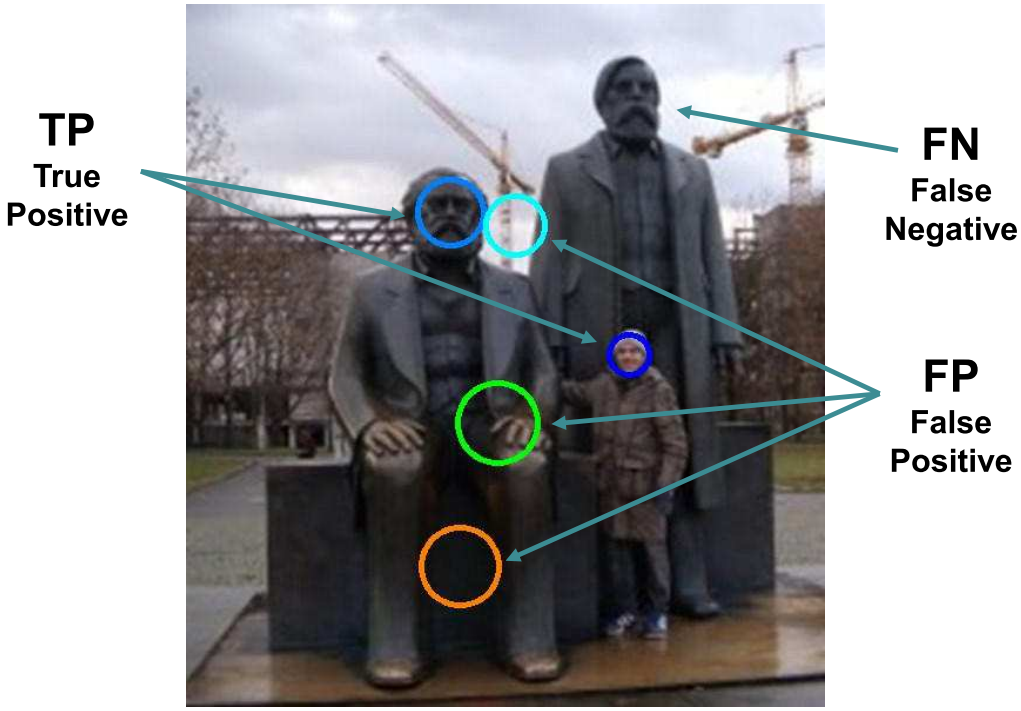
difficulty

slide from 2014 - CVPR Tutorial on Deep Learning for Vision - Object Detection

Face Detection Problem

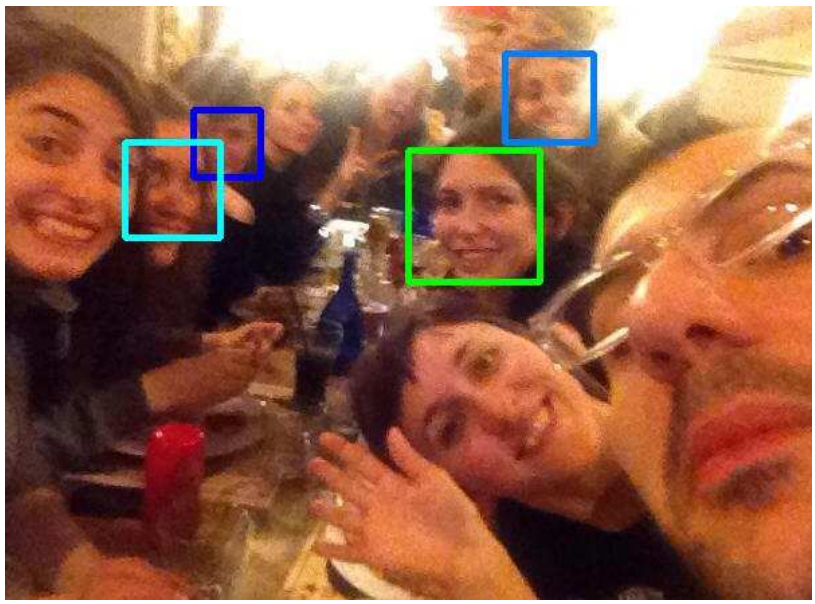
Find regions in the image that contain instances of faces

Detection issues

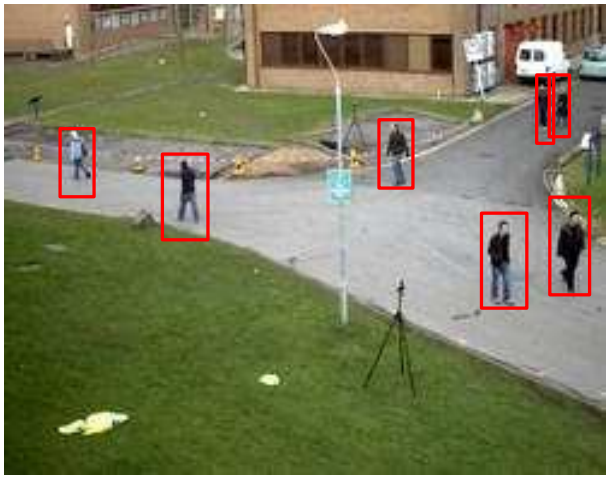


Additional issues

- Rotation
- Blurring
- Variations in illumination
- Occlusions
- Glasses
- …



Detection vs Identification

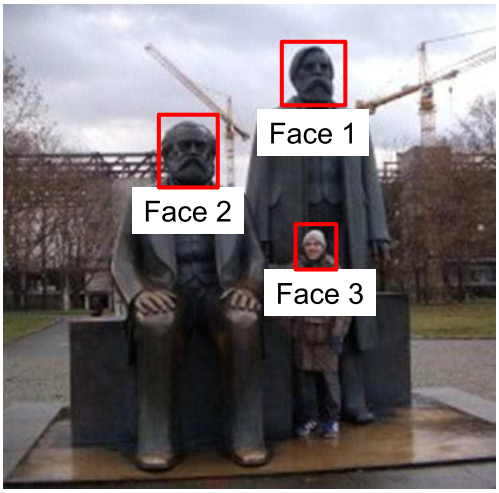


detection

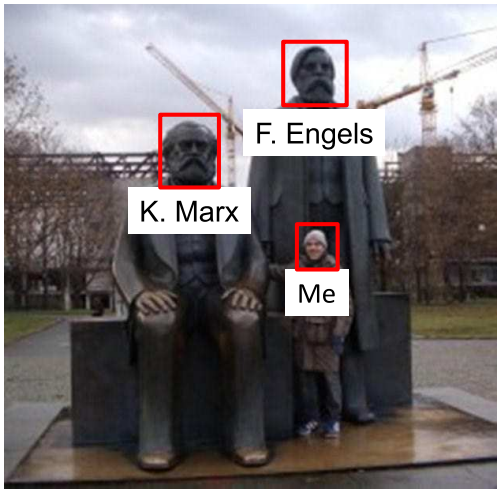
identification

PETS 2009 dataset
http://www.cvg.rdg.ac.uk/PETS2009

Detection vs Recognition

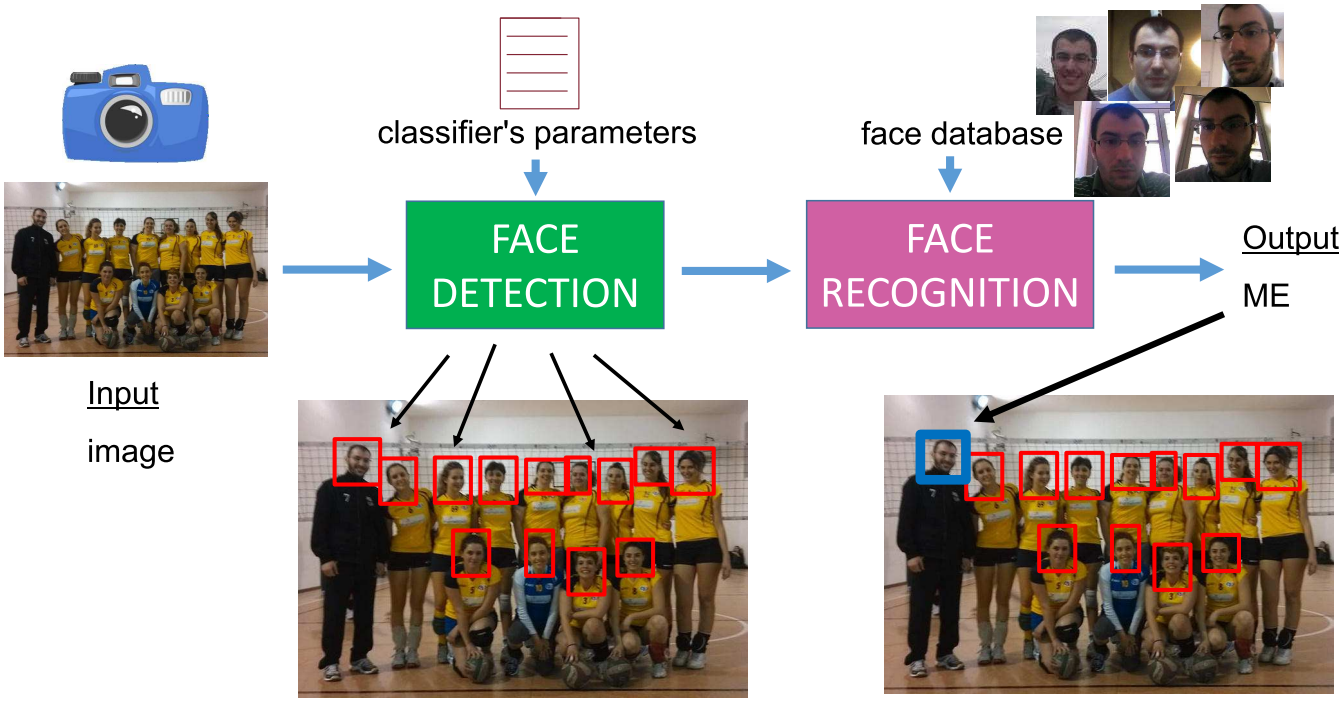


detection

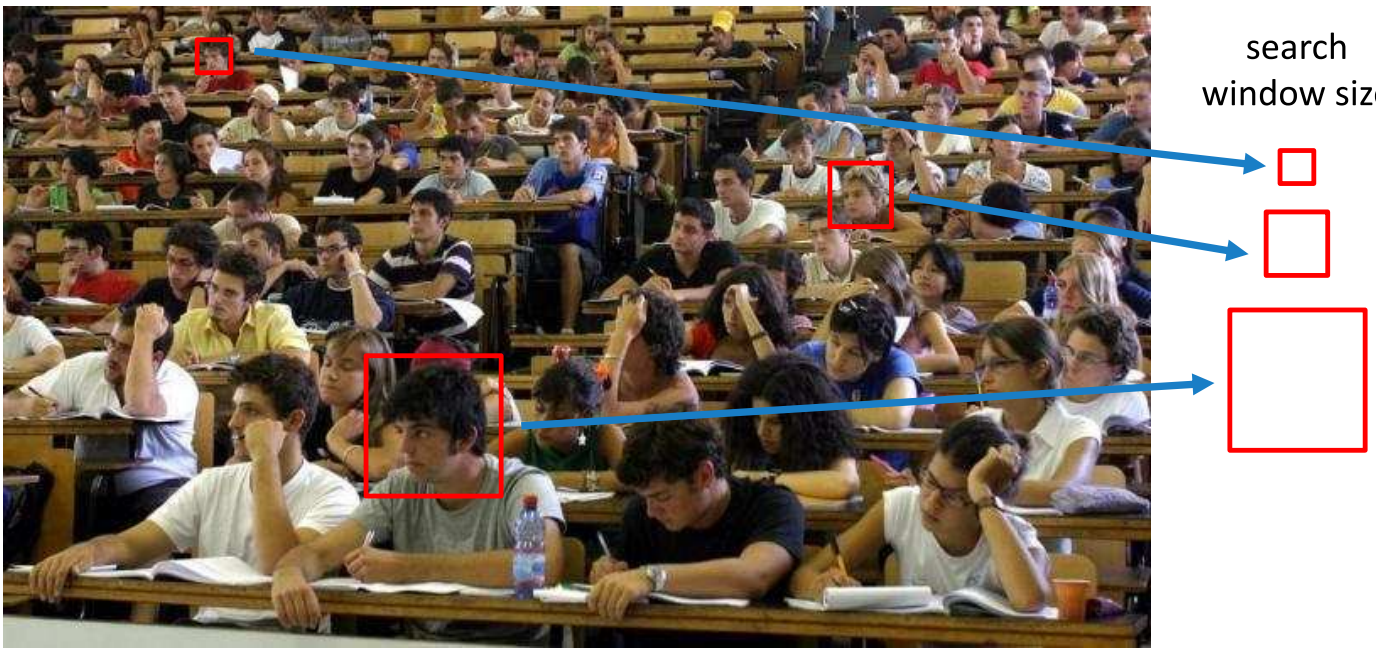


recognition

Detection & Recognition

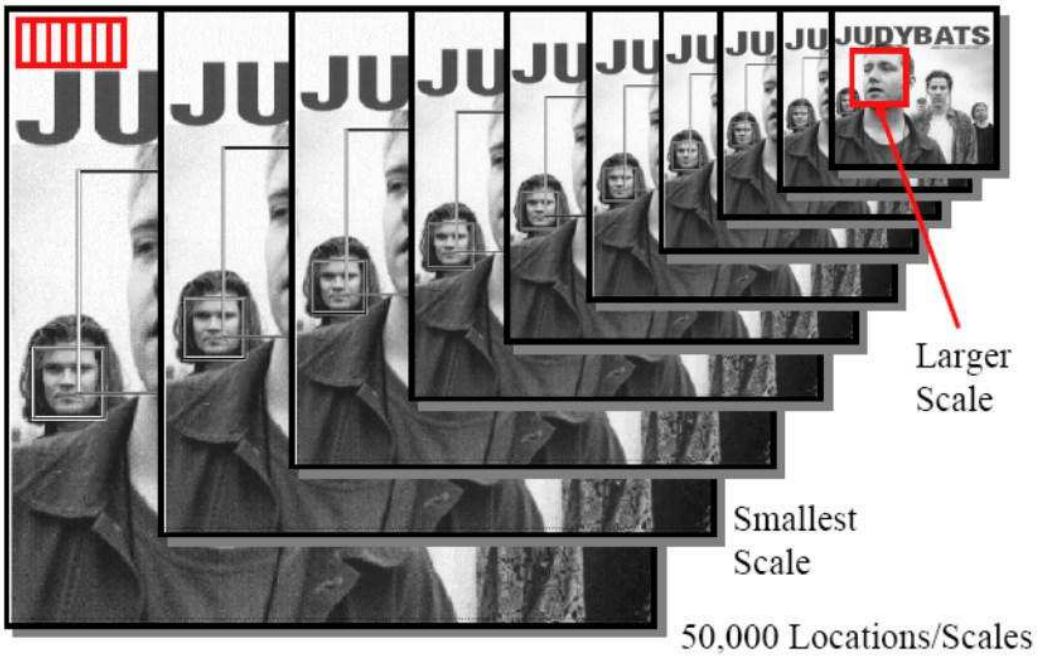


Multiscale search

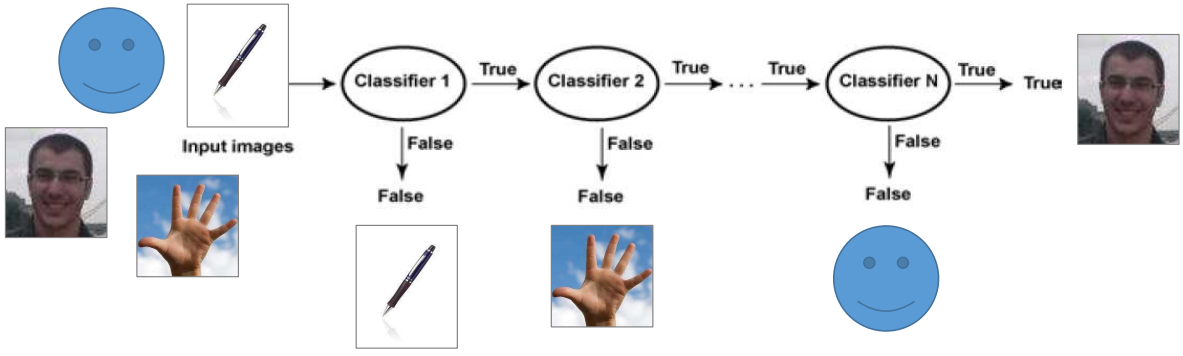


Input resizing

Image Pyramid



Cascade of classifiers



- A chain of classifiers that each reject some fraction of the negative training samples while keeping almost all positive ones
- Each classifier is an AdaBoost ensemble of rectangular Haar-like features sampled from a large pool

Training data

- Training Data
  - 5000 faces
    - All frontal, rescaled to 24x24 pixels
  - 300 million non-faces
    - 9500 non-face images
  - Faces are **normalized**
    - Scale, translation
- **Many variations**
  - Across individuals
  - Illumination
  - Pose

Deep Face

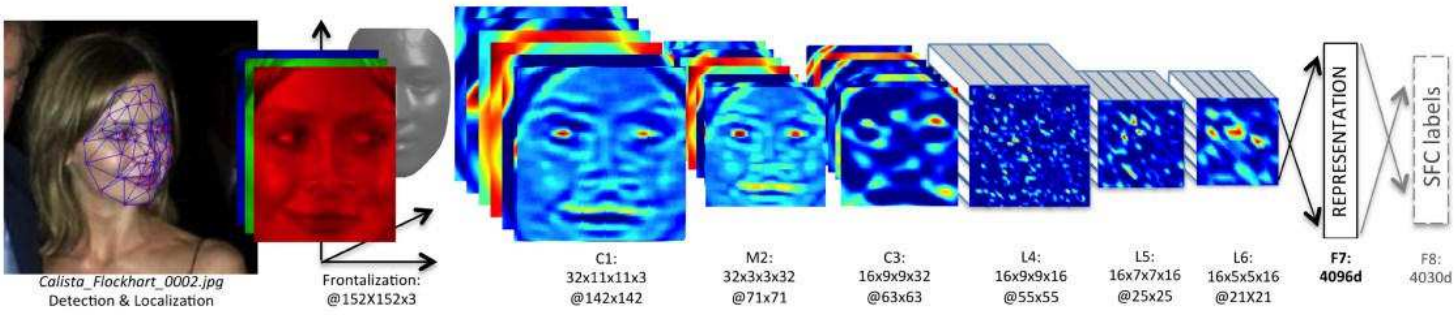


Figure 2. **Outline of the *DeepFace* architecture**. A front-end of a single convolution-pooling-convolution filtering on the rectified input, followed by three locally-connected layers and two fully-connected layers. Colors illustrate feature maps produced at each layer. The net includes more than 120 million parameters, where more than 95% come from the local and fully connected layers.

This figure from
Y. Taigman, M. Yang, M. Ranzato, L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," in IEEE Conference on Computer Vision and Pattern Recognition, pp. 1701-1708, 2014

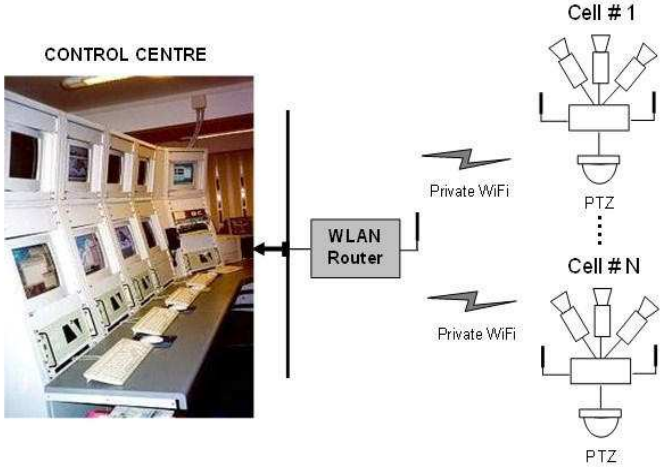
Ball detection in RoboCup SPL Soccer

https://youtu.be/ji0OmkaWh20

ARGOS system

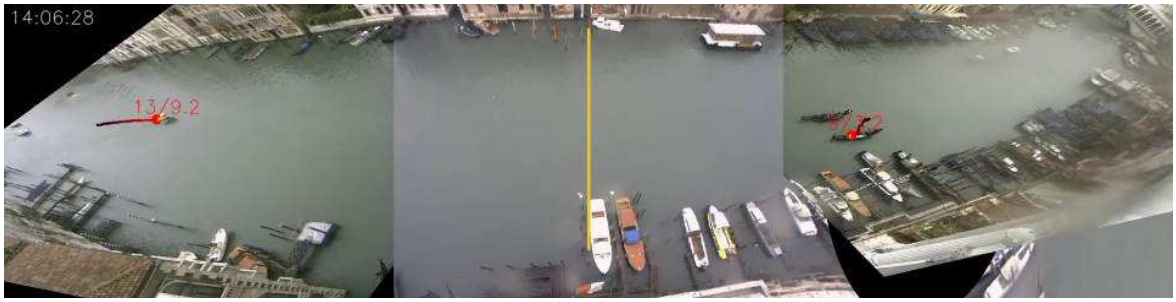
Automatic Remote Grand Canal Observation System

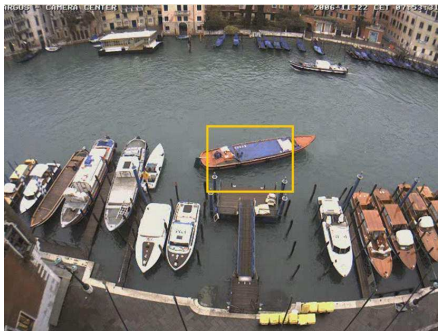
The *ARGOS* system controls a waterway of about 6 km length, 80 to 150 meters width, through 14 observation posts (Survey Cells)



ARGOS system: boat tracking

https://youtu.be/9a70Ucgbi_U



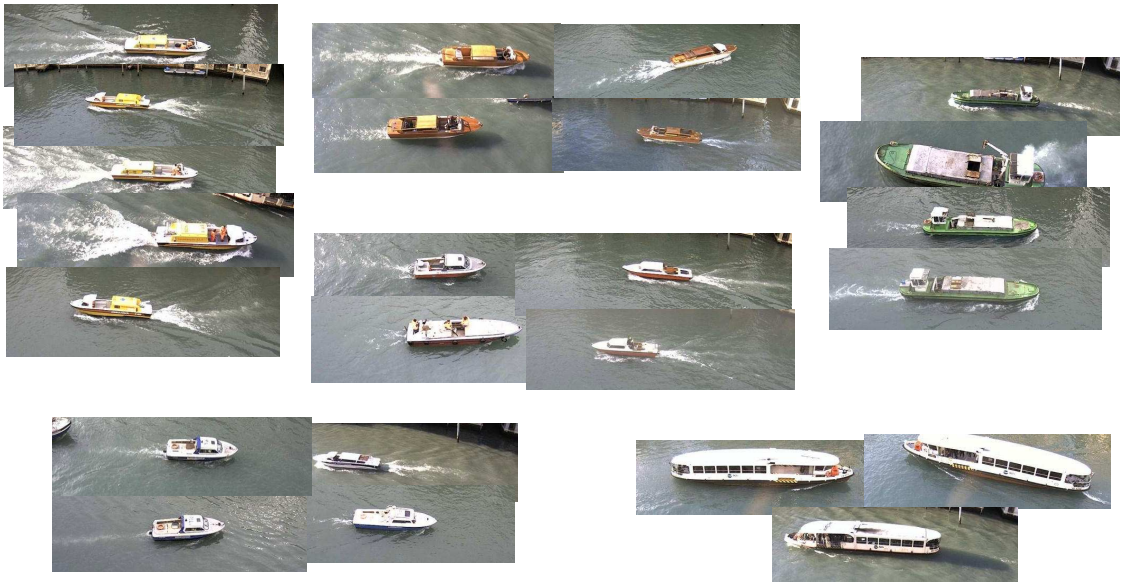


Detecting boats docking in the highlighted area

Speed limit control

ARGOS captures screenshots of the boats when passing through the yellow line

Boat Classification



Boat Categories in Venice

1. Alilaguna
2. Ambulanza
3. Barchino
4. Cacciapesca
5. Caorlina
6. Gondola
7. Lanciafino10m
8. Lanciafino10mBianca
9. Lanciafino10mMarrone
10. Lanciamaggioredi10m
11. Lanciamaggioredi10mBianca
12. Lanciamaggioredi10mMarrone
13. Motobarca
14. Motopontonerettangolare
15. MotoscafoACTV
16. Mototopo
17. Patanella
18. Polizia
19. Raccoltarifiuti
20. Sandoloaremi
21. Sanpierota
22. Topa
23. VaporettoACTV
24. VigilidelFuoco

MarDCT dataset

http://www.diag.uniroma1.it/~labrococo/MAR/

MarDCT available files for detection

Data	Sensor	Location	Time	Camera	Reflections	View	GT Data
venice-1.m4v	EO	Venice, Italy	Afternoon	Static	Medium	Center	
venice-2.m4v	EO	Venice, Italy	Morning	Static	High	Center	
venice-3.m4v	EO	Venice, Italy	Morning	Static	High	Center	
venice-4.m4v	EO	Venice, Italy	Morning	Static	Low	Right	
venice-5.m4v	EO	Venice, Italy	Morning	Static	Low	Center	
venice-6.m4v	EO	Venice, Italy	Afternoon	Static	Low	Right	
venice-7.m4v	EO	Venice, Italy	Afternoon	Static	Medium	Left	
venice-8.m4v	EO	Venice, Italy	Afternoon	Static	Medium	Center	
venice-9.m4v	EO	Venice, Italy	Afternoon	Static	High	Right	
venice-10.m4v	EO	Venice, Italy	Night	Static	High	Center	
ir-1.avi	IR	Northern Europe	Night	Static	Low	Center	
ir-2.avi	IR	Northern Europe	Night	Static	Low	Center	
ir-3.avi	IR	Northern Europe	Night	Static	Low	Center	
IR-validation.zip	IR	Northern Europe	Night	Static	Low	Center	
reflections-1.avi	EO	Italy	Day	Static	High	Center	
reflections-2.avi	EO	Italy	Day	Static	High	Center	
reflections-3.avi	EO	Italy	Day	Static	High	Center	
italy-occlusions-1.avi	EO	Italy	Day	Static	Low	Center	

Examples

References and Credits

- P. Sermanet, “Object Detection with Deep Learning”
- K.H. Wong. “Ch. 6: Face detection”
- P. Viola and T.-W. Yue. “Adaboost for Face Detection”
- D. Miller. “Face Detection & Synthesis using 3D Models & OpenCV”
- S. Lazebnik. “Face detection”
- C. Schmid. “Category-level localization”
- C. Huang and F. Vahid. “Scalable Object Detection Accelerators on FPGAs
- Using Custom Design Space Exploration”
- P. Smyth. “Face Detection using the Viola-Jones Method”
- K. Palla and A. Kalaitzis. “Robust Real-time Face Detection”

Machine Learning

Master in Artificial Intelligence and Robotics

Sapienza University of Rome





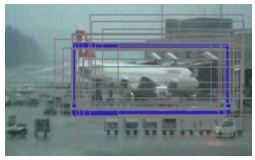
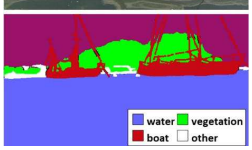
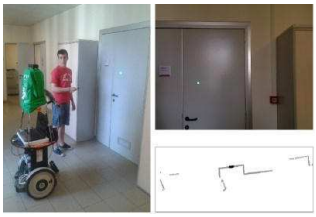
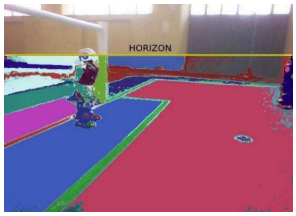


Image Classification and Detection

Domenico Daniele Bloisi





November 2018