TSM_CompVis Computer Vision

Prof. Dr. Thomas Koller

Prof. Dr. Alessandro Giusti

Your instructors

`getShortCV("Thomas Koller", "MSE")`

- Dipl. Informatik Ing. ETH
  - Assistent for Computer Graphics at IPS, ETH
- Dr. sc. techn. ETH Zürich, Computer Vision Lab
  - 3D Medical Image Processing
- Bitplane AG
  - Head of R&D, 3D Image Processing and Visualization
- Zühlke Engineering:
  - Software Engineering Consultant, Project Leader, Architect, Trainer
- HSLU:
  - Computer Vision, Deep Learning, Computer Graphics WebProgramming Lab, MSE Advisor,

Alessandro Giusti

- Website
  https://alessandro-giusti.github.io/
- PhD Computer Science
  Politecnico di Milano
- Professor, Robotics Lab lead
  Dalle Molle Institute for AI (IDSIA), USI–SUPSI, Lugano
  https://idsia-robotics.github.io/
- Applied research projects:
  Computer Vision, Robotics, Applied ML
- 80+ scientific papers
- Lecturer
  USI (Robotics, Machine Learning)
  SUPSI (Data Science)
  Politecnico di Milano (Deep Learning)

idsia

SUPSI

Course Logistics

Logistics

- Moodle: you must be enrolled to the course!
- Materials published normally 48h before class
- 2h lecture + exercise slot (with assistance) after class

Syllabus

- Lecturer: Alessandro Giusti
  - Image Basics. Binarization. Local Filtering. Connected Component Analysis
  - Model Fitting. Hough Transform.
  - Image Classification Pipelines. Data Augmentation.

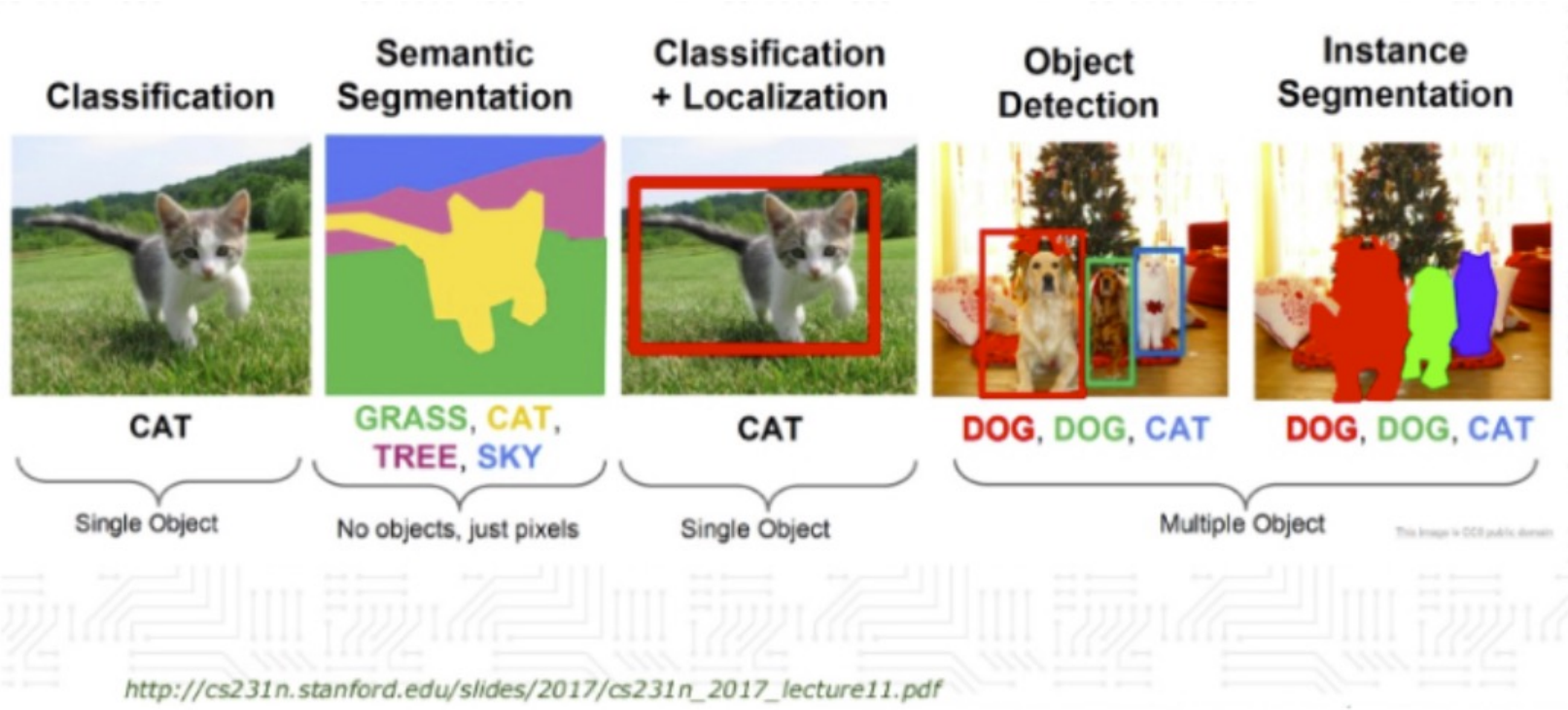
- Lecturer: Thomas Koller
  - Semantic Segmentation
  - Object Detection
  - Image Synthesis with Deep Learning

Evaluation

- Two homeworks during the course
- Final moodle exam (quiz) during exam session

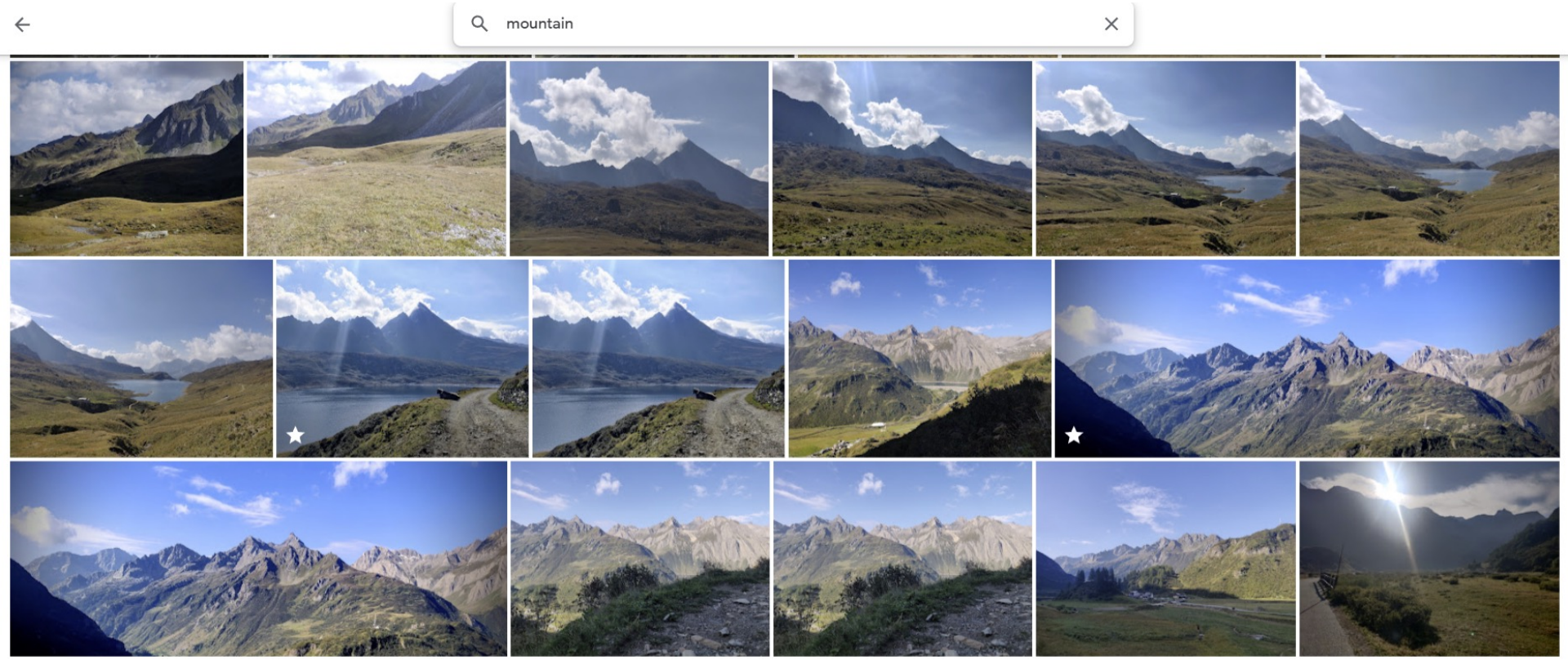
An overview of computer vision and its applications

High-level tasks in Computer Vision

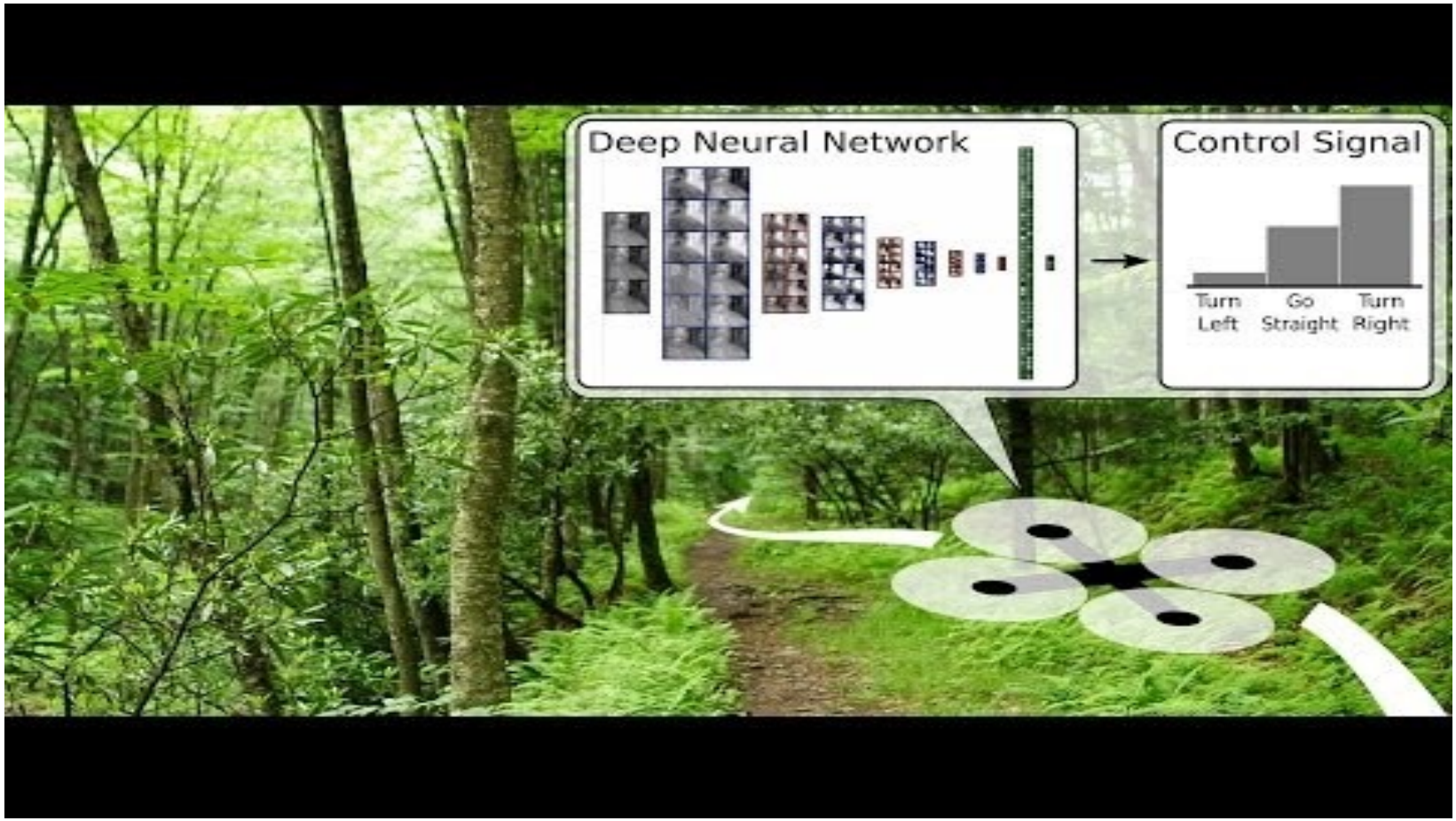


http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf

An application of classification (in google photos)

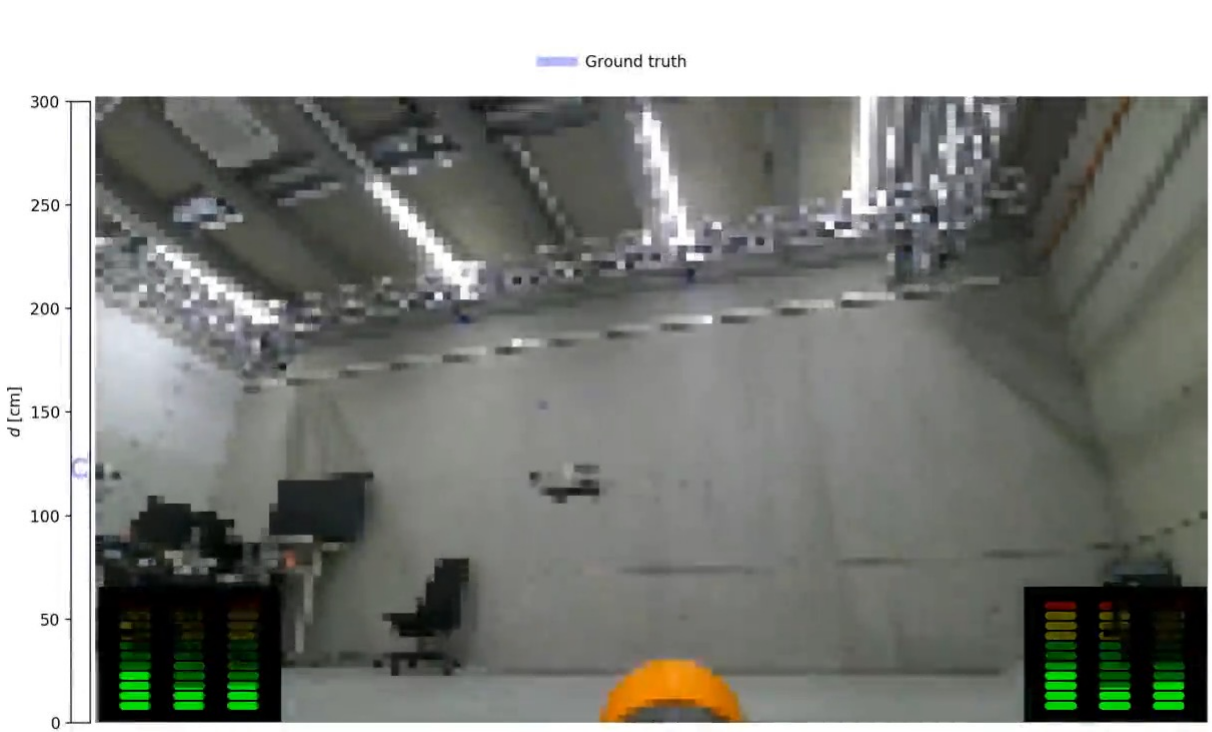


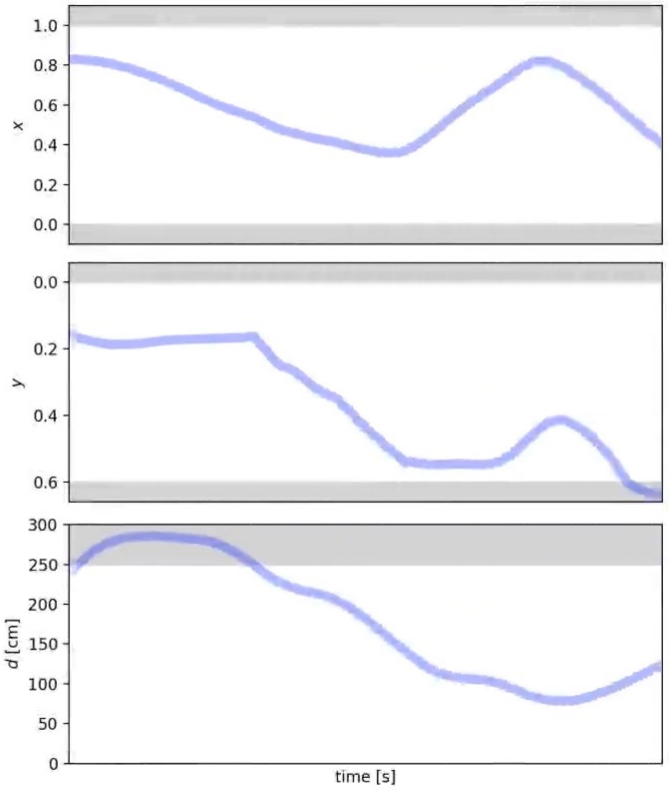
An application of classification to robotics



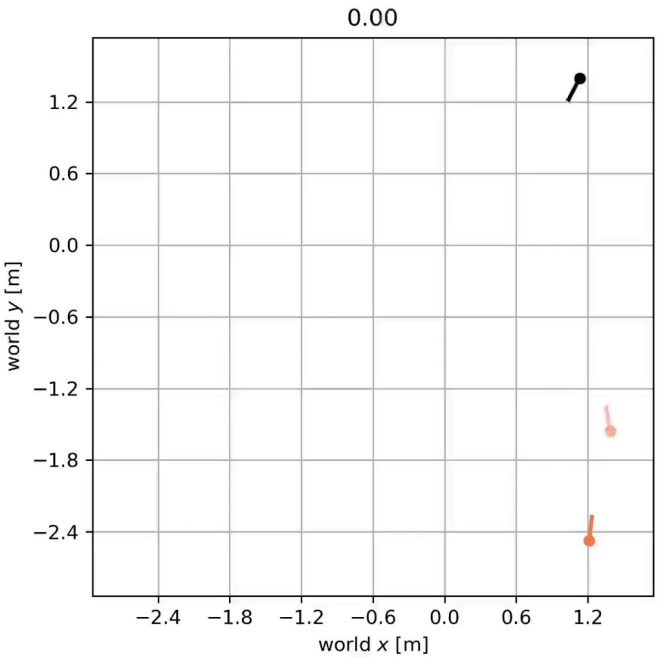
https://www.youtube.com/watch?v=umRdt3zGgpU

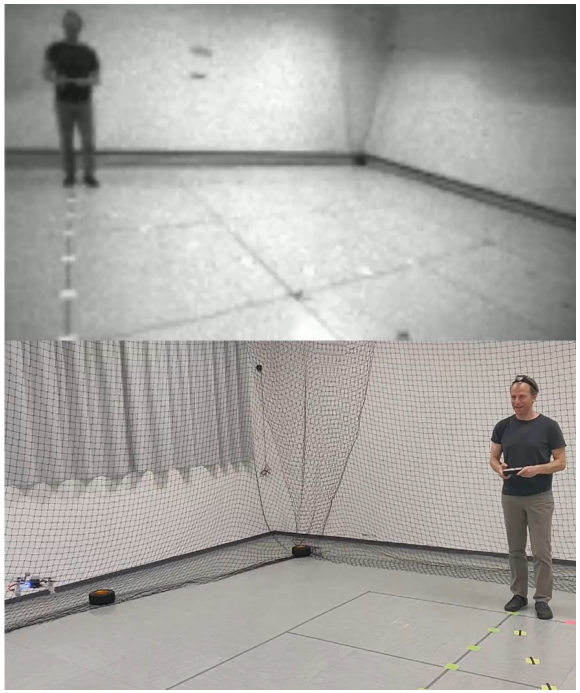
An application of regression to robotics





Another application of regression to robotics



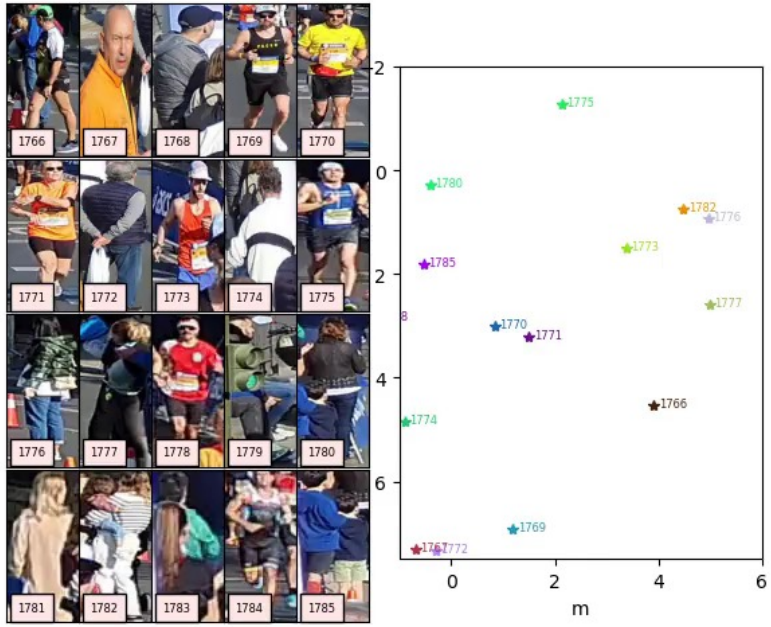


Input: image from onboard camera

Output: relative position, altitude and heading direction of the head of the person visible in the frame (4 scalars)

Runner video tracking and reidentification

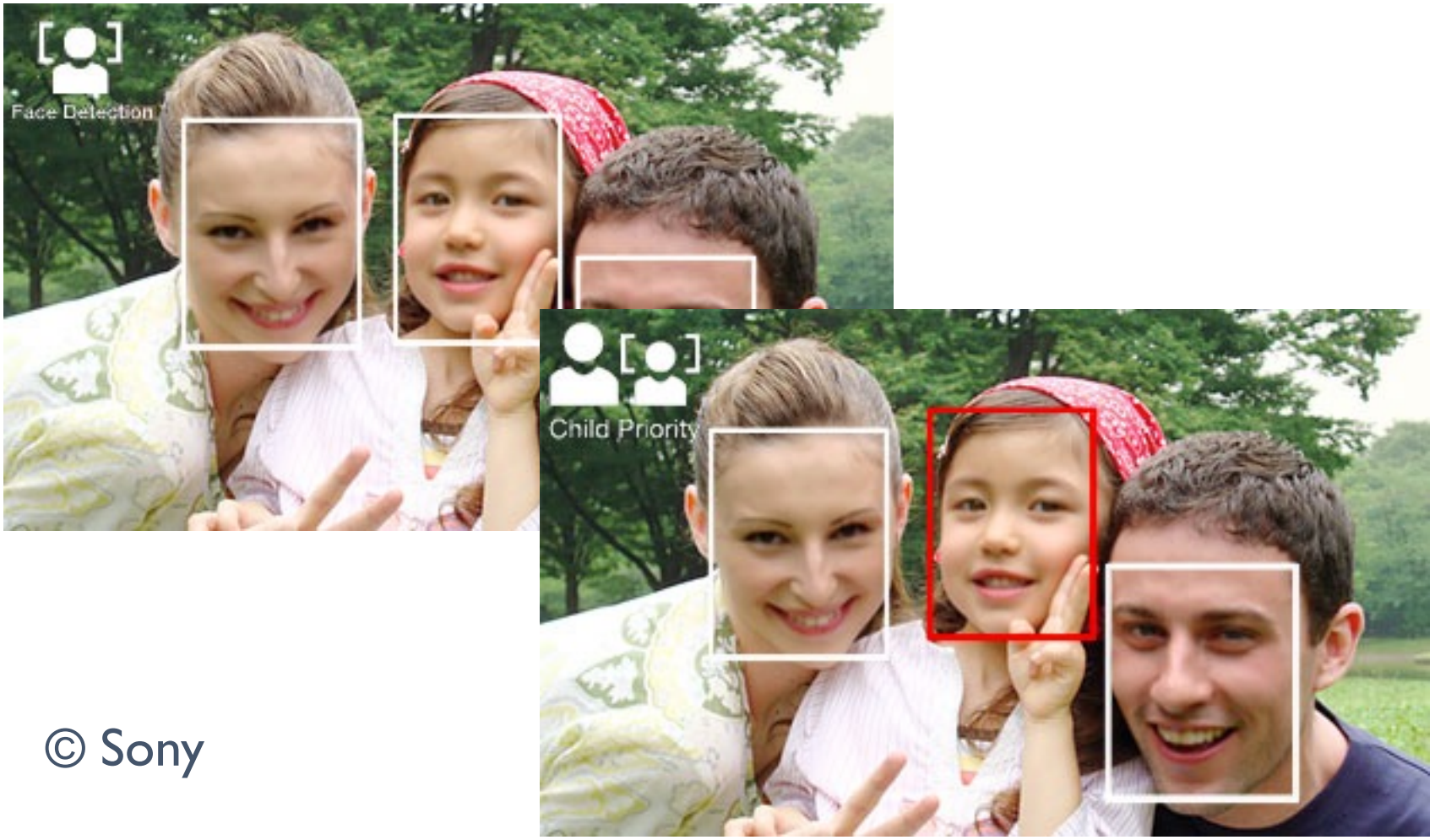
iRewind



Innosuisse project with iRewind

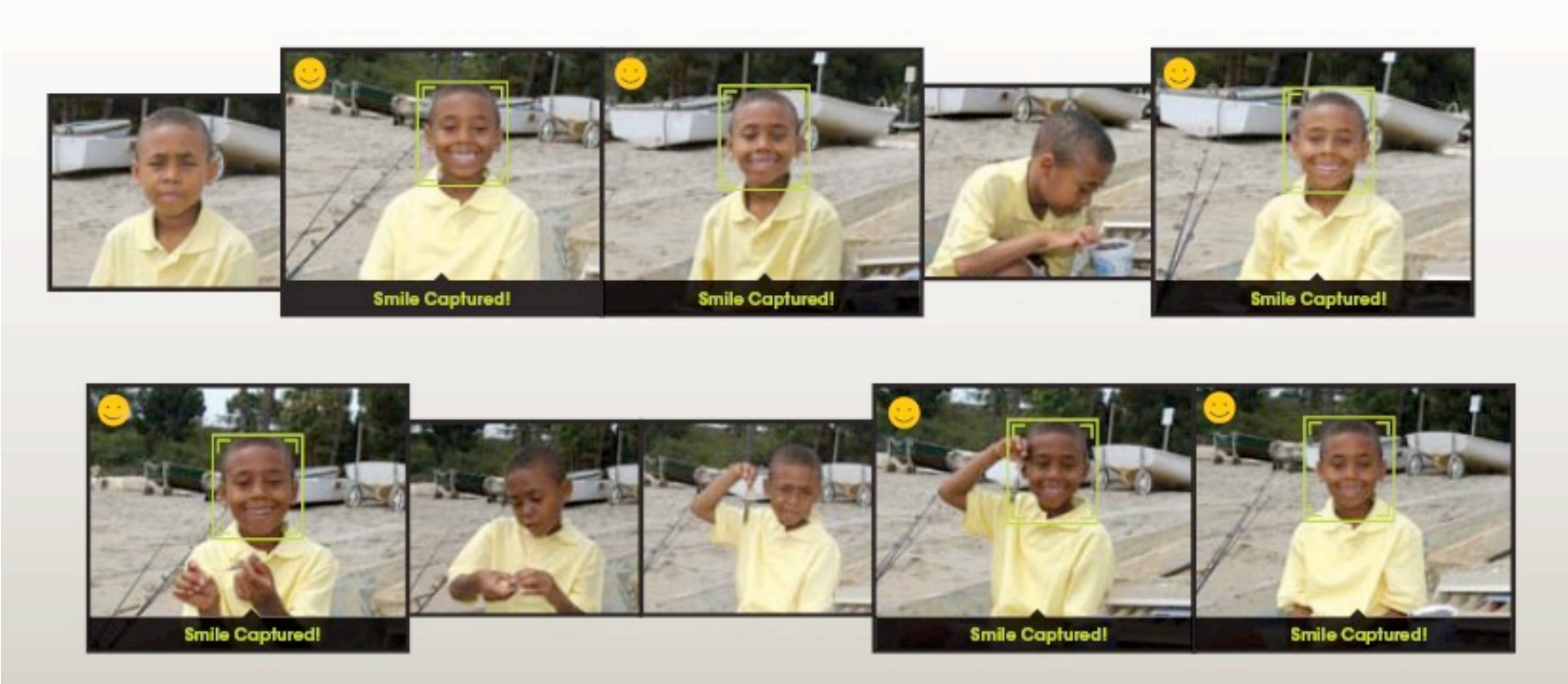
Innosuisse

Face Detection



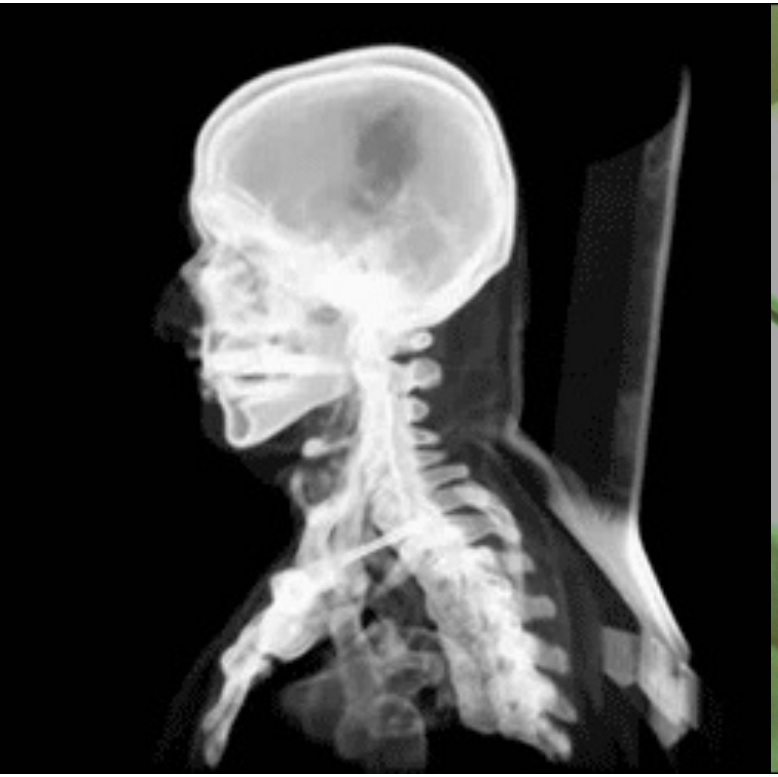
© Sony

Smile Detection!



© Sony

Medical Imaging



3D imaging
MRI, CT

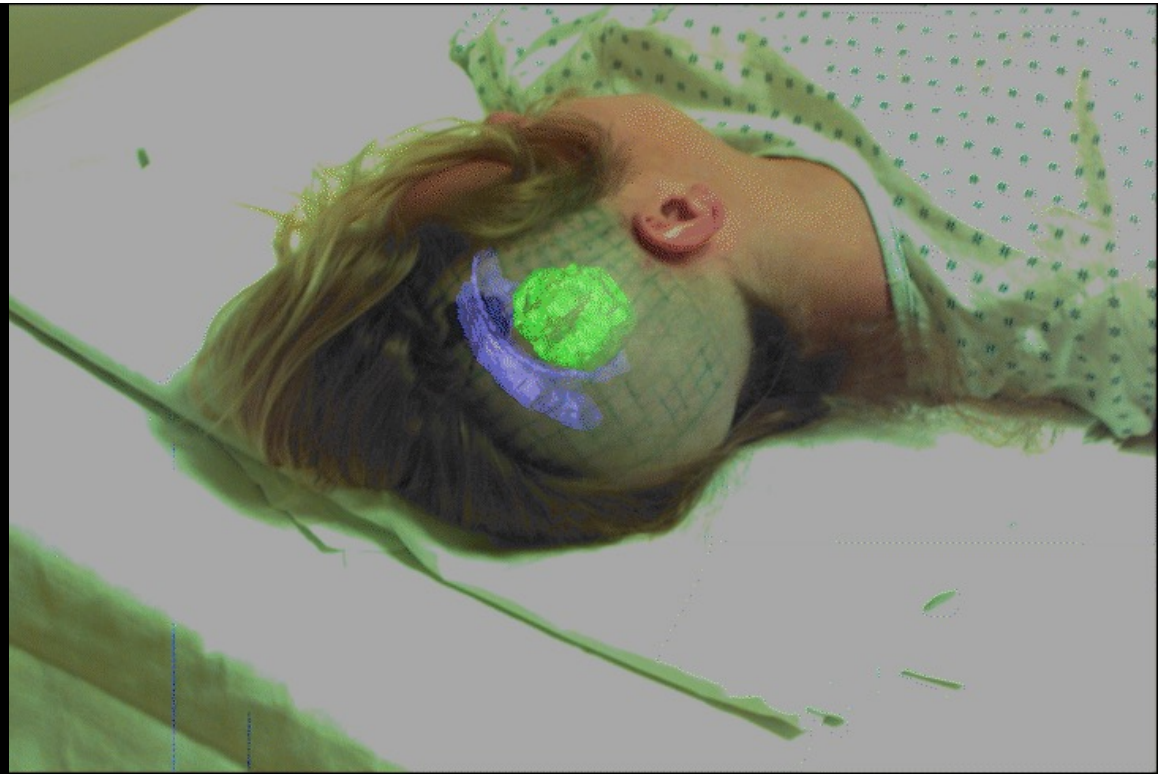


Image guided surgery
Grimson et al., Transactions on Medical Imaging, 1996

Medical Imaging

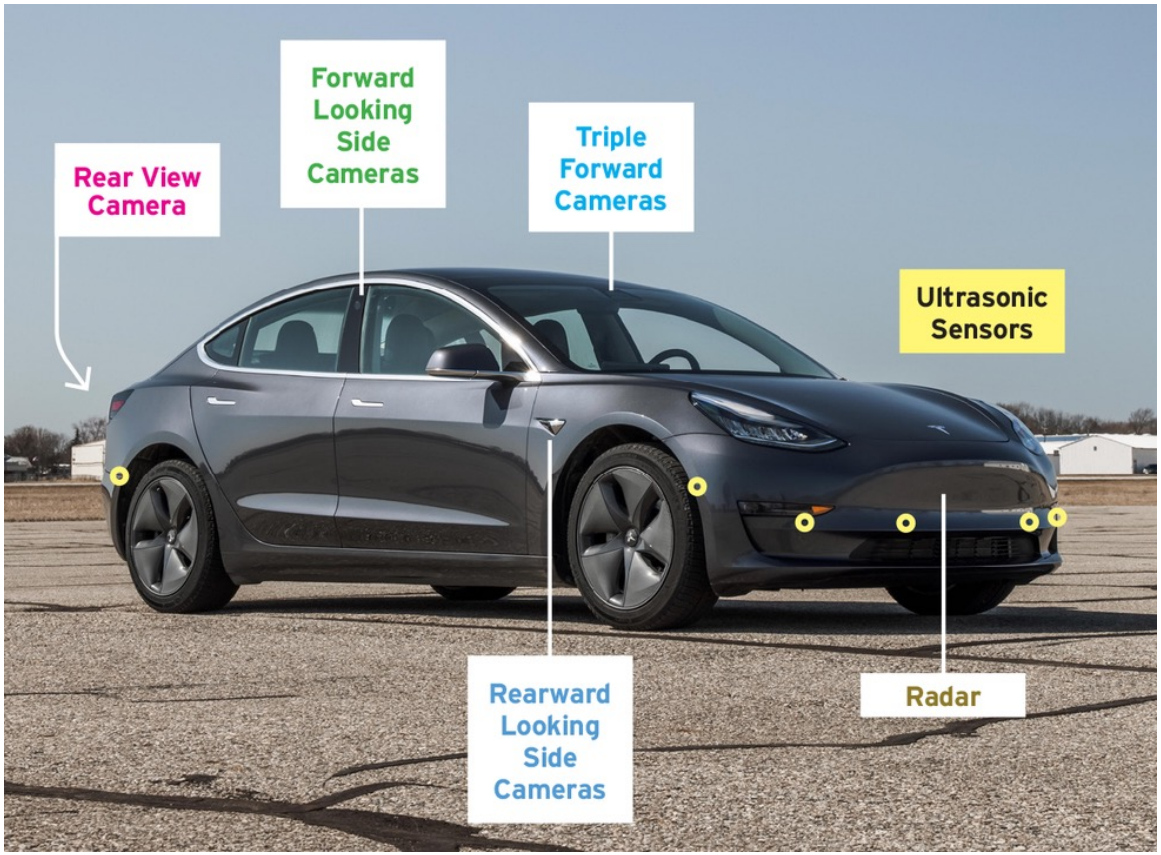
▲ Microsoft's InnerEye project uses AI to make treatment for prostate cancer more efficient. Photograph: Microsoft Project InnerEye Study

http://microsoft.com/en-us/research/project/project-innereye-open-source-software-for-medical-imaging-ai/

Digital Pathology

https://ibex-ai.com/

Autonomous Driving



Tesla's Autopilot system currently uses eight cameras, 12 ultrasonic sensors, and forward radar to read lane lines and detect nearby cars. The new cars will rely mostly on the car's cameras and its computer's processing ability to use Autopilot and the suite of features currently included in the Full Self-Driving (FSD) package. Other automakers use radar for their adaptive cruise control systems, and they benefit by being able to operate in inclement weather and direct sunlight.

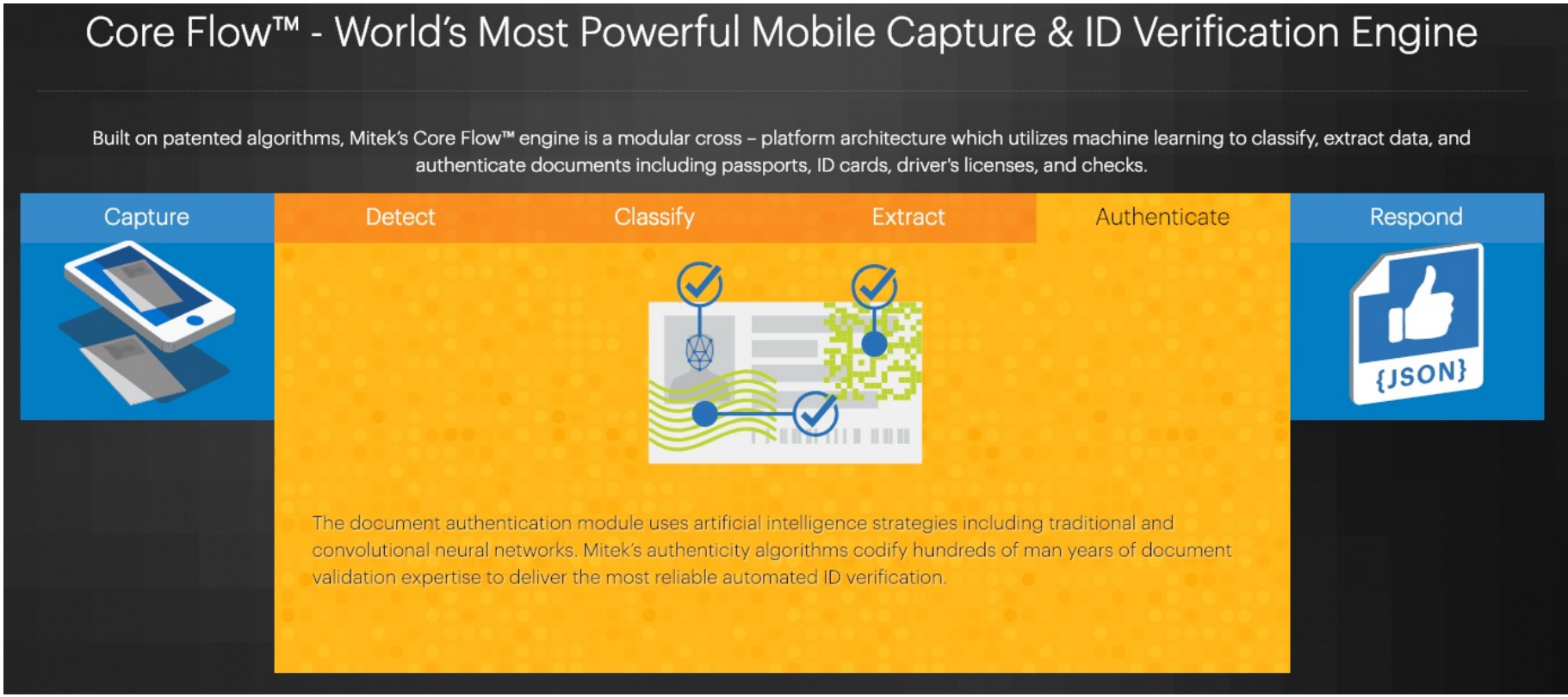
Tesla said that "a vision-only system is ultimately all that is needed for full autonomy."

https://www.caranddriver.com/news/a36542541/tesla-model-3-model-y-pure-vision/

How the car sees the world

Tesla Full Self Driving (FSD) 12.3 -- https://youtu.be/wWt2IPWwSww?si=CfjBuSkUnuf5Whxl

Banking

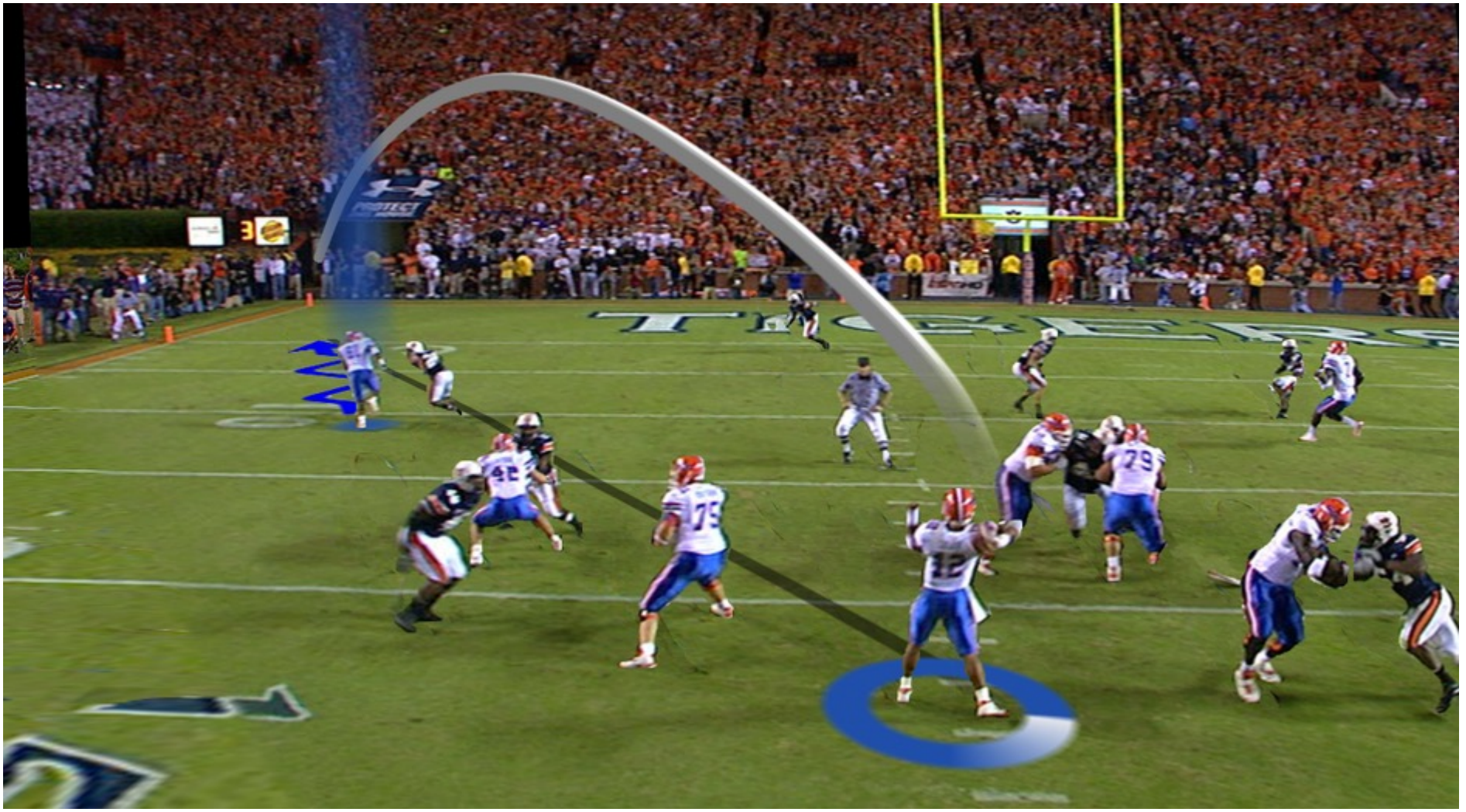


https://www.miteksystems.com/technology

Entertainment

http://www.ilm.com/

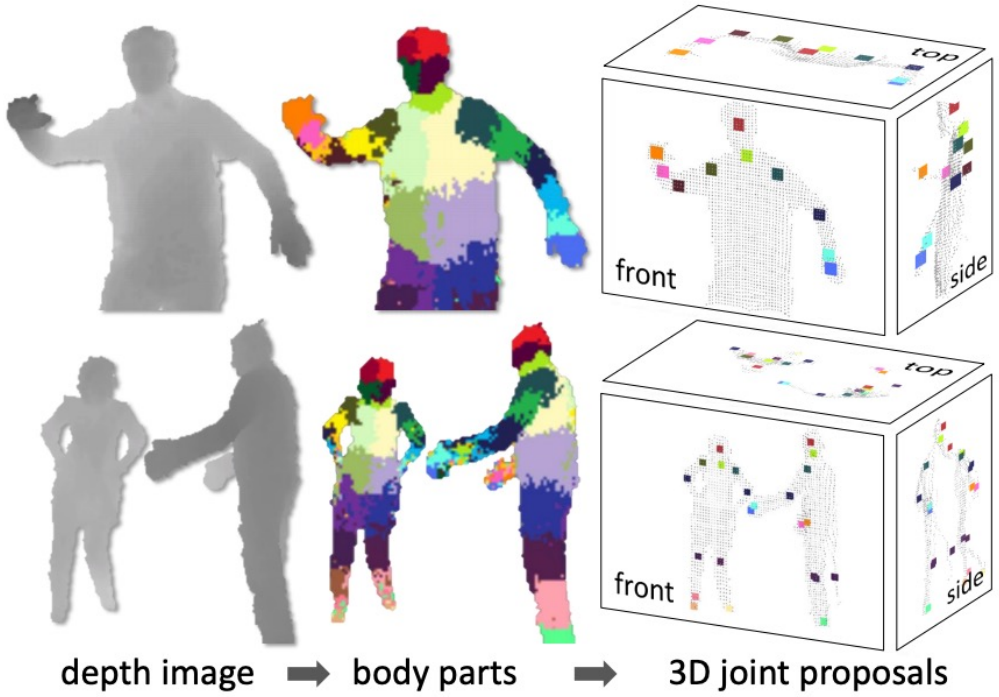
Broadcasting



Kinect Pose Recognition

Researchers from Microsoft Research Cambridge beam during the ceremony in which they won the MacRobert Award: (from left) Mat Cook, Jamie Shotton, Andrew Blake, Andrew Fitzgibbon, and Toby Sharp.

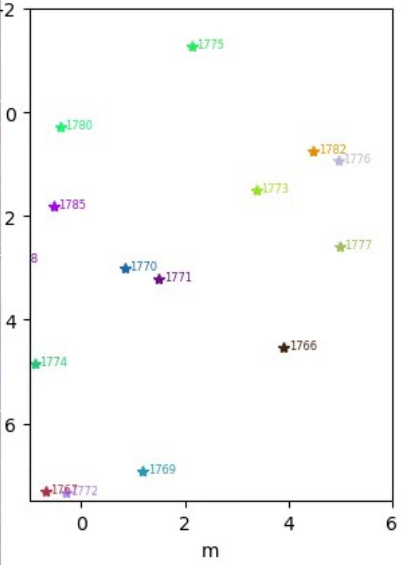
Real-Time Human Pose Recognition in Parts from Single Depth Images



Retail

Sport: tracking of runners

iRewind



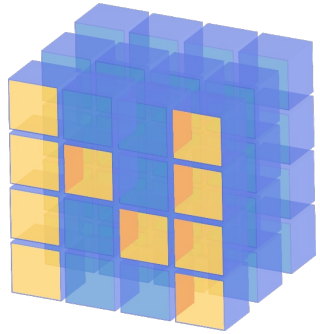
Smarter Farming

SlantRange Advanced Crop Information Solutions

Tools

We will use

- Python 3
- Numpy
- Jupyter notebooks
- Libraries for imaging:
  - scikit-image (skimage)
  - opencv
- Tensorflow (with keras) or pytorch



Your background (quick poll)

- Did you follow...
  - a Machine Learning course?
  - a Deep Learning course?
  - a Computer Graphics course?
  - an image processing course?
- How many lines of python code have you written so far?
  - < 100?
  - < 10k?
  - > 10k?
- How many matplotlib plots have you drawn?
  - < 10?
  - < 100?
  - > 100?
- Given:

  ```
  a = np.array([1, 3, 8, 2, 2, 1])
  ```

  can you write a line of code that:
  - only prints the numbers different than 2?
  - prints the array in reverse order?
  - computes the average of the elements in an even position (i.e. the 2^{nd}, 4^{th}, 6^{th}, 8^{th}, ...)