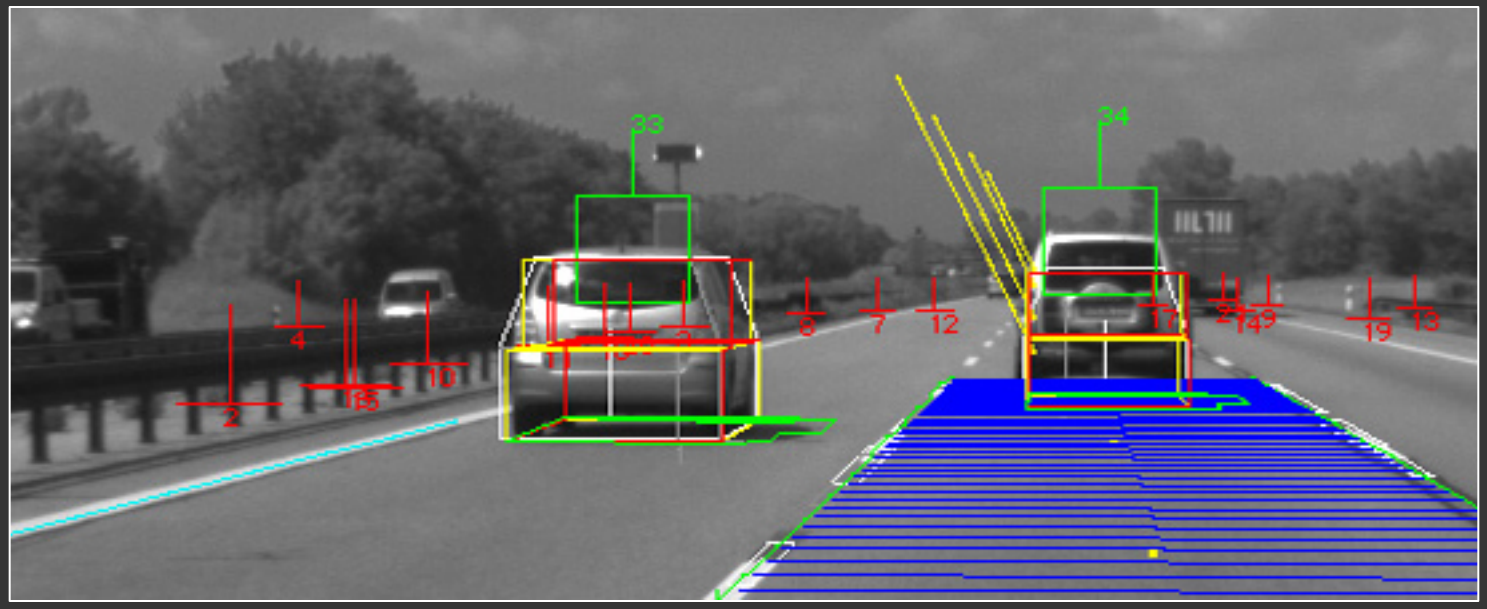
Special Topics in Computer Vision
2P 2019



Course webpage URL:

opilab.utb.edu.co/topics-computer-vision/

Contact Details

Instructor: Andrés G. Marrugo, PhD

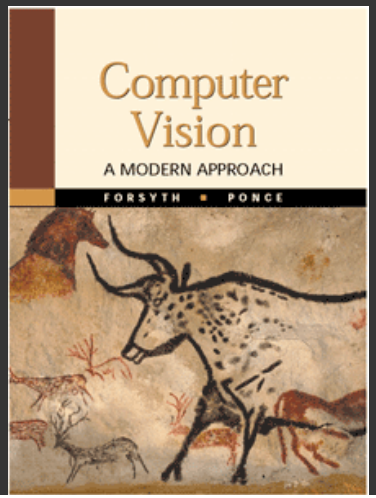
Email: agmarrugo@utb.edu.co Twitter: @agmarrugo

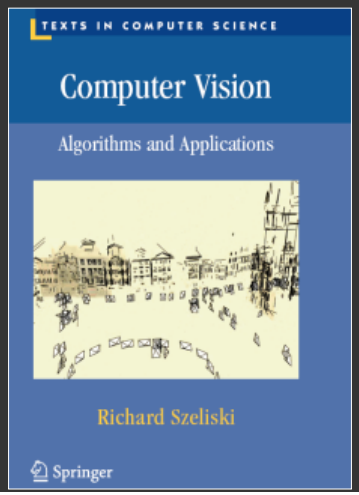
Course website: http://opilab.utb.edu.co/topics-computer-vision/

Suggested textbooks (for fundamentals):

Forsyth & Ponce, *Computer Vision: A Modern Approach*

Richard Szeliski, *Computer Vision: Algorithms and Applications* (available online)





About the course

We will select several topics in Computer vision, read and analyze related papers, and produce a simple implementation.

- Reading assignments (~6, 40%).
- Oral presentation (30%)
- Final project with written paper* (30%)

*For some, the paper is to be submitted to a conference.

Please refer to the syllabus.

Disclosure

Parts of this course have been prepared based on the publicly available information from the courses:

- Special Topics in Computer Vision by Noah Snavely at Cornell University.
- Computer Vision by Svetlana Lazebnik at University of Illinois at Urbana-Champaign.

Credit is given to the corresponding author.

The goal of computer vision

- To extract “meaning” from pixels

What we see

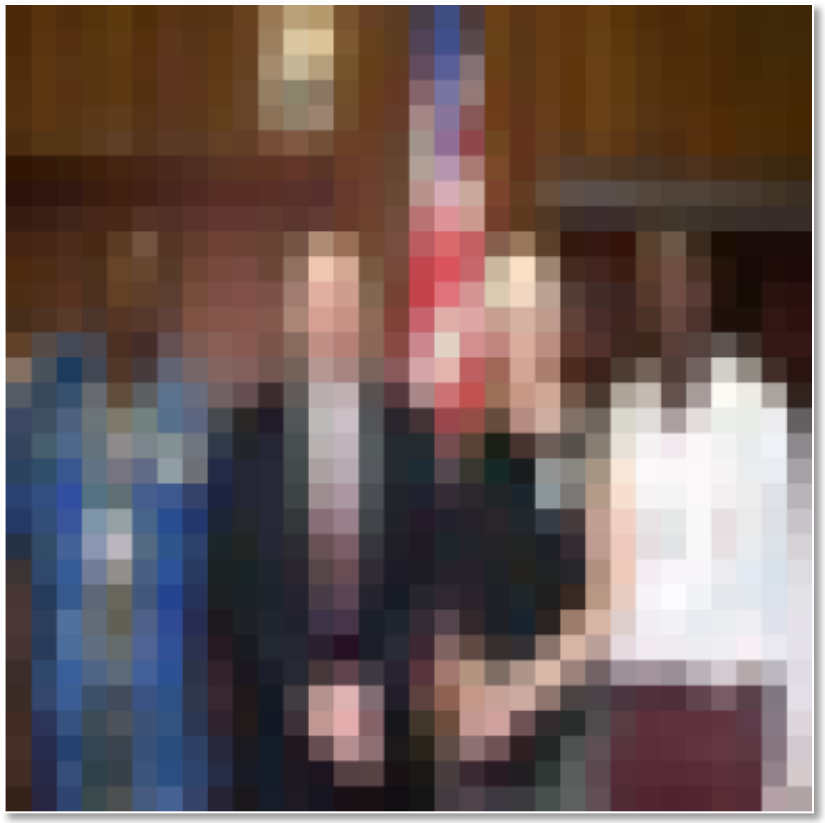
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

The goal of computer vision

- To extract “meaning” from pixels

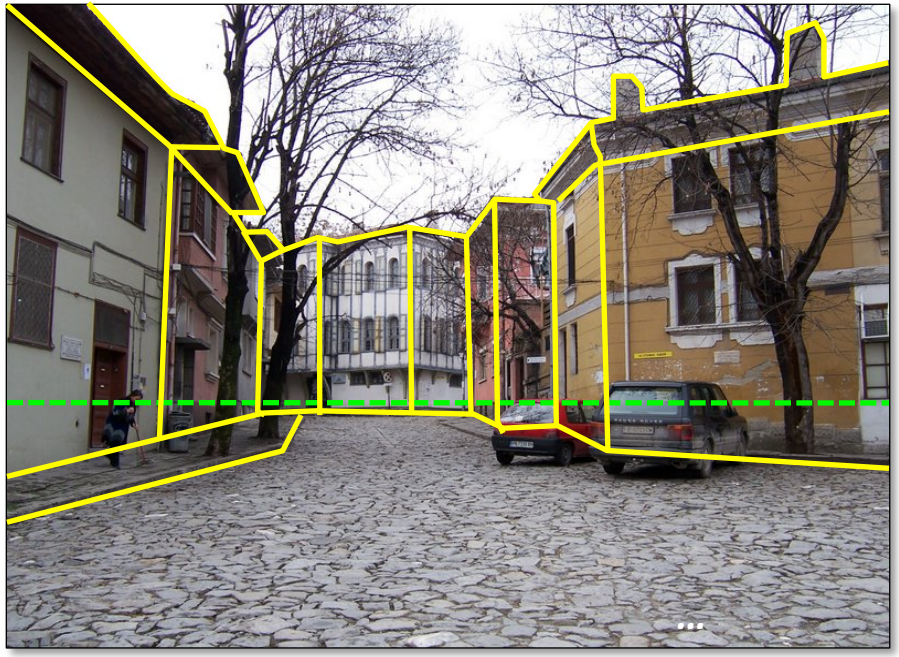


Humans are remarkably good at this…

Source: “80 million tiny images” by Torralba et al.

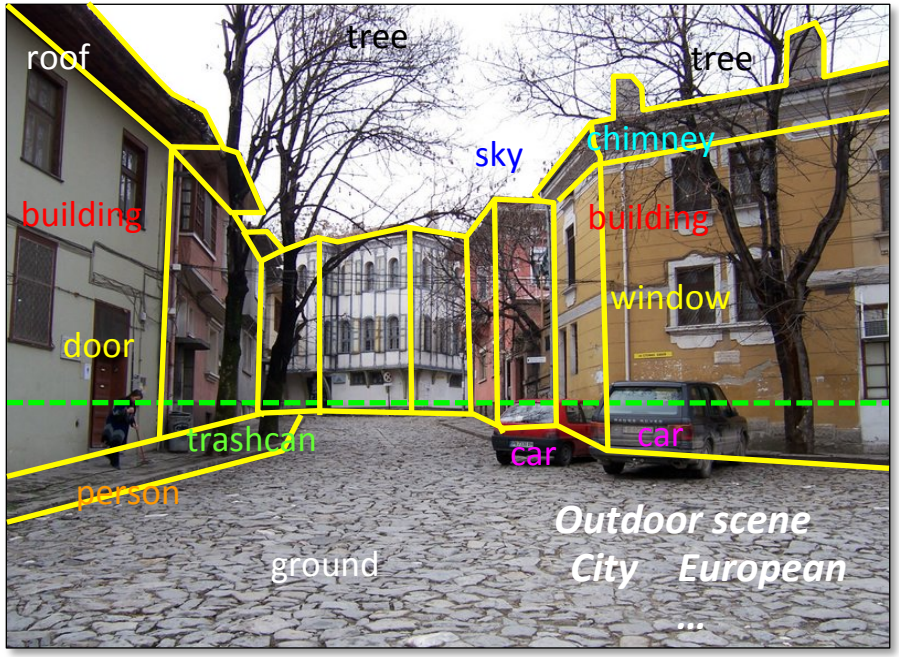
What kind of information can be extracted from an image?

What kind of information can be extracted from an image?



Geometric information

What kind of information can be extracted from an image?



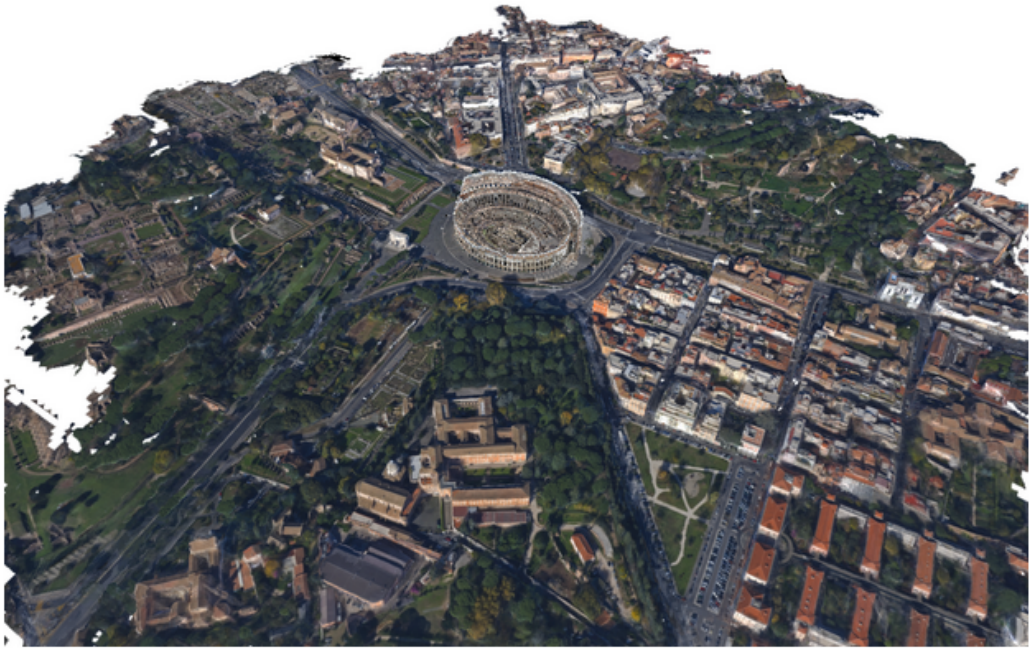
Geometric information

Semantic information

What can computer vision do today?

Reconstruction: 3D from photo collections

Colosseum, Rome, Italy



San Marco Square, Venice, Italy

Q. Shan, R. Adams, B. Curless, Y. Furukawa, and S. Seitz,
The Visual Turing Test for Scene Reconstruction, 3DV 2013

YouTube Video

Reconstruction: 4D from photo collections

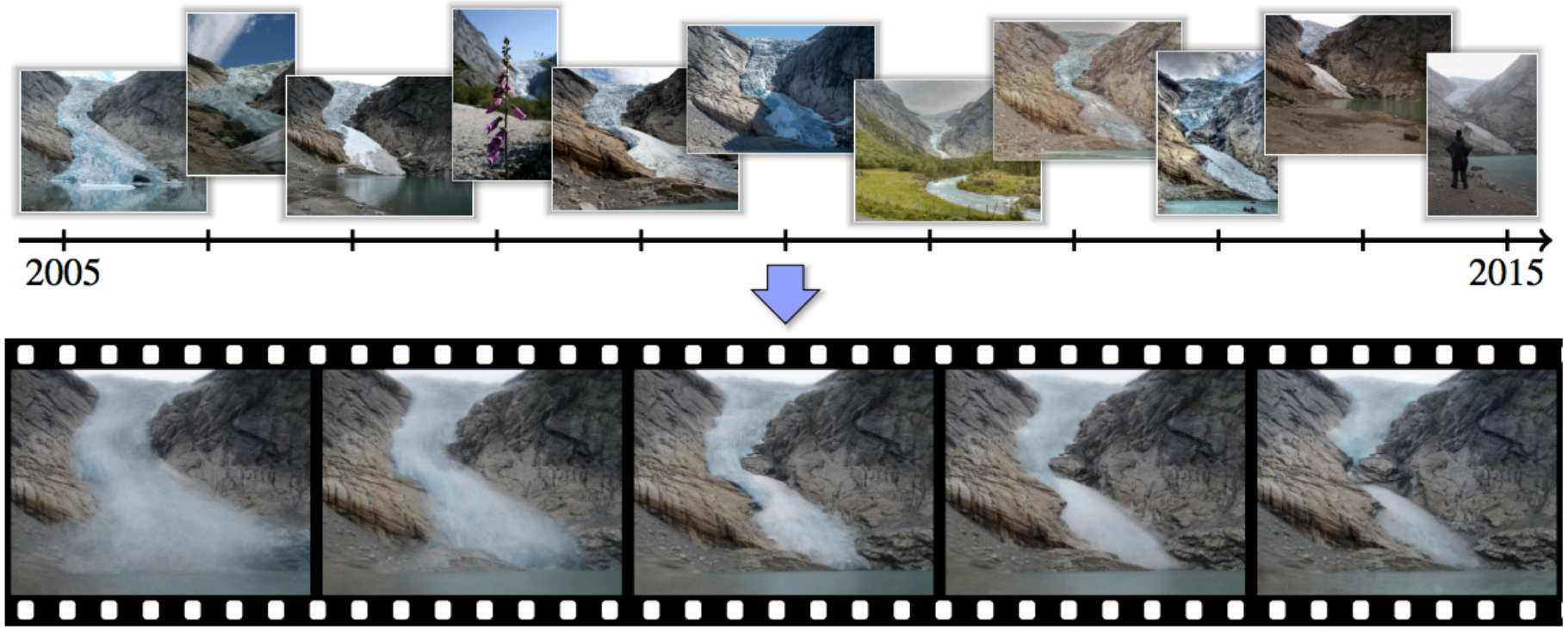


Figure 1: *We mine Internet photo collections to generate time-lapse videos of locations all over the world. Our time-lapses visualize a multitude of changes, like the retreat of the Briksdalsbreen Glacier in Norway shown above. The continuous time-lapse (bottom) is computed from hundreds of Internet photos (samples on top).* *Photo credits: Aliento Más Allá, jirihnidek, mcxurxo, elka_cz, Juan Jesús Orío, Klaus Wißkirchen, Daikrieg, Free the image, dration and Nadav Tobias.*

R. Martin-Brualla, D. Gallup, and S. Seitz,
Time-Lapse Mining from Internet Photos, SIGGRAPH 2015

YouTube Video

Reconstruction: 4D from depth cameras

Figure 1: Real-time reconstructions of a moving scene with DynamicFusion; both the person and the camera are moving. The initially noisy and incomplete model is progressively denoised and completed over time (left to right).

R. Newcombe, D. Fox, and S. Seitz,
DynamicFusion: Reconstruction and Tracking of Non-rigid Scenes in Real-Time, CVPR 2015

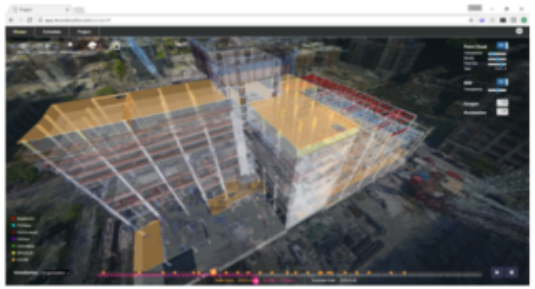
YouTube Video

Reconstruction in construction industry

RECONSTRUCT INTEGRATES REALITY AND PLAN

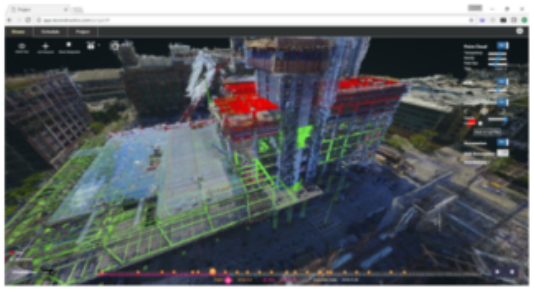
Visual Asset Management

Reconstruct 4D point clouds and organize images and videos from smartphones, time-lapse cameras, and drones around the project schedule. View, annotate, and share anywhere with a web interface.



4D Visual Production Models

Integrate 4D point clouds with 4D BIM, review "who does what work at what location" on a daily basis and improve coordination and communication among project teams.



Predictive Visual Data Analytics

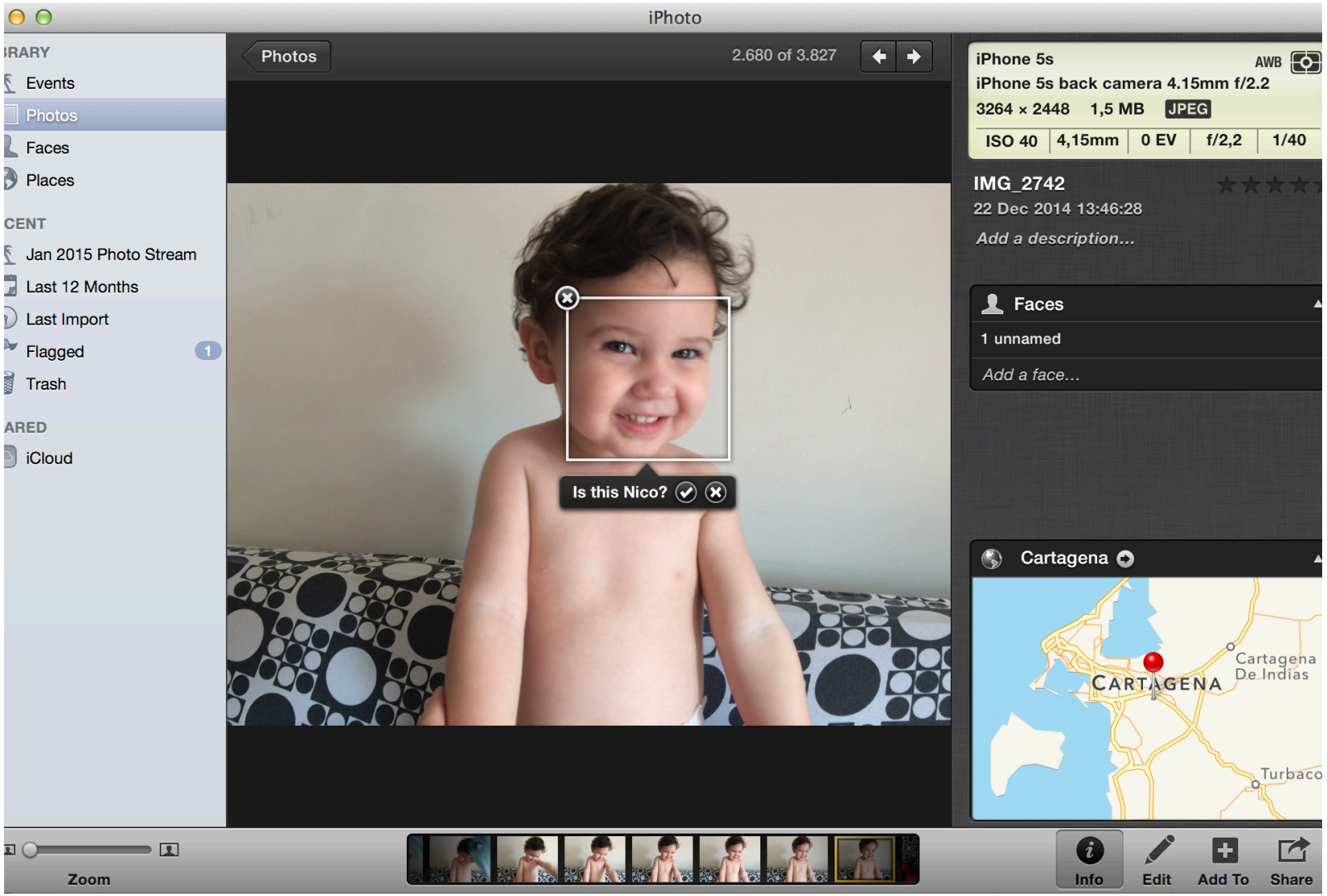
Analyze actual progress deviations by comparing Reality and Plan and predict risk with respect to the execution of the look-ahead schedule for each project location, to offer your project team with an opportunity to tap off potential delays before they surface on your jobsite.

reconstructinc.com

Source: D. Hoiem

Recognition: “Simple” patterns

Face recognition: Apple iPhoto software



Recognition: Apple Face ID

Concerns about face recognition

Beijing bets on facial recognition in a big drive for total surveillance – Washington Post, 1/8/2018

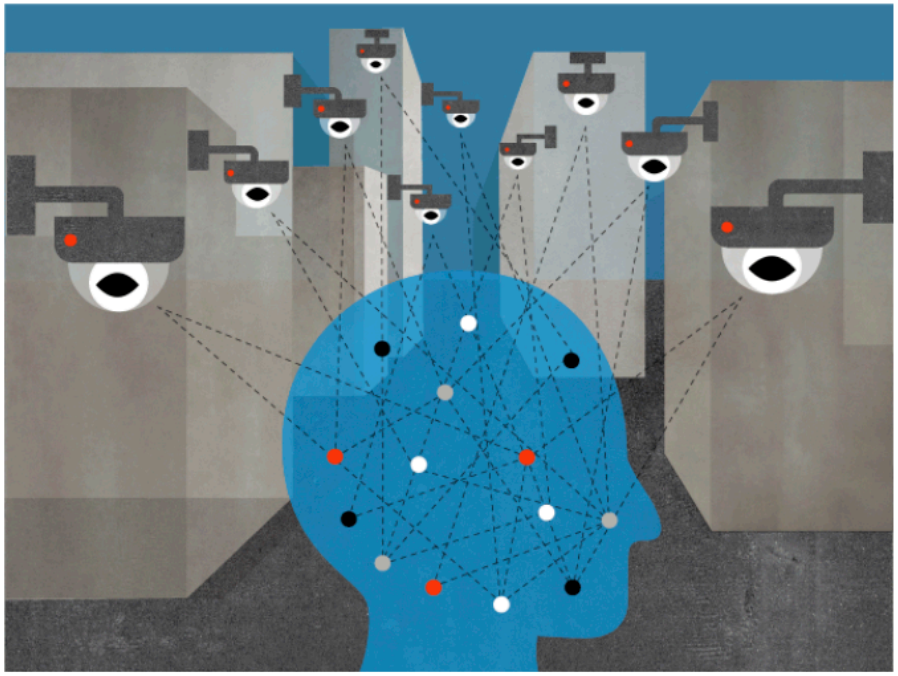
Concerns about face recognition

ANNALS OF TECHNOLOGY DECEMBER 17, 2018 ISSUE

THE NEW YORKER

SHOULD WE BE WORRIED ABOUT COMPUTERIZED FACIAL RECOGNITION?

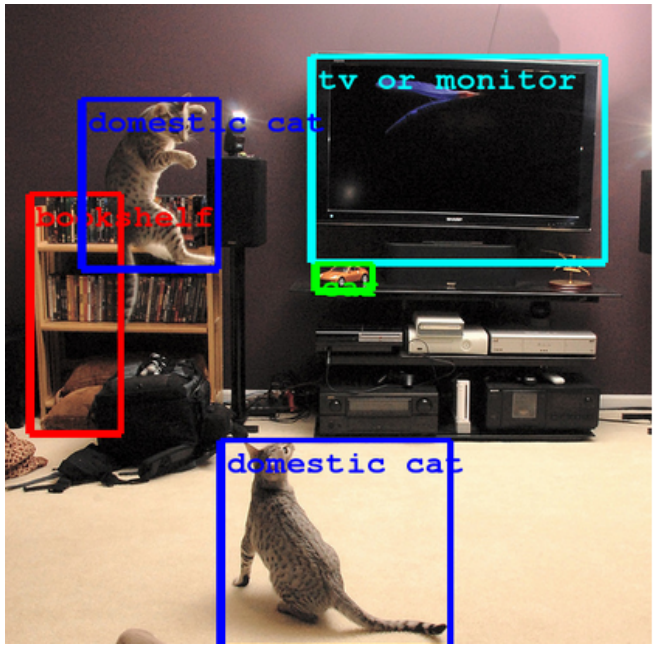
The technology could revolutionize policing, medicine, even agriculture—but its applications can easily be weaponized.

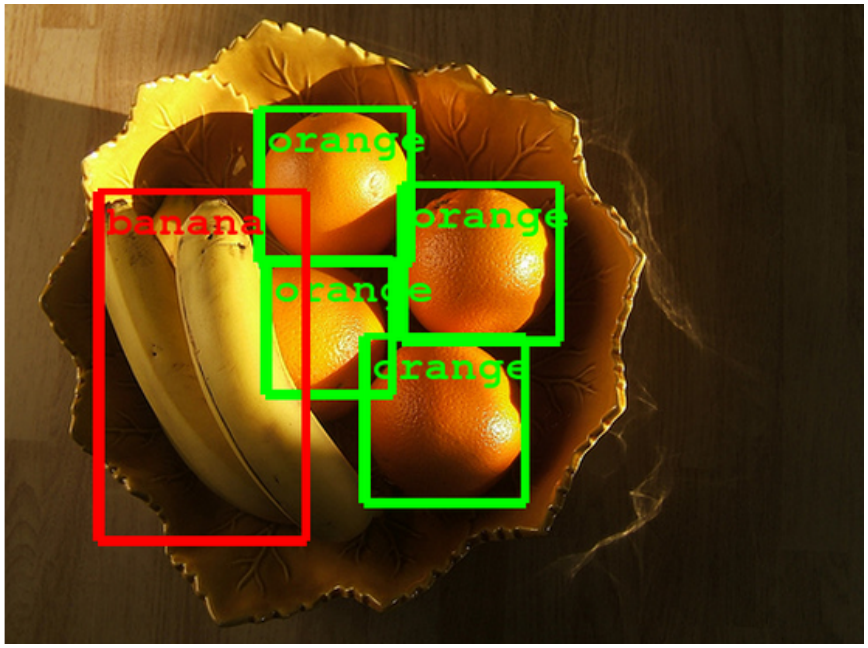


Many U.S. cities won't disclose their police departments' surveillance methods.

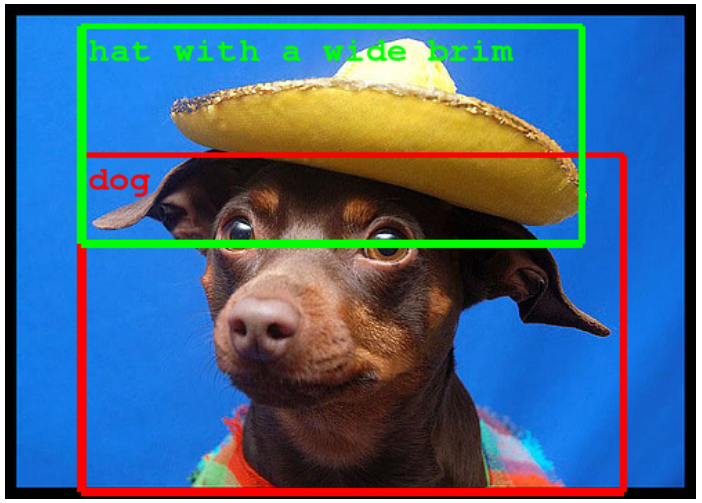
https://www.newyorker.com/magazine/2018/12/17/should-we-be-worried-about-computerized-facial-recognition

Recognition: General categories



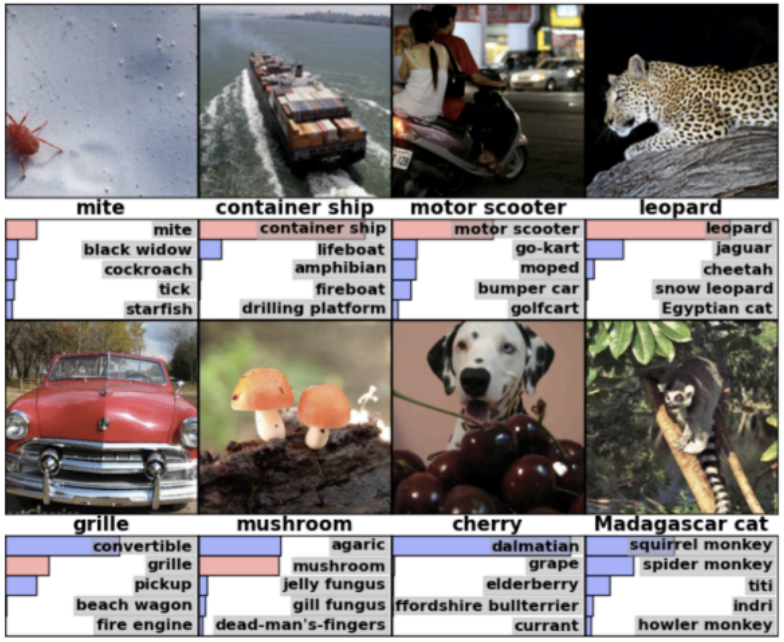


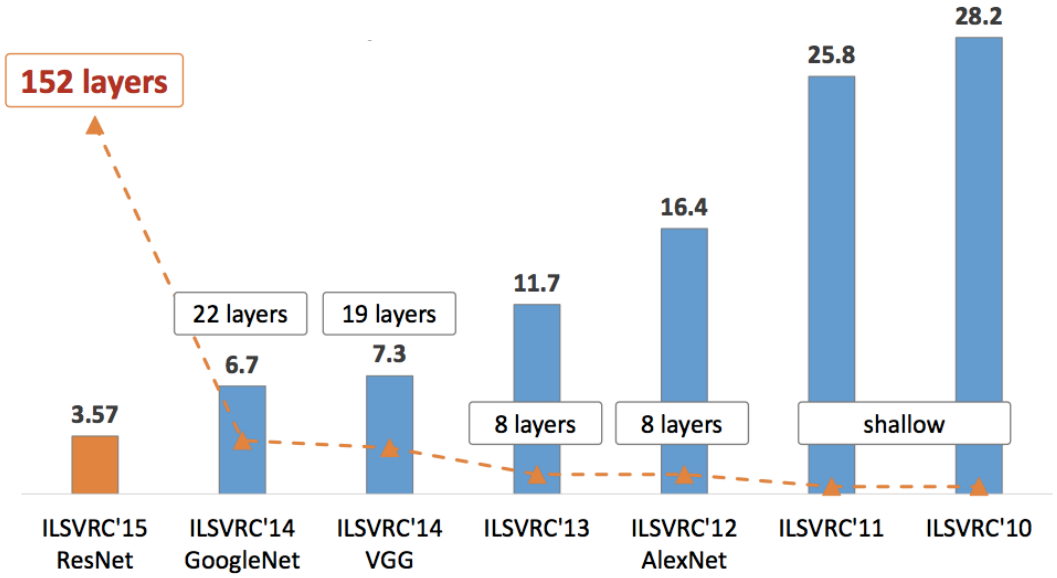
- [Computer Eyesight Gets a Lot More Accurate](), NY Times Bits blog, August 18, 2014
- [Building A Deeper Understanding of Images](), Google Research Blog, September 5, 2014



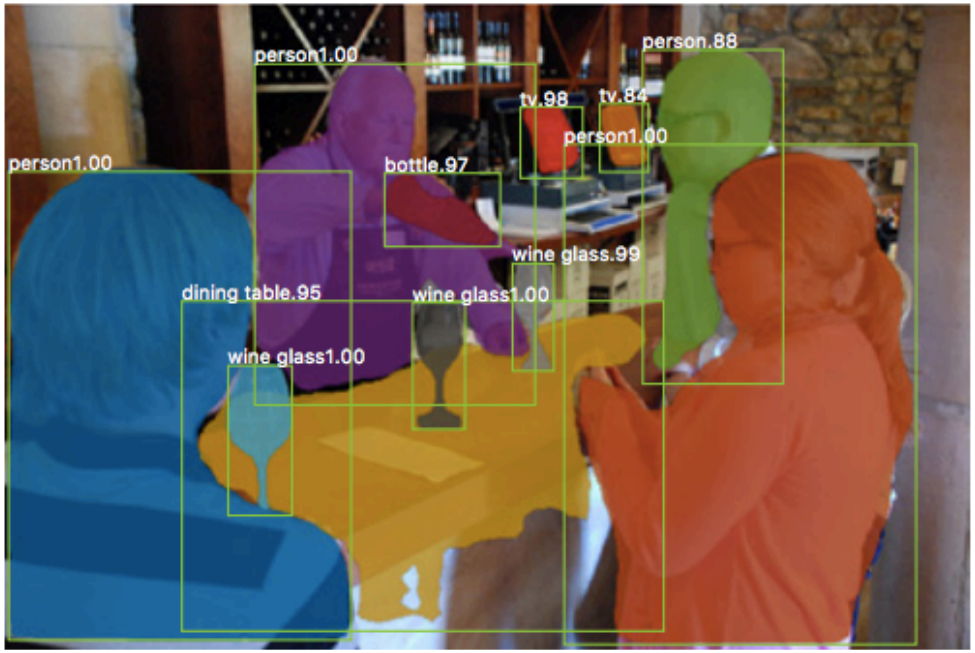
Recognition: General categories

- [ImageNet challenge]





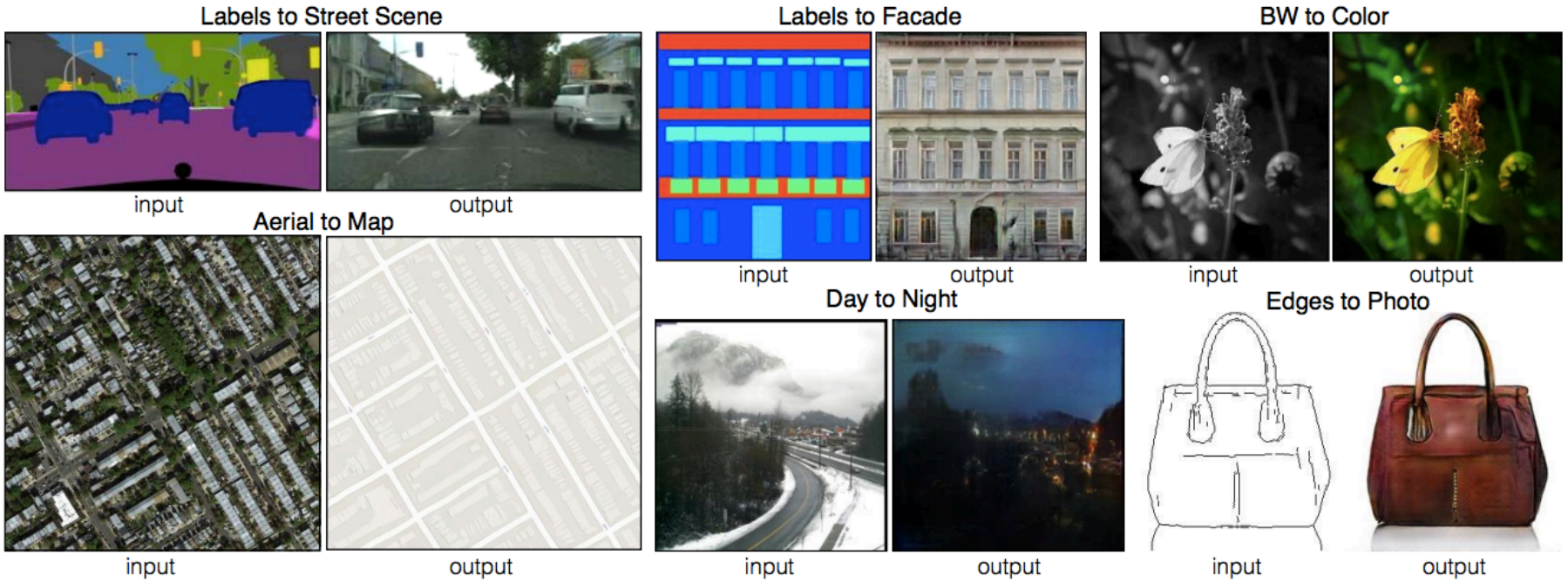
Object detection, instance segmentation



K. He, G. Gkioxari, P. Dollar, and R. Girshick, Mask R-CNN,
ICCV 2017 (Best Paper Award)

Image generation

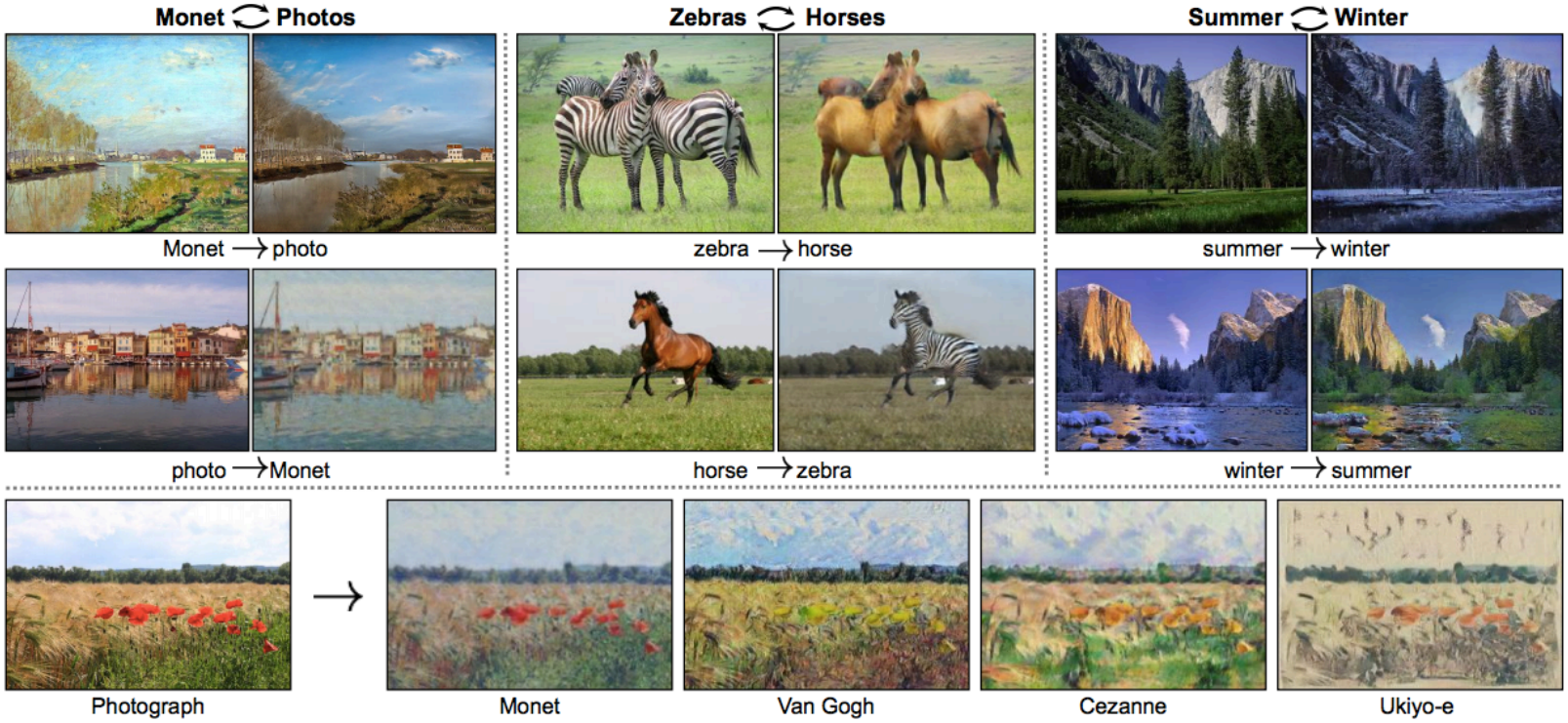
- Image-to-image translation



P. Isola, J.-Y. Zhu, T. Zhou, A. Efros,
Image-to-Image Translation with Conditional Adversarial Networks, CVPR 2017

Image generation

- Unpaired image-to-image translation



J.-Y. Zhu, T. Park, P. Isola, A. Efros, Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks, ICCV 2017

Unsupervised image-to-image translation

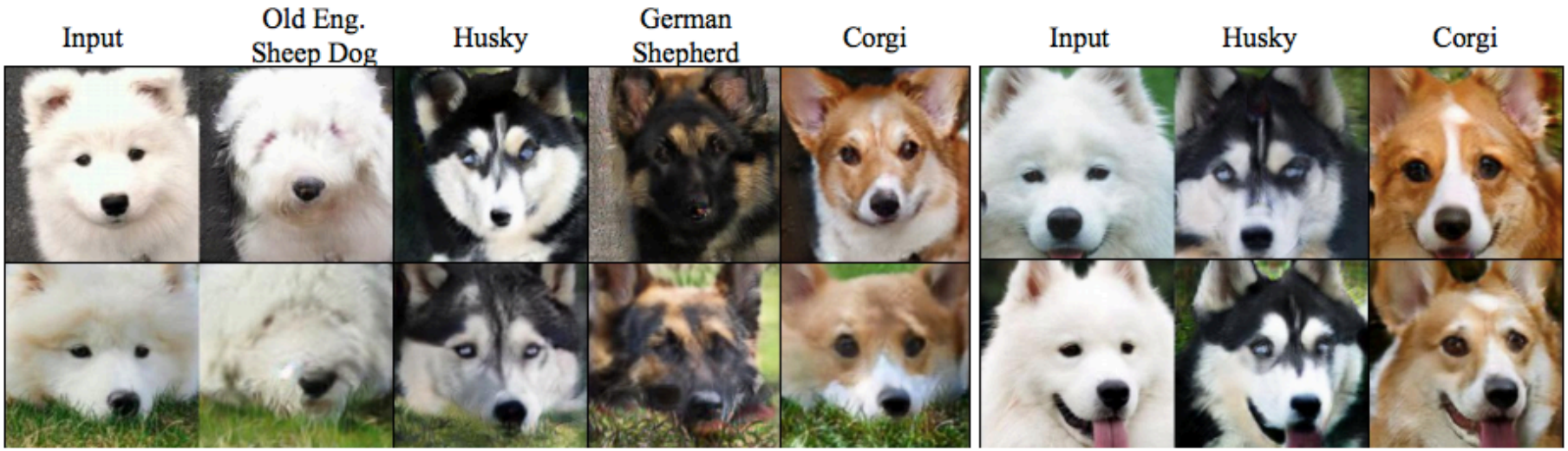


Figure 4: Dog breed translation results.

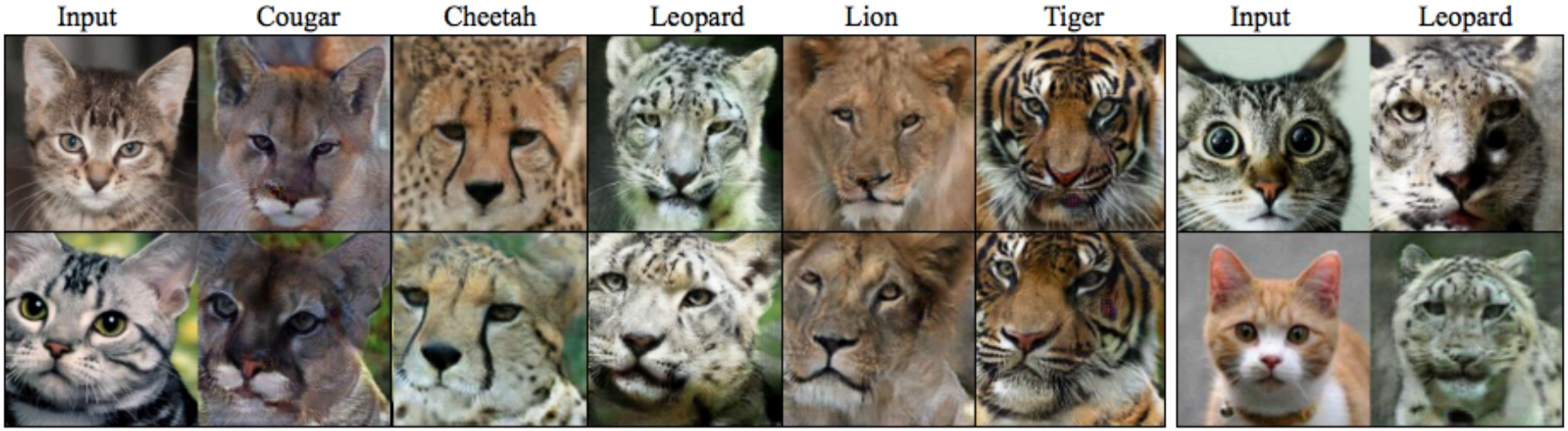


Figure 5: Cat species translation results.

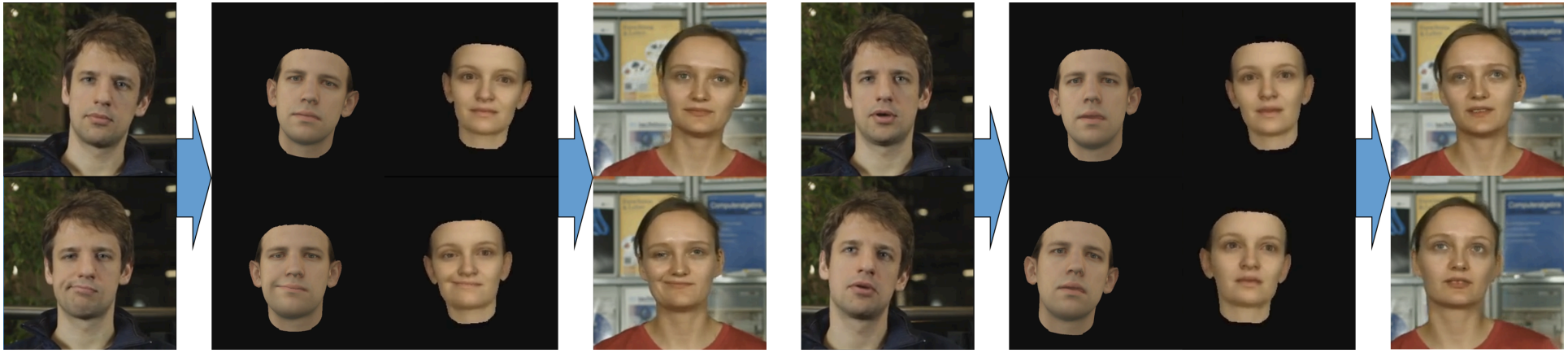
M.-Y. Liu, T. Breuel, and J. Kautz, Unsupervised Image-to-Image Translation Networks, NIPS 2017

Unsupervised image-to-image translation

M.-Y. Liu, T. Breuel, and J. Kautz, Unsupervised Image-to-Image Translation Networks, NIPS 2017

DeepFakes

- Deep video portraits



- *"A quiet wager has taken hold among researchers who study artificial intelligence techniques and the societal impacts of such technologies. They're betting whether or not someone will create a so-called Deepfake video about a political candidate that receives more than 2 million views before getting debunked by the end of 2018"* – IEEE Spectrum, 6/22/2018

DeepFakes

DEPT. OF TECHNOLOGY NOVEMBER 12, 2018 ISSUE

THE NEW YORKER

IN THE AGE OF A.I., IS SEEING STILL BELIEVING?

Advances in digital imagery could deepen the fake-news crisis—or help us get out of it.

By Joshua Rothman

As synthetic media spreads, even real images will invite skepticism.

Illustration by Javier Jaén; photograph by Svetikd / Getty

https://www.newyorker.com/magazine/2018/11/12/in-the-age-of-ai-is-seeing-still-believing

Successes to date

Data-driven approach

nature biomedical engineering

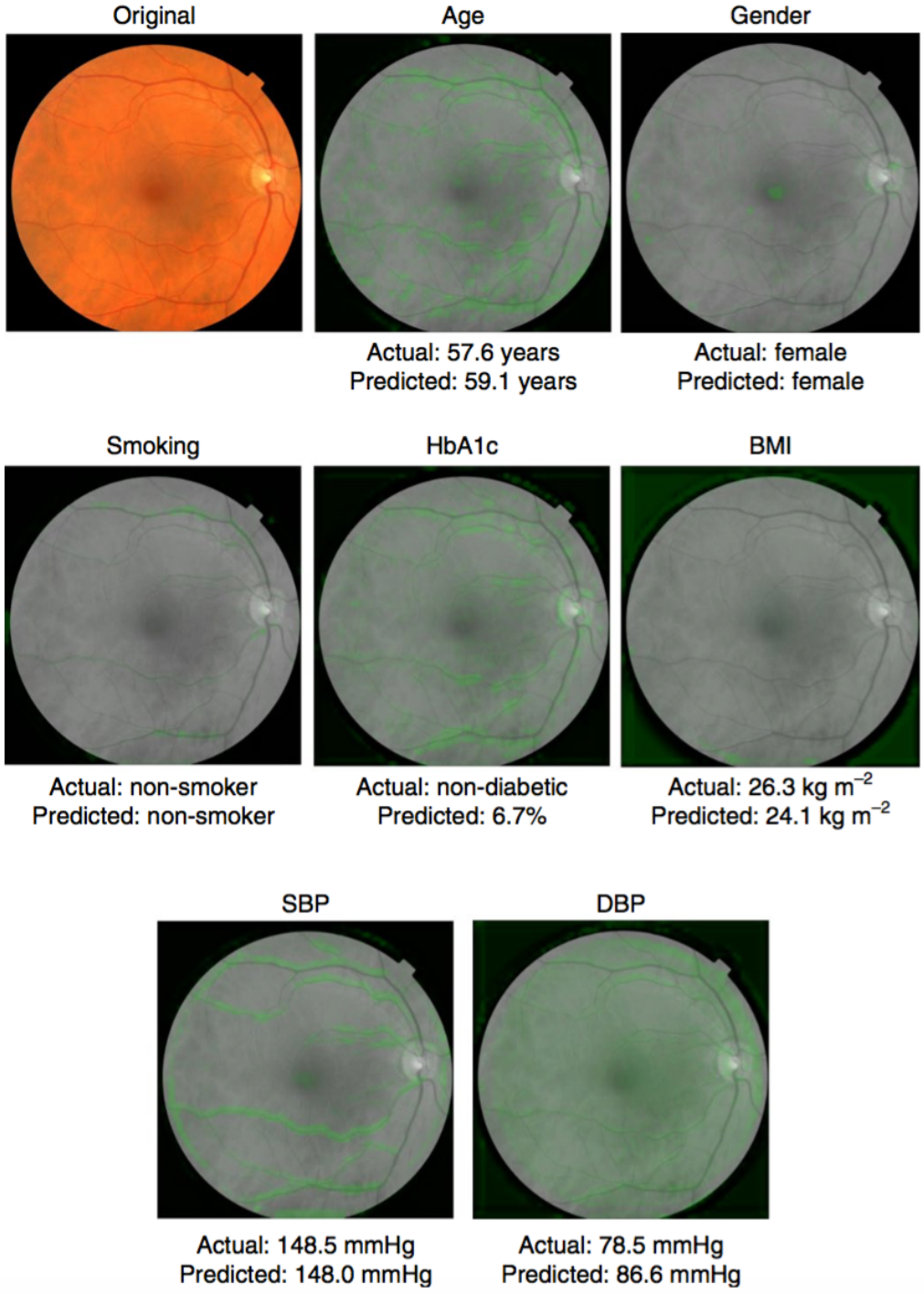
ARTICLES

https://doi.org/10.1038/s41551-018-0195-0

Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning

Ryan Poplin[1,4], Avinash V. Varadarajan[1,4], Katy Blumer[1], Yun Liu[1], Michael V. McConnell[2,3], Greg S. Corrado[1], Lily Peng[1,4]* and Dale R. Webster[1,4]

R. Poplin et al., 2018



nature biomedical engineering

ARTICLES

https://doi.org/10.1038/s41551-018-0195-0

Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning

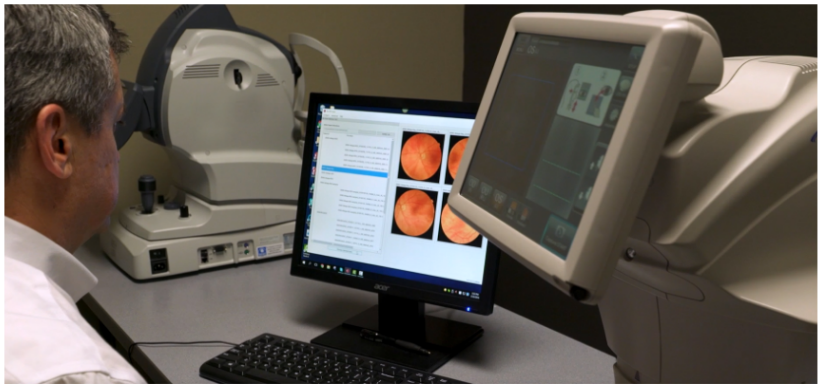
Ryan Poplin[1,4], Avinash V. Varadarajan[1,4], Katy Blumer[1], Yun Liu[1], Michael V. McConnell[2,3], Greg S. Corrado[1], Lily Peng[1,4*] and Dale R. Webster[1,4]

They predicted cardiovascular risk factors not previously thought to be present or quantifiable in retinal images.

R. Poplin et al., 2018

Successes to date

IDX-DR first AI system approved by FDA



IDx-DR

The first ever autonomous AI system cleared by the FDA to provide a diagnostic decision

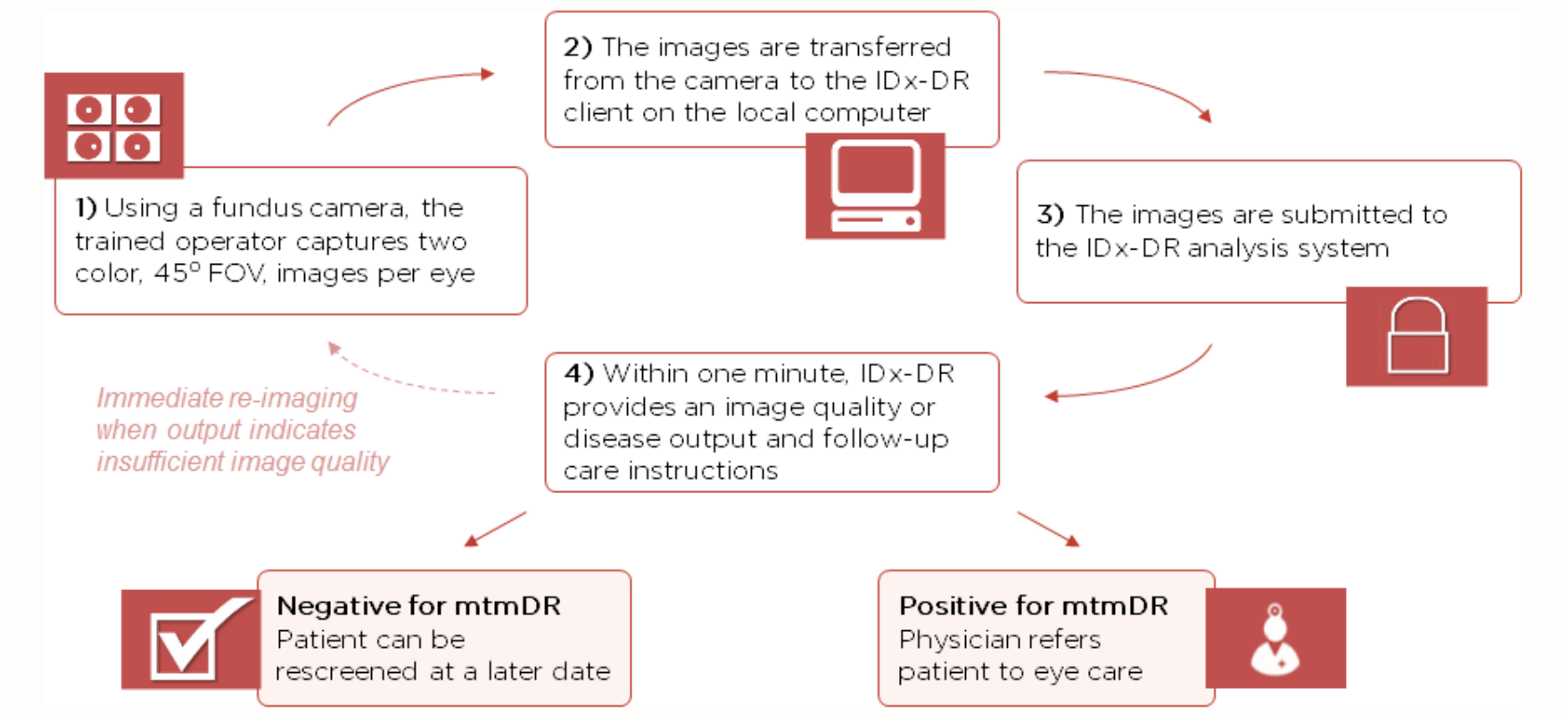
87% Sensitivity
90% Specificity
96% Imageability

Endpoints Exceeded By A Wide Margin

https://www.eyediagnosis.co/

Successes to date

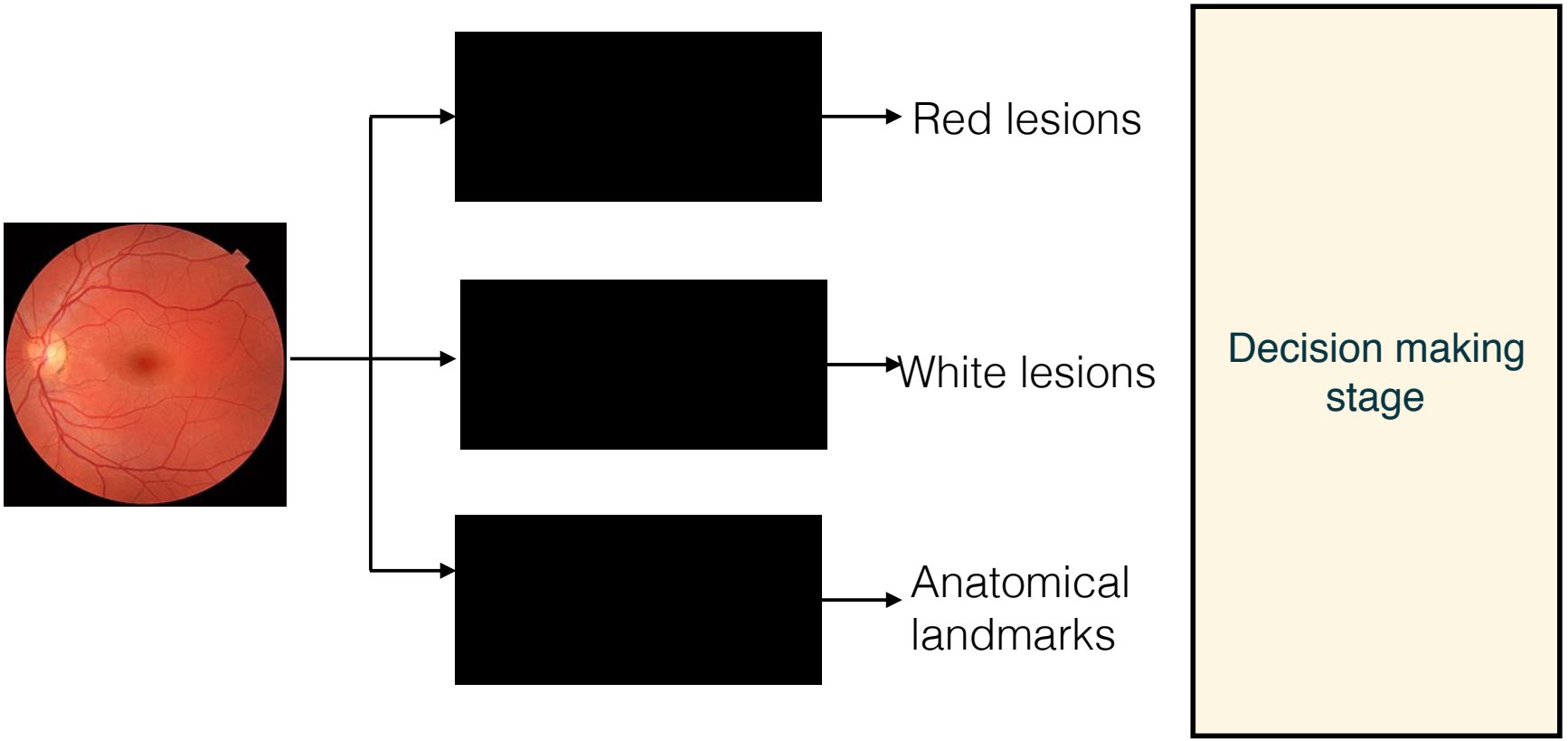
IDX-DR first AI system approved by FDA



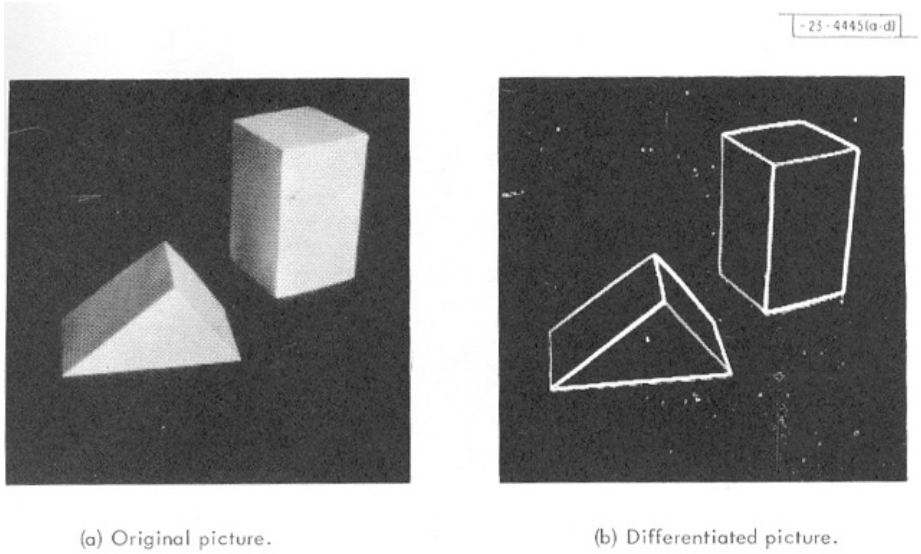
https://www.eyediagnosis.co/

Successes to date

Hybrid approach in DR automated detection



Origins of computer vision



(a) Original picture. (b) Differentiated picture.

(c) Line drawing. (d) Rotated view.

L. G. Roberts
Machine Perception of Three Dimensional Solids

Origins of computer vision

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

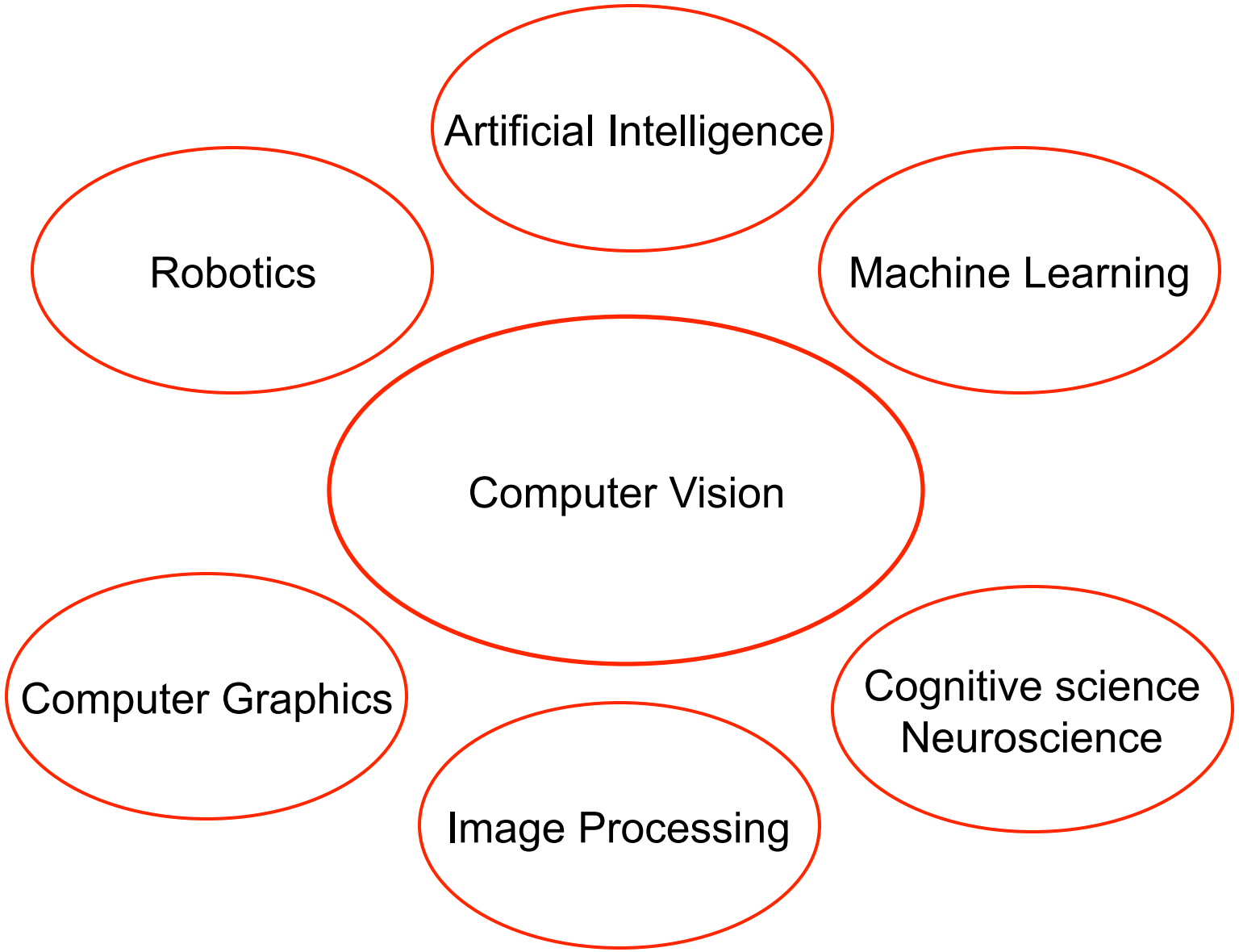
Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

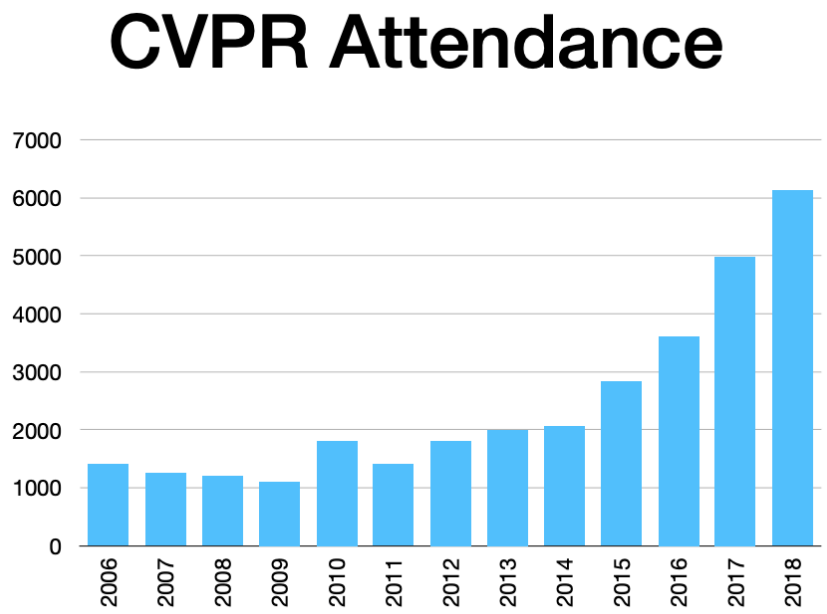
Seymour Papert

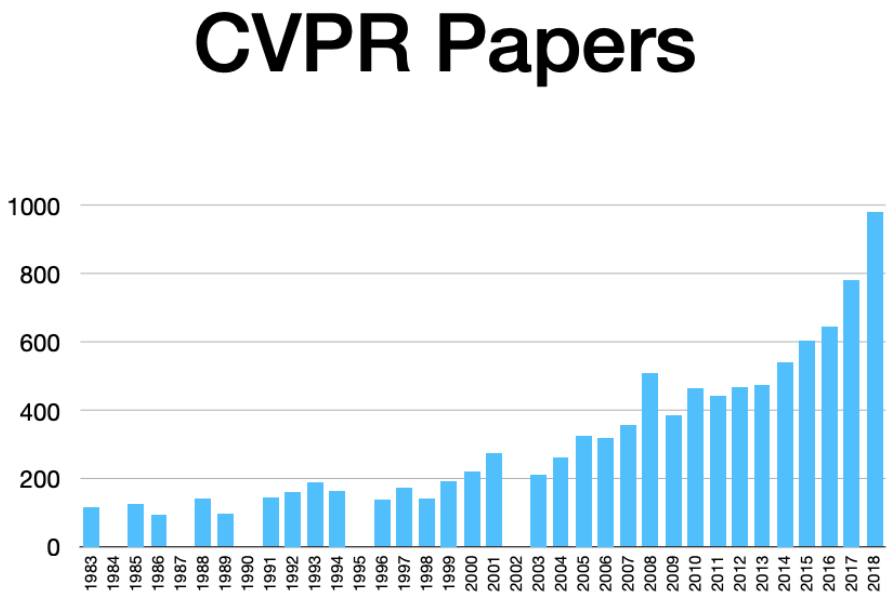
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Connections to other disciplines



Growth of the field





Source

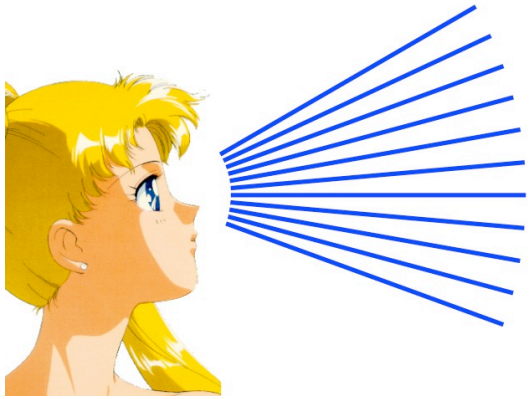
Long list of corporate sponsors

Computer Vision subfields

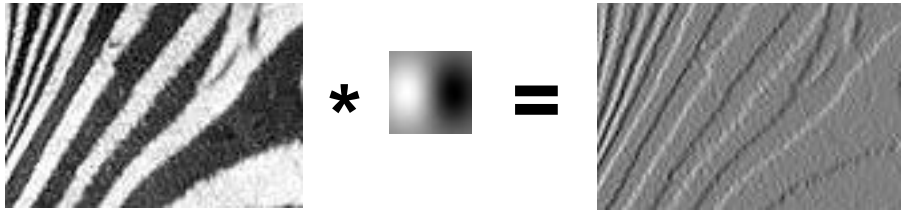
I. Early vision: Image formation and processing
II. Mid-level vision: Grouping and fitting
III. Multi-view geometry
IV. Recognition
V. Additional topics

I. Early vision

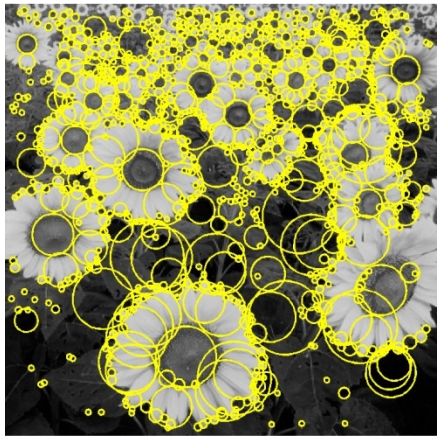
Basic image formation and processing



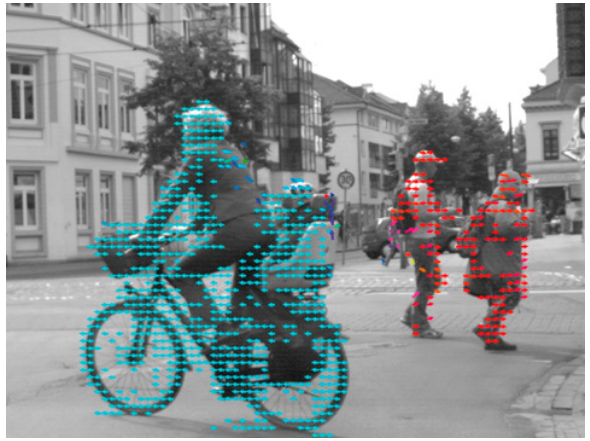
Cameras and sensors
Light and color



Linear filtering
Edge detection



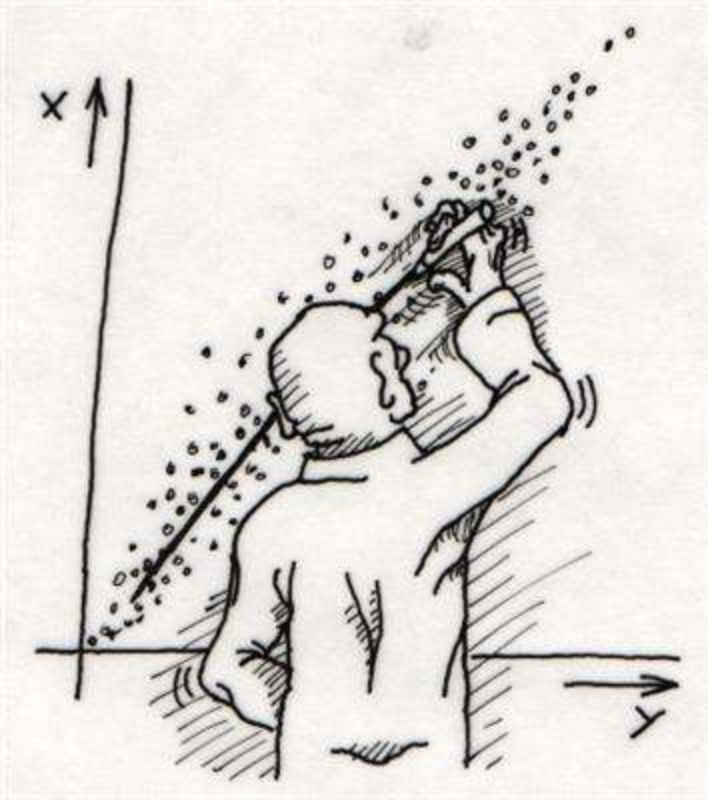
Feature extraction



Optical flow

II. “Mid-level vision”

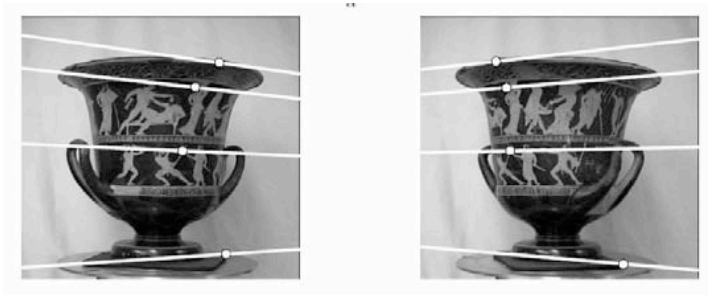
Fitting and grouping



Fitting: Least squares
Voting methods

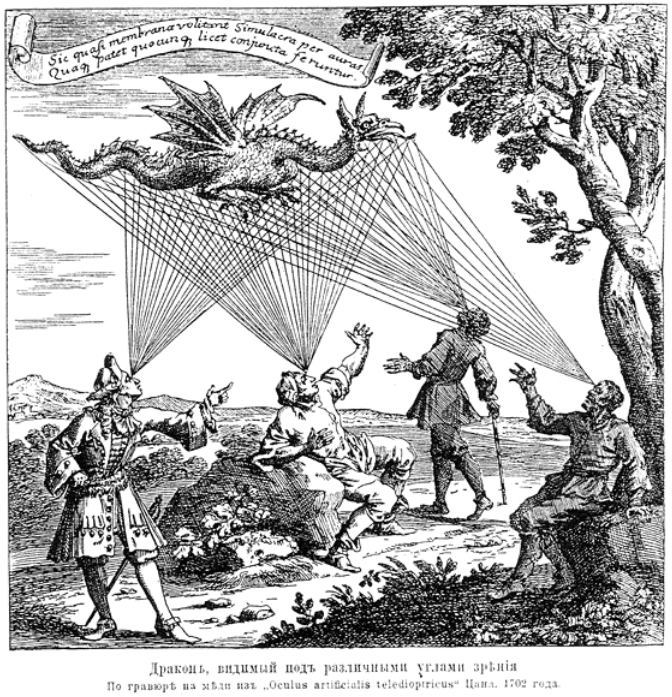
Alignment

III. Multi-view geometry



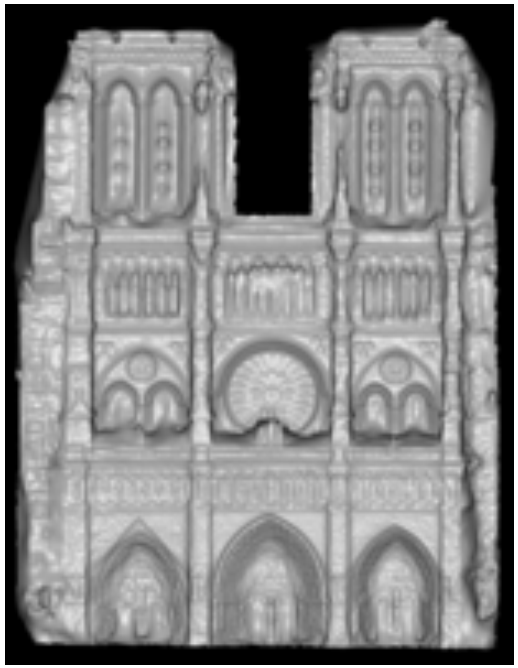
Epipolar geometry

Two-view stereo



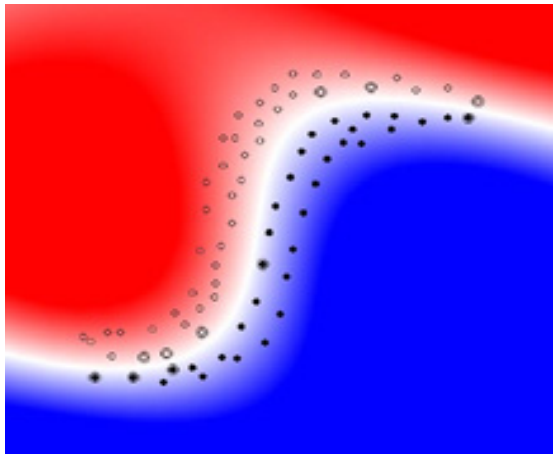
Драконъ, видимый подъ различными углами зрѣнія
По гравюрѣ на мѣди изъ „Oculus artificialis teledioptricus“ Цана. 1702 года.

Structure from motion

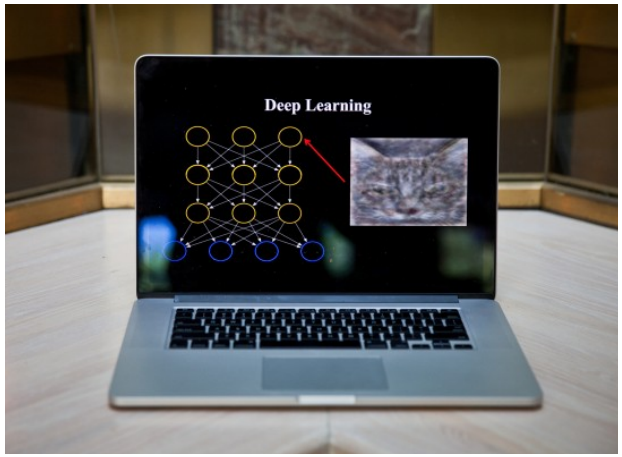


Multi-view stereo

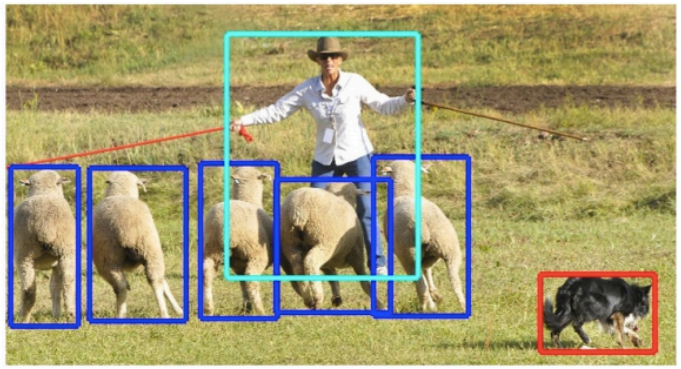
IV. Recognition



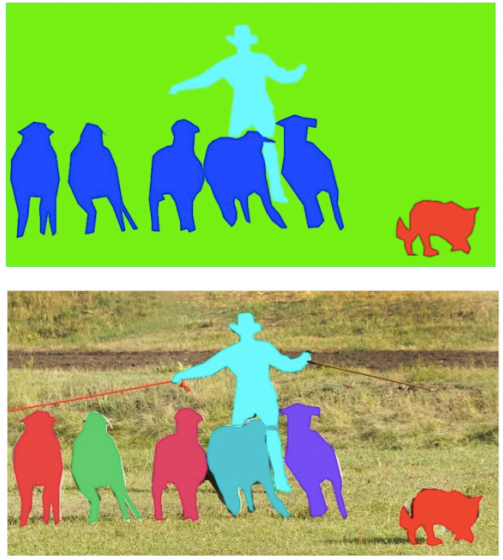
Basic classification



Deep learning



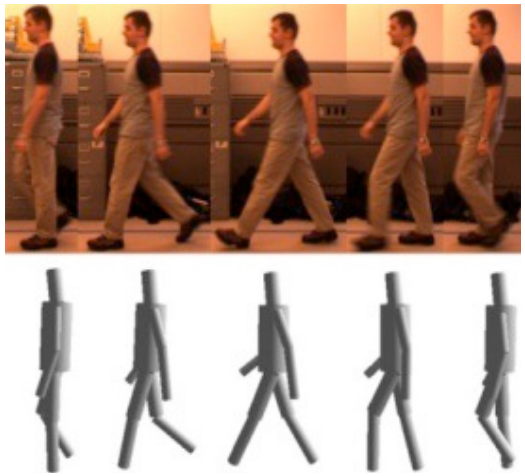
Object detection



Segmentation

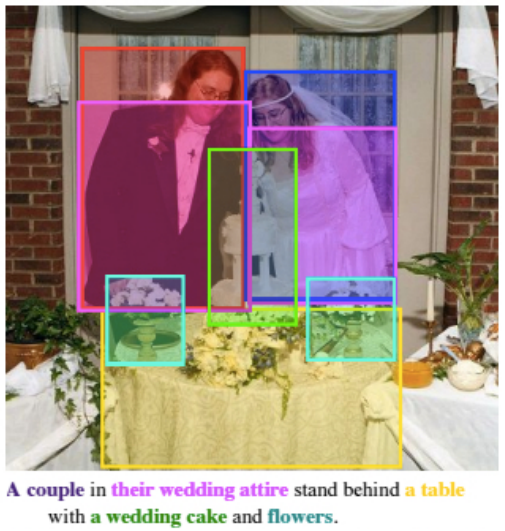
V. Additional Topics

Generation



Video

3D scene understanding



Images and text