

Convolutional Neural Networks

Deep Learning & Applications

Common Applications of CNNs

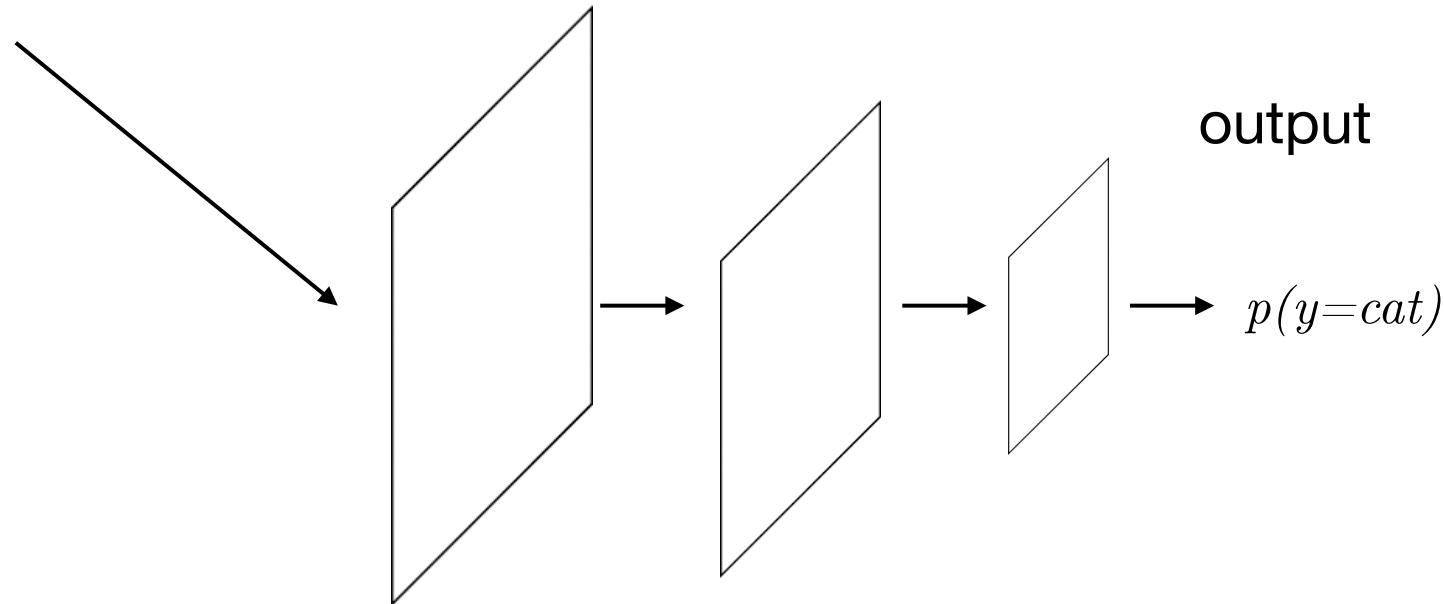
CNNs for Image Classification



Image Source:
twitter.com%2Fcats&psig=AOvVaw30_o-PCM-K21DiMAJQimQ4&ust=1553887775741551



Image Source: <https://www.pinterest.com/pin/244742560974520446>



Object Detection



Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779-788).

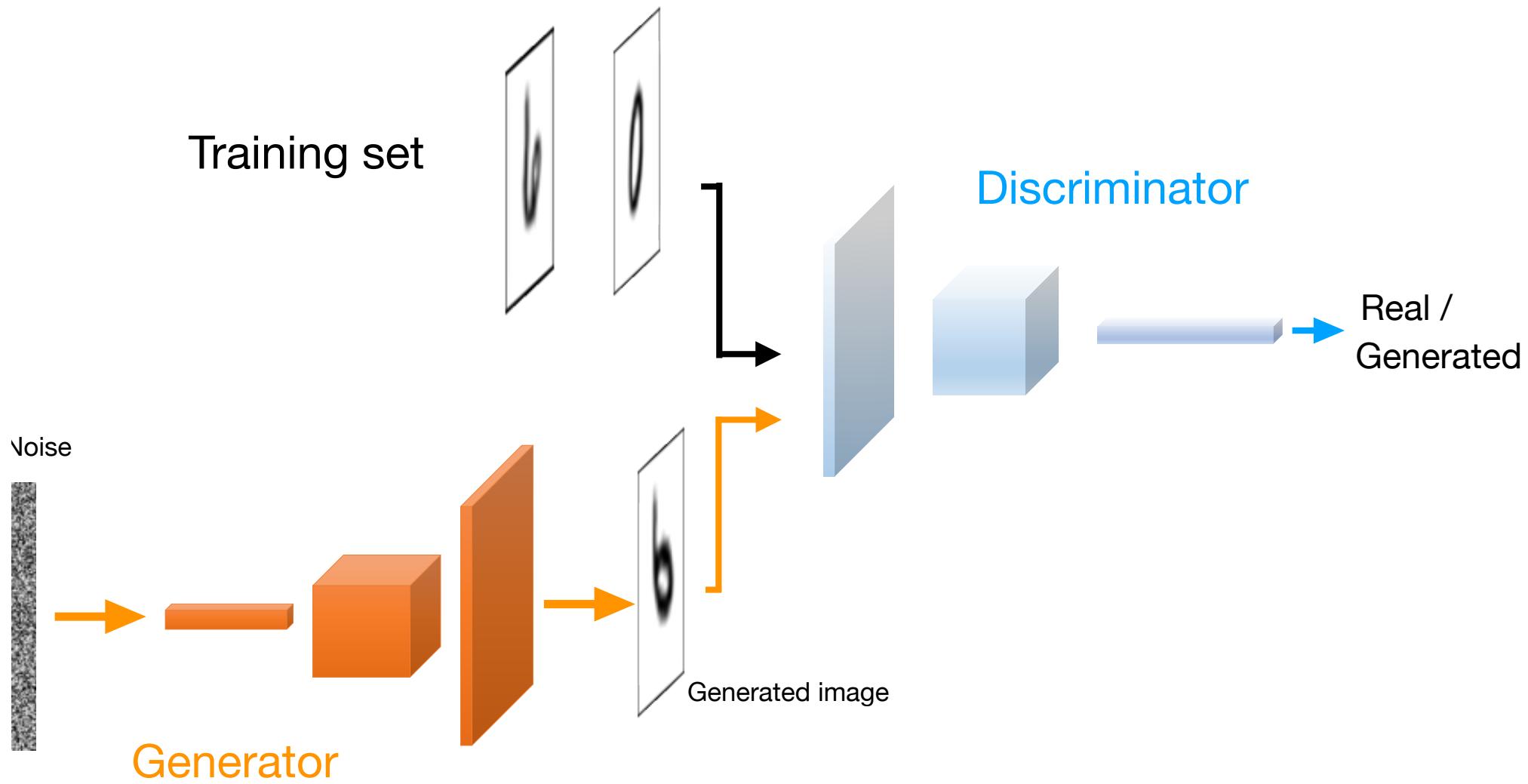
Object Segmentation



Figure 2. **Mask R-CNN** results on the COCO test set. These results are based on ResNet-101 [15], achieving a *mask AP* of 35.7 and running at 5 fps. Masks are shown in color, and bounding box, category, and confidences are also shown.

He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "Mask R-CNN." In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2961-2969. 2017.

Image Synthesis



Human Detection



Face Recognition

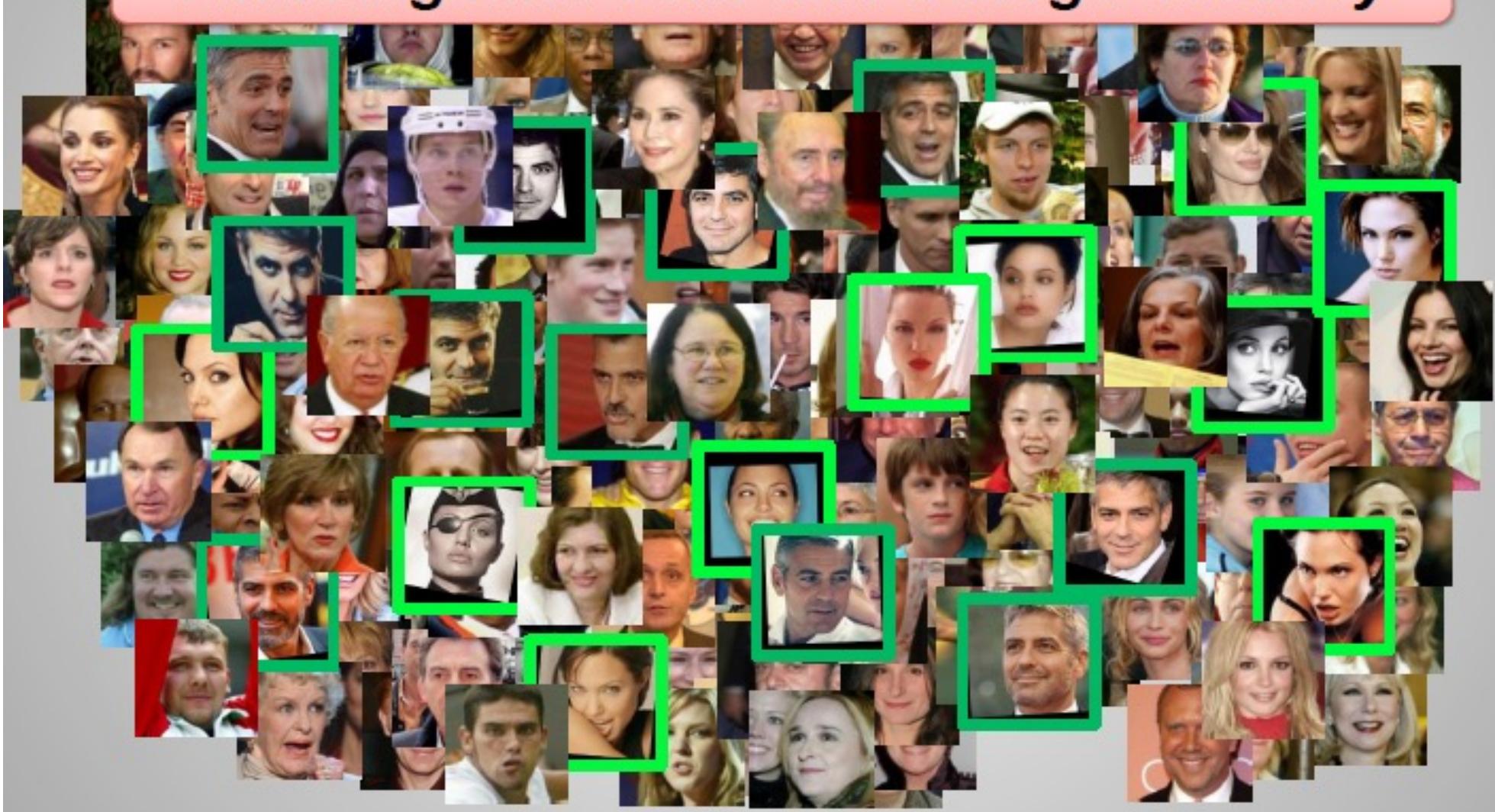


Open-Universe Face Identification



Open-Universe Face Identification

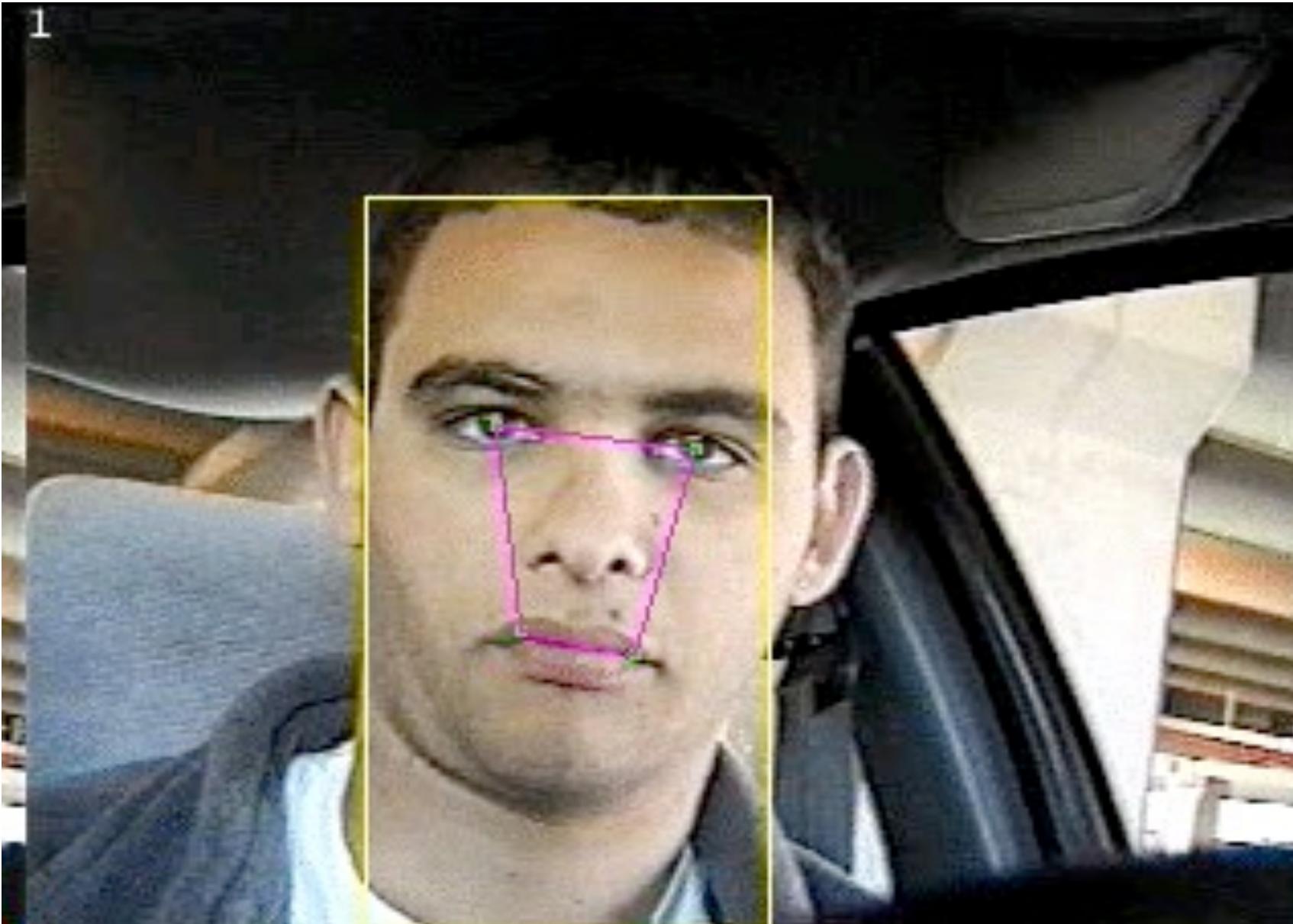
Find Angelina Jolie and George Clooney



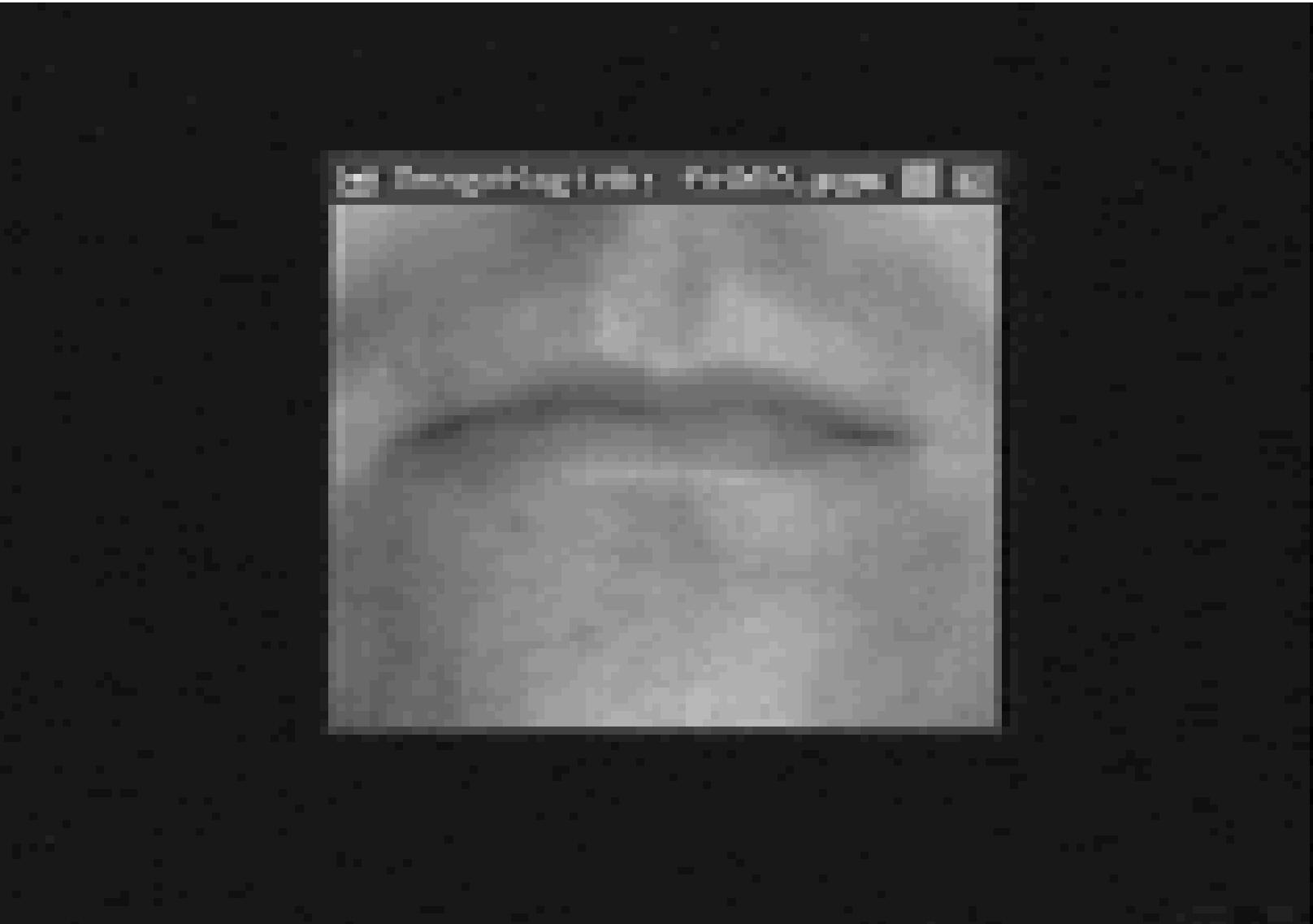
Facial expression



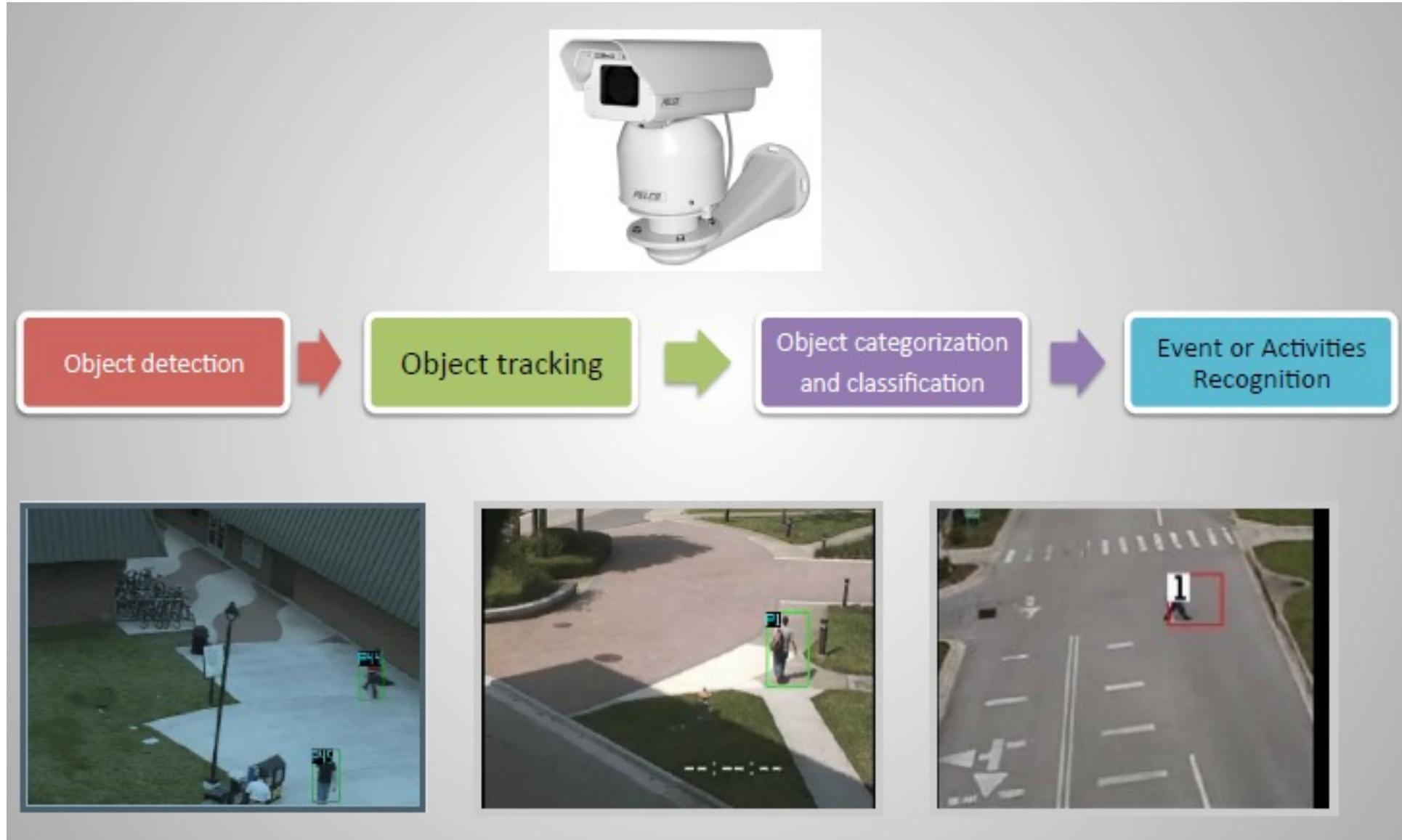
Fatigue detection



Lip-reading



Video Surveillance and Monitoring



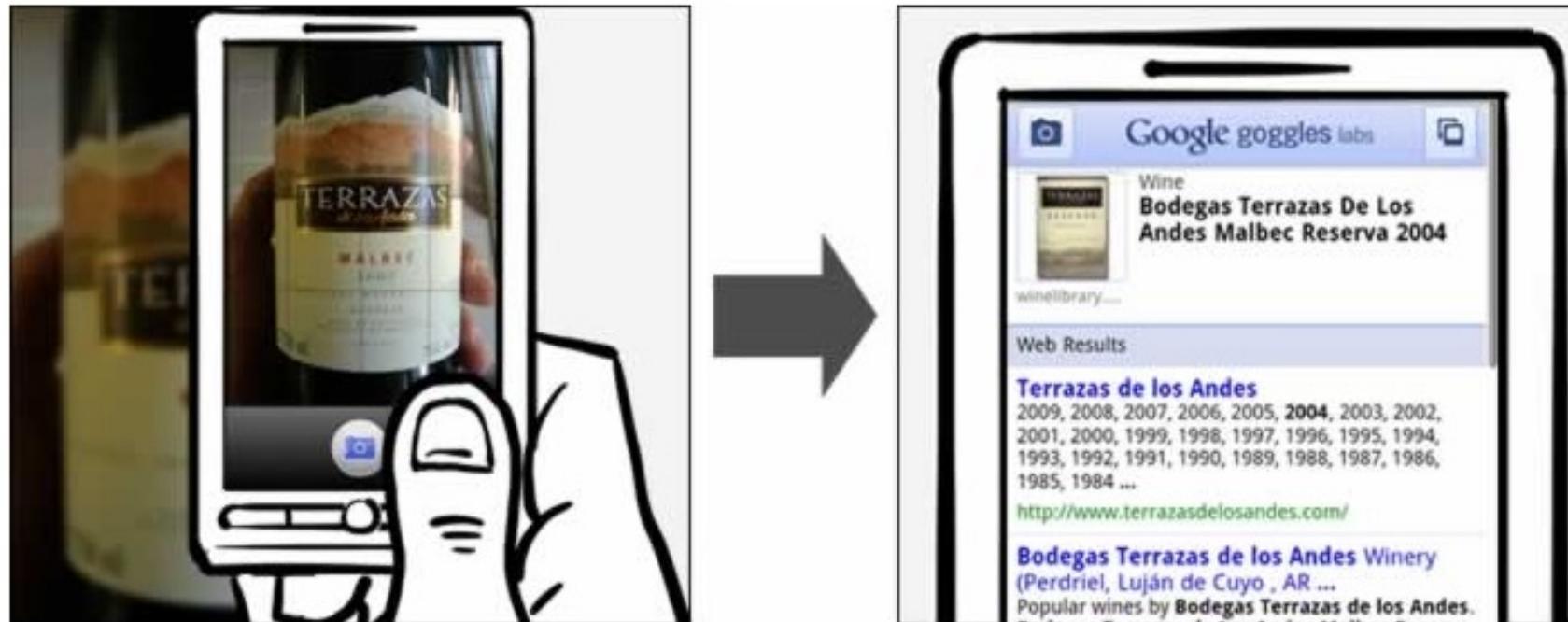
Unmanned Aerial Vehicles (drones)



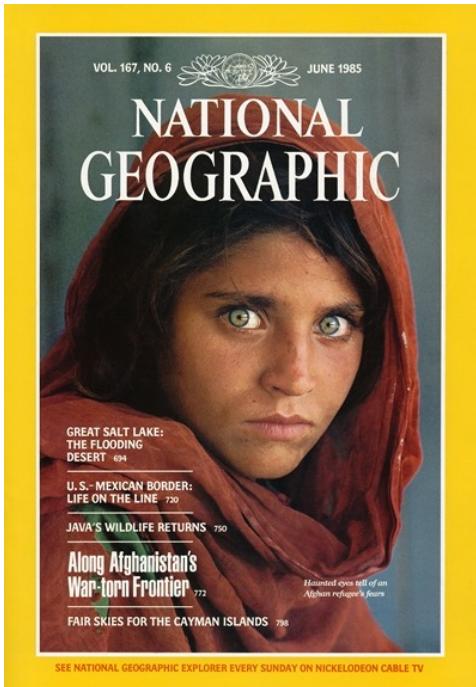
Mobile visual search: Google Goggles

Google Goggles in Action

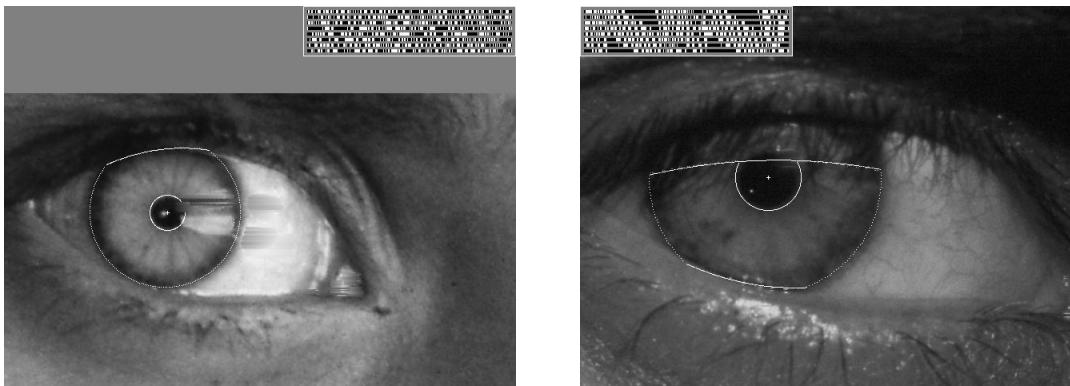
Click the icons below to see the different ways Google Goggles can be used.



Biometrics



How the Afghan Girl was Identified by Her Iris Patterns



Automotive safety

The screenshot shows the Mobileye website interface. At the top, there are navigation tabs for "manufacturer products" and "consumer products". Below this, a main heading reads "Our Vision. Your Safety." with a subtext "A camera-based vision system for advanced driver assistance". A central image shows a car from above with three cameras highlighted: "rear looking camera" at the back, "forward looking camera" at the front, and "side looking camera" on the sides. To the right, there's a "News" section with links to articles about Volvo's collision warning system and a "News" link. Below the main heading, there are three product cards: "EyeQ Vision on a Chip" (with an image of a chip), "Vision Applications" (with an image of a pedestrian crossing), and "AWS Advance Warning System" (with an image of a display screen). Each card has a "read more" link.

- > [manufacturer products](#)
- < consumer products <<
- Our Vision. Your Safety.**
- rear looking camera
- forward looking camera
- side looking camera
- > [EyeQ Vision on a Chip](#)
- > [Vision Applications](#)
- Road, Vehicle, Pedestrian Protection and more
- > [AWS Advance Warning System](#)
- > [Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System](#)
- > [Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end](#)
- > [all news](#)
- > [Mobileye at Equip Auto, Paris, France](#)
- > [Mobileye at SEMA, Las Vegas, NV](#)
- > [read more](#)

- [Mobileye](#): Vision systems in high-end BMW, GM, Volvo models
 - Pedestrian collision warning
 - Forward collision warning
 - Lane departure warning
 - Headway monitoring and warning

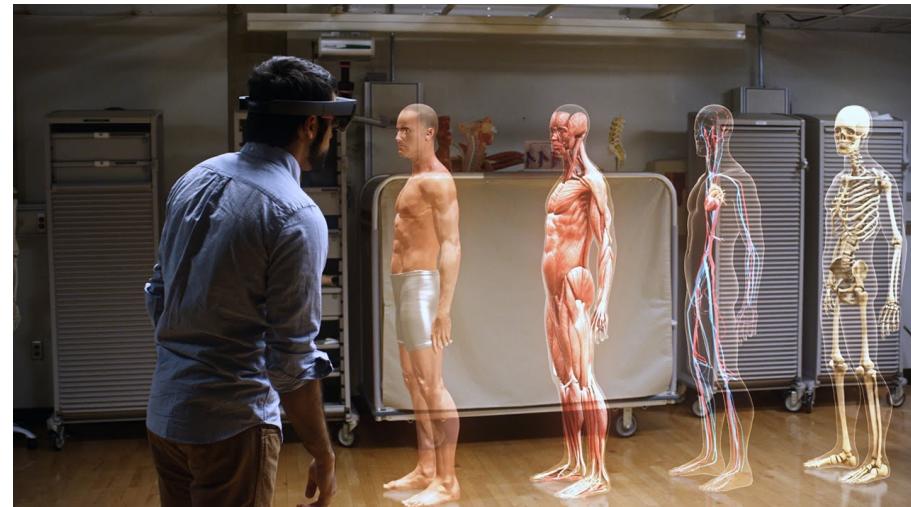
AutoCars - Uber



Mobile robots

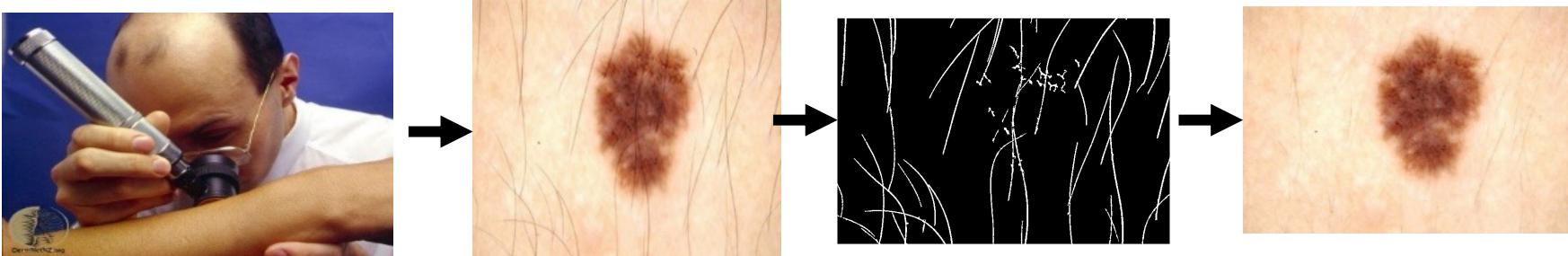
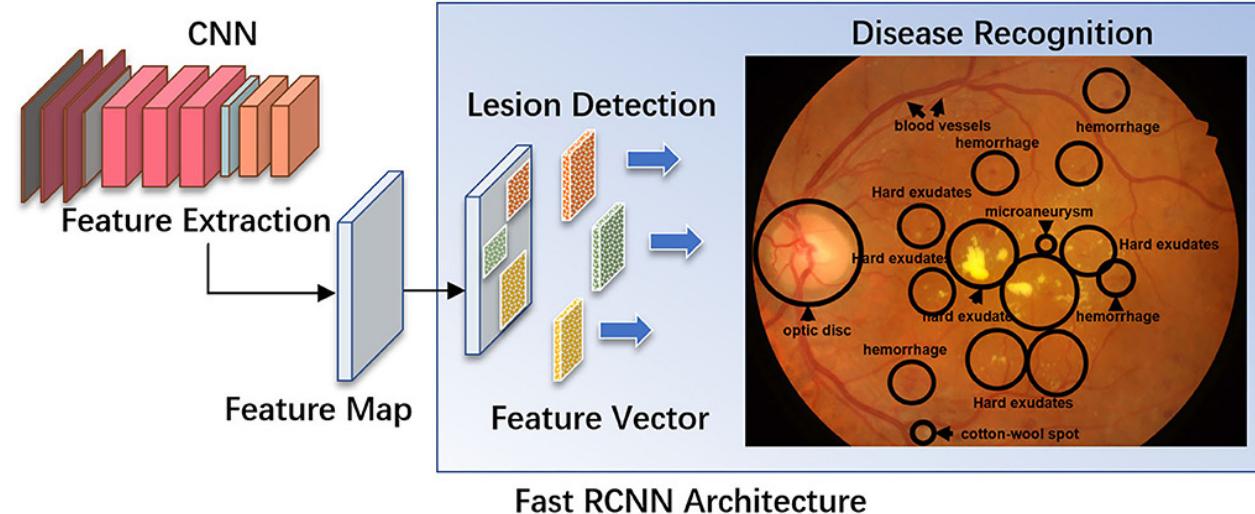
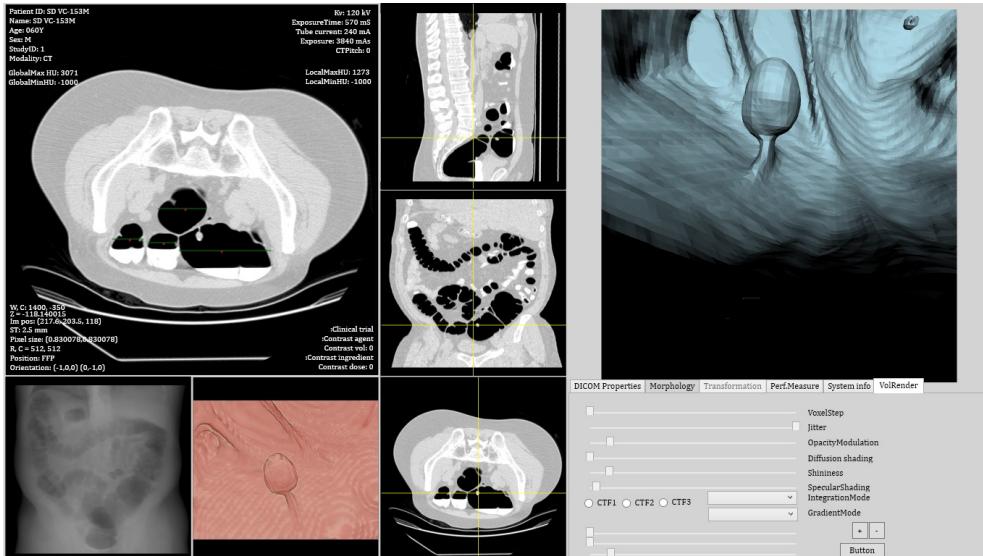


Augmented Reality and Virtual Reality

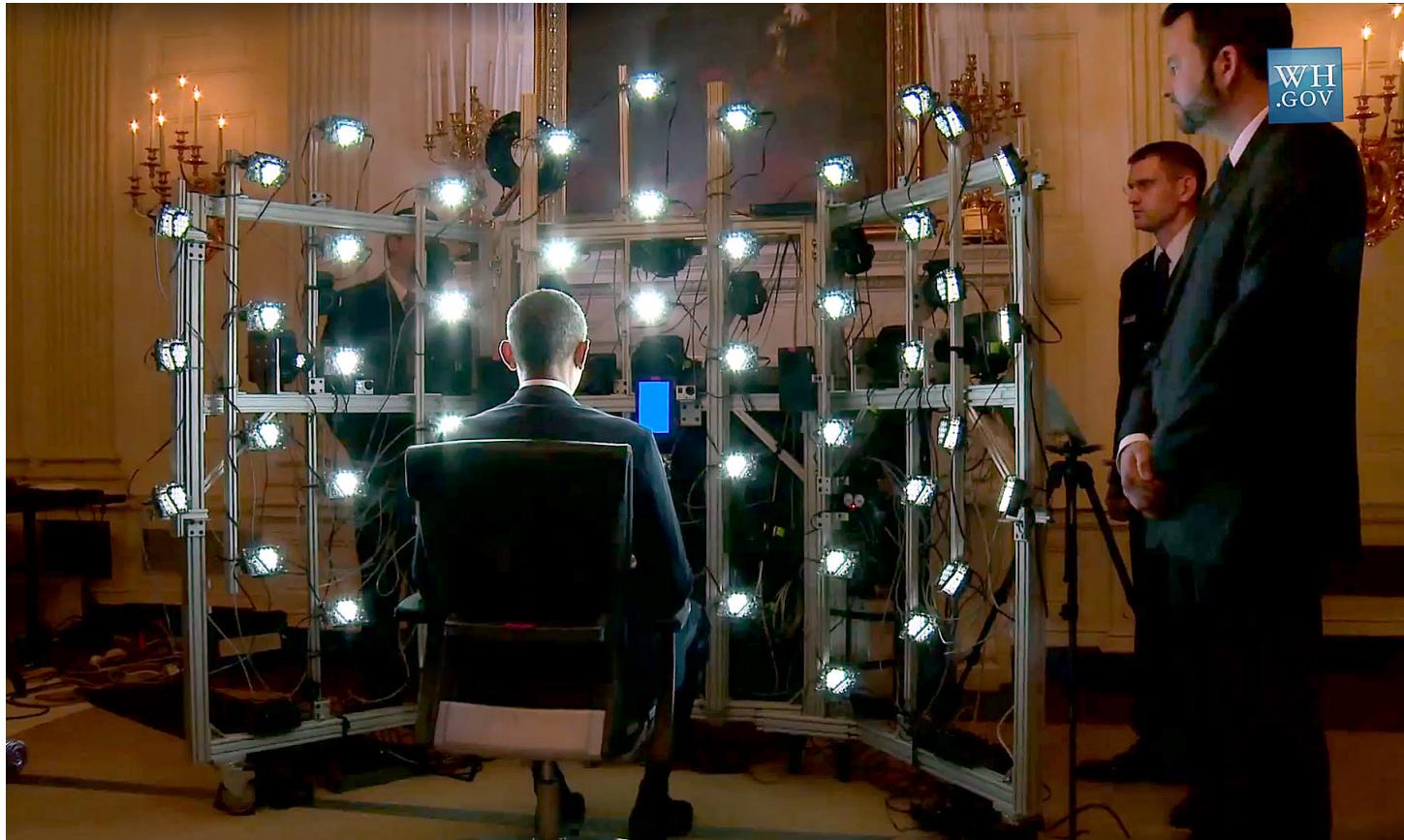


MS HoloLens, Oculus, Magic Leap,

Computer Aided Diagnosis



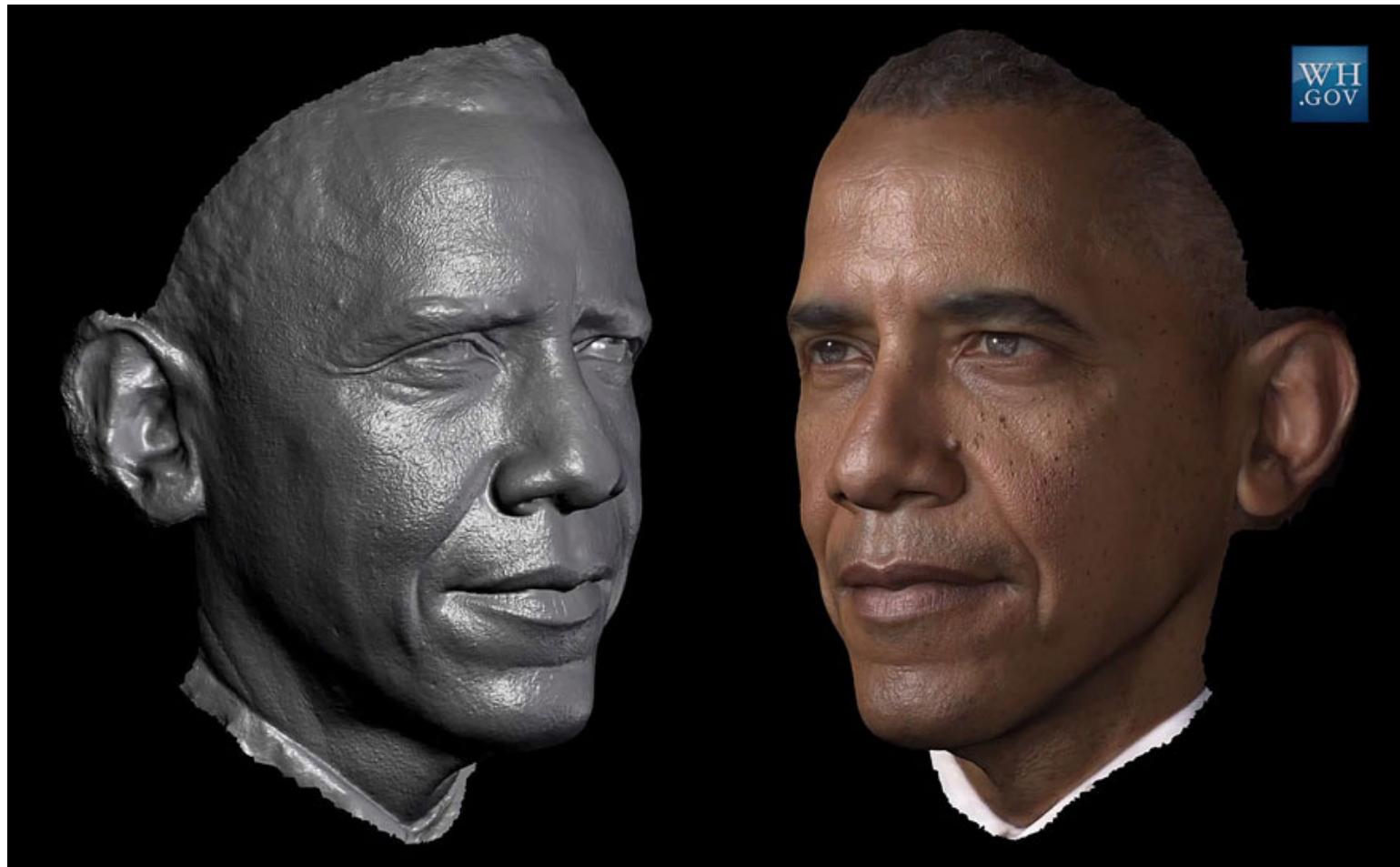
Human shape capture



Human shape capture



Human shape capture



Few Questions!

Can we build hardware as complex as the brain?

- How complicated is our brain?
 - a neuron, or nerve cell, is the basic information processing unit
 - estimated to be on the order of 10^{12} neurons in a human brain
 - many more synapses (10^{14}) connecting these neurons
- Conclusion
 - YES: in the near future we can have computers with as many basic processing elements as our brain, but with
 - far fewer interconnections (wires or synapses) than the brain
 - much faster updates than the brain

Can Computers Talk?

- This is known as “speech synthesis”
- Difficulties
 - sounds made by this “lookup” approach sound unnatural
 - a harder problem is emphasis, emotion, etc
 - humans understand what they are saying
 - machines don’t: so they sound unnatural
- Conclusion:
 - NO, for complete sentences with emotion
 - YES, for individual words

Can Computers Recognize Speech?

- Speech Recognition:
 - mapping sounds from a microphone into a list of words
 - classic problem in AI, very difficult
- Recognizing normal speech is much more difficult
 - speech is continuous: where are the boundaries between words?
 - large vocabularies
 - background noise, other speakers, accents, cold, etc
- Conclusion:
 - NO, normal speech is too complex to accurately recognize
 - YES, for restricted problems (small vocabulary, single speaker)

Can Computers Understand speech?

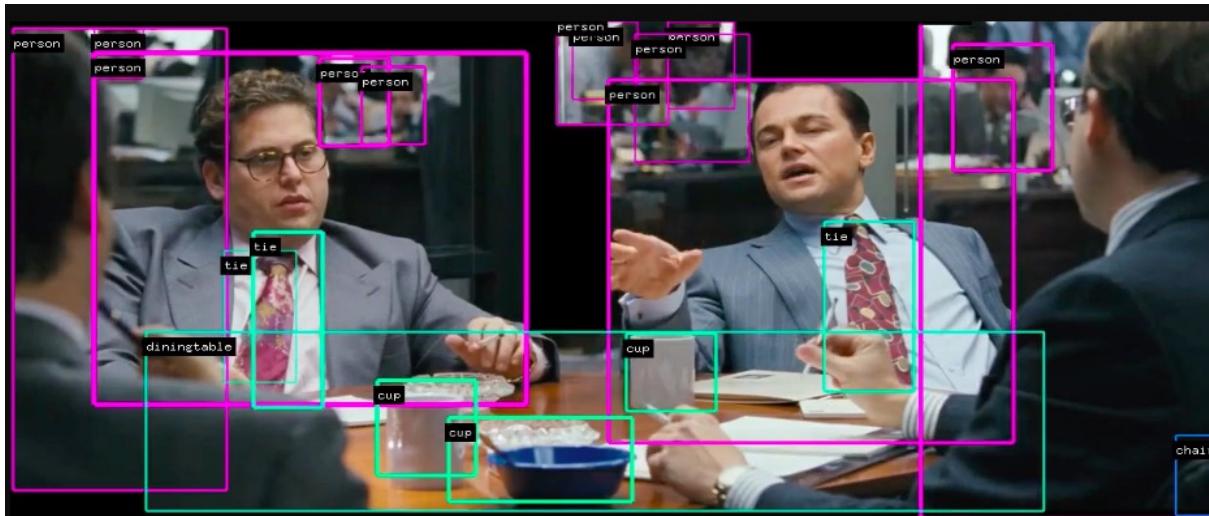
- Understanding is different to recognition:
 - “Time flies like an arrow”
 - assume the computer can recognize all the words
 - how many different interpretations are there?
 1. time passes quickly like an arrow?
 2. command: time the flies the way an arrow times the flies
 3. command: only time those flies which are like an arrow
 4. “time-flies” are fond of arrows
 - only 1. makes any sense,
 - but how could a computer figure this out?
 - clearly humans use a lot of implicit commonsense knowledge in communication
 - Conclusion: NO, much of what we say is beyond the capabilities of a computer to understand at present

Can Computers Learn and Adapt ?

- Learning and Adaptation
 - **Machine learning** allows computers to learn to do things without explicit programming
 - many successful applications:
- Conclusion: YES, computers can learn and adapt, when presented with information in the appropriate way

Can Computers “see”?

- Recognition v. Understanding
 - Recognition and Understanding of Objects in a scene



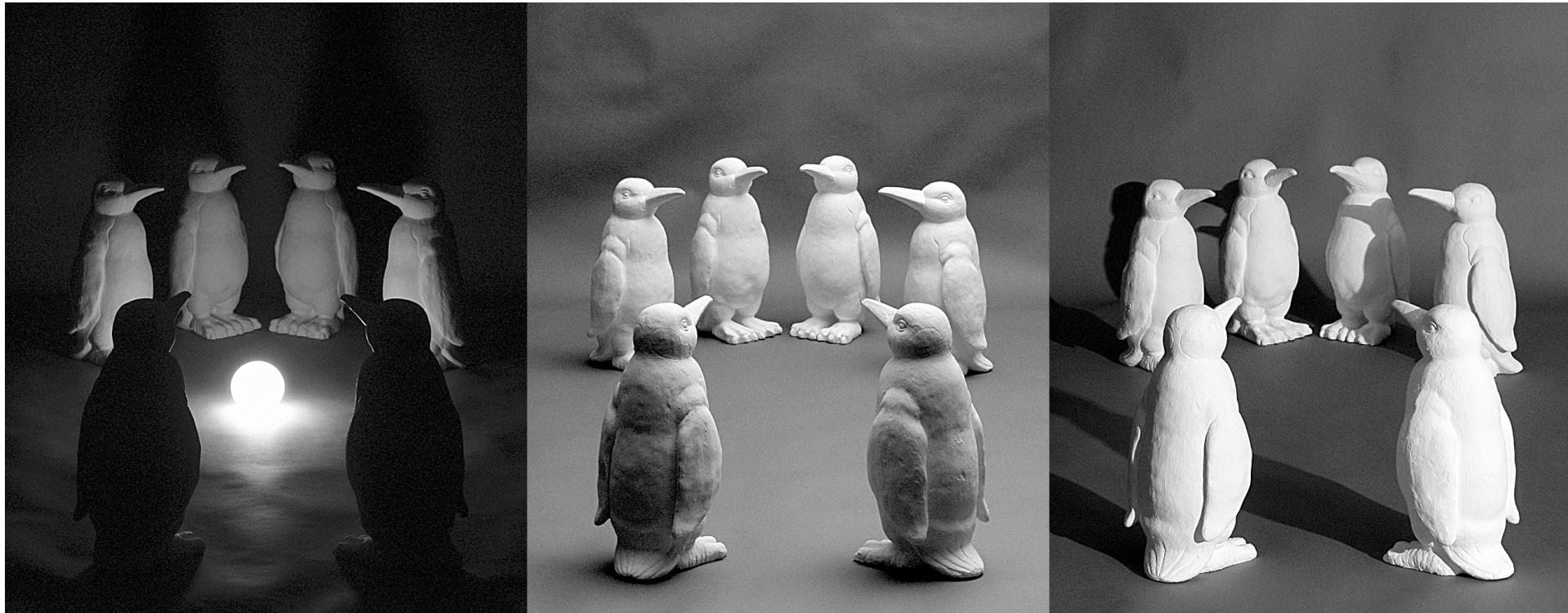
- Why is visual recognition a hard problem?

Challenges: viewpoint variation

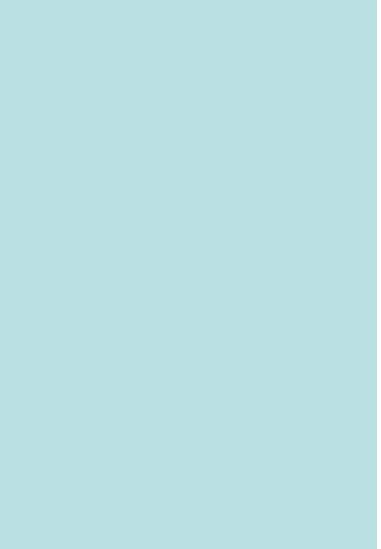


Michelangelo 1475-1564

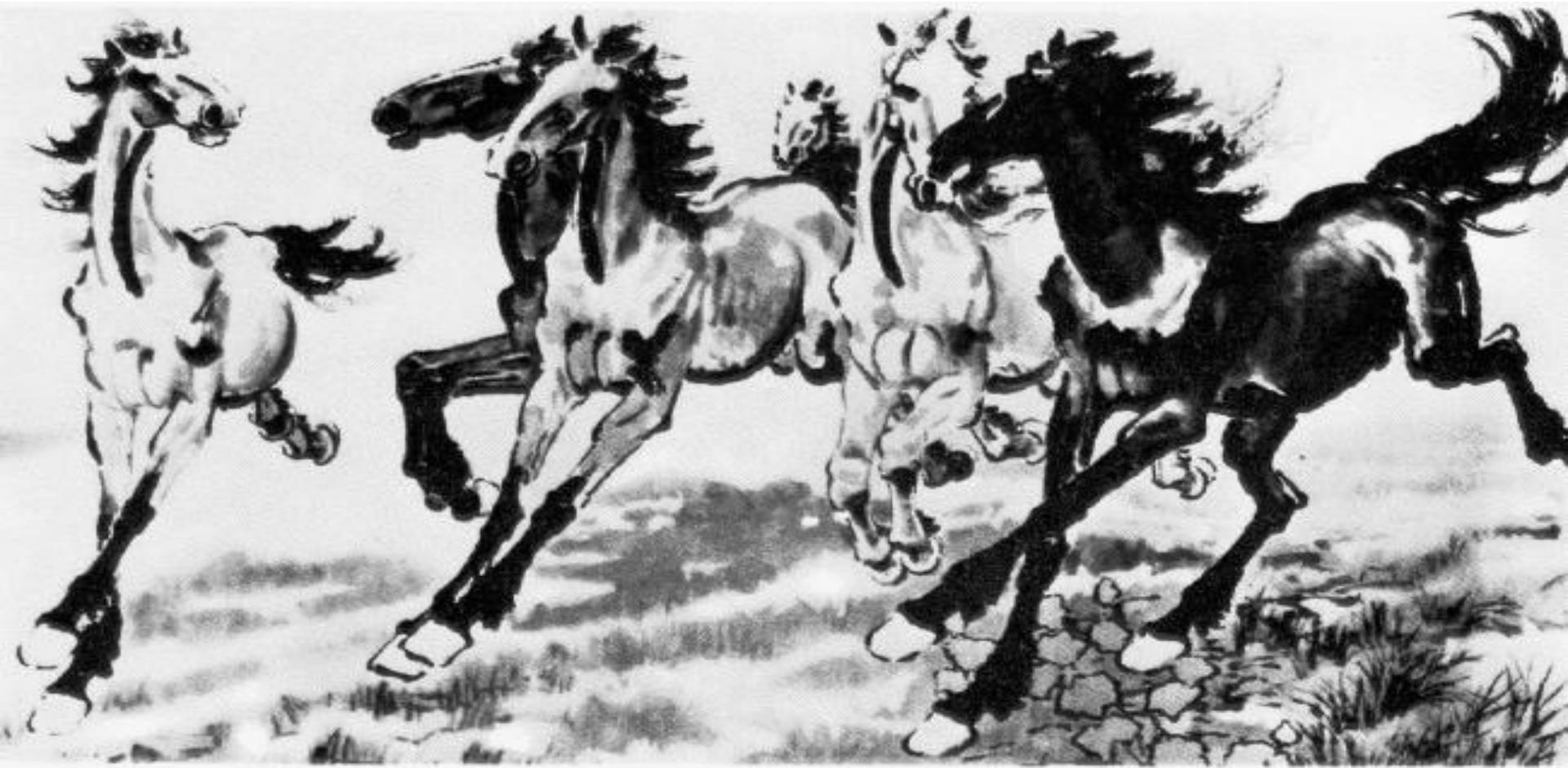
Challenges: illumination



Challenges: scale



Challenges: deformation



Challenges: occlusion



Magritte, 1957

Challenges: background clutter



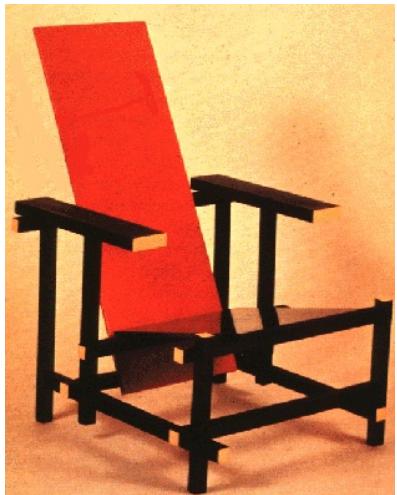
Emperor shrimp and commensal crab on a sea cucumber in Fiji
Photograph by Tim Laman

NATIONAL
GEOGRAPHIC

Challenges: Motion



Challenges: object intra-class variation



slide credit: Fei-Fei, Fergus & Torralba