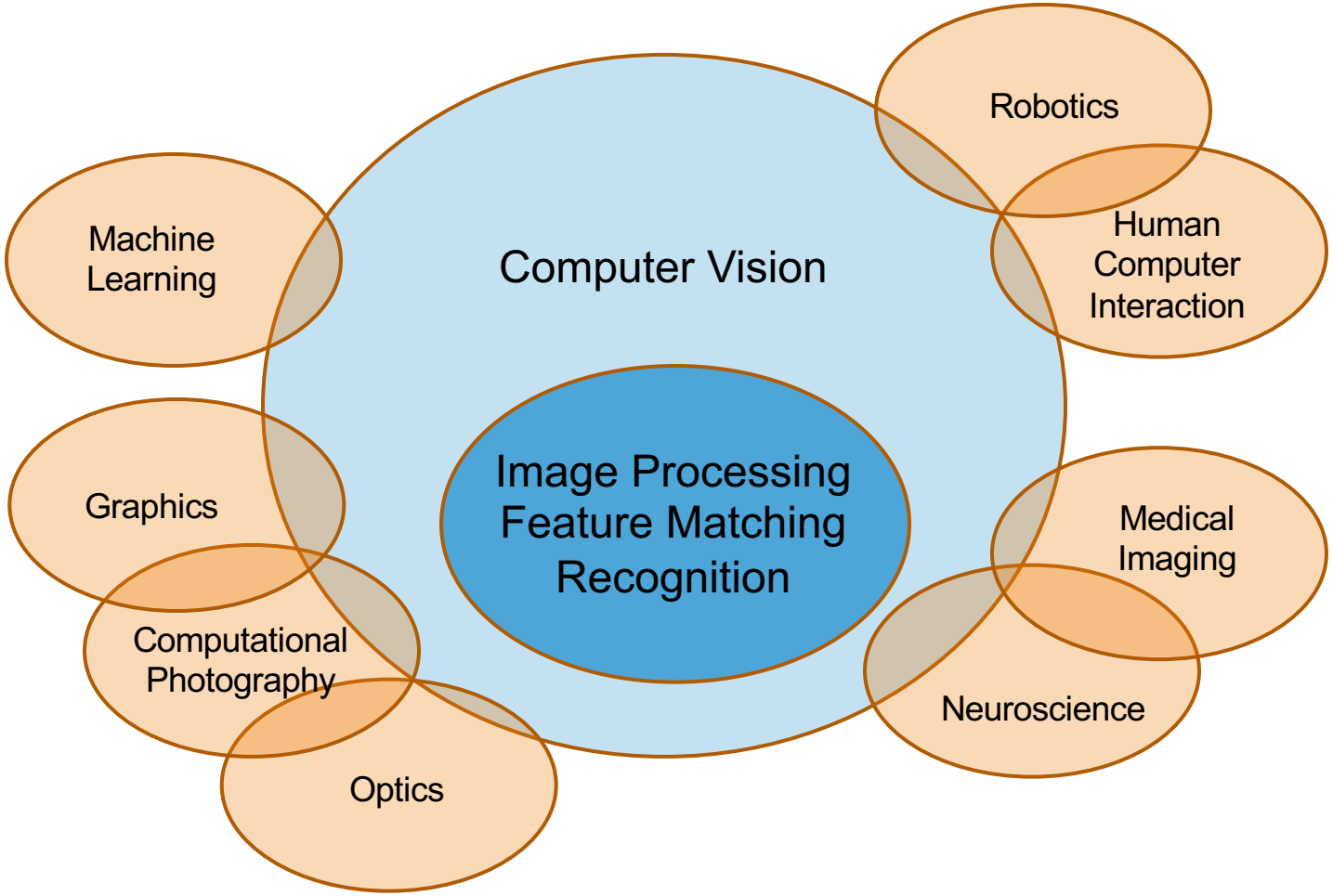
Introduction to Compute Vision

[Some slides were adapted from slides of CV2 course by Prof. Jingya Wang]

Today's Class

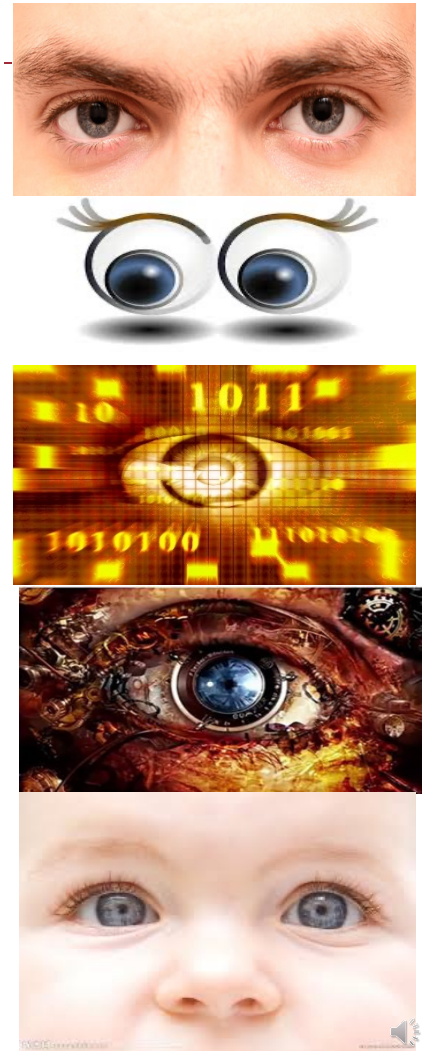
- What is Computer Vision?
- What are the applications of computer vision?
- Some topics.

Relationship with several nearby Research Area



What is Computer Vision?

- What are examples of computer vision being used in the world?



Computer Vision

Make computers understand images and video.

What kind of scene?

Where are the cars?

How far is the building?

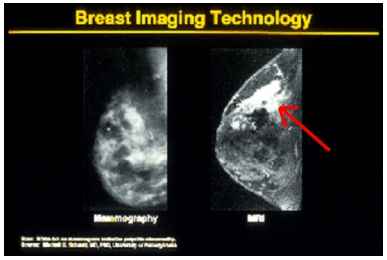
…

Vision is really hard

- Vision is an amazing feat of natural intelligence
  - Visual cortex occupies about 50% of the brain
  - More human brain devoted to vision than anything else

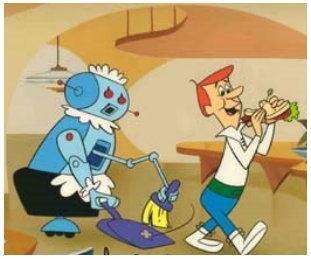
Why computer vision matters

Safety



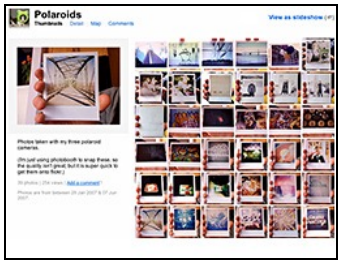
Health

Security



Comfort

Fun



Access

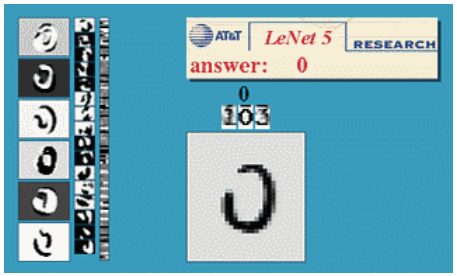
The Applications of Computer Vision

- Optical character recognition
- Face detection
- Image editing
- Medical image analysis
- . . .

Optical character recognition (OCR)

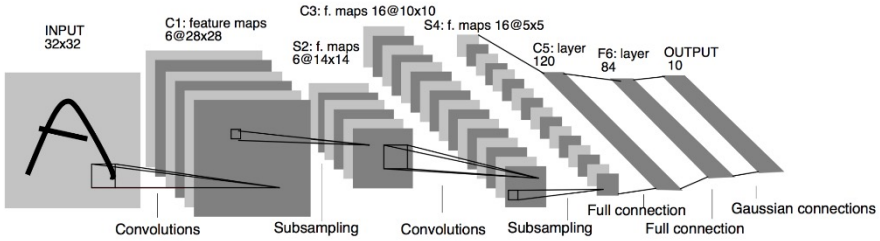
Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

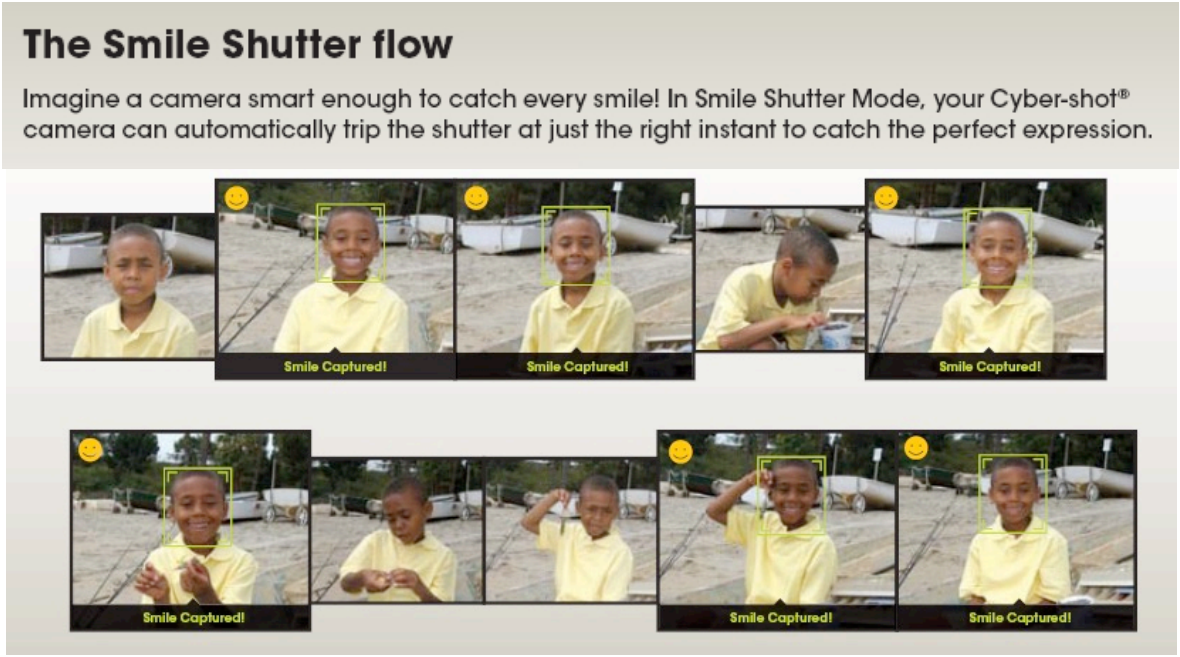


LeNet-5 [1998, paper by LeCun et al.]

Face detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

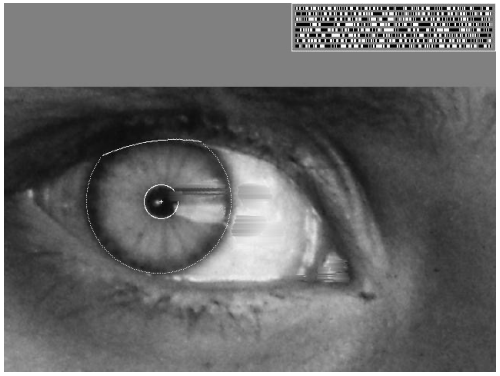
Smile detection

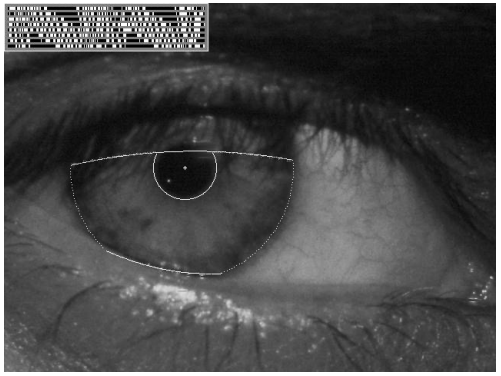


Sony Cyber-shot® T70 Digital Still Camera

Vision-based biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story wikipedia





Login without a password…

Fingerprint scanners on many new laptops, other devices

Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Object recognition (in mobile phones)

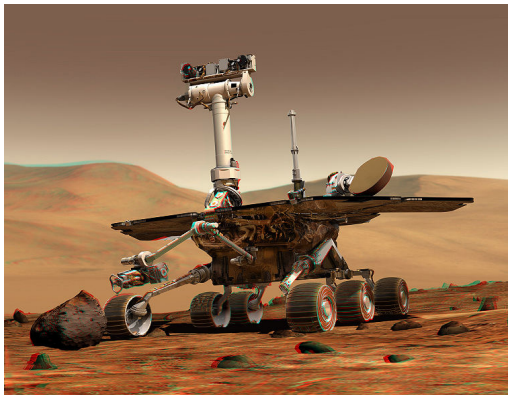
Point & Find, Nokia

Google Goggles

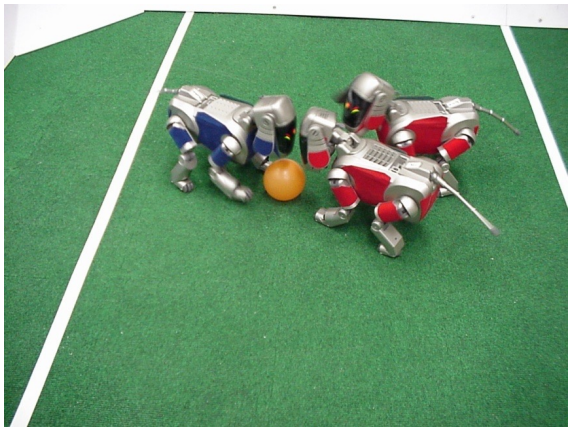
Industrial robots

Vision-guided robots position nut runners on wheels

Mobile robots



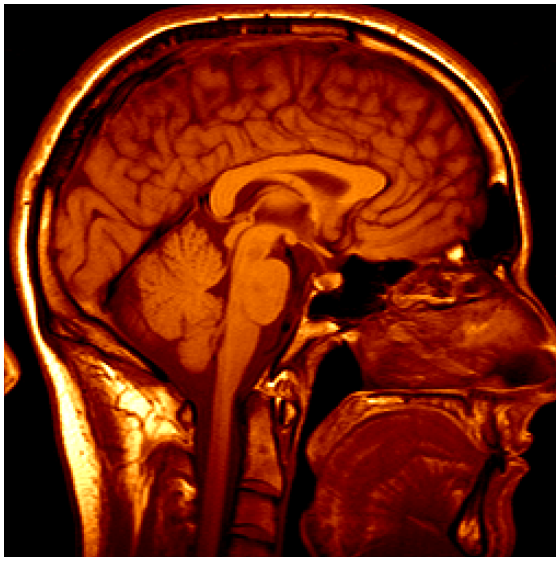
NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

Saxena et al. 2008
STAIR at Stanford

Medical imaging



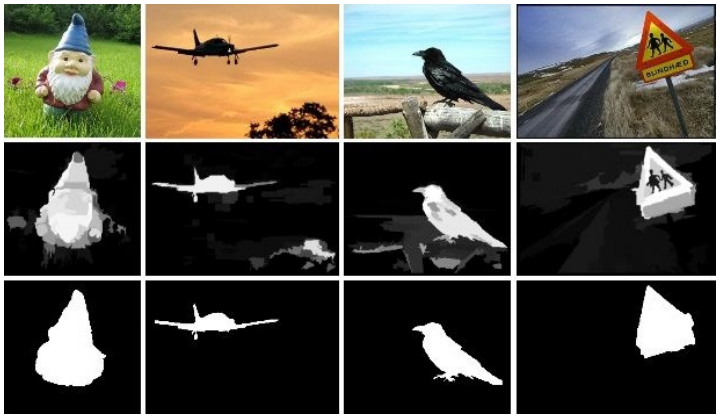
3D imaging
MRI, CT

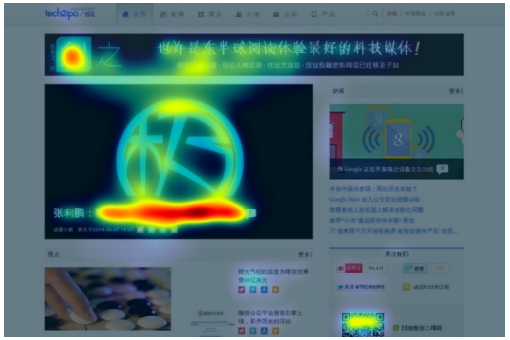
Image guided surgery
Grimson et al., MIT

Topics

- **saliency detection**
  - **unsupervised saliency detection/salient object discovery)**
- **segmentation**
  - **(unsupervised image segmentation, semantic segmentation)**
- **object detection**
  - **(face detection/general object detection, supervised methods)**
- **object recognition**
  - **(face recognition/general image classification, (un)supervised methods)**
- **image generation**
  - **Generative adversarial network, style transfer, diffusion**
- **video processing**
  - **(tracking/event classification(supervised))**

Saliency detection





Application of computer vision

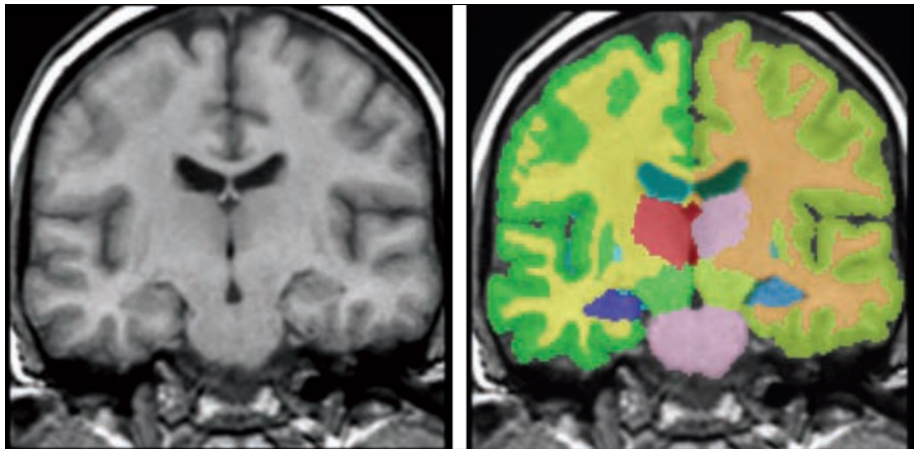
Content based image resizing (Image Retargeting)

https://www.youtube.com/watch?v=6NcIJXTIugc

Website Design

Image/Video Compression

Image Segmentation



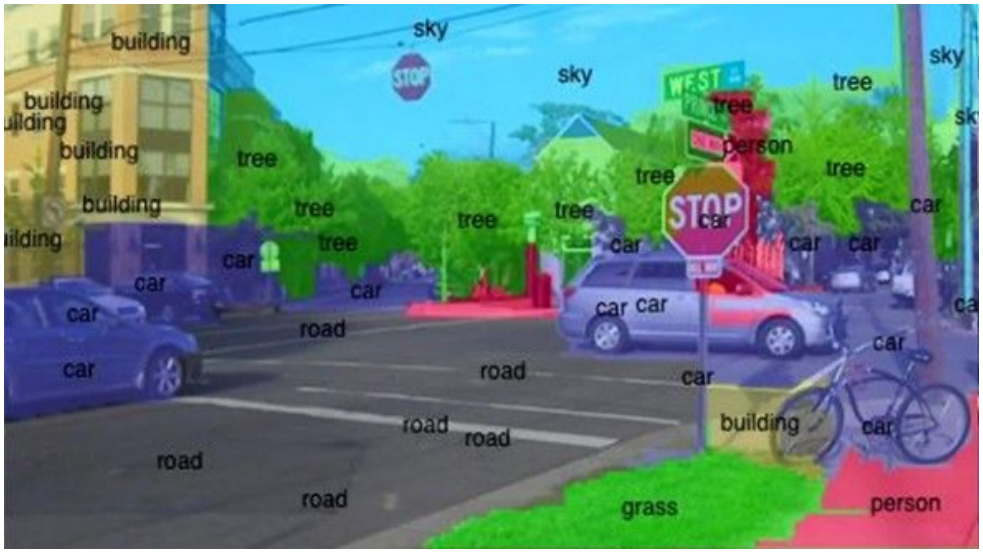
(a) Color Labels (ACA)

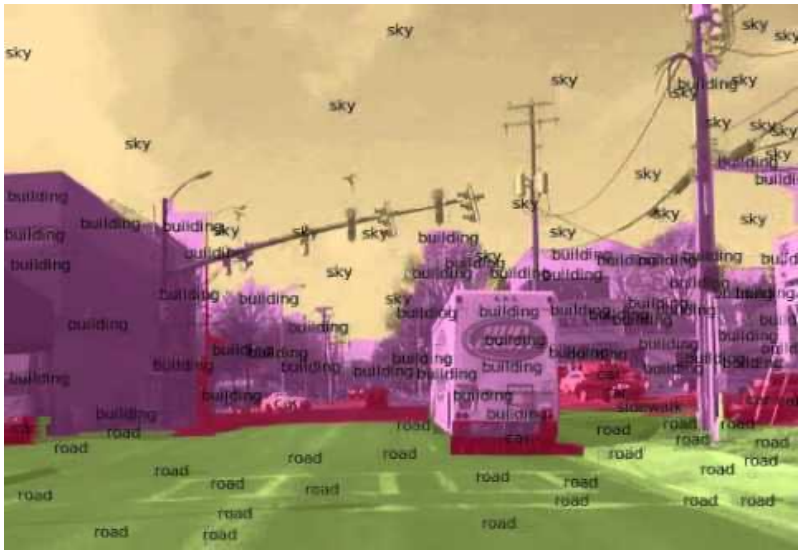
(b) Texture Classes

(c) Crude Segmentation

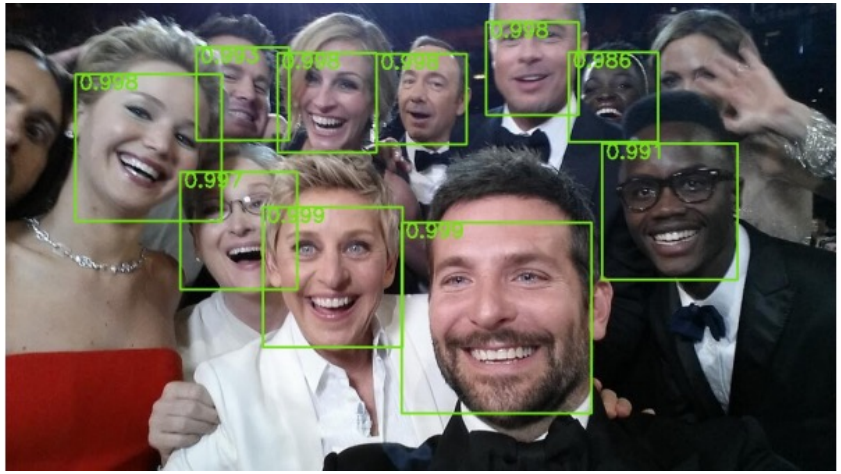
(d) Final Segmentation

Semantic segmentation

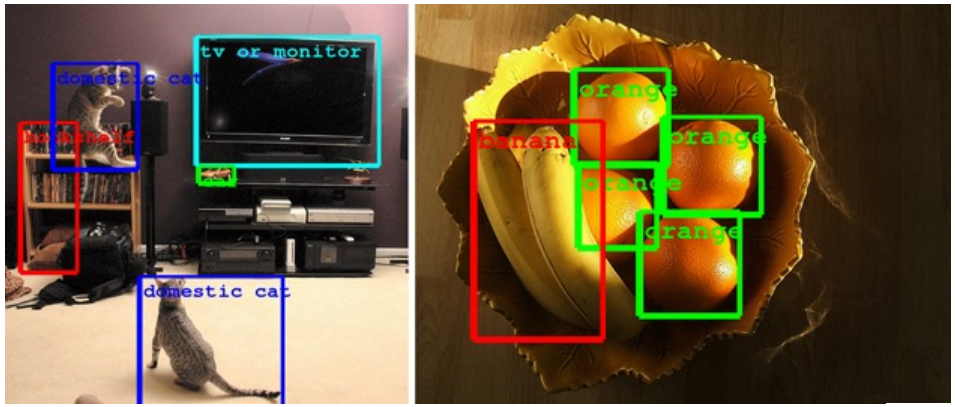




Object Detection

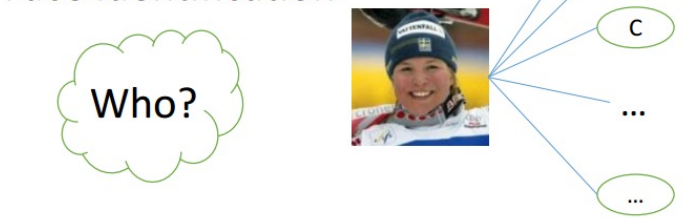


Object Recognition

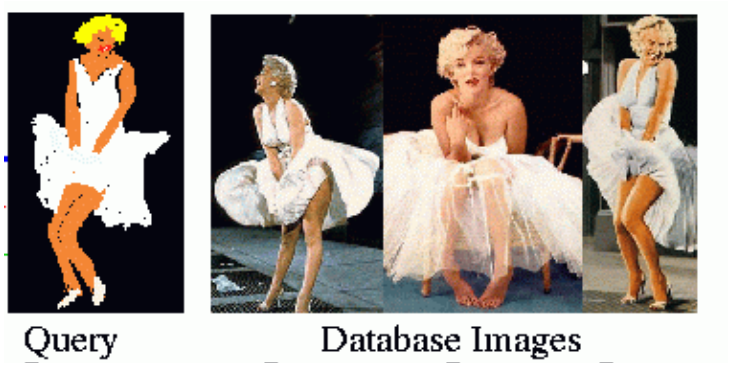


Face recognition

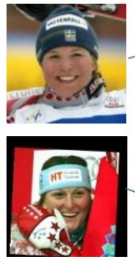
Face identification



Object Identification



Face verification



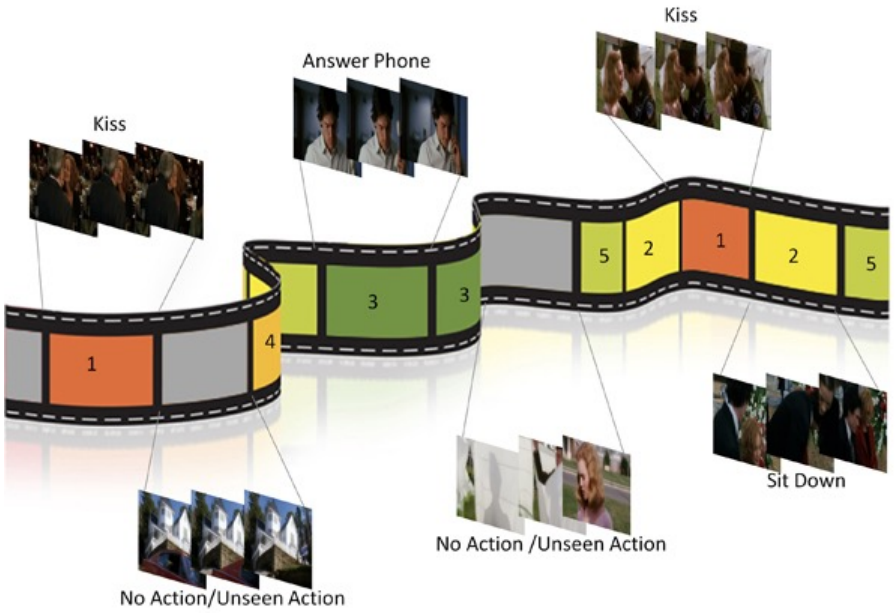
Same person or not.

Binary Result

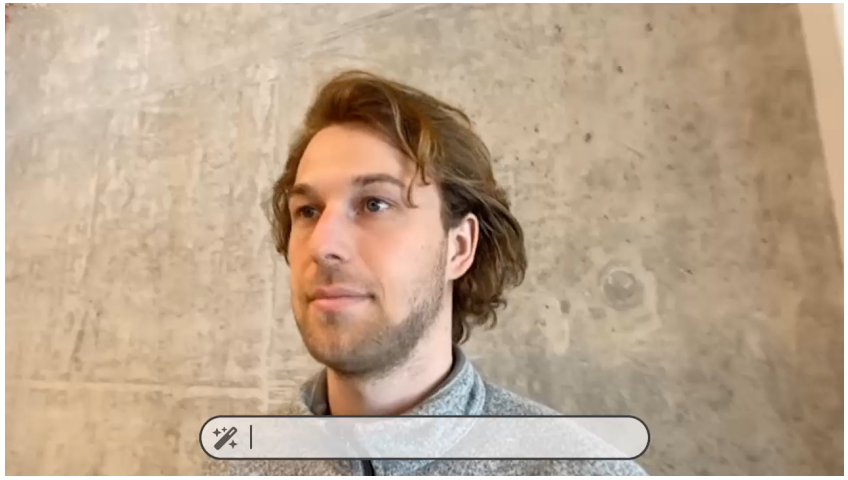
Generative Adversarial Network

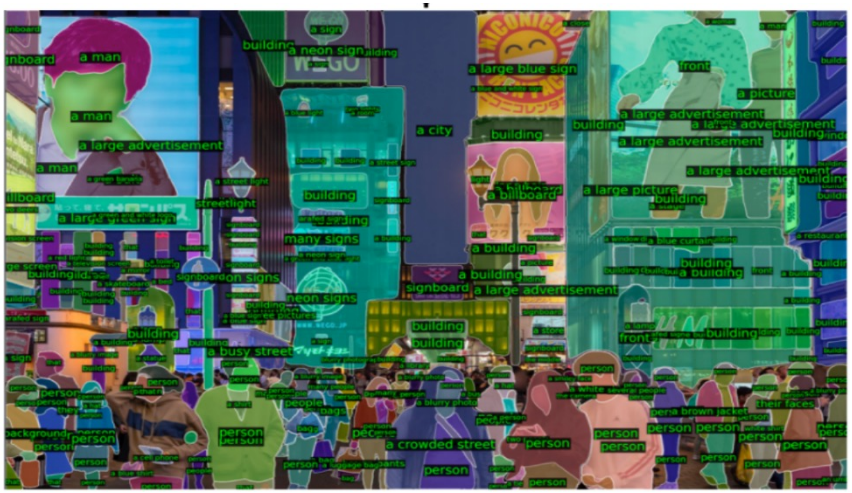
- Face generation
- Style transfer
- Scene generation
- https://www.youtube.com/watch?v=OGGjXG562WU

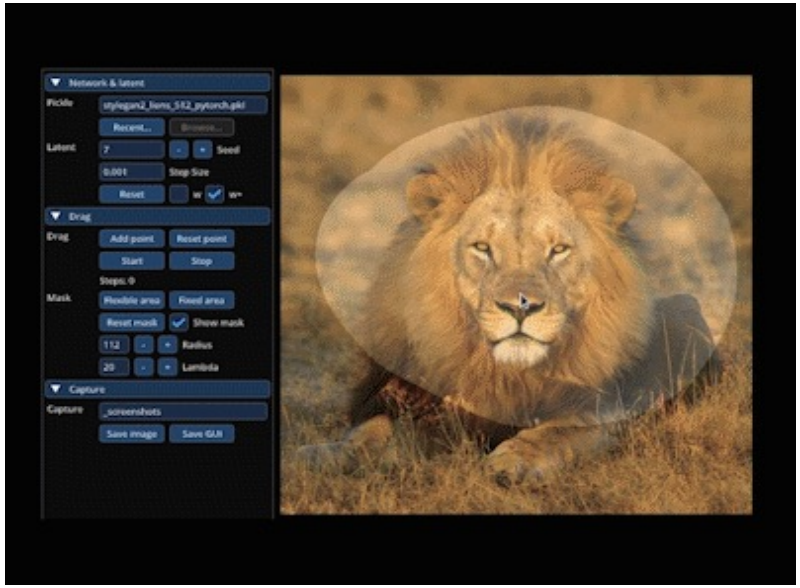
Video Classification



Today's CV

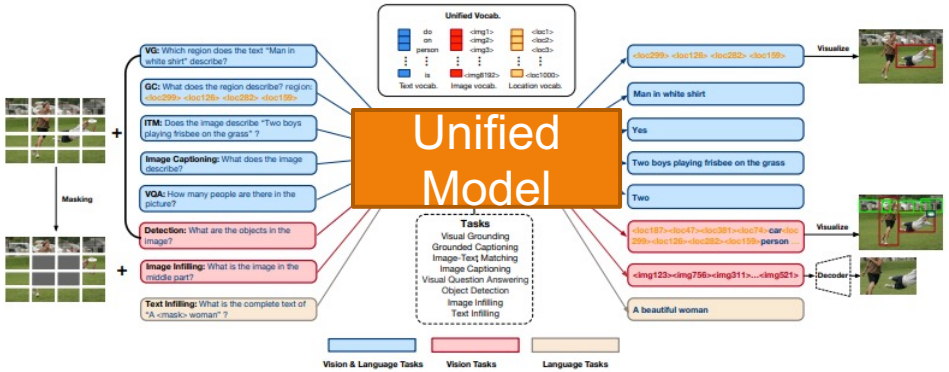



Enter comma-separated queries:

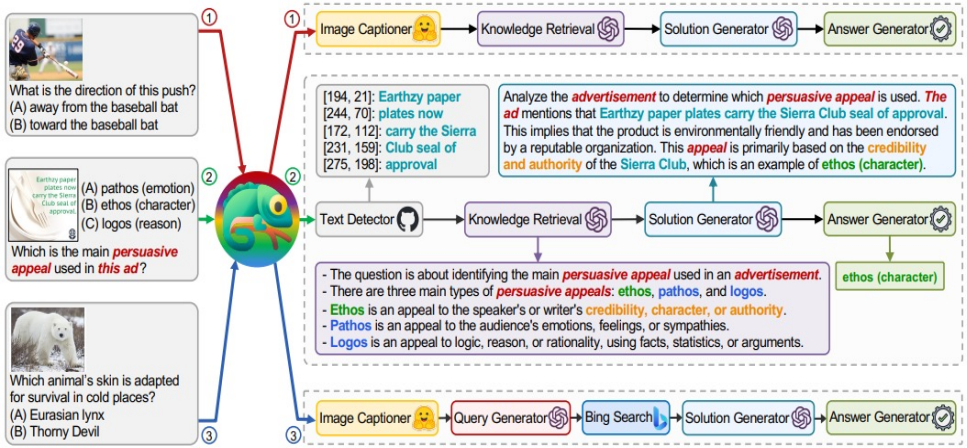

Network & latent
Pickle
stylegan2_lions_512_pytorch.pkl
Recent...
Latent
Seed
Step Size
Reset
Drag
Add point
Reset point
Start
Stop
Steps: 0
Mask
Flexible area
Fixed area
Reset mask
Show mask
Radius
Lambda
Capture
_screenshots
Save image
Save GUI

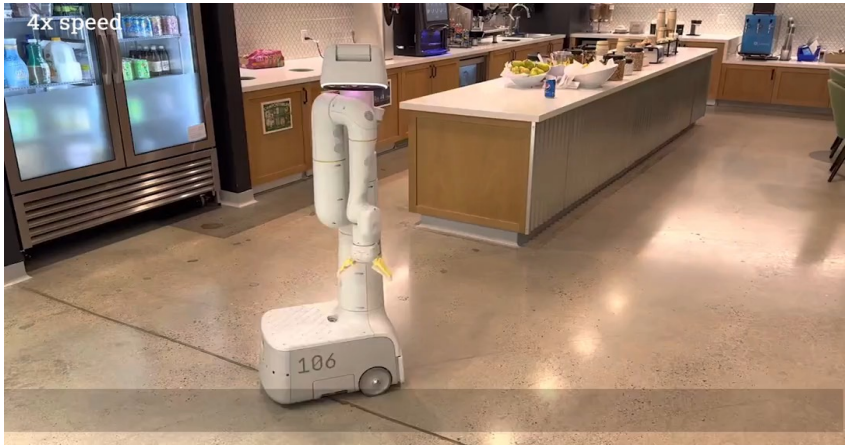
Today's CV

text-image pairs



Today's CV


What is the direction of this push?
(A) away from the baseball bat
(B) toward the baseball bat
(A) pathos (emotion)
(B) ethos (character)
(C) logos (reason)
Which is the main persuasive appeal used in this ad?
Which animal's skin is adapted for survival in cold places?
(A) Eurasian lynx
(B) Thorny Devil
Image Captioner
Knowledge Retrieval
Solution Generator
Answer Generator
[194, 21]: Earthzy paper
[244, 70]: plates now
[172, 112]: carry the Sierra
[231, 159]: Club seal of
[275, 198]: approval
Analyze the advertisement to determine which persuasive appeal is used. The ad mentions that Earthzy paper plates carry the Sierra Club seal of approval. This implies that the product is environmentally friendly and has been endorsed by a reputable organization. This appeal is primarily based on the credibility and authority of the Sierra Club, which is an example of ethos (character).
Text Detector
Knowledge Retrieval
Solution Generator
Answer Generator
ethos (character)
- The question is about identifying the main persuasive appeal used in an advertisement.
- There are three main types of persuasive appeals: ethos, pathos, and logos.
- Ethos is an appeal to the speaker's or writer's credibility, character, or authority.
- Pathos is an appeal to the audience's emotions, feelings, or sympathies.
- Logos is an appeal to logic, reason, or rationality, using facts, statistics, or arguments.
Image Captioner
Query Generator
Bing Search
Solution Generator
Answer Generator


4x speed
106


diandian112205 bilibili

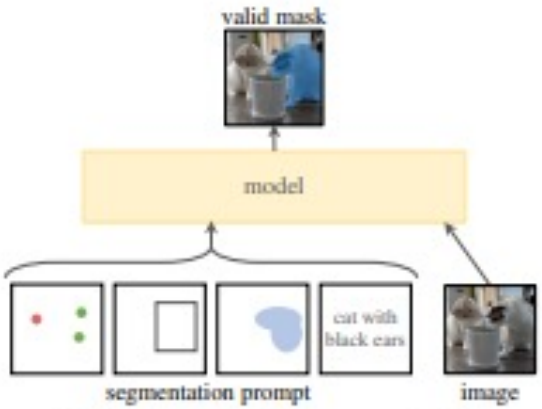
Visual Foundation Model

same basic modules

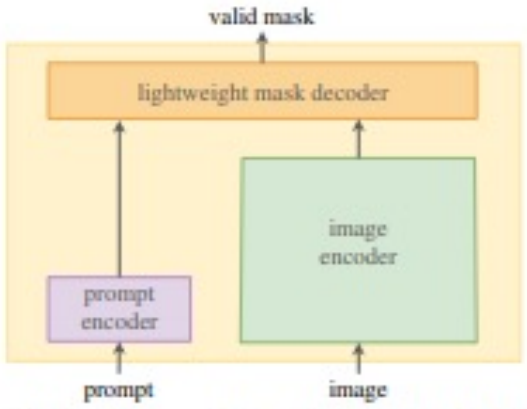
data

scale

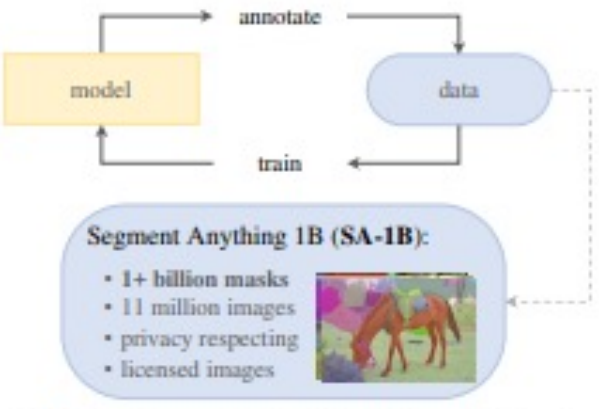
10s to 1000s GPU servers



(a) **Task**: promptable segmentation



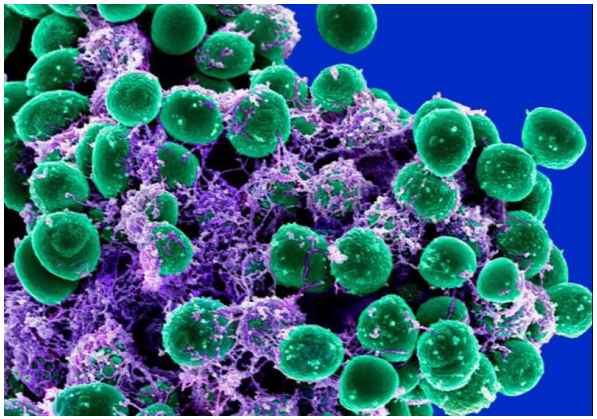
(b) **Model**: Segment Anything Model (**SAM**)



(c) **Data**: data engine (top) & dataset (bottom)

SAM was trained on 256 A100 GPUS for 68 hours.
11M images and 1.1B masks.

SAM



SAM

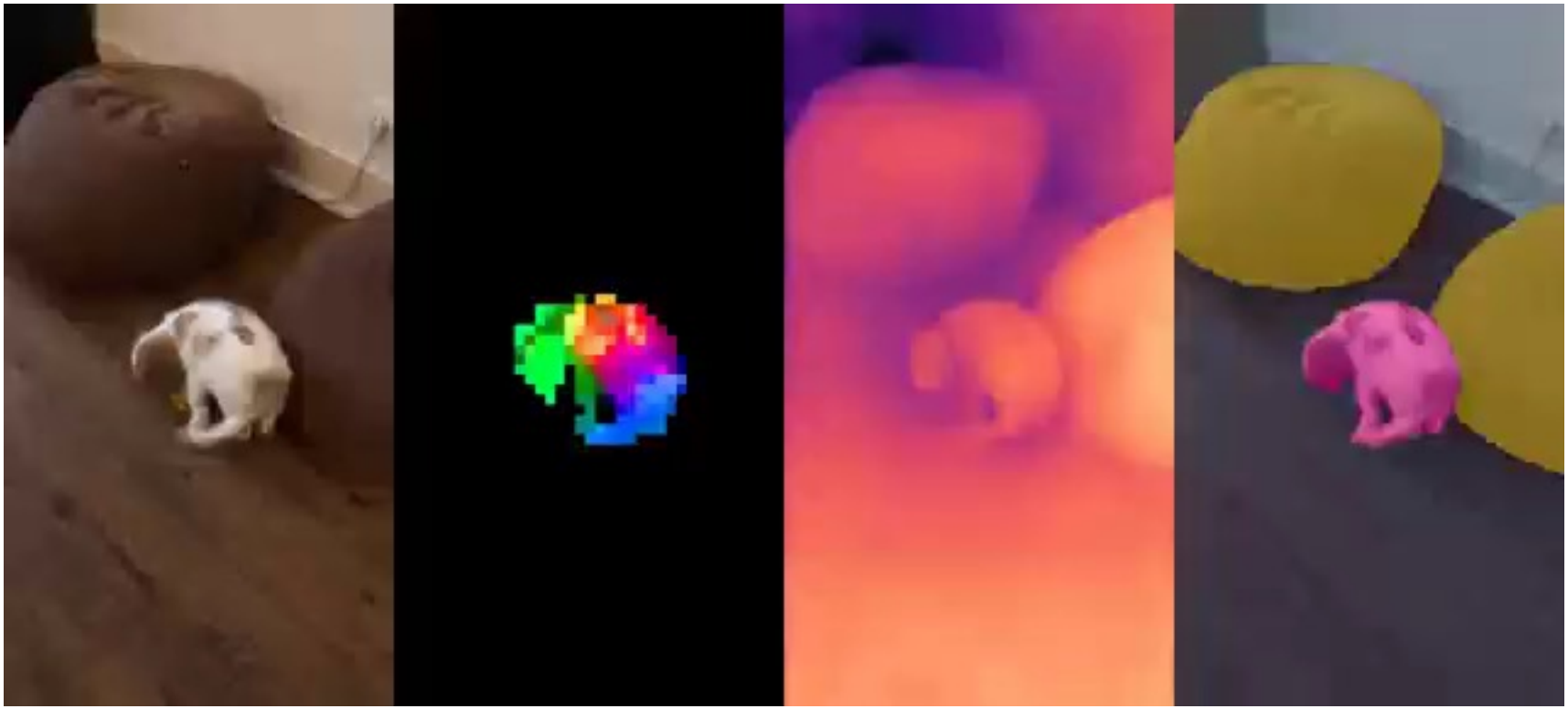
"Change the dog to a monkey"

Stable Diffusion

Whisper

"A paint of spring/summer /autumn/winter field."

Visual Foundation Model



Diffusion