Pattern Recognition and Image Processing

Computer Vision

What is vision

- Recognize objects
  - people we know
  - things we own

- Locate objects in space
  - to pick them up

- Track objects in motion
  - catching a baseball
  - avoiding collisions with cars on the road

- Recognize actions
  - walking, running, pushing

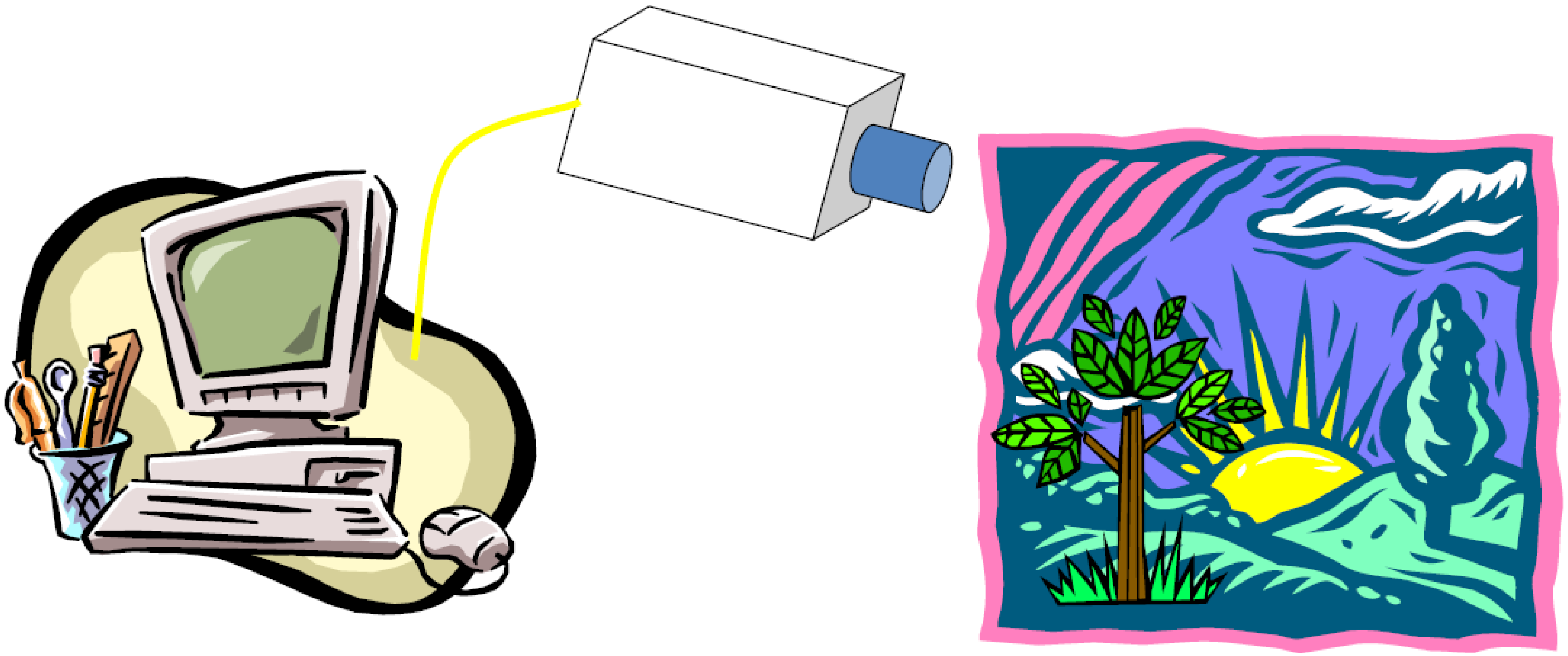
Vision is deceivingly easy

- We see effortlessly
  - seeing seems simpler than “thinking”
  - we can all “see” but only select gifted people can solve “hard” problems like chess
  - however, surprisingly, we use nearly 70%of our brains for visual perception!

- Vision is an exceptionally strong sensation
  - Vision is immediate
  - We perceive the visual world as external to ourselves, but it is a reconstruction within our brains
  - We regard how we see as reflecting the world “as it is;”

Computer vision

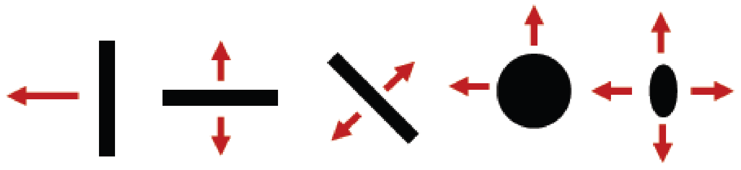
- Developing computer programs to understand the content of images and videos

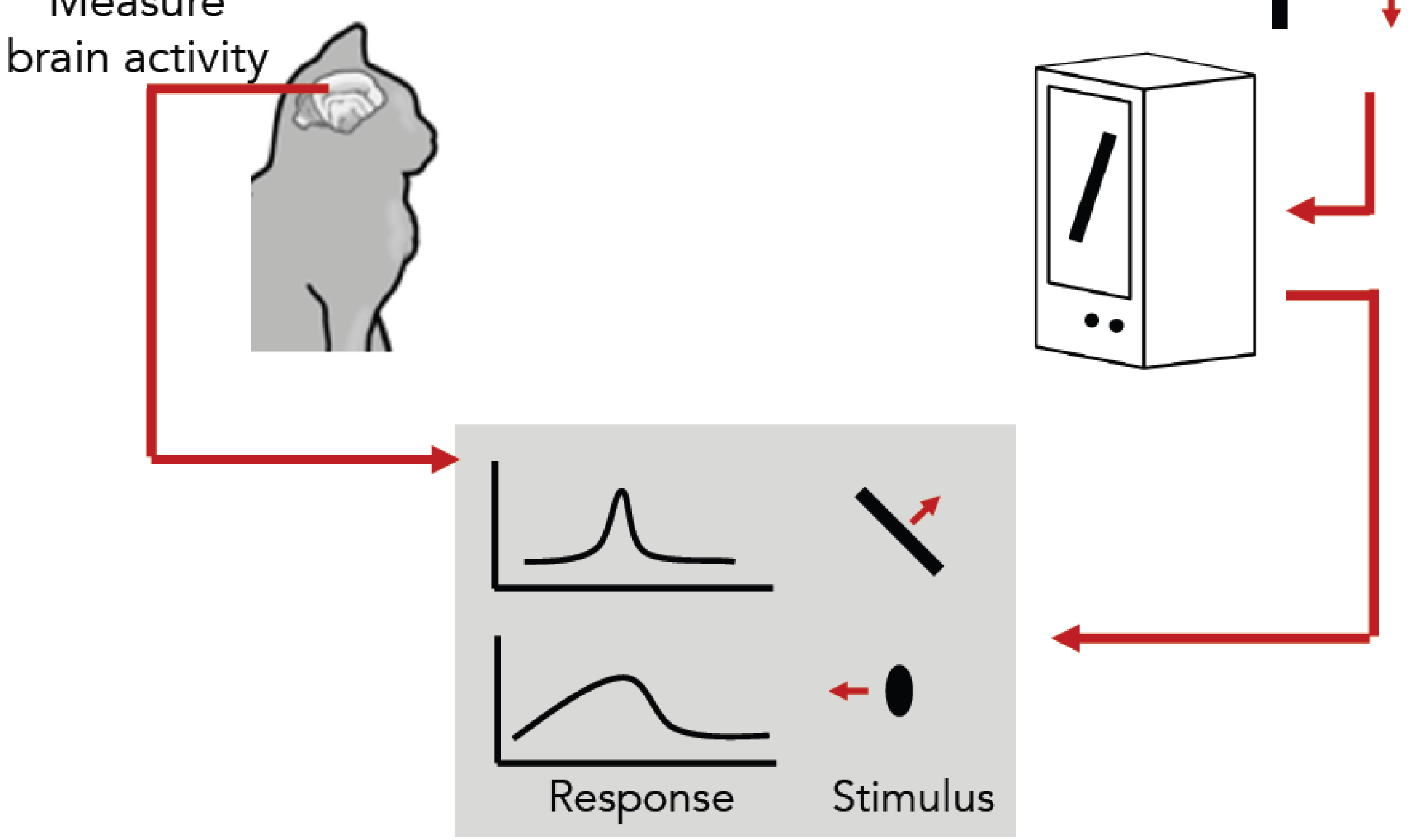


‘‘Computer vision is a field of study and research that focuses on enabling computers to understand, interpret, and analyze visual information from images or videos. It aims to replicate and extend human visual perception capabilities using computational algorithms and models.’’

Computer Vision: a brief history

Hubel and Wiesel, 1959



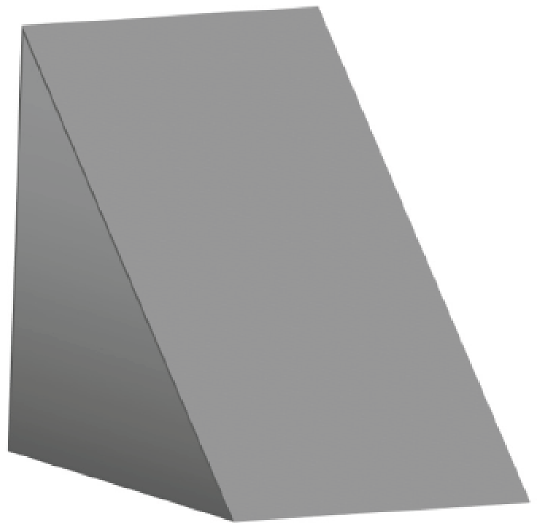


Simple cells:
Response to specific rotation and orientation

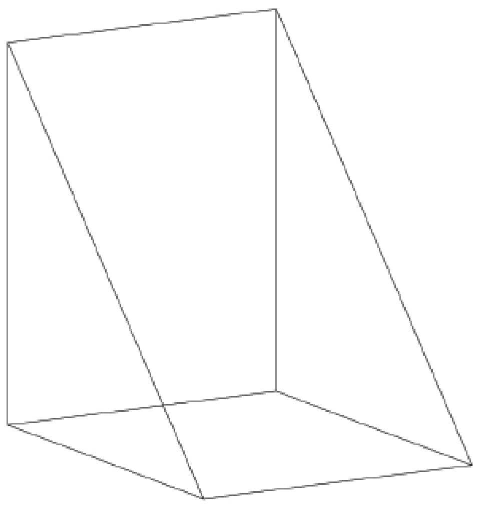
Complex cells:
Response to light orientation and movement, some translation invariance

Computer Vision: a brief history

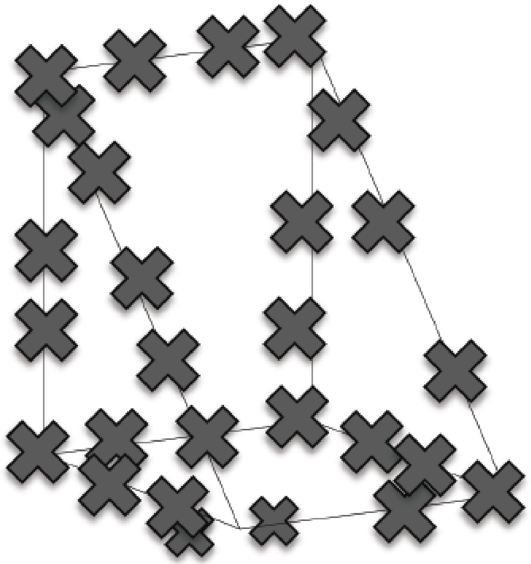
Larry Roberts, 1963



(a) Original picture



(b) Differentiated picture



(c) Feature points selected

Computer Vision: a brief history

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group July 7, 1966

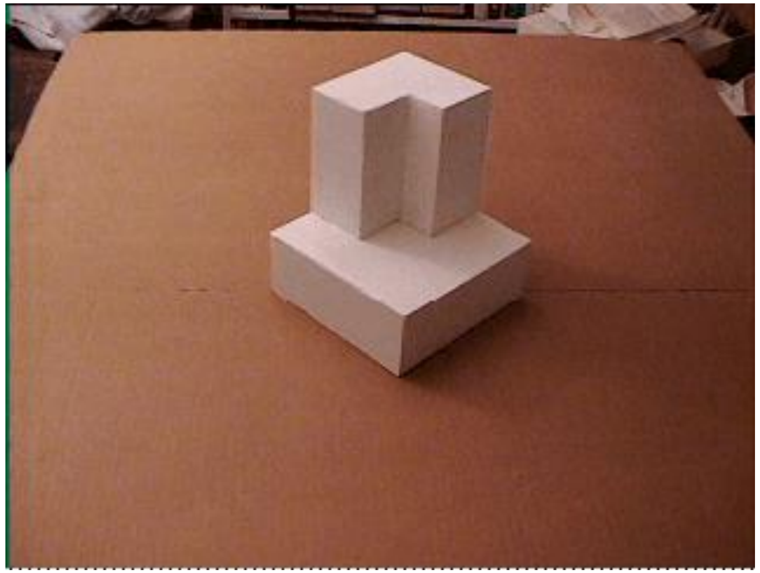
Vision Memo. No. 100.

THE SUMMER VISION PROJECT

Seymour Papert

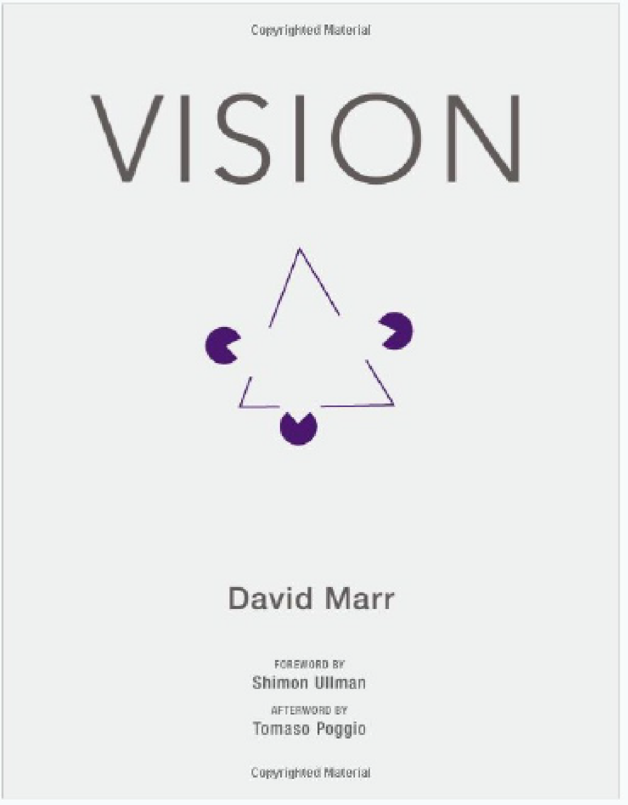
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the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

Computer Vision: a brief history

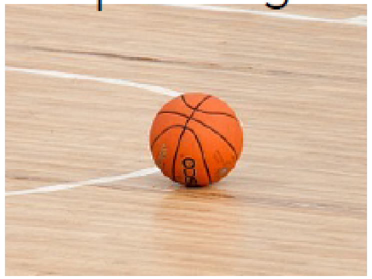


- The M.I.T. summer vision program
  - summer of 1966
  - point TV camera at stack of blocks
  - locate individual blocks
    - recognize them from small database of blocks
  - describe physical structure of the scene
    - support relationships
- Formally ended in 1985
- “ The first great revelation was that the problems are difficult. Of course, these days this fact is a commonplace. But in the 1960s almost no one realized that machine vision was difficult.”

Computer Vision: a brief history

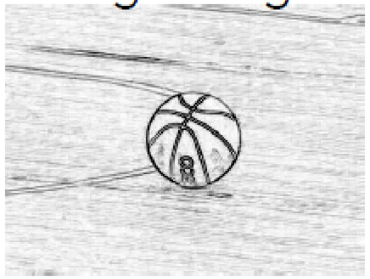


Input image

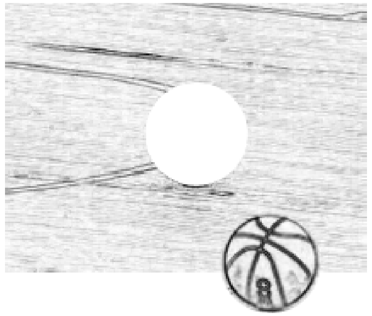


This image is CC0 1.0 public domain

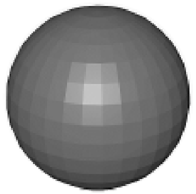
Edge image



2 ½-D sketch



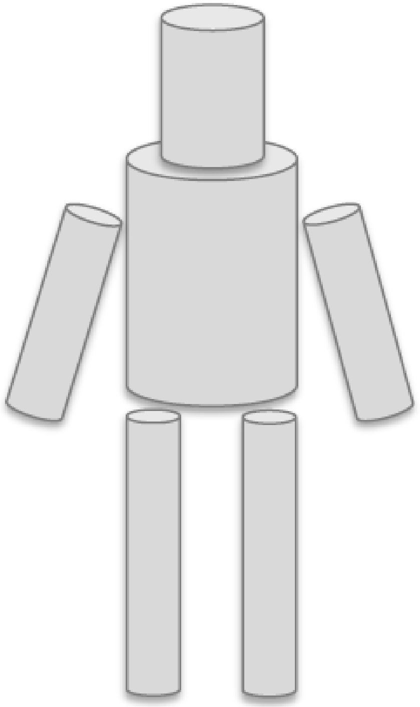
3-D model



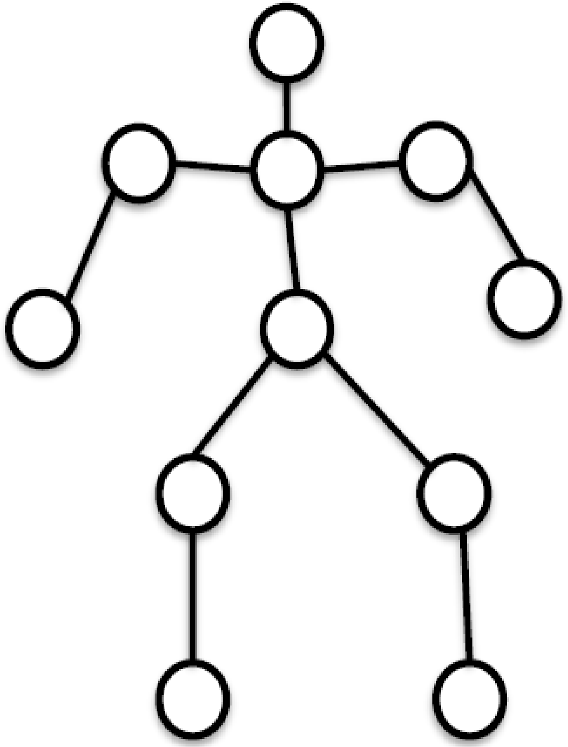
This image is CC0 1.0 public domain

Input Image	Primal Sketch	2 ½-D Sketch	3-D Model Representation
Perceived intensities	Zero crossings, blobs, edges, bars, ends, virtual lines, groups, curves boundaries	Local surface orientation and discontinuities in depth and in surface orientation	3-D models hierarchically organized in terms of surface and volumetric primitives

Computer Vision: a brief history



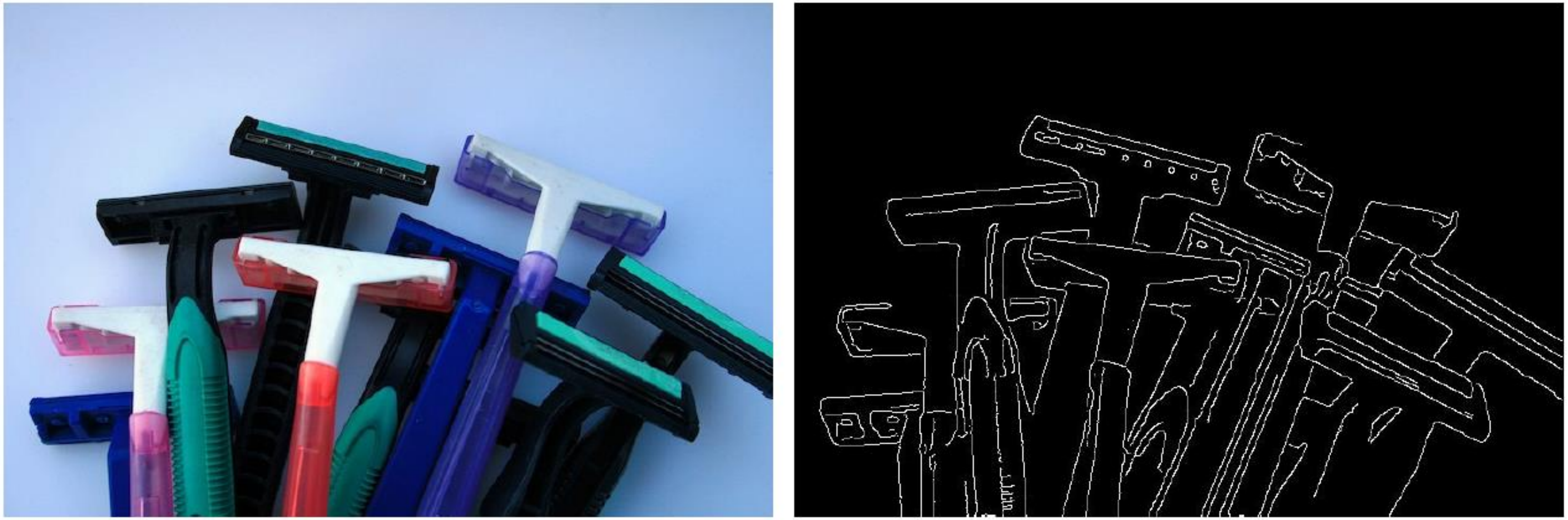
Generalized Cylinders,
Brooks and Binford,
1979



Pictorial Structures,
Fischler and Elshlager, 1973

Computer Vision: a brief history

Recognition via Edge Detection (1980s)



John Canny, 1986
David Lowe, 1987

Computer Vision: a brief history

Computer Vision: a brief history

Image is public domain

Image is public domain

SIFT, David Lowe, 1999

Computer Vision: a brief history

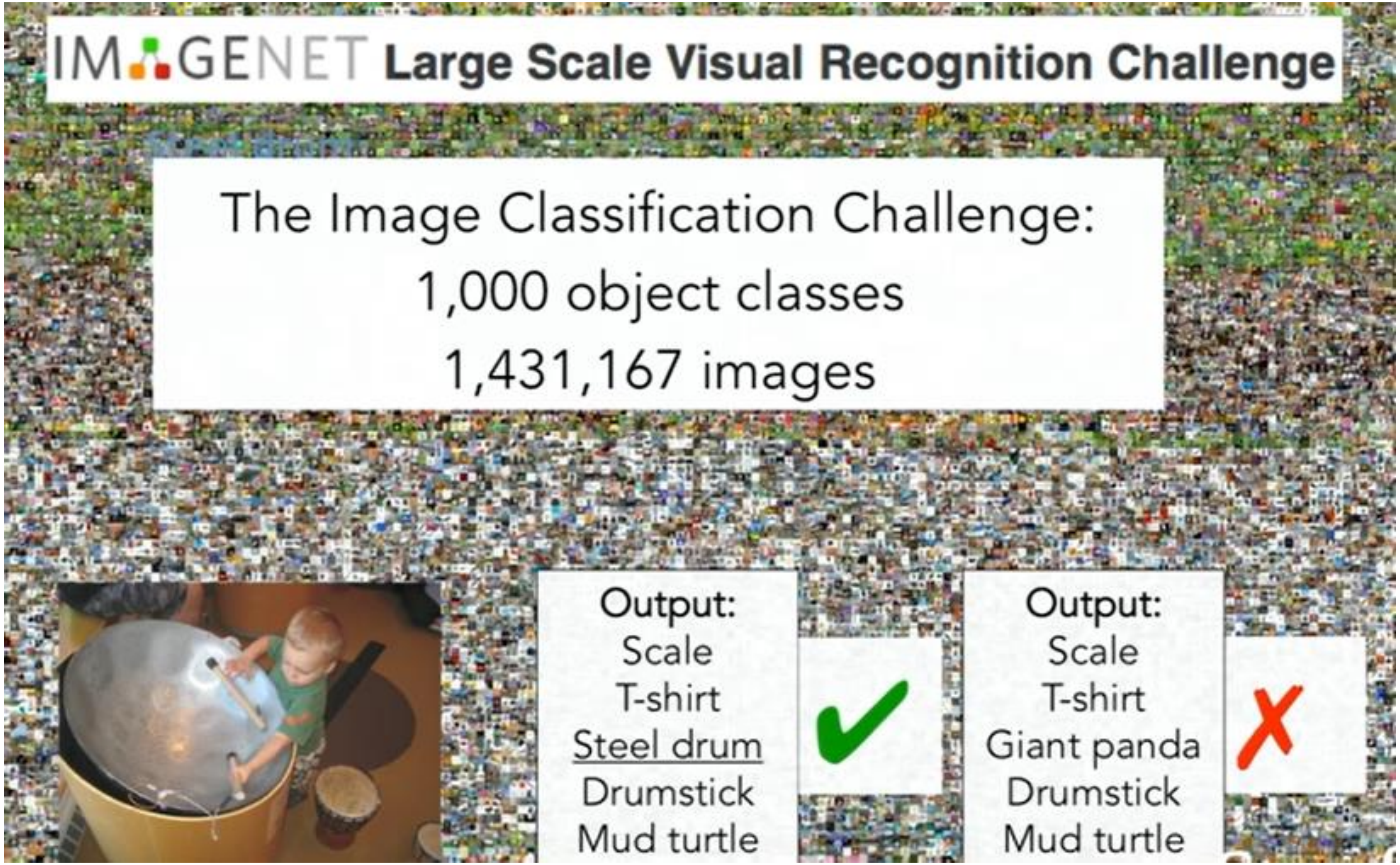
Face Detection

Viola and Jones, 2001

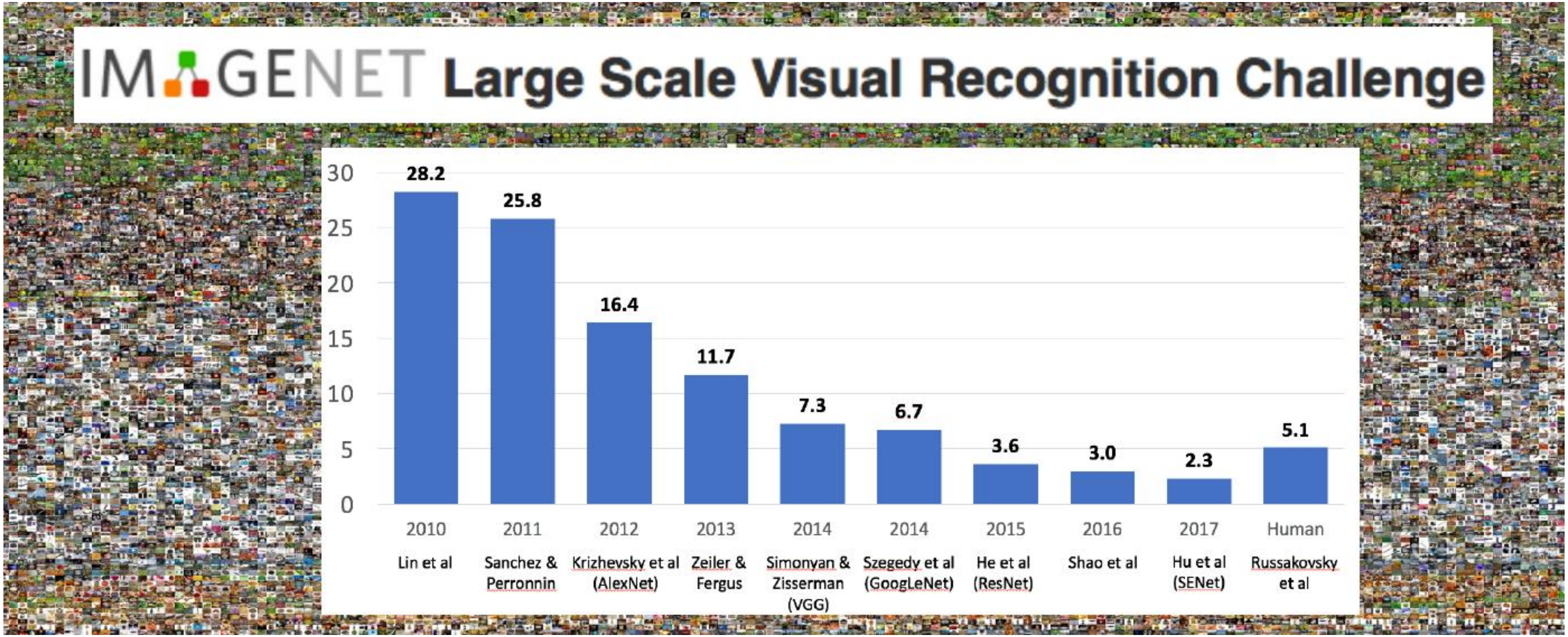
One of the first successful applications of machine learning to vision

https://www.ted.com/talks/fei_fei_li_how_we_re_teaching_computers_to_understand_pictures?language=en

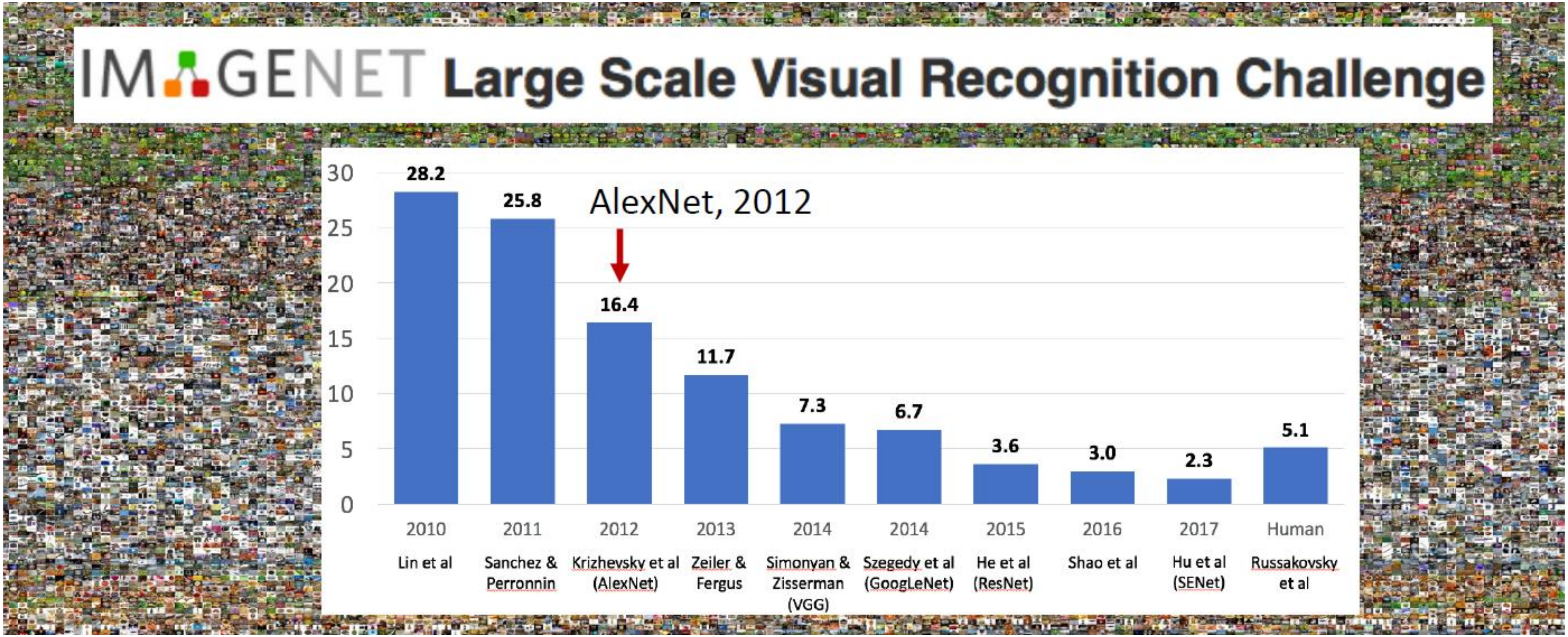
ImageNet challenge



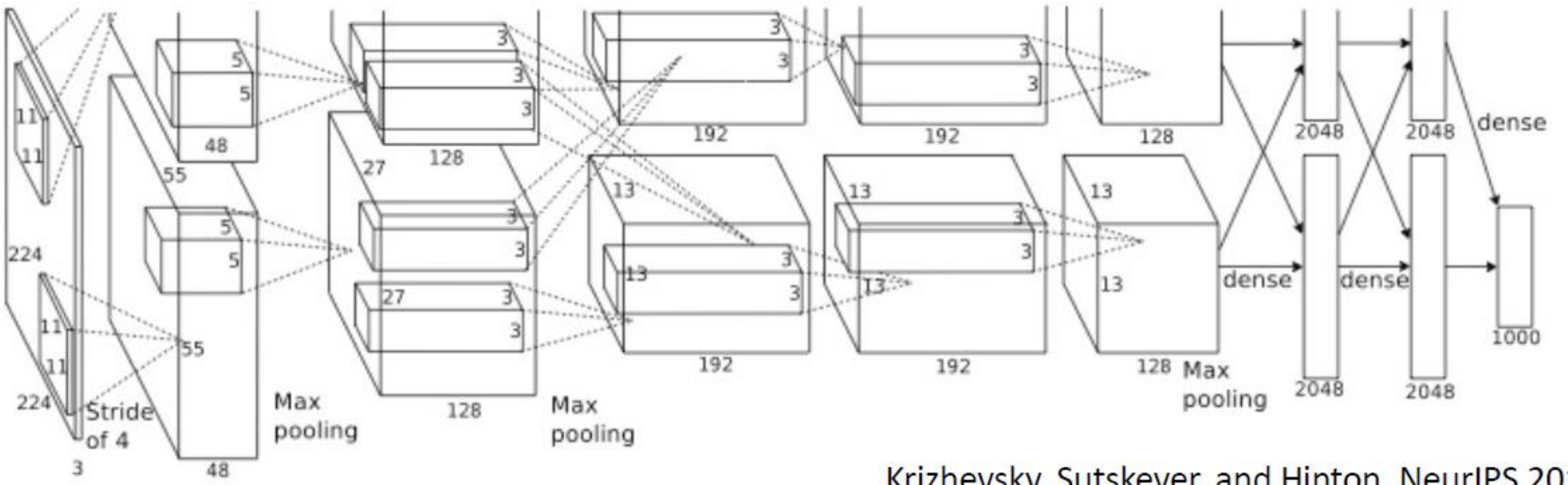
ImageNet challenge



ImageNet challenge



AlexNet: Deep Learning Goes Mainstream

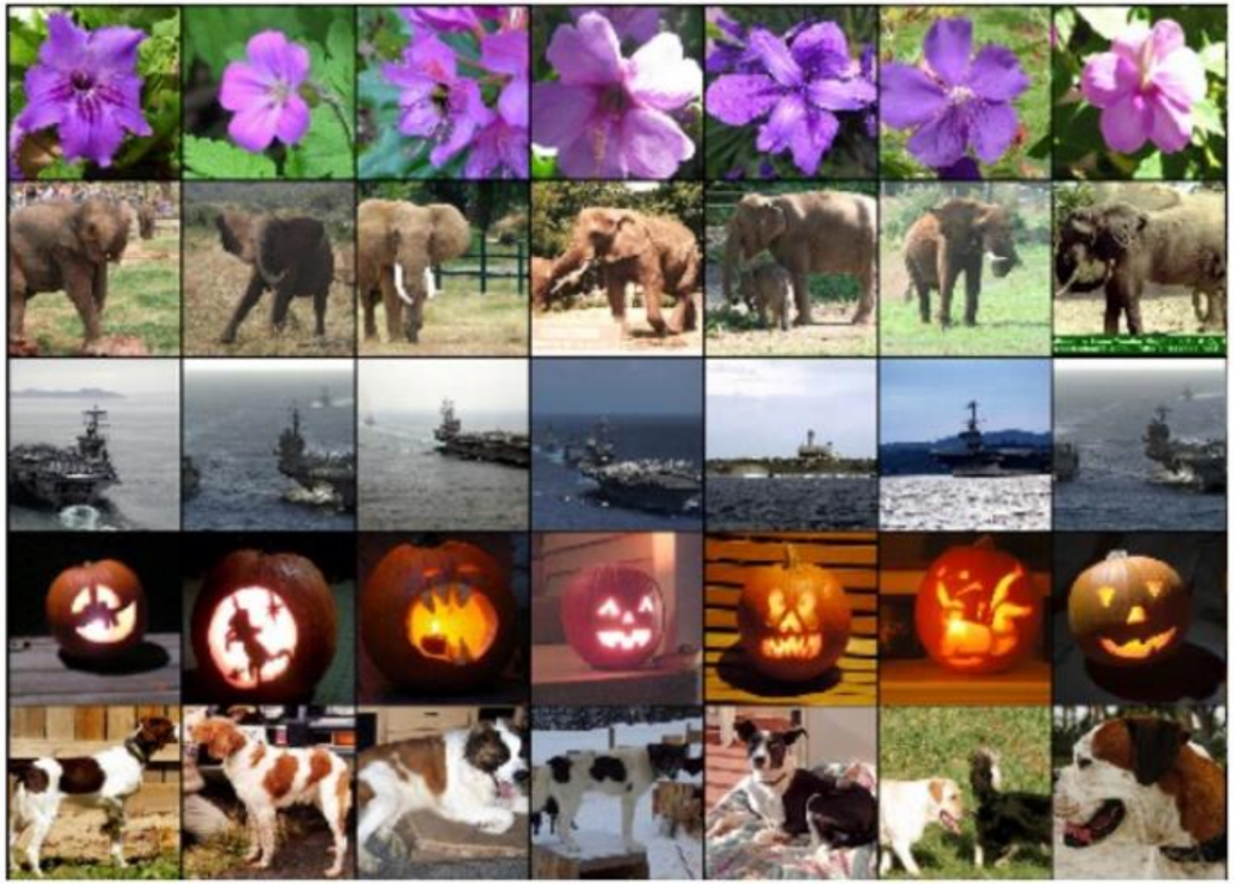


Krizhevsky, Sutskever, and Hinton, NeurIPS 2012

2012 to Present: Deep Learning is Everywhere

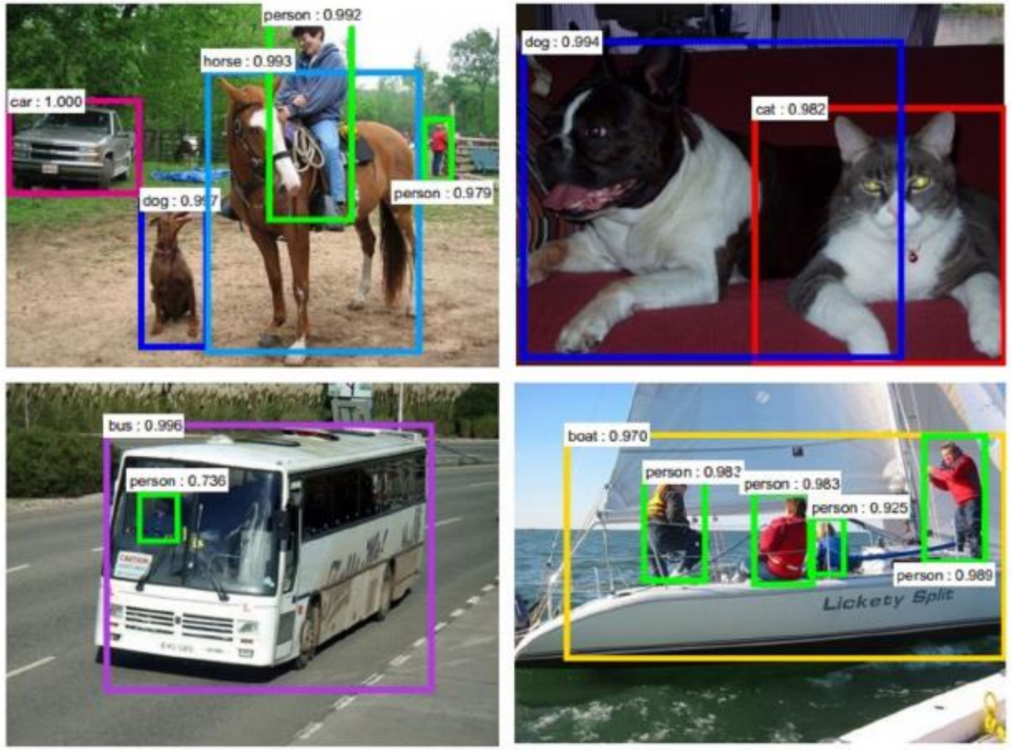
Image Classification

Image Retrieval



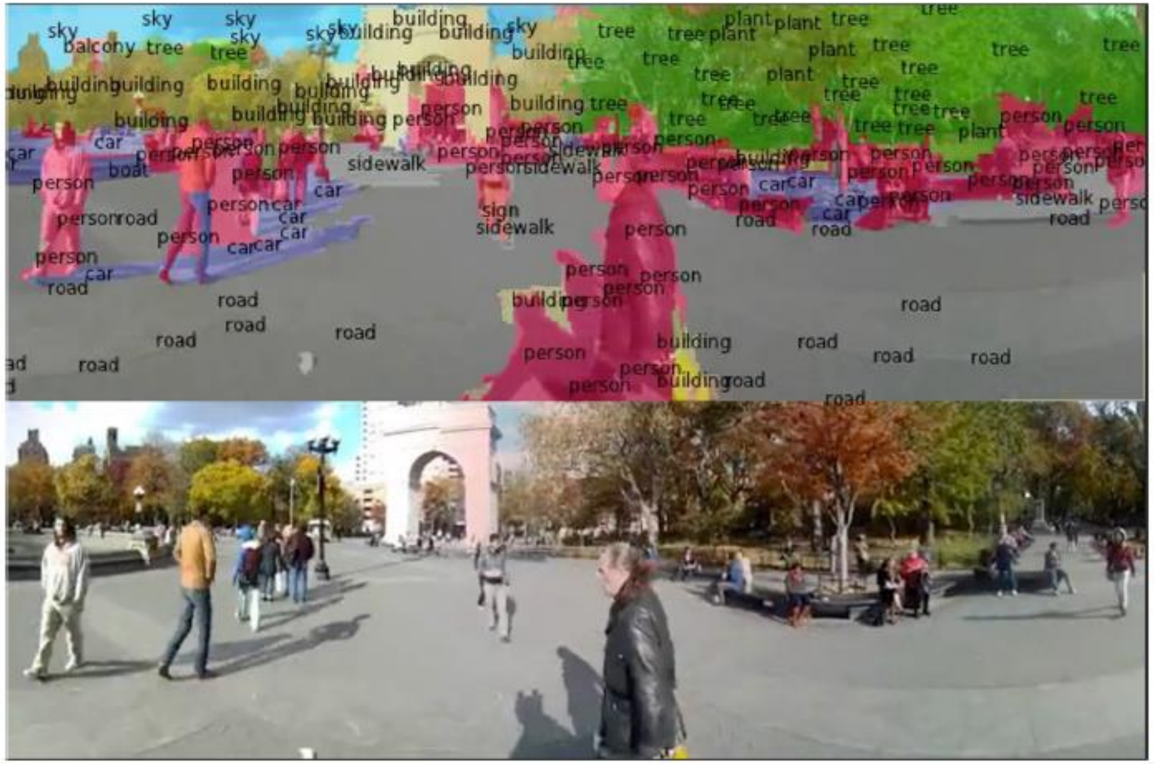
2012 to Present: Deep Learning is Everywhere

Object Detection



Ren, He, Girshick, and Sun, 2015

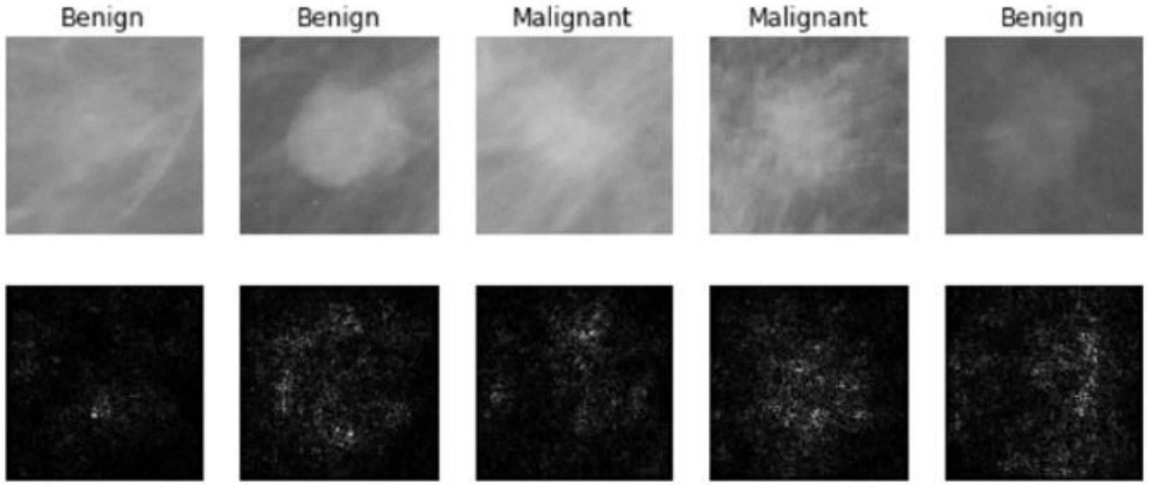
Image Segmentation



Fabaret et al, 2012

2012 to Present: Deep Learning is Everywhere

Medical Imaging



Levy et al, 2016 Figure reproduced with permission

Galaxy Classification

Dieleman et al, 2014

From left to right: public domain by NASA, usage permitted by ESA/Hubble, public domain by NASA, and public domain.

Whale recognition

Kaggle Challenge

This image by Christin Khan is in the public domain and originally came from the U.S. NOAA.

2012 to Present: Deep Learning is Everywhere

Image Captioning
Vinyals et al, 2015
Karpathy and Fei-Fei, 2015

A white teddy bear sitting in the grass

A man in a baseball uniform throwing a ball

A woman is holding a cat in her hand

A man riding a wave on top of a surfboard

A cat sitting on a suitcase on the floor

A woman standing on a beach holding a surfboard

All images are CC0 Public domain:
https://pixabay.com/en/luggage-antique-cat-1643010/
https://pixabay.com/en/teddy-plush-bears-cute-teddy-bear-1623436/
https://pixabay.com/en/surf-wave-summer-sport-litoral-1668716/
https://pixabay.com/en/woman-female-model-portrait-adult-983967/
https://pixabay.com/en/handstand-lake-meditation-496008/
https://pixabay.com/en/baseball-player-shortstop-infield-1045263/

2012 to Present: Deep Learning is Everywhere

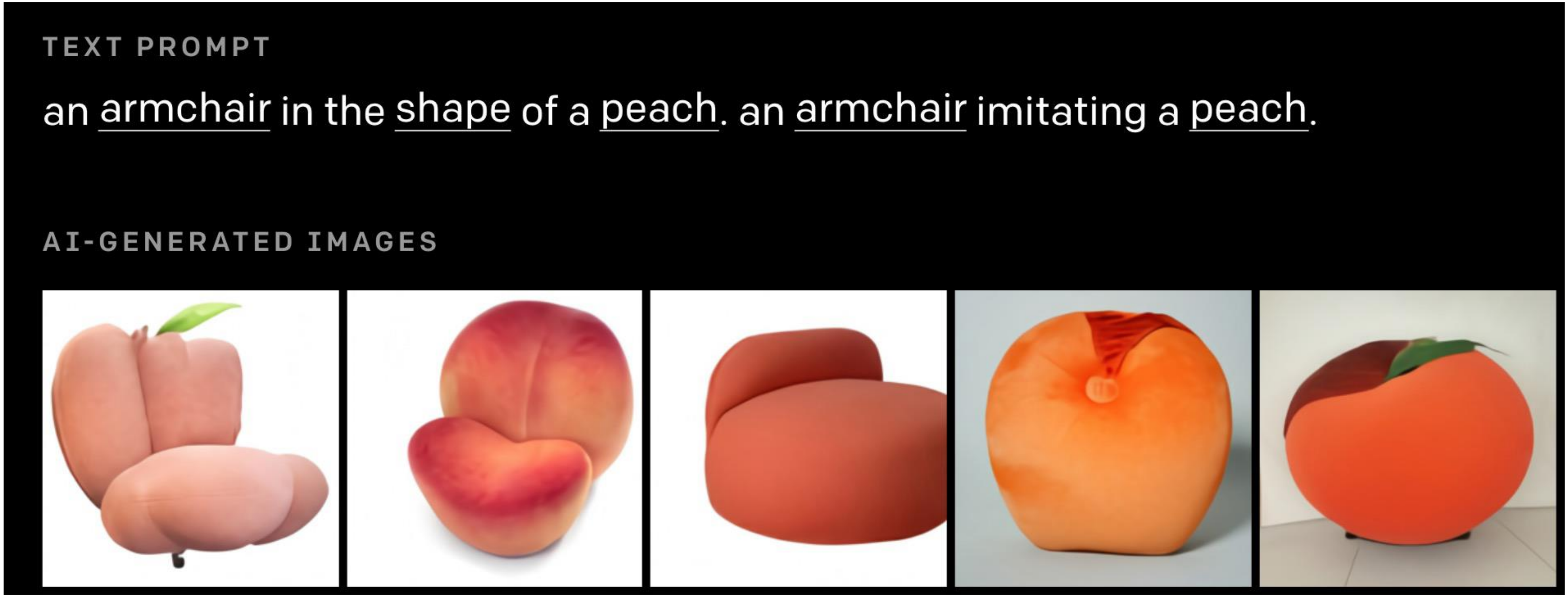
Original image is CC0 public domain
Starry Night and Tree Roots by Van Gogh are in the public domain
Bokeh image is in the public domain
Stylized images copyright Justin Johnson, 2017;
reproduced with permission

Mordvinsev et al, 2015

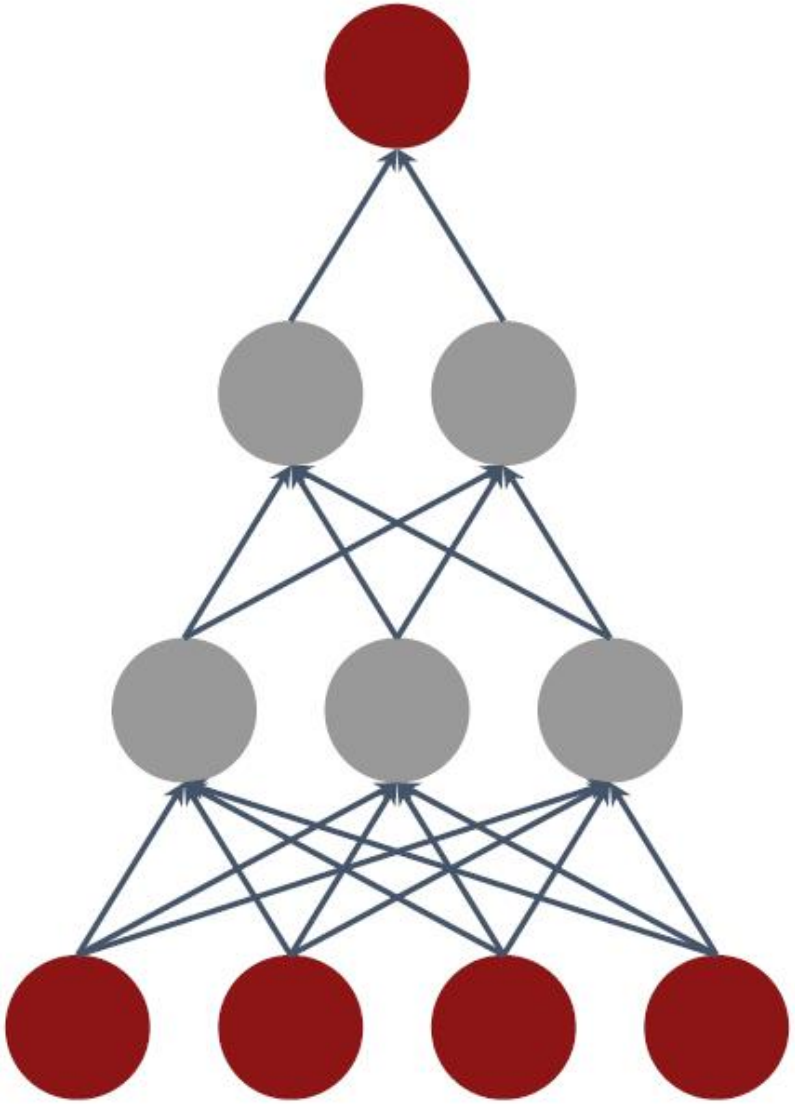
Gatys et al, 2016

2012 to Present: Deep Learning is Everywhere

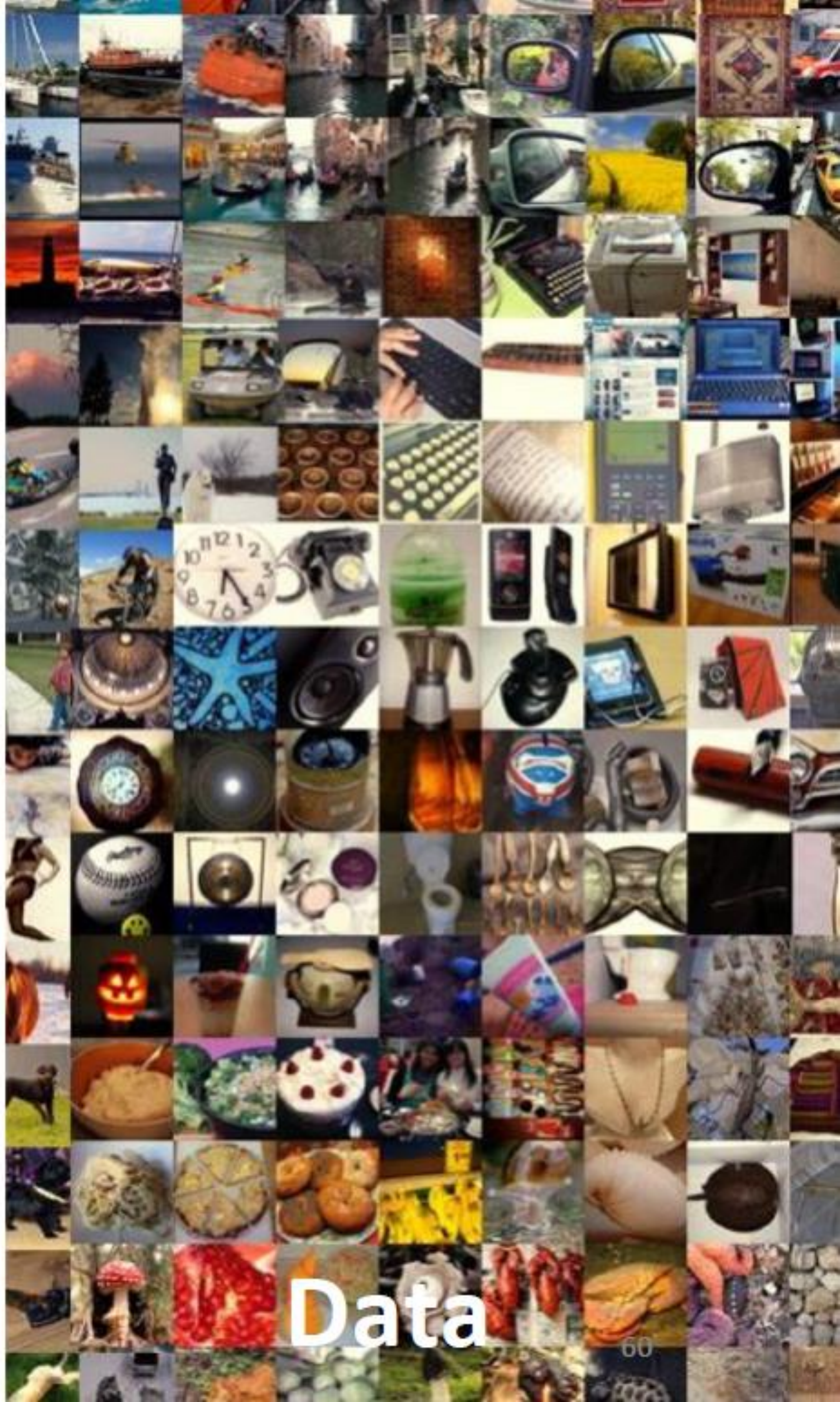
2012 to Present: Deep Learning is Everywhere



Computation



Algorithms



Data

Common computer vision tasks

- Image Classification: Assigning a label or category to an image from a predefined set of classes. For example, classifying images as "cat" or "dog.“
- Object Detection: Identifying and localizing multiple objects within an image. This involves drawing bounding boxes around objects and assigning class labels to them.
- Semantic Segmentation: Assigning a category label to each pixel in an image, enabling a more detailed understanding of the scene. It involves segmenting the image into meaningful regions.

Common computer vision tasks

- Instance Segmentation: Similar to semantic segmentation, but with the additional step of distinguishing between individual instances of objects. This task involves segmenting and labeling each object instance separately.

- Object Tracking: Tracking the movement of objects across frames in a video or a sequence of images. It involves locating and following objects across different frames.

- Pose Estimation: Estimating the position and orientation of objects or human body parts within an image or video. This task is often used in applications like augmented reality and motion tracking.

- Image Captioning: Generating textual descriptions or captions that describe the content of an image. It combines computer vision with natural language processing.

Common computer vision tasks

- Image Super-Resolution: Enhancing the resolution or quality of an image. This task involves generating a higher-resolution version of a low-resolution image.
- Face Recognition: Identifying and verifying individuals based on their facial features. It involves detecting and matching faces in images or videos.
- Image Retrieval: Searching for visually similar images in a large database based on a query image. It involves finding images that share similar visual characteristics.
- Scene Understanding: Inferring high-level scene semantics and context from visual information. This task involves understanding the overall scene, including objects, relationships, and activities.

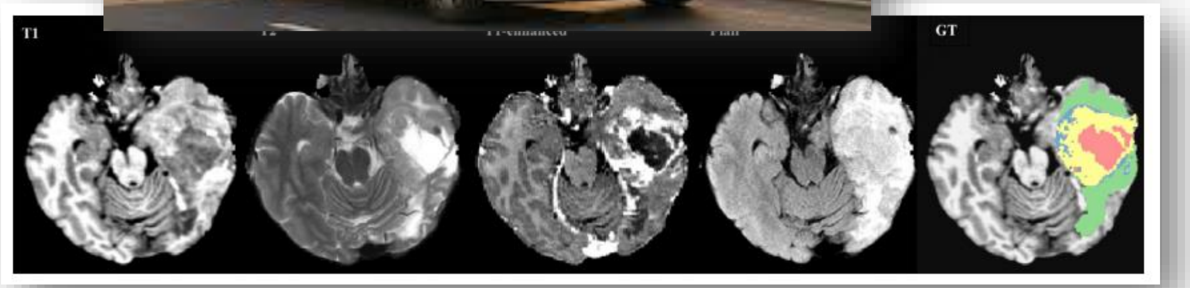
Why study Computer Vision?

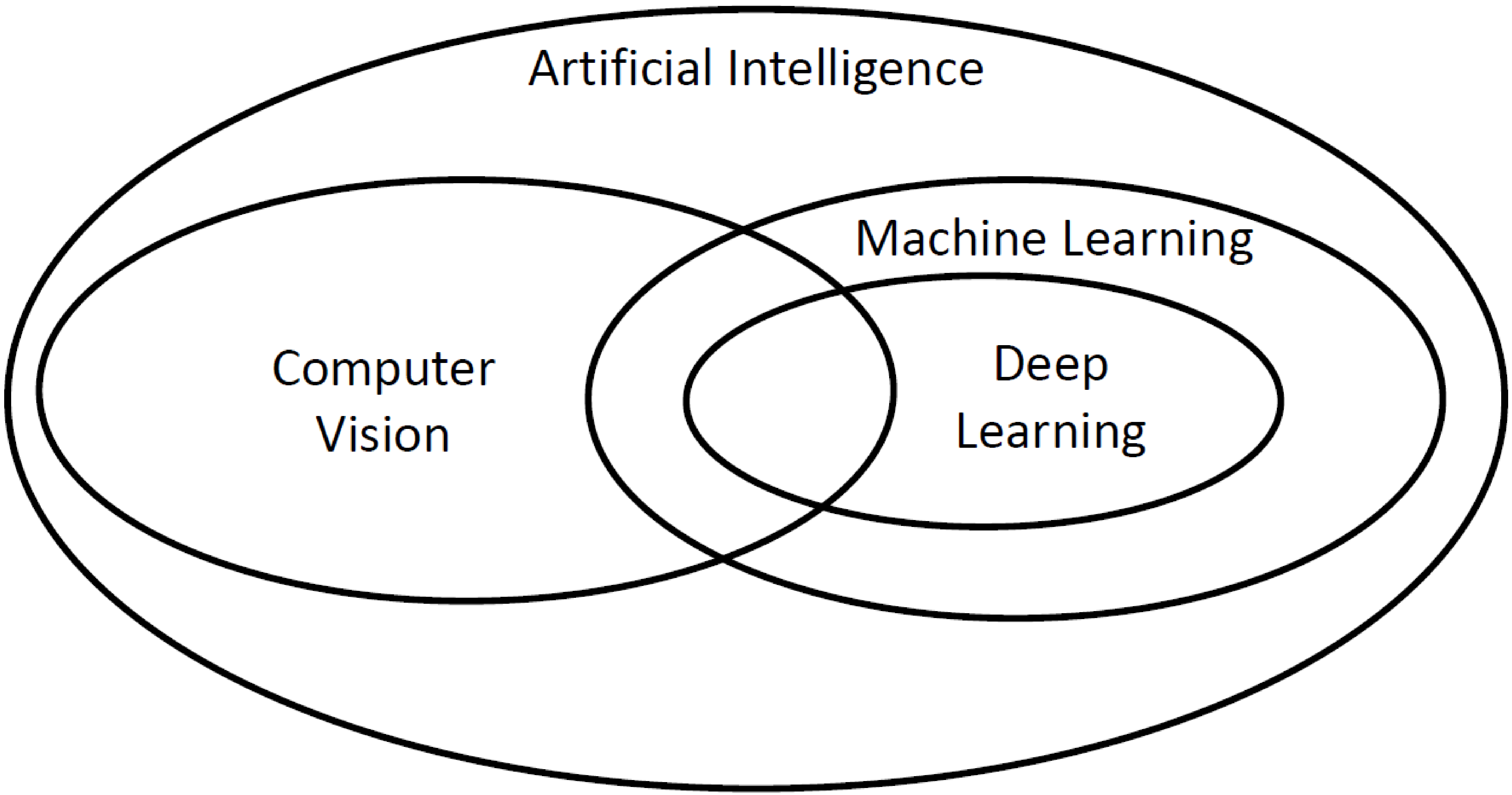
- Images and videos are everywhere
- Fast-growing collection of useful applications
  - building representations of the 3D world from pictures
  - automated surveillance (who’s doing what)
  - movie post-processing
  - face finding
- Various deep and attractive scientific mysteries
- Greater understanding of human vision

Computer vision applications

1. Autonomous Vehicles
2. Surveillance and Security
3. Medical Imaging
4. Augmented Reality (AR)
5. Object Recognition and Classification
6. Robotics
7. Retail and E-commerce
8. Industrial Automation
9. Sports Analytics
10. Human-Computer Interaction

…




Artificial Intelligence
Machine Learning
Computer Vision
Deep Learning

Some Resources

http://cs231n.stanford.edu/

https://cs231n.github.io/python-numpy-tutorial/