Introduction to Computer Vision

Bing Gong
（巩冰）
gongbing@shnu.edu.cn

Introduction to the lecturer

巩冰博士

gongbing@shnu.edu.cn

- 特拉维夫大学（以色列）访问学者　　2022.10-2022.11
  - 计算机学院
- 马德里理工大学（西班牙） 博士　　2012.09-2017.07
  - QS 工程技术专业排名58（2022）
  - 马德里自治大学/康普顿斯大学联合培养
- 密西根州立大学（美国)　　访问学者　　2015.03-2015.08
  - 商学院供应链管理-风险管理（全美专业第一）
- 中国地质大学（中国）　　硕士　　2009.09-2012.06
- 中国地质大学（中国）　　学士　　2005.09-2009.07

➢ **工作经历**

- 于利希超算中心　　博士后研究员　　2019.01 – 2023.09
- 美国康宁公司（世界五百强）　数据科学家　　2017.09-2019.01

Course Information

Time and location:

- Week 1-8 Tuesday, 奉贤科技楼A座312
- Week 1- 16 Thursday, 1-16,奉贤5教楼D207
- 7 Practical courses

The office hours and location:

- Tuesday 15:00 - 17:00 and Thursday 10:00-12:00 (after the class) , 奉贤科技楼A座701

Teaching materials website:

- https://github.com/masak1112/Computer_vision_lecture_shnu

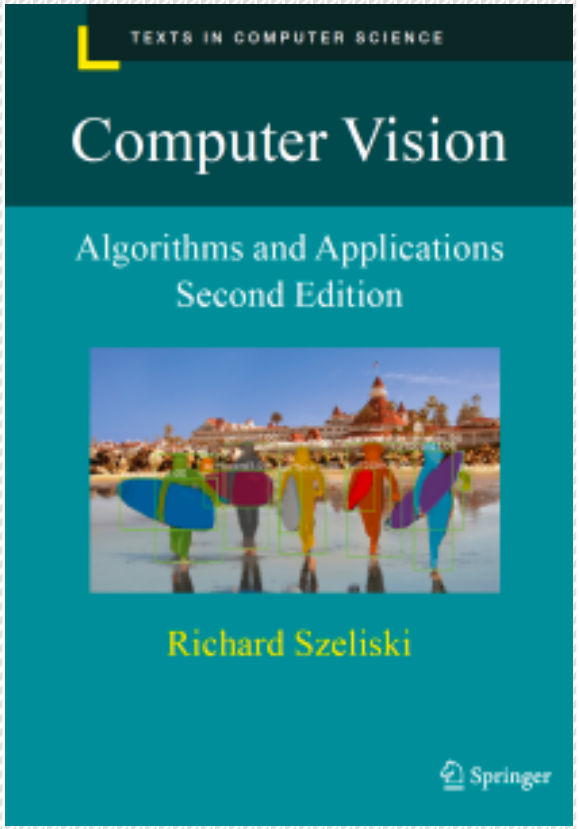
Schedule and topics

- Two lectures + one practical course
- Topics:
  - Lecture 1: Introduction to computer vision
  - Lecture 2 – 9 incl. Practical 3 courses :
    - CV foundations (color, texture, filters, convolution…)
  - Lecture 10 – 14 incl. Practical courses 3-4:
    - Advanced CV (CNN with PyTorch introduction, Generative modeling)
  - Last course (2024-06-13):
    - Final project exam (implement one project, and give presentations)

Grading

- Your final grade:
  - Participation (30%)
  - Assignments (30%):
  - Final Project (40%):
  - NO EXAM!

Text Book



https://szeliski.org/Book/

http://www.computervisionmodels.com/

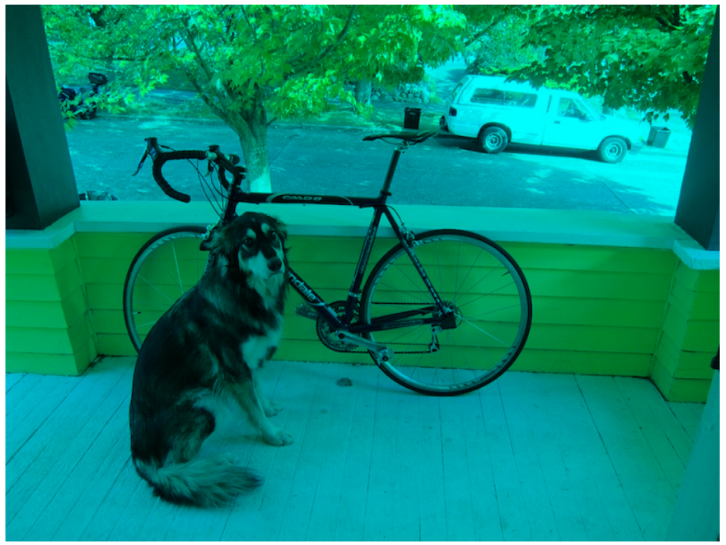
Other references

- Stanford University:
  - http://vision.stanford.edu/teaching/cs131Sfall2122/syllabus.html
- University of Washington:
  - https://courses.cs.washington.edu/courses/cse576/23sp/

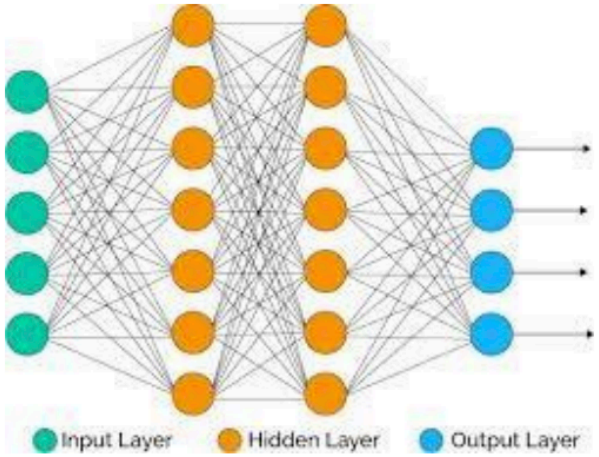
Prerequisites

- A good working knowledge of Python (Numpy) and Jupyter Notebook
- Not required, but good to have knowledge of PyTorch, or TensorFlow
- Not required, but good to have knowledge of Git

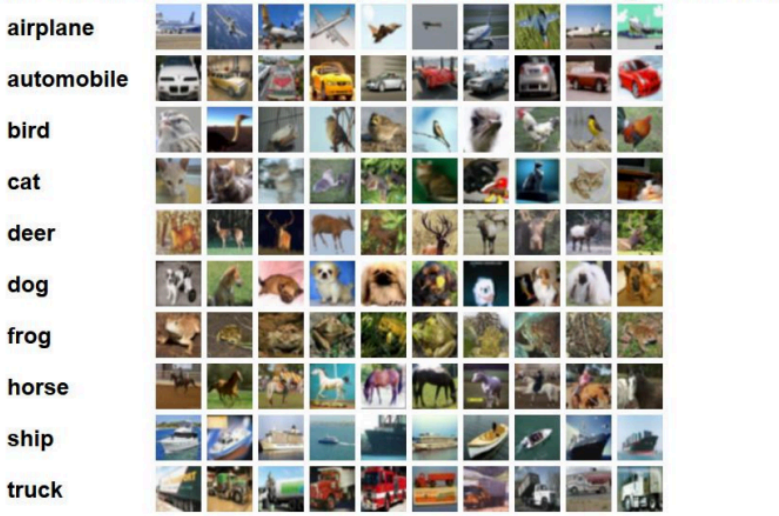
Practice 1&2&3: basic image processing



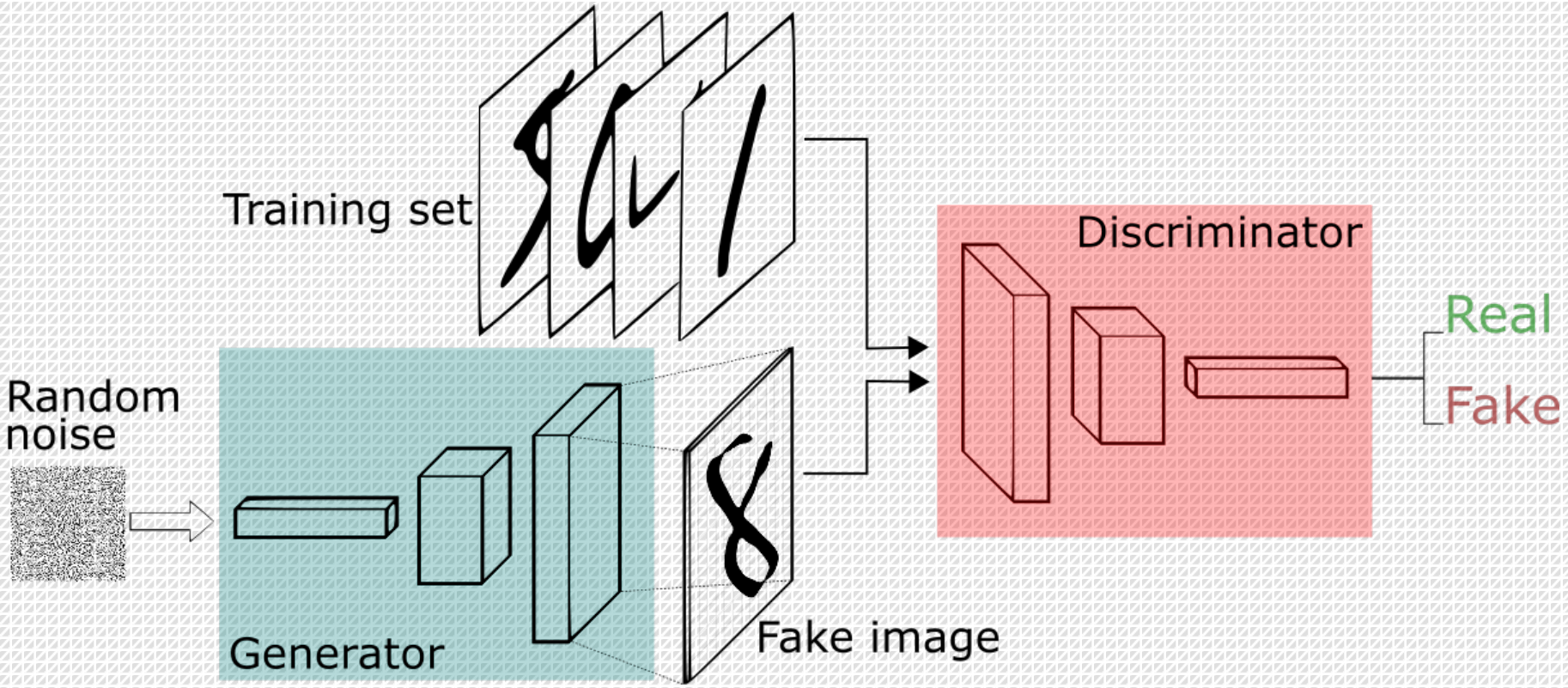
Practice 4&5 – ML pipeline by PyTorch



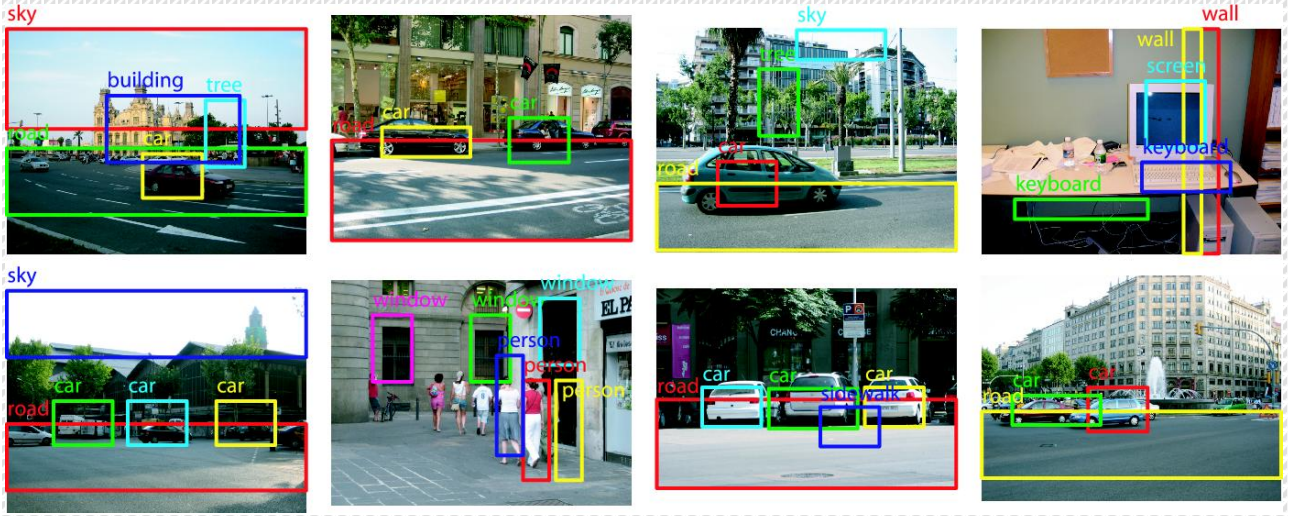
Here are the classes in the dataset, as well as 10 random images from each:

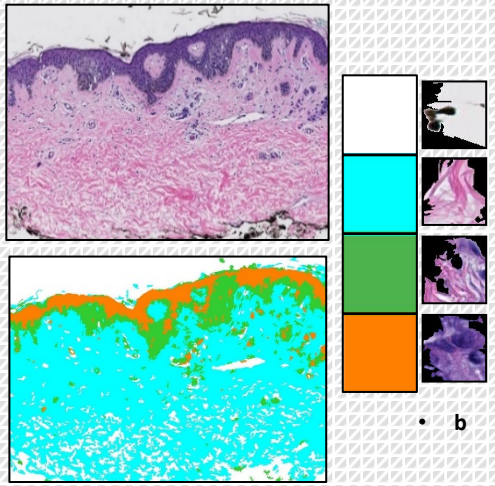


Practice 6&7– Generative models



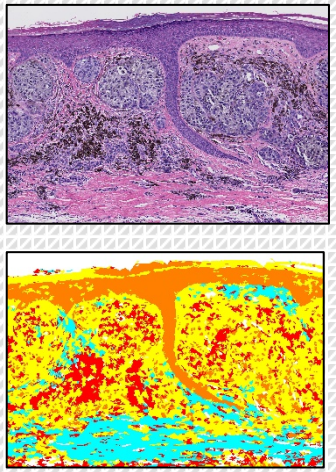
Final project – Machine Learning Application



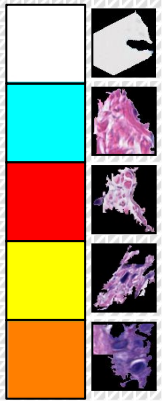


• a

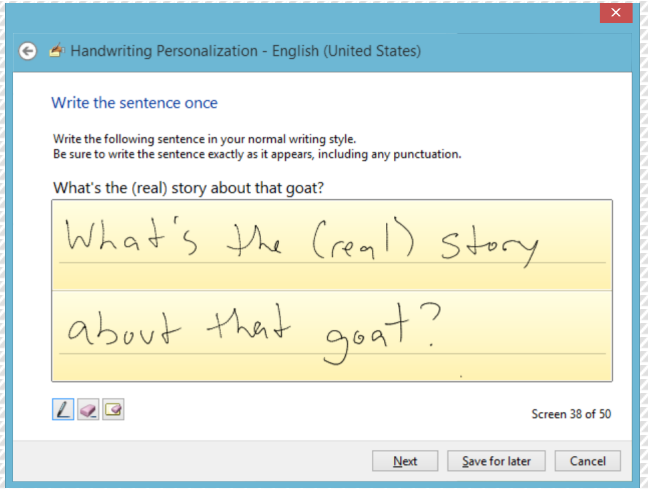
• b



• c



• d



Objectives of Lecture 1

- Get familiar with and understand the basic concepts of Image processing and Computer Vision (CV).
- Learn the history of CV development.
- Know the main applications in CV.
- Top scientist’s talk in CV community.
- Homework.

Questions?

- What is image (video)?
- What is image processing? What are the main tasks for image processing?
- What is (computer) vision? What are the main tasks for computer vision?
- What is the difference between image processing and computer vision?

The definition of image (video)

- What is an ”image” ?

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

An image is a visual representation. An image can be two-dimensional, such as a drawing, painting, or photograph, or three-dimensional, such as a carving or sculpture. Images may be displayed through other media, including projection on a surface, activation of electronic signals, or digital displays; they can also be reproduced through mechanical means, such as photography, printmaking or photocopying. Images can also be animated through digital or physical processes.

The digital image is sampled and mapped as a grid of dots or pictures elements. Each pixel is assigned a tonal value(black, white, shades of gray or color) which is represented by binary code. the binary digits for each pixel are stored in a sequence by a computer. To reduce the storage requirement, the image is often stored in a compressed file format.

Source: https://en.wikipedia.org/wiki/Image

The definition of image (video)

- What are the main attributes for an image?

Image Size: The dimension of the pixel is the horizontal and vertical measurements of an image that is expressed in pixels**.**

Bit Depth: The number of possible shades or tonal degrees that a color can have from black to white.

- Grayscale images are typically 8-bit i.e. 256 values.
- Color images are typically 24 bit colors i.e. 3 colors, 8 bit per color and 16 million values.

Channels:

- Grayscale images have one channel.
- Color images have three channels (e.g. RGB)

File Size: The file size of the image is the digital size of an image file in kilobytes (Kb), megabytes (Mb), or gigabytes (Gb). It is proportional to the pixel dimension of the image. Images with more pixels require more disk space to store. **File Size = (pixel dimensions x bit depth)/8**

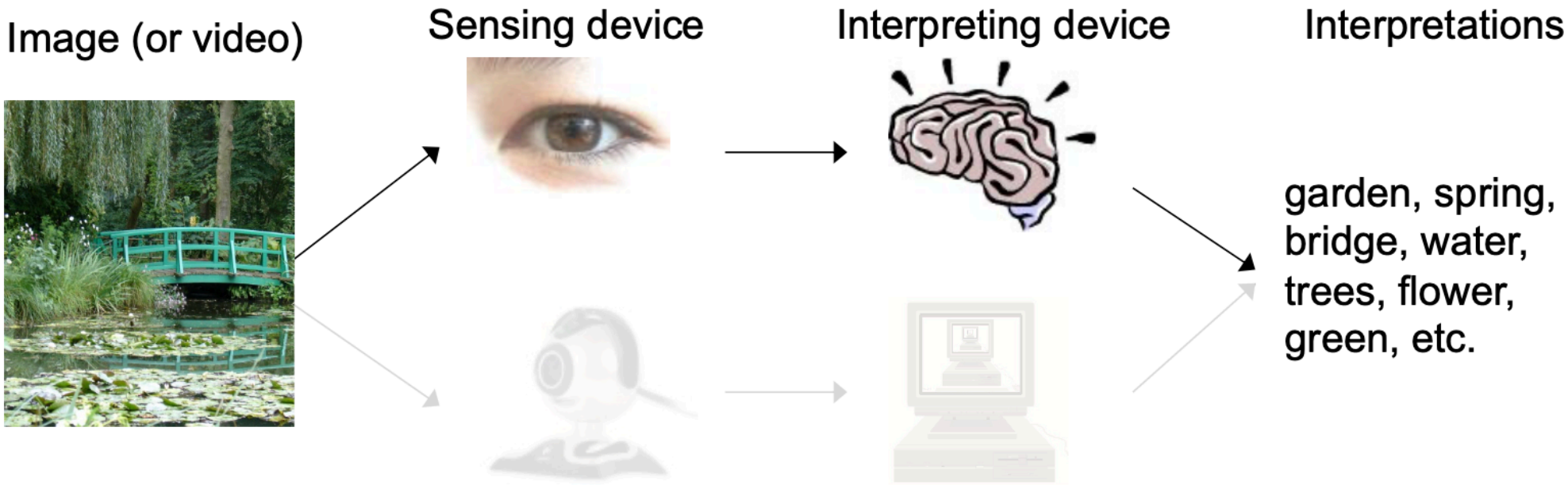
Resolution : It refers to the number of pixels in an image.

What is Image processing

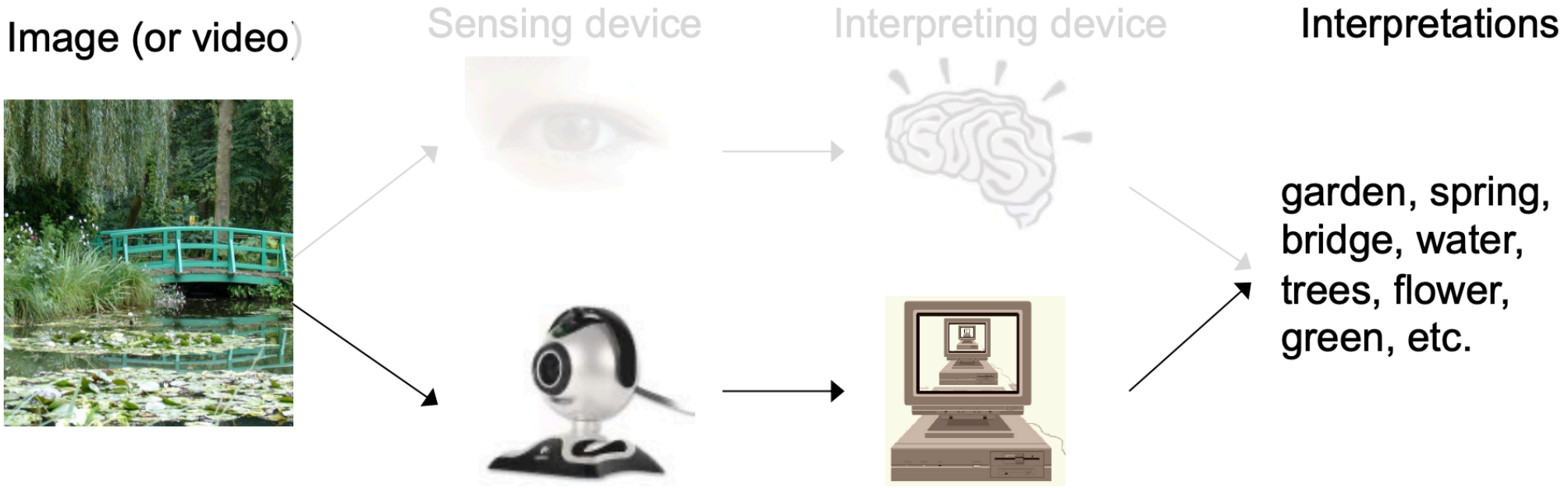
• What is image processing?

- Image processing describes the process of digitally transforming an image and executing specific operations to obtain useful information from it. Image processing systems often treat images as 2D signals when applying some predetermined signal processing approaches.
- Types of image processing include the following:
  - **pattern recognition** to measure various patterns around objects in an image;
  - **recognition** to detect or differentiate objects within an image;
  - **retrieval** to browse or search an extensive image database for an image like the original image;
  - **sharpening and restoration** to create an enhanced image from the original image; and
  - **visualization** to identify objects not visible in an image.
  - …

What is Computer Vision?



What is Computer Vision?

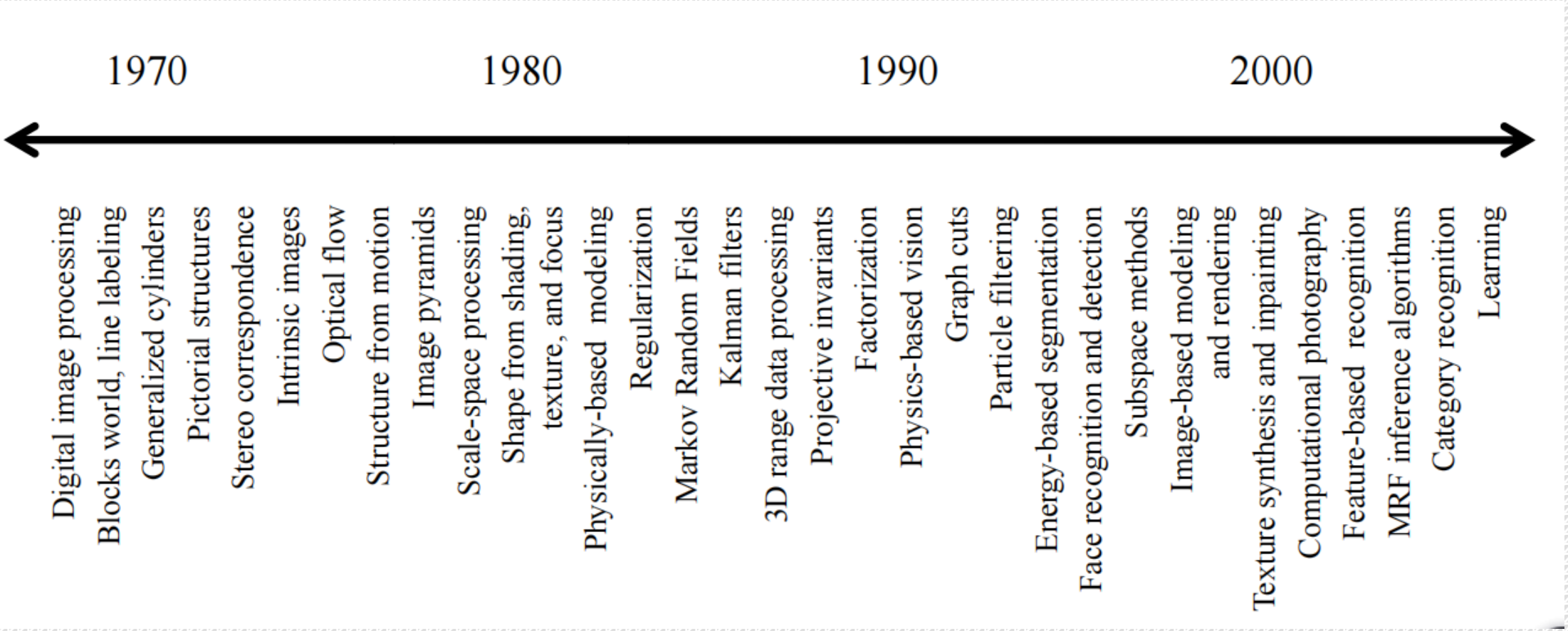


Computer vision VS image processing?

	Image processing	Computer vision
Goals	Manipulating visual aspects of images.	Extracting insights from images and videos.
Input/Out-put	Input and output are both images.	Input can be both image and video. Output can be an interpretation, which is often non-visual.
Scope	Low-level operations that affect pixels within an image.	More comprehensive.
Methods	More straightforward operations.	Complex algorithms and techniques.
Common Applications	Image editing software (like Photoshop), medical imaging enhancements, etc.	Autonomous vehicles, robotics, augmented reality, etc.

Source : https://eastgate-software.com/difference-between-computer-vision-and-image-processing/#:~:text=While%20they%20may%20seem%20similar,on%20enhancing%20and%20altering%20images.

The history of Computer vision



The history of Computer Vision

- **1970s.** AI lab at MIT start his first Machine Vision course. Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw.

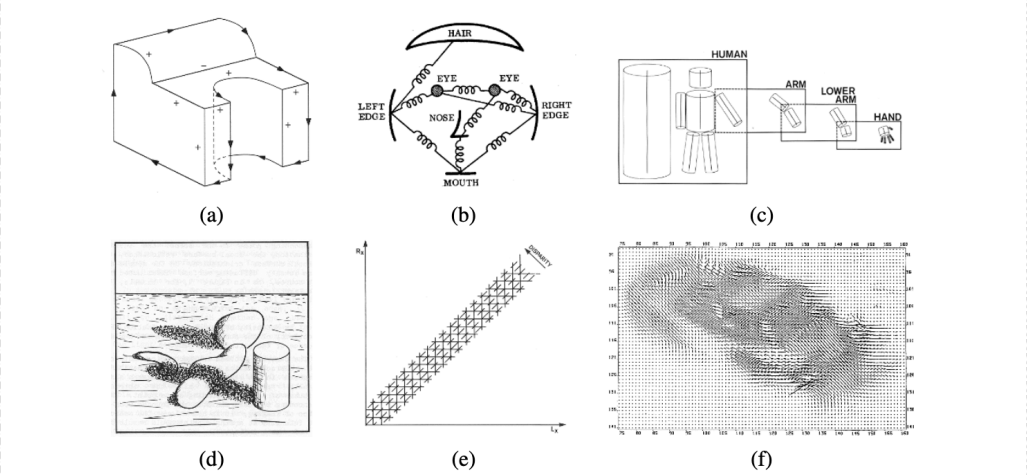
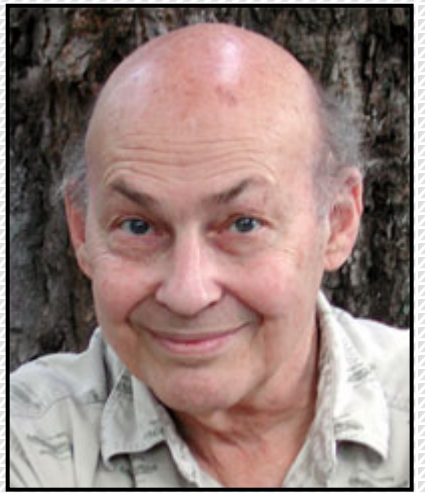


Figure 1.7 Some early (1970s) examples of computer vision algorithms: (a) line labeling (Nalwa 1993) © 1993 Addison-Wesley, (b) pictorial structures (Fischler and Elschlager 1973) © 1973 IEEE, (c) articulated body model (Marr 1982) © 1982 David Marr, (d) intrinsic images (Barrow and Tenenbaum 1981) © 1973 IEEE, (e) stereo correspondence (Marr 1982) © 1982 David Marr, (f) optical flow (Nagel and Enkelmann 1986) © 1986 IEEE.



Marvin Minsky, MIT
Turing award,1969

The history of Computer Vision

- **1980s.** A lot of attention was focused on more sophisticated mathematical techniques for performing quantitative image and scene analysis. Start to be applied in the industry.

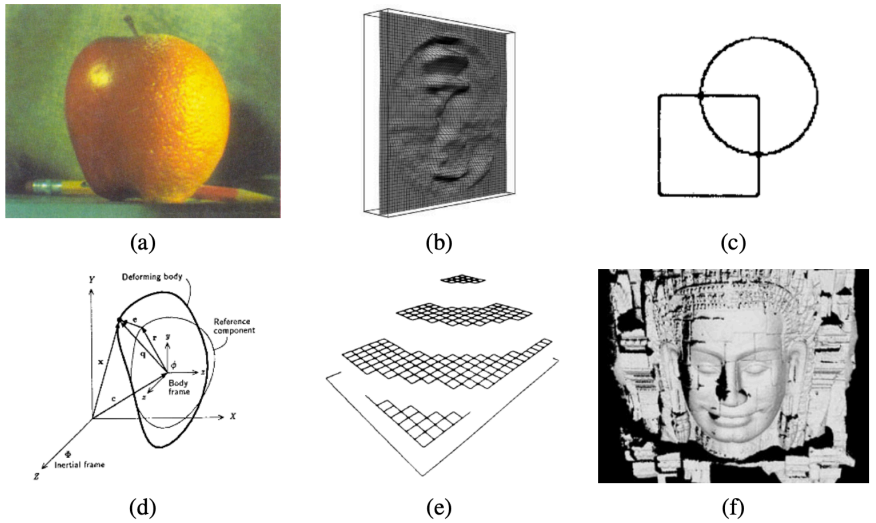


Figure 1.8 Examples of computer vision algorithms from the 1980s: (a) pyramid blending (Burt and Adelson 1983b) © 1983 ACM, (b) shape from shading (Freeman and Adelson 1991) © 1991 IEEE, (c) edge detection (Freeman and Adelson 1991) © 1991 IEEE, (d) physically based models (Terzopoulos and Witkin 1988) © 1988 IEEE, (e) regularization-based surface reconstruction (Terzopoulos 1988) © 1988 IEEE, (f) range data acquisition and merging (Banno, Masuda, Oishi *et al.* 2008) © 2008 Springer.

The history of Computer Vision

- **1990s:** While a lot of the previously mentioned topics continued to be explored, a few of them became significantly more active.
- Nvidia promoted Geforce 256, and the first concept of GPUs.

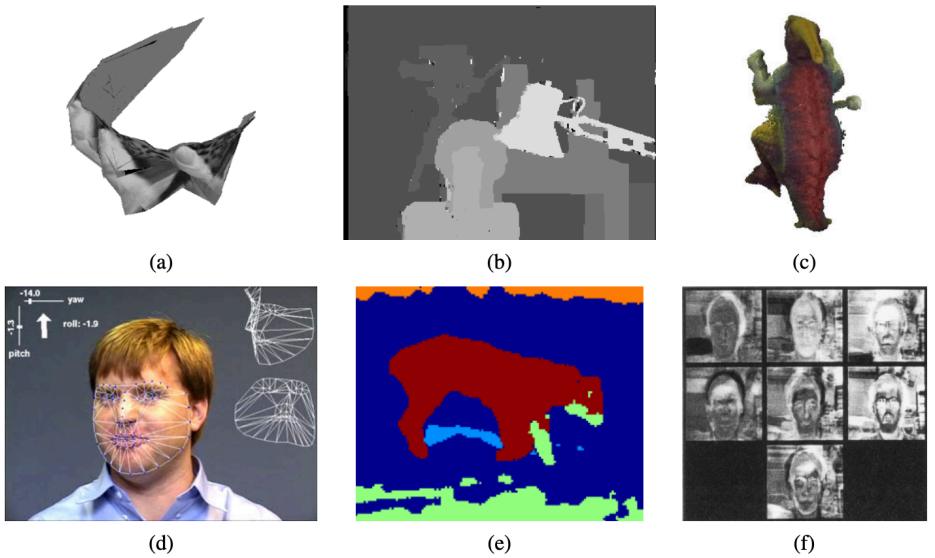


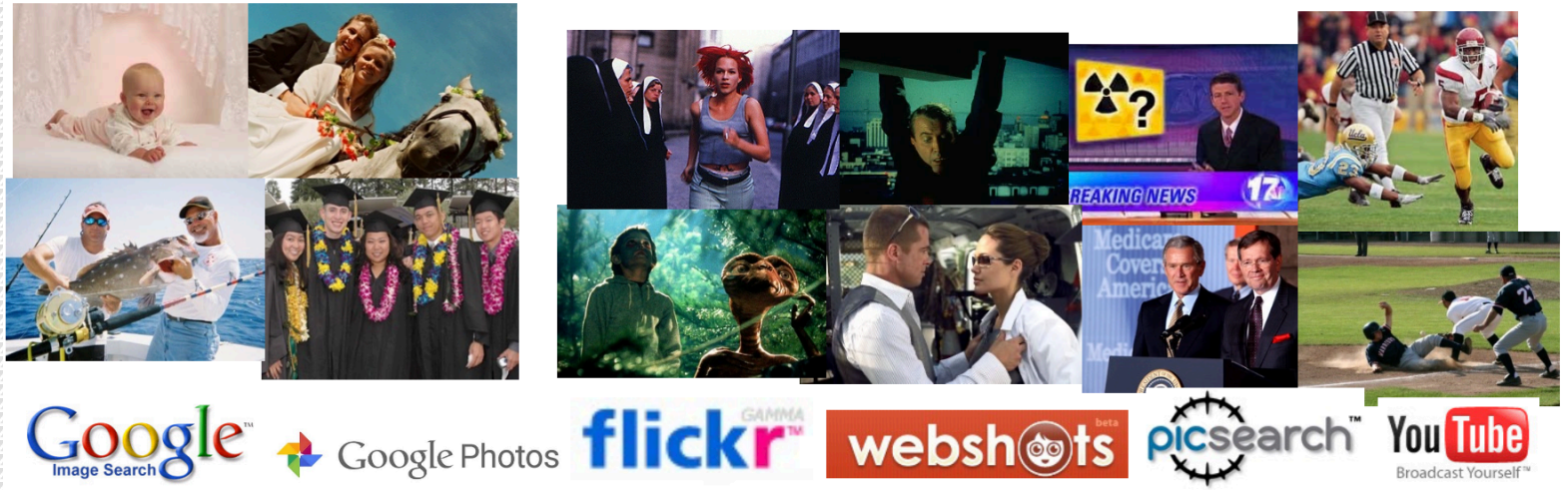
Figure 1.9 Examples of computer vision algorithms from the 1990s: (a) factorization-based structure from motion (Tomasi and Kanade 1992) © 1992 Springer, (b) dense stereo matching (Boykov, Veksler, and Zabih 2001), (c) multi-view reconstruction (Seitz and Dyer 1999) © 1999 Springer, (d) face tracking (Matthews, Xiao, and Baker 2007), (e) image segmentation (Belongie, Fowlkes, Chung *et al.* 2002) © 2002 Springer, (f) face recognition (Turk and Pentland 1991a).

The history of Computer Vision

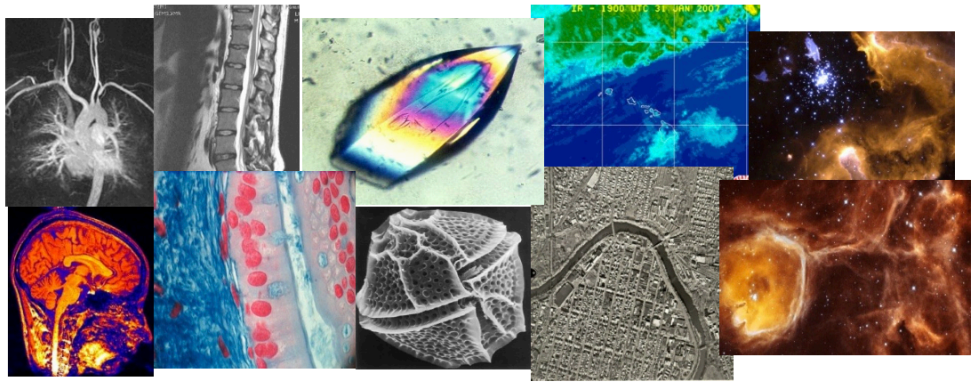
- At the beginning of 21st century:
  - The most notable development in computer vision during this decade was the increased interaction with computer graphics especially in the cross-disciplinary area of image-based modeling and rendering
  - Image-based modeling techniques for automatically creating realistic 3D models from collections of images were also being introduced
- Now: the visual recognition research is the application of sophisticated machine learning techniques to computer vision problems

Why study image processing (Computer Vision)

- Vision is useful: Image and video are everywhere!



Surveillance and security



Medical and scientific images

Why study image processing (Computer Vision)

Special effects: shape and motion capture

Why study image processing (Computer Vision)

- Face detection:

Why study image processing (Computer Vision)

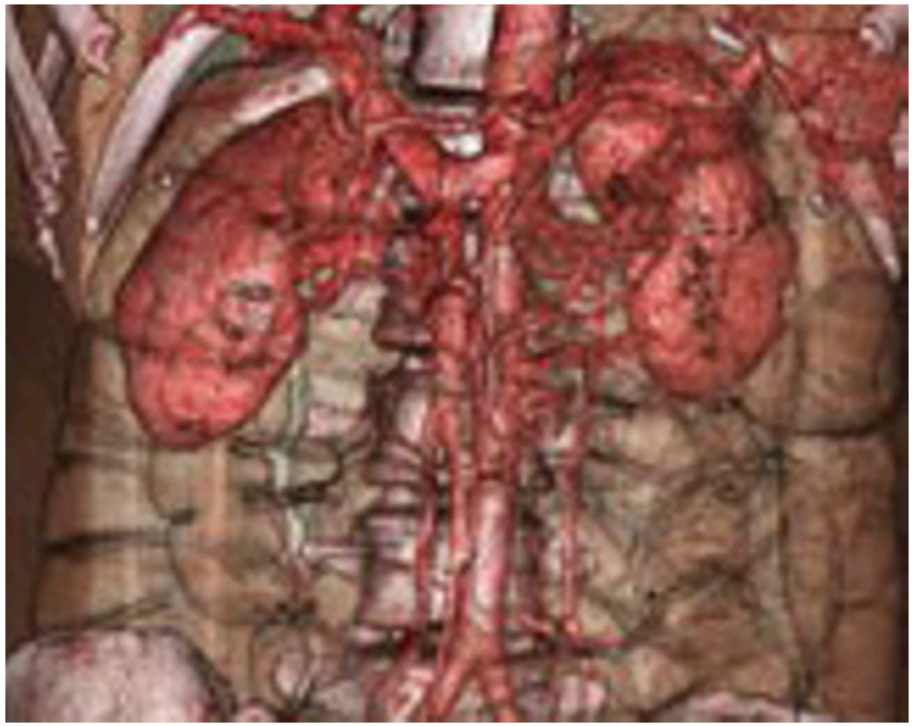
- Mechanical inspection :

http://www.cognitens.com/;

Why study image processing (Computer Vision)

- Retail:

Why study image processing (Computer Vision)

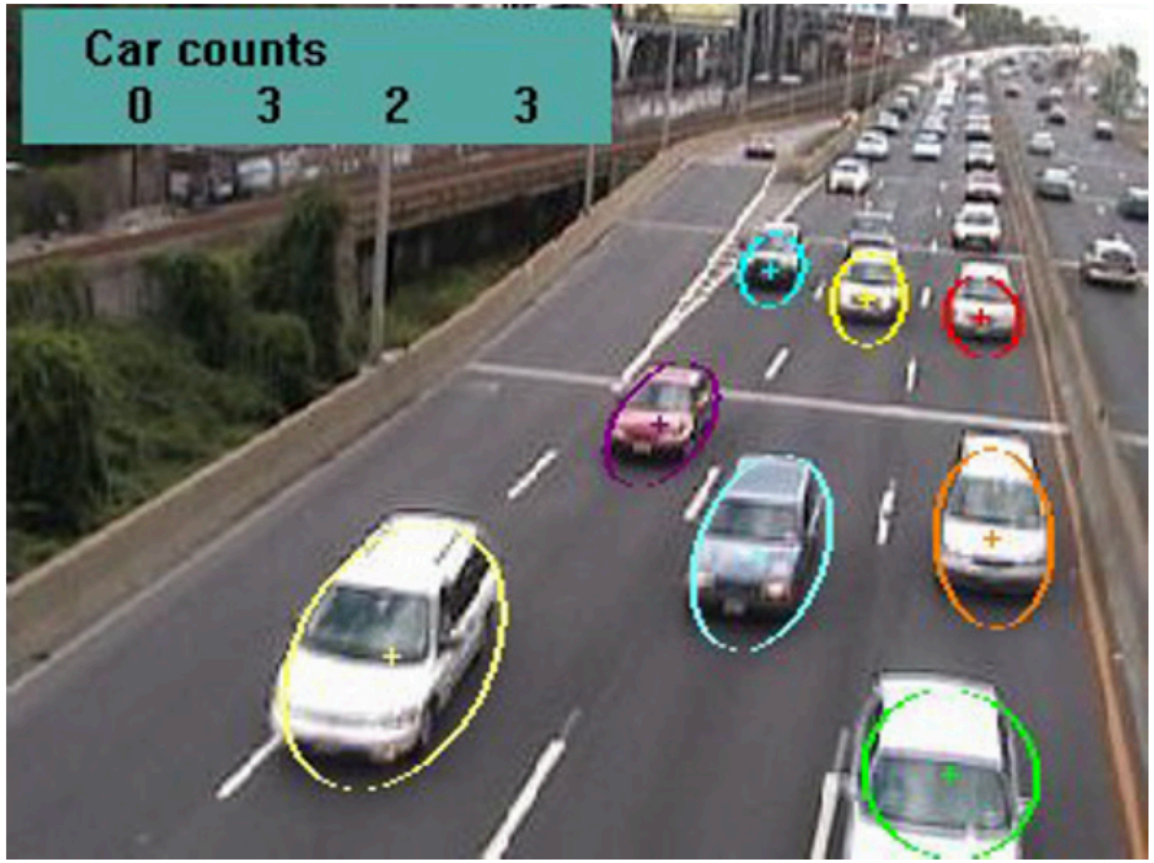


medical imaging http://www.clarontech.com/

Why study image processing (Computer Vision)

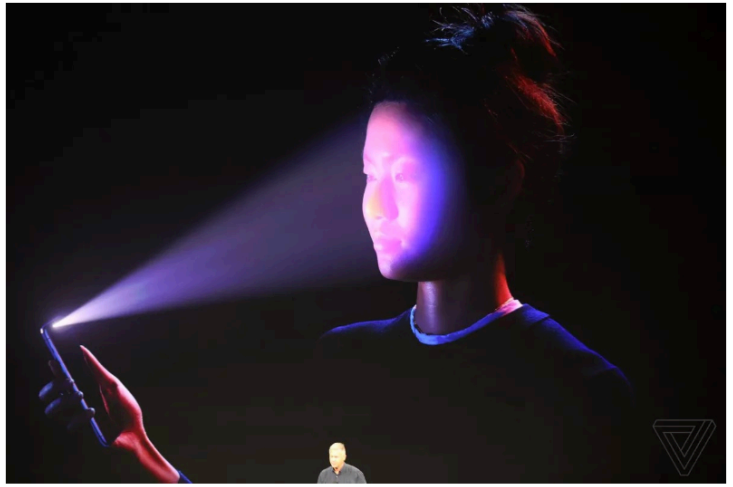
automotive safety http://www.mobileye.com/

Why study image processing (Computer Vision)



surveillance and traffic monitoring

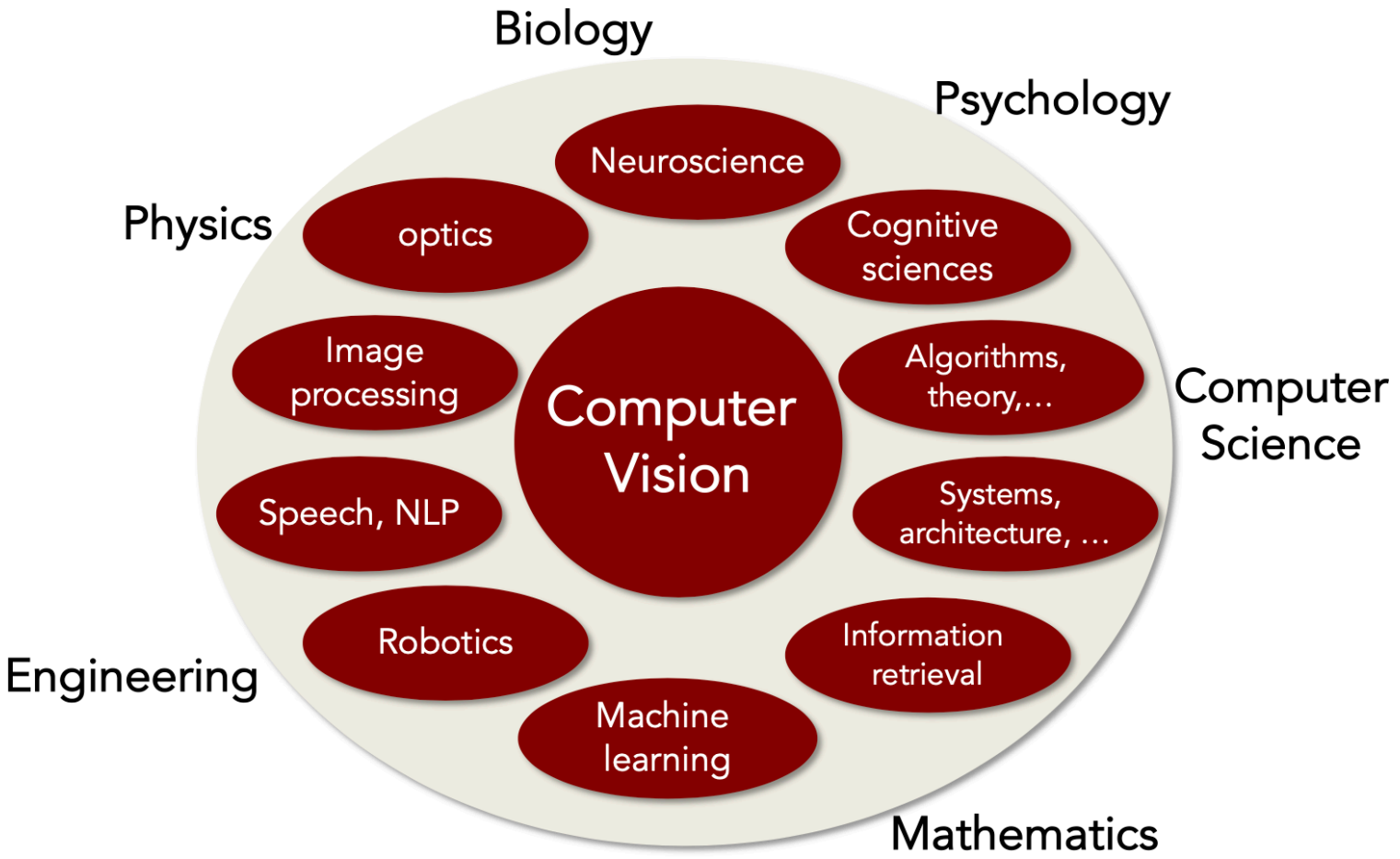
Why study image processing (Computer Vision)



Why study image processing (Computer Vision)

- Others: https://v.youku.com/v_show/id_XNjM3Njc0OTc0NA==.html?playMode=pugv&frommaciku=1

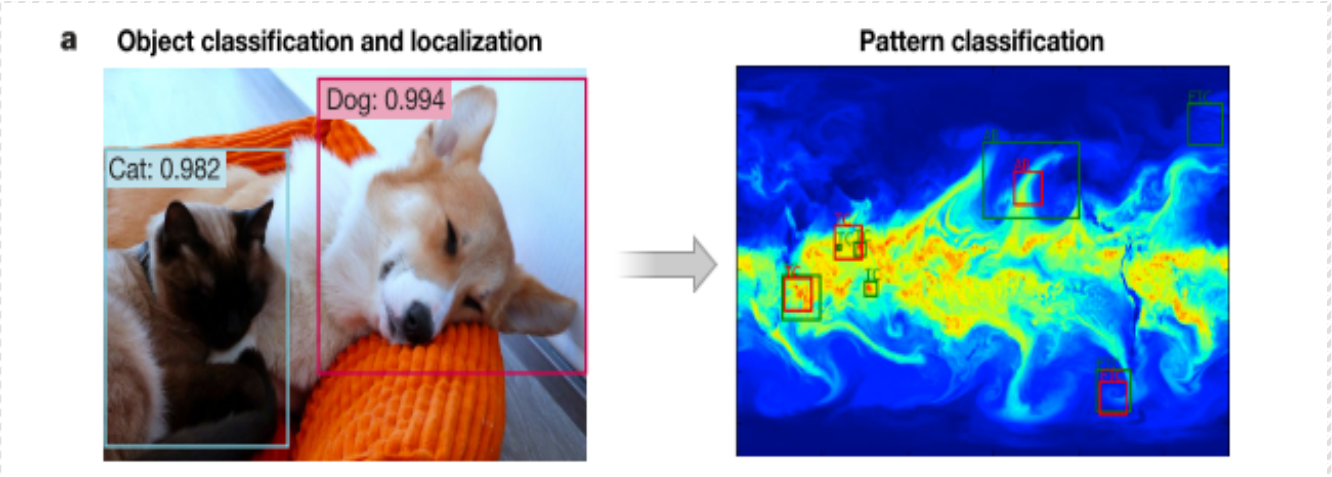
CV + Science

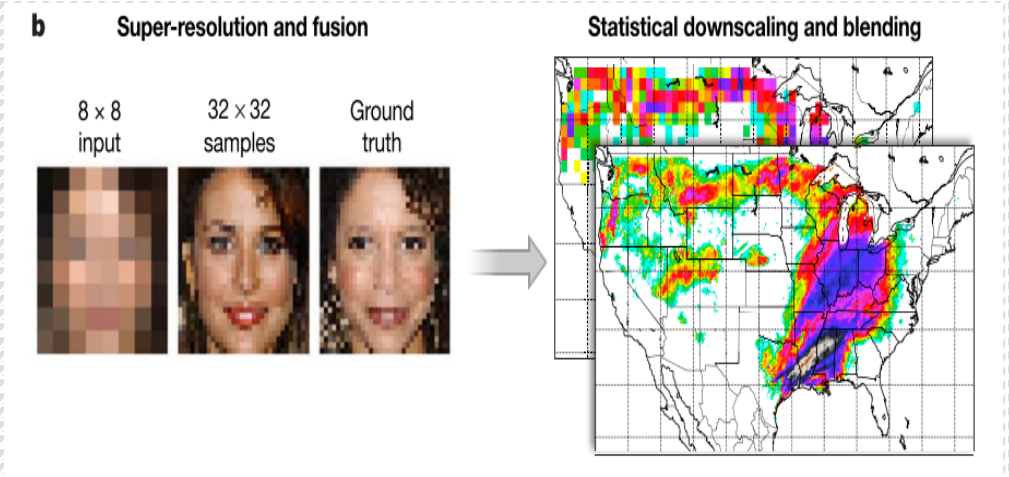


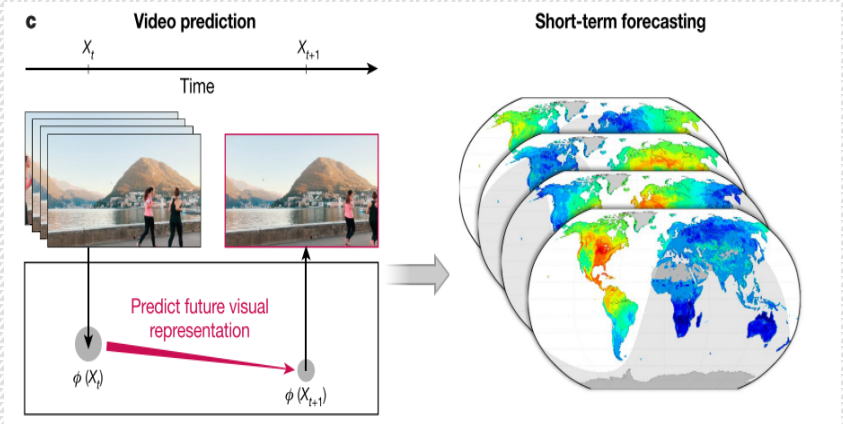
Slides from Stanford University CS131 Computer Vision: Foundations and Applications

CV + Science

- CV in Earth Science

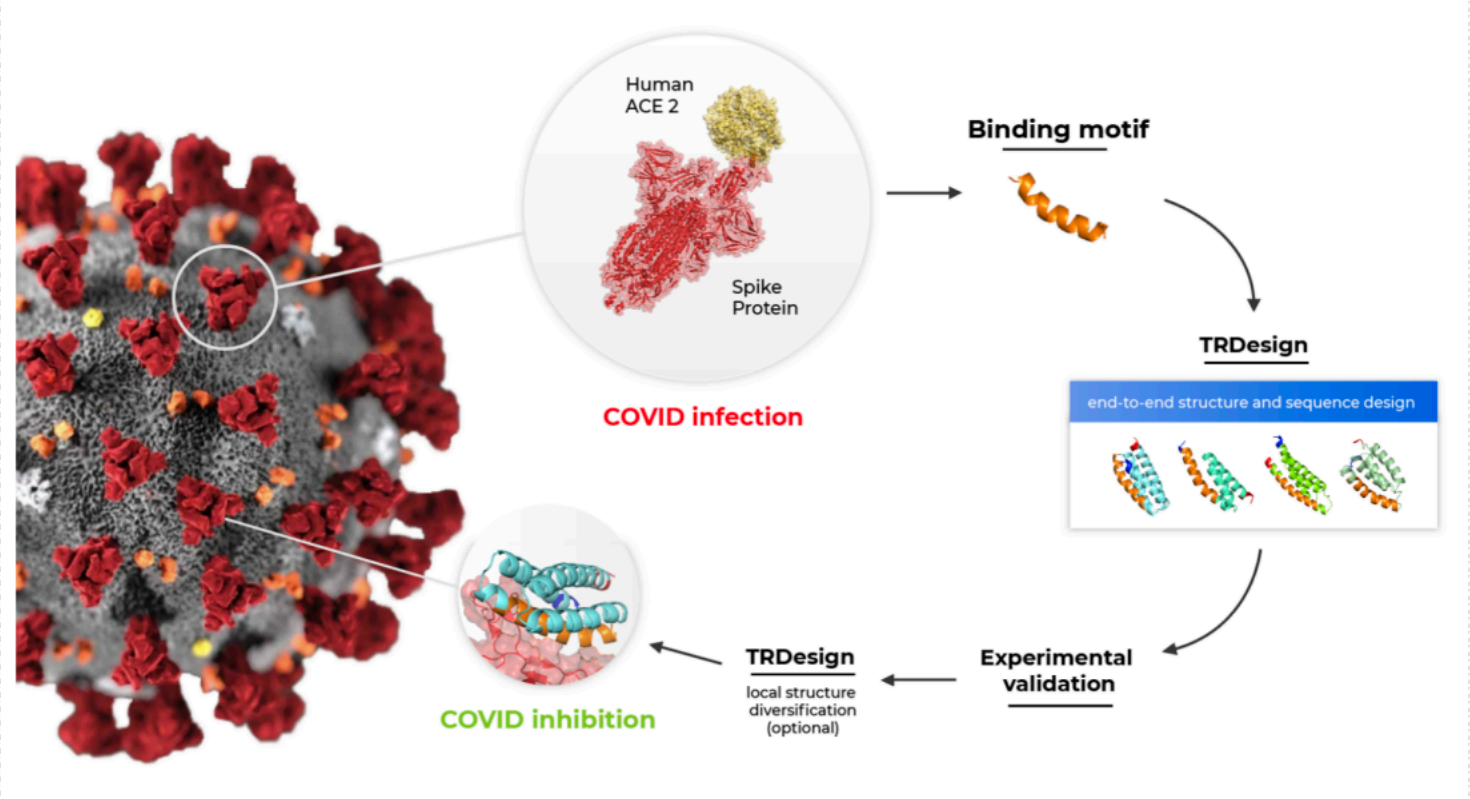






CV + Science

- CV in Biological Science



What are the challenges for computer vision?

- occlusion

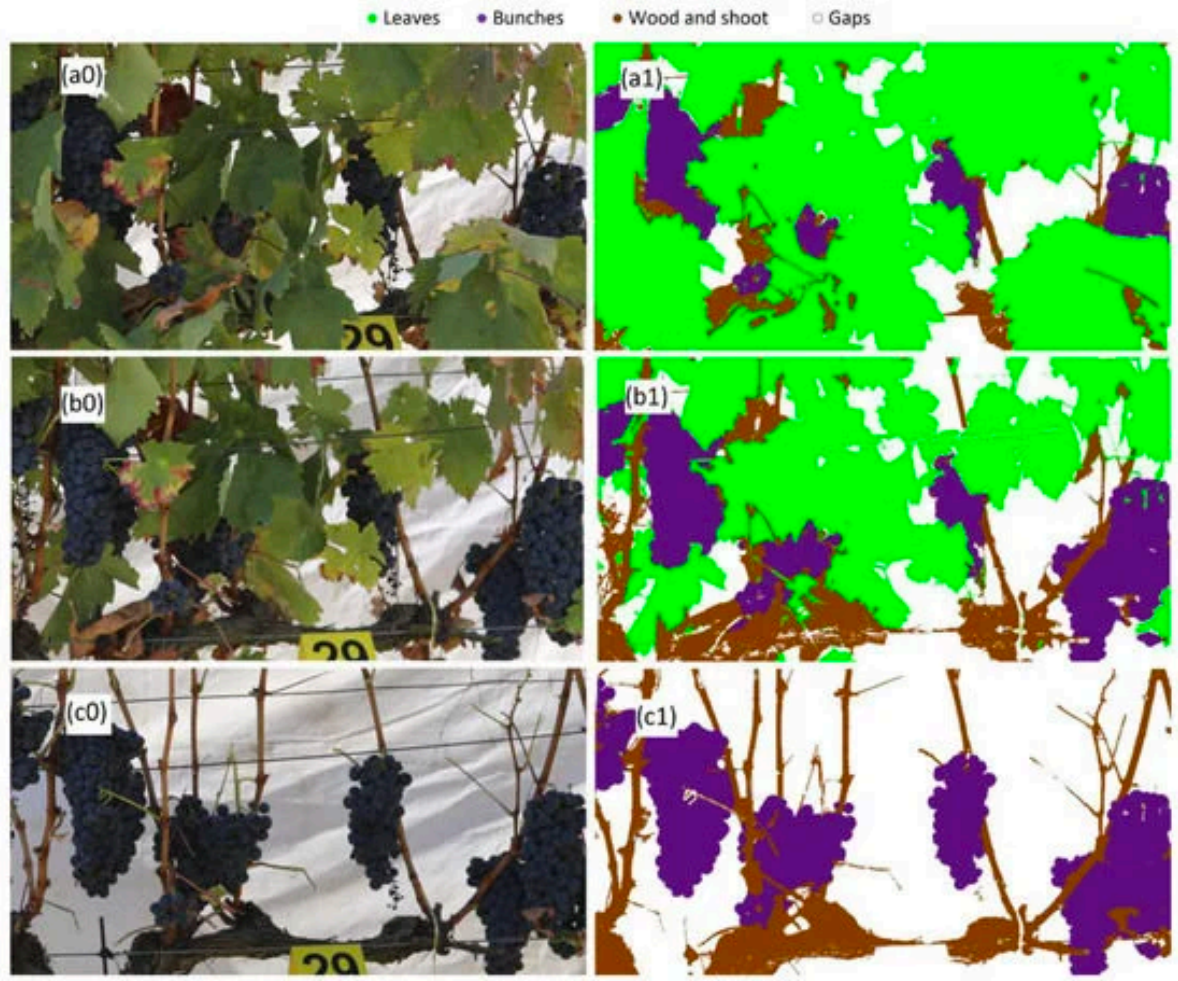


Image from https://www.mdpi.com/2073-4395/11/5/1003

What are the challenges for computer vision?

Slide from Linda Shapiro, University of Washington

What are the challenges for computer vision?

- object intra-class variation






slide by Fei-Fei, Fergus & Torralba

What are the challenges for computer vision?

- object extra-class similarity

What are the challenges for computer vision?

- 1024×768:

  1024×768×3Byte≈2.3MB

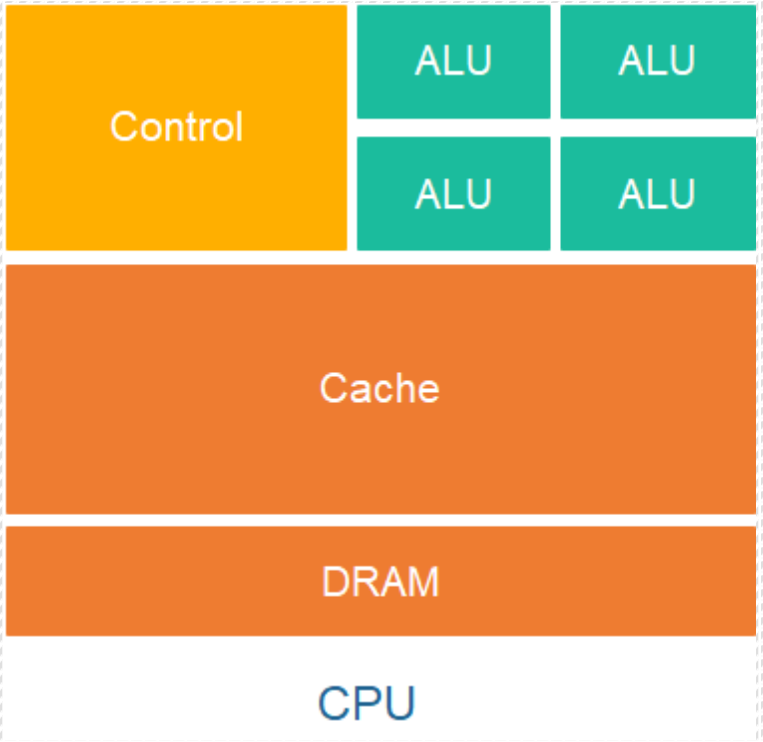
- If video, the size is even bigger。

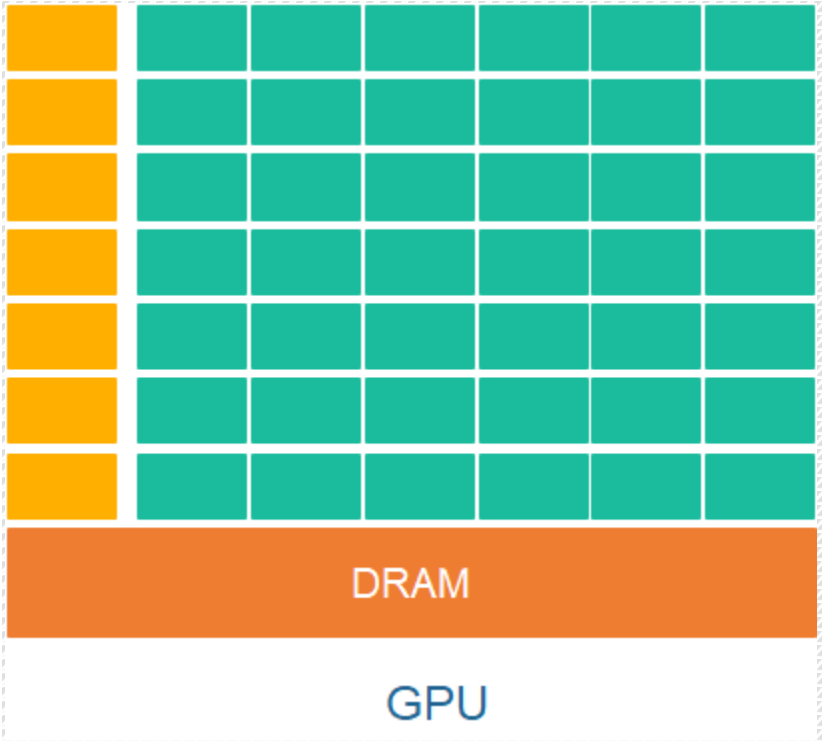
  CIF format, 10 mins video：

  352×288×3Byte×10min×60s/min×30 frame/s ≈5TB

- Big data + very deep learning --> high demand of computing resources

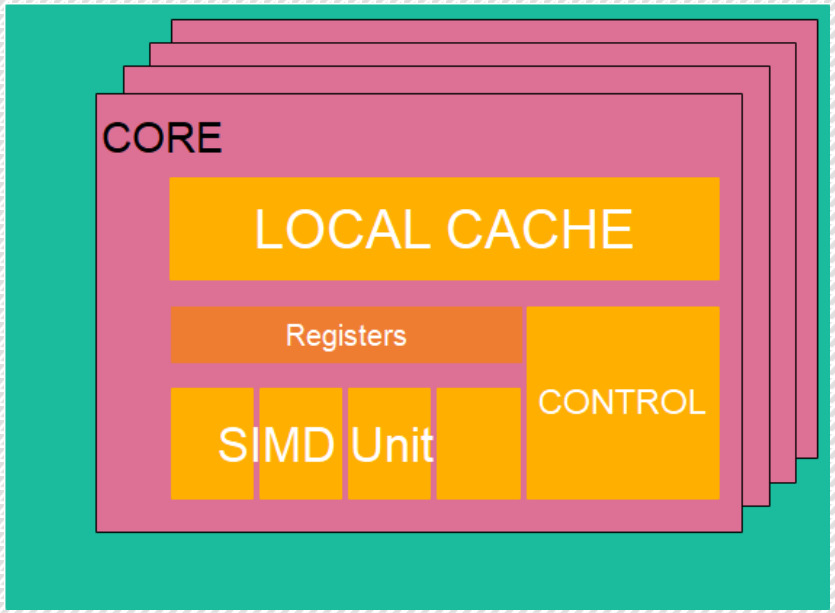
CPU and GPUs

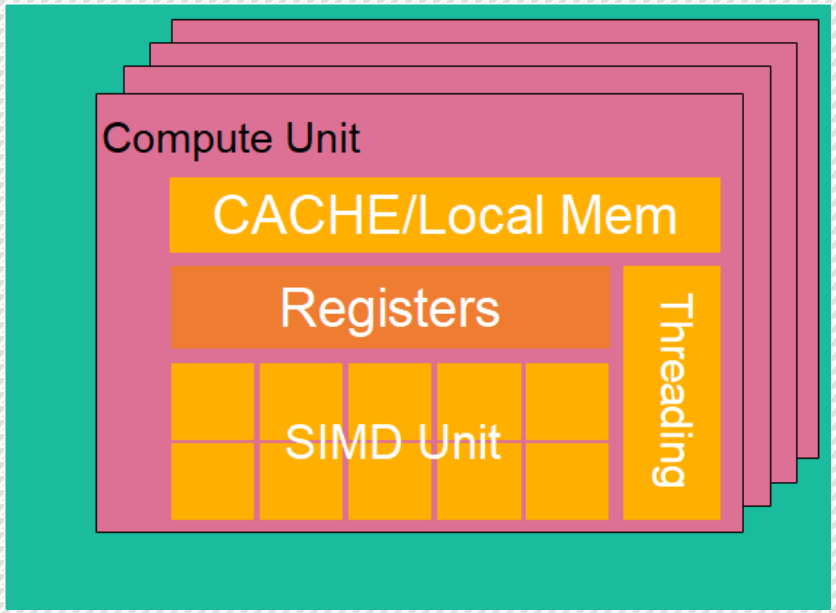




Green：Computing Unit

CPU and GPUs





Cache、Local Memory ： CPU > GPU

Threads（线程数）：GPU > CPU

Registers（寄存器）：GPU > CPU

SIMD Unit（单指令多数据流）：GPU > CPU

Key Journals and Conferences

Key Journals and Conferences

ICCV 2024: 18. International Conference on Computer Vision

September 26-27, 2024 in Vancouver, Canada

Key Journals and Conferences

The 18th European Conference on Computer Vision ECCV 2024

Sun Sep 29th through Fri Oct 4th, 2024
at MiCo Milano

The **European Conference on Computer Vision** (**ECCV**) is a biennial premier research conference in Computer Vision and Machine Learning, managed by the European Computer Vision Association (ECVA). It is held on even years and gathers the scientific and industrial communities on these areas. The first ECCV was held in 1990 in Antibes, France, and subsequently organized all over Europe. Paper proceedings are published by Springer Science+Business Media.

Follow ECCV on Twitter (X)

Key Journals and Conferences

17th Asian Conference on Computer Vision

Key Journals and Conferences

ICLR 2024

The Twelfth International Conference on Learning Representations

Messe Wien Exhibition and Congress Center, Vienna Austria
May 7th, 2024 to May 11th, 2024

Registration

Pricing » | Registration 2024 | Registration Cancellation Policy »

Announcements

- Volunteer Application now **OPEN** until March 15
- Financial Assistance Application now **OPEN** until March 5
- ICLR 2024 Hotel Reservations available **HERE**
- Press Accrediation now **OPEN**
- Authors of eligible TMLR publications can submit a request to present in ICLR conference! Find more here.
- BEWARE of Predatory ICLR conferences being promoted through the World Academy of Science, Engineering and Technology organization.

Key Journals and Conferences

Key Journals and Conferences

NEURAL INFORMATION PROCESSING SYSTEMS

NeurIPS 2024, the Thirty-eighth Annual Conference on Neural Information Processing Systems, will be held at the Vancouver Convention Center
Monday Dec 9 through Sunday Dec 15. Monday is an industry expo.

Talk from the top scientist

- TED talk from Sheperd Doeleman: Inside the black hole image that made history

Homework

- 1. Install Anaconda and Jupyter Notebook
- 2. Familiar with Python (scikit-image)
- 3. Download an image from website
- 4. Read the image and display the image using scikit-image