

Computer Vision

(VO, 911.908, Winter term 23/24)

Roland Kwitt

Department of Artificial Intelligence and Human Interfaces (AIHI)

University of Salzburg, Austria

Schedule for today

- Logistics
- Resources
- What is computer vision?
- Some typical computer vision problems

Logistics

- **E-Mail:** Roland [dot] Kwitt [at] plus [dot] ac [dot] at
- **Course material:** available online at <http://rkwitt.org> (→ Teaching)

Logistics

Grading

1 final exam at the end of the course (see course website for details).

Resources

PyTorch (most recent version v2.0.1)

<https://pytorch.org/>

scikit-learn (Python)

<http://scikit-learn.org/stable/>

scikit-image (Python)

<http://scikit-image.org>

last time I checked, they are pre-installed when running Google Collab.

Resources

“Deep Learning”

Goodfellow, Bengio, Courville

<http://www.deeplearningbook.org/>

Please check the course website for updates on relevant research papers.

Resources

Throughout the lecture, we will primarily cover “**modern**” approaches to computer vision, using (deep) neural networks.

However, I still do want to mention a “classic” vision book that is definitely worth reading.

“Computer Vision: Algorithms and Applications”

R., Szeliski, 2010

<http://szeliski.org/Book/>

What is computer vision?



Courtesy of Szeliski



(Luiz Gomez Photos)

The **human visual system (HVS)** is remarkably good at so many tasks, e.g.,

- detecting people & object's
- perceiving translucency, shape, color (see figure on the left),
- counting people & vehicles (see middle figure)
- recognizing scenes, etc.

What is computer vision?

Remarkable fact:

A substantial fraction of the macaque's total cortical area is devoted to vision (approx. 15% according to [[Hubel, "Eye, Brain and Vision"](#)])

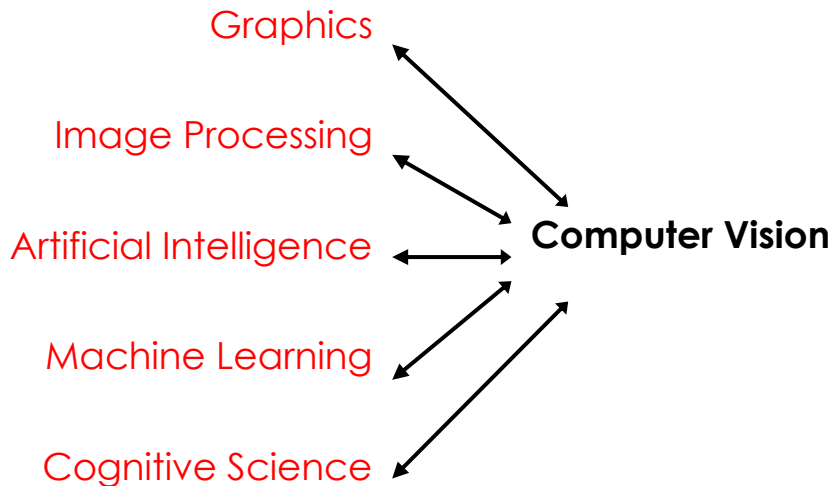
What is computer vision?

We want to build systems for **automatic understanding of images/videos**.

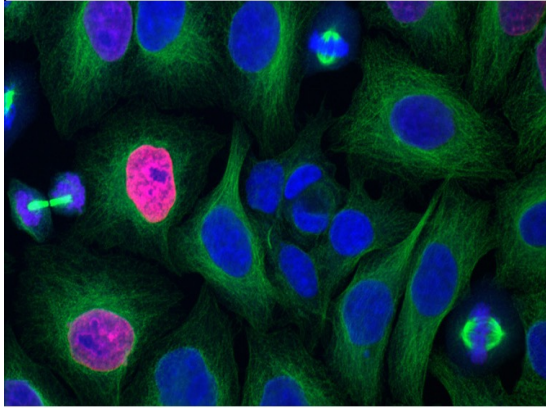
This includes, but is not limited to,

- 1) inferring properties of the 3D world (**measurement**)
- 2) enabling recognition of objects, people, scenes, etc. (**perception**)
- 3) mining, searching and interacting with visual data (**search / organization**)

What is computer vision?



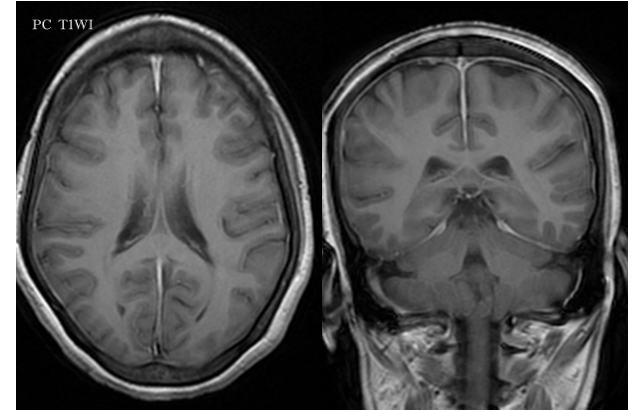
Bio-Imaging



Surveillance



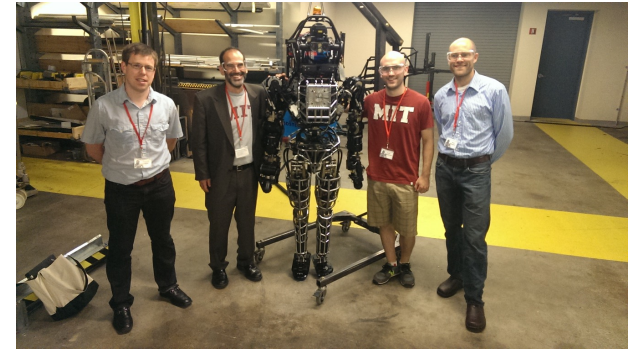
Medical Imaging



Autonomous Cars



Mapping



Robotics

Overview of the field

Conferences

Computer Vision

- Computer Vision and Pattern Recognition (CVPR)
- International Conference on Computer Vision (ICCV)
- European Conference on Computer Vision (ECCV)
- British Machine Vision Conference (BMVC)

Machine learning

- Neural Information Processing Systems (NIPS)
- International Conference on Machine Learning (ICML)
- International Conference on Learning Representations (ICLR)

Overview of the field

Some metrics: checkout Google scholar on that, [here](#)!

Typical vision problems

Recognition



→ “Saiga antelope”

Task: Label an image by its class membership, e.g., class would be “Saiga antelope” in the example above.

Typical vision problems

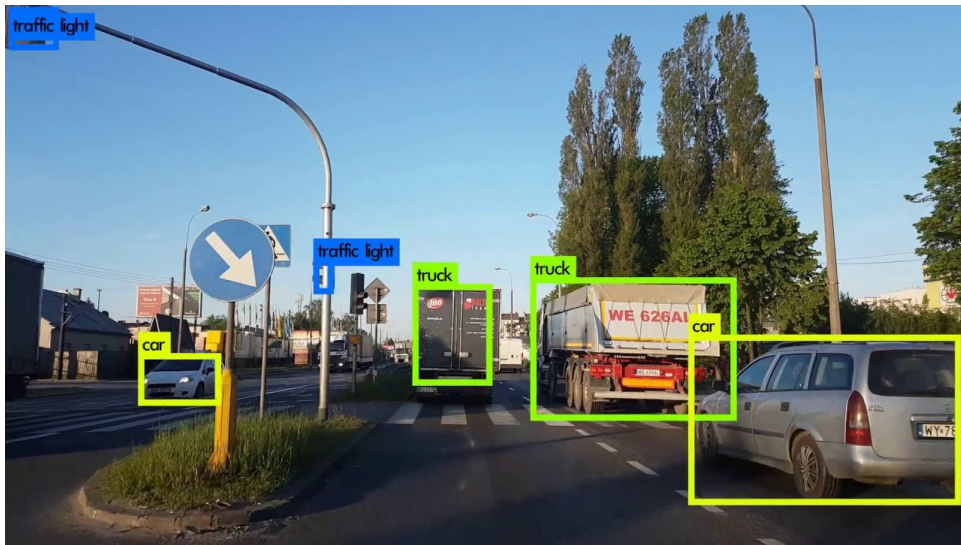
Segmentation



Task: Label every pixel in an image by its class membership, e.g., "Human" vs. "Sheep" vs. "Background" in the example.

Typical vision problems

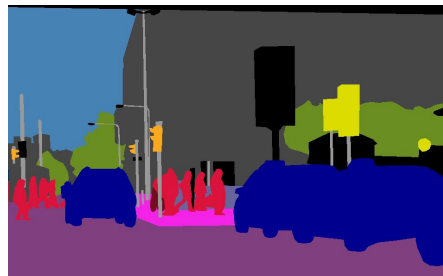
Detection



Task: Locate and label objects. E.g., each object is marked by a bounding box (i.e., coordinates) and each box is assigned a label.

Typical vision problems

Panoptic Segmentation



Semantic
segmentation



Instance
segmentation



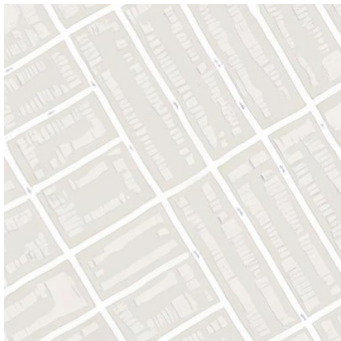
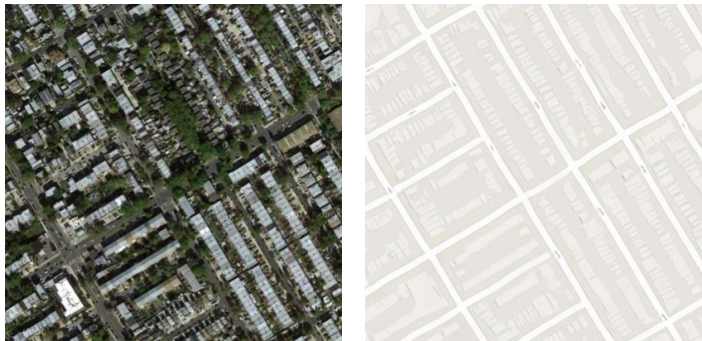
Panoptic
segmentation

Image source: Kirillov et al., "Panoptic Segmentation", In: CVPR '19

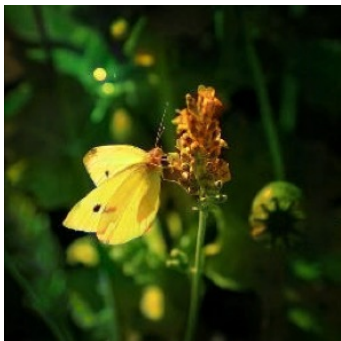
Goal: Unify semantic segmentation and instance segmentation.

Typical vision problems

Image-to-Image translation



Arial \rightarrow Map



BW \rightarrow Color

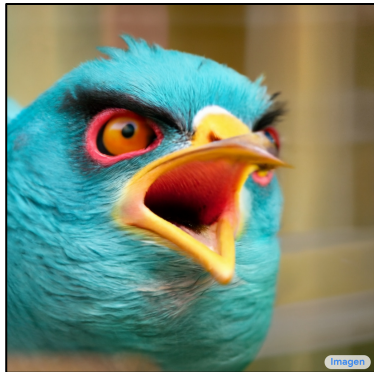


Label map \rightarrow Image



Typical vision problems

Text-to-Image Generation



Prompt: An extremely angry bird.



Prompt: A photo of a Persian cat wearing sunglasses and a red shirt skateboarding on a beach.

Typical vision problems

Synthesizing novel views



Typical vision problems

[Dalle-2](#) (not publicly accessible, you need to sign up and get approved).

Some further interesting resources

Browse D. H. Hubel's online book (Nobel prize winner, 1981)

<http://hubel.med.harvard.edu/book/bcontext.htm>

Read one of his and Torsten Wiesel's seminal papers, e.g.,

Receptive Fields and Functional Architecture of Monkey Striate Cortex,

D. H. Hubel and T. N. Wiesel

J. Physiol., 1968

<http://www.ncbi.nlm.nih.gov/pubmed/4966457>)