Visual AI

Deep learning models for Computer Graphics and Computer Vision

CPSC 533R – 2020/2021 Term 1

Helge Rhodin

Visual AI

CPSC 533R – 2020/2021

Lecture 1. Overview and programming environment

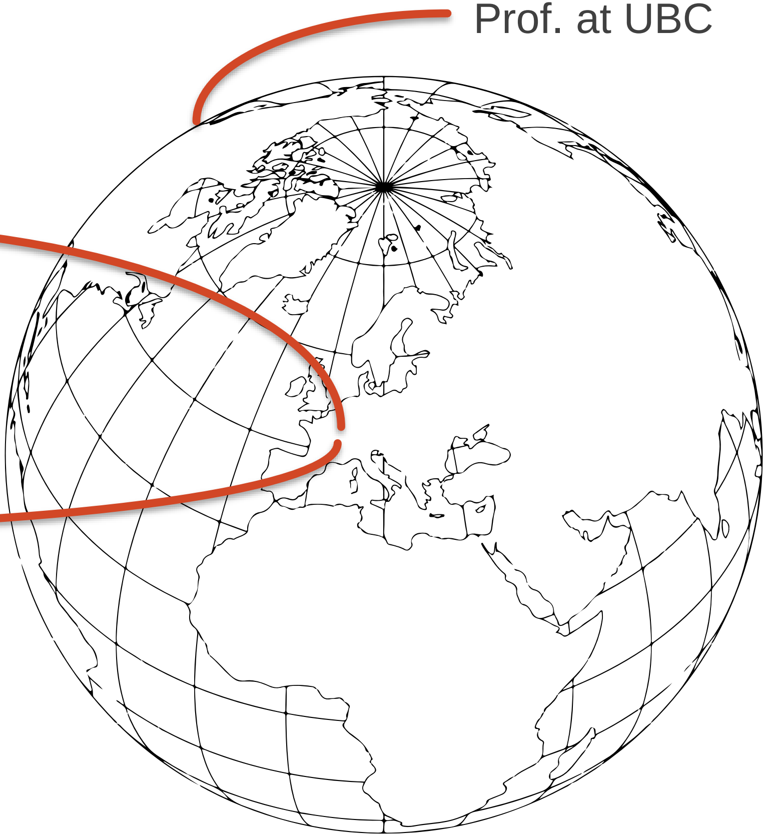
Helge Rhodin

Helge Rhodin

MSc and BSc
Saarland University, Germany

PhD
Max-Plank institute for Informatics

Lecturer and postdoc
EPFL, Switzerland



Prof. at UBC

Organization

Instructor:

Helge Rhodin

rhodin@cs.ubc.ca

Office hours:

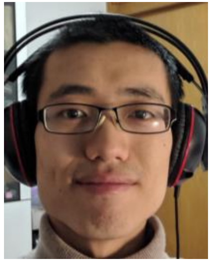
TBD

Room: Zoom (via Canvas)

Teaching assistant:

Yuchi Zhang

yuchi45@cs.ubc.ca



Office hours:

TBD: https://doodle.com/poll/qynygb3daznvqtgy

Room: Zoom (via Canvas)

Course Website

Curriculum	https://www.cs.ubc.ca/~rhodin/2020_2021_CPSC_533R/
Forum	https://piazza.com/ubc.ca/winterterm12020/cpsc533R
Canvas	https://canvas.ubc.ca/courses/53581

Official Disclaimer

Keep in mind that some UBC courses might cover topics that are censored or considered illegal by non-Canadian governments. This may include, but is not limited to, human rights, representative government, defamation, obscenity, gender or sexuality, and historical or current geopolitical controversies. If you are a student living abroad, you will be subject to the laws of your local jurisdiction, and your local authorities might limit your access to course material or take punitive action against you. UBC is strongly committed to academic freedom, but has no control over foreign authorities (please visit http://www.calendar.ubc.ca/vancouver/index.cfm?tree=3,33,86,0 for an articulation of the values of the University conveyed in the Senate Statement on Academic Freedom). Thus, we recognize that students will have legitimate reason to exercise caution in studying certain subjects. If you have concerns regarding your personal situation, consider postponing taking a course with manifest risks, until you are back on campus or reach out to your academic advisor to find substitute courses. For further information and support, please visit: http://academic.ubc.ca/support-resources/freedom-expression

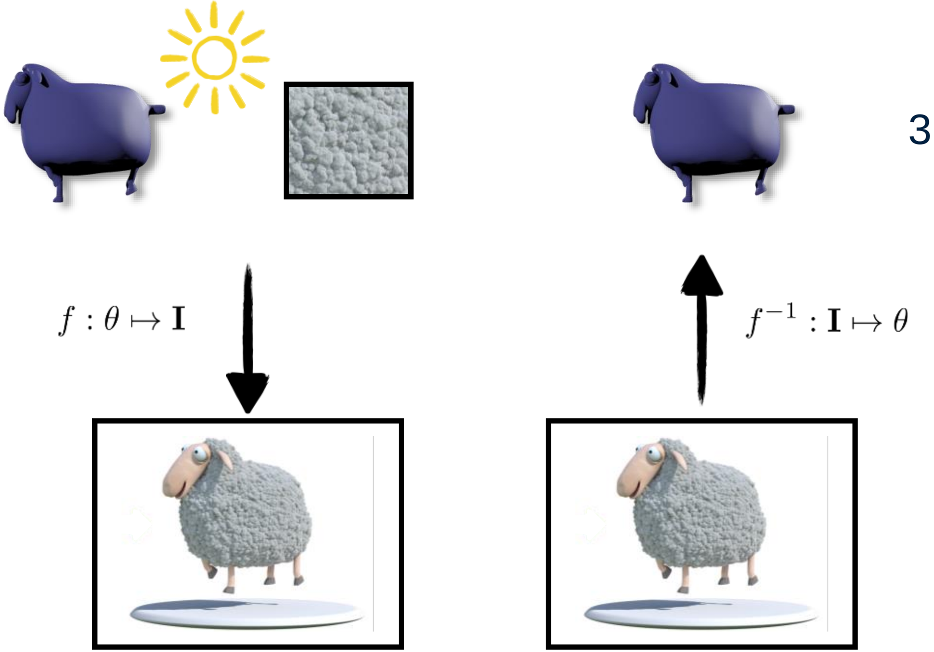
Computer graphics, computer vision and machine learning

Computer graphics

Computer vision

Scene model
(mesh, light source,
texture, …)

3D mesh



$f : \theta \mapsto \mathbf{I}$

$f^{-1} : \mathbf{I} \mapsto \theta$

Image
(pixel intensities)

Image

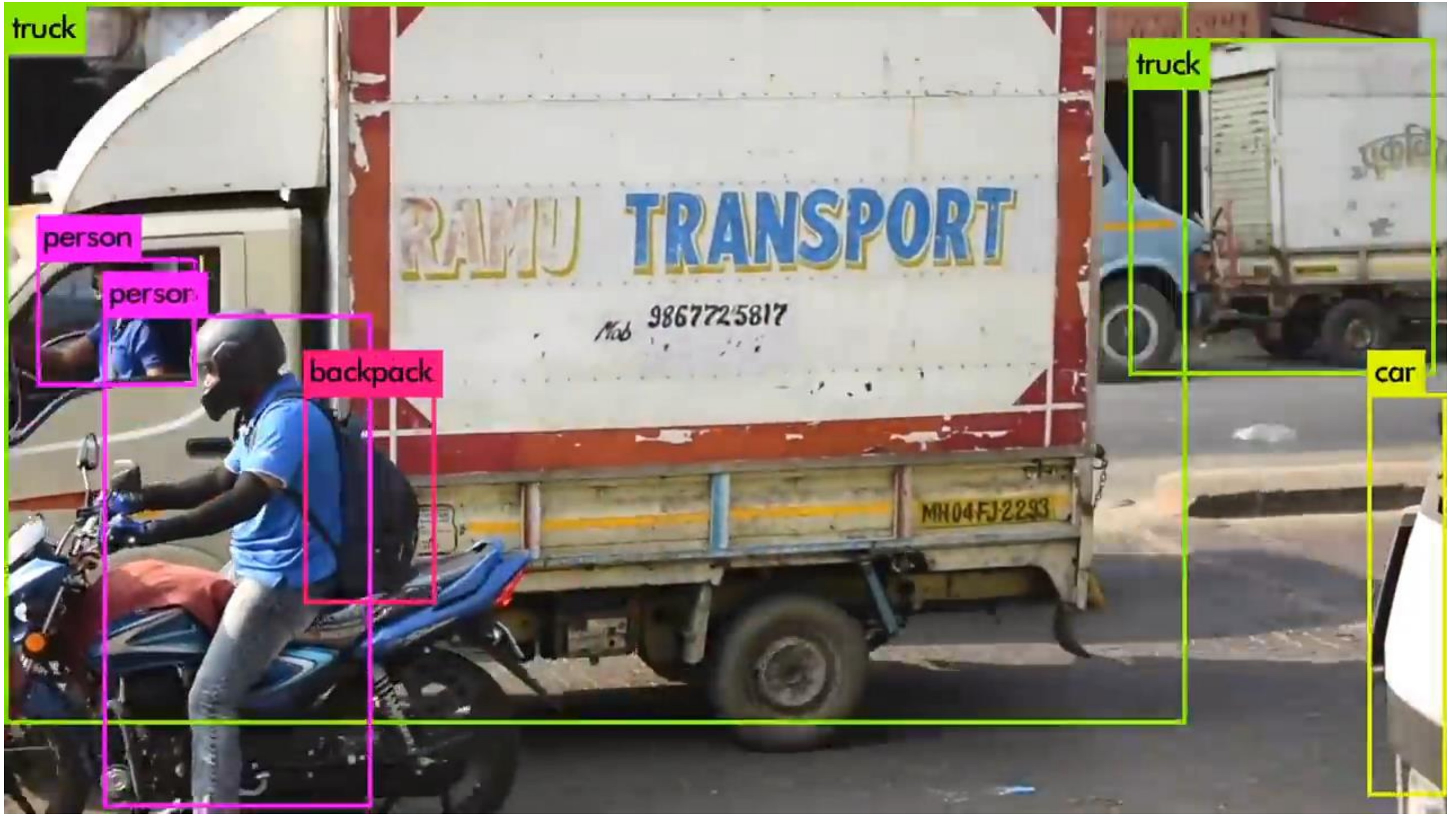
Overview

- 9 Lectures (~ once a week)
  - Introduction
  - Deep learning basics and best practices
  - Network architectures for image processing
  - Representing images and sparse 2D keypoints
  - Representing dense and 3D keypoints
  - Representing geometry and shape
  - Representation learning
  - Unpaired image translation
  - Attention models
- 3x Assignments
  - Playing with pytorch (5% of points)
  - Pose estimation (10% of points)
  - Shape generation (10% of points)
- 1x Paper presentation (Weeks 3 – 12)
  - Presentation, once per student (25% of points) (15 min + 15 min discussion)
  - Read and review one out of the two papers presented per session (10% of points)
- 1x Project (40 % of points)
  - Project pitch (3 min, week 6&7)
  - Project presentation (10 min, week 13&14)
  - Project report (6 pages, Dec 14)

Computer Vision Topics

A few examples

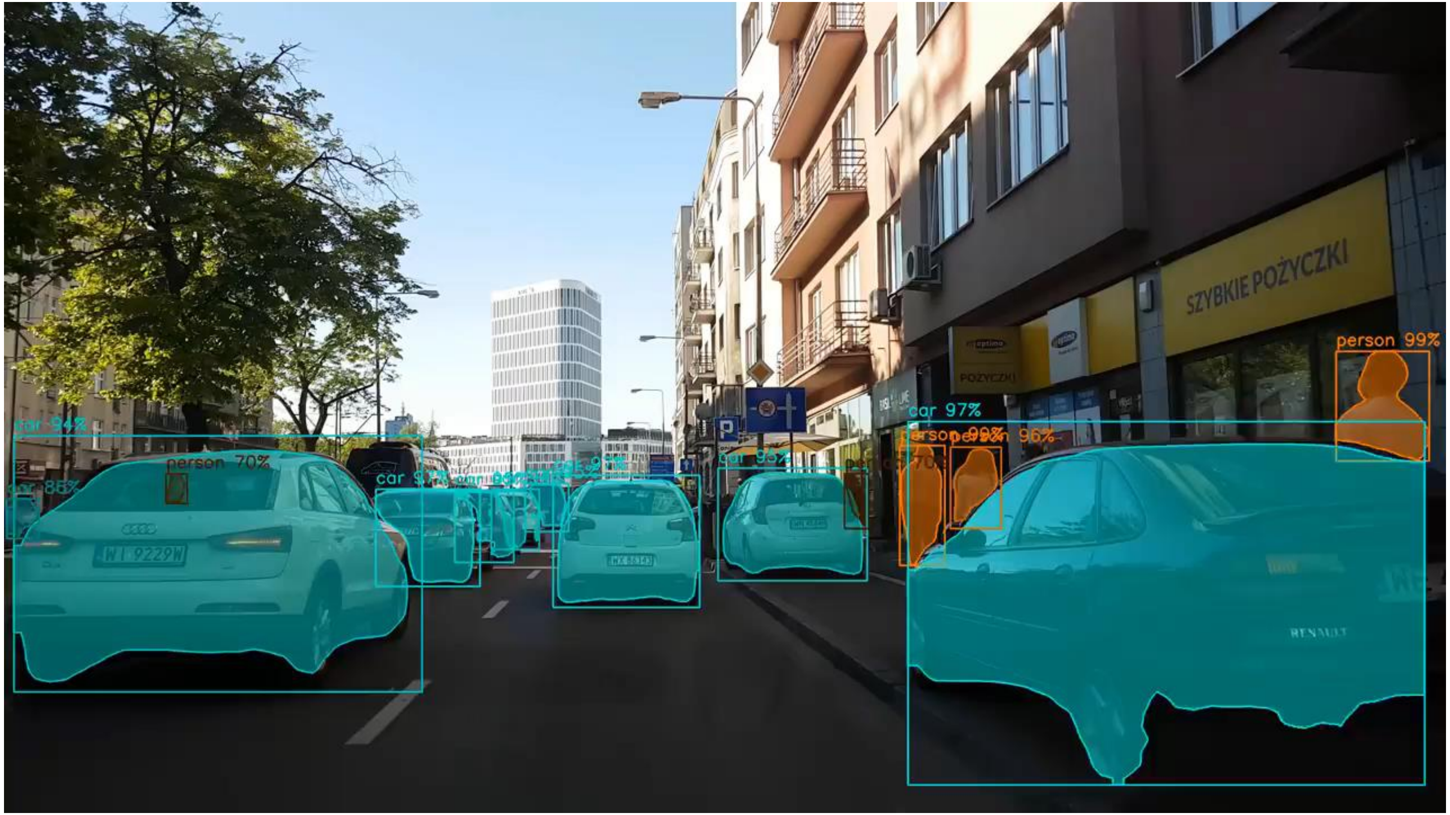
Object detection



[Redmon et al., YOLO: Real-Time Object Detection, CVPR 2016]

detection + offset regression

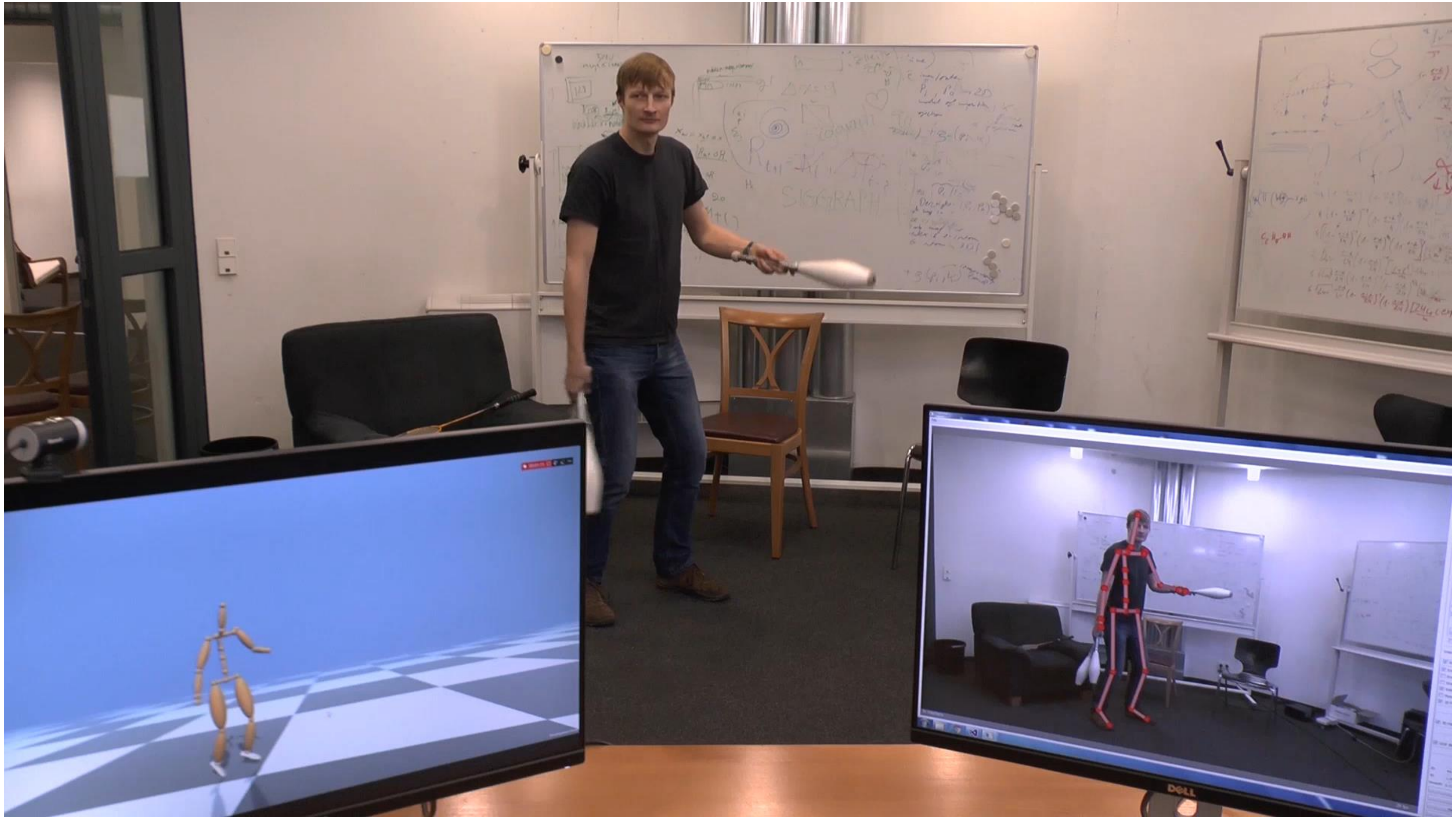
Instance segmentation



[He et al., Mask R-CNN, ICCV 2017]

detection & pixel classification

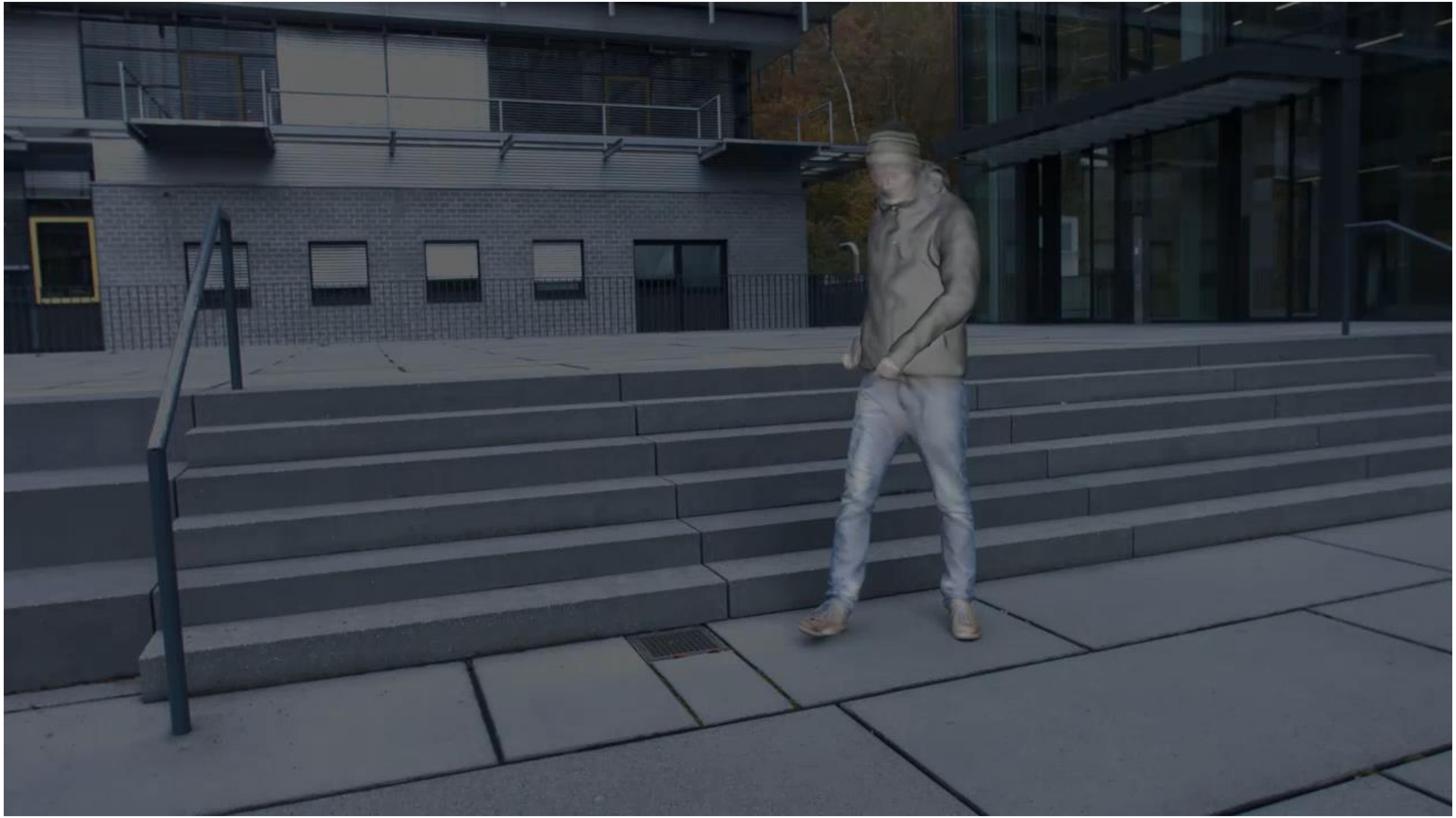
3D Human pose estimation



[VNect: Real-time 3D Human Pose Estimation with a Single RGB Camera, SIGGRAPH 2017]

Regression + 2D&3D skeleton fitting

Surface reconstruction



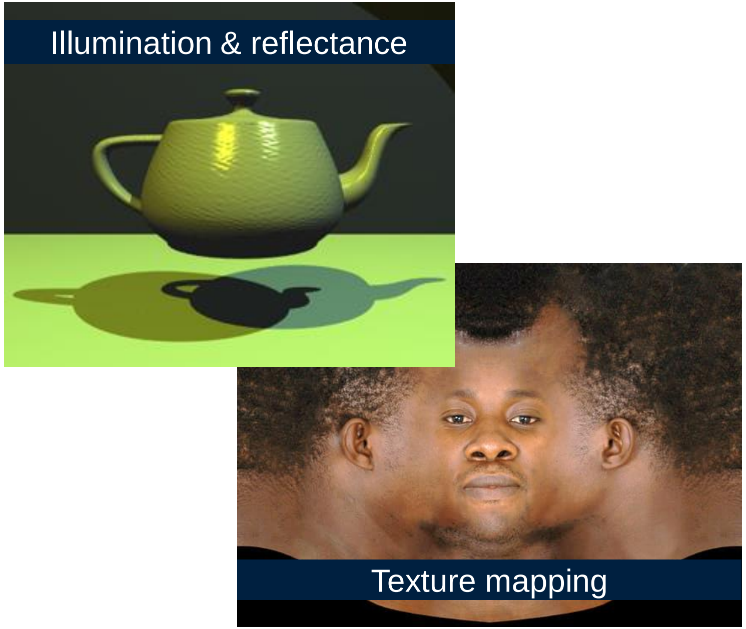
[MonoPerfCap: Human Performance Capture from Monocular Video, TOG 2018]

Computer Graphics Topics

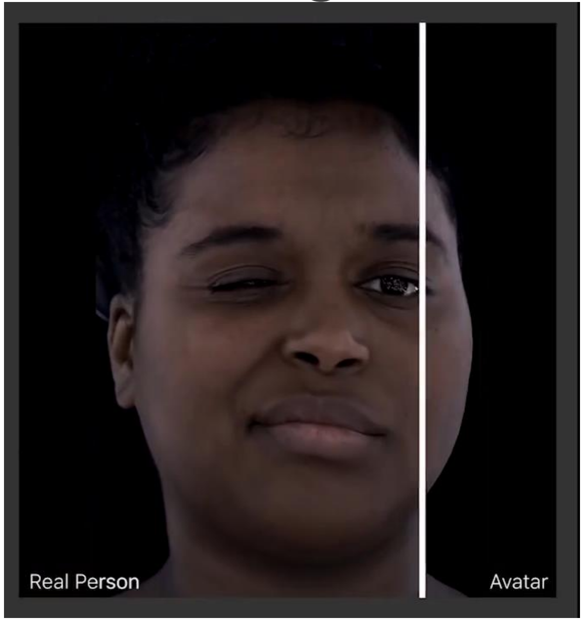
A few examples

Computer Graphics

Classic CG



Learning-based CG



[Facebook AI Research]

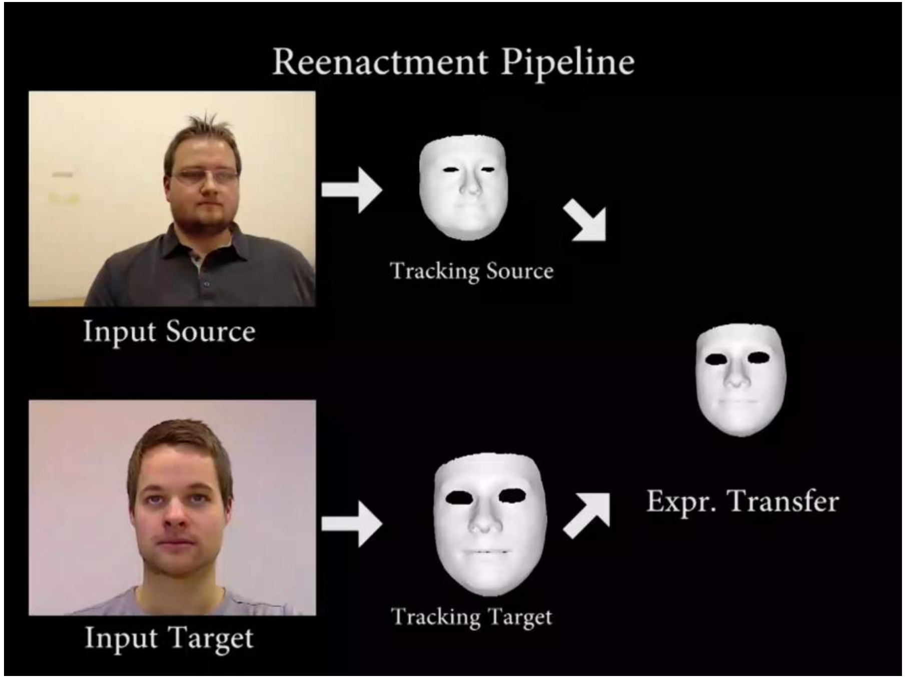
[Utah teapot, www.reallusion.com, www.turbosquid.com]

Facial reenactment

Real-time Facial Reenactment

"THE REMOVAL OF SADDAM HUSSEIN WAS RIGHT DECISION"
Pres. Bush reflects on the Iraq War
LIVE
CNN

Live capture using a commodity webcam

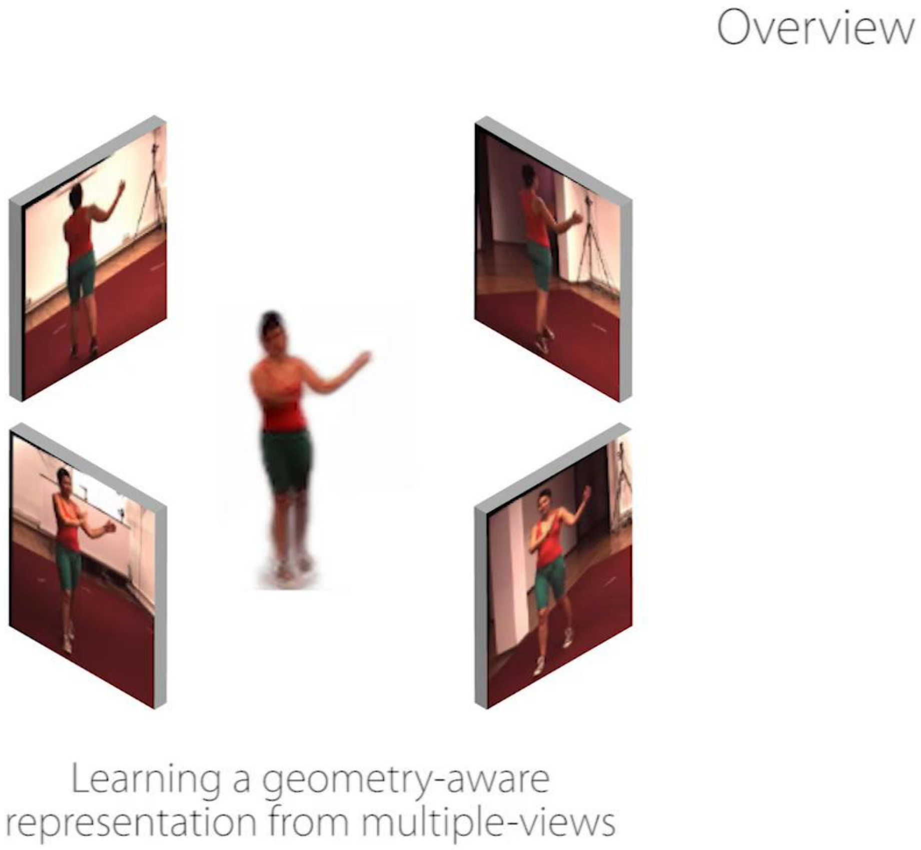


Reenactment Pipeline

[Face2Face: Real-time Face Capture and Reenactment of RGB Videos, CVPR 2016]

Geometry based, no machine learning!

Novel view synthesis

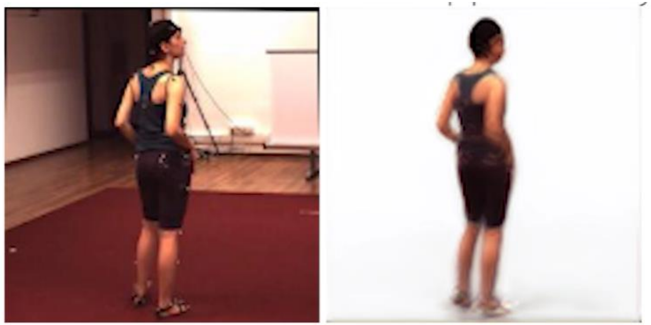


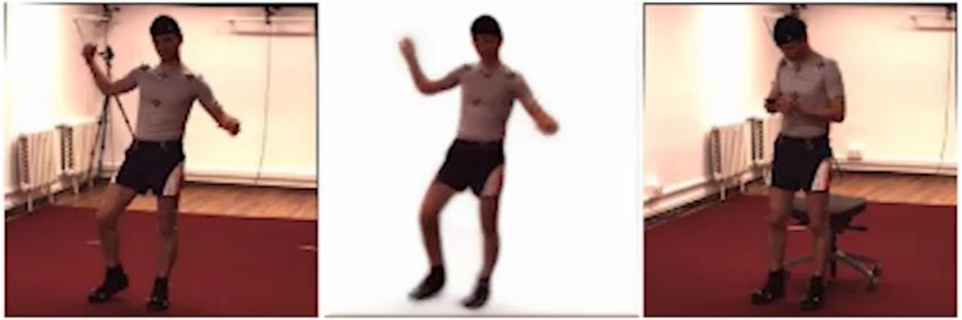
Overview

Learning a geometry-aware
representation from multiple-views

[Unsupervised Geometry-Aware Representation Learning for 3D Human Pose Estimation, ECCV 2018]

Disentangled appearance and pose





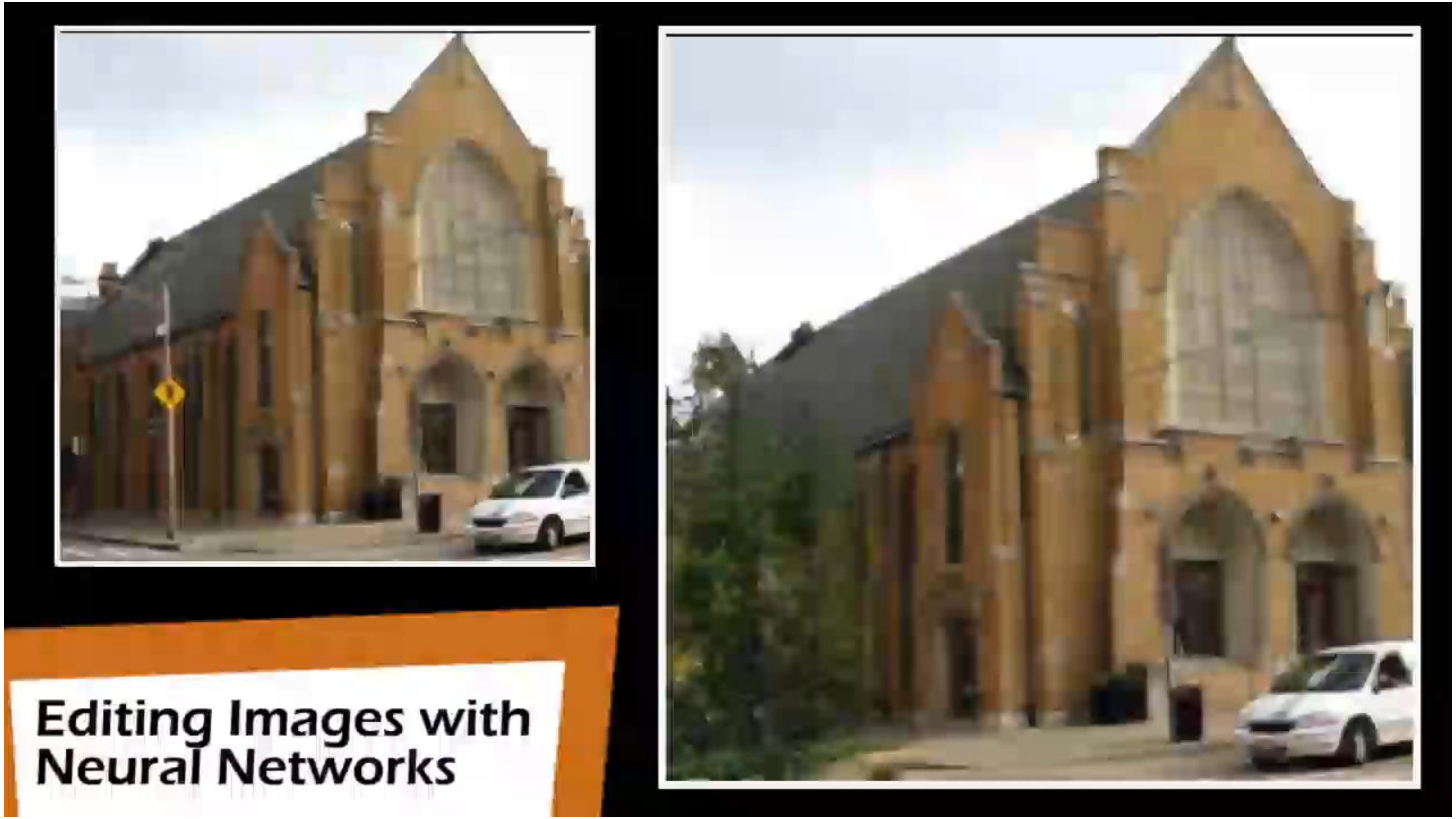
Latent space interpolation

Appearance swap

Image interpolation

GAN Paint – conditional image generation

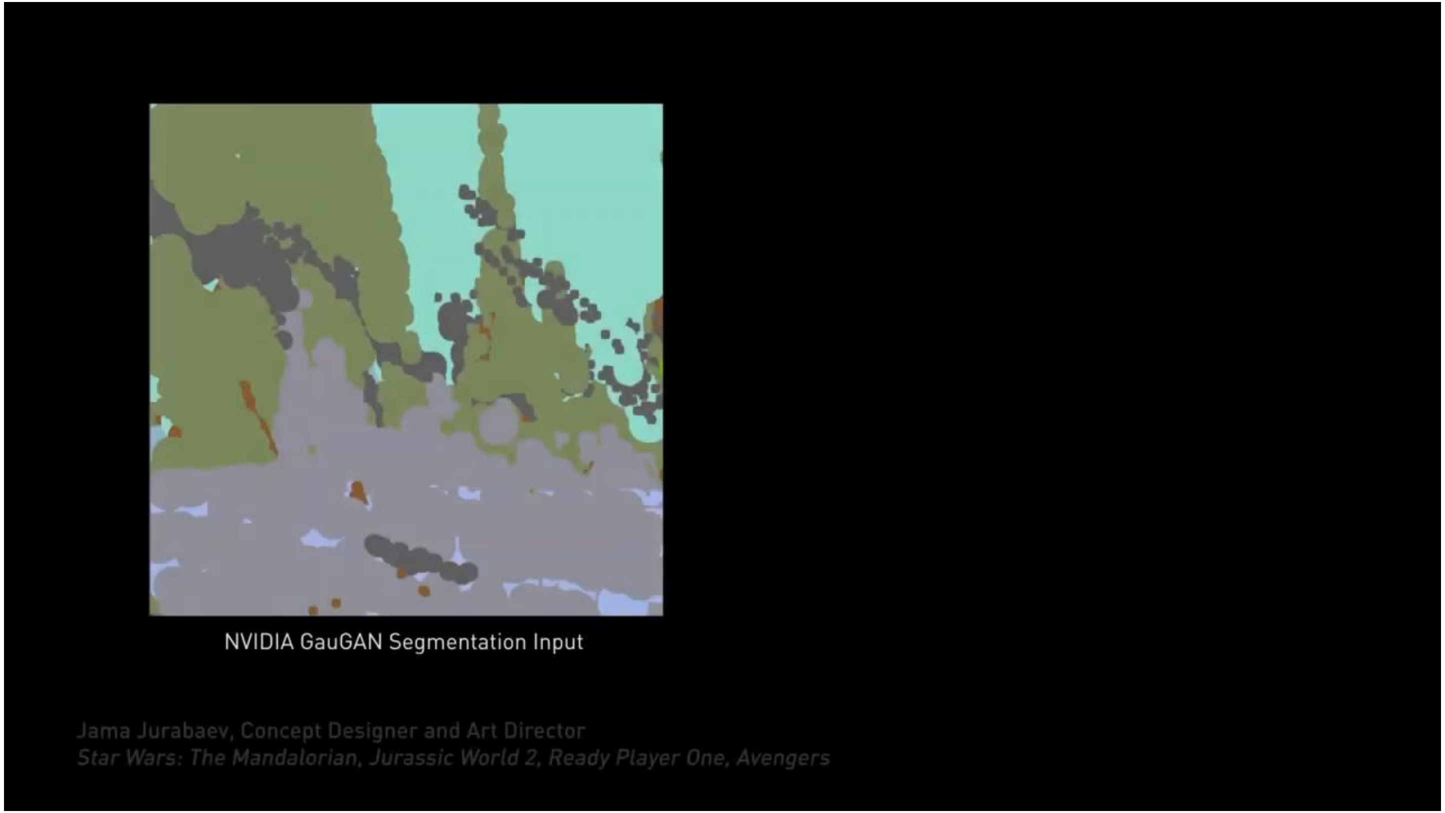
Online demo: http://ganpaint.io/demo/?project=church



[Semantic Photo Manipulation with a Generative Image Prior, SIGGRAPH 2019]

gauGan

Online demo http://nvidia-research-mingyuliu.com/gaugan



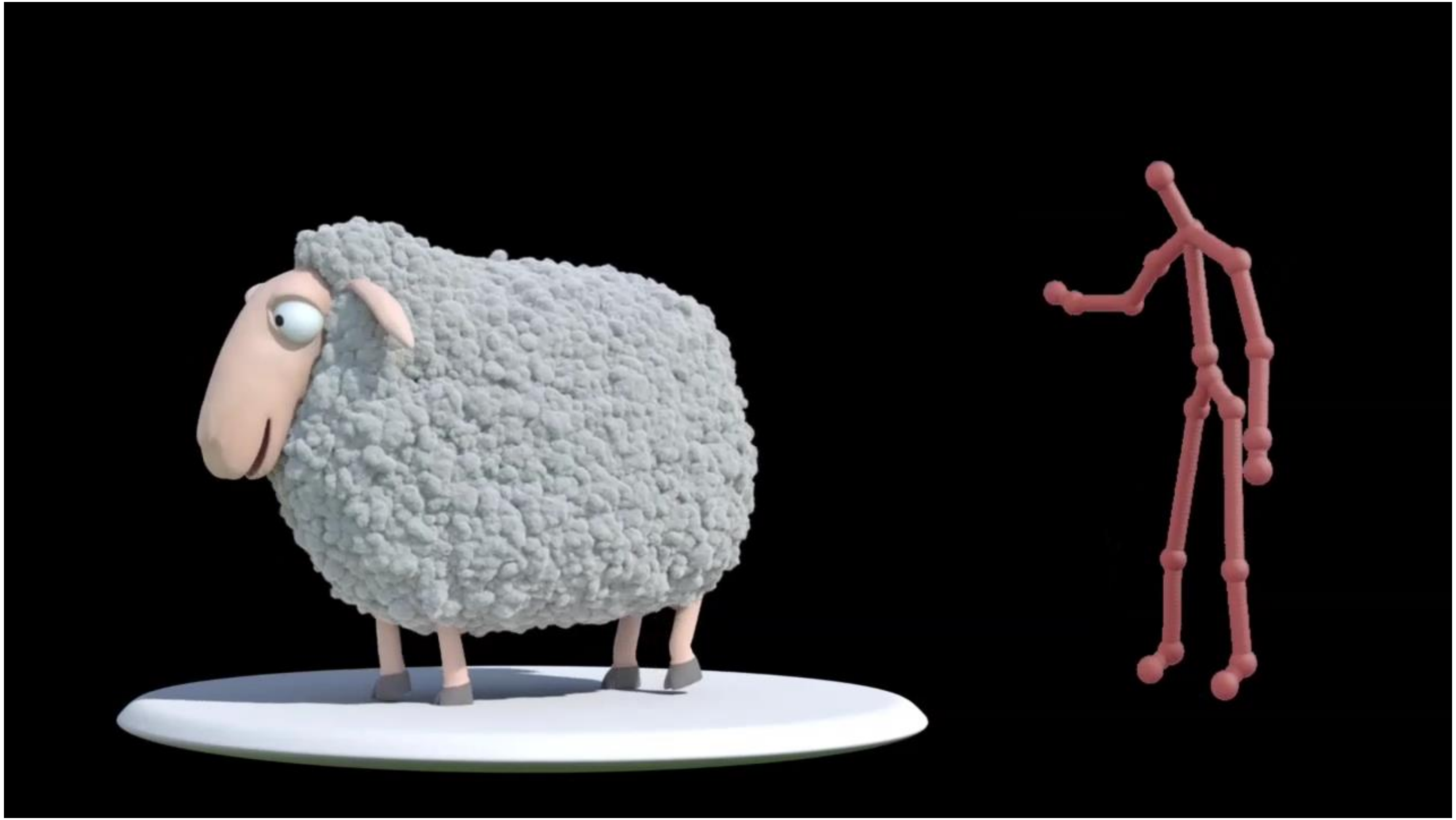
NVIDIA GauGAN Segmentation Input

Jama Jurabaev, Concept Designer and Art Director

Star Wars: The Mandalorian, Jurassic World 2, Ready Player One, Avengers

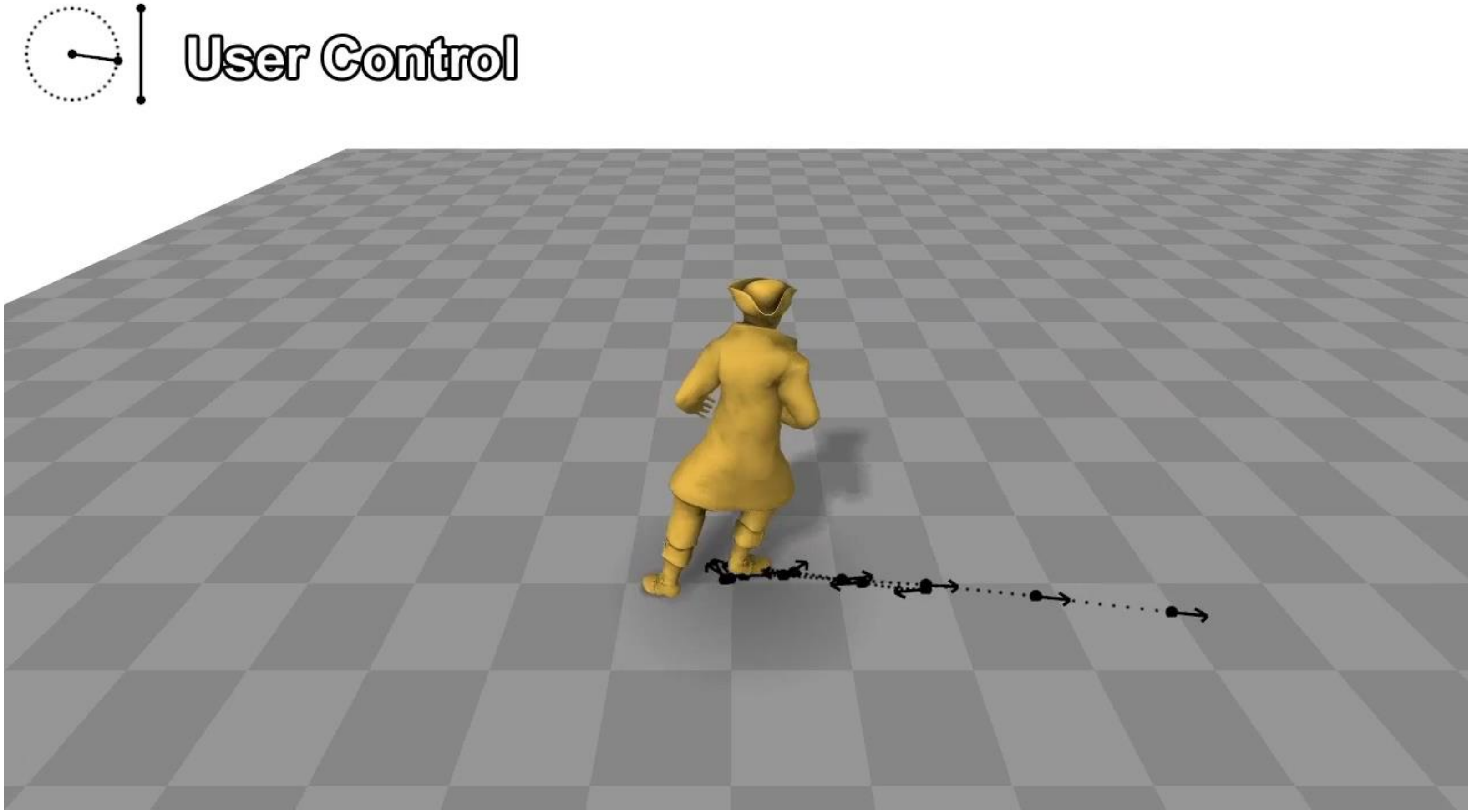
[Semantic Image Synthesis with Spatially-Adaptive Normalization, CVPR 2019]

Puppeteering



[Interactive Motion Mapping for Real-time Character Control, Eurographics 2014]

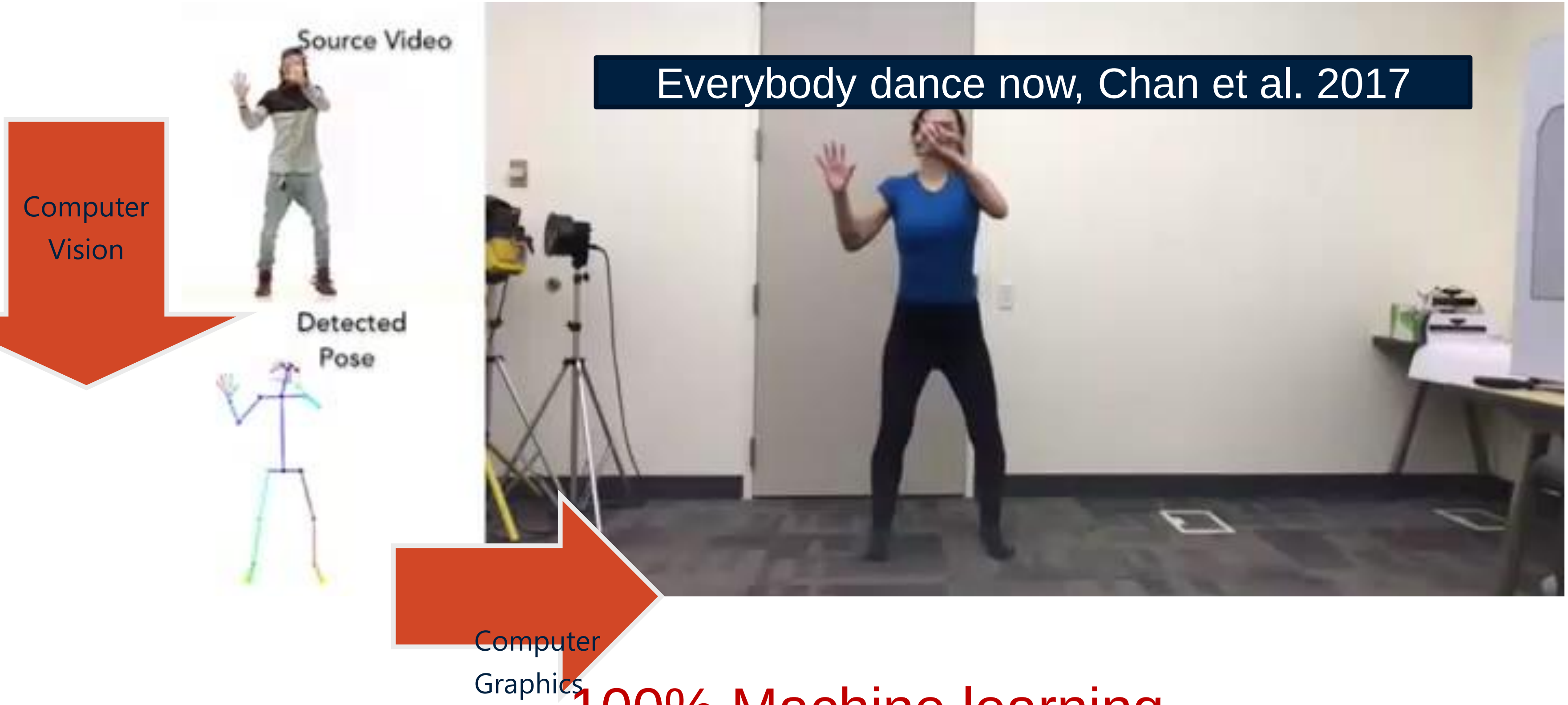
Character animation



[Phase-Functioned Neural Networks for Character Control, SIGGRAPH 2017]

Computer Graphics & Computer Vision Topics

Example topics



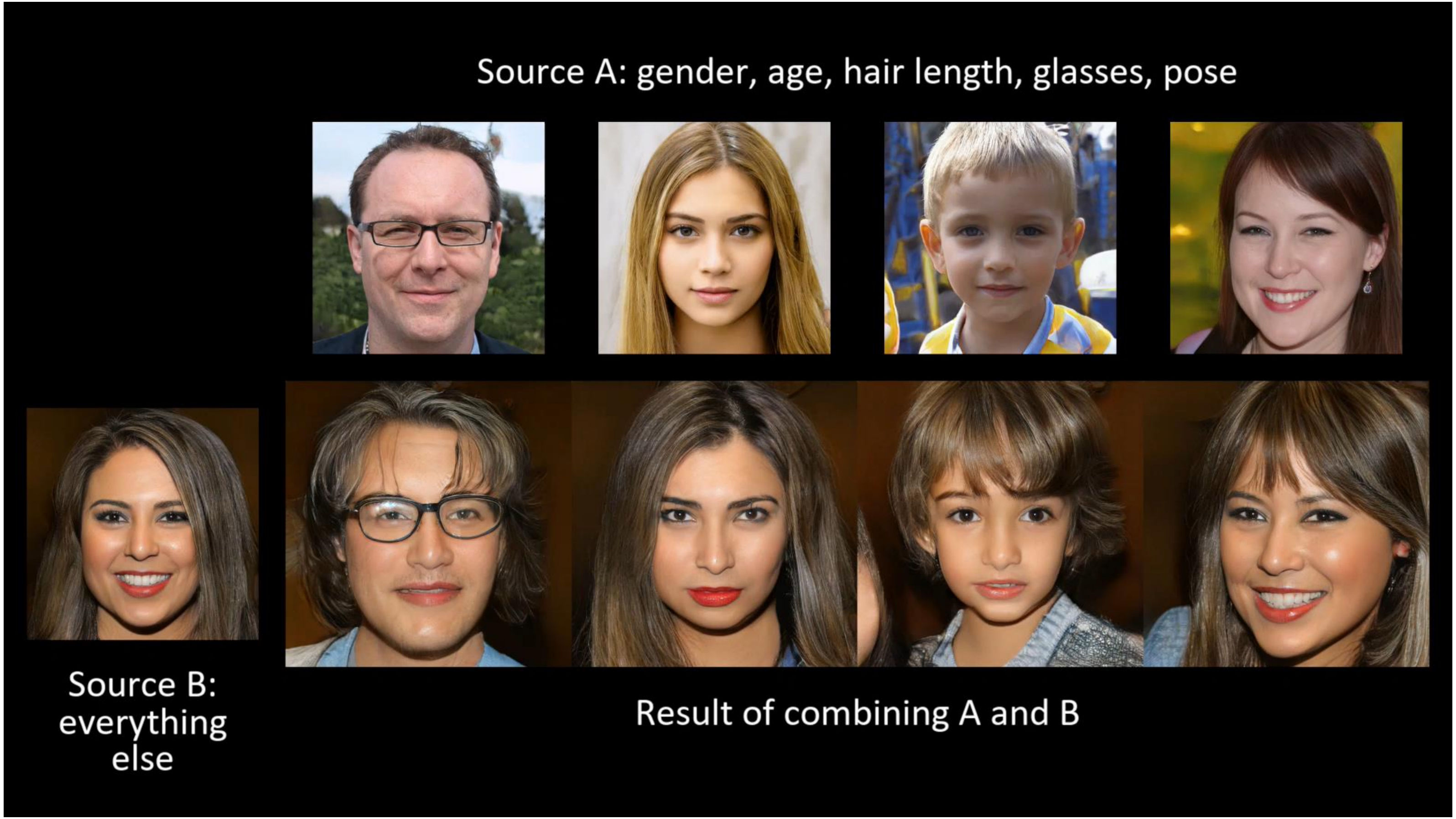
Everybody dance now, Chan et al. 2017

Computer Vision

Computer Graphics

100% Machine learning

StyleGAN – generating images from noise



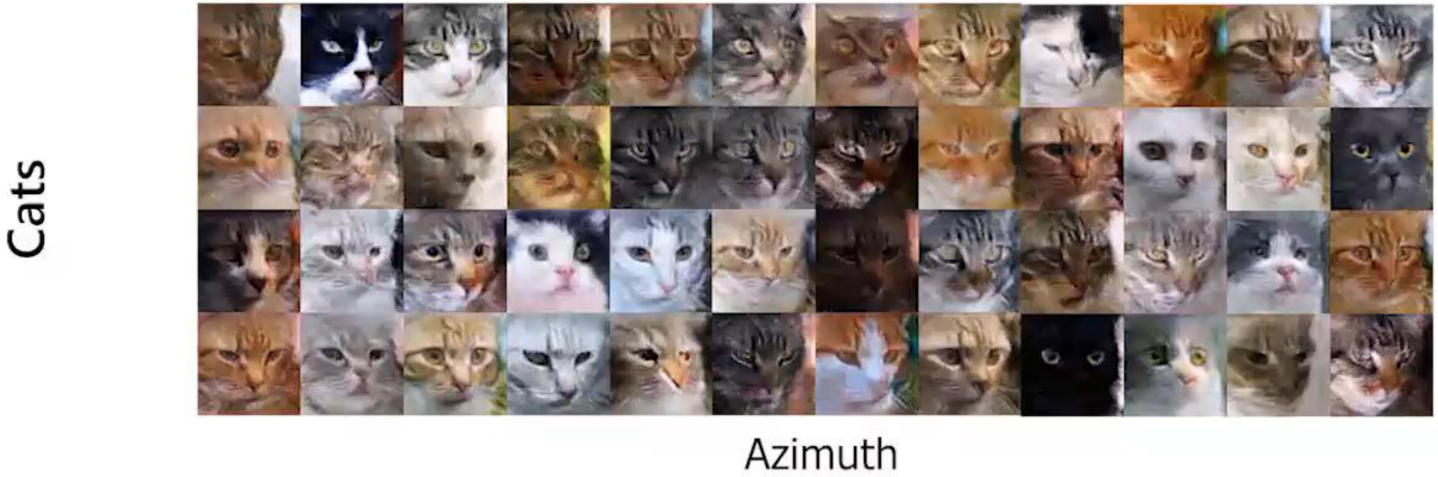
Source A: gender, age, hair length, glasses, pose

Source B: everything else

Result of combining A and B

[StyleGAN: A Style-Based Generator Architecture for Generative Adversarial Networks. CVPR 2019]

HoloGAN – providing viewpoint control



[HoloGAN: Unsupervised learning of 3D representations from natural images, ICCV 2019]

Discussion

- What other topics do you know?

- Which ones are interesting?

- Which ones are hard to solve with machine learning, which ones are easy?

- Black lives matter!

Break

Register on Piazza, get to know your neighbor, discuss

Prime conferences

Graphics

- SIGGRAPH
- SIGGRAPH Asia
- Eurographics (EG)
- Journals: TOG and CFG

Computer vision

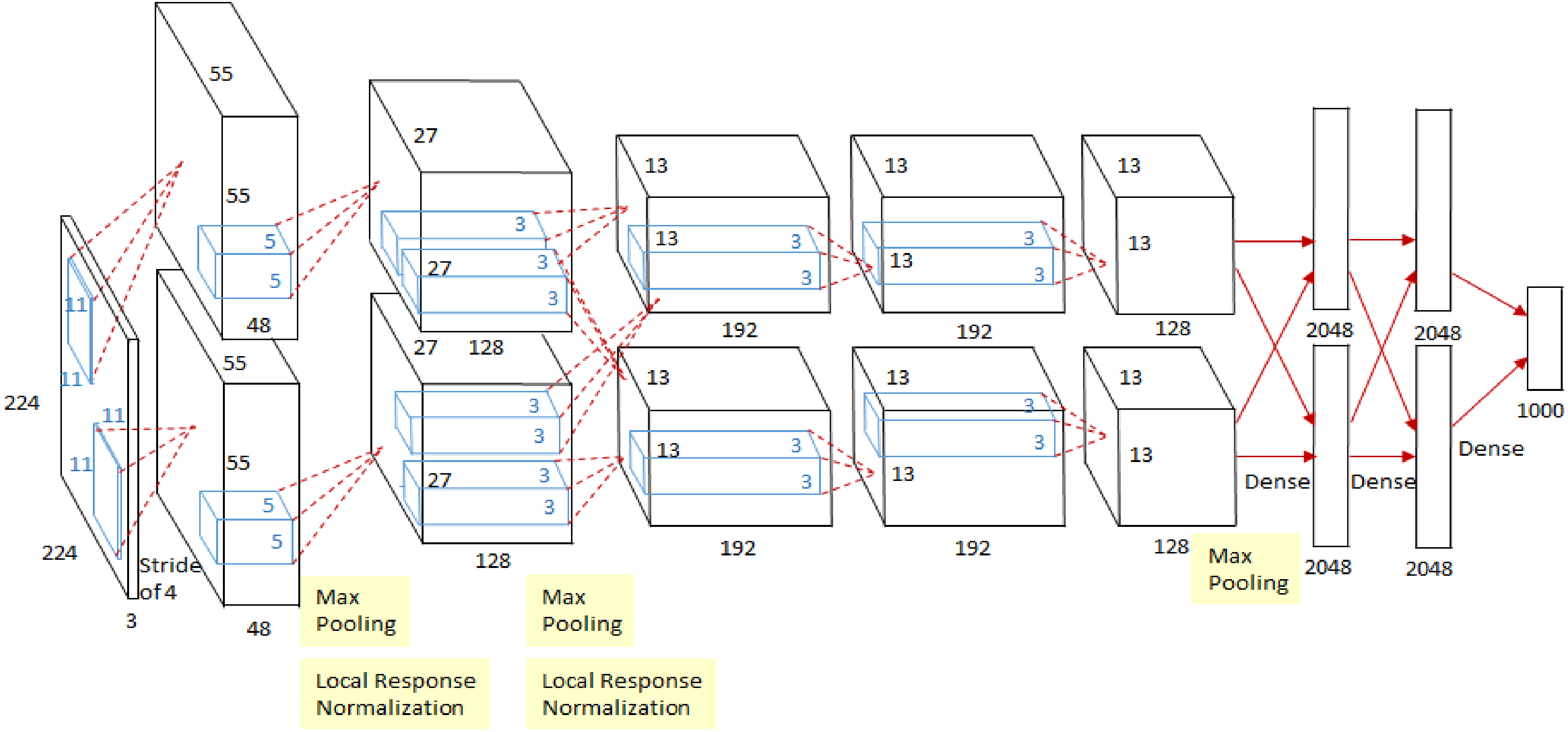
- Conference on Computer Vision and Pattern Recognition (CVPR)
- International Conference on Computer Vision (ICCV)
- European Conference on Computer Vision (ECCV) – every second year, alternates with ICCV
- Journals: IJCV and TPAMI

Machine learning

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

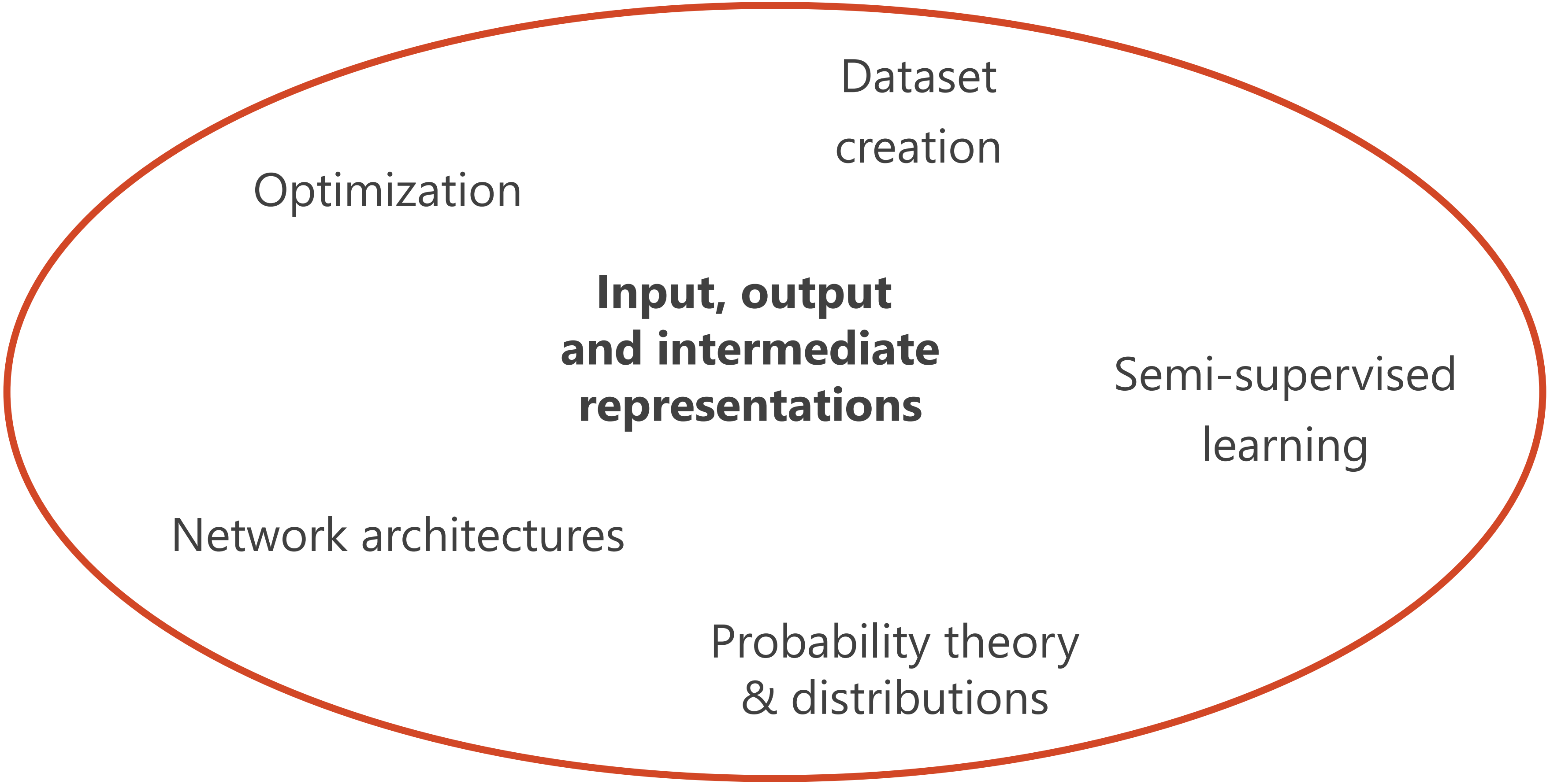
Underlying ML Methods

Deep learning models = architecture?



[Alex Net]

Deep learning models for computer graphics and computer vision



Dataset creation

Optimization

Input, output and intermediate representations

Semi-supervised learning

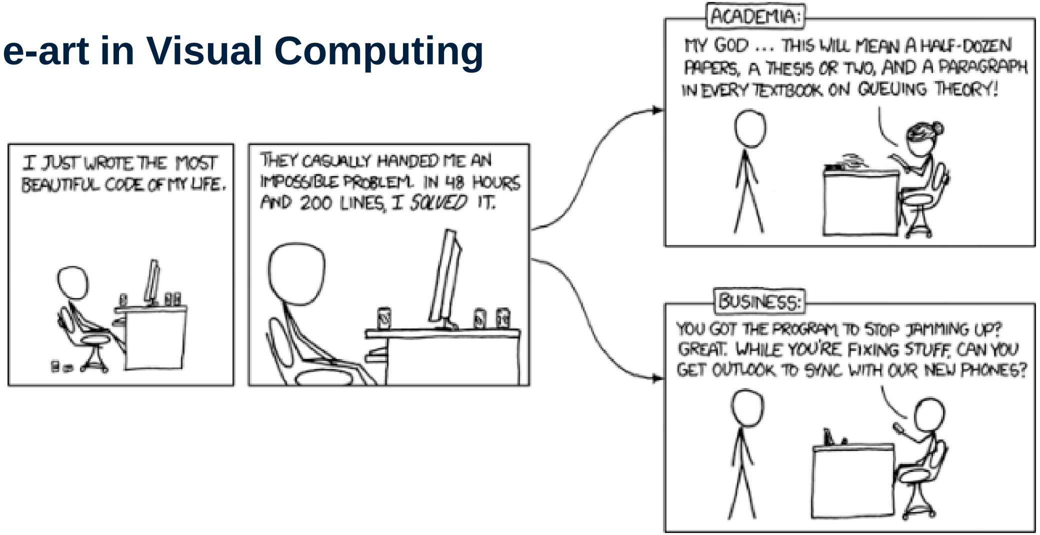
Network architectures

Probability theory & distributions

Logistics

Visual AI - Goals

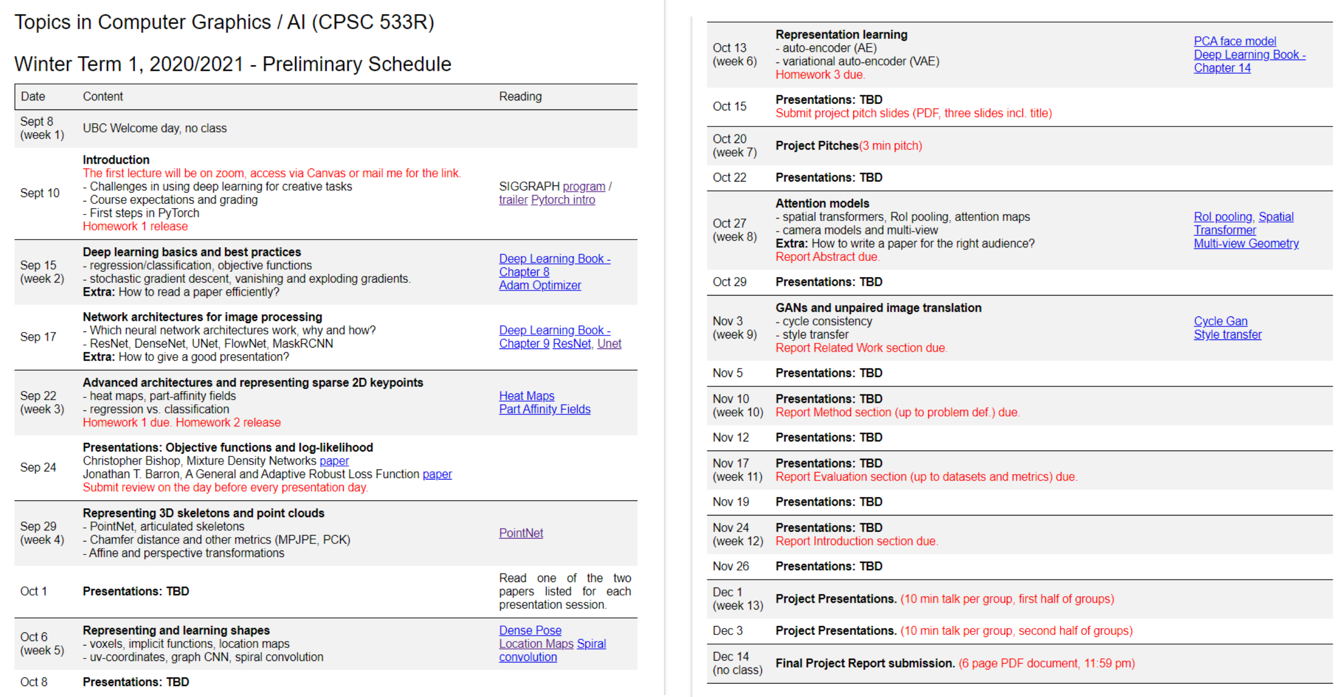
- **Get to know and advance the state-of-the-art in Visual Computing**
  - Computer Graphics
  - Computer Vision
  - towards Visual AI
- Practice machine learning (ML)
  - From design through implementation to evaluation
  - Become a PyTorch and ML expert (PyTorch = deep learning framework)
  - tricks, hacks, gems, best practices, …
- Prepare you for academia (my group?!) and industry
  - **independently** complete a mini research project
  - become a researcher


I JUST WROTE THE MOST BEAUTIFUL CODE OF MY LIFE.
THEY CASUALLY HANDED ME AN IMPOSSIBLE PROBLEM. IN 48 HOURS AND 200 LINES, I SOLVED IT.
ACADEMIA:
MY GOD ... THIS WILL MEAN A HALF-DOZEN PAPERS, A THESIS OR TWO, AND A PARAGRAPH IN EVERY TEXTBOOK ON QUEUING THEORY!
BUSINESS:
YOU GOT THE PROGRAM TO STOP JAMMING UP? GREAT. WHILE YOU'RE FIXING STUFF, CAN YOU GET OUTLOOK TO SYNC WITH OUR NEW PHONES?

https://www.explainxkcd.com/wiki/index.php/664:_Academia_vs._Business

Syllabus

- Talks will be assigned once all students are enrolled
    - Until then: open for paper suggestions
- Choose a project topic before the project pitch (~mid term)



Topics in Computer Graphics / AI (CPSC 533R)

Winter Term 1, 2020/2021 - Preliminary Schedule

Date	Content	Reading
Sept 8 (week 1)	UBC Welcome day, no class	
Sept 10	**Introduction** The first lecture will be on zoom, access via Canvas or mail me for the link. - Challenges in using deep learning for creative tasks - Course expectations and grading - First steps in PyTorch Homework 1 release	SIGGRAPH program / trailer Pytorch intro
Sep 15 (week 2)	**Deep learning basics and best practices** - regression/classification, objective functions - stochastic gradient descent, vanishing and exploding gradients. **Extra:** How to read a paper efficiently?	Deep Learning Book - Chapter 8 Adam Optimizer
Sep 17	**Network architectures for image processing** - Which neural network architectures work, why and how? - ResNet, DenseNet, UNet, FlowNet, MaskRCNN **Extra:** How to give a good presentation?	Deep Learning Book - Chapter 9 ResNet, Unet
Sep 22 (week 3)	**Advanced architectures and representing sparse 2D keypoints** - heat maps, part-affinity fields - regression vs. classification Homework 1 due. Homework 2 release	Heat Maps Part Affinity Fields
Sep 24	**Presentations: Objective functions and log-likelihood** Christopher Bishop, Mixture Density Networks paper Jonathan T. Barron, A General and Adaptive Robust Loss Function paper Submit review on the day before every presentation day.	
Sep 29 (week 4)	**Representing 3D skeletons and point clouds** - PointNet, articulated skeletons - Chamfer distance and other metrics (MPJPE, PCK) - Affine and perspective transformations	PointNet
Oct 1	**Presentations: TBD**	Read one of the two papers listed for each presentation session.
Oct 6 (week 5)	**Representing and learning shapes** - voxels, implicit functions, location maps - uv-coordinates, graph CNN, spiral convolution	Dense Pose Location Maps Spiral convolution
Oct 8	**Presentations: TBD**	
Oct 13 (week 6)	**Representation learning** - auto-encoder (AE) - variational auto-encoder (VAE) Homework 3 due.	PCA face model Deep Learning Book - Chapter 14
Oct 15	**Presentations: TBD** Submit project pitch slides (PDF, three slides incl. title)	
Oct 20 (week 7)	**Project Pitches**(3 min pitch)	
Oct 22	**Presentations: TBD**	
Oct 27 (week 8)	**Attention models** - spatial transformers, RoI pooling, attention maps - camera models and multi-view **Extra:** How to write a paper for the right audience? Report Abstract due.	RoI pooling, Spatial Transformer Multi-view Geometry
Oct 29	**Presentations: TBD**	
Nov 3 (week 9)	**GANs and unpaired image translation** - cycle consistency - style transfer Report Related Work section due.	Cycle Gan Style transfer
Nov 5	**Presentations: TBD**	
Nov 10 (week 10)	**Presentations: TBD** Report Method section (up to problem def.) due.	
Nov 12	**Presentations: TBD**	
Nov 17 (week 11)	**Presentations: TBD** Report Evaluation section (up to datasets and metrics) due.	
Nov 19	**Presentations: TBD**	
Nov 24 (week 12)	**Presentations: TBD** Report Introduction section due.	
Nov 26	**Presentations: TBD**	
Dec 1 (week 13)	**Project Presentations.** (10 min talk per group, first half of groups)	
Dec 3	**Project Presentations.** (10 min talk per group, second half of groups)	
Dec 14 (no class)	**Final Project Report submission.** (6 page PDF document, 11:59 pm)	

Assignments

Assignment 1

Playing with PyTorch

(supervised learning)

Assignment 2:

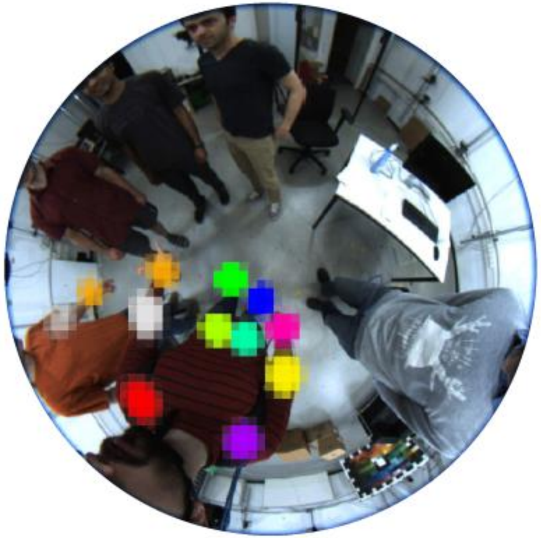
Egocentric human pose estimation

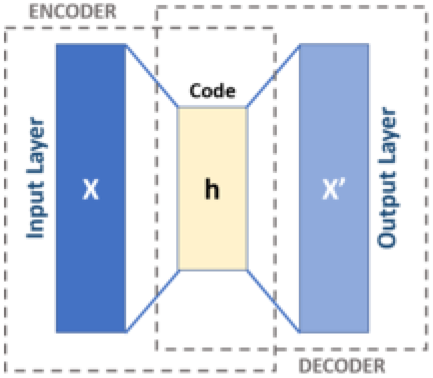
(2D human pose regression)

Assignment 3:

Shape generation

(unsupervised learning)





Assignments - rules

Academic Integrity. Assignments must be solved in teams of two using the available course material and other online sources. You are neither allowed to copy nor look at parts of any of the assignments from anyone outside of your team. Accordingly, it is not allowed to post solutions online or distribute them in (private) forums. The university policies on academic integrity are rigorously applied.

Submission. Solutions must be handed in through the Canvas system and be kept private.

Deadline and grading. Assignments will be due on the dates specified in the schedule, always at 11:59 pm PST. A late submission by one day will still be accepted but reduces the score by -25%. The grading is based on correctness and completeness, as detailed in each exercise description.

Deep Learning with PyTorch, first steps

Deep learning – a new way of programming

Classical programming

- Write down computational rules

  c = a + b

- Requires human programmer

  (domain expert + CS skills)

https://futurism.com/2-whats-the-next-blue-collar-job-coding

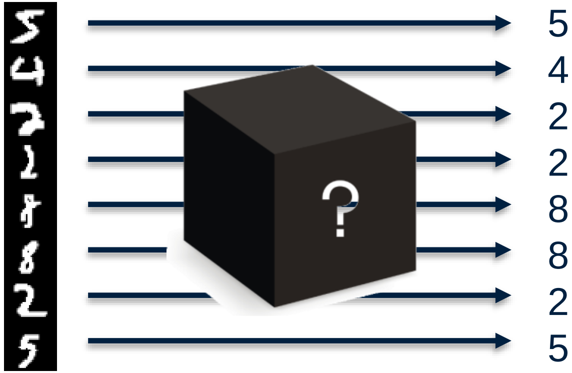
Data driven approach

- Give **lots of** input-output examples

  [(3,4) -> 7, (2,3) -> 5, (100,2) -> 102,

  (2,2) -> 4, (4,3) -> 7, …]

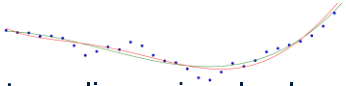
- Requires human annotator (domain expert)
- or Artificial Intelligence (AI) ?
  - Weak supervision, Self-supervision
  - Reinforcement learning …



Deep learning – its curve fitting

Parametric curves

- Polynomial

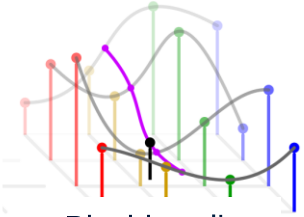


Low-dimensional poly.

$$f(x) = \theta_0 + \theta_1 x + \theta_2 x^2$$

- Spline

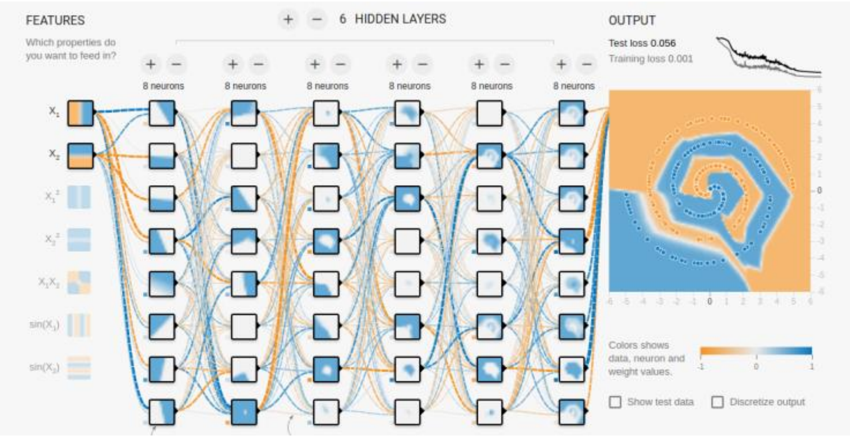
$$f(x) = \begin{cases} f_1(x), & \text{if } x_1 < x \leq x_2 \\ f_2(x), & \text{if } x_2 < x \leq x_3 \\ \vdots & \vdots \\ f_n(x), & \text{if } x_n < x \leq x_{n+1} \end{cases}$$



Bicubic spline

- Neural network

$$f(x) = h(\text{linear}(h(\text{linear}(x, W^{(1)})), W^{(2)}))$$



Multilayer perceptron (fully connected network)
interactive: playground.tensorflow.org

Goal: Find θ that minimizes the objective function on the dataset D

$$\arg\min_{\theta} E(D, \theta)$$

Programming environment - python™

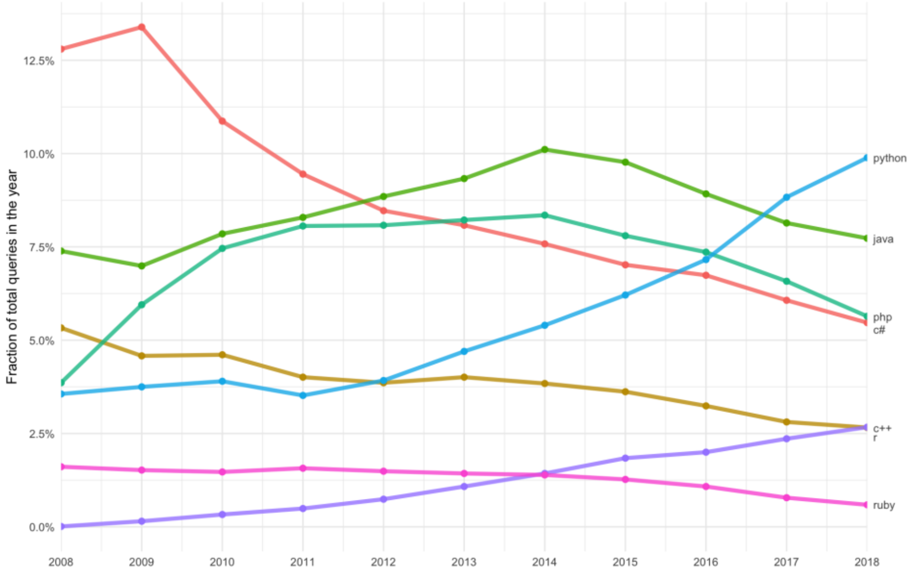
Advantages

- high productivity / quick prototyping
- extensive support libraries
- high performance

  (with libraries linked to programs compiled from FORTRAN, C++, Cuda, …)

- we will use python 3!

Questions per year in Stack Overflow



https://towardsdatascience.com/predicting-the-future-popularity-of-programming-languages-4f28c80bd36f

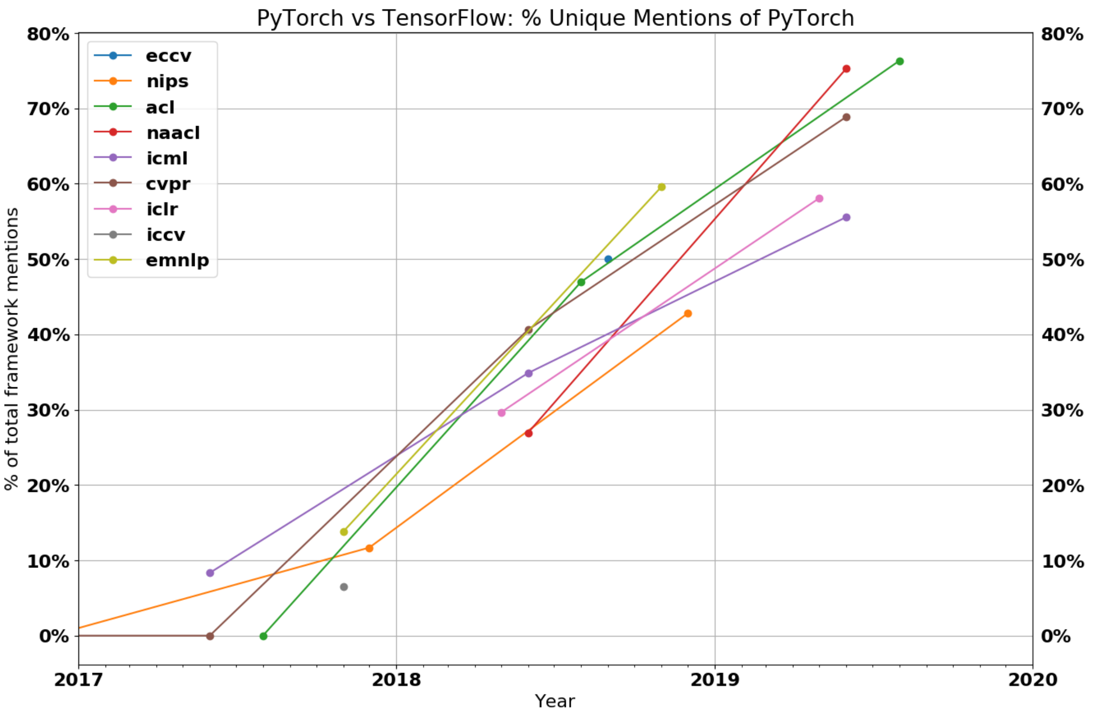
Machine learning framework - PYTORCH

Features

- efficient matrix and tensor operations (like NumPy)
- automatic differentiation (dynamic)
- large number of tutorials
- many open source repositories

Resources

- PyTorch tutorials https://pytorch.org/tutorials/
- PyTorch introduction https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html

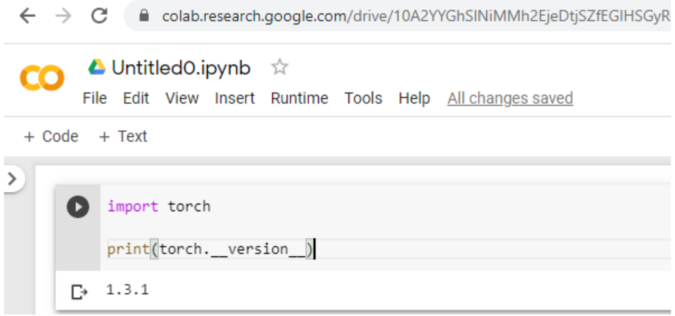


https://thegradient.pub/state-of-ml-frameworks-2019-pytorch-dominates-research-tensorflow-dominates-industry/

Google colab

Cloud computing

- http://colab.research.google.com
- Provides a Jupyter notebook
- Incredible easy to setup
- Provides GPU access (for some time)
- Free of charge
- Interfaces with google drive


colab.research.google.com/drive/10A2YYGhSINiMMh2EjeDtjSZfEGIHSGyR
Untitled0.ipynb
File Edit View Insert Runtime Tools Help All changes saved
+ Code + Text

```
import torch
print(torch.__version__)
```

```
1.3.1
```

Assignment I

“Playing with PyTorch”

- Network architecture
- Dataloaders
- Evaluation
- Visualization

- Posted on course website
- Submit solution on Canvas

- Work in randomly assigned teams