Introduction to Deep Learning

The Team

Lecturer

Prof. Angela Dai

PhDs

Manuel Dahnert

Yujin Chen

Junwen Huang

Student Teaching Assistants

Dan Halperin

Wei Cao

Haoxuan Li

Erik Traise

What is Computer Vision?

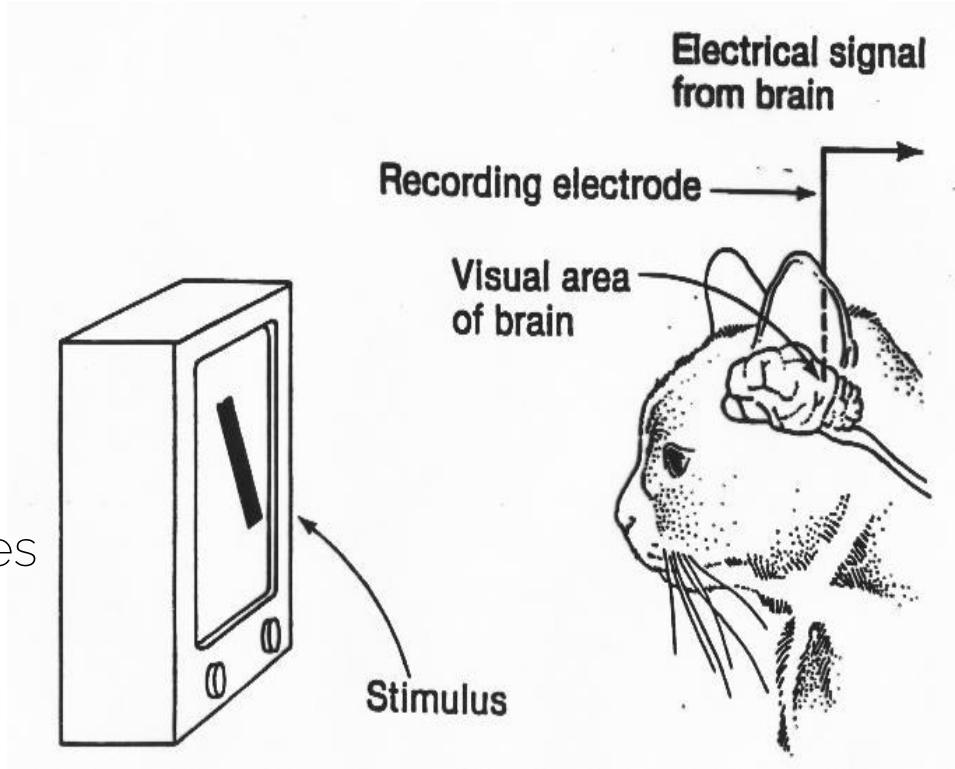
- First defined in the 60s in artificial intelligence groups
- “Mimic the human visual system”
- Center block of robotic intelligence

Hubel and Wiesel

- David Hubel and Torsten Wiesel were neurobiologists from Harvard Medical School
- Experiment revealed several secrets of the human vision system
- Won 2 Nobel prizes

Hubel and Wiesel Experiment

- Recorded electrical activity from individual neurons in the brains of cats.
- Slide projector to show specific patterns to the cats noted specific patterns stimulated activity in specific parts of the brain.
- Results: Visual cortex cells are sensitive to the orientation of edges but insensitive to their position



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

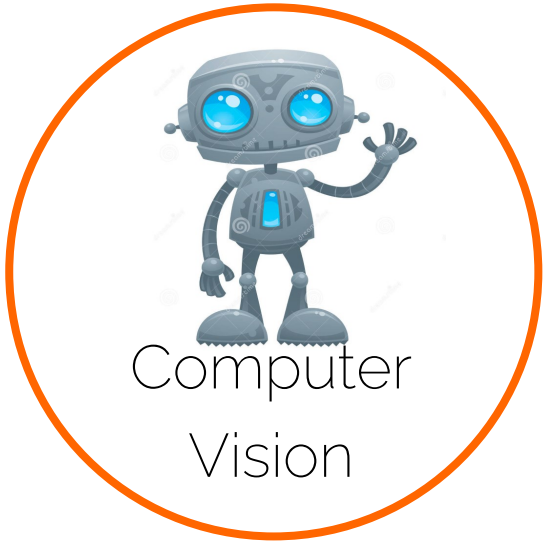
July 7, 1966

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".

A Few Decades Later…



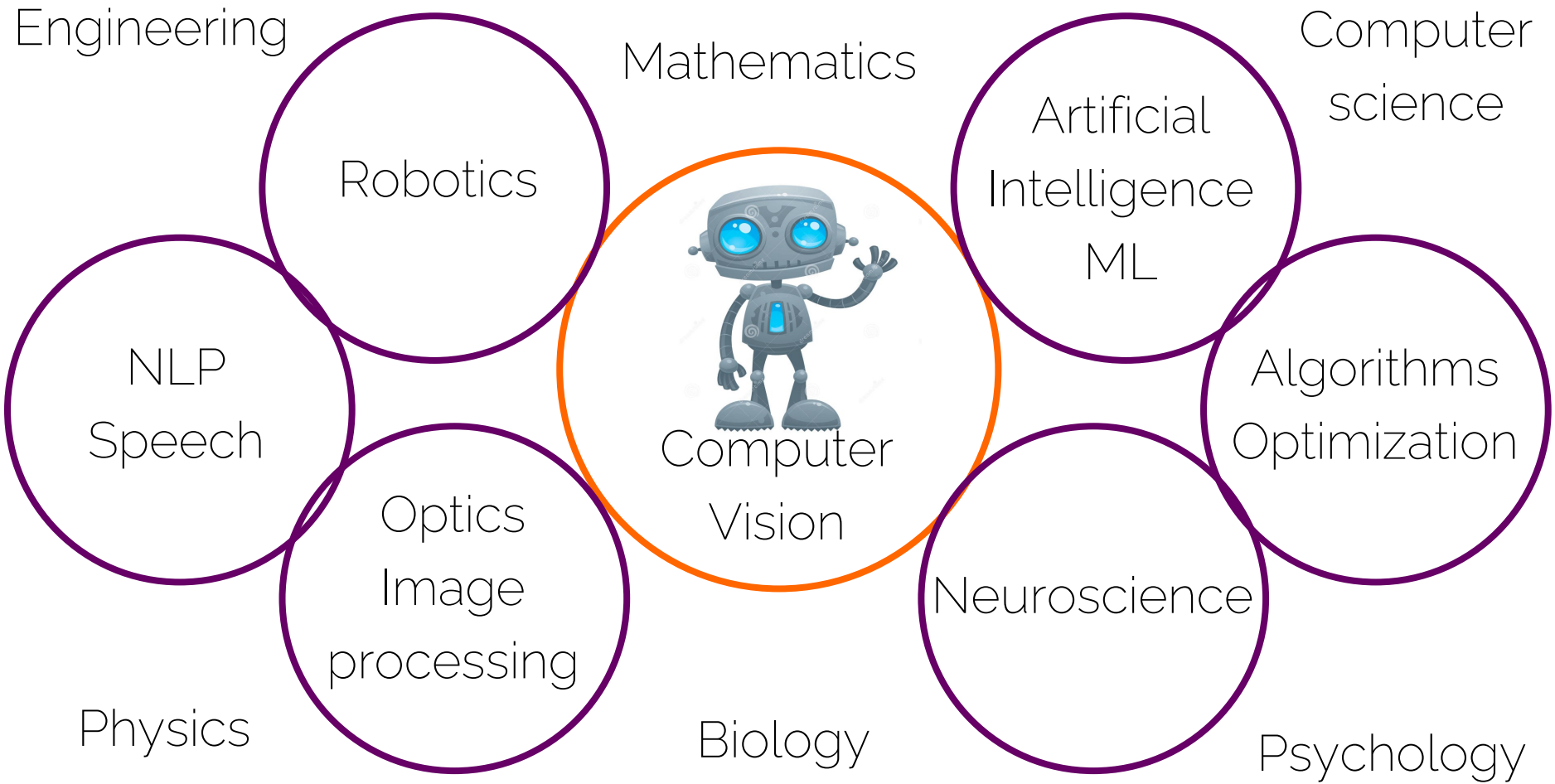
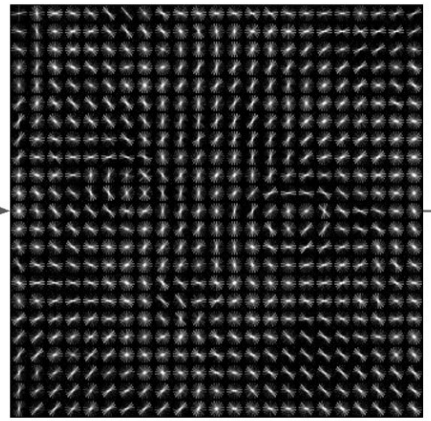

Engineering
Mathematics
Computer science
Robotics
Artificial Intelligence ML
NLP Speech
Computer Vision
Algorithms Optimization
Optics Image processing
Neuroscience
Physics
Biology
Psychology

Image Classification

Input image

Preprocessing



Features : HAAR, HOG, SIFT, SURF

Learning Algorithm

SVM, Random Forests, ANN

Label Assignment

Cat or Background

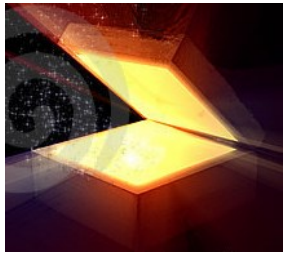
Image Classification

Input image

Awesome magic box

Label Assignment

Cat or Background



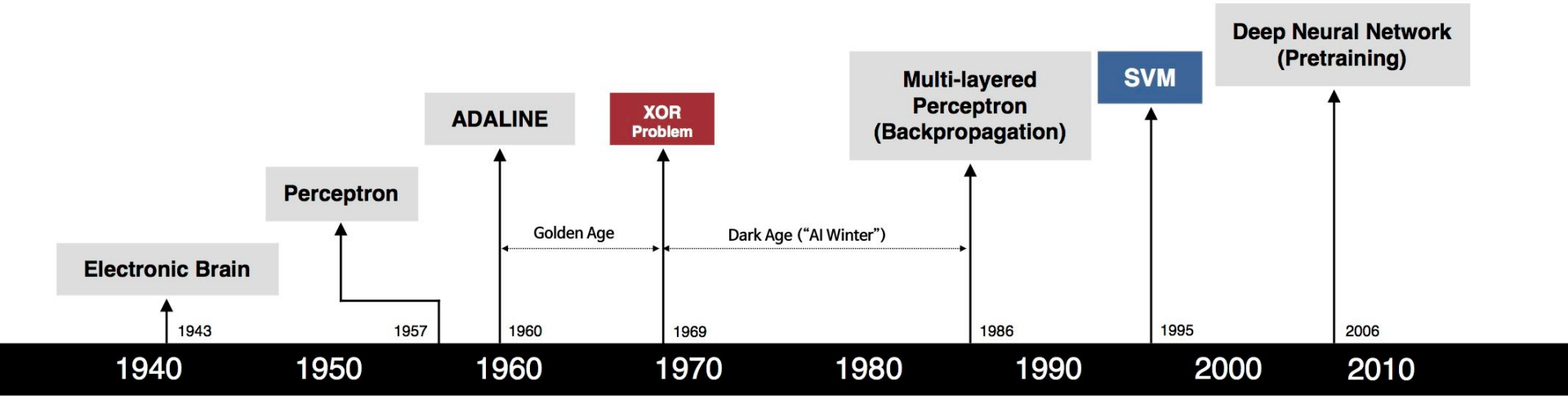
Open the box

Become magicians

Post 2012

Why Deep Learning?

Deep Learning History



S. McCulloch – W. Pitts

F. Rosenblatt

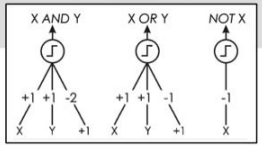
B. Widrow – M. Hoff

M. Minsky – S. Papert

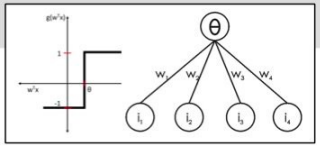
D. Rumelhart – G. Hinton – R. Wiliams

V. Vapnik – C. Cortes

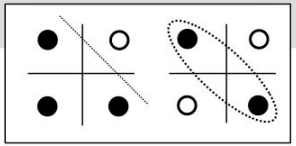
G. Hinton – S. Ruslan



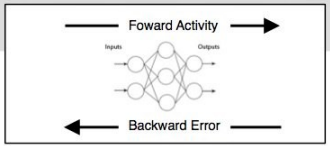
- Adjustable Weights
- Weights are not Learned



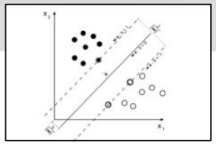
- Learnable Weights and Threshold



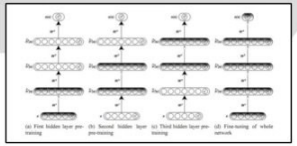
- XOR Problem



- Solution to nonlinearly separable problems
- Big computation, local optima and overfitting

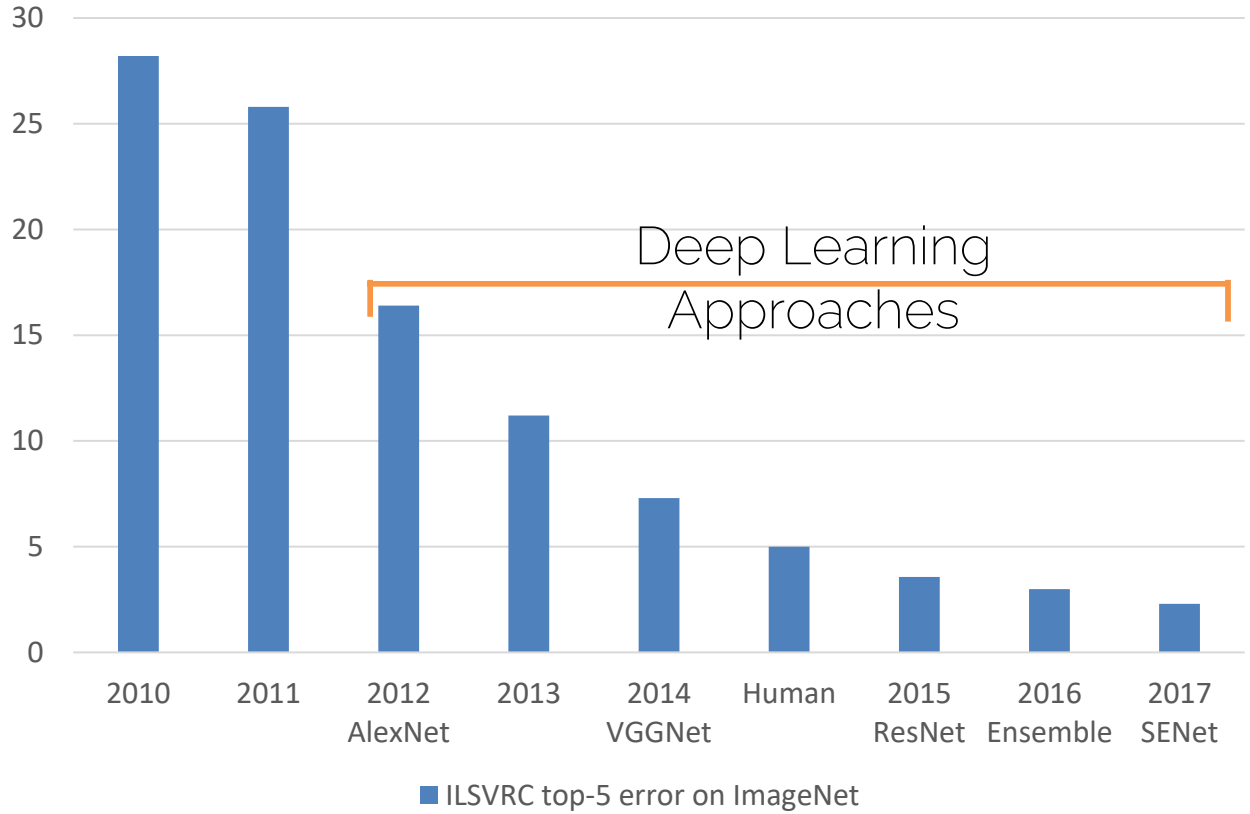


- Limitations of learning prior knowledge
- Kernel function: Human Intervention



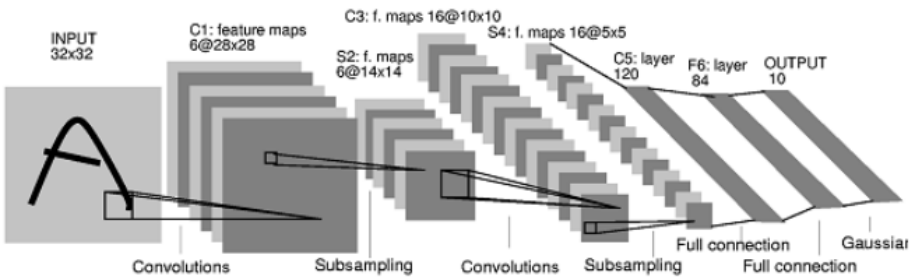
- Hierarchical feature Learning

The Empire Strikes Back



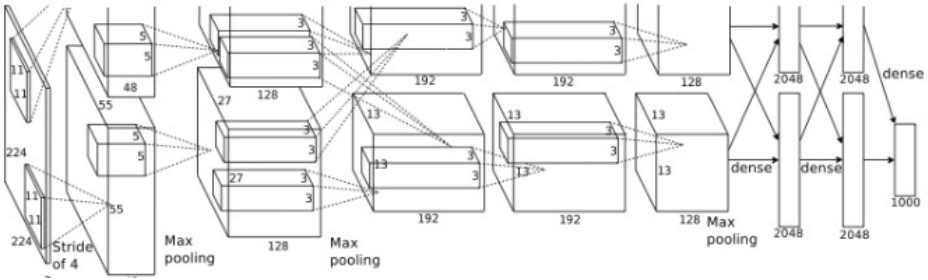
What Has Changed?

1998
LeCun
et al.



- MNIST digit recognition dataset
- 10^7 pixels used in training

2012
Krizhevsky
et al.

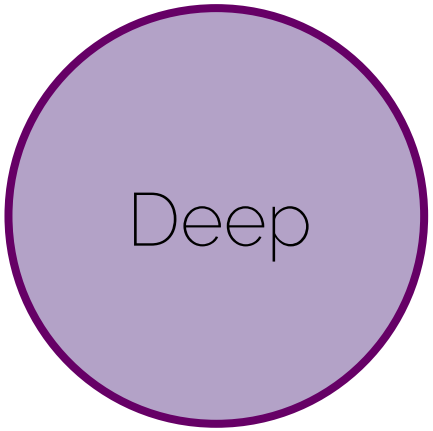


- ImageNet image recognition dataset
- 10^{14} pixels used in training

What Made this Possible?

Models know where to learn from

Models are trainable

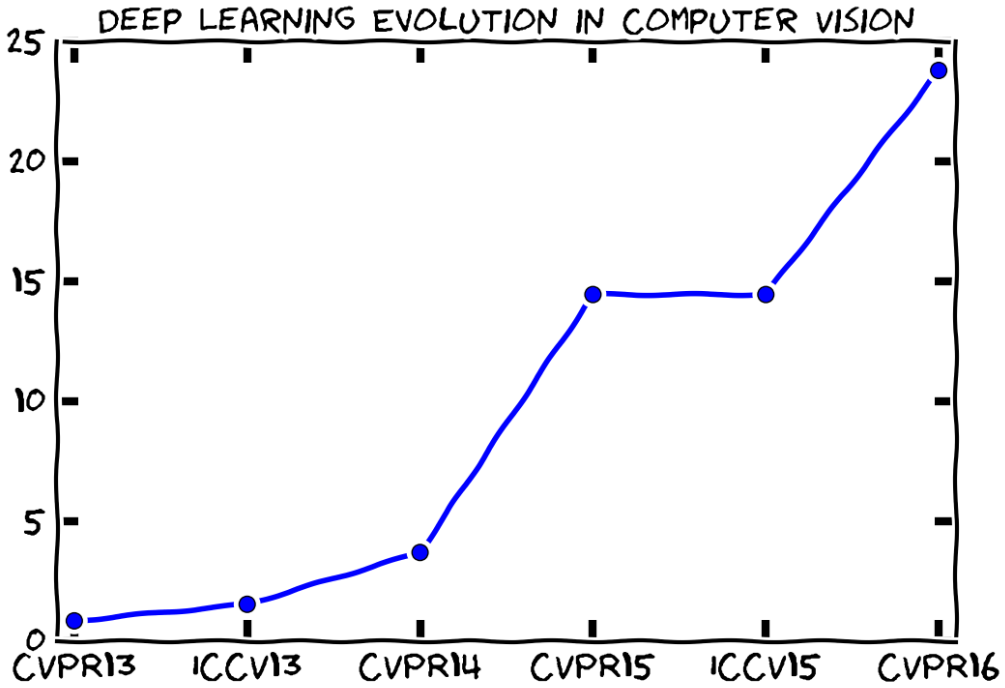


Models are complex

Deep Learning Recognition

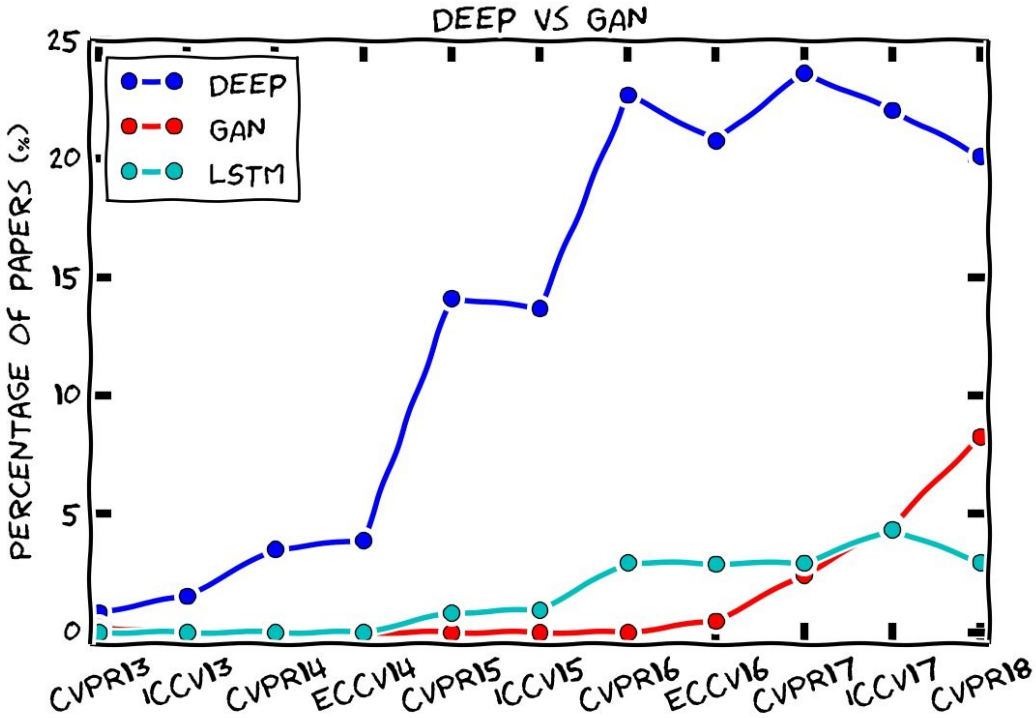
ACM Turing Award 2019 (Nobel Prize of Computing)
Yann LeCun, Geoffrey Hinton, and Yoshua Bengio

Deep Learning and Computer Vision



Credits: Dr. Pont-Tuset, ETH Zurich

Deep Learning and Computer Vision

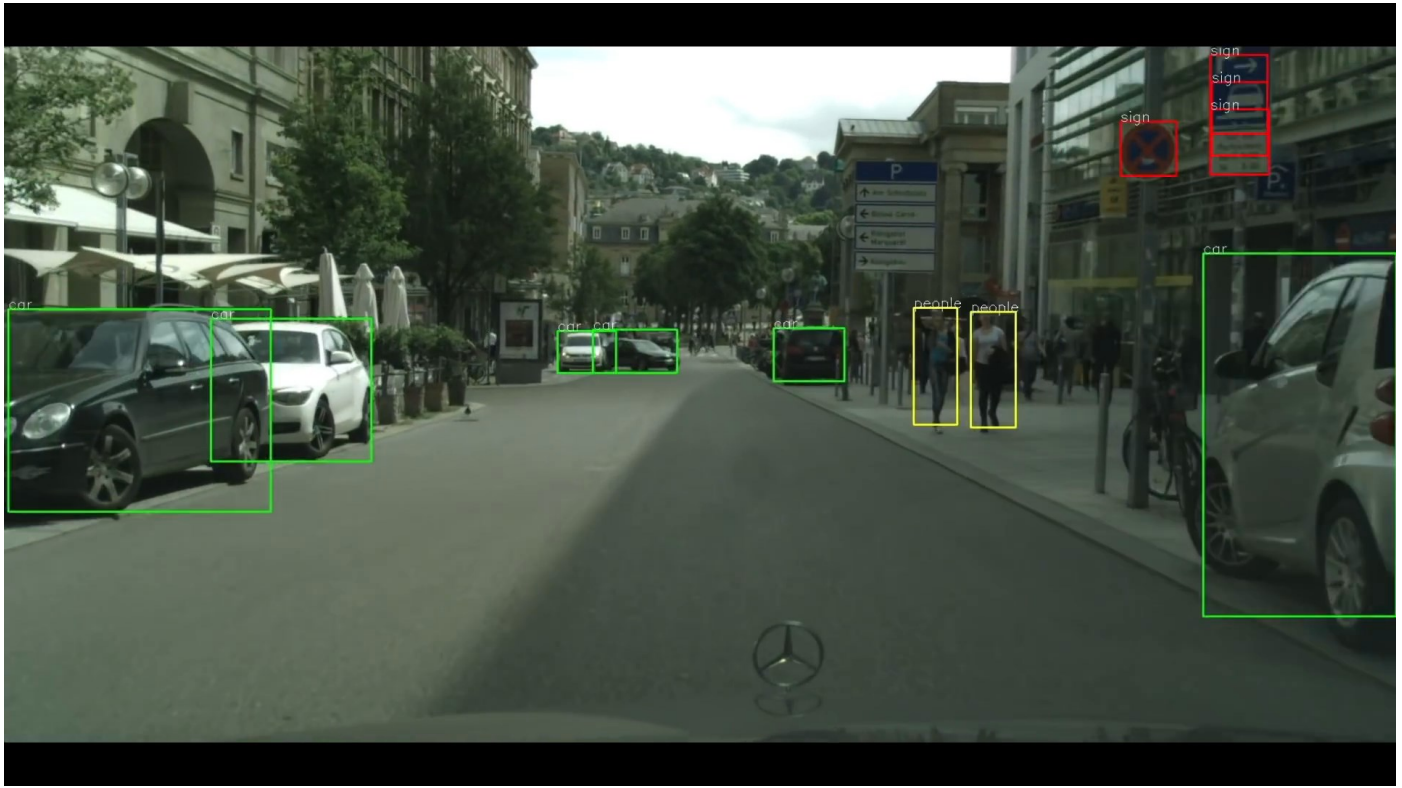


Credits: Dr. Pont-Tuset, ETH Zurich

Deep Learning Today

Object Detection

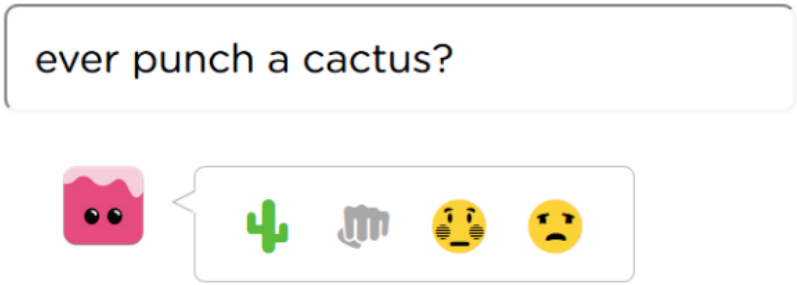
Deep Learning Today



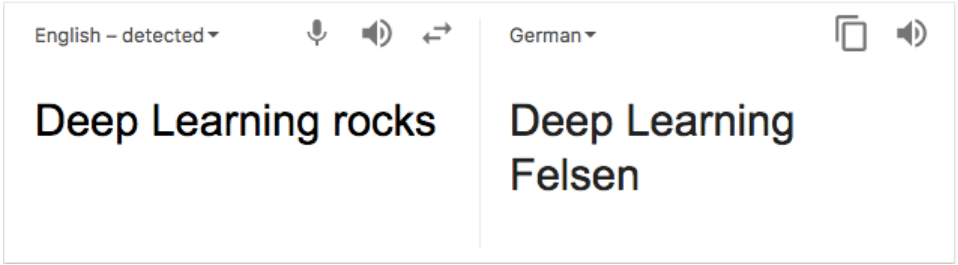
Self-driving cars

Deep Learning Today

AlphaGo

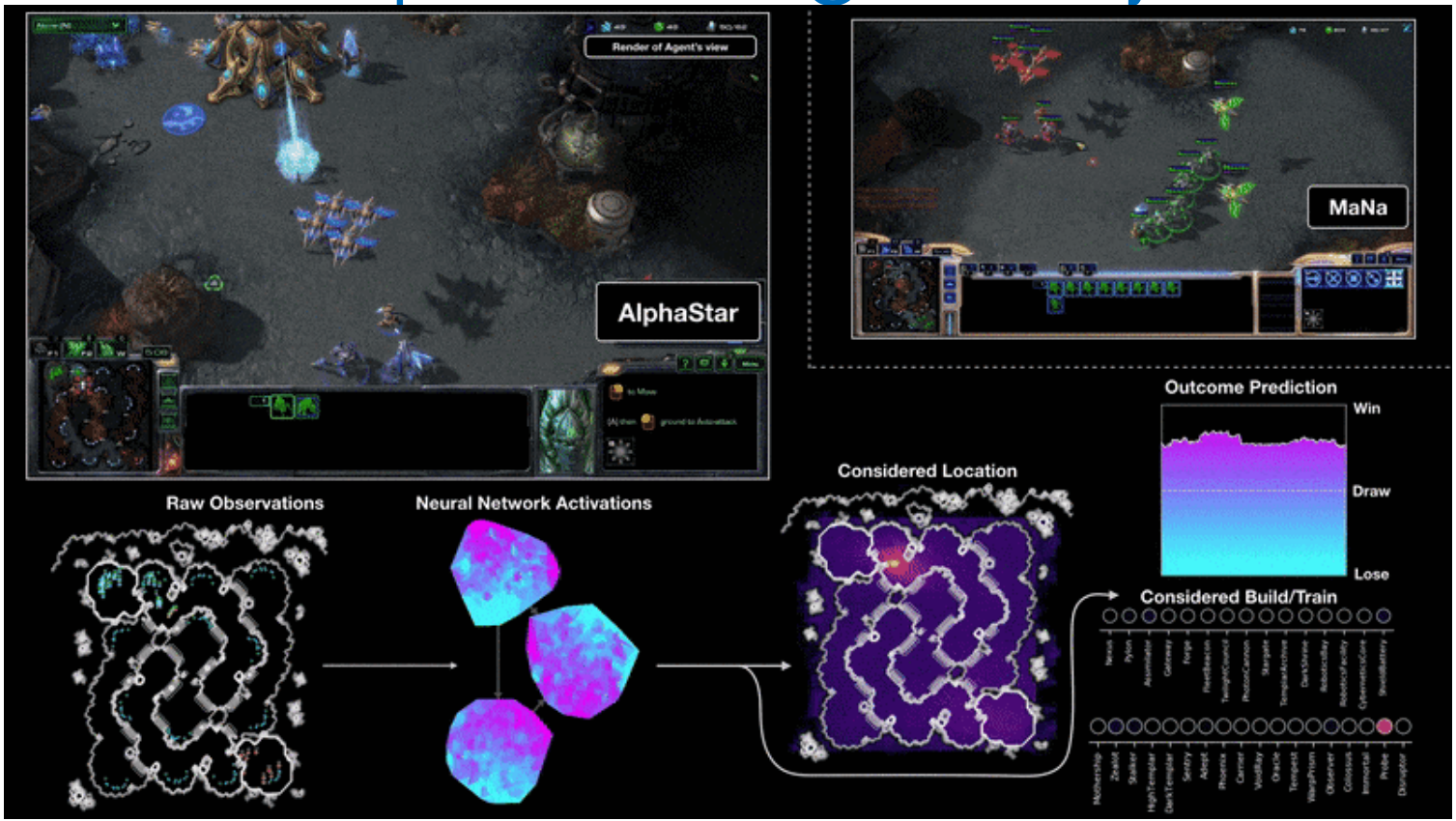


Emoticon suggestion



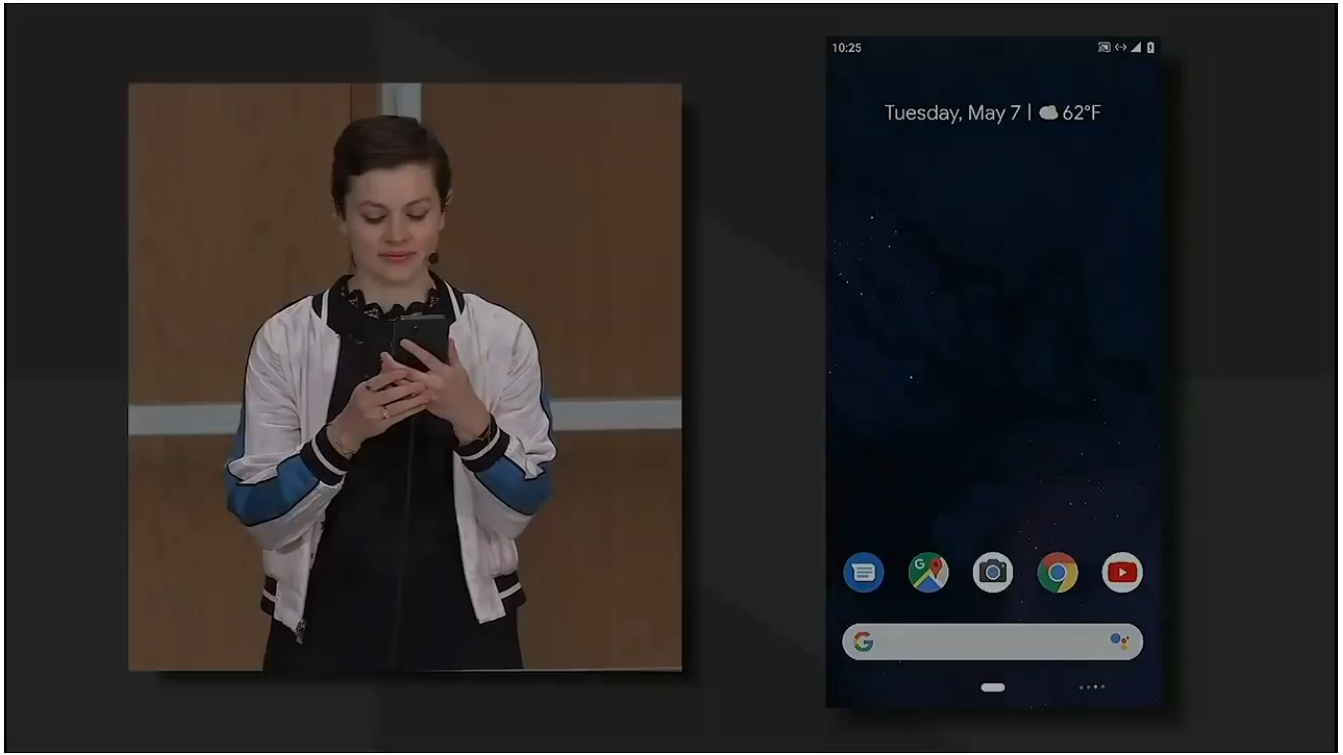
Machine translation

Deep Learning Today



Alpha Star

Deep Learning Today

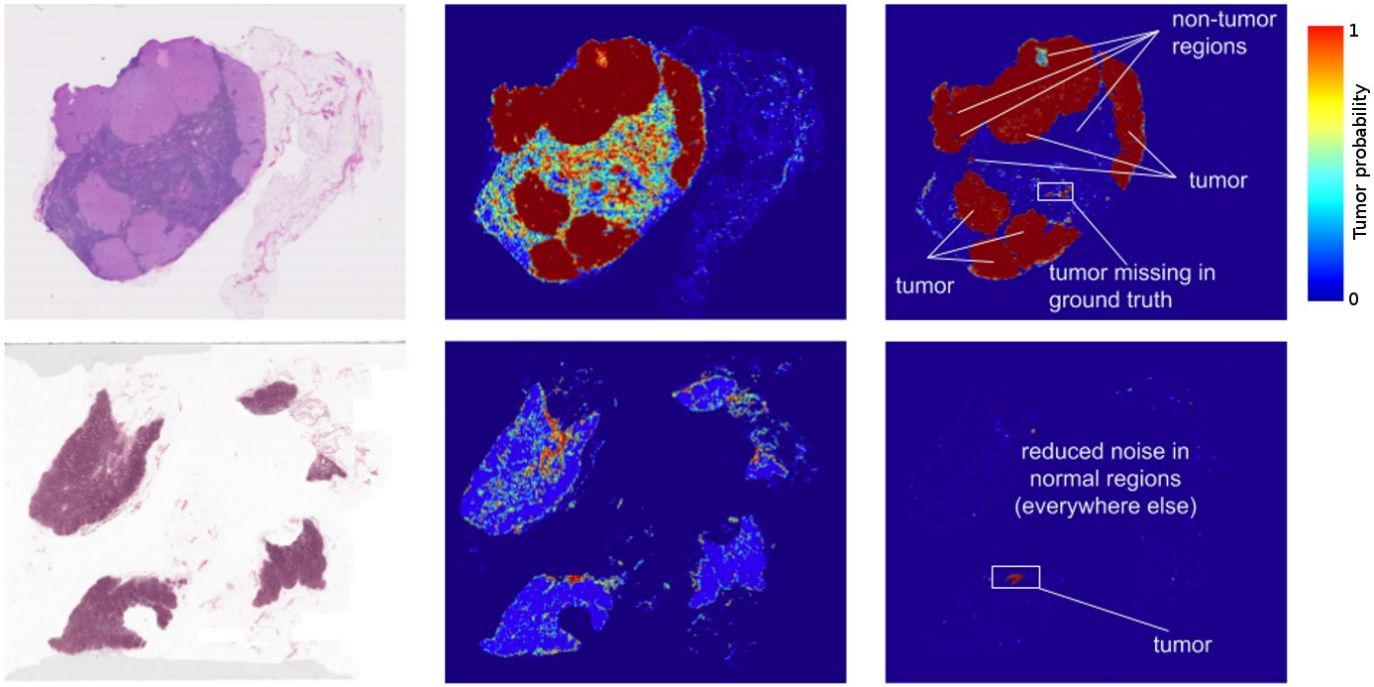


Google Assistant (Google IO'19)

Deep Learning Today

Google LaMDA (Google IO'22)

Deep Learning Today



Healthcare, cancer detection

Deep Learning Today

TEXT DESCRIPTION

An astronaut Teddy bears A bowl of soup

mixing sparkling chemicals as mad scientists shopping for groceries working on new AI research

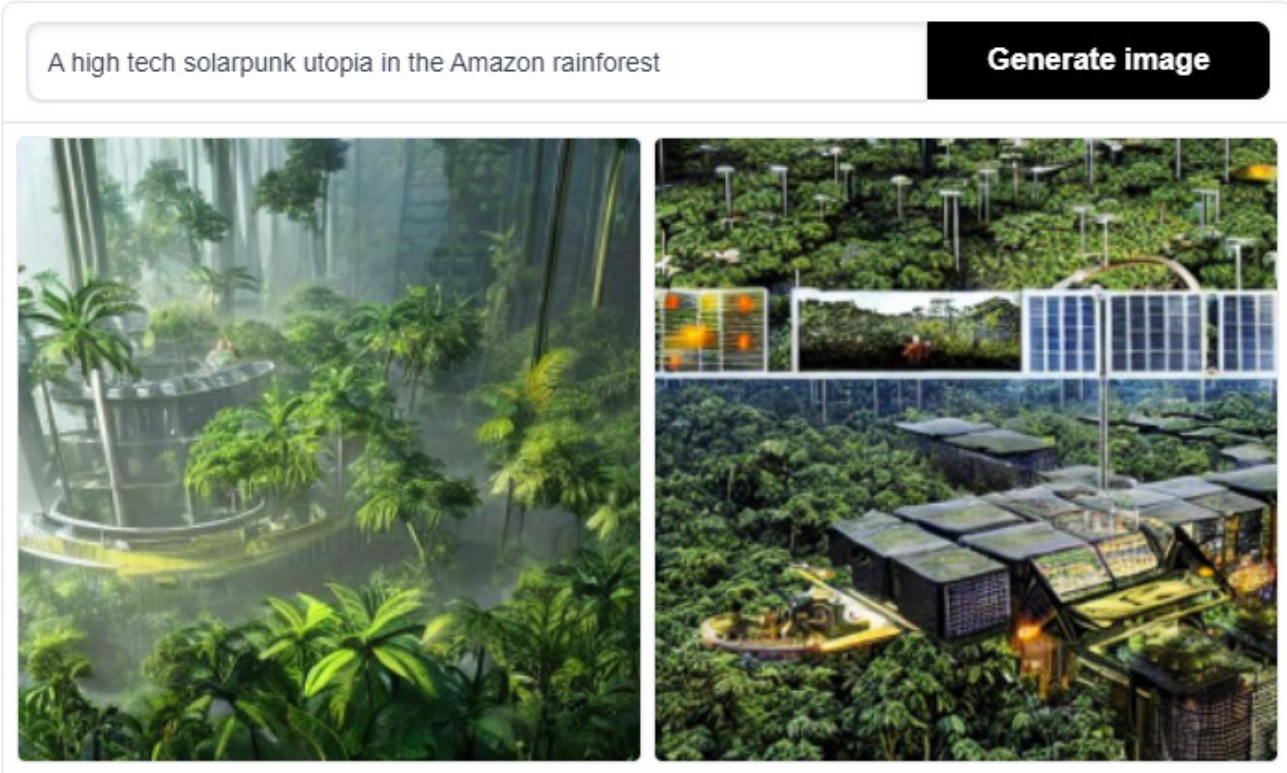
in the style of ukiyo-e as a one-line drawing in ancient Egypt

→

DALL·E 2

[Dall-E 2](#) Image generation from text

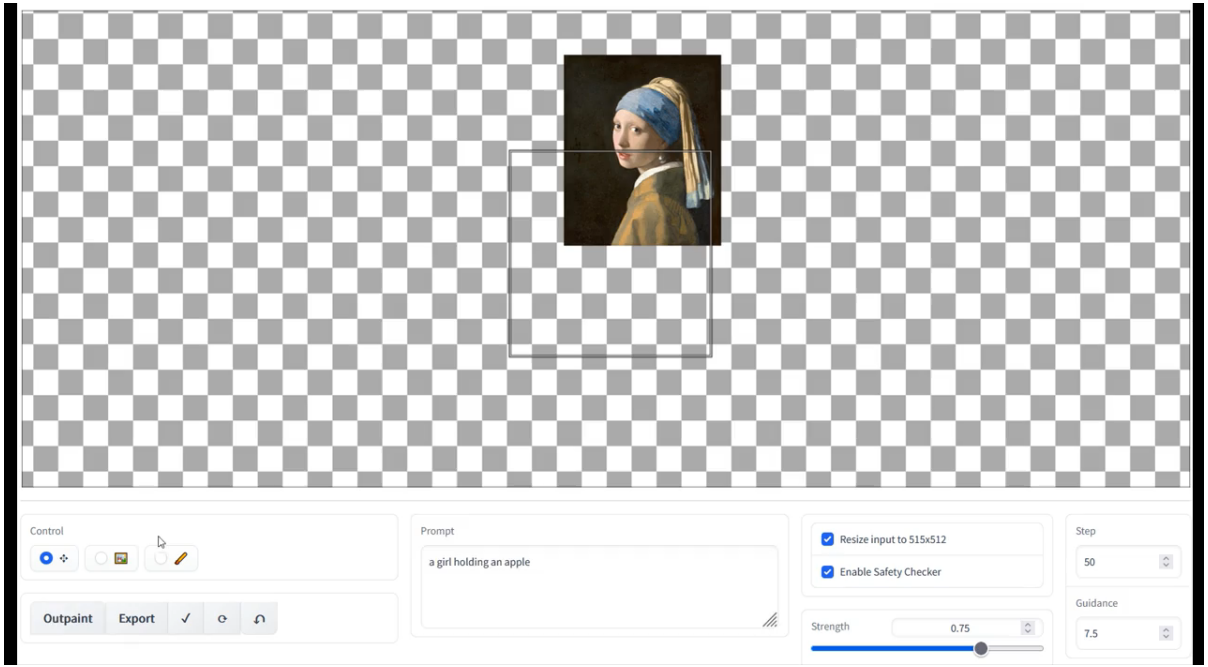
Deep Learning Today



StableDiffusion Image generation from text

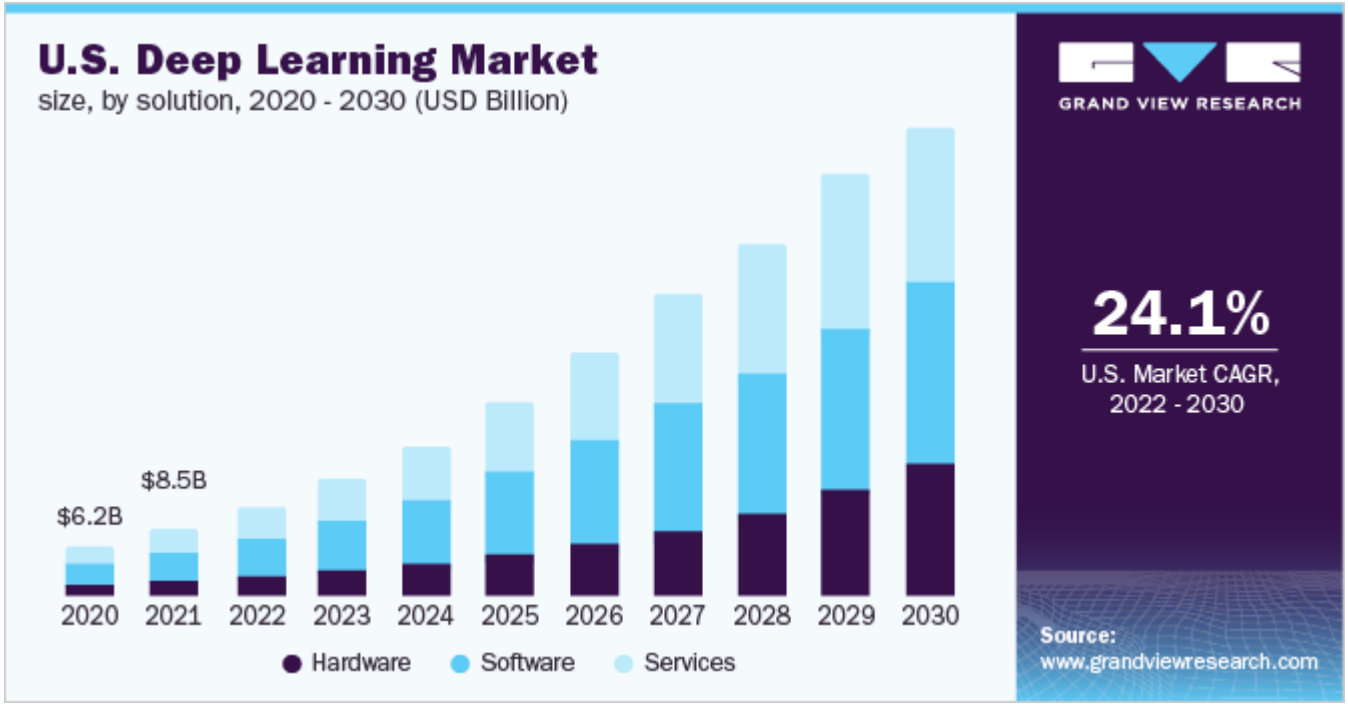
https://huggingface.co/spaces/stabilityai/stable-diffusion

Deep Learning Today



StableDiffusion Image Outpainting

Deep Learning Market



[…] market research report Deep Learning Market […] " the deep learning market is expected to be worth USD 415 Billion by 2030.

Deep Learning Job Perspective

- Excellent Job Perspectives!
  - Automation requires ML/DL -> growth!
  - Top-notch companies will gladly hire you!

- Many industries now:
  - IT-Companies
  - Cars, Logistics, Health Care, etc…
  - Manufacturing / Robotics, etc…

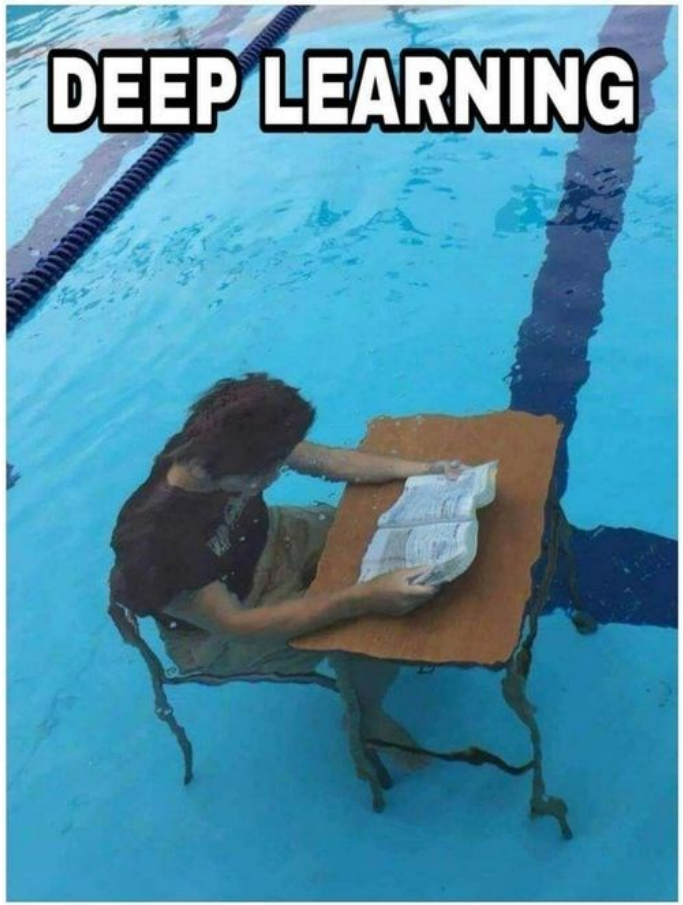
But: Also Challenging!

- High-level understanding is not enough
  - Need proper theory background
  - Need proper practical skillsets

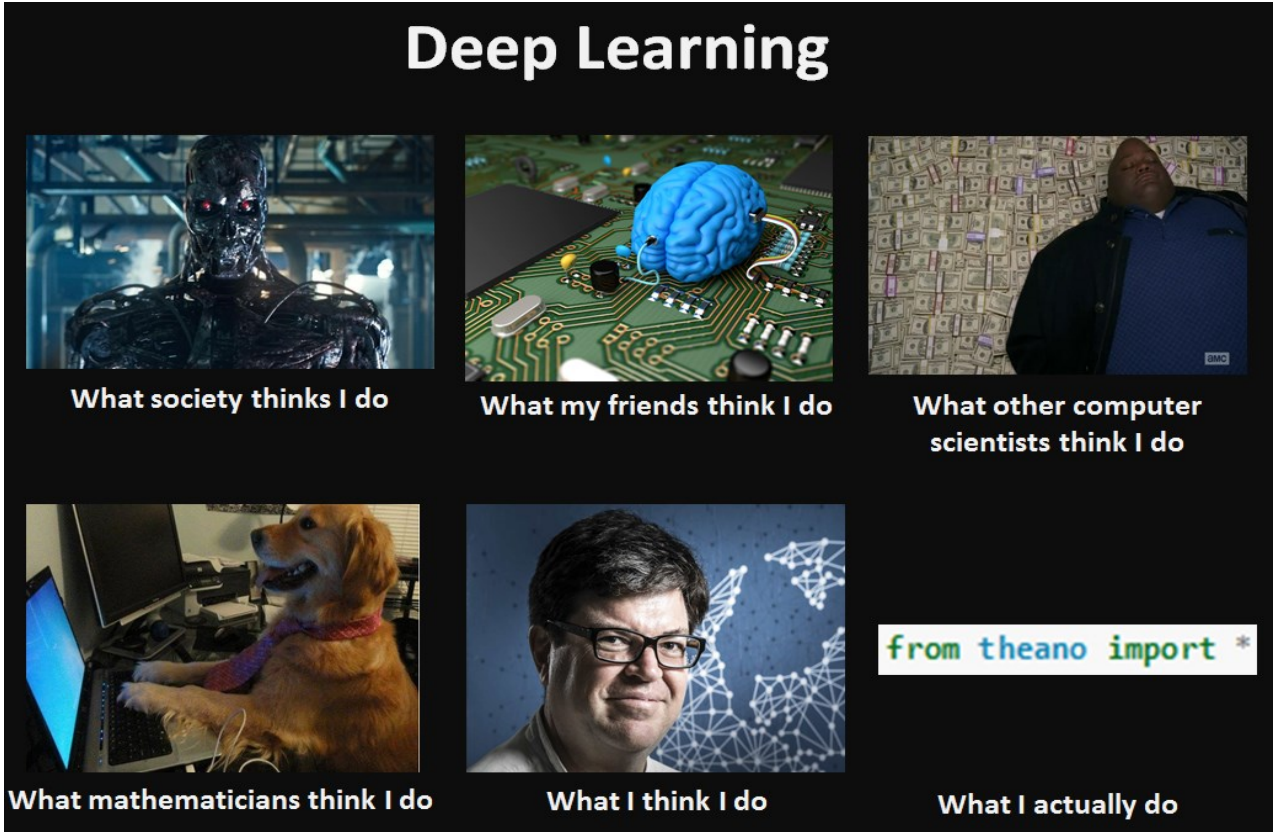
- Can be competitive!
  - Many good people
  - Downloading scripts / running code not enough ☺
  - Deeper understanding often requires PhDs

Deep Learning on the Internet

Deep Learning Memes

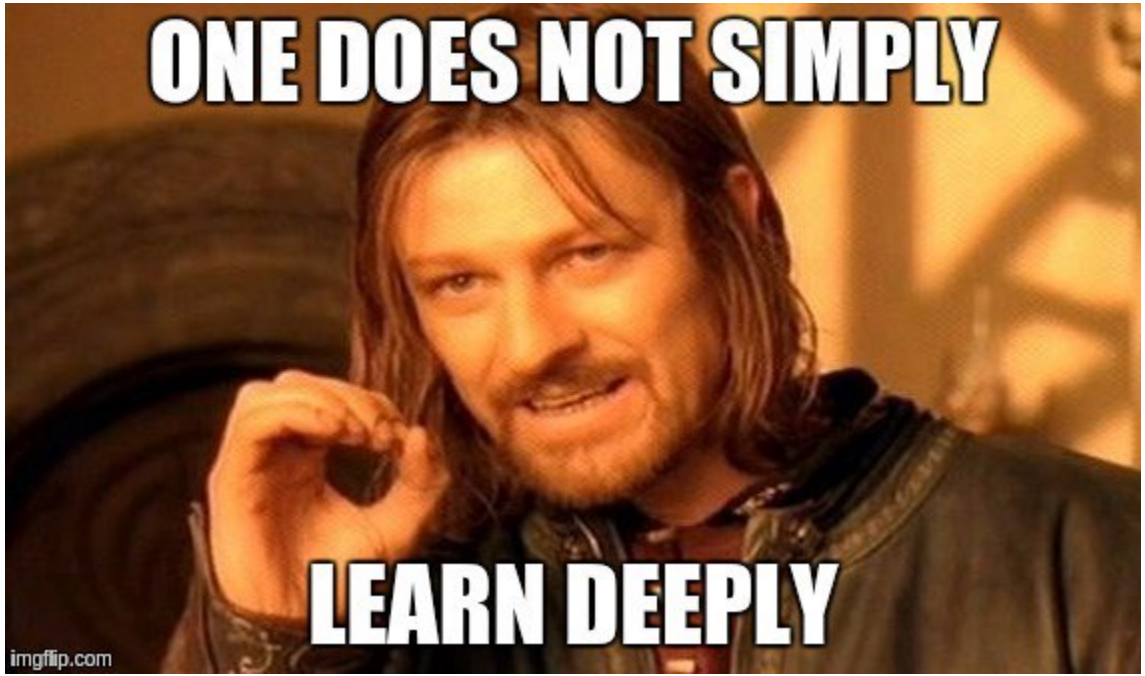


Deep Learning Memes

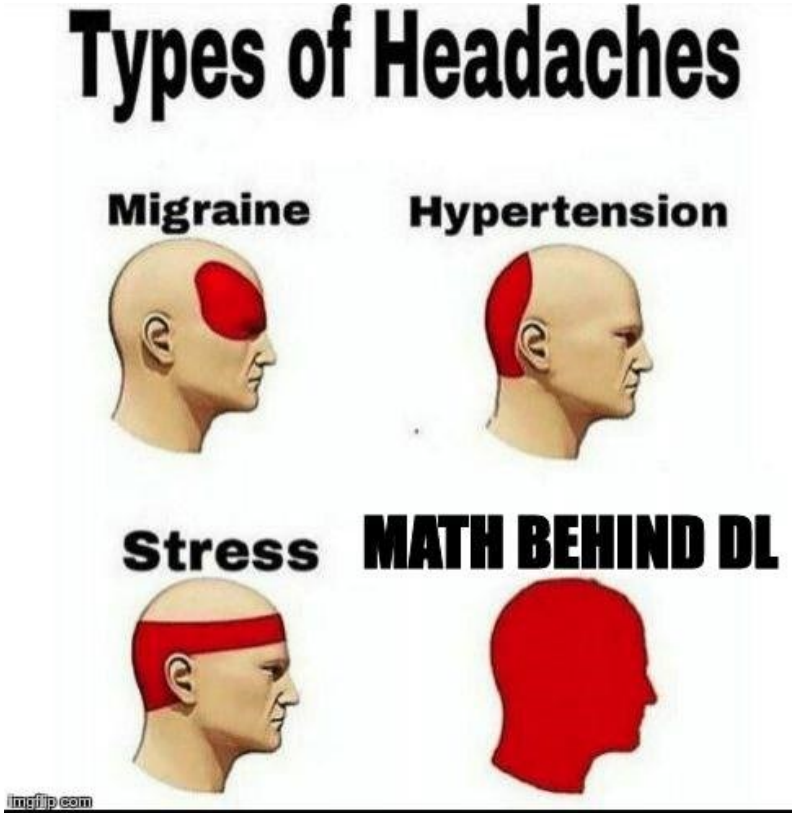


Deep Learning Memes

Deep Learning Memes



Deep Learning Memes



Deep Learning at TUM

Many TUM Research Labs use DL

- 3D AI Lab (Prof. Dai)
  - Research in 3D perception, 3D scene understanding
- Visual Computing Lab (Prof. Niessner):
  - Research in computer vision, graphics, and machine learning
- Computer Vision Group (Prof. Cremers)
  - Research in computer vision and pattern recognition
- Data Mining and Analytics Lab (Prof. Günnemann)
  - Research methods for robust machine learning
- Computer Aided Medical Procedures (Prof. Navab)
  - Research in machine learning for medical applications
- And probably many more ☺

Our Research Lab

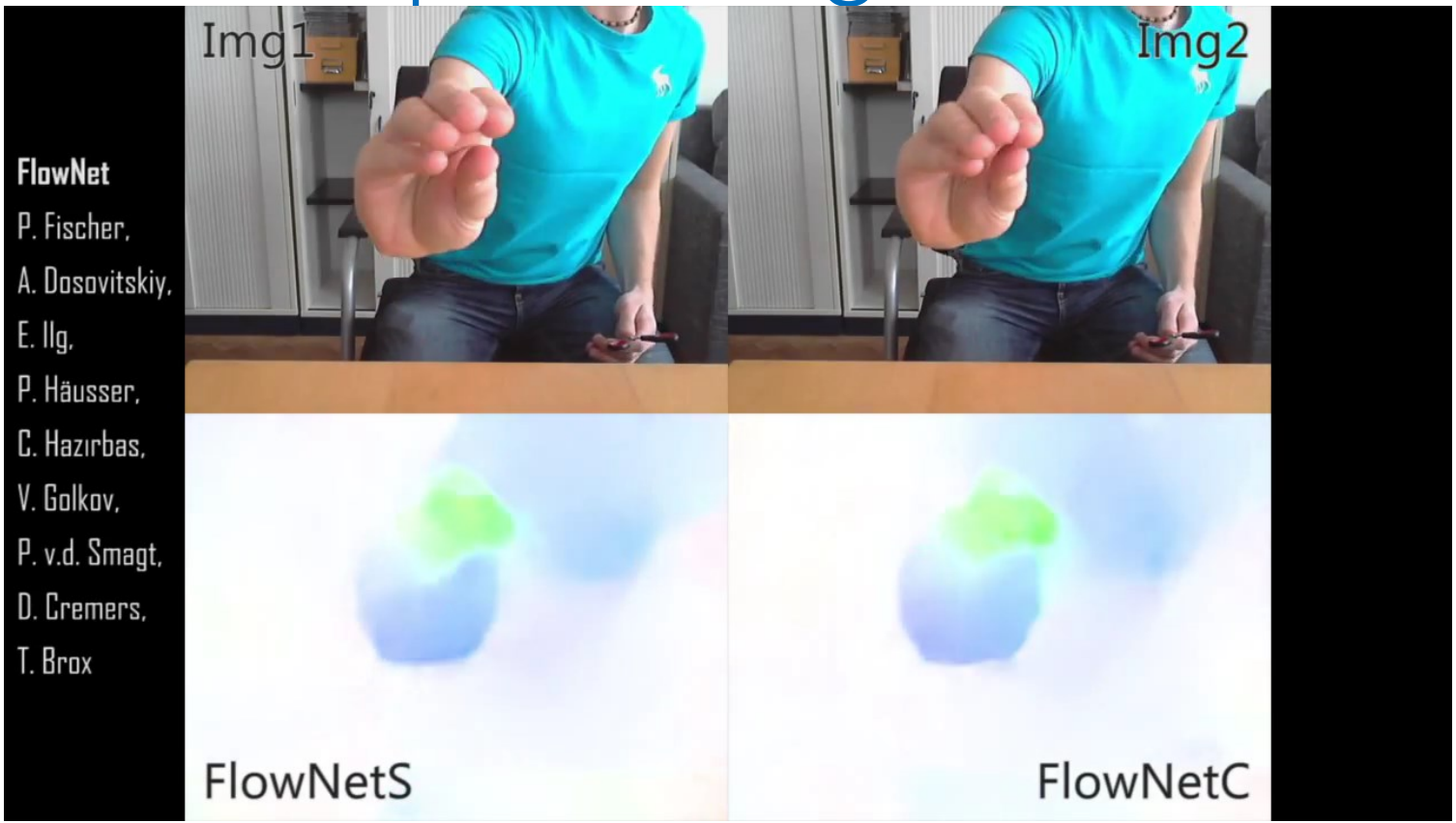
3D AI Lab (Prof. Dai):
https://www.3dunderstanding.org/publications.html

- Twitter: https://twitter.com/angelaqdai
- Youtube: https://www.youtube.com/channel/UC71pxwyTLVK1H0puMmN7pRw

Deep Learning at TUM

[Caelles et al., CVPR' 17] One-Shot Video Object Segmentation

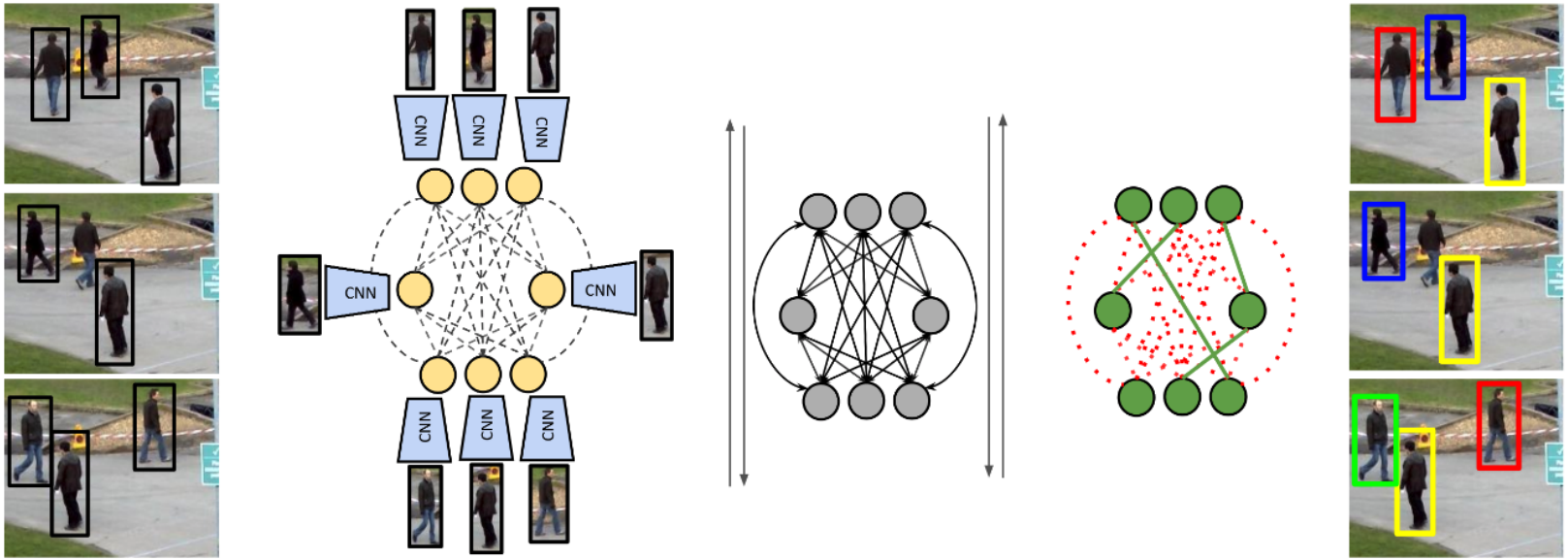
Deep Learning at TUM



[Dosovitskiy et al., ICCV' 15] FlowNet

Deep Learning at TUM

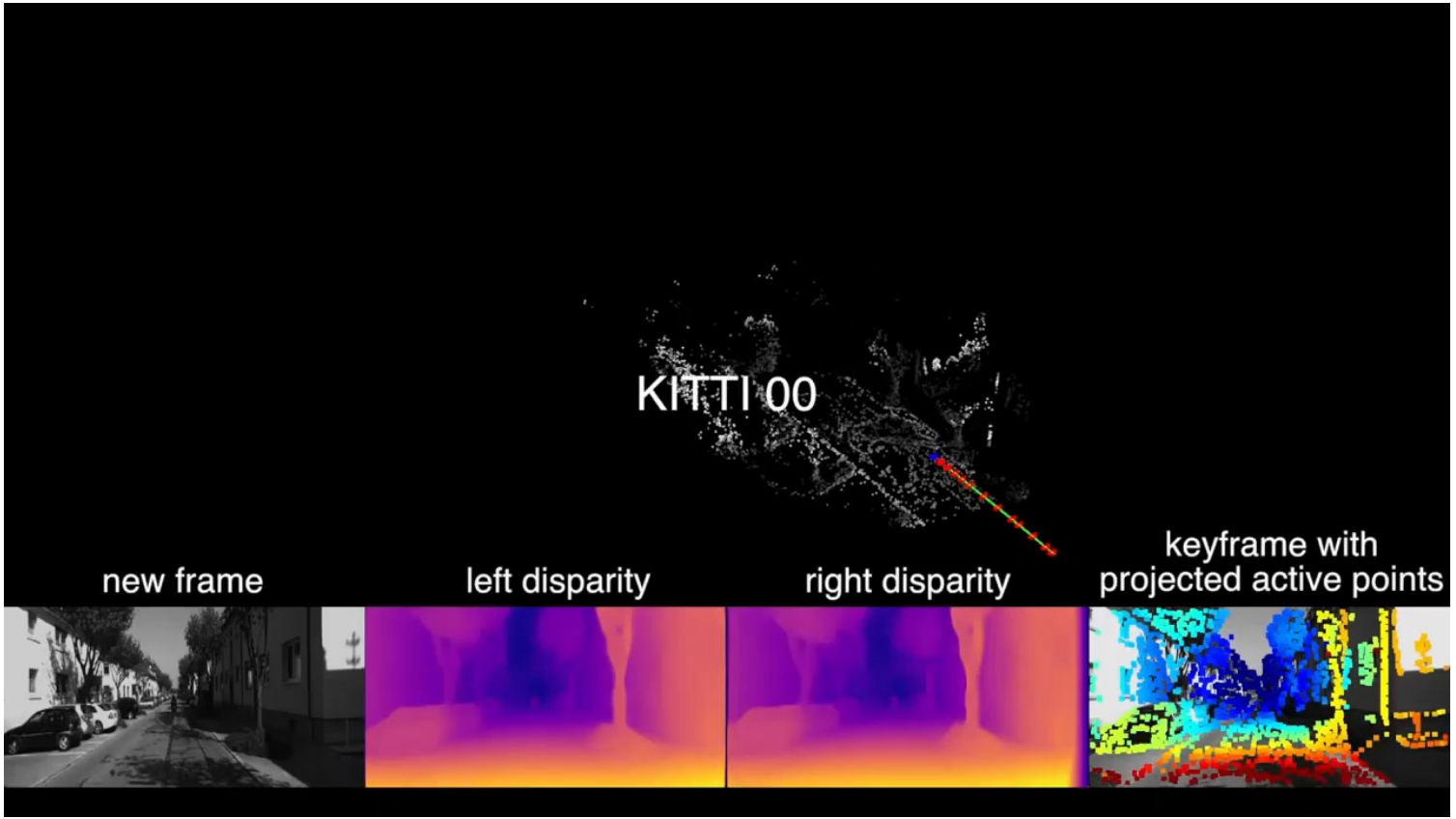
- Multiple object tracking with graph neural networks



(a) Input (b) Graph Construction + Feature Encoding (c) Neural Message Passing (d) Edge Classification (e) Output

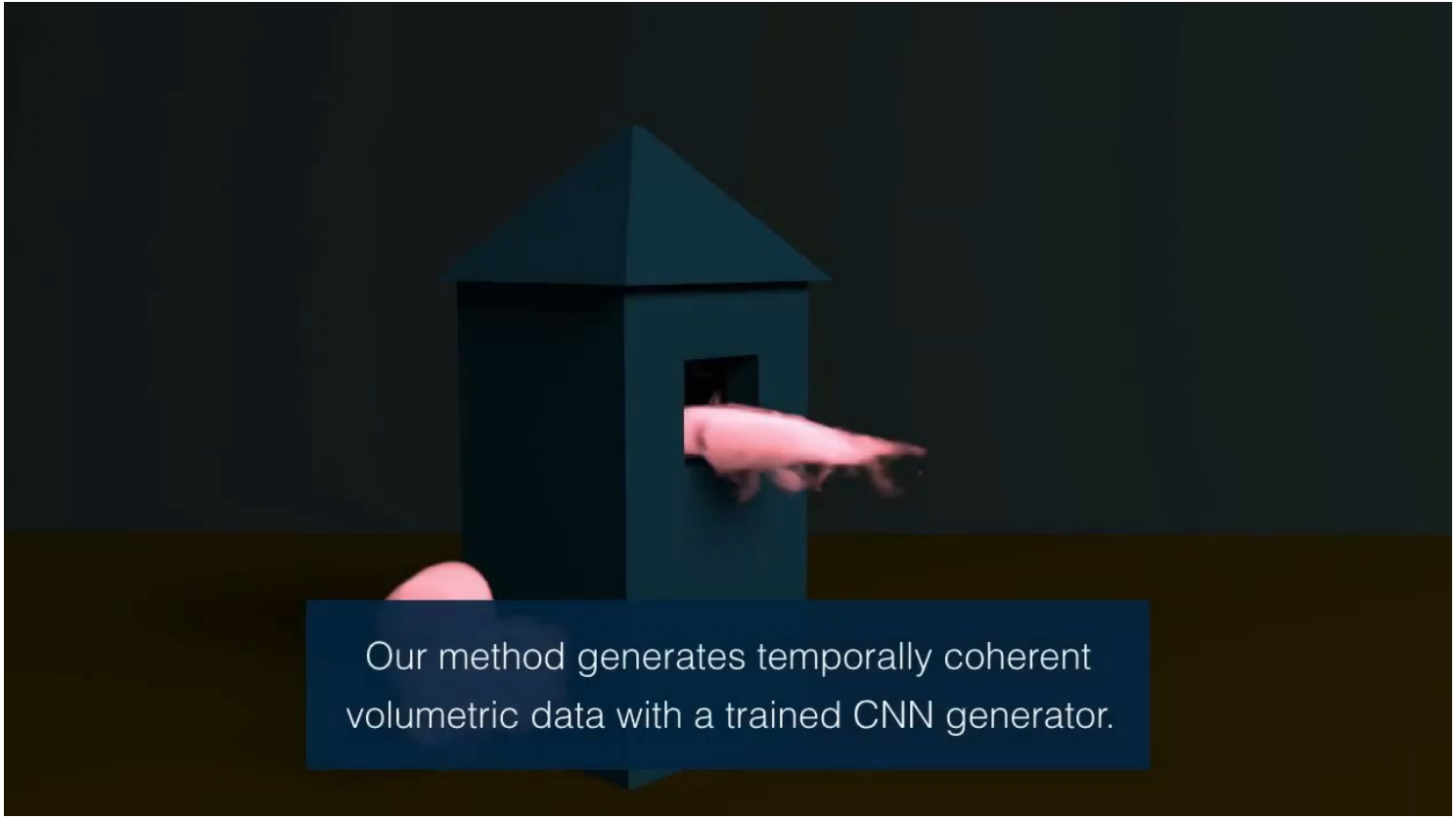
[Brasó and Leal-Taixé, CVPR 2020] Learning a Neural Solver for Multiple Object Tracking.

Deep Learning at TUM



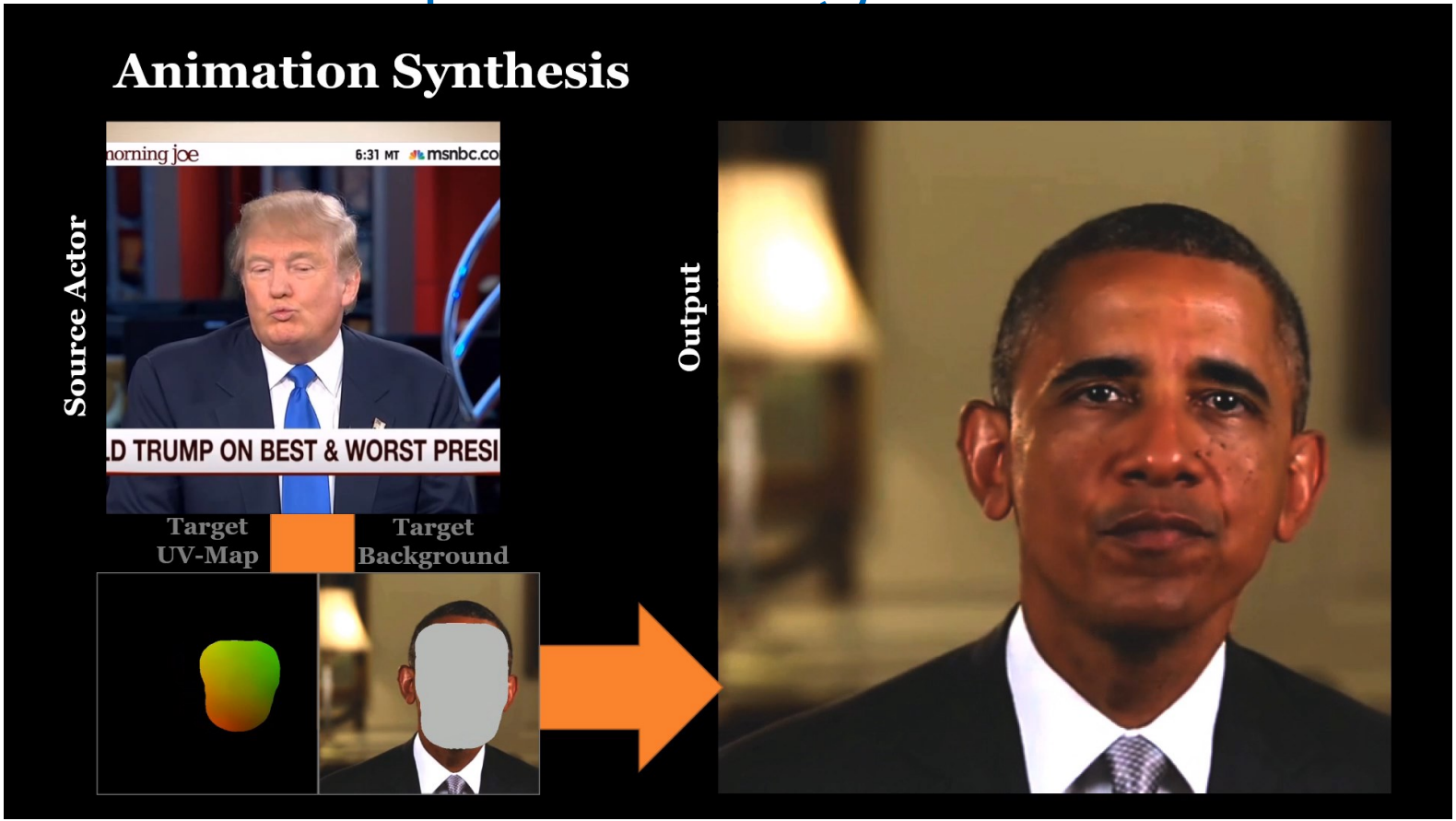
[Yang et al., ECCV' 18] Deep Virtual Stereo Odometry

Deep Learning at TUM



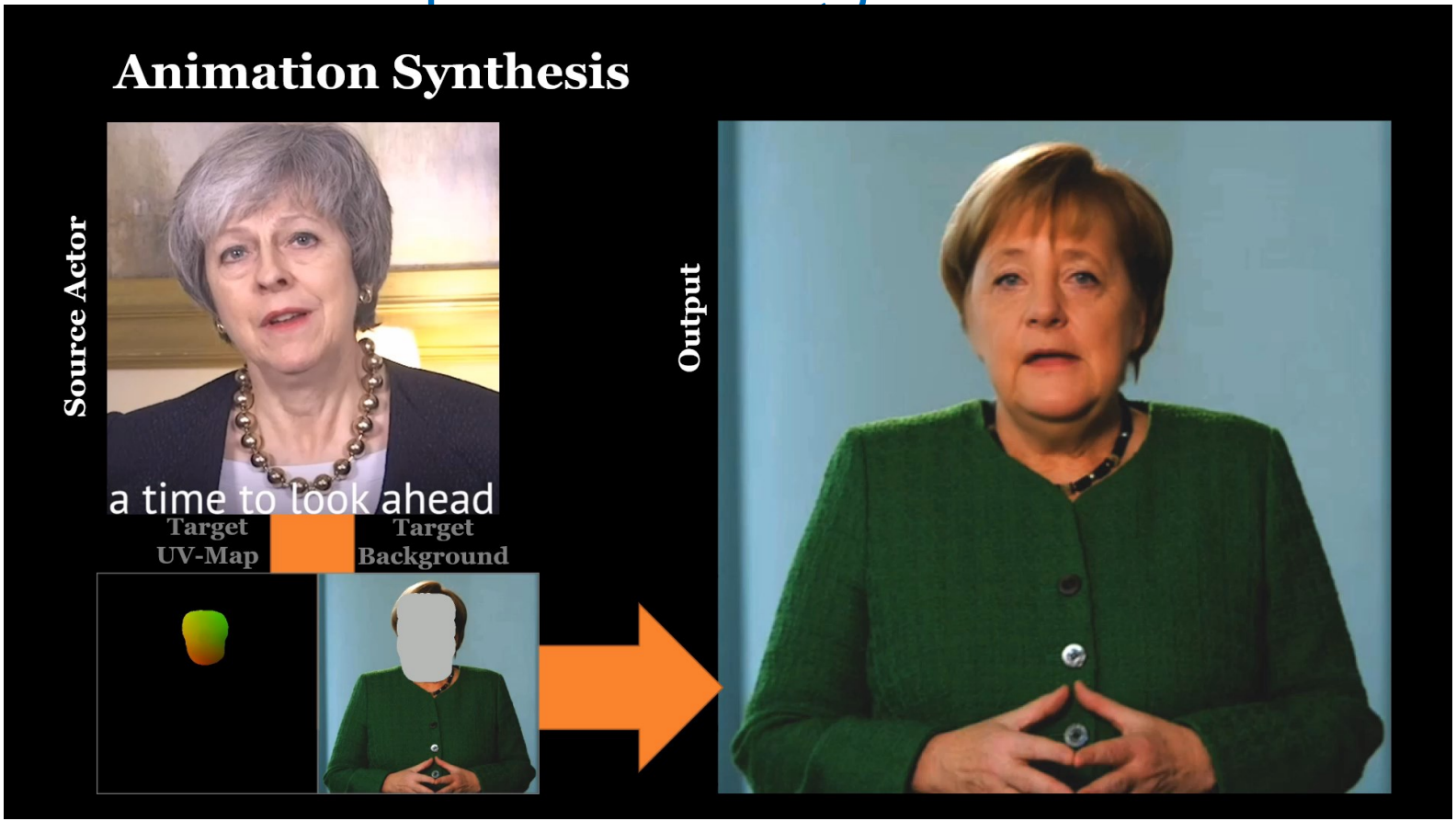
[Xie et al. Siggraph' 18] tempoGAN

Deep Learning at TUM



[Thies et al., Siggraph'19]: Neural Textures

Deep Learning at TUM



[Thies et al., Siggraph'19]: Neural Textures

Deep Learning at TUM

Input Video

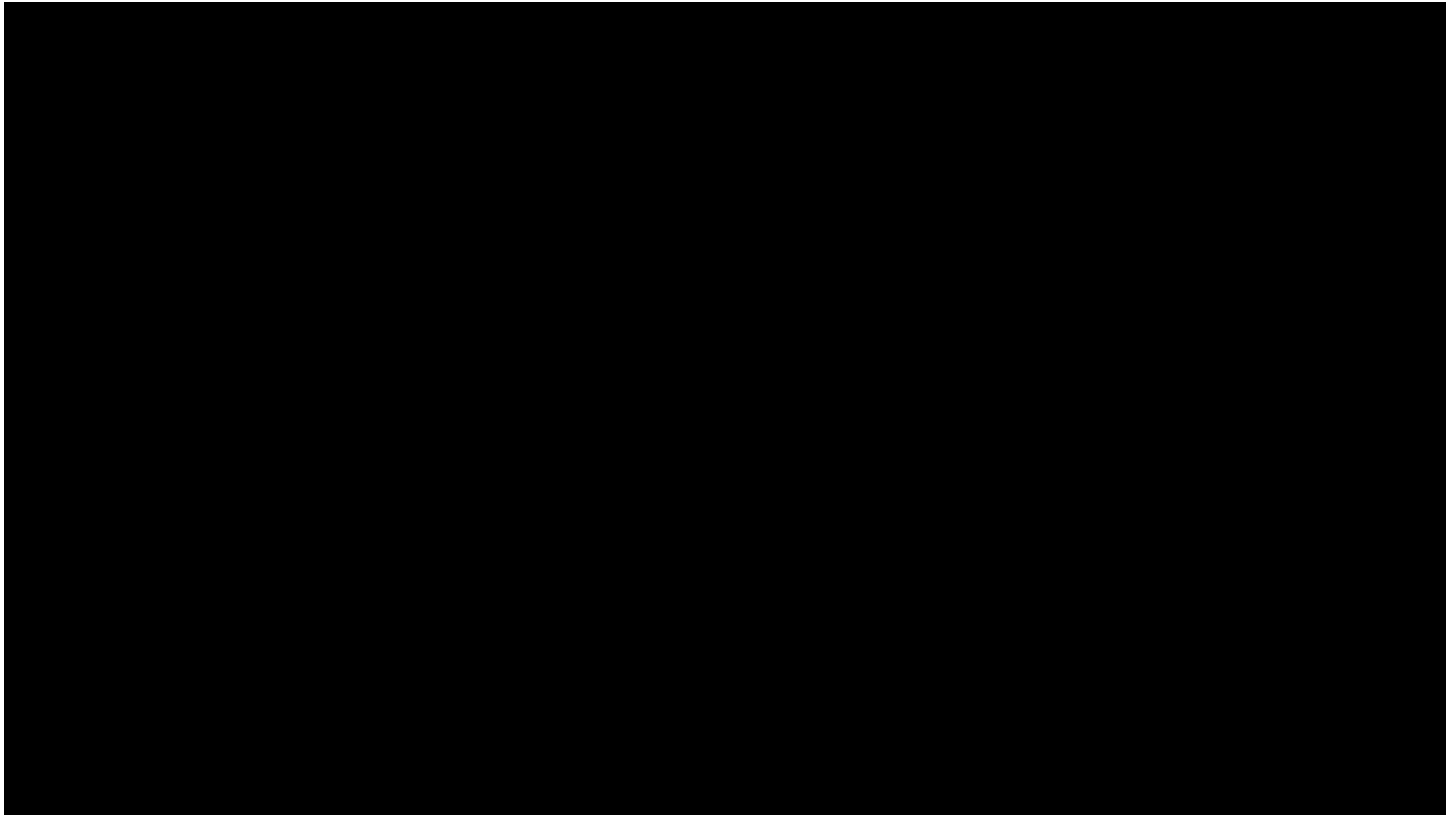
4D Head Avatar

[Grassal et al., CVPR'22]: Neural Head Avatars

Deep Learning at TUM

[Müller et al., CVPR'22]: AutoRF

Deep Learning at TUM



[Božič et al., NeurIPS'21]: TransformerFusion

Deep Learning at TUM

ScanNet Stats:
-Kinect-style RGB-D sensors
-1513 scans of 3D environments
-2.5 Mio RGB-D frames
-Dense 3D, crowd-source MTurk labels
-Annotations projected to 2D frames

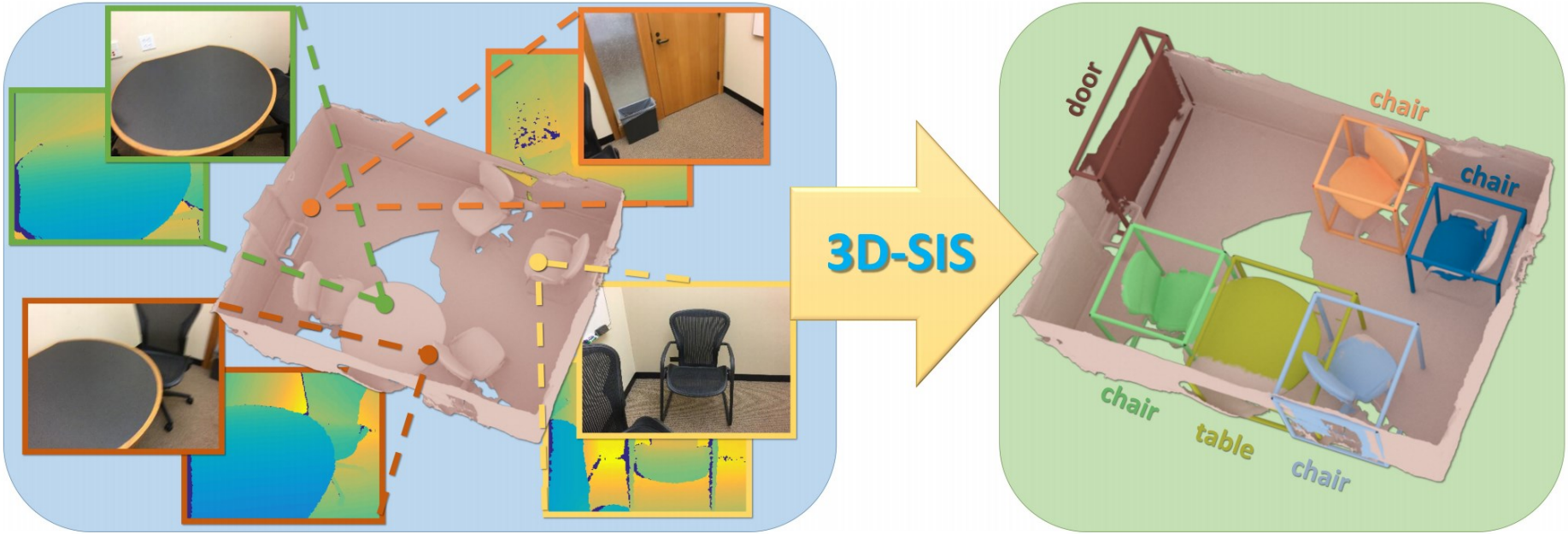
[Dai et al., CVPR'17] ScanNet

Deep Learning at TUM

Learning 3D Using Language

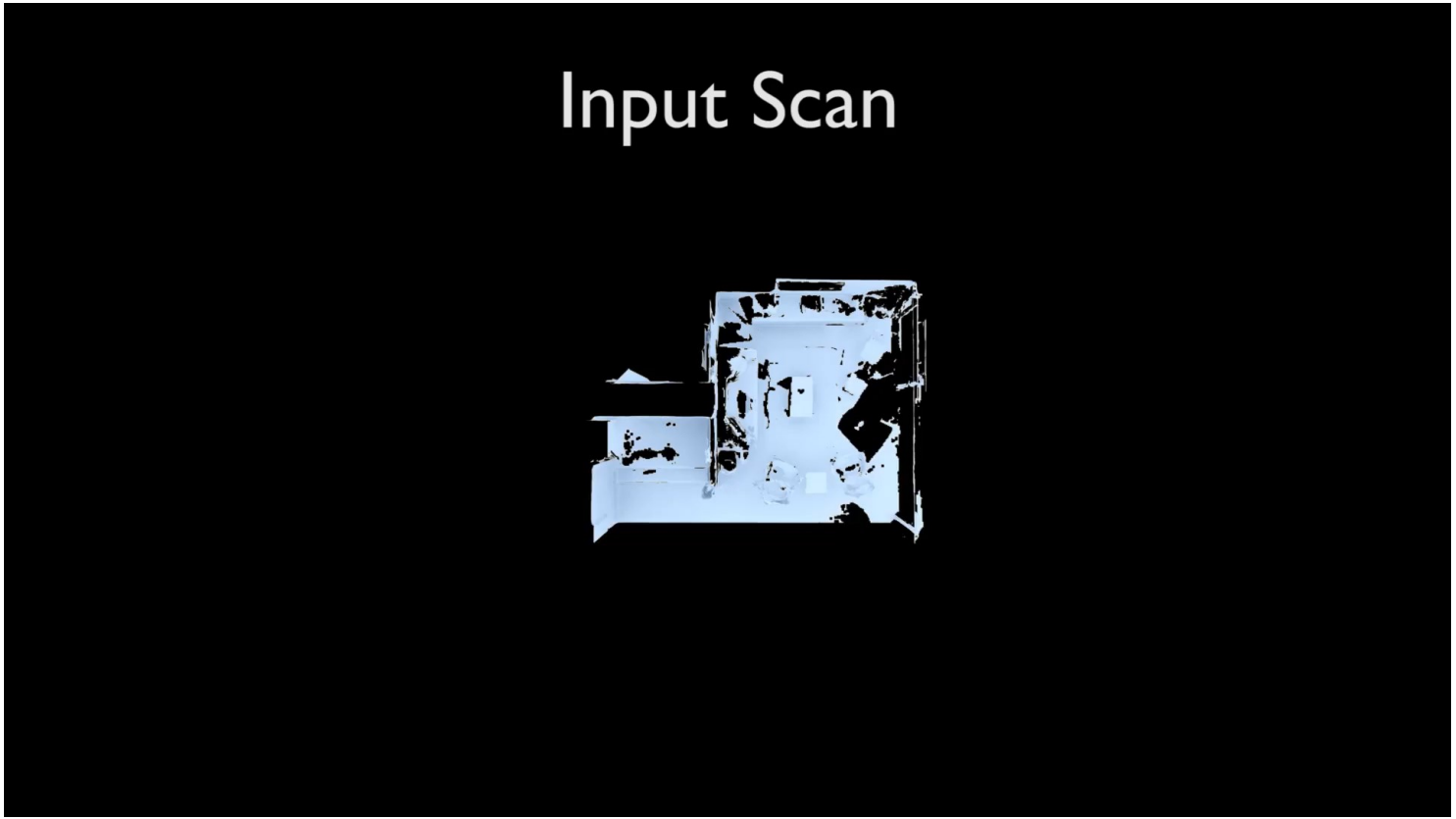
[Rozenberszki et al., ECCV'22]

Deep Learning at TUM



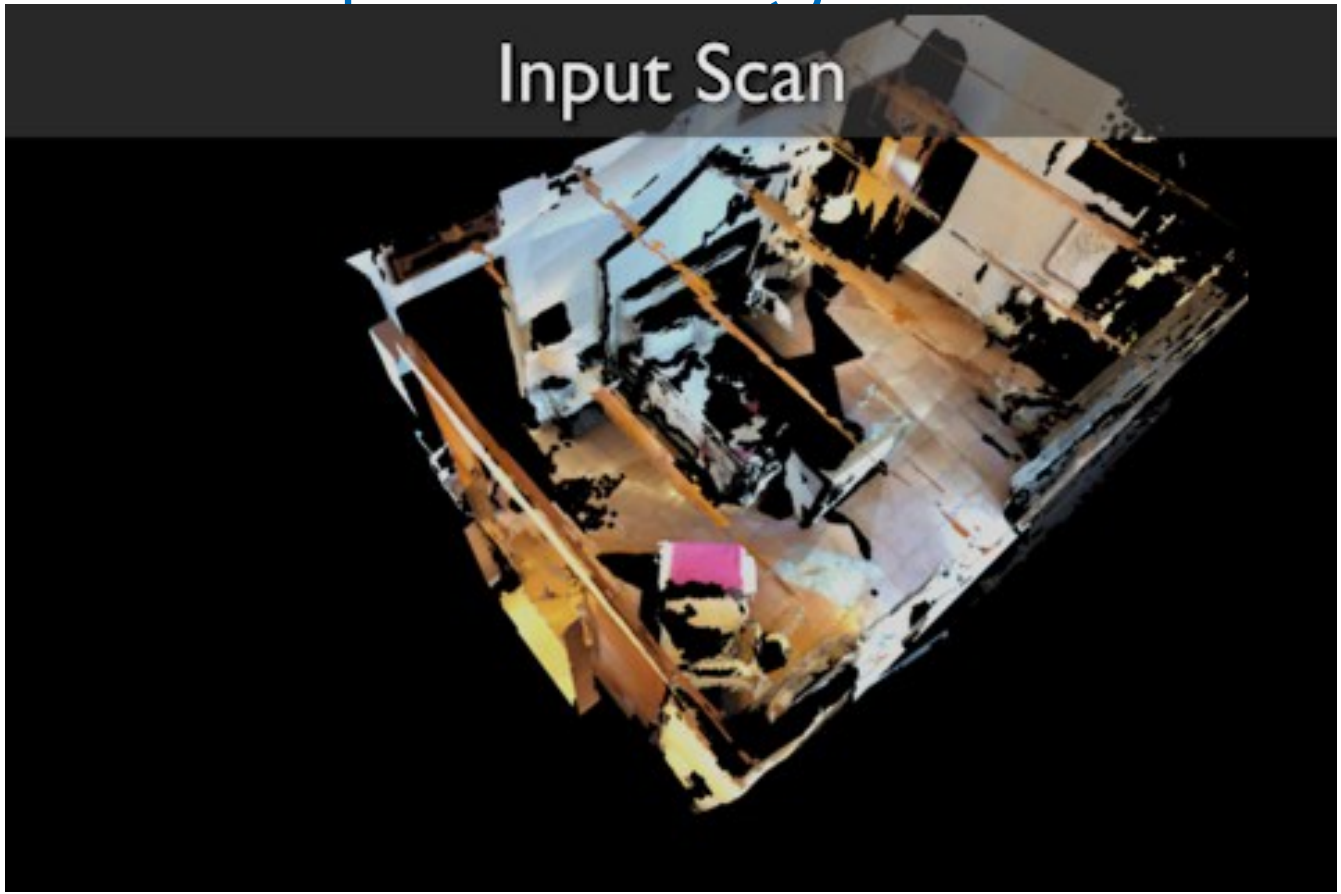
[Hou et al., CVPR'19] 3D Semantic Instance Segmentation

Deep Learning at TUM



[Dai et al., CVPR'20] SG-NN

Deep Learning at TUM

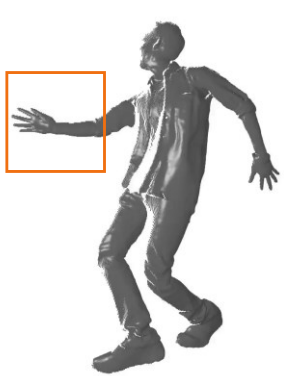


[Dai et al., CVPR'21] SPSG

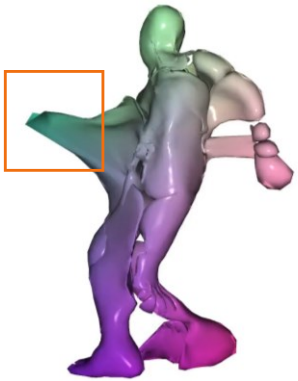
Deep Learning at TUM

[Siddiqui et al., ECCV'22] Texturify

Deep Learning at TUM



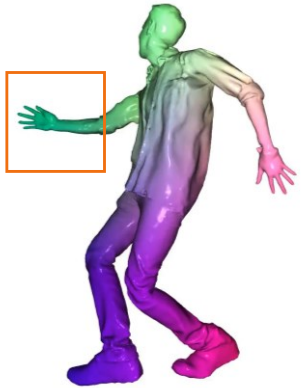
Depth Input
(lateral view)



IPNet
[Bhatnagar et al. 2020]

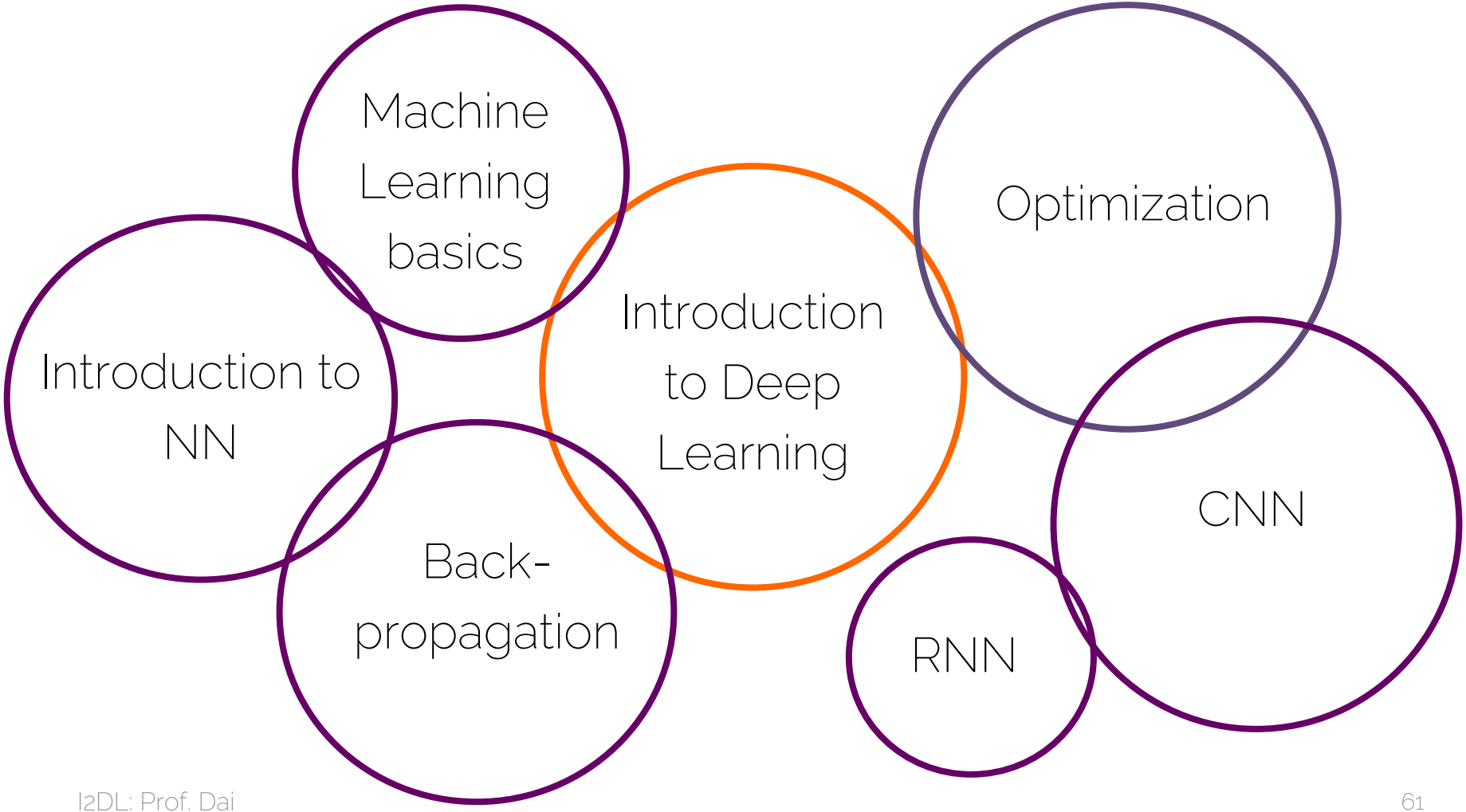
NPMs*
[Palafox et al. 2021]

SPAM (Ours)

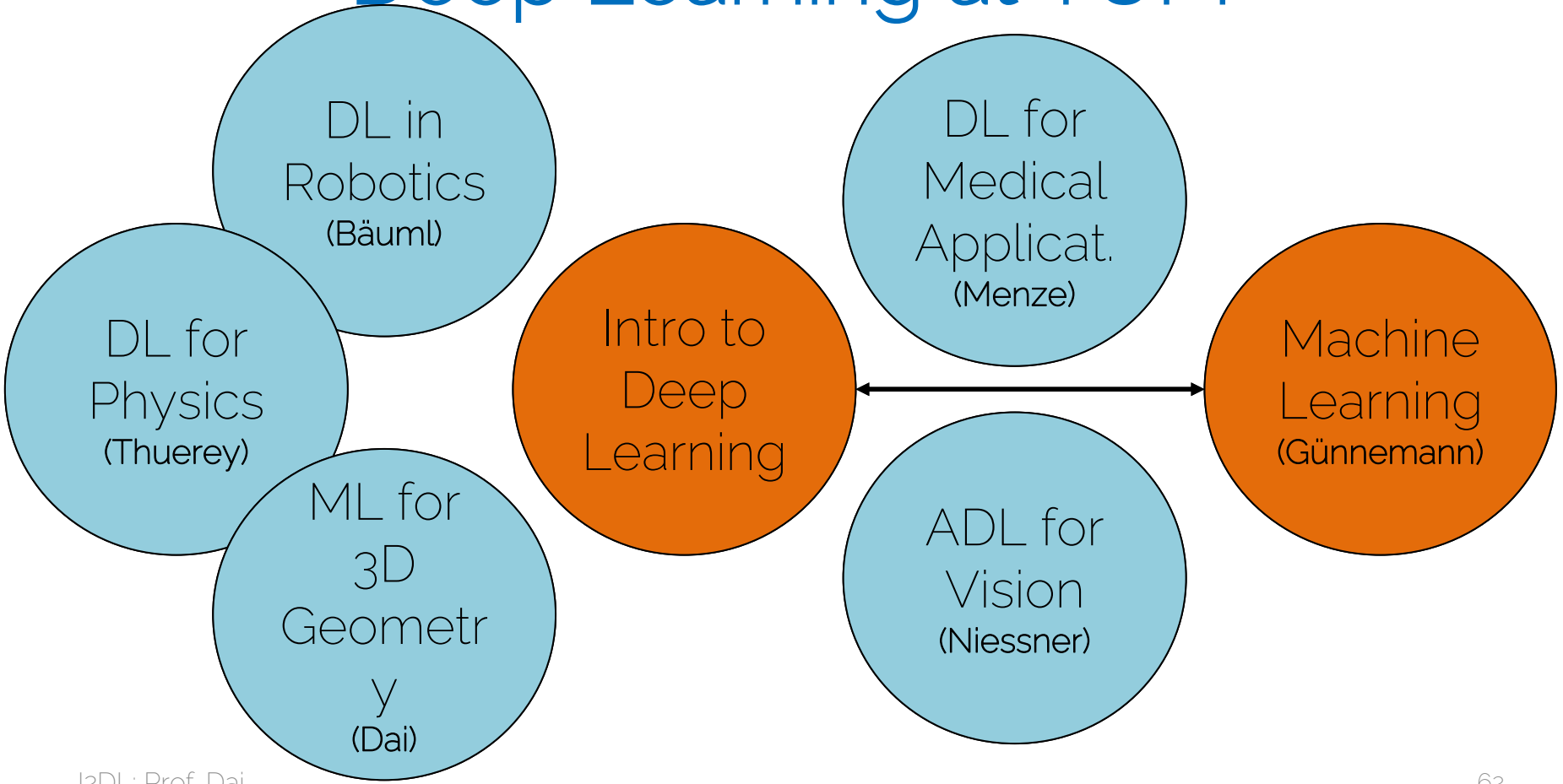


Ground Truth

Context of Other Lectures at TUM


Machine Learning basics
Optimization
Introduction to NN
Introduction to Deep Learning
CNN
Back-propagation
RNN

Deep Learning at TUM



Why is I2DL so Important?

- Many of the other lectures / practical require it!
  - Often only limited spots are available (e.g., in the Advanced Deep Learning for Computer Vision Class)

- Solid preparation for guided research / IDP
  - Most topics require it
  - For career in AI/DL these are the best ways to get into

Introduction to Deep Learning

Logistics

About the Lecture

- Theory lectures
- Every Tuesday 14:00-16:00
  - Lectures are on-site and will be streamed to RGB Live: https://live.rbg.tum.de/course/2022/W/i2dl
- Practical exercises (see later slides)
  - Released every Thursday, 13:00
  - Tutorial: Online videos (uploaded to live.rgb.tum.de as well)
  - Programming exercises

- Guest lecture TBD

Preliminary Syllabus

Lecture 1: Introduction to the lecture, Deep Learning, Machine Learning.
Lecture 2: Machine Learning Basics, Linear regression, Maximum Likelihood
Lecture 3: Introduction to Neural Networks, Computational Graphs
Lecture 4: Optimization and Backpropagation
Lecture 5: Scaling Optimization to large Data, Stochastic Gradient Descent
Lecture 6: Training Neural Networks I
Lecture 7: Training Neural Networks II
Lecture 8: Training Neural Networks III
Lecture 9: Introduction to CNNs
Lecture 10: CNNs architectures;
Lecture 11: Recurrent Neural Networks (RNNs)
Lecture 12: Advanced Deep Learning architectures
Guest Lecture (TBD)

~~Moodle~~ → Piazza

- Announcements via Piazza - IMPORTANT!
  - Sign up online for access: http://piazza.com/tum.de
    - Select "Winter 2023" term, search for IN2346
    - Use your @mytum.de email address
  - We will share common information (e.g., regarding exam)
- Forum
  - Ask and discuss questions
    - Tutors will monitor and answer questions
    - You are very welcome to actively participate
  - Please do not post solutions of the exercises
  - You can post private question visible only to the staff

Email

- Email list:

i2dl@vc.in.tum.de

- Do NOT email us personally!
  - Cannot handle so many emails / hence will be ignored

- Email list for organizational questions only!
  - Content questions -> Piazza or Office Hours
  - Or post the question/issue in a private thread on Piazza

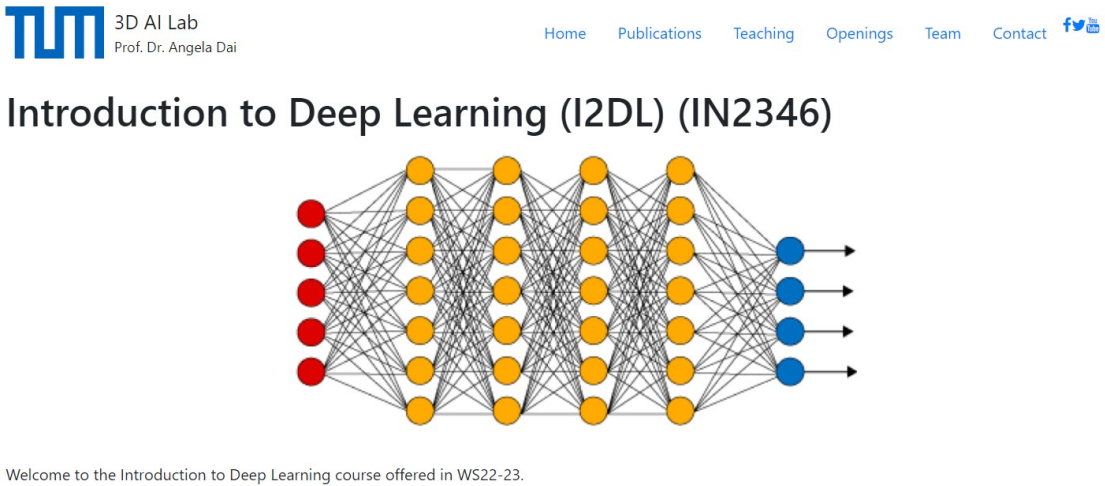
(Virtual) Office Hours

- We will have dedicated office hours regarding
  - Theoretical help (e.g., specific lecture questions)
  - Help on exercises

- Will be starting next week (24/10)
- Zoom links will be posted on Piazza

Website

- Links and slides will be shared on website



https://www.3dunderstanding.org/i2dl-w22/

Exam FAQ

- **Final Exam:** TBA
- **Content:** Lecture & exercises
- **Important:** No retake exam (I2DL is taught every semester)

- Grade Bonus:
  - Solve 8 out of 9 "non-optional" practical exercises
  - Bonus 0.3 on a **passed** final exam
  - Bonus is transferable from previous and future semesters

Other Administrative

- "External" students welcome (LMU, TUM PhD)
  - Fill out registration form and we will add you to the course
  - Will get Certificate / Schein at the end

- Lectures will be recorded and streamed on RGB Live
  - https://live.rbg.tum.de/course/2022/W/i2dl

- Again: check announcements on piazza

Practical Exercises

Exercise – Goal

- Hands-on programming experience (Learning by Doing)
- Reimplementations basic building blocks
- Introduction to common libraries
- (Get grade bonus)
- Ultimately: Gather enough experience to start your own individual (research) deep learning project

Exercise – Format

- Tutorial:
  - Video only
    Posted on live.rbg.tum.de
    Slides uploaded to piazza & website
  - Video length
    Ex02: Full lecture with written exercise
    Ex03-11: Short (~30min) video and coding exercises
- Programming:
  Interactive coding notebooks (~4h each)

Start time: Thursdays 13:00 *
Working Time: 1 week
Deadline: Wednesdays, 15:59

Except for exercise 1:
- No video → this lecture
- Starting time: Today
- Working time: 2 weeks

Exercises – Tech Stack

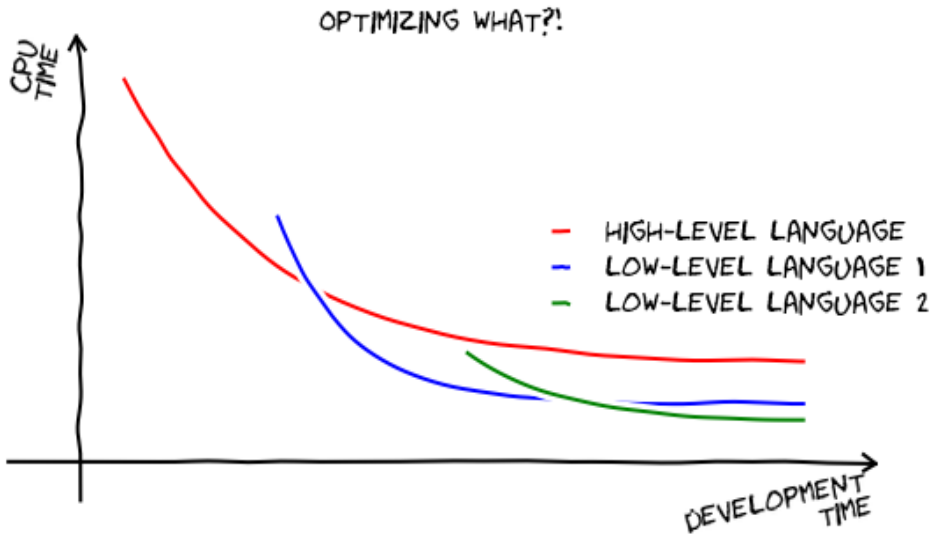
- Python
  Jupter notebooks
  Numpy

- Deep Learning library
  Pytorch

- Hardware requirements
  - Minimum: CPU
  - Preferred: Nvidia GPU
  - Alternative: Google Colab

Why Python?

- Very easy to write development code thanks to an intuitive syntax
- Biggest language used in deep learning research



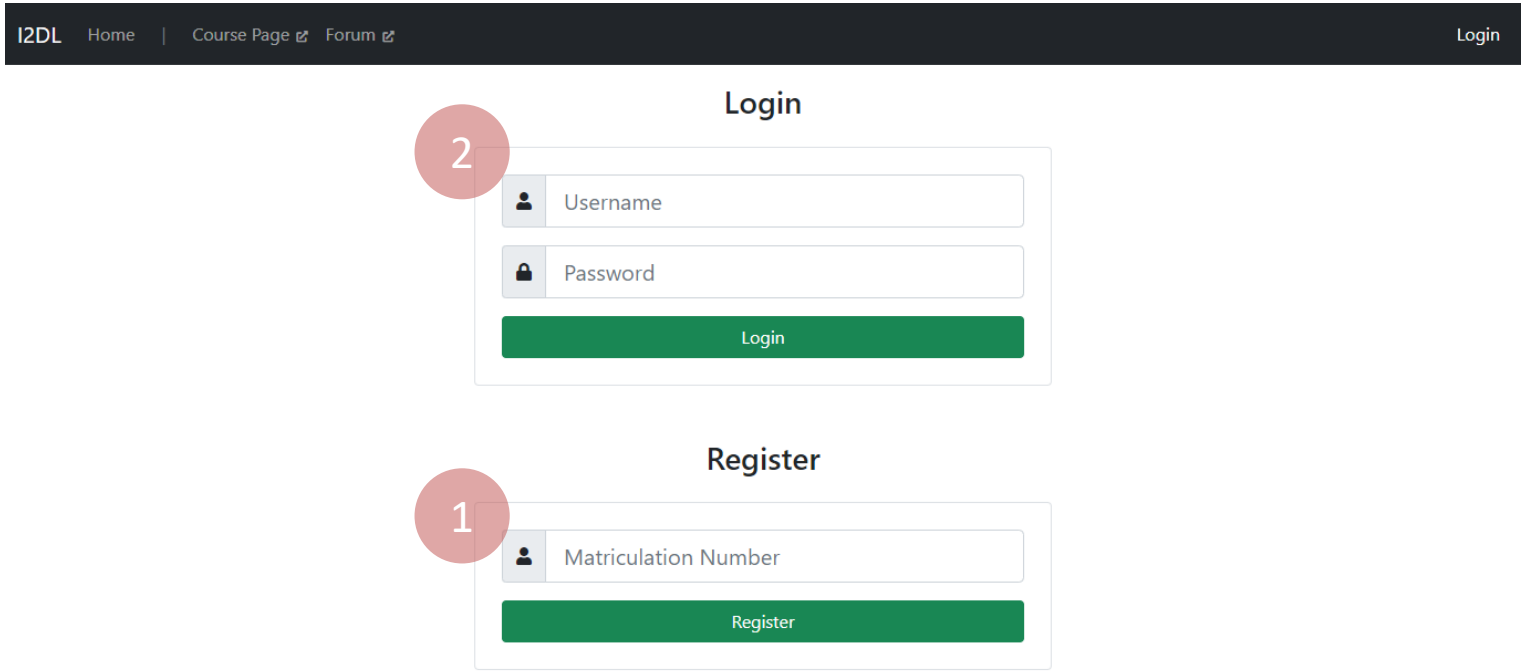
Exercises – Content

Exercise 01: Organization Exercise 02: Math Recap	Intro
Exercise 03: Dataset and Dataloader Exercise 04: Solver and Linear Regression Exercise 05: Neural Networks Exercise 06: Hyperparameter Tuning	Numpy (Reinvent the wheel)
Exercise 07: Introduction to Pytorch Exercise 08: Autoencoder	Pytorch/Tensorboard
Exercise 09: Convolutional Neural Networks Exercise 10: Semantic Segmentation Exercise 11: Recurrent Neural Networks	Applications (Hands-off)

Exercises – 9 Submissions

* **Exercise 01: Organization** Exercise 02: Math Recap	Intro
* **Exercise 03: Dataset and Dataloader** * **Exercise 04: Solver and Linear Regression** * **Exercise 05: Neural Networks** * **Exercise 06: Hyperparameter Tuning**	Numpy (Reinvent the wheel)
Exercise 07: Introduction to Pytorch * **Exercise 08: Autoencoder**	Pytorch/Tensorboard
* **Exercise 09: Convolutional Neural Networks** * **Exercise 10: Semantic Segmentation** * **Exercise 11: Recurrent Neural Networks**	Applications (Hands-off)

Submission System

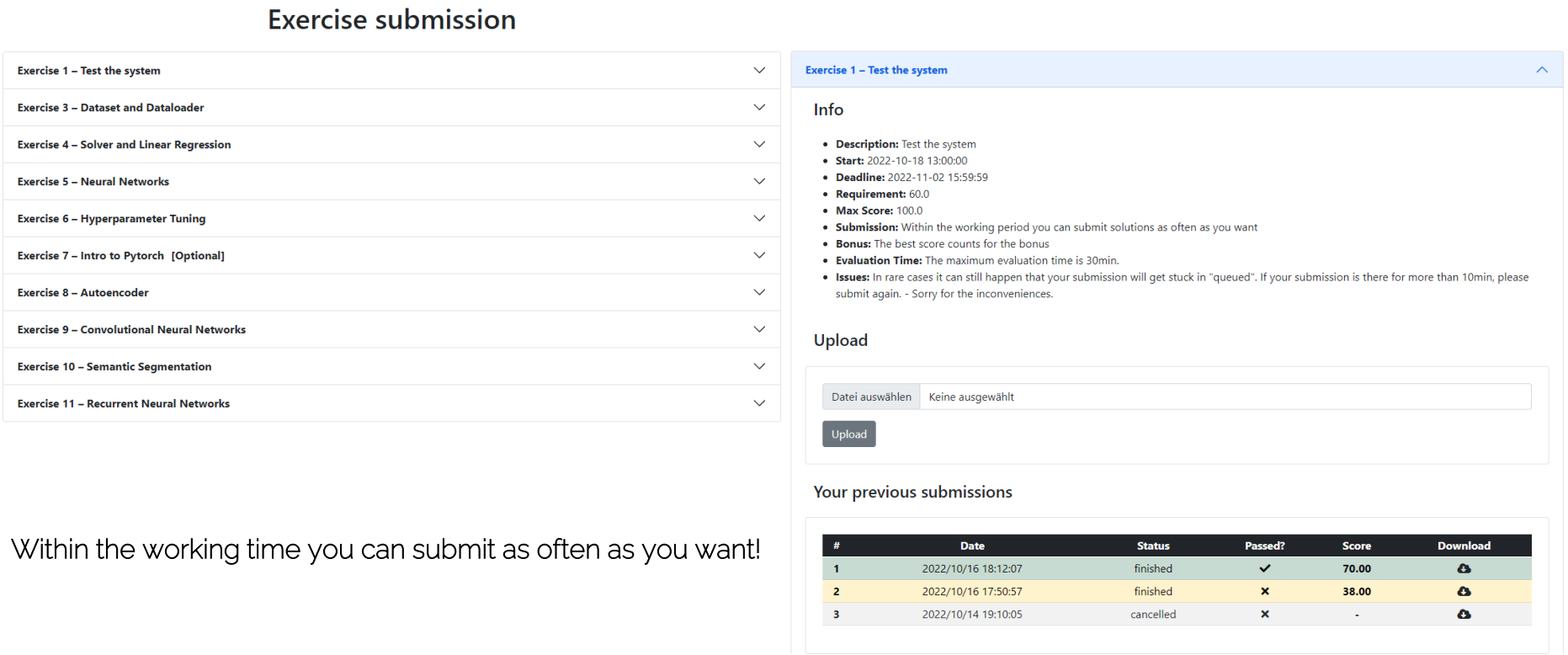


1: Register with your enrolled Matriculation Number

2: Login with your credentials you get via email

https://i2dl.vc.in.tum.de/

Submission System



Within the working time you can submit as often as you want!

Submission System

1

Bonus

Exercise 1	Exercise 3	Exercise 4	Exercise 5	Exercise 6	Exercise 8	Exercise 9	Exercise 10	Exercise 11
✓ (70.00)	✓ (90.00)	✓ (100.00)	✓ (100.00)	✓ (52.02)	✓ (78.00)	× ()	✓ (64.17)	✓ (84.38)

This table gives an overview over your current status regarding the grade bonus.
For each exercise only the best submission is displayed.
To be eligible for the grade bonus of 0.3 (on an already passed exam), you need to pass 8 of the 9 non-optional submissions.

2

Leaderboard

The leaderboard shows for each exercise the highest scoring submission from each user. Only valid submissions are displayed.

Exercise 1 | Exercise 3 | Exercise 4 | Exercise 5 | Exercise 6 | Exercise 8 | Exercise 9 | Exercise 10 | Exercise 11

#	User	Score
1	a0001	90.00
2	a0006	90.00
3	a0007	90.00

Grade Bonus

- Solve 8 out of 9 submissions (bold ones)
  - Pass the required score

- Grade bonus:
  - 0.3 improvement on passed exam
  - Can be transferred from previous and future semesters

* **Exercise 01: Organization**
Exercise 02: Math Recap

* **Exercise 03: Dataset and Dataloader**
* **Exercise 04: Solver and Linear Regression**
* **Exercise 05: Neural Networks**
* **Exercise 06: Hyperparameter Tuning**

Exercise 07: Introduction to Pytorch
* **Exercise 08: Autoencoder**

* **Exercise 09: Convolutional Neural Networks**
* **Exercise 10: Semantic Segmentation**
* **Exercise 11: Recurrent Neural Networks**

Exercise 01 – Overview

- **Starting time:** Today, after the lecture (16:00)
- **Deadline:** Wednesday 02/11, 15:59 (in 2 weeks)

- Content
  - Get familiar with the exercise structure
  - Jupyter Notebook & Python setup (local / Google Colab)
  - Introduction to the submission system
  - Submission:
    - Implement 1 line of code

Upcoming Lecture

- No Lecture next week
- Next Lecture: Lecture 2: Machine Learning basics
- From today: Start of exercise 1

See you next time ☺