

Introduction to Deep Learning



Lecturer





Prof. Angela Dai



Student Tutors



Alexandra Samoylova



Han Wu



Ivan Vasilev



Shaotong Chen



Estevao Gomes



Ignacio Dassori







• Al Methods

.....

- Broad definition!
 - "if" statements"
 if (cold) turn_up_heat();
 - Binary Search
 - Dijkstra, A*, ...
 - Prime, Kurskal, ...
 - Logic algorithms, etc.



- ML Methods
 - Linear/logistic regression
 - Support Vector Machines
 - Random Decision
 Trees, Forests,
 Jungles, ...



- Deep Learning
 - ML-methods leveraging neural networks
 - Multi-layer perceptrons
 - Convolutional neural networks
 - Recurrent neural networks
 - Transformers
 - Generative models, etc...

What do you see?



What does a machine see?





I2DL: Prof. Dai

What does a machine see?



Q: Is the door of the truck cab open?



Q: In this image, how many eyes can you see on the animal?



Q: How many wheels can you see in the image?

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 insights into the primate visual system

• Nobel prize 1981



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.

uly	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...





Image Classification



Image Classification





Why Deep Learning?

I2DL: Prof. Dai

Deep Learning History





S. McCulloch - W. Pitts



· Adjustable Weights · Weights are not Learned



B. Widrow - M. Hoff F. Rosenblatt



· Learnable Weights and Threshold



M. Minsky - S. Papert



XOR Problem



D. Rumelhart - G. Hinton - R. Wiliams



· Solution to nonlinearly separable problems · Limitations of learning prior knowledge · Big computation, local optima and overfitting · Kernel function: Human Intervention





G. Hinton - S. Ruslan

	-@			
-1	-	-	-	
Price 00000	M0000008	~0000000	~	
-1	-	-		
P=0000000	N-0000000	* 0000000	Personal -	
-1	-	-1	-	
1-0000000	-	**D000000	1-000000	
*	-1	-	-	
· 	·00008	·00000	. 00000	
(a) First fidden layer pro-	(h) Second hidden layer pre-training	sci Third hidden layer pro	(d) Fine-tuning of who optimized	

· Hierarchical feature Learning

The Empire strikes Back



ILSVRC top-5 error on ImageNet

What has Changed?





MNIST digit recognition dataset 10⁷ pixels used in training

2012 Krizhevsky et al.



ImageNet image recognition dataset 10¹⁴ pixels used in training

What Made this Possible?



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



Object Detection









AlphaGo

ever punch a cactus?



Emoticon suggestion

English – detected -		German -	\Box		
Deep Learni	ing rocks	Deep Learning Felsen			
Machine translation					



Alpha Star



Google LaMDA (Google 10'22)

• Large Language Models (LLMs), e.g., Chat-GPT



Tell me about deep learning and how I can teach it at TUM.



Deep learning is a subset of machine learning that involves training artificial neural networks \bigcirc \bigcirc to recognize patterns and make predictions or decisions based on input data. The term "deep" refers to the fact that these networks have many layers, which allows them to learn increasingly complex representations of the data as it moves through the network.

To teach deep learning at TUM, you could start with an introductory course that covers the basics of machine learning and neural networks. This could include topics such as:

1. Supervised learning: This involves training a model using labeled examples, so that it can make predictions on new, unseen data.

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

 \rightarrow

DALL-E 2





Dall-E 2 Image generation from text



"a robot paíntíng a colorful sunset on a canvas ín a futurístíc art studío"



"a majestic dragon made of water flying through a colorful, abstract city"

I2DL: Prof. Dai

https://huggingface.co/black-forest-labs/FLUX.1-dev



StableDiffusion Image Outpainting
Deep Learning Today





Deep Learning Today





I2DL: Prof. Dai

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



What society thinks I do



What my friends think I do



What other computer scientists think I do



What mathematicians think I do



What I think I do

from theano import

What I actually do





Deep Learning Memes Types of Headaches

Migraine









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, machine learning and robotics
- 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 many more 🕲



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

FlowNet P. Fischer, A. Dosovitskiy, E. Ilg, P. Häusser, C. Hazırbas, V. Golkov, P. v.d. Smagt, D. Cremers, T. Brox



Multiple object tracking with graph neural networks



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



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

Our method generates temporally coherent volumetric data with a trained CNN generator.

[Xie et al. Siggraph' 18] tempoGAN

Animation Synthesis



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

Animation Synthesis



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



[Kirchstein et al., ToG'23] Nersemble



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



ScanNet++: -1006 high-fidelity Scenes

- 1mm laser scans
- -DSLR images -iPhone RGB-D
- -semantics

[Yeshwanth & Liu et al., ICCV'23] ScanNet++





2DL: Prof. Dai

[Bokhovkin et al., CVPR'25] SceneFactor





[Bokhovkin et al., CVPR'25] SceneFactor



[Siddiqui et al., CVPR'24] MeshGPT



Relation to other lectures at TUM







Introduction to Deep Learning

Logistics

About the Lecture

- Course webpage: <u>https://www.3dunderstanding.org/i2dl-s25</u>
- Theory lectures (every Monday at 14:15)
 - In-person, live-streamed
- Tutorials and exercises (every Thursday at 10:00)
 - Tutorial: Online videos posted to Piazza and the webpage
 - Practical exercises
- Guest Lecture 🕲

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

Moodle → Piazza

- Announcements via Piazza IMPORTANT!
 - Sign up online for access: <u>http://piazza.com/tum.de</u>
 - Select "Summer 2025" 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 questions visible only to the staff





• 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





Home Publications Teaching Openings Team Contact

Introduction to Deep Learning (I2DL) (IN2346)



Welcome to the Introduction to Deep Learning course offered in SS25.

https://www.3dunderstanding.org/i2dl-s25

(Virtual) Office Hours

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

- More info in the first tutorial session
- Zoom links will be posted on Piazza

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" programming 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
- Again:
 - Check announcements on Piazza
 - Check content on website:

https://www.3dunderstanding.org/i2dl-s25



Practical Exercises

Fxercise – Goal

- Hands-on programming experience (learning by doing)
- Reimplement 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 Piazza and course website

- Video length Exo2: Full lecture with written exercise Exo3-12: Short (~30min) video and
 - coding exercises

- Programming:
 - Interactive coding notebooks (~4h each)
- Start time: Thursdays 10:00
- Working Time: 1-2 weeks
- Deadline: Wednesdays, 23:59

Upcoming Lecture

• Next Lecture: Lecture 2: Machine Learning basics

• Thursday: Tutorial 1 and Exercise 1



See you next time 🕲