

# Seminar Vision Systems MA-INF 4208

---

10.07.2024

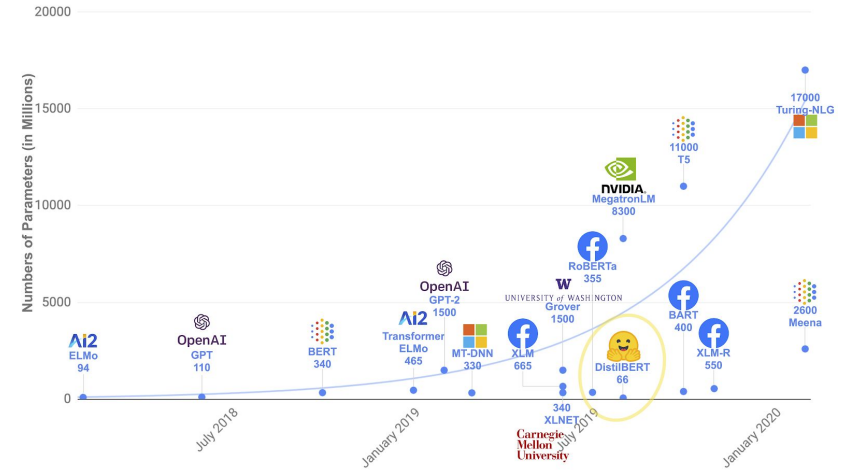
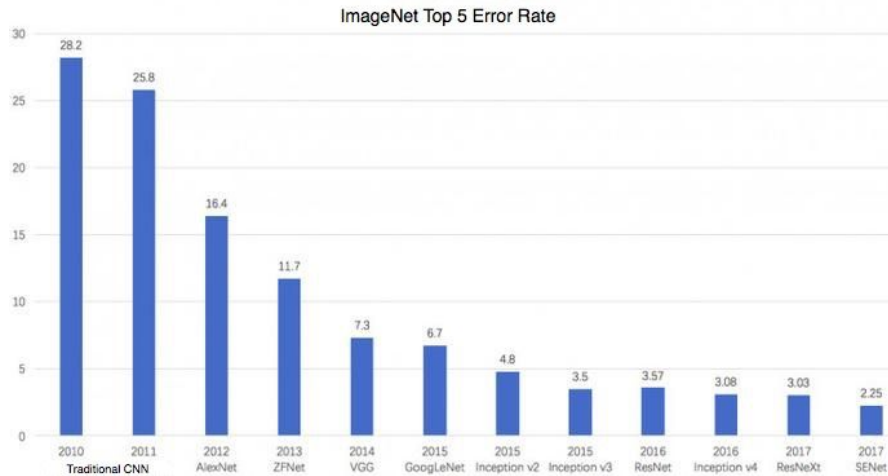
PROF. SVEN BEHNKE, ANGEL VILLAR-CORRALES

Contact: [villar@ais.uni-bonn.de](mailto:villar@ais.uni-bonn.de)

# The Age of Deep Learning



# The Age of Deep Learning



# The Age of Deep Learning

---



HUGGING FACE

Google



DAIMLER

amazon

SIEMENS



TOYOTA  
RESEARCH INSTITUTE



TESLA



Microsoft



# In this seminar...

---

- Acquire/improve ability to:
  - deal with scientific publications (e.g. papers)
  - write a scientific report
  - present a scientific topic to an audience
  - engage technical topics

 Important skills for Master Thesis!

# In this seminar

---

- Discuss trending topics in deep learning and computer vision
- We will cover the following topics
  - Foundation Models and Scene Understanding
  - Advances in Neural Networks
  - Neural Rendering and Applications
  - Unsupervised Learning from Objects and Videos

## Seminar: Vision Systems MA-INF 4208

Prof. Dr. Sven Behnke, Angel Villar-Corrales

### 1 Paper List

#### 1. Foundation and Generative Models

- a) Menapace, Willi, et al. *Promptable Game Models: Text-Guided Game Simulation via Masked Diffusion Models*. Transactions on Graphics. 2024. [Link](#)
- b) Oquab, Maxime, et al. *DINOv2: Learning Robust Visual Features without Supervision*. ArXiv PrePrint. 2023. [Link](#)
- c) Zhou, Hongyu, et al. *HUGS: Holistic Urban 3D Scene Understanding via Gaussian Splatting*. ArXiv PrePrint. 2024. [Link](#)

#### 2. Advances in Network Architectures and Learning Algorithms

- a) Liu, Ziming, et al. *KAN: Kolmogorov-Arnold Networks*. ArXiv Preprint. 2024. [Link](#)
- b) Beck, Maximilian, et al. *sLSTM: Extended Long Short-Term Memory*. ArXiv PrePrint. 2024. [Link](#)
- c) Darcet, Timothée et al. *Vision Transformers Need Registers*. ICLR. 2024. [Link](#)

#### 3. Neural Rendering and Applications

- a) Luiten, Jonathan, et al. *Dynamic 3D Gaussians: Tracking by Persistent Dynamic View Synthesis*. 3DV. 2024. [Link](#)
- b) Kim, Chang, et al. *GARField: Group Anything with Radiance Fields*. CVPR 2024. [Link](#)
- c) Li Zhengqi, et al. *Generative Image Dynamics*. CVPR. 2024. [Link](#)

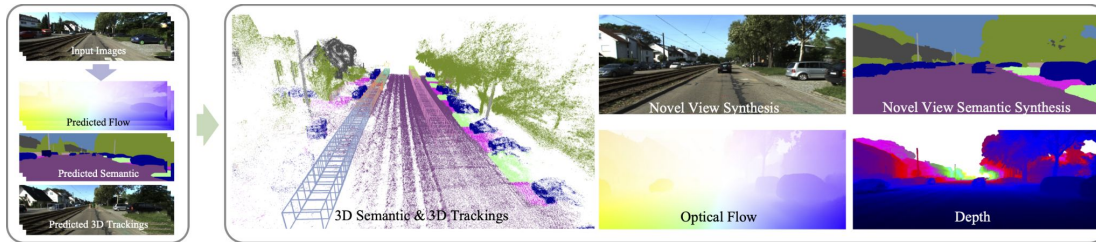
#### 4. Unsupervised Learning from Objects and Videos

- a) Ost, Julian, et al. *Inverse Neural Rendering for Explanable Multi-Object Tracking*. ArXiv PrePrint. 2024. [Link](#)
- b) Jabri, Allan et al *DORSal: Diffusion for Object-centric Representations of Scenes* ICLR 2024. [Link](#)
- c) Garrido, Quentin et al *Learning and Leveraging World Models in Visual Representation Learning*. ArXiv Preprint. 2024. [Link](#)

**Paper List:** <https://www.ais.uni-bonn.de/SS24/SeminarVision/PaperList.pdf>

# Foundation Models and Scene Understanding

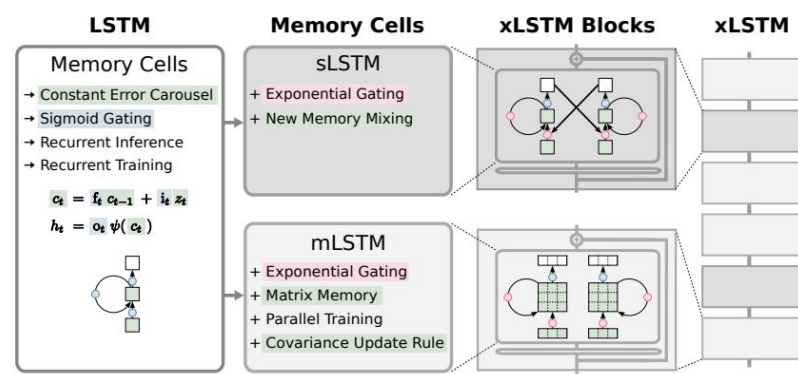
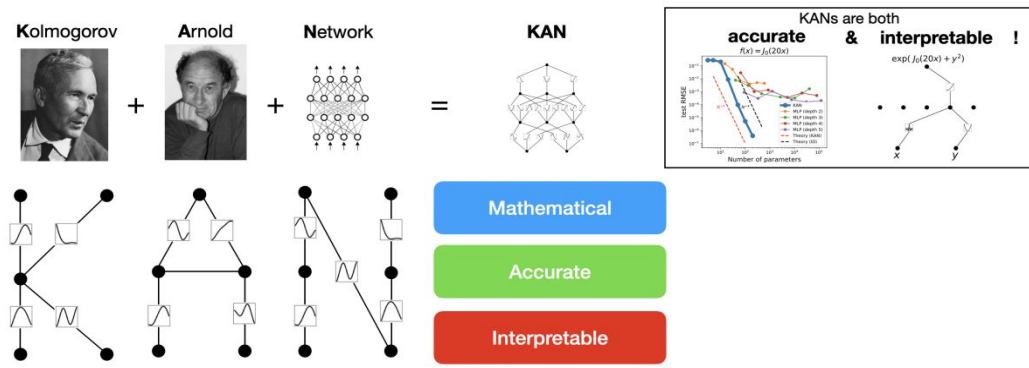
- Models trained on internet-scale data with broad generalization capabilities
- Applications such as:
  - Playable Video Generation
  - Representation Learning
  - 3D Understanding



The player jumps to the right and sends the ball to the no man's land with a forehand

# Advances in Deep Learning Models

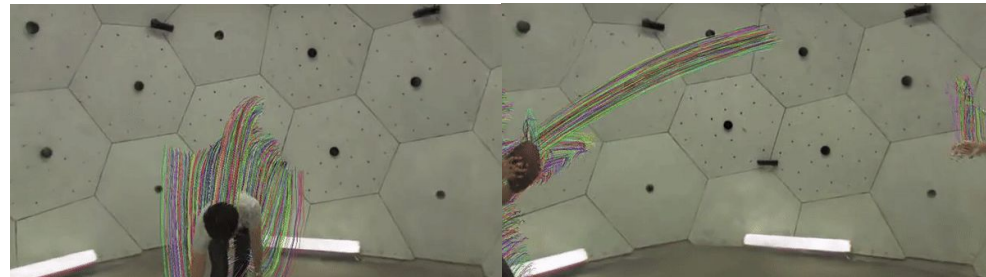
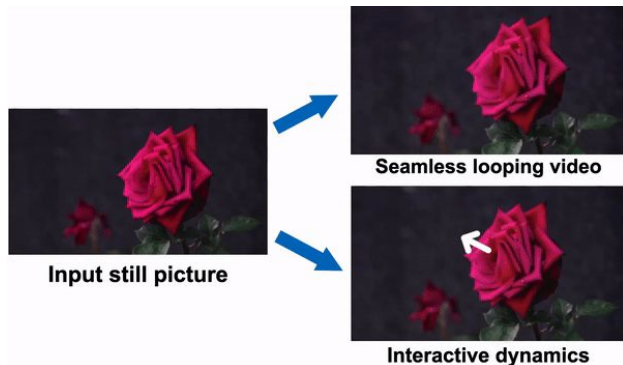
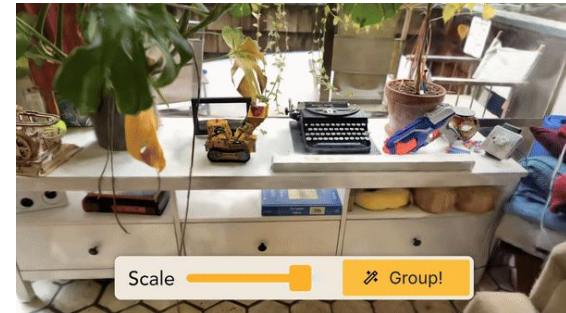
- Novel neural network architectures
  - Kolmogorov-Arnold Networks
  - xLSTM
- Improved training techniques and understanding
  - Vision Transformers need Registers





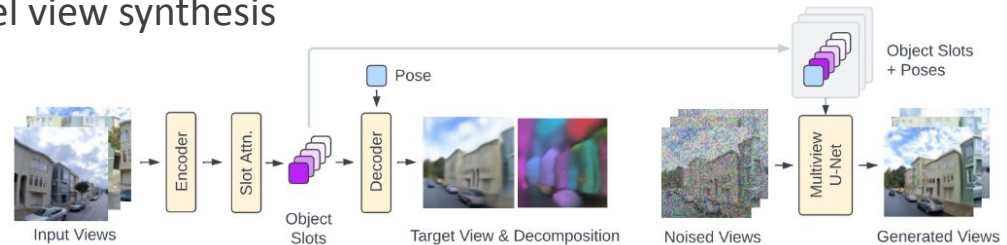
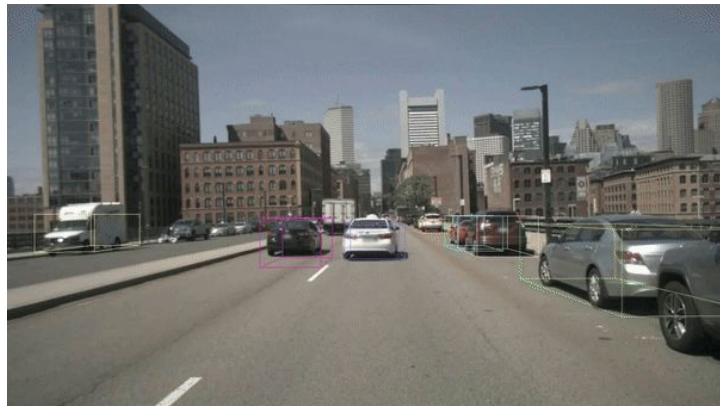
# Neural Rendering and Applications

- Learning 3D-aware representations of a scene given a set of posed images
- Applications such as:
  - Spatio-temporal scene decomposition
  - Generative dynamics
  - Novel-view Synthesis and 3D tracking



# Unsupervised Learning from Videos

- Learning representations from video data without annotations
- Applications such as:
  - Unsupervised 3D object detection & tracking
  - Object-centric learning and novel view synthesis



# Get a Spot and Select your Topic

---


- Fill the following [form](#) before 15.07.2024
  - Your name & email
  - Matriculation number
  - Your three preferred papers
- Based on this form, I will and assign papers
- Upon my confirmation:
  - Register in BASIS
  - Start working on your paper

**BASIS Registration opened until 31.07.2023!**

# Deliverables (preliminary dates)

---

- **Presentation:** Monday 23.09.2024
  - 30 min presentation
  - 15 min discussion
- **Report:** Monday 30.09.2024 (will be one week after presentations)
  - LaTeX template
  - 8-12 pages
  - Brief but readable and informative
  - BibTex citations

 Arrange a meeting with me  $\approx$ 2 weeks before the presentation to check the preliminary materials for the presentation and report.

# Report

---

- Well structured:
  - Abstract
  - Introduction, methods, results, conclusion, ...
  - Tables and figures
  - Correct citations
- Your own scientific opinion:
  - What are the weak points of the paper?
  - What is missing?
  - Are comparisons fair and believable?
  - Possible future steps?

**We don't want a copy of the paper!**

# Grading

---

- 66.7%: Presentation
  - Quality of the presentation slides
  - Presentation skills
  - Ability to answer questions
- 33.3%: Report
  - Overall quality of the report
  - Critical thinking and own discussion
  - Understanding of the concept

# Seminar Alternative

## Seminar Cognitive Robotics: [Link](#)

- Same seminar format
- Papers more robotics related:
  - Grasping and Manipulation
  - Robot perception
  - SLAM
  - Planning and Navigation

➤ Introductory meeting on **12.07.2024**



UNIVERSITÄT BONN AIS

**Institute for Computer Science VI**  
**Autonomous Intelligent Systems**

Home	Seminar Cognitive Robotics (MA-INF 4211)
Persons	Lab Cognitive Robotics (MA-INF 4304) Projektgruppe Kognitive Robotik (BA-INF 051)
Teaching	
Research	Prof. Dr. Sven Behnke, Dr. Raphael Memmesheimer
Publications	First organizational meeting: 12.07.2024 in room 0.107 (Friedrich-Hirzebruch-Allee 6) Registration of interest till 18.07.2024
News	BASIS registration: after the topic association
Jobs	Seminar presentations on TBD Room: 20.09.2024

Cognitive robotics is an active research area at the border between artificial intelligence and robotics. It investigates and tries to implement in technical systems mental functions, which are associated with intelligence. This covers the perception of the environment, action planning, and learning.

In this seminar, we will cover research papers from the area of cognitive robotics. Details will be announced in the organizational meeting.

Some of our robots:





University of Bonn, Institute for Computer Science, Computer Science VI - Intelligent Systems and Robotics | Impressum | Data Privacy Statement

Questions?

