UNIVERSITÄT BONN AIS

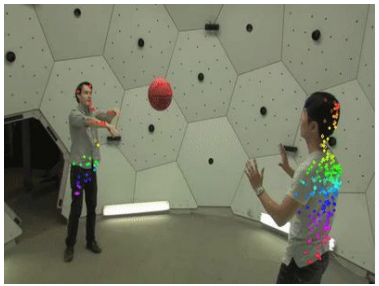
Seminar Vision Systems MA-INF 4208

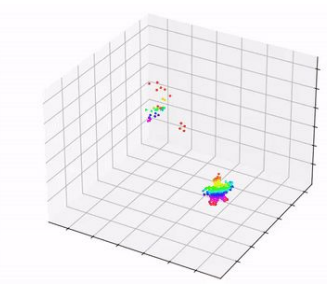
07.02.2025

PROF. SVEN BEHNKE, ANGEL VILLAR-CORRALES

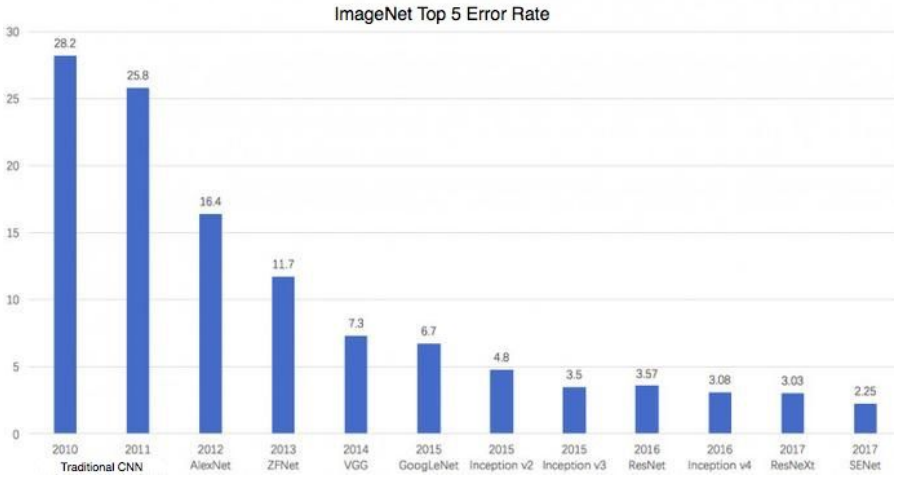
Contact: villar@ais.uni-bonn.de

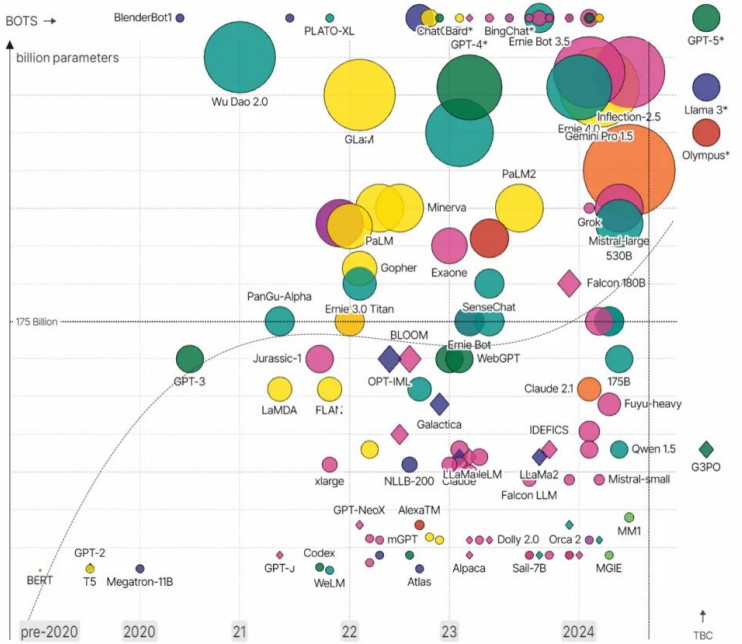
The Age of Deep Learning





The Age of Deep Learning

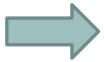




The Age of Deep Learning

In this seminar...

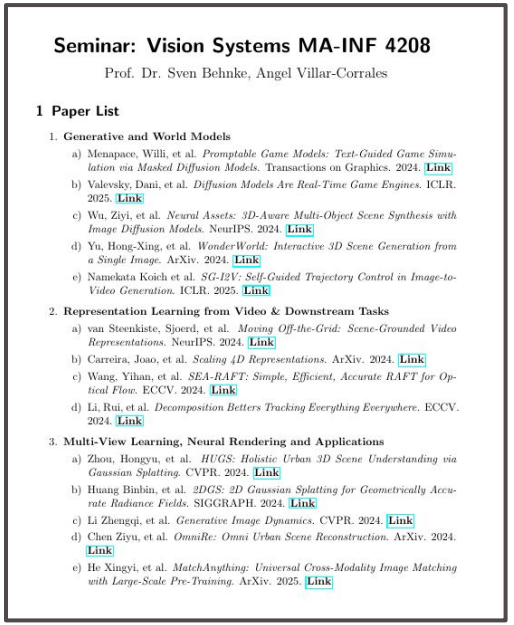
- Acquire/improve ability to:
  - deal with scientific publications (e.g. papers)
  - write a scientific report
  - present a scientific topic to an audience
  - engage technical topics



Important skills for Master Thesis!

In this seminar

- Discuss trending topics in deep learning and computer vision
- We will cover the following topics
  - Generative and World Models
  - Representation Learning from Video & Downstream Tasks
  - Multi-View Learning, Neural Rendering and Applications



Seminar: Vision Systems MA-INF 4208

Prof. Dr. Sven Behnke, Angel Villar-Corrales

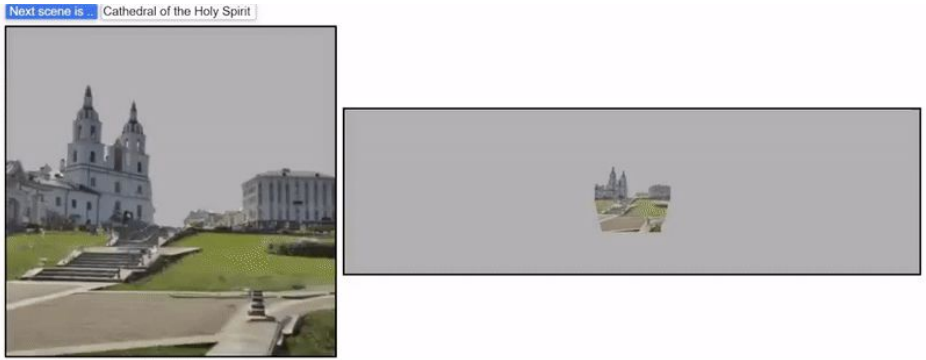
1 Paper List

1. **Generative and World Models**
   a) Menapace, Willi, et al. *Promptable Game Models: Text-Guided Game Simulation via Masked Diffusion Models.* Transactions on Graphics. 2024. Link
   b) Valevsky, Dani, et al. *Diffusion Models Are Real-Time Game Engines.* ICLR. 2025. Link
   c) Wu, Ziyi, et al. *Neural Assets: 3D-Aware Multi-Object Scene Synthesis with Image Diffusion Models.* NeurIPS. 2024. Link
   d) Yu, Hong-Xing, et al. *WonderWorld: Interactive 3D Scene Generation from a Single Image.* ArXiv. 2024. Link
   e) Namekata Koich et al. *SG-I2V: Self-Guided Trajectory Control in Image-to-Video Generation.* ICLR. 2025. Link
2. **Representation Learning from Video & Downstream Tasks**
   a) van Steenkiste, Sjoerd, et al. *Moving Off-the-Grid: Scene-Grounded Video Representations.* NeurIPS. 2024. Link
   b) Carreira, Joao, et al. *Scaling 4D Representations.* ArXiv. 2024. Link
   c) Wang, Yihan, et al. *SEA-RAFT: Simple, Efficient, Accurate RAFT for Optical Flow.* ECCV. 2024. Link
   d) Li, Rui, et al. *Decomposition Betters Tracking Everything Everywhere.* ECCV. 2024. Link
3. **Multi-View Learning, Neural Rendering and Applications**
   a) Zhou, Hongyu, et al. *HUGS: Holistic Urban 3D Scene Understanding via Gaussian Splatting.* CVPR. 2024. Link
   b) Huang Binbin, et al. *2DGS: 2D Gaussian Splatting for Geometrically Accurate Radiance Fields.* SIGGRAPH. 2024. Link
   c) Li Zhengqi, et al. *Generative Image Dynamics.* CVPR. 2024. Link
   d) Chen Ziyu, et al. *OmniRe: Omni Urban Scene Reconstruction.* ArXiv. 2024. Link
   e) He Xingyi, et al. *MatchAnything: Universal Cross-Modality Image Matching with Large-Scale Pre-Training.* ArXiv. 2025. Link

Paper List: https://www.ais.uni-bonn.de/WS/SeminarVision/PaperList.pdf

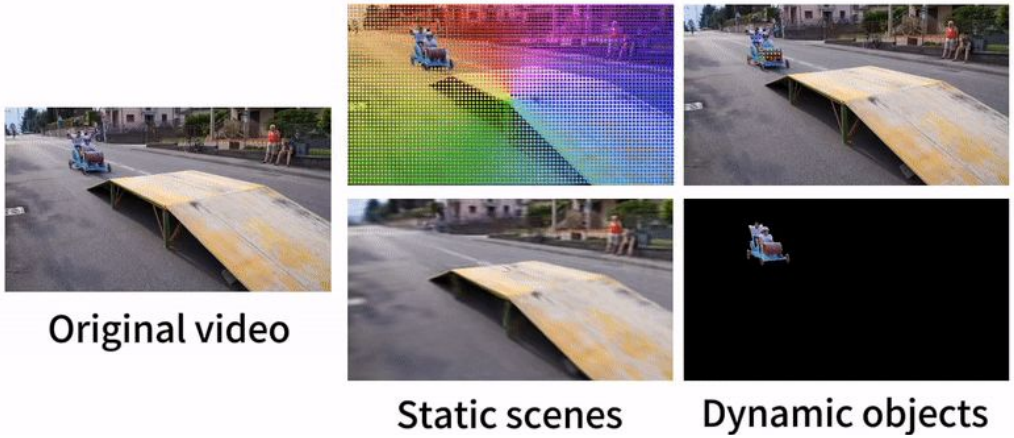
Generative and World Models

- Models trained on large-scale data with broad generalization and controllability
- Applications such as:
  - Playable Video Generation
  - Generating new objects and scenes
  - Control in Image-to-Video Generation



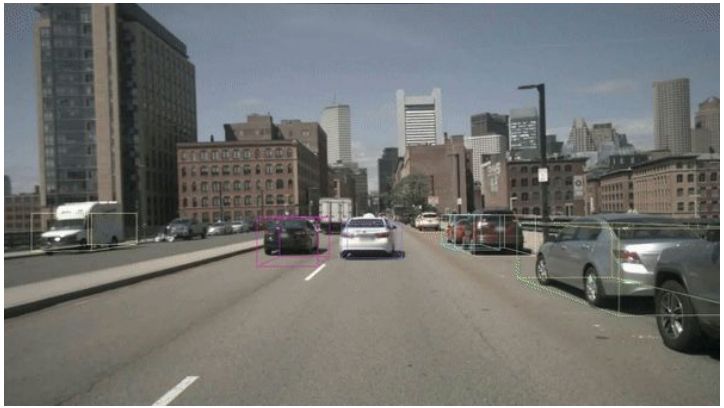
Learning from Videos & Downstream Tasks

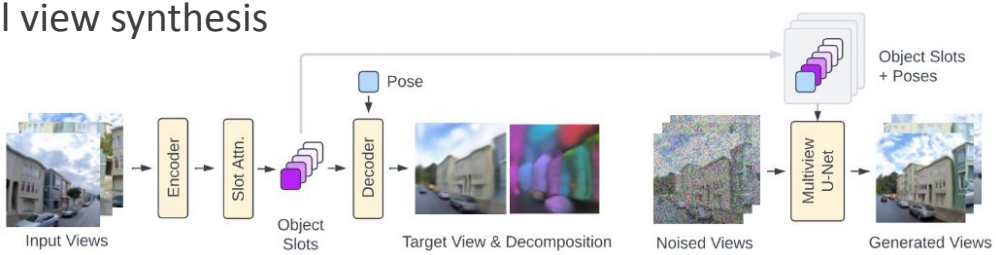
- Learning representations from video data without annotations
- Applications such as:
  - Representation learning
  - Decomposition and Tracking



Unsupervised Learning from Videos

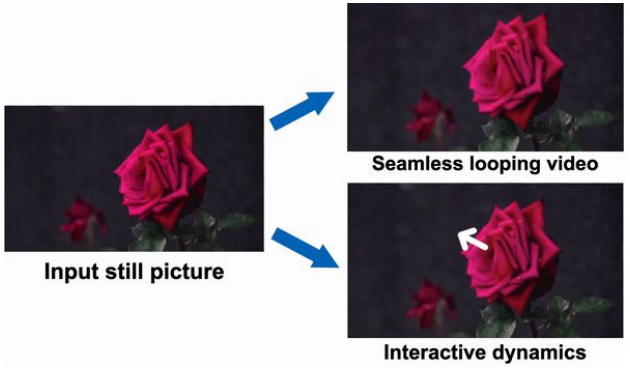
- Learning representations from video data without annotations
- Applications such as:
  - Unsupervised 3D object detection & tracking
  - Object-centric learning and novel view synthesis





Multi-View Learning & Neural Rendering

- Learning 3D-aware representations of a scene given a set of posed images
- Applications such as:
  - Generative dynamics
  - Scene reconstruction and editing
  - Novel-view Synthesis



Get a Spot and Select your Topic

- Fill the following **form** no later than 10.02.2025
  - Your name & email
  - Matriculation number
  - Your three preferred papers
- Based on this form, I will and assign seminar spots and papers to review on Monday 11.02.2025
- Upon my confirmation:
  - Register in BASIS
  - Start working on your paper

BASIS Registration opened until 20.02.2025!

Deliverables (preliminary dates)

- **Presentation**: Thursday 27.03.2025
  - 30 min presentation
  - 15 min discussion

- **Report**: Thursday 03.04.2025 (will be one week after presentations)
  - LaTeX template
  - 8-12 pages
  - Brief but readable and informative
  - BibTex citations

⇨ Arrange a meeting with me ≈1 week 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

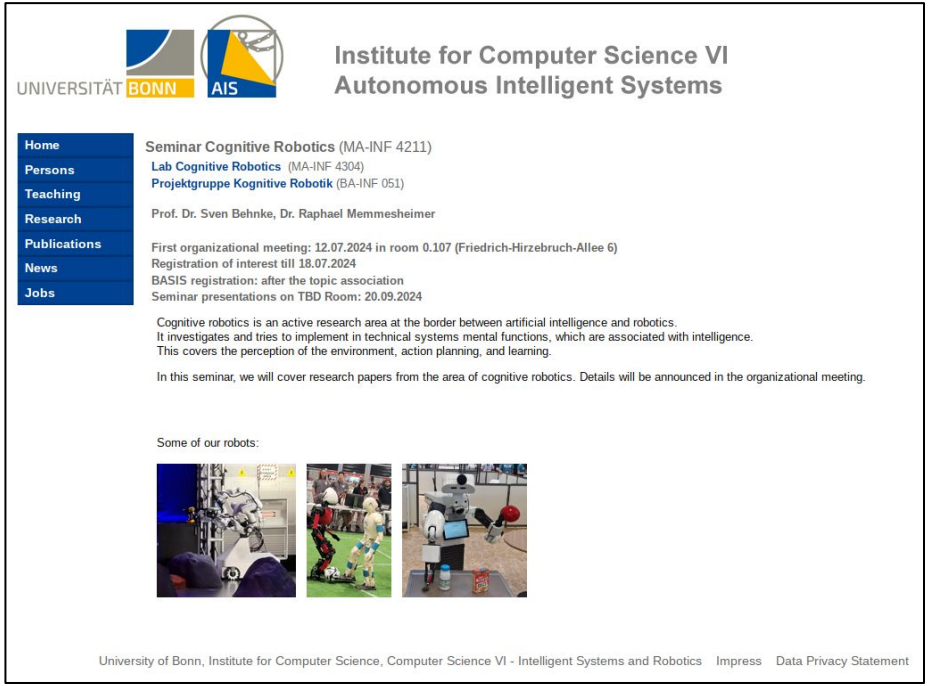
- 33.3%: Report
  - Overall quality of the report
  - Critical thinking and own discussion
  - Understanding of the concept

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**



Questions?