

Seminar Vision Systems MA-INF 4208

23.10.2026

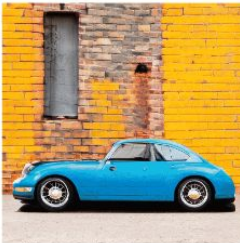
PROF. SVEN BEHNKE, LUCA EICHLER

Contact: eichler@ais.uni-bonn.de

(Beyond) The age of Deep Learning



A living room with a fireplace at a wood cabin. Interior design.



a blue Porsche 356 parked in front of a yellow brick wall.

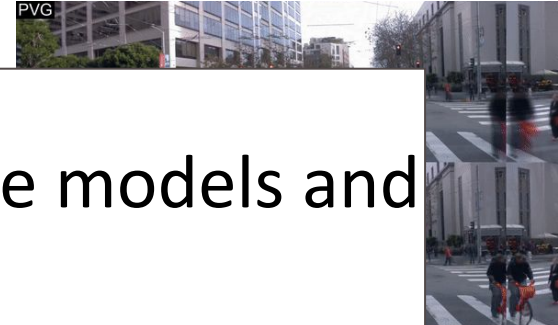


Eiffel Tower, landscape photography



The age of...

...foundation models, vision-language models and multimodal systems



A living room with a fireplace at a wood cabin. Interior design.



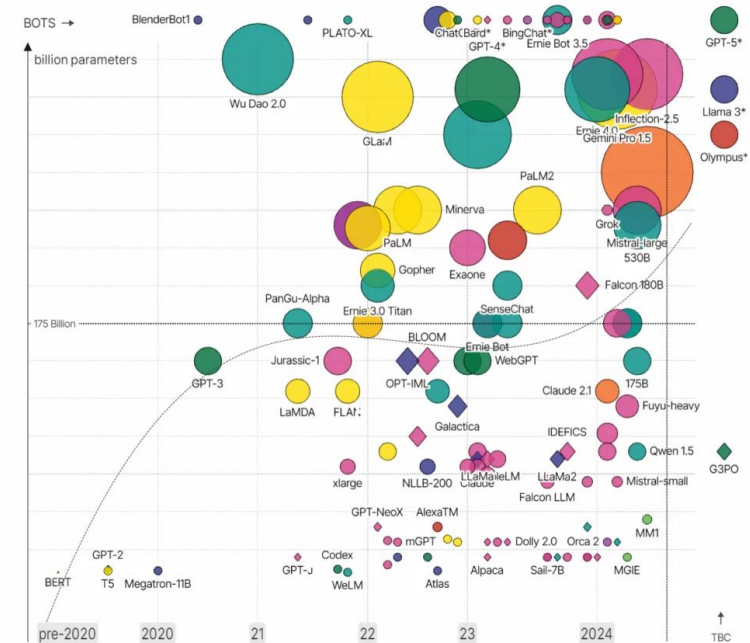
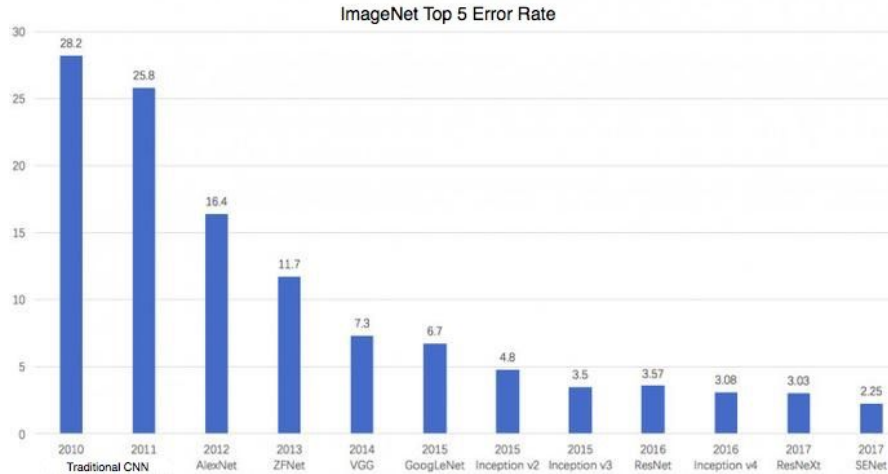
a blue Porsche 356 parked in front of a yellow brick wall.



Eiffel Tower, landscape photography



(Beyond) The age of Deep Learning



(Beyond) 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
 - Object Detection & Instance Segmentation
 - Multimodal Large Language Models (MLLMs)
 - World Models

Seminar: Vision Systems MA-INF 4208

Prof. Dr. Sven Behnke, Luca Eichler

1 Paper List

1. Object Detection & Instance Segmentation

- a) Liu, Shilong, et al. *Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection*. ECCV 2024. [Link](#)
- b) Robinson, Isaac, et al. *RF-DETR: Neural Architecture Search for Real-Time Detection Transformers*. ICLR 2026. [Link](#)
- c) Carion, Nicolas, et al. *SAM 3: Segment Anything with Concepts*. ICLR 2026. [Link](#)

2. Multimodal Large Language Models (MLLMs)

- a) Zheng, Duo, et al. *Learning from Videos for 3D World: Enhancing MLLMs with 3D Vision Geometry Priors*. NeurIPS 2025. [Link](#)
- b) Xu, Jiacong, et al. *Towards Zero-Shot Anomaly Detection and Reasoning with Multimodal Large Language Models*. CVPR 2025. [Link](#)
- c) Zhang, Jianke, et al. *VLM4VLA: Revisiting Vision-Language-Models in Vision-Language-Action Models*. ICLR 2026. [Link](#)

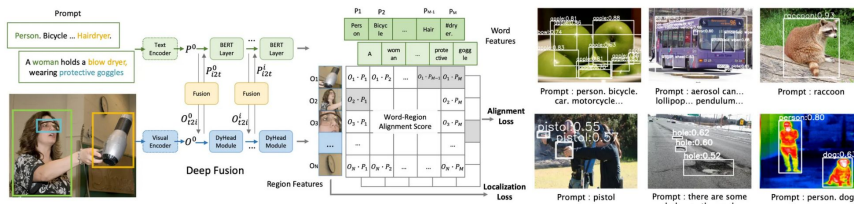
3. World Models

- a) Ye, Seunggyeom, et al. *World Action Models are Zero-shot Policies*. ICLR 2026 (?). [Link](#)
- b) Maes, Lucas, et al. *LeWorldModel: Stable End-to-End Joint-Embedding Predictive Architecture from Pixels*. 2026. [Link](#)
- c) Hafner, Danijar, et al. *Training Agents Inside of Scalable World Models*. 2025. [Link](#)

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

Object Detection & Instance Segmentation

- Given an image, extract pixels belonging to an object of interest
- Applications such as:
 - Robot perception
 - Image / Video editing
 - Manufacturing



Multimodal Large Language Models (MLLMs)

- Combine vision systems and language models
- Applications such as:
 - Image-Language Grounding
 - Anomaly Detection
 - Planning with VLAs

In the wild



Q: Do you notice any anomalies in the image?
A: Yes, the anomaly in the image is a fine, dark crack visible on the ceramic cup, starting from the rim and extending downward.
Q: How can I prevent similar anomalies in the future?
A: Avoid rapid temperature changes, handle the cup carefully to minimize impacts, and consider using cups made with thicker, more durable materials.

Industry

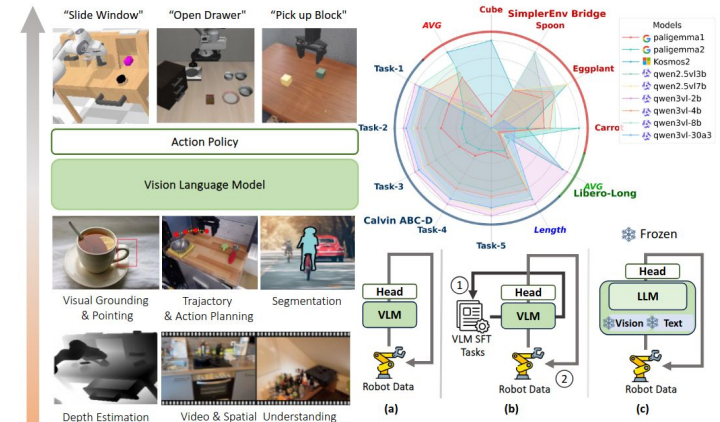


Q: Are there any anomalies in this cable image?
A: Yes, the green wire is missing its internal copper conductors, unlike the blue and brown wires, which contain complete sets of bundled copper strands.
Q: What could be the reason behind this anomaly?
A: The absence of copper in the green wire could be due to manufacturing defects such as failure to insert the strands or a mechanical issue where the strands were removed or misplaced during production or handling.

medical

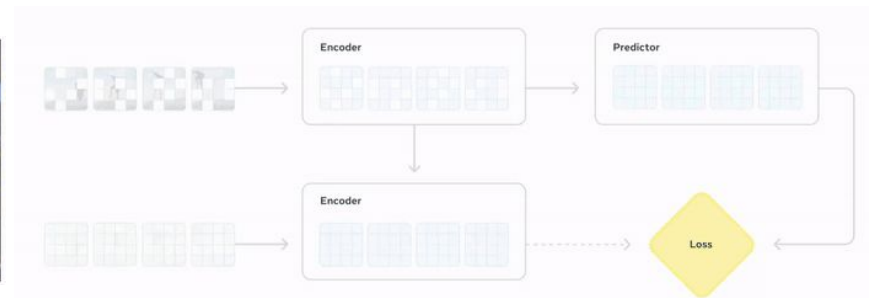
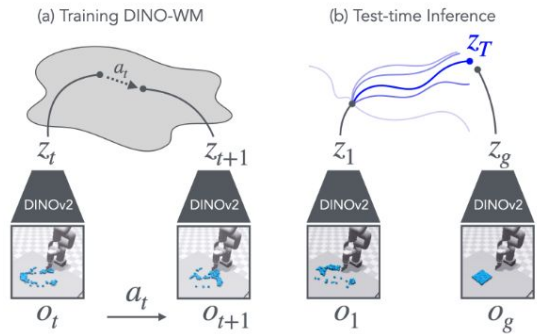


Q: What does the image show?
A: The image shows a densely cellular tissue sample with features indicating potential metastatic breast cancer.
Q: What suggests that there might be malignant activity in the tissue sample?
A: Malignant activity is suggested by irregularly shaped cells, hyperchromatic nuclei, and an increased nuclear-to-cytoplasmic ratio, indicative of abnormal and increased DNA content typical in cancerous cells.



World Models

- Models that learn to simulate environments and predict future possible outcomes
- Applications such as as:
 - Navigation
 - Planning
 - Controllable video generation



Get a Spot and Select your Topic

- Fill the following [form](#) no later than 25.10.2026
 - Your name & email
 - Matriculation number
 - Your three preferred papers
- Based on this form, I will assign seminar spots and papers to review on Monday 26.10.2026
- Upon my confirmation:
 - Register in BASIS
 - Start working on your paper

BASIS Registration opened until 31.10.2026

Deliverables (preliminary dates)

- **Presentation:** end of semester (26.03.2027)

- 30 min presentation
- 15 min discussion

- **Report:** end of semester (31.03.2027)

- 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!

My Tips

- Keep it simple:
 - You know more than your audience
 - Don't get lost in details
- Put the paper into context
 - What basic work does this paper build on? (explain if necessary)
 - Motivation
 - What exactly are the key contributions?

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

- Papers more robotics related:
 - Grasping and Manipulation
 - **Robot vision & perception**
 - SLAM
 - Planning and Navigation




UNIVERSITÄT **BONN** AIS

Institute for Computer Science VI
Autonomous Intelligent Systems

[Home](#)
[Persons](#)
[Teaching](#)
[Research](#)
[Publications](#)
[News](#)
[Jobs](#)

Seminar Cognitive Robotics (MA-INF 4211)

Prof. Dr. Sven Behnke, Dr. Raphael Memmesheimer

First organizational meeting: 25.07.2025 11:00 c.t., Friedrich-Hirzebruch-Allee 6, Room 0.107
 Registration of interest till 01.08.2025
 BASIS registration: 08.09.2025
 Seminar presentations on 26.09.2025 Room: Friedrich-Hirzebruch-Allee 6, Room 0.107

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 Imprint Data Privacy Statement

Questions?



