

Seminar Vision Systems MA-INF 4208

29.06.2023

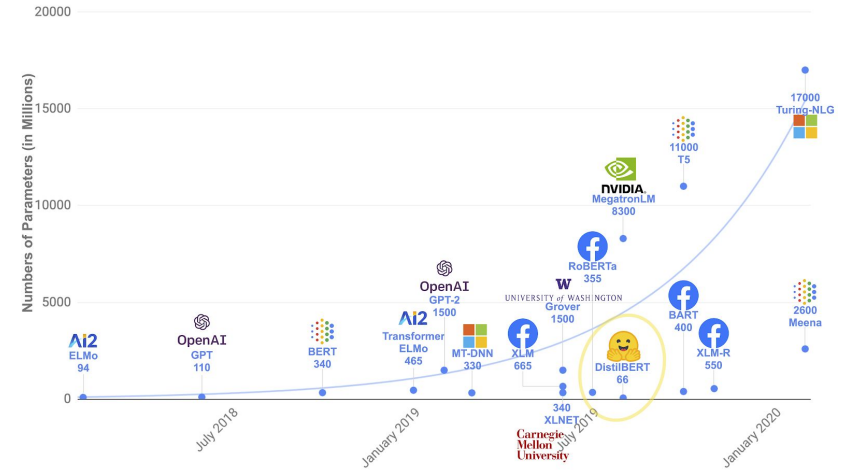
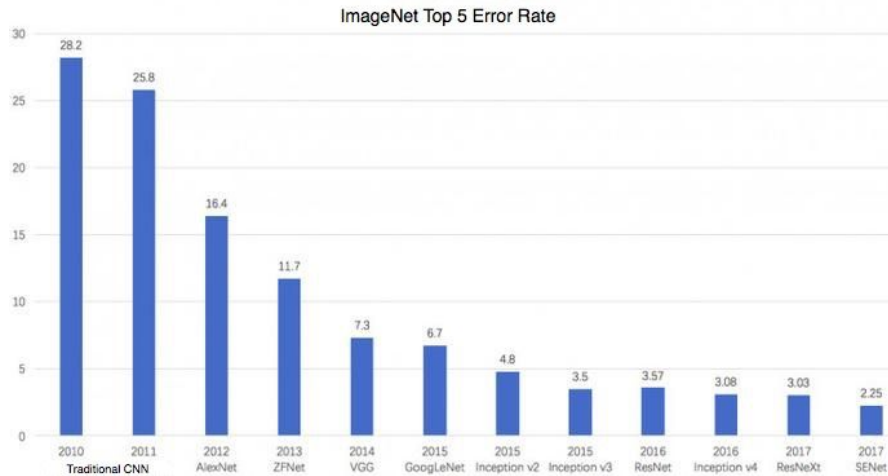
PROF. SVEN BEHNKE, ANGEL VILLAR-CORRALES

Contact: villar@ais.uni-bonn.de

The Age of Deep Learning



The Age of Deep Learning



The Age of Deep Learning



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
 - Neural Rendering and Applications
 - Self-Supervised Learning
 - Transformer-based Models
 - Video Decomposition and Tracking

Seminar: Vision Systems MA-INF 4208

Prof. Dr. Sven Behnke, Angel Villar-Corrales

1 Paper List

1. Neural Rendering and Applications

- a) Sajjadi, Mehdi SM, et al. *Object scene representation transformer*. NeurIPS 2022. [Link](#)
- b) Driess, Danny, et al. *Learning Multi-Object Dynamics with Compositional Neural Radiance Fields*. CoRL 2023. [Link](#)
- c) Siddiqui, Yawar et al. *Panoptic Lifting for 3D Scene Understanding with Neural Fields*. CVPR 2023. [Link](#)

2. Advances in Self-Supervised Learning

- a) Bardes, Adrien, et al. *VICReg: Variance-Invariance-Covariance Regularization for Self-Supervised Learning*. ICLR 2022. [Link](#)
- b) Ziegler, Adrian et al. *Self-Supervised Learning of Object Parts for Semantic Segmentation*. CVPR 2022. [Link](#)
- c) Assran, Mahmoud et al. *Self-Supervised Learning from Images with a Joint-Embedding Predictive Architecture*. ArXiv PrePrint 2023. [Link](#)

3. Transformer-based Model Architectures

- a) Hatamizadeh, Ali, et al. *FasterViT: Fast Vision Transformers with Hierarchical Attention*. ArXiv PrePrint 2023. [Link](#)
- b) Beyer, Lucas, et al. *FlexiViT: One Model for All Patch Sizes*. CVPR 2023. [Link](#)
- c) Dao, Tri, et al. *FlashAttention: Fast and Memory-Efficient Exact Attention with IO-Awareness*. NeurIPS 2022. [Link](#)

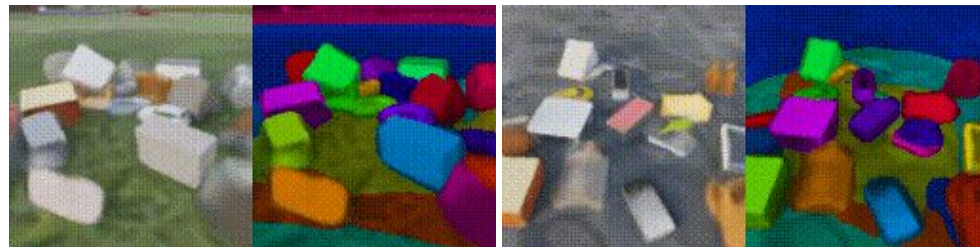
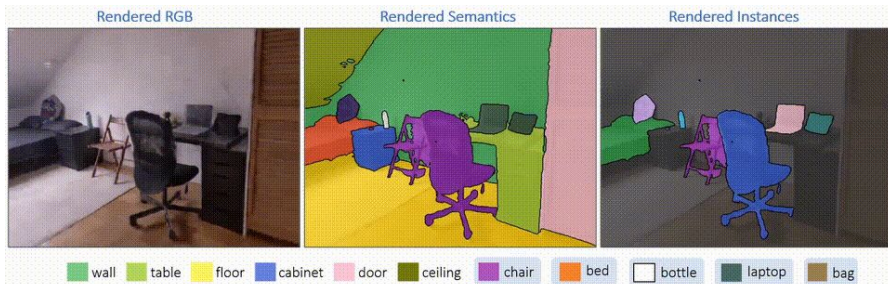
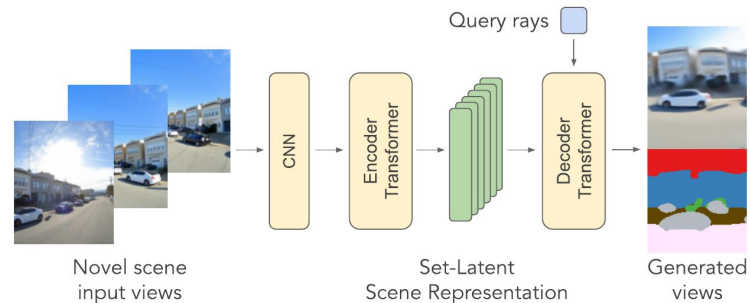
4. Video Decomposition and Tracking

- a) Ye, Vickie, et al. *Deformable Sprites for Unsupervised Video Decomposition*. CVPR 2022. [Link](#)
- b) Wang, Qianqian, et al. *Tracking Everything Everywhere All at Once*. ArXiv Preprint 2023. [Link](#)
- c) Bao, Zhipeng, et al. *Object Discovery from Motion-Guided Tokens*. CVPR 2023. [Link](#)

Paper List: <https://www.ais.uni-bonn.de/SS23/SeminarVision/PaperList.pdf>

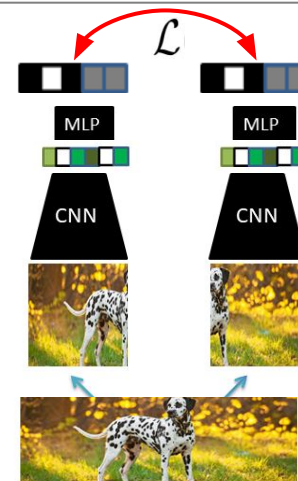
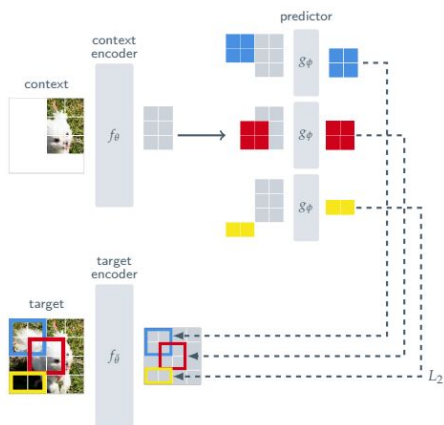
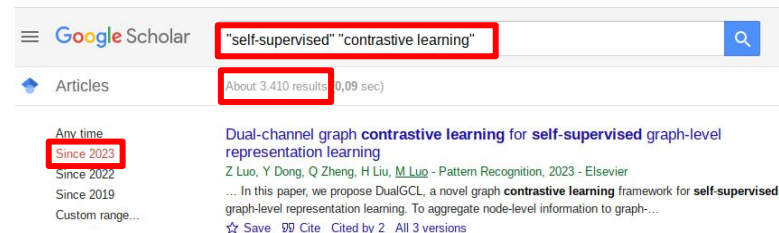
Neural Rendering and Applications

- Learning 3D-aware representations of a scene given a set of posed images
- Applications such as:
 - Novel view synthesis
 - 3D segmentation
 - Object-centric representation learning



Self-Supervised Learning

- Subcategory of unsupervised learning
- Use pretext task to train in a supervised fashion
- Hot-topic in deep learning community
 - Outperforms to supervised pretraining
 - No need for manual annotations

Google Scholar search results for "self-supervised" "contrastive learning".

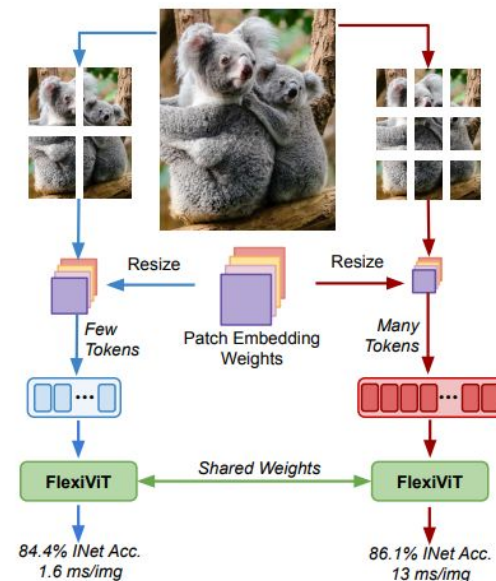
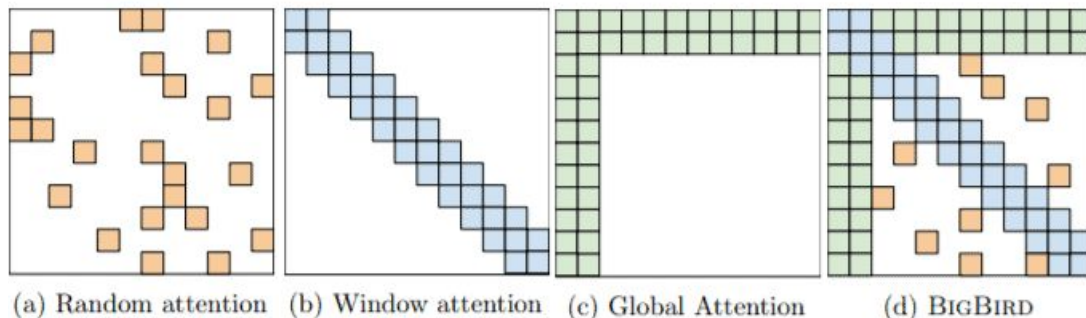
Articles: About 3,410 results (0,09 sec)

Any time
 Since 2023
 Since 2022
 Since 2019
 Custom range...

Dual-channel graph **contrastive learning** for **self-supervised** graph-level representation learning
 Z Luo, Y Dong, Q Zheng, H Liu, M Luo - Pattern Recognition, 2023 - Elsevier
 ... In this paper, we propose DualGCL, a novel graph **contrastive learning** framework for **self-supervised** graph-level representation learning. To aggregate node-level information to graph-...
 ☆ Save ⓘ Cite Cited by 2 All 3 versions

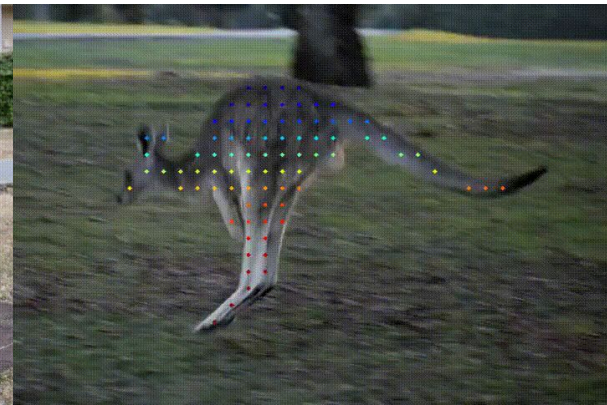
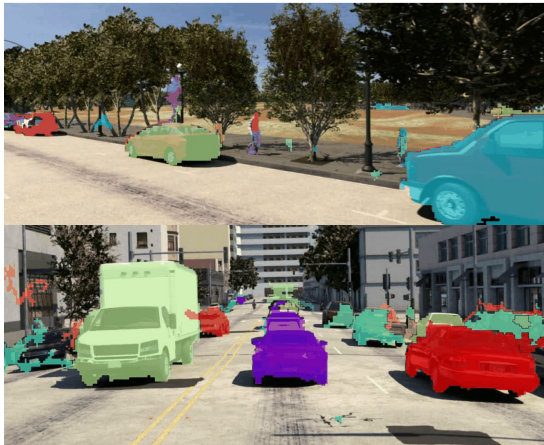
Transformers-based Models

- Transformer-based architectures for computer vision tasks
 - Efficient attention schemes
 - Hybrid attention-convolution models
 - Flexible transformer architectures
- State-of-the-art across many benchmarks



Video Decomposition and Tracking

- **Video decomposition:** decomposing a video sequence into object-centric components in an unsupervised manner
- **Video pixel tracking:** estimating motion trajectories for every pixel in every frame of a video




Select your topic

- Send me an email at villar@ais.uni-bonn.de
 - Your name
 - Matriculation number
 - Your two preferred topics/papers
- Upon my confirmation: Register in BASIS

Registration opened until 12.07.2023!

Deliverables (preliminary dates)

- **Presentation:** Thursday 21.09.2023
 - 30 min presentation
 - 15 min discussion
- **Report:** Thursday 28.09.2023 (will be one week after presentations)
 - LaTeX template
 - 8-12 pages
 - Brief but readable and informative
 - BibTex citations

 Arrange a meeting with me ≈ 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

Slot Assignment Selection

- Six slots for students
 - Assigned at random

Seminar Alternative ([Link](#))




Institute for Computer Science VI
Autonomous Intelligent Systems

Seminar Cognitive Robotics (MA-INF 4211)

Prof. Dr. Sven Behnke, Raphael Memmesheimer

Block course after lecture period.
First organizational meeting: 30.6.2023 12:00 ct, Friedrich-Hirzebruch-Allee 8, Room 0.016.
Presentations will be at the end of the lecture-free period (currently scheduled for 29.09.2023 (all day)).

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, students will give a presentation based on a recent publication from the area of cognitive robotics and write a seminar report.

Some of our robots:











University of Bonn, Institute for Computer Science, Departments: I, II, III, IV, V, VI | Impress | Data Privacy Statement

Questions?

