UNIVERSITÄT BONN AIS

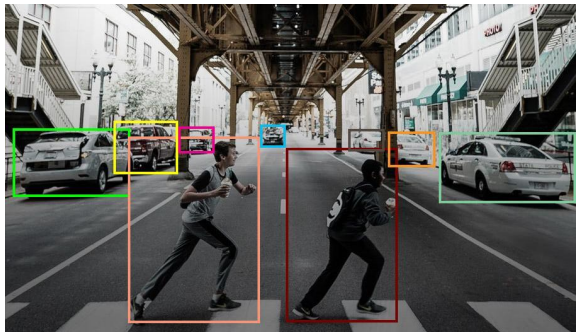
Seminar Vision Systems MA-INF 4208

08.07.2022

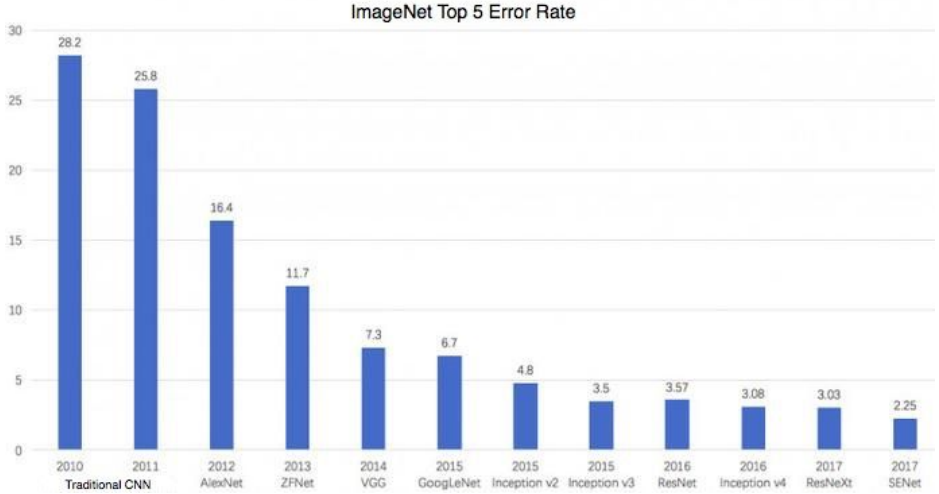
PROF. SVEN BEHNKE, ANGEL VILLAR-CORRALES

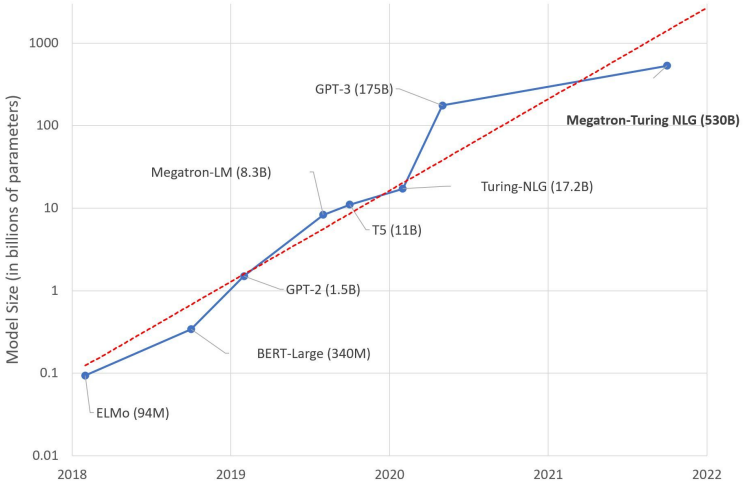
Contact: villar@ais.uni-bonn.de

The Age of Deep Learning



The Age of Deep Learning

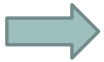




The Age of Deep Learning

In this seminar...

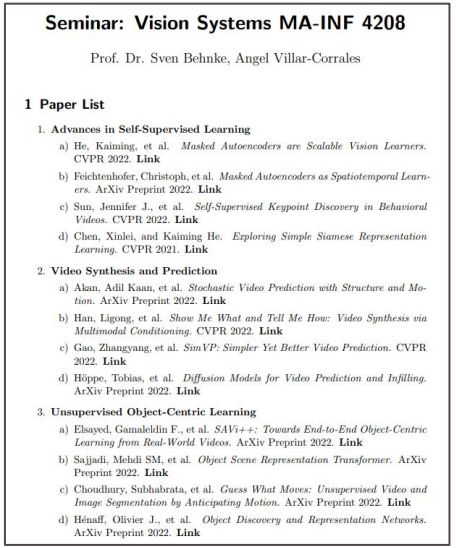
- Acquire/improve ability to:
  - deal with scientific publications (e.g., papers)
  - write a scientific report
  - present a scientific topic to an audience
  - engage technical topics



Important skills for Master Thesis!

In this seminar

- Discuss trending topics in deep learning and computer vision
- We will cover the following topics
  - Self-Supervised Learning
  - Video Synthesis and Prediction
  - Unsupervised Object-Centric Learning



Seminar: Vision Systems MA-INF 4208

Prof. Dr. Sven Behnke, Angel Villar-Corrales

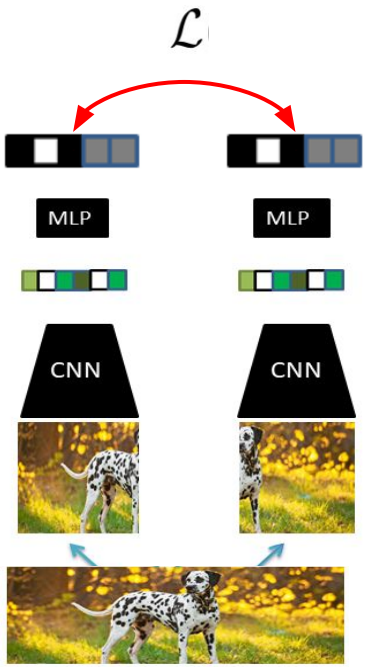
1 Paper List

1. **Advances in Self-Supervised Learning**
   a) He, Kaiming, et al. *Masked Autoencoders are Scalable Vision Learners.* CVPR 2022. **Link**
   b) Feichtenhofer, Christoph, et al. *Masked Autoencoders as Spatiotemporal Learners.* ArXiv Preprint 2022. **Link**
   c) Sun, Jennifer J., et al. *Self-Supervised Keypoint Discovery in Behavioral Videos.* CVPR 2022. **Link**
   d) Chen, Xinlei, and Kaiming He. *Exploring Simple Siamese Representation Learning.* CVPR 2021. **Link**
2. **Video Synthesis and Prediction**
   a) Akan, Adil Kaan, et al. *Stochastic Video Prediction with Structure and Motion.* ArXiv Preprint 2022. **Link**
   b) Han, Ligong, et al. *Show Me What and Tell Me How: Video Synthesis via Multimodal Conditioning.* CVPR 2022. **Link**
   c) Gao, Zhangyang, et al. *SimVP: Simpler Yet Better Video Prediction.* CVPR 2022. **Link**
   d) Höppe, Tobias, et al. *Diffusion Models for Video Prediction and Infilling.* ArXiv Preprint 2022. **Link**
3. **Unsupervised Object-Centric Learning**
   a) Elsayed, Gamaleldin F., et al. *SAVi++: Towards End-to-End Object-Centric Learning from Real-World Videos.* ArXiv Preprint 2022. **Link**
   b) Sajjadi, Mehdi SM, et al. *Object Scene Representation Transformer.* ArXiv Preprint 2022. **Link**
   c) Choudhury, Subhabrata, et al. *Guess What Moves: Unsupervised Video and Image Segmentation by Anticipating Motion.* ArXiv Preprint 2022. **Link**
   d) Hénaff, Olivier J., et al. *Object Discovery and Representation Networks.* ArXiv Preprint 2022. **Link**

Paper List: https://www.ais.uni-bonn.de/SS22/SeminarVision/PaperList.pdf

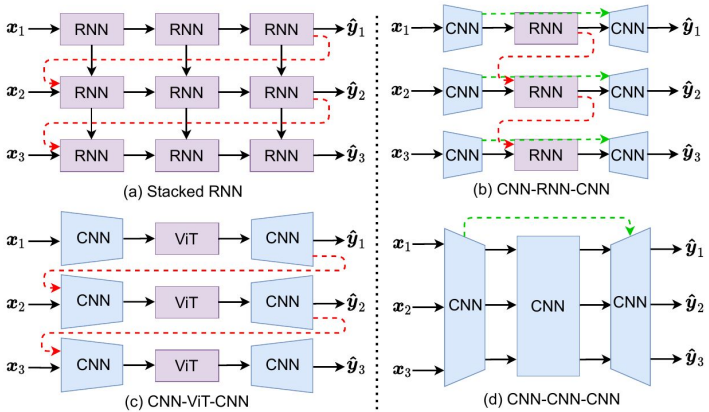
Self-Supervised Learning

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



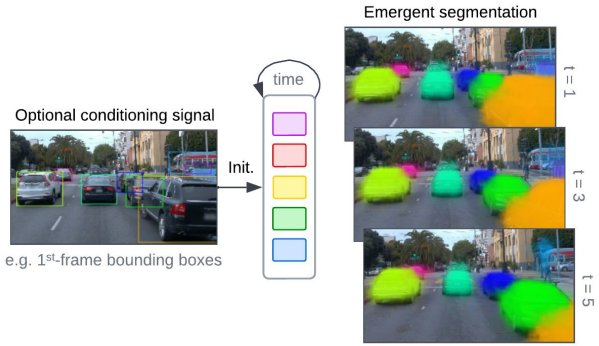
Video Synthesis and Prediction

- **Video prediction**: given C seed frames, predict next N plausible frames
  - Decision making for autonomous agents
  - Representation learning

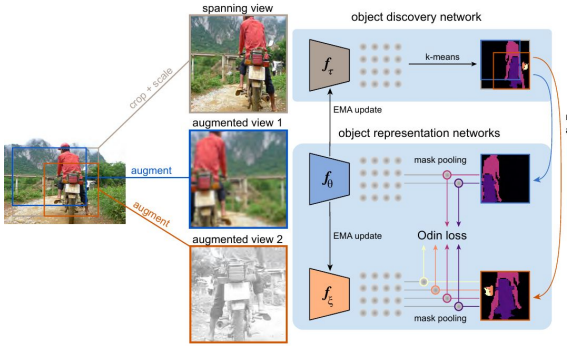


Unsupervised Object-Centric Learning

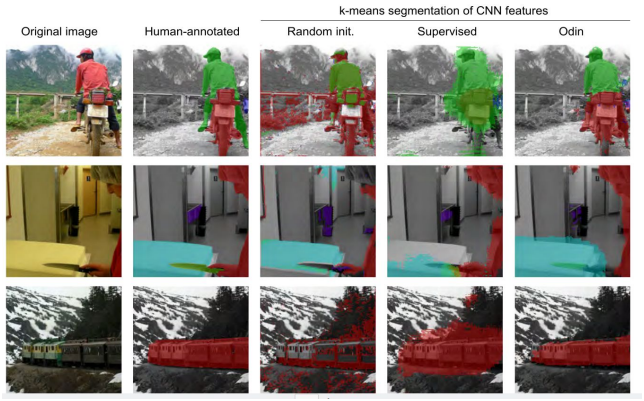
- Unsupervised learning of object representations with structured models:
  - Unsupervised segmentation
  - Object discovery



SAVI



ODIN



Select your topic

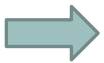
- Send me an email at villar@ais.uni-bonn.de
  - Your name
  - Matriculation number
  - Your selected topic

- Upon my confirmation: Register in BASIS

Deliverables

- Presentation: Thursday 28.09.2022
  - 30 min presentation
  - 15 min discussion

- Report: Thursday 05.10.2022
  - 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

- 60%: Presentation
  - Quality of the presentation slides
  - Presentation skills
  - Ability to answer questions

- 40%: 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

Questions?