UNIVERSITÄT BONN AIS

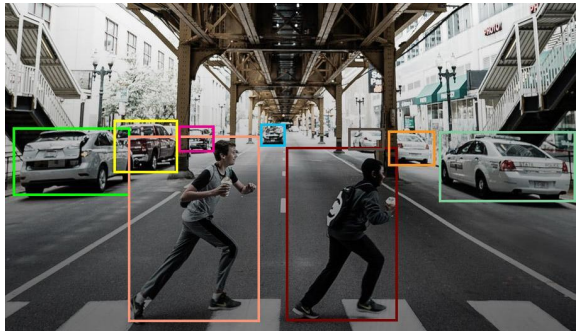
Seminar Vision Systems MA-INF 4208

02.02.2024

PROF. SVEN BEHNKE, ANGEL VILLAR-CORRALES

Contact: villar@ais.uni-bonn.de

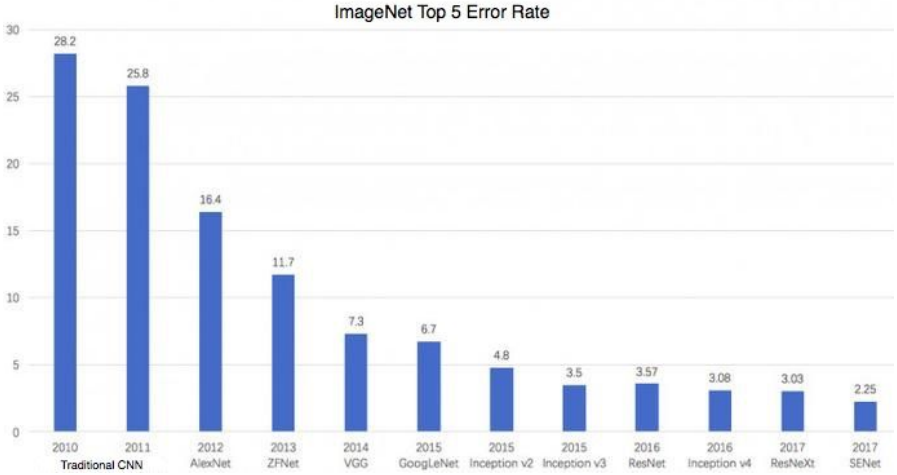
The Age of Deep Learning

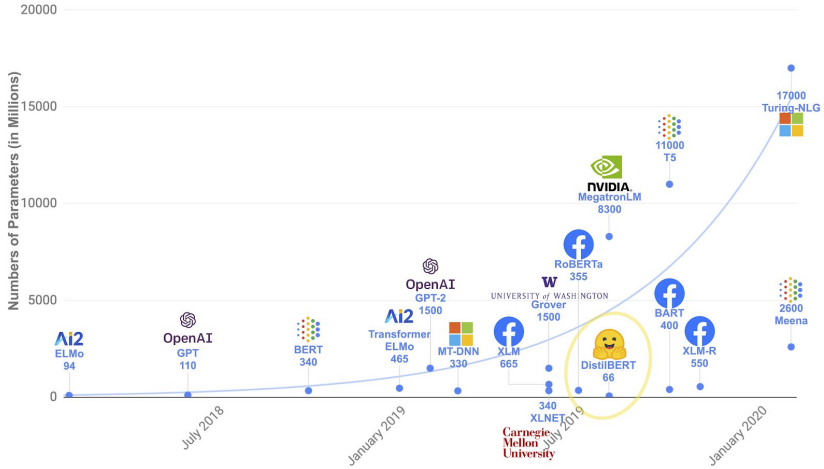


Expressive oil painting of basketball player dunking, causing an explosion of a nebula

The great wall of China in San Francisco

The Age of Deep Learning

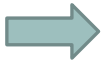




The Age of Deep Learning

In this Seminar...

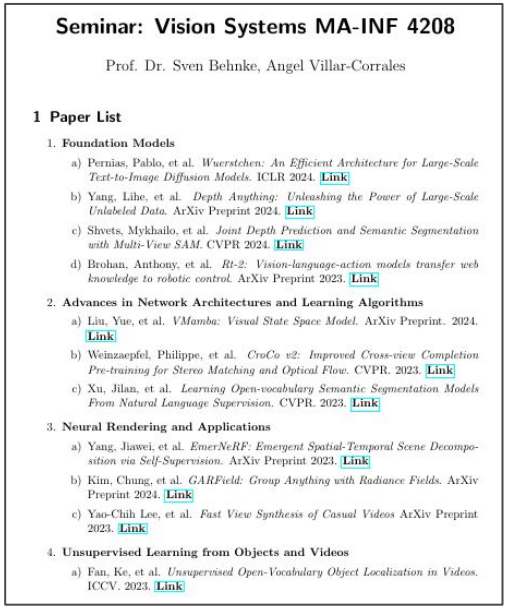
- Acquire/improve ability to:
  - deal with scientific publications (e.g. papers)
  - write a scientific report
  - present a scientific topic to an audience
  - engage technical topics



Important skills for Master Thesis!

In this seminar

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



Seminar: Vision Systems MA-INF 4208

Prof. Dr. Sven Behnke, Angel Villar-Corrales

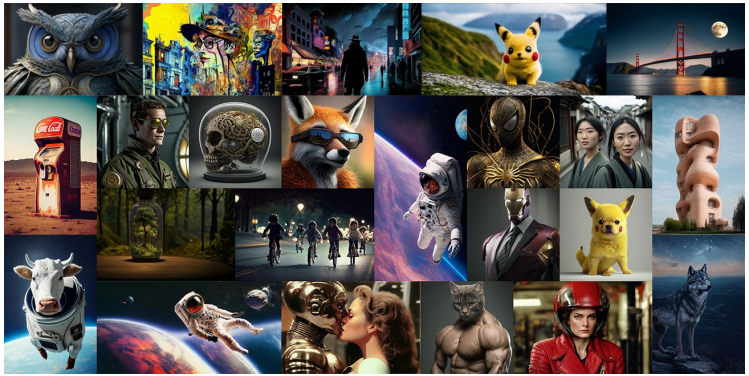
1 Paper List

1. **Foundation Models**
   a) Pernias, Pablo, et al. *Wuerstchen: An Efficient Architecture for Large-Scale Text-to-Image Diffusion Models.* ICLR 2024. Link
   b) Yang, Lihe, et al. *Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data.* ArXiv Preprint 2024. Link
   c) Shvets, Mykhailo, et al. *Joint Depth Prediction and Semantic Segmentation with Multi-View SAM.* CVPR 2024. Link
   d) Brohan, Anthony, et al. *Rt-2: Vision-language-action models transfer web knowledge to robotic control.* ArXiv Preprint 2023. Link
2. **Advances in Network Architectures and Learning Algorithms**
   a) Liu, Yue, et al. *VMamba: Visual State Space Model.* ArXiv Preprint. 2024. Link
   b) Weinzaepfel, Philippe, et al. *CroCo v2: Improved Cross-view Completion Pre-training for Stereo Matching and Optical Flow.* CVPR. 2023. Link
   c) Xu, Jilan, et al. *Learning Open-vocabulary Semantic Segmentation Models From Natural Language Supervision.* CVPR. 2023. Link
3. **Neural Rendering and Applications**
   a) Yang, Jiawei, et al. *EmerNeRF: Emergent Spatial-Temporal Scene Decomposition via Self-Supervision.* ArXiv Preprint 2023. Link
   b) Kim, Chung, et al. *GARField: Group Anything with Radiance Fields.* ArXiv Preprint 2024. Link
   c) Yao-Chih Lee, et al. *Fast View Synthesis of Casual Videos* ArXiv Preprint 2023. Link
4. **Unsupervised Learning from Objects and Videos**
   a) Fan, Ke, et al. *Unsupervised Open-Vocabulary Object Localization in Videos.* ICCV. 2023. Link

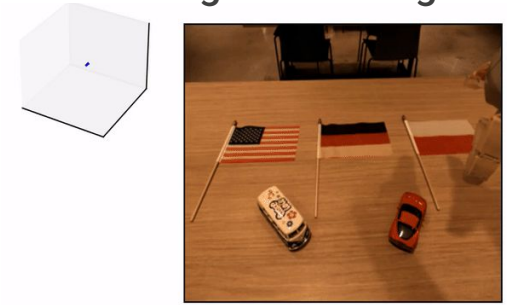
Paper List: https://www.ais.uni-bonn.de/WS2324/SeminarVision/PaperList.pdf

Foundation Models

- Models trained on internet-scale data with broad generalization capabilities
- Applications such as:
  - Text-Image Generation
  - Depth estimation
  - Robotics

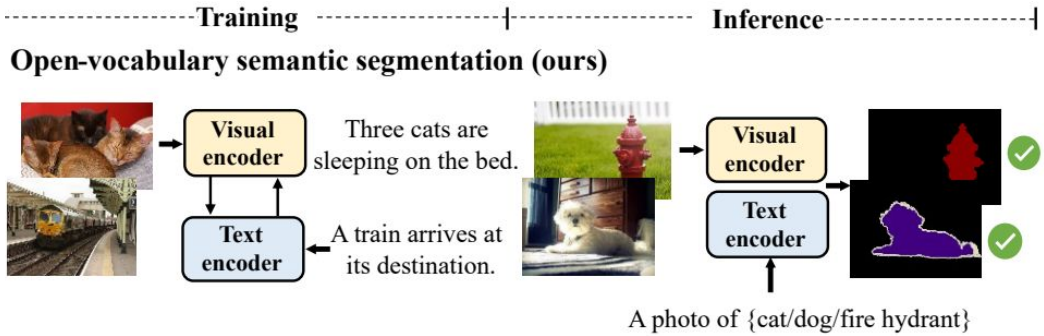


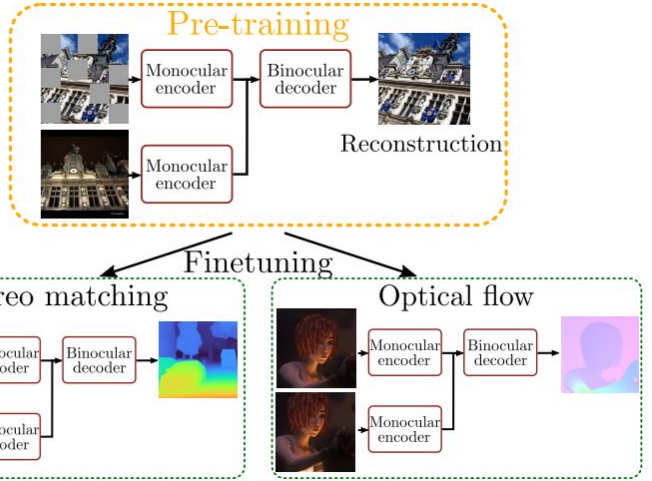
`move vwg van to germany`



Advances in Deep Learning Models

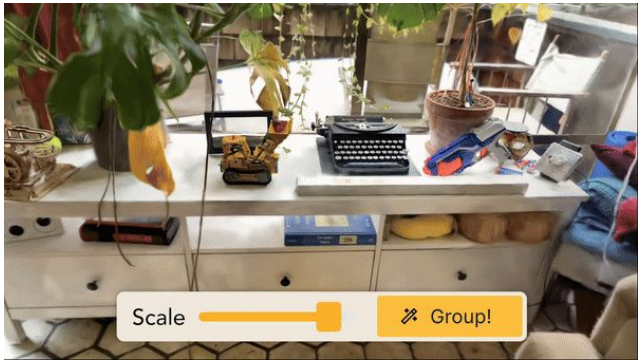
- Novel neural network architectures
  - State-space models
- Improved training techniques
  - Masked Self-Supervised Learning
  - Image-Language Contrastive Learning

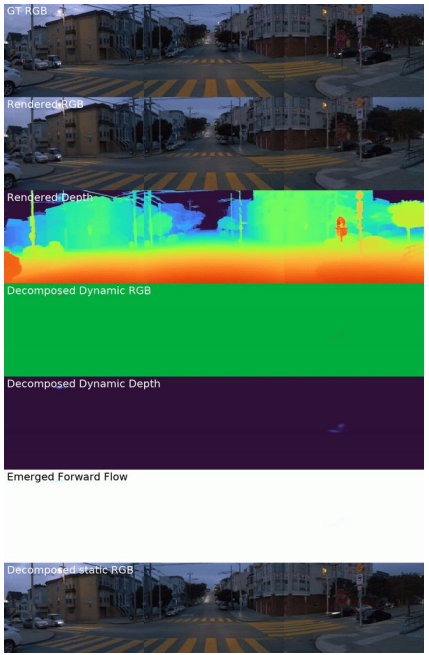




Neural Rendering and Applications

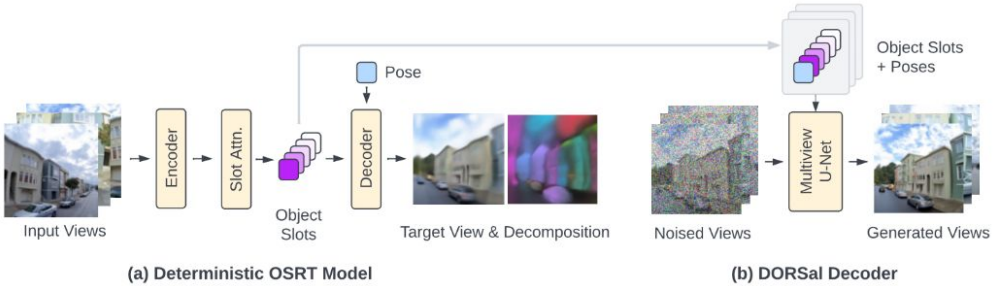
- Learning 3D-aware representations of a scene given a set of posed images
- Applications such as:
  - Spatio-temporal scene decomposition
  - 3D segmentation
  - Novel-view Synthesis

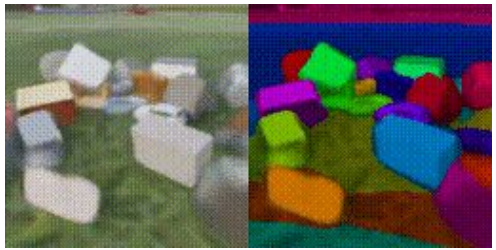


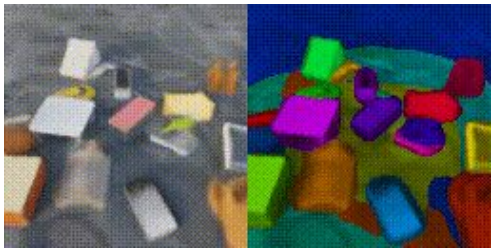


Learning from Objects and Videos

- Learning representations from video data without annotations
- Generalizable models that focus objects and moving segments







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 17.02.2024!

Deliverables (preliminary dates)

- **Presentation**: Thursday 21.03.2024
  - 30 min presentation
  - 15 min discussion

- **Report**: Thursday 28.03.2024 (will be one week after presentations)
  - LaTeX template
  - 8-12 pages
  - Brief but readable and informative
  - BibTex citations

⇨ Arrange a meeting with me ≈1-2 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 and strong points of the paper?
  - What do you think 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
  - Understanding of the concept
  - **Critical thinking and own discussion**

Slot Assignment Selection

- Six slots for students
  - Assigned at random

Seminar Alternative (Link)

Questions?