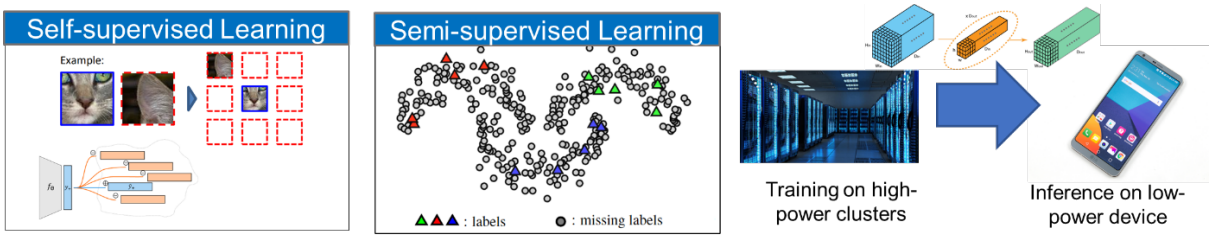
Universität Stuttgart

SEMINAR „LEARNING WITH LESS: RESOURCE- AND DATA-EFFICIENT VISUAL RECOGNITION“



OVERVIEW. Deep Convolutional Neural Networks (CNNs) and Visual Transformers (ViTs) achieve excellent results in many areas of computer vision but face challenges in real-world applications, such as dependence on costly annotated training data and the high amount of computational resources required for training and deploying such models. While the development of recognition algorithms was primarily driven by high recognition rates on large and cleanly annotated datasets for a long time, application-relevant goals such as learning with fewer computational resources, few training examples, imperfect annotations, adaptation to new data appearances and uncertainty estimation for identifying data-demand rapidly gain importance.

This seminar focuses on advanced network architectures, learning methods, and research areas in the field of deep learning for dealing with data scarcity and low-resource conditions in deep learning for computer vision. Students will engage in reading state-of-the-art papers, comparing them, and discussing the results. The seminar will cover, among other topics:

- **Low-resource Architectures**: Exploration of neural network architectures specifically designed to operate under limited computational resources.
- **Network Optimization, Pruning, and Quantization**: Techniques to reduce the size and computational demand of neural networks without significantly compromising performance.
- **Few-shot Visual Recognition**: Methods enabling models training with very few examples.
- **Zero-shot Learning and Knowledge Transfer from Language Models**: Approaches to leveraging pre-trained vision-language models to address new visual recognition tasks with zero training examples.
- **Semi-/Weakly-/Self-supervised Learning**: Strategies that effectively use unlabeled or partially labeled data to improve model performance.
- **Continual Visual Learning**: Techniques to enable models to learn continuously from new data without forgetting previously learned information.
- **Domain Adaptation**: Methods to adapt models trained in one domain to perform well in another, often significantly different, domain.
- **Uncertainty in Computer Vision**: Approaches to quantify and manage uncertainty in model predictions to enhance robustness and reliability.
- **Learning from Simulations and Generative Models for Data Generation**: Utilization of simulations and generative models to create synthetic data.

PREREQUISITES. Students should have a basic understanding of deep learning principles and techniques. Prior experience with neural networks and familiarity with key concepts in computer vision will be highly beneficial.

PROCEDURE. Students will read state-of-the-art papers, compare them, present and discuss the results. We value active participation in the discussion and demonstrating understanding and critical thinking. The procedure includes:

- Initial meeting to discuss the seminar structure and assign specific topics to each student.
- Paper Reading and Meetings with Topic Supervisor: Students will read 2-3 research papers on their assigned topic and meet with the supervisor to discuss the paper understanding and preparation of the presentation and the report.
- Present your findings in a seminar presentation, which may be conducted as a talk or a poster session (format to be determined).
- An 8-page report including a detailed analysis and comparison of the selected papers.

CONTACT:

Jun.-Prof. Dr.-Ing. Alina Roitberg

Intelligent Sensing and Perception, Institute for AI

Email: alina.roitberg@ki.uni-stuttgart.de