

Leon Sick

Research Focus

Five years of research (CVPR, ICCV, NeurIPS) on extracting supervision from unlabeled data with multimodal priors used as unsupervised quality proxies to filter pseudo-labels. Now applying the same principles to multimodal post-training through data-centric self-improvement for VLMs.

Education

Oct 2021 - Dec 2026 **Ph.D. Visual Computing Group, Ulm University, Germany**
Supervised by Timo Ropinski & Pedro Hermosilla (TU Vienna)
Thesis: Unsupervised Scene Perception with Deep Multimodal Priors

Oct 2019 - Aug 2021 **M.Sc. Business Information Technology, HTWG Konstanz, Germany**
Grade average: 1.2 (1.0 is best). With distinction.

Mar 2015 - Aug 2019 **B.A. International Business, Hochschule Aalen, Germany**
Grade average: 1.2 (1.0 is best). Top of my class.

Relevant Experience

Mar 2026 - Today **Student Researcher, Semantic Perception**
Google Research, Zürich, Switzerland
Research on Self-Improving Vision-Language Models:
- Synthetic Curriculum Generation with Unsupervised Quality-Proxies
- On-Policy Self-Distillation for Gemma
- Post-Training for Multimodal Self-Improvement

Oct 2021 - Dec 2024 **Doctoral Student Data & AI**
Liebherr Group, Ulm, Germany
Applied Research on Multimodal Learning & Perception Algorithms

Selected Publications

Google Scholar

Under Review **On-Policy Self-Training: Data-Free Self-Improvement in Large Language Models**
Leon Sick, L. Hoyer, Y. Strümpler, C. Reisswig, T. Ropinski, F. Tombari
Generate synthetic data as privileged information in combination with data filtering to self-improve language models in an iterative loop

NeurIPS 2026 **S2D: Sparse-To-Dense Keymask Distillation For Unsupervised Video Instance Segmentation**
Leon Sick, L. Hoyer, D. Engel, P. Hermosilla, T. Ropinski
Use motion priors as quality-proxy to evaluate noisy frame masks, then construct video pseudo-labels with Sparse-To-Dense Distillation.

ICCV 2025 — **CutS3D: Cutting Semantics in 3D for 2D Unsupervised Instance Segmentation**
Leon Sick, D. Engel, S. Hartwig, P. Hermosilla, T. Ropinski
Use 3D as modality to refine 2D-only pseudo-mask, estimate pseudo-label quality with spatial confidence from 3D and re-weight instance model loss.

CVPR 2025 — **Masked Scene Modeling: Narrowing the Gap Between Supervised and Self-Supervised Learning in 3D Scene Understanding**
P. Hermosilla, C. Stippel, Leon Sick

WACV 2025 Oral Presentation — **Attention-Guided Masked Autoencoders for Learning Image Representations**
Leon Sick, D. Engel, P. Hermosilla, T. Ropinski
Repurpose unsupervised object-discovery as a free guidance signal to improve semantic representation learning during pre-training.

CVPR 2024 — **Unsupervised Semantic Segmentation Through Depth-Guided Feature Correlation and Sampling**
Leon Sick, D. Engel, P. Hermosilla, T. Ropinski
Integrate monocular depth as a guidance signal for semantic segmentation to adapt contrastive visual feature learning to the 3D scene geometry.

References

Google — **Dr. Lukas Hoyer / Yannick Strümpler / Dr. Christian Reisswig**
TU Vienna — **Ass.-Prof. Dr. Pedro Hermosilla**
Ulm University — **Prof. Dr. Timo Ropinski**

Awards & Reviewing

Awards — **Top Reviewer NeurIPS 2024 & 2025**
Reviewer — **NeurIPS (24,25,26), ICLR (25,26), CVPR (25,26), ICML 25, ICCV 25**

Skills and competencies

Languages — Python, Java
ML Frameworks — PyTorch (4 first-author publications)
JAX (Google Internship)
PyTorch Lightning, Numpy, transformers, Detectron2

Spoken Languages — German (C2), English (C1), Italian (B2), French (B2)

Research Talks

Dec 2025 — „Unsupervised Visual Perception in 2D and 3D“ at KAUST, Saudi-Arabia
May 2025 — „Learning Segmentation Without Human Annotations“ at Liebherr Group
Mar 2025 — Oral Presentation of my WACV 2025 paper in Tucson, AZ
Sep 2023 — „Unsupervised Segmentation & Representation Learning“ at TU Vienna