

# Carsten T. Lüth

✉ lth.carsten@gmail.com | 🏠 sten2lu.github.io | 📧 sten2lu | 🌐 carsten-lueth-530798190

“The first principle is that you must not fool yourself — and you are the easiest person to fool.” — Richard Feynman

## Summary

PhD researcher specializing in evaluation methodology and empirical rigor for deep learning systems. Identified and addressed systematic evaluation flaws across Active Learning, Uncertainty Estimation, and Selective Classification. Built the largest empirical Active Learning benchmark for 3D biomedical image segmentation (150k GPU hours). Published 10 papers at top venues including 2 ICLR Orals, 1 NeurIPS Spotlight, 1 MICCAI STAR Award and 1 CVPR (Last Author)–700+ personal citations. Core maintainer of nnU-Net (8k+ GitHub stars). Leading organizer of heidelberg.ai (3,000+ members) with speakers featuring creators of AlphaFold, Stable Diffusion, and leading European LLMs. Expected defense April 2026.

## Education

### Heidelberg University

PHD IN COMPUTER SCIENCE

Heidelberg, Germany

2021/10 – Present

- **Thesis:** Translating Active Learning from Theory to Praxis for 3D Biomedical Segmentation
- **Supervisors:** Prof. Klaus Maier-Hein (DKFZ), Dr. Paul F. Jaeger (Google DeepMind)
- **Examiner:** Prof. Carsten Rother (Heidelberg University)
- **Expected Defense:** April 2026

### University of Amsterdam

ELLIS FOUNDATION MODEL WINTER SCHOOL

Amsterdam, Netherlands

2024 & 2026

- Intensive winter school on foundation models, covering latest research and applications in large language models, vision-language models, and more. Acceptance Rate 25%.

### Heidelberg University

MSc IN PHYSICS, SPECIALIZATION IN COMPUTATIONAL PHYSICS

Heidelberg, Germany

2019/10 – 2021/09

- **Thesis:** Feature-Based Anomaly Detection for MRI Brain Scans (Grade: 1.0)
- **Supervisors:** Prof. Klaus Maier-Hein (DKFZ)
- **Grade:** 1.3

### Heidelberg University

BSc IN PHYSICS

Heidelberg, Germany

2015/10 – 2019/08

- **Thesis:** Conditional Generation of Images with Invertible Neural Networks (Grade: 1.3)
- **Supervisors:** Prof. Ullrich Köthe (Heidelberg University)
- **Grade:** 1.7

## Research Experience

### Division of Medical Image Computing, German Cancer Research Center

PHD RESEARCHER

Heidelberg, Germany

2021/10 – Present

- Research on Deep Learning Evaluation, Active Learning, Uncertainty Estimation, and 3D Biomedical Segmentation
- Leading a team of 4 people for the nnActive & ALEGRA multi-institutional research projects
- Secured €750k in research funding (€200k cash + 250k A100 hours compute €500k value)
- Published 10 papers at top-tier venues (ICLR, NeurIPS, CVPR, TMLR, MICCAI)
- 2 ICLR Oral presentations, 1 NeurIPS Spotlight, 1 MICCAI STAR Award
- Supervised Master student Kim-Celine Kahl (ICLR 2024 Oral paper) and informal mentoring of junior researchers

### Open-Source Project by DKFZ

CORE-MAINTAINER OF NNU-NET

Heidelberg, Germany

2023/06 – Present

- Maintaining and developing nnU-Net, the de-facto standard for 3D biomedical image segmentation, a self-configuring framework with 8,000+ GitHub stars and 8,800+ citations.
- Implemented active learning features, fixed bugs, and provided support to the community of 1,000+ users

### heidelberg.ai

LEADING ORGANIZER

Heidelberg, Germany

2022/01 – Present

- Community building of an AI community with 3,000+ members, organizing events, workshops, and hackathons.
- Scaling platform for scientific discussions and networking in AI research from 1800 to 3000+ members and building a team of 4 volunteers
- Organized 20+ events with speakers from DeepMind, Stability AI, Aleph Alpha, ETH Zurich, and more
- Notable speakers: Simon Kohl (AlphaFold), Robin Rombach (Stable Diffusion), Jonas Andrusis (Aleph Alpha CEO)

- Researching generative models (Normalizing Flows, VAEs, GANs, Energy-Based Models) with specific focus on Invertible Neural Networks (INNs)
- Supervised by both Prof. Ullrich Köthe and Prof. Carsten Rother
- Developed training methods for Invertible Neural Networks in form of conditional INNs (cINNs) and Invertible Autoencoders
- Research informed 'Analyzing Inverse Problems with Invertible Neural Networks' (Ardizzone et al., 2018) (700+ citations)
- Contributed to cINN development (400+ citations)
- Teaching assistant: Fundamentals of Machine Learning (200 students)

**ABB Corporate Research Center**

- Researched and validated multivariate analysis tools for industrial chemometrics for the product BatchInsight
- Implemented production-ready code for real-world machine learning industrial problems in both C# and Python solving previous issues with unreliable statistical methods
- Developed technical report, documentation and tests for chemometric production system
- Presented results to ABB stakeholders and contributed to product roadmap discussions

## Research Funding & Grants

- |      |                                                                                                                                                             |                              |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|
| 2024 | <b>nnActive Compute Grant</b> , 250,000 GPU hours (A100) €500,000 value – Author & Person to Contact                                                        | <a href="#">NHR HoreKa</a>   |
| 2021 | <b>ALEGRA: Active Learning-enabled Generation of (Patho-)Physiological Lung Architectures for Pulmonary Medicine</b> , €200,000 – Author & Executional Lead | <a href="#">Helmholtz AI</a> |

## Awards & Honors

- 2024 **ICLR 2024 Oral Presentation**, ValUES Framework (Top 1% of submissions)
- 2024 **NeurIPS 2024 Spotlight**, Selective Classification Flaws (Top 5% of submissions)
- 2023 **ICLR 2023 Oral Presentation**, Failure Detection Evaluation (Top 1% of submissions)
- 2023 **MICCAI 2023 STAR Award**, cOOpD: COPD Classification as Anomaly Detection

## Skills & Languages

- Programming** Python with Machine Learning Stack (Numpy, SciPy, Pandas,...), Deep Learning Frameworks (PyTorch, TensorFlow), C++, Bash, Git, Linux, Job Scheduling (SLURM, LSF)
- Research Areas** Deep Learning Evaluation, Active Learning, Uncertainty Estimation, Generative Models, Medical Image Segmentation
- Languages** German (Native), English (Fluent, Academic Writing, Presentations)

## Selected Publications

Publications are categorized by research area. For a complete list, please visit:  
<https://scholar.google.com/citations?user=3L6NkkgAAAAJ>

### Evaluation Frameworks & Methodology

- |      |                                                                                                                                                                                                                                                                                                                                                                       |                                    |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------|
| 2026 | <b>Better than Average: Spatially-Aware Aggregation of Segmentation Uncertainty Improves Downstream Performance</b> , Vanessa E Guarino, Claudia Winklmayr, Jannik Franzen, Josef L Rumberger, Manuel Pfeuffer, Sonja Greven, Klaus H Maier-Hein, Dagmar Kainmueller, Christoph Karg, <b>Carsten T Lüth</b> . – Equal contribution: D Kainmueller*, C Karg*, CT Lüth* | <a href="#">CVPR</a>               |
| 2026 | <b>Finally Outshining the Random Baseline: A Simple and Effective Solution for Active Learning in 3D Biomedical Imaging</b> , <b>Carsten T Lüth</b> , Jeremias Traub, Kim-Celine Kahl, Till J Bungert, Lukas Klein, Lars Krämer, Paul F Jaeger, Klaus H Maier-Hein, Fabian Isensee. – Equal contribution: CT Lüth*, J Traub*                                          | <a href="#">TMLR</a>               |
| 2025 | <b>nnActive: A Framework for Evaluation of Active Learning in 3D Biomedical Segmentation</b> , <b>Carsten T Lüth</b> , Jeremias Traub, Kim-Celine Kahl, Till J Bungert, Lukas Klein, Lars Krämer, Paul F Jaeger, Klaus H Maier-Hein, Fabian Isensee. – Equal contribution: CT Lüth*, J Traub*. Compute: 150k A100 GPU hours                                           | <a href="#">TMLR</a>               |
| 2024 | <b>ValUES: A Framework for Systematic Validation of Uncertainty Estimation in Semantic Segmentation</b> , <b>Carsten T Lüth</b> , Kim-Celine Kahl, Maximilian Zenk, Klaus H Maier-Hein, Paul F Jaeger. – Oral Presentation (Top 1% of submissions). Equal contribution: CT Lüth*, KC Kahl*                                                                            | <a href="#">ICLR(Oral)</a>         |
| 2024 | <b>Overcoming Common Flaws in the Evaluation of Selective Classification Systems</b> , Jeremias Traub, Till J Bungert, <b>Carsten T Lüth</b> , Michael Baumgartner, Klaus H Maier-Hein, Lena Maier-Hein, Paul F Jäger. – Spotlight Presentation (Top 5% of submissions)                                                                                               | <a href="#">NeurIPS(Spotlight)</a> |
| 2024 | <b>Navigating the maze of explainable ai: A systematic approach to evaluating methods and metrics</b> , Lukas Klein, Carsten Lüth, Udo Schlegel, Till Bungert, Mennatallah El-Assady, Paul Jäger                                                                                                                                                                      | <a href="#">NeurIPS</a>            |
| 2023 | <b>Navigating the Pitfalls of Active Learning Evaluation: A Systematic Framework for Meaningful Performance Assessment</b> , <b>Carsten T Lüth</b> , Till J Bungert, Lukas Klein, Paul F Jaeger                                                                                                                                                                       | <a href="#">NeurIPS</a>            |
| 2023 | <b>A Call to Reflect on Evaluation Practices for Failure Detection in Image Classification</b> , Paul F Jaeger, <b>Carsten T Lüth</b> , Lukas Klein, Till J Bungert. – Oral Presentation (Top 1% of submissions)                                                                                                                                                      | <a href="#">ICLR(Oral)</a>         |

## Generative Modeling & Representation Learning

- 2019 **Guided Image Generation with Conditional Invertible Neural Networks**, Lynton Ardizzone, **Carsten T Lüth**, Jakob Kruse, Carsten Rother, Ullrich Köthe. – 400+ citations [arXiv](#)

## Vision Language Models

- 2024 **SURE-VQA: Systematic Understanding of Robustness Evaluation in Medical VQA Tasks**, Kim-Celine Kahl, Selen Erkan, Jeremias Traub, **Carsten T Lüth**, Klaus Maier-Hein, Lena Maier-Hein, Paul F Jaeger [TMLR](#)
- 2024 **Why context matters in vqa and reasoning: Semantic interventions for vlm input modalities**, Kenza Amara, Lukas Klein, **Carsten T Lüth**, Paul Jäger, Hendrik Strobelt, Mennatallah El-Assady [arXiv](#)
- 2024 **Interactive Semantic Interventions for VLMs: A Human-in-the-Loop Investigation of VLM Failure**, **Carsten T Lüth**, Lukas Klein, Kenza Amara, Paul F Jaeger. – Equal contribution: L Klein\*, K Amara\*, CT Lüth\* [NeurIPS WS](#)

## Medical Imaging Applications

- 2023 **CRADL: Contrastive Representations for Unsupervised Anomaly Detection and Localization**, **Carsten T Lüth**, David Zimmerer, Gregor Koehler, Paul F Jaeger, Fabian Isenensee, Klaus H Maier-Hein [BVM Workshop](#)
- 2023 **cOOpD: Reformulating COPD Classification on Chest CT Scans as Anomaly Detection Using Contrastive Representations**, **Carsten T Lüth**, Silvia D Almeida, Tobias Norajitra, Tassilo Wald, Marco Nolden, Paul F Jäger, Claus P Heussel, Jürgen Biederer, Oliver Weinheimer, Klaus H Maier-Hein. – STAR Award. Equal contribution: CT Lüth\*, SD Almeida\* [MICCAI\(STAR Award\)](#)
- 2022 **Unsupervised anomaly detection in the wild**, David Zimmerer, Daniel Paech, **Carsten T Lüth**, Jens Petersen, Gregor Köhler, Klaus Maier-Hein [BVM Workshop](#)