

Anna-Maria Halacheva

PhD Student, INSAIT | 3D Computer Vision, 3D Vision-Language Models, Scene Understanding

Sofia, Bulgaria | anna-maria.halacheva@insait.ai | [Google Scholar](#) | [GitHub](#) | [LinkedIn](#)

SUMMARY

PhD student at INSAIT building 3D vision-language models for indoor scene understanding, with first-author publications at NeurIPS (Spotlight), ICCV, and RA-L and a Google DeepMind PhD Fellowship. Experienced across the full research cycle, from 3D dataset creation and annotation tooling to model design, training, and evaluation, and in delivering machine learning and generative AI solutions to enterprise clients at IBM. Interested in research projects on 3D scene understanding, spatial reasoning, and multimodal foundation models.

EDUCATION

PhD in Computer Science (3D Computer Vision) | *INSAIT, Sofia University, Bulgaria*

Sep 2024 – Present · Full-time doctoral program. Supervised by Prof. Luc Van Gool and Prof. Danda Paudel.

M.Sc. Computer Science | *University of Stuttgart, Germany*

GPA 1.2 (1.0 = best)

Oct 2022 – Oct 2024 · First-class with Distinction.

B.Sc. Software Engineering | *University of Stuttgart, Germany*

GPA 1.4 (1.0 = best)

Oct 2019 – Oct 2022 · First-class.

TECHNICAL SKILLS

3D vision: point clouds, 3D Gaussian splatting, object and part segmentation, articulation prediction, Universal Scene Description (USD)

Multimodal ML: 3D vision-language models, language grounding in 3D, open-vocabulary scene understanding, transformer architectures

Data: dataset design, annotation tool development, annotation team management and quality control

Programming and tools: Python, PyTorch, CUDA, Hugging Face, Open3D, Git

Infrastructure: multi-GPU training on Slurm-managed clusters

Languages: Bulgarian (native), English (C1), German (C1)

MAIN PUBLICATIONS

P³-VLM: A Point-based Alternative for Grounded 3D Vision-Language Models [project/pdf](#)

NeurIPS 2026 | Anna-Maria Halacheva, Jan-Nico Zaech, Sombit Dey, Luc Van Gool, Danda Pani Paudel | **Spotlight**

Identifies a shortcut in how 3D vision-language models learn grounding: when asked about an object marked by a bounding box, the box's size and position already give away much of the answer, through both the text prompt and the model architecture. Replaces the box with a single 3D point, under which existing models lose much of their accuracy, and introduces P³-VLM, a 3D VLM designed for point prompts.

GaussianVLM: Scene-Centric 3D Vision-Language Models Using Language-Aligned Gaussian Splats for Embodied Reasoning and Beyond [project/pdf](#)

RA-L 2025 | Anna-Maria Halacheva, Jan-Nico Zaech, Xi Wang, Danda Pani Paudel, Luc Van Gool

Introduces the first 3D vision-language model operating natively on Gaussian splats, without object detectors. A dual sparsifier distills about 40k language-aligned Gaussians into 132 task-aware tokens, achieving state-of-the-art results across 3D VQA benchmarks and fivefold gains out of domain.

Articulate3D: Holistic Understanding of 3D Scenes as Universal Scene Description [project/pdf](#)

ICCV 2025 | Anna-Maria Halacheva*, Yang Miao*, Jan-Nico Zaech, Xi Wang, Danda Pani Paudel, Luc Van Gool | **equal contribution*

Introduces a dataset of 280 indoor scenes in Universal Scene Description format, annotated with object and part segmentation, motion parameters, movable and graspable regions, connectivity, and object mass, so scenes can be used directly in downstream tools. Proposes USDNet, which jointly segments movable and graspable parts and predicts their articulation from scene point clouds, and establishes a benchmark for articulated scene understanding.

OTHER PUBLICATIONS

Task-Adaptive Grounded 3D-Programmers Using 2D VLMs [project/pdf](#)
arXiv 2026 | Arman Raayatsanati, Sombit Dey, Anna-Maria Halacheva, Jan-Nico Zaech, Luc Van Gool, Danda Pani Paudel

From Scan to Action: Leveraging Realistic Scans for Embodied Scene Understanding [project/pdf](#)
CVPRW OpenSUN3D 2025 | Anna-Maria Halacheva, Jan-Nico Zaech, Sombit Dey, Danda Pani Paudel, Luc Van Gool

Expanding Expressivity in Transformer Models with MobiusAttention [project/pdf](#)
arXiv 2024 | Anna-Maria Halacheva, Mojtaba Nayyeri, Steffen Staab

RESEARCH AND INDUSTRY EXPERIENCE

PhD Researcher, 3D Computer Vision | *INSAIT, Sofia University, Bulgaria* Sep 2024 – Present
Focus on 3D computer vision and 3D vision-language models for indoor scene understanding: perceiving physical structure, grounding language in 3D environments, and reasoning about space. Research centered around a generalist 3D vision-language model that understands scenes at both the object level and the fine-grained level of parts, interactions, and affordances, and supports diverse tasks and output formats. Such a model would enable agents that act in physical spaces (robotics), assist people within them (XR assistants, accessibility), and help redesign them (interior design, architecture).

Graduate Research Fellow, 3D Vision | *INSAIT, Sofia University, Bulgaria* Apr 2024 – Aug 2024
Worked in a team on the CVPR 2024 OpenSUN3D workshop challenge on open-vocabulary 3D segmentation of long-tail objects.

AI Engineer, Master@IBM | *IBM Germany, Ehningen* Jan 2023 – Apr 2024
Delivered 7 machine learning and generative AI proof-of-concepts for IBM Client Engineering customers.

FELLOWSHIPS, AWARDS, AND RECOGNITION

Google DeepMind PhD Fellowship | *Google DeepMind* 2024 – Present

Top Reviewer | *NeurIPS 2026, CVPR 2026* 2026

Forbes Bulgaria 30 Under 30 | *Technology & Innovation* Dec 2024

Featured Researcher, EU State of the Union Address | *European Commission* [link](#) 2025

DAAD Undergraduate Fellowship | *German Academic Exchange Service* 2019 – 2022

Bulgarian National Diploma | *Highest Bulgarian distinction for high school graduates* 2019

SERVICE

OpenSUN3D Workshop Organization | *ICCV 2025 / CVPR 2026 / ECCV 2026* [link](#) Dec 2024 – Present

Reviewer for CVPR, ECCV, NeurIPS, IROS, ICRA, AAAI and RA-L 2025 – Present

Master Thesis Supervision | *INSAIT* 2025

EXPLORER@INSAIT Supervisor | *INSAIT; mentoring two undergraduate researchers* Oct 2024 – Present-

SELECTED PUBLIC ENGAGEMENT

The Superhuman Podcast with Georgi Nenov | *AI research in Bulgaria* [link](#) 2026

Bulgarian National Radio | *Bulgaria's place in AI research* [link](#) 2026

Konrad Adenauer Stiftung, Panel Discussion | *Women's empowerment* [link](#) 2025

bTV National Television | *Bulgarian advances in AI* [link](#) 2025