

Shweta Mahajan

🔗 [Personal Website](#) [LinkedIn Profile](#) ✉️ shweta.nith@outlook.com

EDUCATION

Technische Universität Darmstadt, Darmstadt, Germany <i>PhD (summa cum laude), Computer Vision and Machine Learning</i>	2018-2022 GPA: 1.0/1.0, with distinction
Universität des Saarlandes, Saarbrücken, Germany <i>M.Sc, Computer Science, Computer Vision and Machine Learning</i>	2015-2017 GPA: 1.4/1.0, Excellent
National Institute of Technology, Hamirpur, India <i>B.Tech, Computer Science</i>	2009-2013

RESEARCH INTERESTS

Computer Vision. Machine Learning. Probabilistic Models. Deep Generative Models. Multimodal Learning. Vision and Language

EXPERIENCE

Qualcomm AI Research, United States <i>Senior Machine Learning Researcher</i> Topic: Machine Learning and Computer Vision Job Duties: Conduct independent research on diffusion techniques and large multimodal models.	2023 – Present
University of British Columbia, Canada <i>Postdoctoral Research Fellow, Vision Group</i> Topic: Deep Generative Models for Videos and Language Supervisor: Prof. Dr. Leonid Sigal, Dr. Kwang Moo Yi Job Duties: Researcher at the UBC Computer Vision lab. Perform independent research on diffusion models. Publish papers. Advise graduate students.	2022 – 2023
Technical Universität Darmstadt, Germany <i>Doctoral Researcher, Visual Inference Group</i> Topic: Deep Generative Models for Images and Language Supervisor: Prof. Stefan Roth, Ph.D Job Duties: Conduct research in Computer Vision, Machine Learning, and Deep Learning. Publish research in top-tier venues. Teaching and conducting courses. Design homework assignments. Conduct seminars. Organizational activities.	2018 – 2022
Saarland University & MPI Informatics, Germany <i>Research Assistant, Machine Learning Group</i> Topic: Machine Learning, Optimization, Neural Networks Supervisor: Prof. Dr. Matthias Hein Job Duties: Research generalized polynomial neural networks with sparsity constraints. Mathematical problem formulation. Optimization framework. Experimental validation.	2016 – 2017
Samsung Research Institute, India <i>Software Engineer</i> Job Duties: B2B solutions. Develop the education framework for the magic board solutions. Algorithm design. Code and development.	2013 – 2015

HONOURS AND AWARDS

- Awarded Top Reviewer, NeurIPS, 2023
- Awarded Vector Research Grant, Canada, 2023
- Invited to Doctoral Consortium, CVPR, 2023
- Awarded one of the thirty best PhD Thesis in Germany, Switzerland and Austria, Gesellschaft für Informatik Dissertationspreis, 2023
- Best PhD Thesis (Nominated), The Bertha Benz Prize, Germany, 2023
- Awarded PhD with Distinction (summa cum laude) at Technische Universität, Darmstadt, Germany, 2023
- Best Paper Award, Impact on Science at the Computer Graphik Abend, 2021
- Innovative Project award for Bachelor Thesis at National Institute of Technology, 2013

TEACHING

- Deep Learning in Computer Vision, Technische Universität, Darmstadt** | *Teaching Assistant* 2021 – 2022
2020 – 2021
Practical labs on current problems in machine learning and computer vision.
Duties: Design projects on advanced topics in deep learning for computer vision. Supervise students on the practical implementation with modern deep learning frameworks. Results presented in a talk at the end of the lab. Concrete topics follow the current state of the art.
- Computer Vision I, Technische Universität, Darmstadt, Germany** | *Teaching Assistant* 2021 – 2022
Advanced undergraduate/graduate class that is part of a 2 course introductory sequence in computer vision.
Duties: Design assignments. Conduct tutorial sessions. Design examination assignments. Assess students' performance.
- Computer Vision II, Technische Universität, Darmstadt, Germany** | *Teaching Assistant* 2021 – 2022
Advanced undergraduate/graduate class that is part of a 2 course introductory sequence in computer vision.
Duties: Introductory lectures. Design assignments. Conduct tutorial sessions. Design examination assignments. Assess students' performance.

MENTORING & THESIS SUPERVISION

- University of British Columbia, Canada** | *Postdoctoral Supervisor* 2022 – 2023
Mentoring and collaboration with three PhD students on five different projects. All projects presented at top machine learning and computer vision venues including CVPR & NeurIPS.
1. Tanzila Rahman. Research topic: Visual content creation with text-to-image diffusion models. 2022-2023.
 2. Eric Hedlin. Research topic: Semantic correspondences in images with diffusion models. 2022-2023.
 3. Jeong-gi Kwak. Research topic: Diffusion models for novel view synthesis. 2022-2023.
- Technische Universität, Darmstadt, Germany** | *Supervisor* 2020 – 2022
Design master's thesis research topics in Computer Vision. Formalize the research direction. Supervise students on the practical implementation with modern deep learning frameworks. Results presented in a talk at the end of the lab and in a written thesis manuscript.
1. Konstantinos Papa. Multimodal Context-Object Split Latent Spaces for Diverse Multi-Task Learning. Master's Thesis, 2022.
 2. Oliver Hahn. Multimodal Data Augmentation for Image Captioning. Master's Thesis, 2021.
 3. Claudia Lölkes. Quantification of Semantic Spaces in Generative Models for Images. Master's Thesis, 2020.
 4. Franz Klein. Diverse Image Captioning with Style. Master's Thesis, 2020. **Presented at German Conference on Pattern Recognition (GCPR), 2021.**

MEMBERSHIPS

- Computer Vision Foundation: Member
- Vector Institute of Artificial Intelligence: Postdoctoral Affiliate

CONFERENCE REVIEWER

- IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). 2019 – 2024.
- IEEE International Conference on Computer Vision (ICCV). 2019 – 2023.
- Neural Information Processing Systems (NeurIPS). 2019 – 2024.
- International Conference on Machine Learning (ICML). 2020 – 2024.
- European Conference on Computer Vision (ECCV). 2020 – 2024.
- Winter Conference on Applications of Computer Vision (WACV). 2020 – 2024.
- International Conference on Learning Representations (ICLR). 2021 – 2024.
- AAAI Conference on Artificial Intelligence. 2020 – 2022.

JOURNAL REVIEWER

- International Journal of Computer Vision. 2019 – 2024.
- Transactions on Pattern Analysis and Machine Intelligence. 2019 – 2024.
- Computer Vision and Image Understanding. 2019 – 2024.

PRESENTATIONS AND TALKS

- Diffusion models: Theory and Applications** 2023
Tutorial, University of British Columbia, Vancouver, Canada.
- Multimodal Representation Learning with Deep Generative Models** 2023
Invited Talk, National Technical University, Singapore.
Poster Presentation, Doctoral Consortium at IEEE Conference on Computer Vision and Pattern Recognition, Vancouver, Canada.
- Make-A-Story: Visual Memory Conditioned Consistent Story Generation.** 2023
Feature Talk, Borealis AI, Vancouver, Canada.
- PixelPyramids: Exact Inference Models from Lossless Image Pyramids.** 2021
Virtual Presentation, IEEE International Conference on Computer Vision.
Poster Presentation, Simon Fraser University, Burnaby, Canada.
- Diverse Image Captioning with Context-Object Split Latent Spaces.** 2020
Virtual Presentation, Neural Information Processing Systems.
- Latent Normalizing Flows for Many-to-Many Cross-domain Mappings.** 2020
Virtual Presentation, International Conference on Learning Representations.
- Joint Wasserstein Encoders for Aligning Multimodal Embeddings** 2019
Oral Presentation, IEEE ICCV 2019 Workshop on Cross-Modal Learning in Real World, Seoul, Korea.

PUBLICATIONS

- [1] Eric Hedlin, Gopal Sharma, **Shweta Mahajan**, Xingzhe He, Hossam Isack, Abhishek Kar, Helge Rhodin, Andrea Tagliasacchi, and Kwang Moo Yi. “Unsupervised Keypoints from Pretrained Diffusion Models”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, Washington, June 2024, pp. 22820–22830.
- [2] Jeong-gi Kwak, Erqun Dong, Yuhe Jin, Hanseok Ko, **Shweta Mahajan**, and Kwang Moo Yi. “ViVid-1-to-3: Novel View Synthesis with Video Diffusion Models”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, Washington, June 2024, pp. 6775–6785.
- [3] **Shweta Mahajan**, Tanzila Rahman, Kwang Moo Yi, and Leonid Sigal. “Prompting Hard or Hardly Prompting: Prompt Inversion for Text-to-Image Diffusion Models”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, Washington, June 2024, pp. 6808–6817.
- [4] Eric Hedlin, Gopal Sharma, **Shweta Mahajan**, Hossam Isack, Abhishek Kar, Andrea Tagliasacchi, and Kwang Moo Yi. “Unsupervised Semantic Correspondence Using Stable Diffusion”. In: *Advances in Neural Information Processing Systems*. Ed. by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine. Vol. 36. 2023.
- [5] Tanzila Rahman, Hsin-Ying Lee, Jian Ren, Sergey Tulyakov, **Shweta Mahajan**, and Leonid Sigal. “Make-A-Story: Visual Memory Conditioned Consistent Story Generation”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver, British Columbia, June 2023, pp. 2493–2502.
- [6] Franz Klein, **Shweta Mahajan**, and Stefan Roth. “Diverse Image Captioning with Grounded Style”. In: *Proceedings of the 43rd German Conference on Pattern Recognition*. Ed. by Christian Bauckhage, Juergen Gall, and Alexander G. Schwing. Vol. 13024. Lecture Notes in Computer Science. Springer, 2021.
- [7] **Shweta Mahajan** and Stefan Roth. “PixelPyramids: Exact Inference Models from Lossless Image Pyramids”. In: *Proceedings of the Eighteenth IEEE International Conference on Computer Vision*. Virtual, Oct. 2021, pp. 6619–6628.

- [8] Apratim Bhattacharyya, **Shweta Mahajan**, Mario Fritz, Bernt Schiele, and Stefan Roth. “Normalizing Flows With Multi-Scale Autoregressive Priors”. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Virtual, June 2020, pp. 8412–8421.
- [9] **Shweta Mahajan**, Iryna Gurevych, and Stefan Roth. “Latent Normalizing Flows for Many-to-Many Cross-Domain Mappings”. In: *Proceedings of the Eighth International Conference on Learning Representations*. Addis Ababa, Ethiopia, Apr. 2020.
- [10] **Shweta Mahajan** and Stefan Roth. “Diverse Image Captioning with Context-Object Split Latent Spaces”. In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M.-F. Balcan, and H. Lin. Vol. 33. 2020, pp. 3613–3624.
- [11] **Shweta Mahajan**, Teresa Botschen, Iryna Gurevych, and Stefan Roth. “Joint Wasserstein Autoencoders for Aligning Multi-modal Embeddings”. In: *ICCV Workshop on Cross-Modal Learning in Real World*. Seoul, Korea, Oct. 2019, pp. 4561–4570.