

Haoming Cai

Ph.D. Candidate, Computer Science, University of Maryland

hmcai@umd.edu

hm-cai.com

Google Scholar

Research Interests

Foundation models for image and video · Physics-aware generation and 3D reconstruction with diffusion models · Computational imaging · Large-scale benchmark design

Education

2022–2027 **University of Maryland**, College Park, MD

Ph.D. in Computer Science. Advisor: [Christopher Metzler](#).

Physics-aware Gen-AI; computational imaging; image signal processing and end-to-end optimization.

2017–2022 **The Chinese University of Hong Kong, Shenzhen**, Shenzhen, China

B.Eng. in Computer Science and Engineering.

Gap year 2020–2021 at [XPixel](#) as a research assistant, then part time through 2022 while completing the degree, supervised by [Chao Dong](#).

Research Experience

Summer 2026 **Adobe**, San Jose, CA

Ph.D. Research Intern, NextCam team ([Marc Levoy](#)).

Summer 2025 **Adobe**, San Jose, CA

Ph.D. Research Intern, NextCam team ([Marc Levoy](#)), supervised by [Shumian Xin](#) and [Zhoutong Zhang](#).

- Bokeh editing project; outcome: ECCV 2026. Patent filing in process.

Summer 2024 **Dolby Vision Lab**, Sunnyvale, CA

Ph.D. Research Intern, supervised by [Guan-Ming Su](#).

- Portrait lighting control for text-to-image diffusion models without light stage data. Outcome: ICCV 2025, one patent.

2020–2022 **XPixel Group, SIAT**, Shenzhen, China

Undergraduate Research Assistant, supervised by [Chao Dong](#).

- Efficient and controllable image restoration; image quality and aesthetic assessment.
- Outcome: two ECCV papers, two CVPR workshops organized, one patent.

Publications

Reverse chronological. * denotes equal contribution. [Google Scholar](#) lists citation counts.

ECCV 2026 Vid2Bokeh: Synthesizing Large-Scale Light Field Datasets from Videos Enables Joint Focus and Depth-of-Field Image Editing.

[Homepage](#)

[Haoming Cai](#), Zhoutong Zhang, Christopher Metzler, Shumian Xin.

COLM 2026 Enhancing Vision Language Models' Robustness Under Adverse Imaging Conditions.

Tianfu Wang, Mingyang Xie, [Haoming Cai](#), Tianyi Xiong, Xiyao Wang, Dongdong Fu, Guan-Ming Su, Paola Cascante-Bonilla, Christopher Metzler.

- ICCV 2025** Parametric Shadow Control for Portrait Generation in Text-to-Image Diffusion Models.
[Homepage](#) **Haoming Cai**, Tsung-Wei Huang, Shiv Gehlot, Brandon Y. Feng, Sachin Shah, Guan-Ming Su, Christopher Metzler.
- CVPR 2025** Repurposing Pre-trained Video Diffusion Models for Event-based Video Interpolation.
[Homepage](#) Jingxi Chen, Brandon Y. Feng, **Haoming Cai**, Tianfu Wang, Levi Burner, Dehao Yuan, Cornelia Fermüller, Christopher Metzler, Yiannis Aloimonos.
- CVPR 2025** Flash-Split: 2D Reflection Removal with Flash Cues and Latent Diffusion Separation.
[Homepage](#) Tianfu Wang*, Mingyang Xie*, **Haoming Cai**, Sachin Shah, Christopher Metzler.
- NeurIPS 2024** Temporally Consistent Atmospheric Turbulence Mitigation with Neural Representations.
Haoming Cai*, Jingxi Chen*, Brandon Y. Feng, Weiyun Jiang, Mingyang Xie, Kevin Zhang, Cornelia Fermüller, Yiannis Aloimonos, Ashok Veeraraghavan, Christopher Metzler.
- ECCV 2024** Flash-Splat: 3D Reflection Removal with Flash Cues and Gaussian Splats.
[Homepage](#) Mingyang Xie*, **Haoming Cai***, Sachin Shah, Yiran Xu, Brandon Y. Feng, Jia-Bin Huang, Christopher Metzler.
- CVPR 2024** CodedEvents: Optimal Point-Spread-Function Engineering for 3D-Tracking with Event Cameras.
 Sachin Shah, Matthew Albert Chan, **Haoming Cai**, Jingxi Chen, Sakshum Kulshrestha, Chahat Deep Singh, Yiannis Aloimonos, Christopher Metzler.
- NeurIPS 2023** Assessor360: Multi-sequence Network for Blind Omnidirectional Image Quality Assessment.
[Homepage](#) Tianhe Wu, Shuwei Shi, **Haoming Cai**, Mingdeng Cao, Jing Xiao, Yinqiang Zheng, Yujiu Yang.
- ICCV 2023** Snow Removal in Video: A New Dataset and A Novel Method.
[Homepage](#) Haoyu Chen, Jingjing Ren, Jinjin Gu, Hongtao Wu, Xuequan Lu, **Haoming Cai**, Lei Zhu.
- ECCV 2022** Super-Resolution by Predicting Offsets: An Ultra-Efficient Super-Resolution Network for Rasterized Images.
[Homepage](#) Jinjin Gu, **Haoming Cai**, Chenyu Dong, Ruofan Zhang, Yulun Zhang, Wenming Yang, Chun Yuan.
- ECCVW 2022** Efficient Image Super-Resolution Using Vast-Receptive-Field Attention.
[Homepage](#) **Haoming Cai***, Lin Zhou*, Jinjin Gu, Zheyuan Li, Yingqi Liu, Xiangyu Chen, Yu Qiao, Chao Dong.
- CVPRW 2022** Blueprint Separable Residual Network for Efficient Image Super-Resolution.
[Homepage](#) Zheyuan Li, Yingqi Liu, Xiangyu Chen, **Haoming Cai**, Jinjin Gu, Yu Qiao, Chao Dong.
- CVPRW 2021** Toward Interactive Modulation for Photo-Realistic Image Restoration.
[Homepage](#) **Haoming Cai**, Jingwen He, Yu Qiao, Chao Dong.
- ECCV 2020** PIPAL: A Large-Scale Image Quality Assessment Dataset for Perceptual Image Restoration.
[Homepage](#) Jinjin Gu, **Haoming Cai**, Haoyu Chen, Xiaoxing Ye, Jimmy S. Ren, Chao Dong.

Preprints and Manuscripts

- arXiv 2026** Real2SAM2Real: Generative 3D Caches as Complementary Context for Video Diffusion.
[Homepage](#) Jiayi Wu*, **Haoming Cai***, Cornelia Fermüller, Christopher Metzler, Yiannis Aloimonos.
- In preparation** What If? Guiding Tropical Cyclones to Low-Likelihood High-Impact Scenarios.
Haoming Cai, Sunny Sharma, Dean Calhoun, Kyle J. C. Hall, Jhayron S. Perez-Carrasquilla, Jonathan Starfeldt, Christopher A. Metzler, Maria J. Molina.

Academic Service

- 2021 & 2022 **Co-organizer**, NTIRE Challenge on Perceptual Image Quality Assessment, CVPR Workshop. Jinjin Gu, [Haoming Cai](#), Chao Dong, Jimmy S. Ren, Yu Qiao, Shuhang Gu, Radu Timofte.
- Addressed generative distortion, a failure mode existing IQA metrics did not capture.
 - Methods introduced through the challenge include IQT (150+ citations) in 2021, and MANIQA (740+) and AHIQ (130+) in 2022.
 - Benchmark integrated into IQA toolboxes including *PyTorch-Image-Quality* and *IQA-PyTorch* (4,700+ combined GitHub stars); 800+ participants.
 - [NTIRE 2021](#) · [NTIRE 2022](#)
- Reviewer CVPR (2022–2026) · ECCV (2022, 2024, 2026) · NeurIPS (2024–2025) · ICCV (2025) · ICLR (2025) · ICML (2025) · AAAI (2023–2024) · WACV (2023, 2025) · TMM · TMLR

Honors and Awards

- 2026 Google Fellowship Finalist, university-wide. Graduate School, University of Maryland.
- 2025 CVPR Outstanding Reviewer. [Link](#)
- 2024 Finalist, Qualcomm Innovation Fellowship (18% acceptance). [Link](#)
- 2022–2024 Ph.D. Dean Fellowship Award. Graduate School, University of Maryland.
- 2022 Winner, Efficient Image Super-Resolution Challenge, model complexity track. NTIRE Workshop, CVPR 2022.

Technical Skills

- Languages Python · C++ · CUDA · PyTorch · NumPy · OpenCV · Git
- Methods Machine learning · Deep learning · Diffusion models · Vision-language models · 3D reconstruction and Gaussian splatting · Neural representations · Self-supervised learning · Model efficiency
- Domains Computational imaging and photography · Image signal processing (ISP) · HDR · Event cameras · Image quality assessment