

Gauri Jagatap

Senior Researcher, Dolby Laboratories

gauri.jagatap26@gmail.com

gaurijagatap.github.io

San Francisco, CA

RESEARCH INTERESTS

Multimodal representation learning, video understanding, vision-language models, computer vision, generative models, adversarial learning, computational imaging.

EXPERIENCE

Senior Researcher, Dolby Laboratories, San Francisco, CA Jan 2022 – Present
Multimodal Experiences Lab, Advanced Technology Group

- Transformer architectures that fuse video, audio and text for moment retrieval and video question answering.
- Open-world, event-level video scene graph generation for long-form content understanding.
- Submodular-optimization frame sampling for long-form video QA; point-supervised online temporal action localization.
- Agentic framework using event-level video scene graphs for personalized content curation.
- Lightweight audio-visual active speaker detection for real-time audio capture on mobile devices.

Research Intern, Google, New York, NY May – Aug 2021
Using external knowledge to improve language models.

Data Science Research Intern, Adobe, San Jose, CA May – Aug 2020
Generative models for lossless image compression.

Graduate Research Assistant, New York University Jan 2020 – Dec 2021
Computational imaging, deep learning, adversarial ML, HDR imaging, compression.

Research Intern, Mitsubishi Electric Research Laboratories, Cambridge, MA May – Aug 2018
Algorithms for multi-modal active imaging.

Graduate Research Assistant, Iowa State University Aug 2016 – Dec 2019
Computational imaging, deep learning, super-resolution, HDR imaging, compression.

EDUCATION

Ph.D. in Electrical Engineering, New York University (Advisor: Chinmay Hegde) 2020 – 2022

M.S. in Electrical Engineering, Iowa State University (Advisor: Chinmay Hegde) 2016 – 2019

B.E. in Electrical and Electronics Engineering and M.Sc. in Physics, BITS Pilani, India 2010 – 2015

SELECTED RESEARCH PROJECTS

- **Multimodal semantic video understanding**: moment retrieval, video QA, event-level scene graphs, long-form VQA sampling, online action localization.
- **Agentic workflows for personalized video curation** built on event-level video scene graphs.
- **Real-time audio-visual active speaker detection** for mobile devices.
- **Adversarial attacks and defenses**: sparse token attacks on vision transformers; entropy-regularized robust training via stochastic gradient Langevin dynamics.
- **Provable inverse imaging with trained and untrained generative priors**: compressed sensing, phase retrieval and HDR imaging with convergence guarantees.
- **Structured-sparsity phase retrieval** and **super-resolution via Fourier ptychography** with sparse and low-rank priors.

CONFERENCE PUBLICATIONS

2026 S. Reza, **G. Jagatap**, M. Moghaddam, O. Camps, A. Fanelli, "OnPoint: Offline-to-Online Multi-Level Distillation for Point-Supervised Online Temporal Action Localization", *European Conference on Computer Vision*, 2026.

- 2026 T. Nguyen, S. Zhang, S. Shankar, **G. Jagatap**, D. Chandran, A. Fanelli, M. Fiterau, "Video2Reaction: Training Foundation Video Models to Predict Audience Reaction", *European Conference on Computer Vision, 2026* (Also at the CVPR 2026 Workshop on Emerging Directions in Data for Multimodal Foundation Models).
- 2026 A. Agrawal, G. KV, R. Aralikatti, **G. Jagatap**, J. Yuan, S. Baskar, V. Kamarshi, A. Fanelli, F. Huang, "Towards Mitigating Hallucinations in Large Vision-Language Models by Refining Textual Embeddings", *Findings of the Association for Computational Linguistics, 2026* (Also at the ICLR 2026 Workshop on Multimodal Intelligence).
- 2022 T. Nguyen, **G. Jagatap**, C. Hegde, "Inverse Imaging with Generative Priors via Langevin Dynamics", *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2022*.
- 2020 **G. Jagatap**, C. Hegde, "High Dynamic Range Imaging using Deep Image Priors", *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2020*.
- 2019 **G. Jagatap**, C. Hegde, "Algorithmic Guarantees for Inverse Imaging with Untrained Network Priors", *Advances in Neural Information Processing Systems, 2019* (Acceptance rate: 21.18%).
- 2019 **G. Jagatap**, C. Hegde, "Linearly Convergent Algorithms for Learning Shallow Residual Networks", *IEEE International Symposium on Information Theory, 2019*.
- 2018 **G. Jagatap**, Z. Chen, C. Hegde, N. Vaswani, "Model Corrected Low Rank Ptychography", *IEEE International Conference on Image Processing, 2018*.
- 2018 **G. Jagatap**, C. Hegde, "Towards Sample-Optimal Methods for Solving Random Quadratic Equations with Structure", *IEEE International Symposium on Information Theory, 2018*.
- 2018 **G. Jagatap**, Z. Chen, C. Hegde, N. Vaswani, "Sub-diffraction Imaging using Fourier Ptychography and Structured Sparsity", *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2018* (Oral presentation).
- 2018 Z. Chen, **G. Jagatap**, S. Nayer, C. Hegde, N. Vaswani, "Low Rank Fourier Ptychography", *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2018*.
- 2017 **G. Jagatap**, C. Hegde, "Fast, Sample Efficient Algorithms for Structured Phase Retrieval", *Advances in Neural Information Processing Systems, 2017* (Acceptance rate: 20.93%).

JOURNAL ARTICLES

- 2022 T. Nguyen, **G. Jagatap**, C. Hegde, "Provable Compressed Sensing with Generative Priors via Langevin Dynamics", *IEEE Transactions on Information Theory, 2022*.
- 2021 **G. Jagatap**, A. Joshi, A. Chowdhury, S. Garg, C. Hegde, "Adversarially Robust Learning via Entropic Regularization", *Frontiers in Artificial Intelligence, 2021* (Also at the ICML 2021 Workshop on Adversarial Machine Learning).
- 2019 **G. Jagatap**, Z. Chen, S. Nayer, C. Hegde, N. Vaswani, "Sample Efficient Fourier Ptychography for Structured Data", *IEEE Transactions on Computational Imaging, 2019*.
- 2019 **G. Jagatap**, C. Hegde, "Sample Efficient Algorithms for Recovering Structured Signals from Magnitude-Only Measurements", *IEEE Transactions on Information Theory, 2019*.

WORKSHOP PAPERS

- 2026 T. Lin, D. Chandran, **G. Jagatap**, C. Chen, A. Fanelli, D. Gunawan, J. Kimball, "Semantic Action Graph: A Shared Representation for Agent Grounding and Human Interpretation of Sports Highlights", *IEEE VIS Workshop on GenAI, Agents, and the Future of VIS, 2026*.
- 2025 M. Chasmai, **G. Jagatap**, G. KV, G. Van Horn, S. Maji, A. Fanelli, "Moment Sampling in Video LLMs for Long-Form VQA", *CVPR Workshop on Video Large Language Models, 2025*.
- 2023 A. Joshi, S. Akula, **G. Jagatap**, C. Hegde, "A Few Adversarial Tokens Can Break Vision Transformers", *CVPR Workshop on Adversarial Machine Learning on Computer Vision, 2023*.
- 2023 P. Mehta, **G. Jagatap**, K. Gallagher, B. Timmerman, P. Deb, S. Garg, R. Greenstadt, B. Dolan-Gavitt, "Can Deepfakes be created on a whim?", *International Workshop on Cyber Social Threats (CySoc), ACM Web Conference, 2023*.
- 2022 A. Joshi, **G. Jagatap**, C. Hegde, "Adversarial Token Attacks on Vision Transformers", *CVPR Workshop on Transformers for Vision, 2022*.
- 2019 **G. Jagatap**, C. Hegde, "Phase Retrieval using Untrained Neural Network Priors", *NeurIPS Workshop on Solving Inverse Problems with Deep Neural Networks, 2019*.

UNDER REVIEW

2026 A. Majee, **G. Jagatap**, D. Chandran, R. Iyer, A. Fanelli, "SCOUT: Segment Curation for Optimal Understanding of Long-Form Videos", *under review*.

PATENT APPLICATIONS (FILED BY DOLBY LABORATORIES)

2025 A. Majee, **G. Jagatap**, D. Chandran, A. Fanelli, "SCOUT: Segment Curation for Optimal Understanding of Long-Form Videos" (D25113USP1).

2024 G. Cengarle, D. Scaini, Y. Ma, Z. Li, **G. Jagatap**, B. Crockett, "Object extraction and lock in position for complex spatial scenes" (D24113ES).

2024 **G. Jagatap**, K. V. Gouthaman, J. Serra, A. Fanelli, D. Chandran, "System for audio-visual content trimming: highlight generation, moment retrieval and summarization for user generated content creators" (D23160USP2).