

Gauri Jagatap

gauri.jagatap26@gmail.com | gaurijagatap.github.io

EDUCATION

JAN 2020	Doctor of Philosophy (PhD) in ELECTRICAL ENGINEERING
-JAN 2022	New York University (Advisor: Chinmay Hegde)
AUG 2016	Master of Science (MS) in ELECTRICAL ENGINEERING
-DEC 2019	Iowa State University (Advisor: Chinmay Hegde)
AUG 2010	Bachelor of Engineering (BE) in ELECTRICAL AND ELECTRONICS ENGINEERING
-MAY 2015	Master of Science (MSc) in PHYSICS
	Birla Institute of Technology and Science, India

RESEARCH INTERESTS

Deep Learning, Multimodal Representation Learning, Video Understanding, Computer Vision, Generative Models, Adversarial Learning, Computational Imaging,

WORK EXPERIENCE

JAN 2022	Senior Researcher at Dolby Laboratories , San Francisco, California.
-Present	Developing deep learning algorithms for multi-modal video applications.
MAY 2021	Research Intern at Google , New York, New York.
-AUG 2021	Using external knowledge for improving language models.
MAY 2020	Data Science Research Intern at Adobe , San Jose, California.
JAN 2020	Graduate Research Assistant at New York University
-DEC 2021	Computational imaging, deep learning, adversarial ML, high dynamic range imaging, compression.
-AUG 2020	Developing generative models for lossless image compression.
MAY 2018	Research Intern at Mitsubishi Electric Research Laboratories (MERL) , Cambridge, Massachusetts.
-AUG 2018	Developing algorithms for multi-modal active imaging.
AUG 2016	Graduate Research Assistant at Iowa State University
-DEC 2019	Computational imaging, deep learning, super-resolution, high dynamic range imaging, compression.

RESEARCH PROJECTS

-
- **Multimodal semantic video understanding.**
 - Developed a new transformer based architecture that leverages video, audio and text modalities for (1) retrieving moments of interest in video content, and (2) video question answering.
 - Developed an open-world event-level video scene graph generation framework for long-form video content understanding.
 - Designed a new sampling strategy based on sub-modular optimization for long-form video question-answering.
 - Designed an online temporal activation localization method using point-supervised training labels.
 - **Agentic workflows for personalized video curation.**
 - Designed an agentic framework that leverages an event-level video-scene graph representation for personalized content curation based on user preferences.
 - **Real-time audio-visual active speaker detection.**
 - Developed a light-weight audio-visual active speaker detection model for improved real-time audio capture on mobile devices.
 - **Adversarial attacks and defenses on deep networks.**
 - Developed a block sparsity based token attack to highlight sensitivity of vision transformers to specially structured attacks.
 - Devised a new algorithm for training neural networks robust to adversarial perturbation, with better generalization properties, using entropy and stochastic gradient langevin dynamics.
 - **Provable inverse imaging using deep trained and untrained generative priors.**
 - Introduced a new provably convergent algorithm for solving inverse problems such as compressed sensing with pre-trained generative priors using stochastic gradient langevin dynamics.
 - Used deep untrained CNNs as priors for inverse imaging problems such as compressed sensing, phase retrieval and HDR imaging, showed superior empirical performance. Provided theoretical guarantees for convergence of gradient descent based solution.
 - **Inverse imaging from magnitude-only measurements using structured sparsity priors.**
 - Phase retrieval using structured sparsity: used underlying structure (such as block and tree sparsities) in images to develop fast and memory efficient algorithms to reconstruct images from absolute-valued measurements with theoretical guarantees.
 - **Image and video super-resolution via ptychography.**
 - Developed algorithms for super-resolution of multiplexed microscopic images by using sparsity and low rank priors.

JOURNAL ARTICLES

MAY 2022	T. Nguyen, G. Jagatap and C. Hegde, "Provable compressed sensing with generative priors via langevin dynamics", IEEE Transactions on Information Theory , 2022.
OCT 2021	G. Jagatap , A. Joshi, A. Chowdhury, S. Garg and C. Hegde, "Adversarially robust learning via entropic regularization", Frontiers in Artificial Intelligence , 2021.
AUG 2019	G. Jagatap , Z. Chen, S. Nayer, C. Hegde and N. Vaswani, "Sample efficient Fourier ptychography for structured data", IEEE Transactions on Computational Imaging , 2019.
JAN 2019	G. Jagatap and C. Hegde, "Sample-efficient algorithms for recovering structured signals from magnitude-only measurements", IEEE Transactions on Information Theory , 2019.

CONFERENCE PUBLICATIONS

APR 2026	A. Agrawal, G. KV, R. Aralikatti, G. Jagatap , J. Yuan, S. Baskar, V. Kamarshi, A. Fanelli and F. Huang, "Towards Mitigating Hallucinations in Large Vision-Language Models by Refining Textual Embeddings", <i>to appear</i> , ACL Findings, 2026.
APR 2026	A. Agrawal, G. KV, R. Aralikatti, G. Jagatap , J. Yuan, S. Baskar, V. Kamarshi, A. Fanelli and F. Huang, "Towards Mitigating Hallucinations in Large Vision-Language Models by Refining Textual Embeddings", ICLR Workshop on Multimodal Intelligence, 2026.
JUN 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 LLMs, 2025.
MAY 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?", Companion Proceedings of the ACM Web Conference 2023.
MAR 2023	A. Joshi, S. Akula, G. Jagatap and C. Hegde, "A Few Adversarial Tokens Can Break Vision Transformers", CVPR Workshop on Adversarial Machine Learning on Computer Vision, 2023.
JUN 2022	A. Joshi, G. Jagatap and C. Hegde, "Adversarial token attacks on vision transformers", CVPR Workshop on Transformers for Vision, 2022.
MAY 2022	T. Nguyen, G. Jagatap and C. Hegde, "Inverse Imaging with Generative Priors Via Langevin Dynamics", ICASSP, 2022.
JUL 2021	G. Jagatap , A. Joshi, A. Chowdhury, S. Garg and C. Hegde, "Adversarially robust learning via entropic regularization", ICML Workshop on Adversarial Machine Learning, 2021.
MAY 2020	G. Jagatap and C. Hegde, "High dynamic range imaging using deep image priors", ICASSP, 2020.
DEC 2019	G. Jagatap and C. Hegde, "Algorithmic guarantees for inverse imaging with untrained network priors", NeurIPS, 2019. (Acceptance rate: 21.18%).
DEC 2019	G. Jagatap and C. Hegde, "Phase retrieval using untrained neural network priors", NeurIPS Workshop on Solving Inverse Problems with Deep Networks, 2019.
JUL 2019	G. Jagatap and C. Hegde, "Linearly convergent algorithms for learning shallow residual networks", ISIT, 2019.
OCT 2018	G. Jagatap , Z. Chen, C. Hegde and N. Vaswani, "Model corrected low rank ptychography", ICIP, 2018.
JUN 2018	G. Jagatap and C. Hegde, "Towards sample-optimal methods for solving random quadratic equations with structure", ISIT, 2018.
APR 2018	G. Jagatap , Z. Chen, C. Hegde and N. Vaswani, "Sub-diffraction imaging using Fourier ptychography and structured sparsity", ICASSP, 2018 (Oral presentation).
APR 2018	Z. Chen, G. Jagatap , S. Nayer, C. Hegde and N. Vaswani, "Low rank Fourier ptychography", ICASSP, 2018.
DEC 2017	G. Jagatap and C. Hegde, "Fast, sample-efficient algorithms for structured phase retrieval", NeurIPS, 2017. (Acceptance rate: 20.93%).

PREPRINTS UNDER REVIEW

MAR 2026	S. Reza, G. Jagatap , M. Moghaddam, O. Camps and A. Fanelli, "OnPoint: Offline-to-Online Multi-Level Distillation for Point-Supervised Online Temporal Action Localization", <i>under review</i> 2026.
MAR 2026	A. Majee, G. Jagatap , D. Chandran, R. K. Iyer and A. Fanelli, "SCOUT: Segment Curation for Optimal Understanding of Long-Form Videos", <i>under review</i> , 2026.
MAR 2026	T. Lin, D. Chandran, G. Jagatap , C. Chen, F. Robinson, A. Fanelli, D. Gunawan and J. Kimball, "SportSAGE: Reliving the Game through Semantic Action Graphs and Narrative-Driven AI Highlights", <i>under review</i> , 2026.
MAR 2026	T. Nguyen, S. Zhang, S. Shankar, G. Jagatap , D. Chandran, A. Fanelli and M. Fiterau, "Video2Reaction: Mapping Video to Audience Reaction Distribution in the Wild", <i>under review</i> , 2026.

SUBMITTED PATENT APPLICATIONS

APRIL 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), <i>application filed by Dolby Laboratories</i> .
AUG 2025	A. Majee, G. Jagatap , D. Chandran and A. Fanelli, "SCOUT: Segment Curation for Optimal Understanding of Long-Form Videos", (D25113USP1) <i>application filed by Dolby Laboratories</i> .
AUG 2024	G. Cengarle, D. Scaini, Y. Ma, Z. Li, G. Jagatap and B. Crockett, "Object extraction and lock in position for complex spatial scenes", (D24113ES) <i>application filed by Dolby Laboratories</i> .