

Hritam Basak

Stony Brook, New York, 11790

✉ hbasak@cs.stonybrook.edu | [G Google Scholar](#) | [in LinkedIn](#) | [🔗 hritam-98.github.io](https://github.com/hritam-98)

INDUSTRIAL EXPERIENCE

Applied Scientist II (×2)

Amazon Robotics

California, USA

June, 2026 – Present

- Working on Humanoid manipulation and whole body control.
- Closely collaborating with Fauna Robotics group for VLA pose-training.
- Supervising interns for dexterous manipulation and gripper grasping in cluttered environment.

Applied Scientist Intern (×2)

Amazon

California, USA

June, 2025 – Present || May, 2024 – Oct. 2024

- Researching on Vision-Language Action (VLA) models for 3D affordance estimation and manipulation. Work under review at NeurIPS'26.
- Developing a comprehensive end-to-end video generation pipeline for exocentric-to-egocentric viewpoint translation, enabling transformative progress in imitation learning and advancing the frontier of embodied AI. Work accepted at ECCV'26.
- Worked on image-guided 3D generation using Gaussian Splatting using 3D and 2D Diffusion priors. [IROS'25]
- Generated 3D assets used for VLM-based articulated object understanding and manipulation.
- Collaborated on uncertainty-guided 6-DoF pose estimation for robotics. [CVPR'25]

Data Scientist

Tata Neu

Mumbai, India

Jun. 2021 – Jul. 2022

- Engineered a visual search engine for fashion recommendations using RCNN for foreground extraction and pre-trained ResNet for feature extraction, achieving over 96% accuracy.
- Developed an automated human-in-the-loop system to annotate Tata Group's native 20M+ fashion image dataset.
- Designed a promotion recommendation algorithm for 80M customer groups employing Churn and CLTV, and collaborative filtering.

RESEARCH EXPERIENCE

Graduate Research Assistant

Stony Brook University – Advisor: Dr. Zhaozheng Yin

New York, USA

Aug. 2023 – Present

- Devised a novel Visual Iterative Self-Prompting algorithm with Neural Affordance Field supervision for language-guided 3D affordance estimation. [NeurIPS'25]
- Developed a novel Dynamic 4D Gaussian Splatting with Diffusion guidance and dual-depth supervision for Surgical Scene Reconstruction. [MICCAI'25]
- Proposed the first VLM-based Semi-supervised Domain Adaptation paradigm, along with a novel dynamic class-balancing loss (DyCE) for semantic segmentation. [CVPR'25]
- Devised an episodic learning paradigm for unlearning domain-specific features followed by learning domain-agnostic ones for domain adaptation. [ECCV'24]
- Developed robust and adaptive Generative Latent Search (GLS) for semi-supervised domain-adaptive semantic segmentation. [MICCAI'24]
- Proposed novel semi-supervised algorithm by utilizing pseudo-labels in contrastive learning, which outperformed state-of-the-art methods by ~5% DSC score. [CVPR'23]
- Participated in medical applications of computer vision encompassing semi-supervised and self-supervised learning using very few annotations ($\leq 10\%$). [MICCAI'23]

Research Intern

ETH Zurich – Advisor: Dr. Luc Van Gool

Zurich, Switzerland

May 2020 – Aug 2020

- Collaborated on cross-image pixel contrast project to enforce pixel embeddings belonging to the same semantic class to be more similar than embeddings from different classes. [ICASSP'23]
- Guided a team of 5 undergraduate students to execute cross-image context mining for label-efficient semantic segmentation employing neural co-attention.
- Composed over 1000 lines in the ContrastiveSeg repository for understanding contextual dependencies among pixels.

Research Intern

Sorbonne University – Advisor: Dr. Daniel Racoceanu

Paris, France

Nov 2019 – Mar 2020

- Engaged with a team of 2 Postdocs and 3 Ph.D. students, utilizing Deep Learning techniques for clustering and segmentation of densely packed cells.
- Designed a MATLAB toolbox for tracking of NPC cells in culture medium with an accuracy of 93%. [\[SIRS'20\]](#)
- Coordinated with Paris Brain Institute to analyze the brain activity of human vs. non-human primates with segmentation IoU of 80%.

RESEARCH INTERESTS

Broad interest: Generative Models, Multimodal Understanding, VLM, Medical Image Analysis

Specific interest: Diffusion Models, 3D Vision & Robotics, Manipulation, RL, Domain Adaptation

EDUCATION

Stony Brook University

Doctor of Philosophy (Ph.D.) in Computer Science | Grade : 3.89/4

New York, USA

Aug. 2022 – Jan. 2026 (expected)

Jadavpur University

Bachelor of Engineering in Electrical Engineering | Grade : 8.9/10

Kolkata, India

Jul. 2017 – May 2021

TECHNICAL SKILLS

Languages: Python, MATLAB, Java, C/C++, SQL, JavaScript, HTML/CSS

Libraries: Pytorch, TensorFlow, OpenCV, Pandas, NumPy, Matplotlib

SELECTED PUBLICATIONS

- [H Basak](#), H Tabatabaee, S Gayaka, X Yang, CH Kuo, Z Yin, M Sun. Exo2EgoPolicy: Pose-Aligned Cross-View Policy Learning, **ECCV, 2026**
- [H Basak](#), Z Yin. Fixed Reality, Diffused Possibility: Disentangling Stochastic and Deterministic Latent for Cluttered Grasping, **ECCV, 2026**
- [H Basak](#), Z Yin. ViSPLA: Visual Iterative Self-Prompting for Language-Guided 3D Affordance Learning, **NeurIPS, 2025** [\[Paper\]](#)
- [H Basak](#), Z Yin. SemiDAViL: Semi-supervised Domain Adaptation with Vision-Language Guidance for Semantic Segmentation, **CVPR, 2025** [\[Paper\]](#)
- M Li, X Yang, F Wang, [H Basak](#), Y Sun, S Gayaka, M Sun, C Kuo. UA-Pose: Uncertainty-Aware 6D Object Pose Estimation and Online Object Completion with Partial References, **CVPR, 2025** [\[Paper\]](#)
- [H Basak](#), Z Yin. D⁴Recon: Dual-stage Deformation and Dual-scale Depth Guidance for Endoscopic Reconstruction, **MICCAI'25** [\[Paper\]](#)
- [H Basak](#), H Tabatabaee, S Gayaka, MF Li, X Yang, CH Kuo, A Sen, M Sun, Z Yin. Enhancing Single Image to 3D Generation using Gaussian Splatting and Hybrid Diffusion Priors. **IROS'25** [\[Paper\]](#)
- [H Basak](#), Z Yin. Forget More to Learn More: Domain-specific Feature Unlearning for Semi-supervised and Unsupervised Domain Adaptation, **ECCV, 2024** [\[Paper\]](#)
- [H Basak](#), Z Yin. Quest for Clone: Test-time Domain Adaptation for Image Segmentation by Searching the Closest Clone in Latent Space, **MICCAI, 2024** [\[Paper\]](#)
- [H Basak](#), Z Yin. Pseudo-label Guided Contrastive Learning for Semi-supervised Medical Image Segmentation, **CVPR, 2023** [\[Paper\]](#)
- [H Basak](#), Z Yin. Semi-supervised Domain Adaptive Medical Image Segmentation through Consistency Regularized Disentangled Contrastive Learning, **MICCAI, 2023** [Early Accept: top 13%][\[Paper\]](#)

- *H Basak*, S Chattopadhyay, R Kundu, S Nag, R Mallipeddi. Ideal: Improved Dense Local Contrastive Learning For Semi-Supervised Medical Image Segmentation, **IEEE ICASSP, 2023** [[Paper](#)]
- *H Basak*, S Ghosal, R Sarkar. Addressing Class Imbalance in Semi-supervised Image Segmentation: A Study on Cardiac MRI, **MICCAI, 2022** [[Paper](#)]
- *H Basak*, R Bhattacharya, R Hussain, A Chatterjee. An Exceedingly Simple Consistency Regularization Method For Semi-Supervised Medical Image Segmentation, **IEEE ISBI, 2022** [[Paper](#)]
- *H Basak*, R Kundu, R Sarkar. MFSNet: A Multi Focus Segmentation Network for Skin Lesion Segmentation, **Pattern Recognition, Elsevier** [IF: 8.518] [[Paper](#)]
- *H Basak*, Singh PK, Ahmadian A, Ferrara M, Sarkar R. Fuzzy rank-based fusion of CNN models using Gompertz function for screening COVID-19. **Scientific reports, Nature** [IF: 5.516] [[Paper](#)]
- *H Basak*, R Hussain, A Rana. DFENet: A novel dimension fusion edge guided network for brain MRI segmentation. **SN Computer Science**. [IF: 1.55] [[Paper](#)]

ACHIEVEMENTS/AWARDS

- **ICCV/CVF Travel Grant** for attending ICCV 2025 (Acceptance Rate $\sim 10\%$)
- **IEEE RAS Travel Award, 2025** for presenting work at IROS 2025 (Acceptance Rate $\sim 10\%$)
- **ECCV DEI Award, 2024** for presenting work at ECCV 2024 (Acceptance Rate $\sim 11\%$)
- **NIH Travel Grant, 2024** for presenting work at MICCAI 2024 (Acceptance Rate $\sim 15\%$)
- **IEEE SPS Scholarship, 2023 & 2024** for contributions in computer vision (Acceptance Rate $\sim 10\%$)
- **IEEE Travel Grant** for presenting work at IEEE ICASSP 2022 (Acceptance Rate $\sim 4\%$)
- **MICCAI STAR Award** for presenting work at MICCAI 2022 (Acceptance Rate $\sim 11\%$)
- **Gandhi Fellowship, 2021** by Primal Foundation India (Acceptance Rate $\sim 9\%$)
- **Charpak Fellowship, 2020** for Summer Research Internship in France (Acceptance Rate $\sim 28\%$)

ACADEMIC SERVICES

- **Teaching Assistant:** CSE377: Medical Imaging, CSE214: Intro. to Data Structure
- **Reviewer:** CVPR, ECCV, MICCAI, ICASSP, IJCAI, TMI, TIP, TNNLS, TCSVT, Pattern Recognition