Kathakoli Sengupta

Computer Science Ph.D. Student at Stony Brook University

(+1) 8573914685 | ksengupta@cs.stonybrook.edu | diasengupta.github.io | DIASENGUPTA | kathakolisengupta

Education

Stony Brook University *Stony Brook, NY*

Doctor of Philosophy in Computer Science *August 2025–May 2030(Expected)*

- **Cumulative GPA : 4.00/4.00**
- Relevant Coursework: Independent Study (Vision and Language Model) with SPELLAB, Machine Learning, Computer Vision, Database Systems, Data Science Fundamentals, Human-Computer Interaction

Boston University *Boston, MA*

Master of Science in Artificial Intelligence *September 2022–May 2024*

- **Cumulative GPA : 3.96/4.00**
- Relevant Coursework: Machine Learning, Artificial Intelligence, Data Science, Image and Video Computing, Geometric Processing, Robot Learning and Vision for Navigation, Introduction to NLP, Directed Study (Research) & Thesis with Human to Everything (H2X) Lab

Vellore Institute of Technology, Vellore *Vellore, India*

Bachelor of Technology in Electronics and Communication Engineering *July 2018–May 2022*

- **Cumulative GPA : 9.33/10.0**
- Relevant Coursework: Neural Networks and Fuzzy Control, Robotics and Automation, Sensors and Instrumentation

Publications

- **Kathakoli Sengupta**, Kai Ao, Paola Cascante-Bonilla. SceneCritic: A Symbolic Evaluator for 3D Indoor Scene Synthesis., presented at **Workshop on Open-World 3D Scene Understanding with Foundation Models (OpenSun3D), the 1st Workshop on Multi-Agent Robotic Systems (MARS), and the 2nd Workshop on Knowledge-Intensive Multimodal Reasoning @ CVPR 2026** (non-archival) ***Preprint*** [*Paper*] [*Website*]
- **Kathakoli Sengupta**, Zhongkai Shagguan, Sandesh Bharadwaj, Sanjay Arora, Eshed Ohn-Bar, Renato Mancuso. Unified Local-Cloud Decision-Making via Reinforcement Learning. ***ECCV 2024*** [*Paper*] [*Website*]
- Hee Jae Kim, **Kathakoli Sengupta**, Masaki Kuribayashi, Hernisa Kacorri, Eshed Ohn Bar. Text to Blind Motion. ***NeurIPS 2024*** [*Paper*] [*Website*]
- Hee Jae Kim, **Kathakoli Sengupta**, Masaki Kuribayashi, Hernisa Kacorri, Eshed Ohn Bar. A Multi-Modal Dataset for Urban Navigation by Blind Individuals. ***The Future of Urban Accessibility Workshop, ASSETS 2024***
- **Kathakoli Sengupta**, Ayushi Srivastava, Satyam Shreyansh, Siddhi Aggarwal, Vaegae Naveen Kumar. Driver Sleep Detection: A New and Accurate Approach. ***i-PACT 2021*** [*Paper*]
- **Kathakoli Sengupta**. Stress Detection: A Predictive Analysis. ***ASIANCON 2021*** [*Paper*]

Research Experience

Ph.D. Student | Computer Vision Lab, SBU [***Lab***] *Stony Brook, NY*

Compositional Understanding and Human Motion in 3D scenes | Supervised by Dr. Paola Cascante Bonilla *August 2025 - Ongoing*

- Training a Vision-Language Model to perform **3D scene-aware human motion generation with understanding of multistep referential language commands**.
- Developed a **novel evaluation benchmark for spatial semantics and compositionality** of 3D scenes.

Research Assistant | Human to Everything (H2X) Lab, Boston University [***Lab***] *Boston, MA*

Text to Blind Motion | Unified Local-Cloud Decision-Making | Supervised by Dr. Eshed Ohn-Bar *March 2023 - May 2024*

- Introduced a **3D multi-modal navigation dataset** for visually-impaired individuals with **high-level and low-level text annotations** by experts, a novel benchmark for future motion models.
- Trained a motion prediction model featuring a **transformer-based motion diversification module** to generate diversity among subtle actions with **RGB and textual context integration**.
- Fine-tuned SOTA **motion-language models like MotionGPT** to generalize for rare human conditions, such as lack of vision.
- Collaborated with **RedHat** to build a conditional routing module in CARLA for **dynamic task offloading** that prioritizes controls provided by a local imitation policy, saving communication and computation energy, while offloading the task to cloud for complex scenarios.
- Proposed **a multi-objective reward function** designed to optimize **energy efficiency** while **maintaining navigation performance** in a custom-designed crowded navigation environment.
- Demonstrated SOTA performance across multiple CARLA-based navigation evaluation metrics and **35% Ecological Navigation Score boost** over baselines.

Research Assistant | Bio-Imaging and Informatics (BIL) Lab, Boston University [*Lab*] *Boston, MA*

Lobe-wise Brain Age Prediction for Gulf-war affected patients | Supervised by Dr. Bang-bon Bon Koo *October 2022 - May 2023*

- Designed a **dual-channel (intensity and ravens map) Global-Local Transformer (GLT)** architecture for feature fusion resulting in improved accuracy for brain age prediction.
- Implemented **Transfer Learning** for limited dataset of Gulf-war Patients and used **3D segmentation masks** for lobe-based analysis to determine the predominant one.

Machine Learning Intern | University of Calcutta, Kolkata *Kolkata, India*

Signal Quality Analysis of ECG using Machine Learning Techniques | Supervised by Dr. Rajarshi Gupta *September 2021 - November 2021*

- Segregated noisy ECG signals using a random forest classifier with features like Shannon- Entropy and Log Energy entropy.
- Increased **9.2% performance** by balancing data and choosing proper feature space with **Adaptive Synthetic Sampling (ADASYN)** and **Principal Component Analysis (PCA)**.

Data Science and Machine Learning Intern | Indian Institute of Technology, Delhi *Delhi, India*

Prediction of Ultraviolet Radiation protection of Woven Fabrics | Supervised by Dr. Abhijit Majumdar *October 2021*

- Modeled a Ultraviolet Protection Factor(UPF) predictor and classified woven fabrics based on safety-quotient using **Random Forest** algorithm.

Industry and Teaching Experience

Teaching Assistant *Stony Brook, NY*

CSE 590 - Vision-Language Foundation Models *August 2025 - December 2025*

- Course Topics: Foundations & Representation Learning, Compositionality & Visual Reasoning, Domain Shift & Generalization, Cultural Bias & Fairness, Embodied AI & Agents, Generation & Diffusion, World Models, Efficiency & Scaling Strategies

Computer Vision Scientist *Boston, MA*

Face Live/Spoof Detection | Wicket LLC *July, 2024 - May 2025*

- Addressed generalization issues in Spoof Detector by implementing a **spoof equalization loss** to prevent overfitting to a particular type and introducing **test-time style transformations** for evaluation.
- Demonstrated a **3x performance boost** by optimizing initial layer freezing and careful hyper-parameter tuning.
- Achieved performance **improvements of 5.8%** with a patch-based spoof detector trained using **semi-supervised knowledge distillation and contrastive learning**, effectively capturing facial features and textures.

Teaching Assistant *Boston, MA*

EC-518 Robot Learning *September 2023 - December 2023*

- Course Topics: Advanced Robotics, Deep Imitation Learning, Semantic Scene Understanding, Object Detection and Tracking, Reinforcement Learning, Sim2Real, Human Robot Interaction, CARLA basics

Achievements

- Recognized as **Top Reviewer** at NeurIPS 2026
- **Reviewer** at CVPR 2026, CVPRW MAR 2026, ECCV 2026, NeurIPS 2026
- **John Hennessy Scholarship** for Doctoral Program at Stony Brook University (2025-2028)
- **Poster Presentation** at NeurIPS 2024
- **Presented** at The Future of Urban Accessibility Workshop, ASSETS 2024
- **Poster Presentation** at ECCV 2024
- **Thesis Defense**, Boston University (2024)
- **Merit Scholarship Winner** for academic excellence at VIT, Vellore (2019-2020)

Projects

Learning Spatial Representation for Efficient Robot Navigation (Masters Thesis) *September 2023 – March 2024*

Supervised by Dr. Eshed Ohn Bar | Department of Electrical and Computer Engineering

- **Improved the imitation policy for robot navigation in CARLA by 17%** with action correction and final-layer fine-tuning using RL.
- **Predicted diversified motions** for pedestrians in close proximity of the robot with **generative modeling approach (VQ-VAEs) and latent transformer** to avoid collisions.

Person Following LIMO Robot *March 2023 – September 2023*

Supervised by Dr. Eshed Ohn Bar | Department of Electrical & Computer Engineering

- Established person tracking with **Prototypical Cross Attention Network (PCAN), yielding a 36.2% IDF1 improvement** over Yolov8 with DeepOCSort, while achieving a **75.7% MOTSP score** for the ”person” class.
- Trained our LIMO bot in Gazebo to perform **image-based visual servoing with Reinforcement Learning** to achieve the person following nature.

NeRF Editing with Geometric Processing

January 2023 – April 2023

Supervised by Dr. Edward Chien | Department of Computer Science

- **Reproduced "Geometry Editing of Neural Radiance Fields"** by applying TetWild and volumeARAP to propagate user edits to the tetrahedral proxy created around the triangular mesh.
- Achieved **faster mesh generation** by using NeUS2, 95% faster algorithm than NeUS.
- Implemented **NeRF relighting** to handle the dark spots generated in unseen areas revealed during user edits to produce a fully developed edited representation.

Early-Exit Inspired Dynamic Neural Network For OOD Satellite Imaging

October 2023 - December 2023

Supervised by Dr. Iddo Drori | Department of Computer Science

- Designed **early-exit inspired neural network** with One-class SVM classifiers distinguishing Out of Distribution sample by extracting features of So2sat satellite dataset at each layer of a VGG16 backbone.
- Proved layer 2 as best for model-specific OOD detection with **90.14% AUROC** and **84.52% AUPR** and OOD confidence score as threshold for early exit.

Driver Accident Prevention System (Bachelors Thesis)

January 2022 - May 2022

Supervised by Dr. Vaegae Naveen Kumar | Department of Communication Engineering

- Forecasted potential collision with **96.52% accuracy** on Youtube driving videos using an **Inception with LSTM** model trained on CARLA simulator collision data.
- Crafted Residual architectures for drowsy and drunk driver face detection with OpenCV-captured faces and SVM based fire hazard detector for gaseous component analysis.

Extracurricular Activities

- **Chairperson** of IEEE Signal Processing Society, VIT and organizer of HackX hackathon (2021)
- **Technical Head** of IEEE Robotics and Automation Society, VIT and coordinator of Reboot hackathon (2020)
- **Coordinator** of Building Chatbot: Expertise from Scratch (GraVITas2k20)
- **Program Representative**, School of Electronics Engineering, VIT (2019-2020)