

EDUCATION

Johns Hopkins University, MD, USA

Ph.D. in Electrical and Computer Engineering

August 2026 - Present

University at Buffalo, NY, USA

MS in Computer Science and Engineering(Research Track)

GPA - 4.0/4.0

August 2024 - May 2026

Indian Institute of Technology(IIT), Jodhpur, India

B.Tech in Artificial Intelligence and Data Science

GPA - 3.7/4.0

December 2020 - May 2024

RESEARCH EXPERIENCE

Johns Hopkins University, Baltimore, MD

Research Assistant - VIU, **Advisor:** Prof. Vishal M. Patel

[Aug 2026 - Present]

- Leading research efforts for the IARPA VideoLINC grant, focusing on general object re-identification.

University at Buffalo, Buffalo, NY

Graduate Student Assistant - CUBS, **Advisor:** Prof. Nalini. K. Ratha

Best Student Paper Award

[Aug 2024 - Present]

- Reduced inference time of Resnet-18 by **12.75%** by replacing convolution layer with box-filters and integral image.
- Created a Binary CNN for face recognition using custom CUDA-kernels, thereby leading to a 128-bit Fully Homomorphic Encryption(FHE)-based secure convolution operator reducing inference time by **31.46%** over SOTA.
- Introduced a novel dataset for secure multimodal entity resolution with FHE and achieved an EER of **4.08%**.
- Formulated token reduction problem grounded in optimal transport for long-video understanding.
- Worked on layout-to-image generation with **Diffusion Models** emphasizing failures due to memorization.
- Technologies Used: **Pytorch-Lightning**, **Jupyter**, **MMAction**, **Docker**, **ffmpeg**, **HElib**

Indian Institute of Technology(IIT), Jodhpur, India

Undergraduate Student Researcher - IAB Lab, **Advisors:** Prof. Richa Singh & Prof. Mayank Vatsa

Rising Star Award, BP Award

[Aug 2022 - July 2024]

- Established a comprehensive study for the impact of White and Black-Box Adversarial Attacks on different SOTA CNNs and Transformers (ViT & Swin-T) upon various Vision Benchmarks like CelebA, LFW & ImageNet200.
- Designed a vision transformer based detection network to differentiate adversarial and unintentional noise via MMD and Center loss achieving **~ 100%** detection accuracy on seen and unseen adversarial attacks on biometric data
- Explored the asymmetry of **diffusion models** for generating universal adversarial perturbations in a black-box setting achieving ASR of **97.47%** on WideResnet-28-10 and reducing generation time by **94.73%** on ImageNet.
- Technologies Used: **PyTorch**, **ART**, **HuggingFace**, **RobustBench**, **CleverHans**

University of Siegen, Siegen, Germany

Research Intern - Learning2Sense, **Advisor:** Prof. Michael Moeller

Github

[August 2023 - Jan 2024]

- Formulated a non-uniform, intensity-dependent noise model for CT sinograms using Beer-Lambert law, replacing the standard additive noise assumption with one that accounts for variations in rays across angles and offsets.
- Optimized a softmax-parameterized distribution of ray intensities under a fixed radiation budget and used filtered back-projection to assess how this mathematically constrained allocation improves reconstruction quality.
- Technologies Used: **PyTorch**, **ASTRA Toolbox**, **Matplotlib**

University of Alberta, Edmonton, Canada

MITACS Visiting Researcher - Vision and Learning Lab, **Advisor:** Prof. Li Cheng

SGA.ai

[May 2023 - August 2023]

- Improved object detection models like *EfficientDet* through self-attention mechanism and selective feature propagation to the Bi-FPN for accurate wheat-head growth stage detection by **37%** through analysis of physical properties and Finetuned *Yolov8* by large number of experiments for high-precision canola flower count.
- Spearheaded the deployment of the trained models through AWS S3 for real-time usage and performed data collection, annotation, and analysis to create a extensive and varied plant dataset for the company.
- Technologies Used: **AWS**, **Pytorch**, **Javascript**, **XML**, **Ultralytics**, **SCVAT**

Indian Institute of Technology, Guwahati, India

Research Intern - Vision Intelligence Lab, **Advisor:** Dr. Santosh K Vipparthi

Github

[May 2022 - Jan 2023]

- Proposed an inter-frame movement detector module, which extracts the movement information between the consecutive frames and helps integrate temporal information with spatial visual features for high fps videos.
- Developed a novel CNN-based architecture using an encoder-decoder module, DeepLab-V3 model and BERT thereby facilitating late feature fusion, reducing the parameter space and increasing Jaccard metric by **~2%**.
- Technologies Used: **PyTorch, OpenCV, Pycocotools**

PUBLICATIONS

- **Coreview: Compact Yet Complete Video Representation** (paper, code)
Roy, S.*, Kaushik, A.*, Ratha, N., Govindaraju, V.
2026. In IEEE International Conference on Image Processing (**IEEE ICIP'26**).
- **Efficient and Secure Convolutions on Encrypted Data** (paper, code)
Roy, S.*, Yalavarthi, B., Ratha, N.
2026. In IEEE International Conference on Image Processing (**IEEE ICIP'26**).
- **Multimodal Privacy-Preserving Entity Resolution with Fully Homomorphic Encryption** (Oral Accepted — paper, code)
Roy, S.*, Ratha, N.
2026. In IEEE International Conference on Acoustics, Speech and Signal Processing (**IEEE ICASSP'26**).
- **TAIGen: Training-Free Adversarial Image Generation via Diffusion Models** (Oral Accepted — paper, code, slides)
Roy, S.*, Jain, A., Vatsa, M., Singh, R.
2025. In Proceedings of IEEE/CVF International Conference on Computer Vision (**IEEE/CVF ICCVw'25**).
- **EffBin: Efficient Face Recognition via Binary Neural Networks** (Oral Accepted — paper, slides)
Roy, S.*, Yalavarthi, B., Ratha, N.
2025. In IEEE Conference on Artificial Intelligence (**IEEE CAI'25**).
- **Efficient Convolutions using Box-Filters for Low-Latency Image Processing** (Oral Accepted — paper, slides)
Roy, S.*, Yalavarthi, B., Ratha, N.
2025. In IEEE Western New York Image and Signal Processing Workshop (**IEEE WNYISPW'25**).
- **Discerning the Chaos: Detecting Adversarial Perturbations while Disentangling Intentional from Unintentional Noises** (Oral Accepted — paper, code, slides, poster)
Jain, A.*, Roy, S.*, Vatsa, M., Singh, R.
2024. In Proceedings of the 8th IEEE International Joint Conference on Biometrics (**IEEE IJCB'24**).
- **RefMOS: A Robust Referred Moving Object Segmentation framework based on text query** (paper, code, poster, slides)
Saxena, PP.*, Roy, S.*, Tyagi, DK., Vipparthi, SK., Balasubramanian, R., Murala, S.
2024. 20th IEEE International Conference on Advanced Video and Signal-Based Surveillance (**IEEE AVSS' 24**).

PROJECTS

- **AVDDetector: Audio-Visual Multimodal Deepfake Detector and Temporal Localizer** Leaderboard
ACM Multimedia'25 AVDeepfake Challenge, University at Buffalo — **Advisor:** Prof. Nalini Ratha [May 2025 – Jul 2025]
 - Led algorithmic development and data engineering of multimodal deepfake detector achieving a **top-10** rank.
 - Used VideoMAEv2, AudioSSL models and magnified video frames increasing **mAP** by **25%** and **AP** by **40%**.
- **MagFormer: Parameter-Efficient Modular Approach for Person Re-Identification** Github
Deep Learning Course Project, University at Buffalo — **Advisor:** Prof. Kaiyi Ji [Jan 2025 – April 2025]
 - Developed a two-stage architecture using Resnet-50 and **LoRA**-based approaches for multi-head self-attention
 - Used triplet and center loss functions, achieving mAP & top-10 acc of **92.7%** & **99.5%** respectively on Market-1501
- **Scalable and Efficient LLM-RAG Service for Conversational AI in Pepper Robot** Github
Self Project, University at Buffalo [Jan 2025 – April 2025]
 - Implemented end-to-end RAG with **LangChain** using document chunking, **BGE** and Instructor embeddings, **FAISS** vector store, and a similarity-threshold conversational retriever with grounded sources, served via **FastAPI**.
 - Deployed **Llama-2 13B** using **4-bit bitsandbytes** quantization and multi-GPU device mapping, tuned decoding for latency/quality, and integrated **OpenAI** tool-calling as fallback; increasing contextual response quality by **60%**.
- **A Sequential Memory Preserving Approach for Few-Shot Image Classification** Github
Adv. ML Course Project, IIT Jodhpur — **Advisor:** Prof. Mayank Vatsa [Aug 2023 - Dec 2023]
 - Modelled the meta-training set by combining all task-specific training sets to learn a richer embedding space.
 - Used Matching Feature Hierarchy and Memory Wrap Modules for hierarchical feature refinement, achieving **~80%** on 5-shot MiniImageNet and **~72%** on 1-shot CIFAR-FS, outperforming 2020 SOTA baselines like MetaOptNet.

TECHNICAL SKILLS

- **Languages:** Python, C/C++, Java, JavaScript, SQL, R, ROS2, Go, CUDA, VueJS, PHP, Matlab, GIT, Pandas
- **Technologies:** PyTorch, Tensorflow, Hadoop, Scikit-Learn, OpenCV, PySpark, Docker, Kubernetes, AWS, MongoDB
- **Software Tools:** Visual Studio Code, Android Studio, Jupyter Notebook, MLlib, Cursor, Tensorboard, W&B

TEACHING, VOLUNTEERING & AWARDS

- **Teaching Assistantship:** Served as a teaching assistant at IIT Jodhpur for 200+ Sophomore & Junior year students, conducting weekly lab and viva sessions, preparing assessments and grading them for the following courses:
 - Deep Learning(Under Dr. Angshuman Paul) [Jan '24 - May '24]
 - Pattern Recognition and Machine Learning(Under Dr. Richa Singh and Dr. Pratik Mazumder) [Jan '23 - May '23]
- **Volunteering:** **Reviewer** at **BioSig'26**, **BMVC'26**, **CVPRw'25** and Conference Volunteer at **IJCB'24**.
- **Awards:** Received the **Patricia J. Eberlein Master's Thesis Award** at UB, **Best Student Paper Award** at IEEE WNYISPW'24, **Broadening Participation Award** at ICCV'25, **Rising Star Award** at IJCB'24 and **MITACS Research Scholarship** for summer 2023.