Susim Mukul Roy

Email : roysusim7@gmail.com
Contact : +1(716)-939-8558

Linkedin — Github — Portfolio — Google Scholar

EDUCATION

University at Buffalo, NY, USA	**GPA - 4.0/4.0**
MS in Computer Science and Engineering(Research Track)	*August 2024 - May 2026*

Indian Institute of Technology(IIT), Jodhpur, India	**GPA - 3.7/4.0**
B.Tech in Computer Science and Engineering	*December 2020 - May 2024*

RESEARCH EXPERIENCE

- **University at Buffalo, Buffalo, NY** *Best Student Paper Award*
  Graduate Research Assistant - CUBS, **Advisor:** *Prof. Nalini. K. Ratha* *[Aug 2024 - Present]*
  - Reduced inference time of Resnet-18 by **12.75%** with box-filters and created an accurate Binary CNN-based face recognition system. Created a 128-bit FHE-based secure convolution operator reducing inference time by **31.46%**.
  - Developed a multimodal deepfake detection and localization model using SoTA VideoMAEv2, AudioSSL pre-trained extractors and custom attention mechanisms increasing AUC by **40%** on the AVDeepfake1M++ dataset.
  - Working on visual reasoning in **VLMs**, controllable image generation with **DMs** and long-video understanding.
  - Technologies Used: **Pytorch-Lightning, Jupyter, HuggingFace, Docker, ffmpeg, InsightFace, ONNX**

- **Indian Institute of Technology(IIT), Jodhpur, India** *Rising Star Award, BP Award*
  Undergraduate Student Researcher - IAB Lab, **Advisors:** *Prof. Richa Singh & Prof. Mayank Vatsa* *[Aug 2022 - May 2024]*
  - Established a comprehensive study for the impact of White/Black-Box Adversarial Attacks on different SOTA CNNs and Transformers (ViT & Swin) upon various Vision Benchmarks (CelebA, LFW & ImageNet200).
  - Designed a vision transformer based detection network to differentiate adversarial and unintentional noise via MMD and Center loss achieving $\sim$ **100%** detection accuracy on seen and unseen adversarial attacks on biometric data
  - Explored the asymmetry of **foundation 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 Alberta, Edmonton, Canada** *SGA.ai*
  MITACS Visiting Researcher - Vision and Learning Lab, **Advisor:** *Prof. Li Cheng* *[May 2023 - August 2023]*
  - Improving 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 finetuning *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** Github
  Research Intern - Vision Intelligence Lab, **Advisor:** *Dr. Santosh K Vipparthi* *[May 2022 - Jan 2023]*
  - Proposed an inter-frame movement detector (IFCD) 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 J&K metric by **10%**.
  - Technologies Used: **PyTorch, OpenCV, Pycocotools**

- **University of Seigen, Seigen, Germany** Github
  Research Intern - Learning2Sense, **Advisor:** *Prof. Michael Moeller* *[August 2023 - Jan 2024]*
  - Did a theoretical analysis on non-uniform noise modelling for the *sinogram* generated from forward propagation algorithm for CT Images via the Beer-Lambert law and researched on existing methods to model the same.
  - Worked on suitable reconstruction technique like *FPB* on noisy images(*Poisson/Gaussian*) for ULDCT with a constrained loss function and a parameterized version of the pixel intensities for a smoother convergence.
  - Technologies Used: **PyTorch, ASTRA Toolbox, Matplotlib**

PUBLICATIONS

- *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(**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(**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(**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(**IJCB'24**).

- *RefMOS: A Robust Referred Moving Object Segmentation framework based on text query* (Poster Accepted — 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(**AVSS' 24**).

PROJECTS

- **MagFormer: Parameter-Efficient Modular Approach for Person Re-Identification** *Github*
  *Deep Learning, 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** function/tool-calling as fallback; increasing contextual response quality.

- **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, Rust, 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, Xcode, Cursor, Tensorboard, W&B

TEACHING AND VOLUNTEERING EXPERIENCE

- **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:
  - Pattern Recognition and Machine Learning(Under Dr. Richa Singh and Dr. Pratik Mazumder) *[Jan '23 - May '23]*
  - Deep Learning(Under Dr. Angshuman Paul) *[Jan '24 - May '24]*

- **Volunteering Experience:**
  - **Reviewer** at **CVPRw'25**, **NeurIPS'25**, **ACM MM'25** and Conference Volunteer at **IJCB'25**.
  - **Captain** of *Robotics Society* *[Dec '21 - Feb '23]*
    Overall Management of society by organising sessions on ML, ROS, Mechanics of robots etc. along with leading different teams for inter-collegiate competitions and mentoring society members in club projects.
  - **Subsection Head** of Team IITJ for *ABU-Robocon'22* *[Nov '21 - Apr '22]*
    Integrating AI with ROS for proper detection of disks to perform the necessary path planning of the Lagori Robots.