Sandeep Kumar Routray

sroutra2@cs.cmu.edu | (412) 450-2068 | linkedin.com/in/sandeep-routray | sroutray.github.io

EDUCATION

Carnegie Mellon University **Dec. 2025**
Master of Science in Machine Learning *GPA: 4.00 / 4*
Relevant Coursework: Deep Reinforcement Learning, Probabilistic Graphical Models, Multimodal Learning

Indian Institute of Technology Kanpur **May 2021**
Bachelor of Technology in Electrical Engineering | Department Rank: 2 *GPA: 4.00 / 4*
Relevant Coursework: Data Structures, Algorithms, NLP, Digital Signal Processing, Optimization Algorithms
Honors: Dean's List (all semesters), Summa Cum Laude, Gold Medal for Undergrad Project

PUBLICATIONS

[1] S. Routray, H. Pan, U. Jain and D. Pathak, **"ViPRA: Video Prediction for Robot Action",** in International Conference on Learning Representations (ICLR), April 2026 [*Under Review*]
[2] S. R. Dash*, S. Routray*, P. Varshney* and A. Modi, **"CS-NET at SemEval-2020 Task 4: Siamese BERT for ComVE"**, in Proceedings of the Fourteenth Workshop on Semantic Evaluation, International Committee for Computational Linguistics (COLING), Dec 2020

EXPERIENCE

Skild AI **May 2025 - Present**
Research Intern, Humanoid Team *Pittsburgh, PA*
- Developing a robotics foundation model using robot demos and unlabeled human videos, with efficient inference for real-world deployment

Carnegie Mellon University | [ViPRA] **Oct. 2024 - May 2025**
Graduate Research Assistant, Prof. Deepak Pathak's Lab *Pittsburgh, PA*
- Adapted **multimodal LLMs** and **video diffusion models** for predicting robot actions from internet-scale human and robot videos
- Developed ***ViPRA***, enabling few-shot generalization via **latent action** learning with VQVAE and **flow matching policy** for robot control
- Outperformed baselines with **16%** SIMPLER benchmark gain and **14%** real-world task improvement with minimal demonstrations

Samsung Research | [CES 2024] **Sep. 2021 - June 2024**
Machine Learning Engineer, SmartThings Team *Seoul, South Korea*
- Spearheaded 3D home layout reconstruction project. Showcased at **CES 2024** and deployed across **1 million** homes globally
- Trained **ConvNeXt with focal loss** for structure detection; optimized with TF Lite and quantization for **4×** smaller, **3×** faster inference
- Enabled single-image 3D reconstruction using **NeRFs** trained on multiple synthetic views generated from a **video diffusion model**

Vector Institute for Artificial Intelligence | [Report] **Oct. 2020 - July 2022**
Research Fellow, Prof. Sanja Fidler's Lab *Toronto, Canada*
- Leveraged inter-image relationships in a **Slot Attention** framework to learn object-centric features with self-supervised learning (SSL)
- Created an image context aware score function to mine positives and negative slots for **contrastive loss** to improve feature consistency
- Ablated **vision transformers** training with SSL losses on multi-GPU clusters, obtained **2 % mIoU** improvements over existing baselines

Samsung Research | [Report] **May 2020 - July 2020**
Software Engineer Intern, 6G Research Team *Seoul, South Korea*
- Built an LTE scheduler using **DDPG reinforcement learning** and custom rewards to optimize throughput, delay, and user fairness
- Obtained **80% lower delay** and better **user scalability** than prevalent Proportional Fair scheduler without compromising data rates

PROJECTS

Simulator-based Scaling of Inference Time Compute for Robotics **Jan. 2024 - Present**
- Scaling inference-time compute for robotics by combining **Chain-of-Thought** reasoning with diffusion transformer **world model** rollouts
- Attained **30%** improvement with **model-based RL** and **reward modeling** to optimize trajectory search and policy performance at test time

Common Sense Validation And Explanation | [Paper] **June 2020 - Dec. 2020**
- Proposed a **Siamese architecture** and **Mixture-of-Experts** with encoder based **LLMs** for efficient inter-relational information extraction
- Coupled with cross attention, achieved **94.8%** accuracy for Validation task and **89%** for Explanation task. Results published in **COLING '20**

SKILLS

Languages: C, **C++, Python**, MATLAB, SQL
Technologies: Docker, Git, **PyTorch**, **JAX**, **CUDA**, ONNX, TF Lite, NetworkX, OpenCV, Fast APIs