Sandeep Kumar Routray

MS in Machine Learning | Carnegie Mellon University

(+1) 412-450-2068 | sroutra2@cs.cmu.edu | sroutray.github.io | sroutray | sandeep-routray

Interests

Computer Vision, Vision Language Models (VLMs), World Models, Multi-modal Learning, Reinforcement Learning, Robotics

Education

Carnegie Mellon University *August 2024 - December 2025*

Master of Science | Machine Learning *GPA: 4.00/4*

- **Key Courses -** Intermediate Statistics, Advanced ML, Deep Reinforcement Learning

Indian Institute of Technology Kanpur *July 2017 - May 2021*

Bachelor of Technology | Electrical Engineering | Department Rank 2 *GPA: 9.84/10*

- Awarded **Academic Excellence Award 2017-18, 2018-19, 2019-20** Equivalent to Dean's List
- Awarded **Prof. Samares Kar Memorial Gold Medal** for the best undergraduate project
- **Key Courses -** Data Structures, Algorithms, Machine Learning, Probabilistic ML, NLP, Optimization Algorithms, Digital Signal Processing

Publications

[1] 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. [Paper]

[2] N. V. Deshpande*, S. Routray* and A. K. Gupta, "Spectral Efficiency in Poisson Cluster Based HetNets with Users and Base stations Correlation", in *IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, Dec 2020. [Paper], [Video]

Work Experience

Carnegie Mellon University *PA, US*

Graduate Research Assistant | *Prof. Deepak Pathak's Lab* *October 2024 - Present*

- Research on leveraging VLMs to predict actions from unlabeled human and robot videos, enabling task and environment generalization.

Samsung Research [Press] *Seoul, South Korea*

Machine Learning Engineer | *SmartThings Team* *September 2021 - June 2024*

- Spearheaded *Map View* project to convert home layouts to 3D model. Showcased at **CES 2024** and deployed across **1 million homes** globally.
- Trained a **ConvNext** model using **focal loss** to identify rooms, walls, doors and junctions and designed a custom raster to vector pipeline.
- Performed integer quantization for mobile deployment using TF Lite C API achieving **4x reduction** in size and **3x increase** in inference speed.
- Worked on 3D reconstruction from single image by training a **neural radiance field** (NeRF) on multi-views generated from a diffusion model.

Vector Institute for Artificial Intelligence *Toronto, Canada*

Research Fellow | *Prof. Sanja Fidler's Lab* *October 2020 - July 2022*

- Leveraged inter-image relationships in a **Slot Attention** framework to learn object-centric features in a self-supervised manner.
- Designed image context aware score function to mine positives/negative slots from a queue for **contrastive loss**, improving feature consistency.
- Fine-tuned **vision transformers** on multi-GPU clusters using our framework, obtained **1 % mIoU improvements** over existing benchmarks.

Samsung Research [Report] *Seoul, South Korea*

Intern | *6G Research Team* *May 2020 - July 2020*

- Implemented a reinforcement learning based resource scheduler for LTE system using **Deep Deterministic Policy Gradient** (DDPG) algorithm.
- Devised two reward mechanisms to **maximize throughput** while maintaining QoS requirements of delay and fair allocation among users.
- Achieved lower delay (upto **80% lower**) and better scalability than the prevalent Proportional Fair scheduler without compromising data rates.

Key Projects

Joint Learning of Dense Representations And Object-Part Relationships *Vector Institute, Canada*

Research Intern | *Prof. Sanja Fidler* *August 2020 - December 2021*

- Incorporated **geometrical cues** from 2D-mesh decomposition of image, and performed **hierarchical grouping** for object-part relationships.
- Performed self-supervised pre-training on multiple HPC nodes in distributed data parallel mode to learn dense representations.

Common Sense Validation And Explanation [Paper] *IIT Kanpur, India*

Undergraduate Project | *Prof. Ashutosh Modi* *January 2020 - May 2020*

- Designed a **Siamese architecture** with encoder based **LLMs** like BERT, RoBERTa to enable efficient inter-relational information extraction.
- Coupled with cross attention, achieved **94.8%** accuracy for Validation task and **89%** for Explanation task. Results published in **SemEval-2020**.

Minimax Optimization in Non-Euclidean Space Using Bregman Divergence [Slides] *IIT Kanpur, India*

Undergraduate Project | *Prof. Ketan Rajawat* *August 2020 - December 2020*

- Designed a novel restarting algorithm to minimize smooth, strongly convex functions in non-Euclidean space using Nesterov's AGD.
- Proposed a new algorithm for smooth minimax optimization using the above result. Improved convergence rate by **order of 2** in both cases.

Skills

Programming Languages/Tools Python, C, C++, MATLAB, Bash, Linux, Git, SQL, HPC, Fast API

Libraries PyTorch, CUDA, NetworkX, Keras, TensorFlow, OpenCV, NumPy, ONNX, TF Lite