

# Ranajoy Sadhukhan

Carnegie Mellon University

☎ 412-430-1472 ✉ [rsadhukh@andrew.cmu.edu](mailto:rsadhukh@andrew.cmu.edu) [in ranajoy-sadhukhan](https://www.linkedin.com/in/ranajoy-sadhukhan) [🌐 ranonrkm](https://github.com/ranonrkm)

## RESEARCH INTEREST

---

Efficient Machine Learning, Information Retrieval, Natural Language Processing.

## EDUCATION

---

- **Carnegie Mellon University** 2023-present  
*PhD Student, Electrical and Computer Engineering* GPA: 3.92/4.0  
*Advisor: Dr. Beidi Chen*
- **Indian Institute of Technology Kharagpur** 2016-2021  
*Dual-degree(B.Tech+M.Tech) in Electrical Engineering, Specialization in Signal Processing* GPA : 9.39/10  
*Minor in Computer Science & Engineering* GPA : 9.92/10

## PUBLICATIONS

---

- **MagicDec: Breaking the Latency-Throughput Tradeoff for Long Context Generation with Speculative Decoding**  
by **Ranajoy Sadhukhan\***, Jian Chen\*, Vashisth Tiwari, Zhuoming Chen, Ruihang Lai, Jinyuan Shi, Ian En-Hsu Yen, Avner May, Tianqi Chen, Beidi Chen  
*In Submission to International Conference on Learning Representations 2025* [\[Paper\]](#) [\[Code\]](#)
- **MagicPIG: LSH Sampling for Efficient LLM Generation**  
by Zhuoming Chen, **Ranajoy Sadhukhan**, Zihao Ye, Yang Zhou, Jianyu Zhang, Niklas Nolte, Yuandong Tian, Matthijs Douze, Léon Bottou, Zhihao Jia, Beidi Chen  
*In Submission to International Conference on Learning Representations 2025* [\[Code\]](#)
- **Memory Mosaics**  
by Jianyu Zhang\*, Niklas Nolte\*, **Ranajoy Sadhukhan**, Beidi Chen, Léon Bottou  
*In Submission to International Conference on Learning Representations 2025* [\[Paper\]](#) [\[Code\]](#)
- **Taxonomy Driven Learning Of Semantic Hierarchy Of Classes**  
by **Ranajoy Sadhukhan**, Ankita Chatterjee, Jayanta Mukhopadhyay, Amit Patra  
*In IEEE International Conference on Image Processing 2022* [\[Paper\]](#) [\[Code\]](#)
- **Knowledge Distillation Inspired Fine-Tuning of Tucker Decomposed CNNs and Adversarial Robustness Analysis**  
by **Ranajoy Sadhukhan**, Abhinav Saha, Jayanta Mukhopadhyay, Amit Patra  
*In IEEE International Conference on Image Processing 2020* [\[Paper\]](#) [\[Code\]](#)

## RESEARCH EXPERIENCE

---

- **Memory-Efficient Approximate Nearest Neighbor Search (ANNS)** Aug '21 - Jan '23  
*Advisors - Dr. Harsha Vardhan Simhadri & Dr. Pratyush Kumar* Microsoft Research India
  - Developed retrieval metric-aware learnable Product Quantization (PQ) for memory-efficient and accurate dense retrieval (64× compression ratio).
  - Devised a novel graph ANNS-adaptive PQ learning algorithm for low-latency retrieval.
- **Improving Graph-based ANNS** Feb '22 - July '23  
*Advisors - Dr. Harsha Vardhan Simhadri & Dr. Manik Varma* Microsoft Research India
  - Improved index-building strategy for DiskANN, a C++ based fast, scalable graph ANNS algorithm, to enable better performance on out-of-distribution queries. The proposed method offers up to 45% latency improvement over the state-of-the-art methods at comparable recall on 100 Million scale databases. This work was submitted to **Neurips'23 Big-ANN-Benchmarks** competition track.

- **Depth Estimation & Panoptic Segmentation** May - July '20  
*Advisor - Pankaj Kumar Bajpai* *Samsung R&D Institute Bangalore*
  - Developed an efficient and light-weight (6 MB) Deep Neural Network for the joint task of monocular depth estimation and panoptic segmentation of road scene images.
  - Devised adaptive batch sampling strategy and novel ways of combining task specific losses to alleviate the issue of asymmetrically annotated datasets with an uneven number of annotations per modality.
- **Medical Image Analysis** May - July '19  
*Advisor - Dr. Mirza Faisal Beg* *Simon Fraser University — MITACS Globalink Internship*
  - Developed a high throughput image Quickcheck generation pipeline for displaying information within CT images and for helping the raters to check the images and their segmentation annotations for quality control.

## ACADEMIC RESEARCH

---

- **HSD-CNN : Hierarchically Self Decomposing CNN** Feb '20 - May '21  
*Advisor - Dr. Jayanta Mukhopadhyay* *Master's Thesis Project*
  - Explainable hierarchical decomposition of pre-trained CNNs using a novel semantic loss function, creating efficient sub-networks that can achieve upto 4x speedup and 75% parameter reduction for limited-class tasks without fine-tuning.
- **Model Compression of Deep Neural Networks** Dec '18 - Jan '20  
*Advisor - Dr. Jayanta Mukhopadhyay* *Bachelor's Thesis Project*
  - Effective fine-tuning of Tensor Decomposed CNNs based on Knowledge Distillation, providing better accuracy and robustness against white-box adversarial attacks than the state-of-the-art tensor decomposition methods.

## SOFTWARE SKILLS

---

- **Programming Languages:** C/C++, Python, MATLAB, Bash, CUDA
- **Libraries:** Pytorch, OpenMP, AVX, MKL
- **Frameworks:** Git, Simulink, Ruby on Rails, LTspice

## KEY COURSES UNDERTAKEN

---

- **Computer Science:** Algorithms-I, Advanced Machine Learning, Advanced Image Processing and Computer Vision, Computer networks, Computer architecture and Operating Systems, Parallel Computer Architecture and Programming, Deep Learning Systems.
- **Mathematics:** Advanced Statistics, Intermediate Statistics, Linear Algebra, Probability and Stochastic processes, Transform Calculus.
- **Electrical:** Analog Electronics, Digital Electronics, Embedded systems, Digital Signal Processing, Statistical Signal Processing, Analog Signal Processing, Mixed Signal Circuits and Systems on-chip.

## TEACHING EXPERIENCE

---

- **18-789: Deep Generative Modeling** Jan '24 - May '24  
*Instructors: Dr. Beidi Chen, Dr. Giulia Fanti*
- **EE19001: Electrical Technology Lab** Dec '20 - Mar '21  
*Instructors: Dr. Souwik Chattopadhyay, Dr. Dipankar Debnath*

## ACHIEVEMENTS

---

- Selected for the highly decorated **Mitacs Globalink Research Internship Program** and **DAAD Research Internship Program** (2019).
- Awarded **Best Project Award** for **Masters Thesis Project** in Electrical Engineering at IIT Kharagpur.
- Awarded **Merit-cum-Means scholarship, IIT Kharagpur**, for academic excellence among 1400 students (2016).
- Secured an All-India-Rank of 3730 in JEE Advanced amongst 150,000 candidates (2016).
- Secured an All-India-Rank of 1457 in JEE Mains amongst 13,00,000 candidates (2016).
- Stood 4th in Advanced Mathematical Ability Test organized by **Calcutta Mathematical Society** (2015).
- Part of Bronze winning team, unit 3 at National Social Services(NSS) Annual Camp, 2016.