Sreetama Sarkar
2638 Portland Street, Room 206C
Los Angeles, California 90007
Email-id : sreetama@usc.edu / sreetama1994@gmail.com
Mobile No.: +12135744381/ +919432822839

ACADEMIC DETAILS

Examination/Degree	University	Year	GPA/%
PhD: *Computer Engineering*	University of Southern California	Ongoing	3.92/4.0
Master of Science: *Communications Engineering*	Technical University of Munich	2021	1,2 (German Grading System)
Bachelor of Technology: *Electronics & Communication Engineering*	NIT Durgapur, India	2017	9.63/10
12th Std. (Higher Secondary)	WBCHSE	2013	88.2
10th Std. (Madhyamik)	WBBSE	2011	88.5

AREAS OF INTEREST

Energy-efficient ML for edge devices, Adversarial Attacks and Defense, Secure and Trustworthy AI

EXPERIENCE

Graduate Research Assistant, *University of Southern California* (August 2022 - Present)

– On-device Training for Vision Transformers for transfer learning tasks
– Robust linearization of Convolutional Neural Networks for Efficient Private Inference
– Intelligent region-skipping for object detection as part of the project *Energy-Efficient Event-based SNNs for Complex CV Applications*, funded by Intel; results showcased in monthly presentation to Intel
– Intelligent frame-skipping for object detection and tracking as part of the RPIXELS project funded by DARPA; results showcased in demo presentation to DARPA
– Investigating efficient ML methods for bad pixel detection and correction as part of *In-pixel Computing* project, funded by Samsung
– Helped in writing research grants for NASA and NSF

Graduate Teaching Assistant, *University of Southern California* (January 2024 - Present)
EE354L: Introduction to Digital Circuits

– Digital system design and implementation; synchronous design of datapath and control; schematic/Verilog-based design, simulation, and implementation in FPGAs; timing analysis; semester-end project.

Intern, *Intel Labs, Munich* (May 2021 - September 2021)

– Investigating performance metrics for object detection in safety-critical scenarios
– Identified shortcomings of existing metrics like mean average precision (mAP)
– Proposed new metrics considering object distance, safety-critical clustering and object tracking, with a scope for further refinement considering other safety aspects

Master's Thesis Student, *Autonomous Driving Group, BMW Munich* (September 2020 - April 2021)
Robustness aware Pruning methods for Convolutional Neural Networks (CNNs)

– Developing a Pruning framework that simultaneously optimizes the three objectives: task-specific performance, adversarial robustness and resource-awareness
– Combined adversarial training and model pruning in a joint formulation of the fundamental learning objective during training, termed In-train Pruning
– Achieved state-of-the-art adversarial robustness and an improvement in natural accuracy without any additional overhead of GPU hours; results published in the CVPR Workshop SAIAD 2021

Working Student, *Intel Communications and Devices Group (iCDG), Intel Munich* (March 2019 - August 2019)

– Analyzing traces in System Trace Tool (STT) and automating process to decode events and identify potential problems in the Protocol Processing Unit (PPU)

Research Assistant, *School of Computer Science and Engineering (SCSE), Nanyang Technological University, Singapore* (November 2017 - September 2018)

– Part of a team responsible for the design of a state-of-the-art Secure Processor using RISCV ISA that implements features like secure boot, memory encryption etc.

Summer Intern, CERN, Geneva, Switzerland (May 2016–July 2016)
  – Quality Assurance of GEM Detectors at ALICE (A Large Ion Collider Experiment)

ACADEMIC PROJECTS

On-device Training in Vision Transformers *(October 2023 – Present)*
Energy Efficient Secure Sustainable Computing Group, University of Southern California
  – Memory limitations on edge devices pose challenges for on-device fine-tuning while fine-tuning in the cloud leads to increased communication latency and compromise in user data privacy
  – We devise an approach for selectively fine-tuning a small fraction of blocks of a pre-trained models coupled with token dropping resulting in reduced compute, training time and activation memory
  – Compared to the existing alternatives, our approach yields additional training memory saving of up to $1.4\times$ and compute cost by up to $2\times$

Robust Linearization for Efficient Private Inference*(January 2023 – August 2023)*
Energy Efficient Secure Sustainable Computing Group, University of Southern California
  – Linearization approaches like reducing ReLU in CNNs and Softmax in ViTs are devised to reduce latency of cryptographic primitives in private inference (PI) frameworks, but their robustness remains unexplored
  – We propose latency-efficient models with improved performance on clean, natural as well as adversarially perturbed images that outperform SoTA models on all three fronts

In-Pixel Computing for Efficient Low-energy Heterogeneous Systems *(August 2022 – Present)*
Energy Efficient Secure Sustainable Computing Group, University of Southern California
  – Vast amount of data in videos creates a bottleneck in storage as well as data-transfer bandwidth from cameras to computer vision systems for real-time object detection and tracking
  – To address this, we devise an intelligent approach, considering relative motion of objects between frames, to skip frames and regions in videos without degrading performance

Hardware aware Automated Pruning using Genetic Algorithms *(March 2020 – August 2020)*
Project Lab Integrated Systems, Technical University of Munich
  – Analyzing effectiveness of genetic pruning compared to RL based pruning like AMC pruning
  – Detailed study of various automated Model Compression techniques, specially pruning methods available in literature like heuristic approaches, RL based pruning and comparison with genetic pruning
  – Genetic pruning proves to be faster for large datasets like ImageNet; also NSGA provides multi-objective optimization and pareto-optimal solutions

Impact and Sustainability of Government Policy Decisions on CO_2 Emissions during COVID-19 pandemic *(March 2020 – August 2020) Applied Machine Intelligence Course, Technical University of Munich*
  – Investigation of government policies and social trends during the Covid-19 pandemic that can lead to a sustained reduction in CO_2 emissions
  – Time-series analysis of carbon emissions data and its correlation with stringency of government policies (quantified by Oxford Coronavirus Government Response Tracker, OxCGRT)
  – Data collection, cleaning and building a Python software that implements a comprehensive data analysis pipeline with a web frontend for interactive visualization

Reliable Communication of IoT Devices using D2D Technology *(January 2017 - August 2017)*
Final Year Project, NIT Durgapur
  – Investigation and improvement of D2D link reliability
  – IoT devices are assigned the resources of licensed CUEs through Maximum Bipartite Matching; Reliability of the links thus established is investigated and the weak links are discarded; Disconnected IoT devices are rerouted to the IoT gateway using a two-hop scheme in order to improve access rate and sum throughput of the system
  – Results indicate improvement in access rate and sum throughput of the system, published at COMSNETS 2018

PUBLICATIONS

- Sreetama Sarkar et al. *Block Selective Reprogramming for On-device Training of Vision Transformers*, Efficient Computer Vision (ECV) Workshop, CVPR 2024 (acceptance rate: $< 33\%$).
- Sreetama Sarkar et al. *RLNet: Robust Linearized Networks for Efficient Private Inference*, Fair, Data-Efficient, and Trusted Computer Vision (TCV) Workshop, CVPR 2024.

- Pierpaolo Mori et al. *Accelerating and Pruning CNNs for Semantic Segmentation on FPGA*, DAC 2022.
- Manoj Vemparala et al. *Adversarial Robust Model Compression using In-Train Pruning*, Safe Artificial Intelligence for Automated Driving (SAIAD) Workshop, CVPR 2021.
- Anish Pradhan et al. *Implementation of Relay Hopper Model for Reliable Communication of IoT Devices in LTE Environment through D2D Link*, COMSNETS 2018.
- Sreetama Sarkar et al. *Performance of Different Power Control Schemes for a Hybrid LTE System with Channel Impairment*, iEECON2017.
- Swagata Mandal et al. *Criticality Aware Soft Error Mitigation in the Configuration Memory of SRAM based FPGA*, VLSID 2019.

SCHOLASTIC ACHIEVEMENTS

- **DAC Young Fellow**, 2024
- **Annenberg Endowed Fellowship**, University of Southern California, 2024 (awarded to 2 students in the department)
- **MHI Fellowship**, University of Southern California, 2022 (awarded to less than 10% of incoming PhD students in the department)
- **Grace Hopper Celebration (GHC) Student Scholarship**, 2021
- **Deutschlandstipendium**, The German National Scholarship at TUM, 2019
- **Institute Gold Medal**, NIT Durgapur, 2017 (highest marks in B. Tech in Electronics and Communication Engineering)
- **Parpatidevi Chandumal Memorial Gold Medal**, NIT Durgapur, 2017 (Scholastic excellence among girl students)
- **Scholarship for College and University Students, Government of India**, 2013

TECHNICAL SKILLS

- **Programming Languages:** Python (Packages: Pytorch, Tensorflow), MATLAB, C, C++, VHDL, Verilog
- **Tools:** LaTeX, Git, Xilinx ISE, Vivado, Multisim
- **OS:** Linux, Windows

EXTRA ACADEMIC ACTIVITIES

- Viterbi Graduate Student Mentor, Fall 2023 (best mentor nomination)
- Mentored Master's students at USC for Directed Research courses
- Reviewed papers for conferences like DATE, DAC, IJCNN, ICASSP, KDD, VLSID and journals like TCAD, TCAS
- Poster Mentor and Social Host, WiML Workshop co-located with NeurIPS 2020
- Poster presenter and volunteer, WiML Un-Workshop co-located with ICML 2021
- Interested in *Painting, Indian Classical Dance and Yoga*