Souvik Kundu

https://www.linkedin.com/in/souvik-kundu
Personal-web-home

Email : souvikku@usc.edu
Mobile : +1-213-431-9768

EDUCATION

- **University of Southern California** — Los Angeles, CA
  Doctor of Philosophy (Ph.D.) in Electrical Engineering; Major GPA: 4.00 /4.00 — *August 2017 – present*
  Major courses taken: VLSI Design, Probability for Electrical Engg, Algorithms Design, Computer Architecture, Deep Learning, Advanced Computer Architecture Design, Mathematical Model for Cyber-physical System Design, Machine Learning, Database, Affective Computing.

- **Indian Institute of Technology** — Kharagpur, India
  Master of Technology in Microelectronics and VLSI design; GPA: 9.44 /10.0 (DR 1) — *August 2013 – May 2015*

- **West Bengal University of Technology** — Kolkata, India
  Bachelor of Technology in Electronics and Communication Engineering; GPA: 9.02 /10.0 (DR 2) — *August 2005 – July 2009*

EXPERIENCE

- **Intel Labs** — Santa Clara, USA
  Deep Learning Research Intern — *May 2020 - Dec 2020*
  - **Self-attention for vision**:
    1. As a part of the Intel AI Labs team I am closely involved in various distillation and computer-vision related projects and having the opportunity to work on SOTA CV models and enhance my skills on Pytorch, efficient model training.
    2. Worked towards development of various computer vision models for joint image upsampling tasks.

- **University of Southern California** — Los Angeles, CA
  Ph.D. Graduate Student (Advisors: Prof. Peter A. Beerel and Prof. Massoud Pedram.) — *August 2017 - present*
  - **Machine Learning**: **Pre-defined sparsity based network model search**
    1. Currently involved in real time machine learning network model design for energy and storage constrained application to make the model deploy able fully or partially in embedded systems or edge devices for real-time training and inference. We have also **developed** a more **efficient** form of **compressed sparse representation** scheme to represent our notion of sparsity. This representation leverages the advantage of lower transfer of data from high cost DRAM to processing elements (PEs) where the multiply-accumulation operation takes place. Thus it ensures to minimize the cost of data transfer in our proposed sparse representation for domain specific ASIC based application.
  - **Machine Learning: Training framework in beyond CMOS technology**:
    1. Pre-defined sparse neuromorphic ex-situ training framework for Memristive accelerators for MLPs.
  - **Machine Learning: Novel network pruning driven by optimization**:
    1. Currently working on an energy efficient model pruning framework design driven by pruning and quantization. Future scope involves, architecture searches for efficient training and inference.
  - **Machine Learning: Adversarially robust NN systems**:
    1. Developed a novel training scheme to yield robust yet compressed DNN models via a single training iteration.
  - **Algorithm-hardware Co-design for Energy-efficient Event-driven Inference**:
    1. Developed novel neuro-inspired learning algorithms with reduced latency which can yield orders of magnitude improvement in energy-efficiency for image classification tasks compared to traditional deep neural networks.
    2. Working on a novel architecture that can process a set of input spikes using an output-stationary dataflow model, along with a hybrid on-chip memory to accelerate the processing of Spiking Neural Networks (SNNs).
  - **Beyond CMOS: Architecture Design for Single Flux Quantum based ALU**:
    1.We have proposed and successfully demonstrated through simulations a novel block-skewed arithmetic logic unit for single flux quantum technology (SFQ) based processors, to alleviate the issue of penalty in instruction throughput for data-dependent operation raised due to gate-level pipelined nature of such technology gates.

- **Texas Instruments** — Bangalore, India
  Digital Design Engineer — *June 2016 - July 2017*
  - **System Verilog, UVM and MATLAB**: As a part of high speed converter group I was responsible for designing an automatic gain controller (AGC) block in Verilog. This block was responsible for controlling the gain of the low noise amplifier (LNA) which captures the input analog signal, to avoid signal amplitude saturation for the filter blocks which is after LNA in the datapath.

- **Synopsys** — Bangalore, India
  R and D Engineer II — *June 2015 - May 2016*
  - **VCS and Assertion**: Developed skills in VCS and System Verilog assertion.

- **Bharat Sanchar Nigam Limited** — Kolkata, India
  Junior Telecom Officer — *March 2010 - June 2013*
  - **Skills developed**: Managed the billing team of Call Data Record Project, Worked on PL-SQL based DBMS.

SERVICES

- **Summer Mentoring** : Mentored and guided students from IIT Gandhinagar through Viterbi IUSSTF Summer intern program in 2018, 2020 (remote).
- **Reviewing and membership services**:
  Journals: More than 10 journals including IEEE Transactions on Circuits and Systems I and II, Computers, Neural Networks and Learning Systems, CAD, MICRO, MDPI. **Conferences:** More than 15 including ISCAS (2020), DAC (2020), BMVC (2020), EMNLP (2020) [outstanding reviewer], WACV (2021).

SELECTED PUBLICATIONS [Google Scholar: SOUVIKKUNDU]

- **S. Kundu***, H. Mostafa*, S. N. Sridhar*, S. Sundaresan*, "Attention-based Image Upsampling", under review.[paper]
- **S. Kundu**, S. Sundaresan, "AttentionLite: Towards Efficient Self-Attention Models for Vision", under review, ***ICASSP**, 2021.* [paper]
- **S. Kundu**, G. Datta, M. Pedram, P. A. Beerel, "Spike-Thrift: Towards Energy-Efficient Deep Spiking Neural Networks by Limiting Spiking Activity via Attention-Guided Compression", ***WACV**, 2021.* [paper]
- **S. Kundu**, M. Nazemi, P. A. Beerel, M. Pedram, "DNR: A Single-Shot Tunable Robust Pruning Framework Through Dynamic Network Rewiring of DNNs", ***ASP-DAC**, 2021.* [paper]
- **S. Kundu**, M. Nazemi, M. Pedram, K. M. Chugg, P. A. Beerel, "Pre-defined Sparsity for Low-Complexity Convolutional Neural Networks", ***IEEE Transactions on Computers**, July 2020*, [paper]
- **S. Kundu**, S. Prakash, H. Akrami, P. A. Beerel, K. M. Chugg, "pSConv: A Pre-defined Sparse Kernel Based Convolution for Deep CNNs," ***Allerton Conference**, 2019.* [paper]
- **S. Kundu**, G.Datta, P.A. Beerel, M. Pedram, "qBSA: Logic Design of a 32-bitBlock-Skewed RSFQ Arithmetic Logic Unit," ***International Superconducting Electronics Conference**, 2019.* [paper]
- G.Datta, H. Cong, **S. Kundu**, P.A. Beerel, "qCDC: Metastability-Resilient Synchronization FIFO for SFQ Logic," ***International Superconducting Electronics Conference**, 2019.* [paper]
- **S. Kundu***, A. Fayyazi*, S. Nazarian, P.A. Beerel, M. Pedram, "CSrram: Area-Efficient Low-Power Ex-Situ Training Framework for Memristive Neuromorphic Circuits Based on Clustered Sparsity," ***ISVLSI**, 2019.* [paper]

  (*= equal contribution)

PROJECTS

- **Deep Learning Mini Prject**: A bag of feature based convolution with feature of interpretability (won **best Research project** in EE599 course consisting of nearly 100 students). (April 2019 - May 2019)
- **VLSI Design Mini Project**: Design of a General Purpose 5-stage Pipelined Microprocessor with Software and Hardware Components in Cadence. (Nov 2017 - Dec 2017)
- **DFT Mini Project**: Implementation of ATPG and Fault Simulator for combinatorial circuits. (Nov 2017 - Dec 2017)

ACADEMIC ACHIEVEMENTS

- **USC MHI Scolar Finalist, 2020**: Was one of the 11 finalists out of all the Ph.D. students in the department of Electrical and Computer Engineering.
- **Best poster award**: Won the best poster out of 136 posters in the USC MHI research festival event, held on 11-08-2019. The poster topic was: Toward low complexity CNN models for both training and inference.
- **USC Annenberg Fellowship**: Top few incoming Ph.D. students at USC.
- **Academic Excellence**: Dept. rank of 1 and 2 during M.Tech from IIT, Kharagpur and B.Tech from WBUT respectively.
- **Govt. of India Fellowship during M.Tech**: Was within top 1.2% of students in Engg. Services Exam 2011, GATE 2012, 2013 .

SKILLS & ACTIVITIES

- **Programming Languages**: Python (Experienced with API, viz. pyTorch, Keras on Tensorflow), C, C++; worked on online cloud platform, viz. AWS;
  HDL related: Verilog, System-verilog, UVM, DFT, Digital Design.
  Language: English, Hindi, Bengali.

Last updated: January 21, 2021.