Gobinda Saha

2450 Sycamore Ln Apt. 36A, West Lafayette, Indiana, 47906
765-409-1042 gobinda.eee.buet@gmail.com Linkedin GitHub Website

Objective

I am actively seeking a full-time engineering or research role with a primary focus on AI/ML algorithm design. With 6 years of dedicated research experience, I possess a robust foundation in deep learning algorithm development and proficiency in coding with Python and PyTorch. I am deeply passionate about advancing the field of AI, particularly in the areas of computer vision, natural language processing, and generative AI applications, with a strong emphasis on scalability and efficiency.

Education

Purdue University, West Lafayette, Indiana, USA
PhD in Electrical and Computer Engineering (Deep Learning) **Aug. 2017 – Dec. 2023 (expected)**

Bangladesh University of Engineering and Technology (BUET), Bangladesh
Master of Science (M.Sc.) in Electrical and Electronic Engineering **May 2013 – Aug. 2015**

Bangladesh University of Engineering and Technology (BUET), Bangladesh
Bachelor of Science (B.Sc.) in Electrical and Electronic Engineering **Jan. 2008 – Feb. 2013**

Domain Skills

Deep Learning (DL) Algorithms: Continual & Online Learning, Meta/Few-shot Learning, Transfer & Multi-task Learning, Decentralized & Federated Learning, Supervised/Self-supervised Learning, Out-of-Distribution Detection, Adversarial Training
DL Optimization: Orthogonal, Scaled & Natural Gradient Descent, Second-order & Bi-level Optimization
DL Techniques: Model Compression/Pruning, Distillation, Quantization, Regularization, Hardware-aware Training
Text Analytics: Natural Language Processing (Language Models, Transformers), Vision-Language (Multimodal) Models
Visualization Tools: Saliency Maps (Grad-CAM, Smooth-Grad), Loss-landscapes and Input/Embedding Space Visualization
Applications: Computer Vision (Object Localization, Detection, Image Synthesis), Reinforcement Learning (Model-free RL)
Data Analytics: Principal Component Analysis (PCA), Singular Value Decomposition (SVD)
Programming Languages: Python (Expert), C (Intermediate), C++ (Basic), Julia (Basic)
DL Frameworks/Technologies: PyTorch (Expert), TensorFlow (Basic), AWS (Amazon Elastic Compute Cloud)

Research Experience

Nanoelectronics Research Laboratory & Center for Brain-Inspired Computing, Purdue University
Graduate Research Assistant | Advisor : Professor Kaushik Roy **Aug. 2017 – Present**
Skills: Python, PyTorch, torchvision, torchmeta, opencv, numpy, panda, scikit-learn, matplotlib, gym, CUDA, Anaconda

- **Overcoming Catastrophic Forgetting during Continual Learning (CL) in Neural Networks**
  - Invented Gradient Projection Memory (GPM)[4] to store important gradient spaces of each task during CL
  - Obtained near zero forgetting on continual object classification tasks through orthogonal gradient descent to GPM
  - Proposed Scaled Gradient Projection (SGP)[1] algorithm relaxing strict orthogonal gradient constraints in GPM
  - Attained up to 2% higher accuracy in continual image classification and 12% more reward in reinforcement learning (Atari games) tasks than SOTA with minimal forgetting in SGP
  - Introduced SPACE[5] algorithm to optimally allocate resources for each CL task by network pruning and growth
  - Achieved zero forgetting with SPACE and up to 5x energy efficiency during inference due to emerging sparsity
- **Explainable Learning and Meta-Learning for Fast Online Continual Learning (OCL)**
  - Invented an experience replay method (EPR)[2] with episodic memory utilizing explainable AI (XAI) tools - saliency maps for memory selection
  - Attained up to 5% accuracy improvement over SOTA on OCL object classification benchmarks with tiny memories
  - Proposed a meta-optimization algorithm for episodic memory-free online fast continual learning
  - Obtained up to 4% higher accuracy on computer vision tasks with meta-learned scaled gradient update
- **Communication Efficient and Data Private Decentralized Continual Learning (DCL)**
  - Collaborated in a team of 4 to design CoDeC[9] - a DCL algorithm - combining orthogonal gradient projections with gossip-averaging among distributed agents ensuring inter-agent data privacy
  - Achieved SOTA accuracy on 3 DCL image classification benchmarks with up to 4.8x reduced communication cost

Internship Experience

GlobalFoundries, Santa Clara, CA, USA

Research Intern, Memory Solution Team | Mentor : M. Ahosan Ul Karim **Jun. 2019 – Aug. 2019**

Skills: Python, PyTorch, CUDA, HSPICE, Cadence, AWS (Amazon Elastic Compute Cloud)

- **Software Framework for Hardware-Algorithm Co-design for Deep Learning Applications**
  - Proposed a 10T bit-cell based IMC primitive for accelerating binary neural network inference as part of a team of 6
  - Developed a Python-HSPICE based software framework for hardware-algorithm co-design and measured performance of in-memory computing (IMC) arrays on DL workloads under nonidealities and process variations [6]

Teaching Experience

Bangladesh University of Engineering and Technology (BUET), Bangladesh

Faculty Member, Dept. of Electrical and Electronic Engineering **Aug. 2014 – Aug. 2017**

- Taught electrical circuits and digital electronics theory courses to undergraduate classes of 65 students each
- Supervised over 200 students in VLSI laboratory courses and projects with step-by-step instructions, guiding them to learn good circuit/hardware design practices with Cadence and HSPICE tools
- Complied coding assignments and guided 100 students to solve problems with numerical techniques in MATLAB

Projects

Interior Point Linear Program (IPLP) Solver | Skill: *Julia* **May 2020**

- Implemented IPLP solver in Julia based on Mehrotra's Predictor-Corrector algorithm with modified Cholesky factorization

Hardware Accelerator Design for Neural Network Inference on FPGA | Skills: *Verilog, Quartus* **Dec. 2018**

- Built a hardware accelerator for MLP on Altera DE2-115 FPGA board to reduce inference time for MNIST classification

Photonic IMC Primitives Design for Spiking Neural Networks (SNNs) | Skills: *COMSOL, MATLAB* **Jan. 2018**

- Designed spiking neuron and synaptic arrays based on GST (phase change material) embedded optical micro-ring resonators to realize energy-efficient and fast photonic computing primitives [7,8]
- Developed a device-circuit-algorithm co-design framework to evaluate their performance as SNN inference engine

Relevant Coursework

- Machine Learning (Python, TensorFlow)
- Computational Methods in Optimization (Julia, Python)
- Natural Language Processing (Online Courses)
- Data Structures and Algorithms (Python)
- MOS VLSI Design (Cadence, HSPICE)
- Advanced VLSI Design
- System On Chip Design (Verilog, Quartus)
- Solid State Devices

Selected Publications

1. G. Saha, K.Roy, "Continual Learning with Scaled Gradient Projection", AAAI, vol. 37, no. 8, pp. 9677-9685, Jun. 2023.
2. G. Saha, K.Roy, "Saliency Guided Experience Packing for Replay in Continual Learning", Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 5273-5283, 2023. (**Award Finalist - Top 1% paper**)
3. G. Saha, C Wang, A Raghunathan, K Roy, "A cross-layer approach to cognitive computing", Proceedings of the 59th ACM/IEEE Design Automation Conference (DAC), 1327-1330, 2022.
4. G. Saha, I.Garg, K.Roy, "Gradient Projection Memory for Continual Learning", International Conference on Learning Representations (ICLR), 2021. (**Oral Presentation - Top 1% Papers**)
5. G. Saha, I.Garg, A.Ankit, K.Roy, "SPACE: Structured Compression and Sharing of Representation Space for Continual Learning", IEEE Access, 9, 150480-15094, 2021.
6. G. Saha, Z.Jiang, S.Parihar, C.Xi, J.Higman and M.Ahosan Ul Karim, "An Energy-Efficient and High Throughput in-Memory Computing Bit-Cell With Excellent Robustness Under Process Variations for Binary Neural Network", IEEE Access, 8, 91405-91414, 2020.
7. I. Chakraborty, G. Saha, K.Roy, "Photonic In-Memory Computing Primitive for Spiking Neural Networks Using Phase-Change Materials", Phys. Rev. Applied, 11, 014063, Jan 2019.
8. I. Chakraborty, G. Saha, A. Sengupta, K.Roy, "Toward fast neural computing using all-photonic phase change spiking neurons", Scientific Reports, 8, 12980, Aug 2018.
9. S. Choudhary, S. Aparna Aketi, G. Saha, K.Roy, "CoDeC: Communication-Efficient Decentralized Continual Learning", arXiv:2303.15378, 2023. (Under review)
10. G. Saha, K.Roy, "Online continual learning with saliency-guided experience replay using tiny episodic memory", Machine Vision and Applications 34, 65 (2023).