Abinash Mohanty

Linkedin

Email : amohant4@asu.edu
Mobile : +1-480-265-6556

Summary

PhD. student with 3+ years of professional experience and excellent understanding on deep learning frameworks, hardware design & acceleration, FPGA prototyping & ASIC design, algorithms and data-structures.

Education

- Arizona State University, USA — *CGPA: 3.91/4.00*
  Doctor of Philosophy in Electrical Engineering; — *Aug. 2013 – May. 2018 (expected)*
- National Institute of Technology, India — *CGPA: 8.18/10.0*
  Bachelor of Technology in Electronics & Comm. Engineering; — *May. 2006 – May. 2010*

Professional Experience

- Arizona State University, USA — *Advisor: Prof. Yu Cao*
  Graduate Research Assistant — *Dec 2013 – Present (3.5+ years)*
  - Design & optimization of deep learning algorithms for power efficient and high performance hardware acceleration.
  - Custom IC design and FPGA prototyping for training and inference of deep learning algorithms.
  - Compact modeling of device and circuit Level reliability issues.
- Intel Corporation, Santa Clara, CA — *Manager: Mrs. Tanuja Rao*
  Graduate Intern — *May 2016 – Oct 2016 (5 months)*
  - System level modeling and performance benchmarking of hardware accelerator for speech processing.
  - Development and optimization of algorithms and hardware architecture for efficient online training in embedded systems.
- Samsung Research Institute, Noida, India — *Manager: Mr. Amit Jaiswal*
  Senior Software Engineer II — *June 2010 – July 2013 (3 years)*
  - Developed embedded application platform with stereoscopic 3D graphics rendering capabilities for HTML based widgets.
  - Designed and implemented NPAPI Plugins & JavaScript Engine for running widgets on smart televisions.

Technical Skills

- Languages: Python, Verilog, C++, C, MATLAB, JavaScript, OpenCL, LaTeX
- Deep Learning Frameworks: TensorFlow, Caffe

Research Projects & Publications

- Random Sparse Adaptation of Deep Learning Algorithms for Hardware Acceleration using Analog Devices.
  - Developed a novel bifurcated deep network to finetune pre-trained parameters and compensate for variations in devices.
  - Achieved $30\times \sim 100\times$ speed improvements with deep networks on CIFAR10 dataset compared to existing methodologies.
  - **Mohanty, A.**, Du, X., Chen, P., Seo, J.S., Yu, S., Cao, Y.,"Random Sparse Adaptation for Accurate Inference with Inaccurate Multi-level RRAM Arrays," in *IEEE International Electron Devices Meeting (IEDM),* 2017, *Accepted.*
- Machine Learning for Bio-medical Sensing and HealthCare Applications.
  - Designed a deep learning algorithm for cough detection from audio data recorded using VitaloJAK wearable microphone.
  - Demonstrated an average leave-one-out cross-validation specificity and sensitivity of 93.7% and 97.6% respectively.
  - Kadambi, P., **Mohanty, A.**, Ren, H., Smith, J., McGuinnes, K., Holt, K., Furtwaengler, A., Slepetys, R., Yang, Zheng., Seo, J.S., Vrudhula, S., Chae, J., Cao, Y., Berisha, V.,"Towards a wearable cough detector based on deep neural networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* 2017, *Submitted*
- Hardware-Software Co-optimization for real-time operation of Cascaded Random Forest Algorithms.
  - Designed a cascaded week classifier optimized for efficient hardware acceleration ($30\times \sim 1000\times$ performance improvement).
  - Demonstrated real-time face and traffic signs detection applications with hardware accelerators on ASIC and FPGAs.
  - **Mohanty, A.**, Suda, N., Kim, M., Vrudhula, S., Seo, J. S., and Cao, Y.,"High-performance face detection with CPU-FPGA acceleration," in *IEEE International Symposium Circuits and Systems (ISCAS),* 2016, pp. 117-120.
  - Kim, M., **Mohanty, A.**, Kadetotad, D., Suda, N., Wei, L., Saseendran, P., He, X., Cao, Y. and Seo, J.S.,"A real-time 17-scale object detection accelerator with adaptive 2000-stage classification in 65nm CMOS," in *Design Automation Conference (ASP-DAC), 2017 22nd Asia and South Pacific,* 2017, pp. 21-22.

- Hardware Acceleration of Deep Neural Networks for Machine Learning.
  - Analyzed operation of deep networks & modified the layers to improve efficiency & optimize from hardware point of view.
  - Implemented a low-precision neural network accelerator of VGG and AlexNet on FPGA using OpenCL framework.
  - Suda, N., Chandra, V., Dasika, G., **Mohanty, A.**, Ma, Y., Vrudhula, S., Seo, J.S. and Cao, Y.,"Throughput-optimized OpenCL-based FPGA accelerator for large-scale convolutional neural networks," in *Proceedings of the ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA),* 2016, pp. 16-25.
- True Random Number Generation using Random Telegraph Noise in MOSFETs.
  - Conceptualized a novel method to efficiently generate on-chip true random numbers using RTN and validated using 180nm, 65nm and 28nm technology nodes transistor measurement data.
  - **Mohanty, A.**, Sutaria, K., Awano, H., Sato, T., and Cao, Y.,"RTN in Scaled Transistors for On-chip Random Seed Generation," in *IEEE Transactions on Very Large Scale Integration Systems (TVLSI),* 2017 Apr 12.
- Attention based Object Detection using Saliency algorithm for Search Space Reduction.
  - Designed a saliency algorithm with adaptive grid based normalization to be used as a precursor for detection algorithms.
  - Achieved $\sim 20\%$ reduction in search space directly translating to $\sim 20\%$ reduction of run-time of succeeding algorithms.
- Parallel architecture with resistive cross-point array for dictionary learning.
  - Conceptualized a hardware architecture using cross-point arrays of RRAM devices targeted towards learning algorithms.
  - Designed and implemented peripheral circuits necessary for dictionary update operations needed for online learning.
  - Xu, Z., **Mohanty, A.**, Chen, P.Y., Kadetotad, D., Lin, B., Ye, J., Vrudhula, S., Yu, S., Seo, J.S. and Cao, Y.,"Parallel programming of resistive cross-point array for synaptic plasticity," in *BioTL,* 2014, pp.126-133.
  - Kadetotad, D., Xu, Z., **Mohanty, A.**, Chen, P.Y., Lin, B., Ye, J., Vrudhula, S., Yu, S., Cao, Y. and Seo, J.S.,"Parallel architecture with resistive crosspoint array for dictionary learning acceleration," in *IEEE Journal on Emerging and Selected Topics in Circuits and Systems,* 2015, pp.194-204.
  - Chen, P.Y., Kadetotad, D., Xu, Z., **Mohanty, A.**, Lin, B., Ye, J., Vrudhula, S., Seo, J.S., Cao, Y. and Yu, S.,"Technology-design co-optimization of resistive cross-point array for accelerating learning algorithms on chip," in *Design, Automation & Test in Europe Conference and Exhibition (DATE),* 2015, pp. 854-859.
  - Kadetotad, D., Xu, Z., **Mohanty, A.**, Chen, P.Y., Lin, B., Ye, J., Vrudhula, S., Yu, S., Cao, Y. and Seo, J.S.,"Neurophysics-inspired parallel architecture with resistive crosspoint array for dictionary learning," in *Biomedical Circuits and Systems Conference (BioCAS),* 2014, pp. 536-539.
  - Seo, J.S., Lin, B., Kim, M., Chen, P.Y., Kadetotad, D., Xu, Z., **Mohanty, A.**, Vrudhula, S., Yu, S., Ye, J. and Cao, Y.,"On-chip sparse learning acceleration with CMOS and resistive synaptic devices," in *IEEE Transactions on Nanotechnology (TNANO),* 2015, pp.969-979.
- Compact modeling of device and circuit level reliability issues.
  - Modeled device and circuit level reliability phenomenons (NBTI, PBTI) for prediction of aging issues in CMOS circuits.
  - Sutaria, K.B., **Mohanty, A.**, Wang, R., Huang, R. and Cao, Y.,"Accelerated aging in analog and digital circuits with feedback," in *IEEE Transactions on Device and Materials Reliability (TDMR),* 2015, pp.384-393.
  - Sutaria, K.B., Ren, P., **Mohanty, A.**, Feng, X., Wang, R., Huang, R. and Cao, Y.,"Duty cycle shift under static/dynamic aging in 28nm HK-MG technology," in *IEEE International Reliability Physics Symposium (IRPS),* 2015, pp. CA-7.

Training Experience

- IBM TrueNorth BootCamp — *2nd Aug. 2015 – 23rd Aug, 2015*
  Algorithm design for deep learning on IBM Synapse platform — *IBM Research, Almaden*
- Altera OpenCL Optimization Training — *15th Oct. 2015*
  Hardware-software co-optimization using openCL SDK for Altera FPGA boards — *Phoenix, Arizona*

Relevant Poster Presentations

- "Random Sparse Adaptation for Accurate Inference: A Hybrid Approach with RRAM and On-chip Memory," in *IBM CAS Symposium,* T.J. Watson Research Center, Yorktown, New York, 2017.
- "High-performance face detection with CPU-FPGA acceleration," in *HALO, International Conference on Computer Aided Design (ICCAD),* Austin, Texas, 2016.
- "Duty cycle shift under static/dynamic aging in 28nm HK-MG technology," in *IEEE International Reliability Physics Symposium (IRPS),* Monterey, CA, 2015.