# Guest Editorial: IEEE TC Special Issue On Software, Hardware and Applications for Neuromorphic Computing

Yiran Chen, *Fellow, IEEE* and Qinru Qiu, *Senior Member, IEEE*

Inspired by biological neural systems, neuromorphic computing has drawn much attention for its great potential of achieving machine intelligence at extremely low energy dissipation. Bio-inspired computing models have been investigated for information encoding, sparse representation, event driven communication/computation, and online learning. This new computing paradigm triggered a recent wave of innovations in software and hardware architecture and emerging device technology, which consequently enabled many novel applications. This is an exemplar of research area where the application, computing model, architecture and circuit level design are tightly coupled to deliver unprecedented functionality and energy efficiency. This special issue of IEEE Transactions on Computers will explore academic and industrial research on all topics related to neuromorphic computing, from computing model, software, and hardware architecture to application design.

The research topics that are covered by the special issue include but not limited to: 1) Network, neuron and synapse models of bio-inspired and spiking neural networks (SNNs); 2) Non-von Neumann computing architectures; 3) Emerging devices and hardware implementations; 4) Supervised or unsupervised learning in neuromorphic computing systems; 5) Biologically-inspired network structure and computing models; 6) Applications where neuromorphic system have the potential to outperform state-of-the-art techniques; 7) Online learning and real-time applications of neuromorphic systems; and 8) Application, computing model, and hardware architecture co-design for neuromorphic systems.

The Special Issue of IEEE Transactions on Computers on Software, Hardware and Applications for Neuromorphic Computing attacted 15 submissions from authors in more than five countries on vastly diverse topics. We were supported by more than 40 reviewers who provided very constructive feedbacks and helped the authors to improve their contributiosn. Each submitted manuscript has been reviewed by at least three reviewrs. Conflict-of-Interest is strictly handled during the whole review process. Eventually, 10 manuscripts have been selected to form this special issue. The accepted contributions span over many important and emerging fields in neuromorphic computing such as spiking neural networks, biologically inspired systems, unsupervised learning, emerging memory based neurmophic computing, hyperdimensional computing, memristive systems, spiking generative aderversial networks, time-coded spiking systems, imary visual cortex, and process-in-memory. Below, we provide a summary of the key contributions of the papers included in the special issue.

The paper entitled "SpikeBASE: Spiking Neural Learning Algorithm with Backward Adaptation of Synaptic Efflux" by Stauffer et al. proposes a novel algorithm, namely, SpikeBASE, to globally, supervisedly, and comprehensively coordinate the synaptic dynamics including both synaptic strength and responses. SpikeBASE can learn synaptic strength by backpropagating the error through the predefined synaptic responses and enable synaptic response adaptation through backpropagation, to mimic the complex dynamics of neural transmissions. Through comprehensively coordinating the learning of synaptic strength, synaptic responses, and multi-scale temporal memory formation, SpikeBASE has demonstrated its effectiveness on end-to-end SNN training.

Inspired by how the retina helps human brain incept new information efficiently, the paper entitled "VisualNet: An End-to-End Human Visual System Inspired Framework to Reduce Inference Latency of Deep Neural Networks" by Wang et al. presents an end-to-end structured framework built using any existing convolutional neural network (CNN) as the backbone. The proposed framework, called VisualNet, can create task parallelism for the backbone during the inference of a single image. Experiments on GPUs and CPUs showthat the proposed VisualNet reduces the latency of the regular network it builds on by up to 80.6% when both are fully parallelized with state-of-the-art acceleration libraries. At the same time, VisualNet can achieve similar or slightly higher accuracy.

An event-based vision sensor allows for sparse and asynchronous events to be produced that are dynamically related to the scene. However, many traditional machine learning (ML) and computer visioin (CV) approaches are not applicable to event-based spatial-temporal visual streams. The paper entitled "Unsupervised Spiking Instance Segmentation on Event Data using STDP Features" by Kirkland et al. presents a novel approach that can perform instance segmentation using just the weights of a Spike Time Dependent Plasticity (STDP) trained Spiking CNN

Digital Object Identifier no. 10.1109/TC.2022.3208389

that was trained for object recognition. This exploits the spatial and temporal aspects of the network's internal feature representations adding this new discriminative capability.

"Exploring Model Stability of Deep Neural Networks for Reliable RRAM-based In-Memory Acceleration" by Krishnan et al. investigates a new metric, model stability, from the loss landscape to help shed light on accuracy loss under variations and model compression, which guides an algorithmic solution to maximize model stability and mitigate accuracy loss. The authors characterize wafer-level RRAM variations based on a CMOS/RRAM 1T1R test chip at 65nm and develop a cross-layer benchmark tool that incorporates quantization, pruning, device variations, model stability, and IMC architecture parameters to assess post-mapping accuracy and hardware performance.

Hyperdimensional computing (HDC) has emerged as an alternative lightweight learning solution to deep neural networks. In "OpenHD: A GPU-Powered Framework for Hyperdimensional Computing," Kang et al. present OpenHD, a flexible and high-performance GPU-powered framework for automating the mapping of general HDC applications including classification and clustering to GPUs. OpenHD takes advantage of memory optimization strategies specialized for HDC, minimizing the access time to different memory subsystems, and removing redundant operations. The authors also propose a novel training method to enable data parallelism in the HDC training.

Time Delay Reservoir (TDR) is a hardware-friendly machine learning approach from two perspectives. However, it performs poorly on tasks that involve long-term dependency. In "Adaptive Memory-enhanced Time Delay Reservoir and Its Memristive Implementation" by Shi et al., the authors first introduce a higher-order delay unit, which is capable of accumulating and transferring the long history states in an adaptive manner to further enhance the reservoir memory. Then they propose a memristive implementation of our adaptive memory-enhanced TDR that achive high effectiveness, smaller circuit area, and lower power consumption simultaneously.

In order to fully leverage the time-encoding capacity of spikes during the training of SNNs, "Spiking Generative Adversarial Networks With a Neural Network Discriminator: Local Training, Bayesian Models, and Continual Meta-Learning" by Rosenfeld et al. proposes to train SNNs so as to match distributions of spiking signals rather than individual spiking signals. In particular, the work introduces a novel hybrid architecture comprising a conditional generator, implemented via an SNN, and a discriminator, implemented by a conventional artificial neural network (ANN).

"Time-Coded Spiking Fourier Transform in Neuromorphic Hardware" by Ĺopez-Randulfe et al., proposes a time-based SNN that is mathematically equivalent to the Fourier transform. The authors implemented the network in the neuromorphic chip Loihi and conducted experiments on five different real scenarios with an automotive frequency modulated continuous wave radar. Experimental results validate the algorithm, and it is expected these results can prompt the design of ad hoc neuromorphic chips that can improve the efficiency of state-of-the-art digital signal processors and encourage research on neuromorphic computing for signal processing.

Inspired by the working mechanism of the primary visual cortex, pulse-coupled neural networks (PCNNs) can exhibit the characteristics of synchronous oscillation, refractory period, and exponential decay. However, electrophysiological evidence shows that the neurons exhibit highly complex nonlinear dynamics when stimulated by external periodic signals. This chaos phenomenon, also known as the "butterfly effect", cannot be explained by all PCNN models. "The Butterfly Effect in Primary Visual Cortex" by Liu et al., analyzes the main obstacle preventing PCNN models from imitating a real primary visual cortex and consider neuronal excitation as a stochastic process. The authors then propose a novel neural network of the primary visual cortex, called a continuous-coupled neural network (CCNN). Theoretical analysis indicates that the dynamic behavior of the CCNN is distinct from the PCNN. Numerical results show that the CCNN model exhibits periodic behavior under a DC stimulus, and exhibits chaotic behavior under an AC stimulus, which is consistent with the testing results of primary visual cortex neurons.

Emerging device-based digital processing-in-memory (PIM) architectures have emerged as a promising neuromorphic computing solution due to their energy and area efficiency derived from analog to digital converter (ADC)-less PIM hardware. However, digital PIM architectures generally need large extra memories to copy parameters, and they also suffer from low computation per memory-cycle efficiencies. In "SOT-MRAM Digital PIM Architecture with Extended Parallelism in Matrix Multiplication" by Kim et al., The authors present a novel spin-orbit torque magnetic random access memory (SOT-MRAM) based digital PIM architecture to alleviate the extra memory size burden and computation cycle issues. A set of architecture and circuit-level solutions are designed and validated using 28nm CMOS process.

We would like to thank all authors who submitted their work to this special issue. We also would like to thank all our reviewers for their tremendous efforts to evaluate and judge the submissions. Further on, we would like to express our sincere gratidude to the Editor-in-Chief Ahmed Louri, Associate Editor James C. Hoe, Chao Li, and Kristin Falco FaFeleur from IEEE for continuous help and support during the organization and preparation of the special issue.

Yiran Chen
Qinru Qiu
*Guest Editors*

*Yiran Chen is with Duke University, Durham, NC, 27708. E-mail: yiran.chen@duke.edu.*
*Qinru Qiu is with Syracuse University, Syracuse, NY, 13244. E-mail: qiqiu@syr.edu.*
*F.A. is with the National Institute of Standards and Technology, Boulder, CO 80305. E-mail: author@boulder.nist.gov.*
*S.B. Jr. is with the Department of Physics, Colorado State University, Fort Collins, CO 80523. E-mail: author@colostate.edu.*
*T.C. is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309. On leave from the National Research Institute for Metals, Tsukuba, Japan E-mail: author@nrim.go.jp.*