

2025 IEEE 7th International Conference on Artificial Intelligence Circuits and Systems (AICAS 2025)

**Bordeaux, France
28-30 April 2025**



**IEEE Catalog Number: CFP25R18-POD
ISBN: 979-8-3315-2425-8**

**Copyright © 2025 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved**

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP25R18-POD
ISBN (Print-On-Demand):	979-8-3315-2425-8
ISBN (Online):	979-8-3315-2424-1
ISSN:	2834-9830

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

A1L-A

Computer Vision Algorithms & Architectures

Date: Monday Apr 28th, 2025

Time: 09:00-10:20

Room: Lecture Hall G

Chair: Amélie Gruel

Optimizing the Semantic Segmentation CNN SalsaNext on a Vertical Vector Processor

Oliver Renke, Holger Blume

Institute of Microelectronic Systems, Leibniz University Hannover, Germany

Page: 30 : [Click here to get to the paper](#)

AiRDUNet: All-in-One Residual Dense U-Net with Automated Humidity Quality Assessment for Partial Wet Fingerprint Restoration and Recognition

Jui Min Hsu², Ching-Te Chiu², Mao Hsiu Hsu¹, Che Wei Liao²

¹National Formosa University, Taiwan; ²National Tsing Hua University, Taiwan

Page: 35 : [Click here to get to the paper](#)

Efficient Coarse-to-Fine Diffusion Models with Time Step Sequence Redistribution

Yu-Shan Tai, An-Yeu Wu

National Taiwan University, Taiwan

Page: 40 : [Click here to get to the paper](#)

A Real-Time 3DGS-Slam with Voxel-Based Memory Management and Region of Interest-Based Pixel-Wise Skipping Architecture on Edge Device

Hyungnam Joo, Junha Ryu, Seryeong Kim, Jongjun Park, Hoi-Jun Yoo

Korea Advanced Institute of Science and Technology, Korea

Page: 45 : [Click here to get to the paper](#)

A1L-B

In-Memory Computing

Date: Monday Apr 28th, 2025

Time: 09:00-10:20

Room: Lecture Hall H

Chair: Pierre Lewden

Optimizing and Exploring System Performance in Compact Processing-in-Memory-Based Chips

Peilin Chen, Xiaoxuan Yang

University of Virginia, United States

Page: 50 : [Click here to get to the paper](#)

A 107 TOPS/W Instruction Controlled Scalable 8b Precision Analog In-Memory Computing Core with Flexible Kernel Routing for Embedded Applications

Bijoy Kundu^{1,2}, Roland Müller^{1,2}, Yogesh Ramesh Patil¹, Ralf Brederlow^{1,2}

¹*Fraunhofer Institute for Integrated Circuits IIS, Germany*; ²*Technical University of Munich, Germany*

Page: 55 : [Click here to get to the paper](#)

A CIM Crossbar Array Data Mapping Methodology for Unstructured Sparse CNN

Yan-Lin Hung, Ping-Han Liu, Zun-Sheng Wu, Bo-Cheng Lai, Shyh-Jye Jou

National Yang Ming Chiao Tung University, Taiwan

Page: 60 : [Click here to get to the paper](#)

AnalogAI: A Generalized Accuracy Recovery Framework for DNNs Under Analog Computing

Yinghao Li¹, Jiaqi Lv¹, Xuelling Wang¹, Xu Dai¹, Xufeng He¹, Lei Wang¹, Yiran Zhang², Xin Si², Yang Hu³, Shouyi Yin³

¹*Shanghai Artificial Intelligence Laboratory, China*; ²*Southeast University, China*; ³*Tsinghua University, China*

Page: 65 : [Click here to get to the paper](#)

A2L-A

Design Automation for AI Systems / AI for Circuits & Systems

Date: Monday Apr 28th, 2025

Time: 13:30-15:10

Room: Lecture Hall G

Chair: Patricia Desgreys

Optimal Mapping of Monte Carlo Analysis for Analog Circuit Designing Using Bayesian Neural Networks

Yuto Moriguchi, Nobukazu Takai

Kyoto Institute of Technology, Japan

Page: 70 : [Click here to get to the paper](#)

Piecewise NARX Behavioral Model for RF Power Amplifiers in 5G Applications

Trong Thuy Pham², Dang-Kiên Germain Pham², Reda Mohellebi², Pierre Almairac¹, Carolina Pedrosa¹, Patricia Desgreys²

¹*NXP Toulouse, France*; ²*LTCI, Télécom Paris, Institut Polytechnique de Paris, France*

Page: 75 : [Click here to get to the paper](#)

GEMMify: A Unified Model for Layer-Wise DNN Performance Prediction on Edge GPUs

Mewe-Hezoudah Kahanam^{1,2}, Karol Desnos², Guillaume Fortier¹, Maxime Pelcat²

¹*Inetum France, France*; ²*Université de Rennes, INSA Rennes, IETR, CNRS UMR6164, France*

Page: 80 : [Click here to get to the paper](#)

Hardware-Aware Neural Architecture Search for Memory Constrained Embedded Neural Networks Accelerators

Andrea Castagnetti², Alain Pegatoquet², Benoît Miramond², Olivier Montfort¹, Vincent Huard¹

¹*Dolphin Design, France*; ²*Université Côte d'Azur, LEAT, Sophia Antipolis, France*

Page: 85 : [Click here to get to the paper](#)

Accelerating the Self-Consistent Quantum Transport Simulation Model with Advanced Trained Neural Networks

Mani Shankar Yadav, Akhilesh Rawat, Brajesh Rawat

Indian Institute of Technology Ropar, Rupnagar, India

Page: 90 : [Click here to get to the paper](#)

A3L-A

Neuromorphic & Bioinspired Circuits, Processors, Systems & Their Applications I

Date: Monday Apr 28th, 2025

Time: 15:30-17:10

Room: Lecture Hall G

Chair: Antoine Frappé

To Spike or Not to Spike, That Is the Question

Sanaz Mahmoodi Takaghaj, John Sampson
Pennsylvania State University, United States
Page: 95 : [Click here to get to the paper](#)

Neuromorphic Robotics on Conventional Low-Power CPU Hardware

Nicola Russo, Thomas Madsen, Konstantin Nikolic
University of West London, United Kingdom
Page: 100 : [Click here to get to the paper](#)

Explore Activation Sparsity in Recurrent LLMs for Energy-Efficient Neuromorphic Computing

Ivan Knunyants¹, Maryam Tavakol², Manolis Sifalakis¹, Yingfu Xu¹, Amirreza Yousefzadeh^{1,3},
Guangzhi Tang⁴
¹*imec Netherlands, Netherlands*; ²*Eindhoven University of Technology, Netherlands*; ³*University of Twente, Netherlands*; ⁴*Maastricht University, Netherlands*
Page: 105 : [Click here to get to the paper](#)

Enhancing Ant-Inspired Visual Compass with Focused Visual Scan in a Compact Robot

Gabriel Gattaux¹, Antoine Wystrach², Franck Ruffier¹, Julien Serres^{1,3}
¹*Aix Marseille Univ, CNRS, ISM, Marseille, France*; ²*Univ Toulouse, CRCA, CBI, UMR CNRS-UPS 5169, Toulouse, France*; ³*Institut Universitaire de France, IUF, France*
Page: 110 : [Click here to get to the paper](#)

Optimizing Mixed-Precision DNN Scheduling on Heterogeneous SoCs for Enhanced Robustness and Efficiency

Yulong Song, Qiang Tao, Jinlun Ji, Congyi Sun, Yuxiang Fu, Li Li
Nanjing University, China
Page: 115 : [Click here to get to the paper](#)

A4P-C

Poster Session

Date: Monday Apr 28th, 2025
Time: 17:10-19:00
Room: Engineers area
Chair: Pierre Lewden, Sylvain Saïghi

UPE: A Device-Edge DNN Inference Artificial Intelligence Processor With Supporting Reconfigurable Training

Zhou Wang⁷, Haochen Du⁵, Yanqing Xu², Zhou Shu¹⁰, Jiuren Zhou⁴, Liyuan Guo³, Baoyi Han⁹, Xiaonan Tang¹, Shushan Qiao⁸, Shouyi Yin¹¹, Anil A. Bharath⁶, Manos Mic Drakakis⁶
¹Beijing Wisemay Science and Technology Co.,LTD, China; ²Chinese University of Hong Kong, China; ³Dresden University of Technology, Germany; ⁴Hangzhou Institute of Technology, Xidian University, China; ⁵Hong Kong University of Science and Technology, China; ⁶Imperial College London, Imperial Global Singapore, United Kingdom; ⁷Imperial College London, Imperial Global Singapore, Nanyang Technological University, United Kingdom; ⁸Institute of Microelectronics of Chinese Academy of Sciences, University of CAS, China; ⁹Institute of Semiconductors of Chinese Academy of Sciences, University of CAS, China; ¹⁰National University of Singapore, China; ¹¹Tsinghua University, China
Page: 120 : [Click here to get to the paper](#)

UMS-Net: A Multi-Scale Difference Network for Cuffless Blood Pressure Estimation

Ganji Zhang^{1,2}, Hao Wu^{1,2}
¹Shenzhen University, China; ²Shenzhen Ninenovo Technology Limited, China
Page: 125 : [Click here to get to the paper](#)

Epileptic Seizure Prediction: Transfer Learning with Compact AI Models

Hongliu Yang, Ronald Tetzlaff
Dresden University of Technology, Germany
Page: 130 : [Click here to get to the paper](#)

SAPIENS: Towards Energy-Efficient Deep Spiking Neural Networks via the Approximation of Power-Expensive Neurons

Cheol-Min Kang, Nguyen-Dong Ho, Dongyoung Lee, Yonghwan Kwon, Donghyun Kim, Inhan Kang, Ik-Joon Chang
Kyung Hee University, Korea
Page: 135 : [Click here to get to the paper](#)

Comparison of Hardware-Friendly, Audio-to-Spikes Cochlear Encoding for Neuromorphic Processing

Valentin Meunier, Amélie Gruel, Adrien F. Vincent, Sylvain Saïghi
Univ. Bordeaux, CNRS, Bordeaux INP, IMS, UMR 5218, France
Page: 140 : [Click here to get to the paper](#)

A Reconfigurable Event Driven Accelerator for SNN Inference

Yuankang Zhao, Arne Heitmann, Emre Neftci
¹Forschungszentrum Juelich, Germany
Page: 145 : [Click here to get to the paper](#)

Performance Evaluation of Neuromorphic Visual Attention Based on Proto-Objects

Amélie Gruel¹, Adrien F. Vincent¹, Chip Hong Chang^{2,3}, Sylvain Saïghi^{1,3}

¹Univ. Bordeaux, CNRS, Bordeaux INP, IMS, UMR 5218, France; ²Nanyang Technological University, Singapore; ³CNRS@CREATE LTD, Singapore

Page: 151 : [Click here to get to the paper](#)

Hybrid Neural Networks: Integrating Spikes for Energy Efficient Computation

Thomas Louis^{1,2}, Alain Pegatoquet¹, Benoît Miramond¹, Adrien Girard²

¹LEAT / Université Côte d'Azur / CNRS UMR 7248, France; ²IRT Saint Exupéry, France

Page: 156 : [Click here to get to the paper](#)

MemQuant: Towards Memory-Aware Post-Training Quantization for Multimodal Transformers

Kazi Barria Nine, Foroozan Karimzadeh, Zishen Wan, Arijit Raychowdhury

Georgia Institute of Technology, United States

Page: 161 : [Click here to get to the paper](#)

A 26.7 TOPS/W Multiplier-Less Digital In-Memory Computing Macro with Low-Cost Multi-Layer Inference in 28nm FDSOI for Edge AI

Antoine Gautier², Soumya Rank¹, Benoit Larras², Antoine Frappé²

¹Indian Institute of Technology Dharwad Karnataka, India, India; ²University of Lille, CNRS, Centrale Lille, Junia, Université Polytechnique Hauts-de-France, IEMN UMR8520, France

Page: 166 : [Click here to get to the paper](#)

Addition of CXL Memory Pool to Increase LLM Training Check-Pointing Frequency

Kiran Gunnam, Prasad Chandrashekara, Murali Kothandapani, Rajesh Bhagwat, Priyanshi Jain

Micron Technology

Page: 171 : [Click here to get to the paper](#)

Range Encoding: A Low-Power Spiking Neural Network Encoding Scheme for ECG Classification

Xinyu Kang, Xingbo Wang, Yuru Li, Terry Tao Ye

Southern University of Science and Technology, China

Page: 175 : [Click here to get to the paper](#)

Locality-Aware Adaptive Threshold Scaling for Efficient Binary Neural Networks Inference

Tae-Hwan Kim, Su-Jung Lee, Sohye Lee, Jiyoung Lee

Korea Aerospace University, Korea

Page: 179 : [Click here to get to the paper](#)

Computation-Efficient Hardware Accelerator for Unstructured-Pruned Spiking Convolutional Neural Networks

Jiaxiang Li, Youhua Shi

Waseda University, Japan

Page: 184 : [Click here to get to the paper](#)

Design of an Analog Integrated Circuit Solving the Two-Armed Bandit Problem in 180-nm CMOS

Kohei Mori¹, Kazuyuki Wada¹, Kawori Sekine¹, Shinsuke Hara², Satoru Tanoi², Akifumi Kasamatsu²

¹Meiji University, Japan; ²National Institute of Information and Communications Technology, Japan

Page: 189 : [Click here to get to the paper](#)

Design of a Full CMOS Computer Vision Sensor with an Embedded Analog Convolutional Neural Network Processor for Human Face Recognition

Yoochan Yoon, Hyunjin Song, Soo Youn Kim, Minkyu Song

Dongguk University, Korea

Page: 194 : [Click here to get to the paper](#)

PBNNs: Plastic Bayesian Neural Networks for Bio-Inspired Bayesian Learning

Nicolas Ramos, Hai Li

Duke University, United States

Page: 198 : [Click here to get to the paper](#)

Efficient Neuromorphic QUBO Solver with Spintronic Resonators

Guangyu Jiang², Xiaoyue Hu², Weiling Li¹, Yan Fang²

¹*Dongguan University of Technology, China*; ²*Kennesaw State University, United States*

Page: 203 : [Click here to get to the paper](#)

B1L-A

Circuits & Systems for Neural Network-based Language Modeling / Hardware-Software co-design

Date: Tuesday Apr 29th, 2025

Time: 08:40-10:20

Room: Lecture Hall G

Chair: Frank Mizrahi

HEADS: Head-Wise Efficient and Adaptive Sparsification for Transformer-Based Models

Yu-Sung Lee², Jyun-Siou Huang², Yao-Hua Chen¹, Chih-Tsun Huang²

¹*Industrial Technology Research Institute, Taiwan*; ²*National Tsing Hua University, Taiwan*

Page: 208 : [Click here to get to the paper](#)

End-to-End Fully-Binarized Network Design: From Generic Learned Thermometer to Block Pruning

Thien Nguyen, William Guicquero

Smart Integrated Circuits for Imaging Laboratory, CEA-LETI, France

Page: 213 : [Click here to get to the paper](#)

QoS-Nets: Adaptive Approximate Neural Networks

Elias Trommer¹, Bernd Waschneck¹, Akash Kumar²

¹*Infineon Technologies, Germany*; ²*Ruhr-Universität Bochum, Germany*

Page: 218 : [Click here to get to the paper](#)

Enhancing Training Efficiency: A Novel Approach to Handling GPU Failures in Large-Scale Distributed System for LLM Training

Kiran Gunnam, Pranshu Mandal, Shivam Khandelwal, Rajesh Bhagwat

Micron Technology

Page: 223 : [Click here to get to the paper](#)

LogicCraft: LLM-Assisted Optimization of Netlist to Layout for Complex Custom Standard Cell Designs

Minwei Lin¹, Duy-Hieu Bui², Chao Wang¹, Wangzilu Lu¹, Ruoyu Tang¹, Qing Zhang¹, Yuhang Zhang¹, Jian Zhao¹, Xuan-Tu Tran², Yongfu Li¹

¹*Shanghai Jiao Tong University, China*; ²*Vietnam National University - Information Technology Institute, Vietnam*

Page: 228 : [Click here to get to the paper](#)

B2L-A

Emerging Applications I

Date: Tuesday Apr 29th, 2025

Time: 10:40-12:00

Room: Lecture Hall G

Chair: Antoine Gautier

Ring-BP: Using a Wearable Smart Ring to Cufflessly Estimate Blood Pressure with Mobile and Efficient Net

Bin Liu², Hao Wu³, Guoxing Wang², Jiarong Chen², Cheng Chen², Grantham K. H. Pang¹

¹RingConn LLC, United States; ²Shanghai Jiao Tong University, China; ³Shenzhen Ninenovo Technology Limited, China

Page: 233 : [Click here to get to the paper](#)

EnhancePPG: Improving PPG-Based Heart Rate Estimation with Self-Supervision and Augmentation

Luca Benfenati¹, Sofia Belloni¹, Alessio Burrello^{1,5}, Panagiotis Kasnesis^{2,3}, Xiaying Wang⁴, Luca Benini^{4,5}, Massimo Poncino¹, Enrico Macii⁶, Daniele Jahier Pagliari¹

¹DAUIN, Politecnico di Torino, Italy; ²DEEE, University of West Attica, Egaleo, Greece; ³ThinGenious PC, Marousi, Greece; ⁴IIS, ETH Zürich, Switzerland; ⁵DEI, University of Bologna, Bologna, Italy; ⁶DIST, Politecnico di Torino, Italy

Page: 238 : [Click here to get to the paper](#)

A 94.76%-Accuracy, <16 μ s/Classification-Latency STT-MRAM CiM-Based Classifier for Closed-Loop Peripheral Nerve Modulation Implants

Seoyoung Lee¹, Donghyeon Yi¹, Edward Choi¹, Vincent Lukito¹, Jongho Kim¹, Ryan G.L. Koh², Sohmyung Ha³, Ik-Joon Chang⁴, Minkyu Je¹

¹Korea Advanced Institute of Science and Technology, Korea; ²KITE Research Institute, Toronto Rehabilitation Institute- University Health Network, Toronto, Canada; ³New York University Abu Dhabi, U.A.E.; ⁴Kyung Hee University, Korea

Page: 243 : [Click here to get to the paper](#)

Continuous Blood Pressure Measurement Guided by Physiological Mechanisms

Jinbo He², Hao Wu¹

¹RingConn LLC / Shenzhen University / Shenzhen Ninenovo Technology Limited, China; ²Shenzhen University, China

Page: 248 : [Click here to get to the paper](#)

B2L-B

Advances in Hardware and Software Co-Optimization for Intelligent Systems

Date: Tuesday Apr 29th, 2025

Time: 10:40-12:00

Room: Lecture Hall H

Chair: Mao Wei

REN: A Reconfigurable End-to-End NeuroSLAM Hardware Accelerator for Micro Mobile Robots

Zixuan Shen¹, Jian Xiao¹, Heng Ping¹, Yuhao Mei¹, Bowen Shi¹, Zhengzhe Wei², Boyi Dong², Xinglong Ji³, Yuanjin Zheng², Chao Wang¹

¹Huazhong University of Science and Technology, China; ²Nanyang Technological University, Singapore;

³Tsinghua University, China

Page: 253 : [Click here to get to the paper](#)

ANP-L: A 22nm 0.72pJ/SOP Asynchronous Snn-Based Edge-AI Processor Enabling One/Few-Shot On-Chip Incremental Learning

Dexuan Huo, Jilin Zhang, Hong Chen

Tsinghua University, China

Page: 258 : [Click here to get to the paper](#)

RSPE: A High Energy Efficient SNN Inference Processor with RISC-V Based Dynamic Pruning Mechanism

Zhou Wang⁶, Haochen Du⁴, Zhou Shu⁹, Vivek Mohan⁵, Liyuan Guo², Wei Mao³, Baoyi Han⁸, Xiaonan Tang¹, Shushan Qiao⁷, Shouyi Yin¹⁰, Anil A. Bharath⁵, Manos Mic Drakakis⁵

¹Beijing Wisemay Science and Technology Co.,LTD, China; ²Dresden University of Technology, Germany;

³Hangzhou Institute of Technology, Xidian University, China; ⁴Hong Kong University of Science and Technology, China; ⁵Imperial College London, Imperial Global Singapore, United Kingdom; ⁶Imperial College London, Imperial Global Singapore, Nanyang Technological University, United Kingdom;

⁷Institute of Microelectronics of Chinese Academy of Sciences, University of CAS, China; ⁸Institute of Semiconductors of Chinese Academy of Sciences, University of CAS, China; ⁹National University of Singapore, China; ¹⁰Tsinghua University, China

Page: 263 : [Click here to get to the paper](#)

Hardware-Aware Optimization of Large Language Models for Enhanced Throughput on Arm CPUs

Xingyu Zhu¹, Longhao Chen², Cheng Zhang¹, Pengcheng Yang¹, Tingjie Yang¹, Guosheng Yu³, Evens Pan⁴, Xiguang Wu¹, Bo Li¹, Wei Mao¹, Genquan Han¹

¹Hangzhou Institute of Technology, Xidian University, China; ²Hangzhou Dianzi University, China;

³T-HEAD Semiconductor Co., LTD, China; ⁴Arm Technology (China) Co., LTD, China

Page: 268 : [Click here to get to the paper](#)

B3L-A

Emerging Applications II

Date: Tuesday Apr 29th, 2025

Time: 14:30-15:50

Room: Lecture Hall G

Chair: Damien Querlioz

Transfer Learning for Keypoint Detection in Low-Resolution Thermal TUG Test Images

Wei-Lun Chen³, Chia-Yeh Hsieh¹, Yu-Hsiang Kao¹, Kai-Chun Liu⁴, Sheng-Yu Peng², Yu Tsao³

¹National Taiwan University, Taiwan; ²National Taiwan University of Science of Technology, Taiwan;

³Research Center for Information Technology Innovation, Academic Sinica, Taiwan; ⁴University of Massachusetts Amherst, United States

Page: 273 : [Click here to get to the paper](#)

Aiding Convergence in Federated Learning via Local Perturbation and Mutual Similarity Information

Emanuel Buttaci, Giuseppe Calafiore

Politecnico di Torino, Italy

Page: 278 : [Click here to get to the paper](#)

AI-Driven Traffic Scene Understanding Using Static LiDAR Sensors

Elham Binshafout², Chaima Zaghouni¹, Charalampos Antoniadis², Hakim Ghazzai², Nawfal

Guefrachi³, Ahmad Alsharoa³, Sameh Najeh¹, Gianluca Setti²

¹Higher School of Communication of Tunis Sup'Com, Tunisia; ²King Abdullah University of Science and Technology, Saudi Arabia; ³Missouri University of Science & Technology, United States

Page: 283 : [Click here to get to the paper](#)

An AutoEncoder-Based Semi-Supervised Learning for Efficient Blood Pressure Estimation with Minimal Labeled Data

Huimin Rong³, Hao Wu², Ruifang Liu¹

¹RingConn LLC, China; ²RingConn LLC / Shenzhen University / Shenzhen Ninenovo Technology Limited, China; ³Shenzhen University, China

Page: 288 : [Click here to get to the paper](#)

B3L-B

Pushing the Boundaries of LLMs for Cutting-Edge Circuit & System Applications

Date: Tuesday Apr 29th, 2025

Time: 14:30-15:50

Room: Lecture Hall H

Chair: Bo Li, Yongfu Li

Optimizing Inference Performance for Large Language Models on ARMv9 Architecture

Longhao Chen¹, Cheng Zhang², Huiyuan Zhang¹, Chi Wang³, Yina Zhao⁴, Xiaoxi Li², Xiguang Wu², Wei Mao², Genquan Han²

¹Hangzhou Dianzi University, China; ²Hangzhou Institute of Technology, Xidian University, China;

³Hangzhou Dianzi University, China; ⁴Wuhan University, China

Page: 293 : [Click here to get to the paper](#)

ChatDNA: An End-to-End Genetics Analysis Pipeline with Embedding Databases and Prompt Learning

Shaobo Luo³, Albert Yu³, Yiming Zhu³, Huiyuan Sun³, Zhiyuan Xie³, Ziyi Guan^{1,2}, Hong Huang¹, Ao Shen¹, Qiyuan Zhang¹, Mei Yan³, Hao Yu¹

¹Southern University of Science and Technology, China; ²University of Hong Kong, China; ³GeneSense Technology Inc., China

Page: 298 : [Click here to get to the paper](#)

Circuit-Agent: A Large Language Model Based Circuit Agent Framework for Analog/Mixed-Signal Circuit Design Automation

Tingjie Yang¹, Bo Li¹, Yongfu Li², Wei Mao¹, Genquan Han¹

¹Hangzhou Institute of Technology, Xidian University, China; ²Shanghai Jiao Tong University, China

Page: 304 : [Click here to get to the paper](#)

D2D-LLM+: Unified Translation Between Design Rules/Manuals and DRC - Bridging Inconsistencies for Accurate Implementation

Ruoyu Tang^{1,2}, Chao Wang^{1,2}, Jiajun Yap³, Zixian Guo³, Yuhang Zhang¹, Jian Zhao¹, Minghui Yin², Zhiqiang Li², Fakhrol Zaman Rokhani³, Yongfu Li^{1,2}

¹Shanghai Jiao Tong University, China; ²Institute of Microelectronics of Chinese Academy of Sciences, China; ³University Putra Malaysia, Malaysia

Page: 309 : [Click here to get to the paper](#)

B4L-A

Neuromorphic & Bioinspired Circuits, Processors, Systems & Their Applications II

Date: Tuesday Apr 29th, 2025

Time: 16:10-17:50

Room: Lecture Hall G

Chair: Jan Finkbeiner

On-Chip Learning via Transformer In-Context Learning

Jan Finkbeiner¹, Emre Neftci²

¹Fakultät für Elektrotechnik und Informationstechnik, , Germany; ²Peter Grünberg Institut, Forschungszentrum Jülich GmbH, RWTH Aachen, Germany

Page: 314 : [Click here to get to the paper](#)

ViT-LCA: A Neuromorphic Approach for Vision Transformers

Sanaz Mahmoodi Takaghaj

Pennsylvania State University, United States

Page: 319 : [Click here to get to the paper](#)

Spatiotemporal Trigger on Event Data for Backdoor Attack on Spiking Neural Networks

Shengyu Gao^{1,2}, Weiyang He¹, Chip Hong Chang^{1,2}, Amélie Gruel³, Sylvain Saïghi^{2,3}

¹Nanyang Technological University, Singapore; ²CNRS@CREATE LTD, Singapore; ³Univ. Bordeaux, CNRS, Bordeaux INP, IMS, UMR 5218, France

Page: 324 : [Click here to get to the paper](#)

A Minimal Hardware Implementation of a Neural Integrator with Graded Persistent Activity

Jiaming Wu, Adrien d'Hollande, Haoran Du, Marcelo Rozenberg

Laboratoire de Physique des Solides, Université Paris-Saclay, CNRS, France

Page: 329 : [Click here to get to the paper](#)

Non-Uniform Memory Partitioning for Low-Power Spiking Neural Networks

Simon Richter, Darío Fernández Khatiboun, Maryam Sadeghi, Milad Zamani, Farshad Moradi

Aarhus University, Denmark

Page: 334 : [Click here to get to the paper](#)

C1L-A

Frontiers of AI Hardware: Navigating the Challenges of Emerging Device Technologies

Date: Wednesday Apr 30th, 2025

Time: 09:00-10:20

Room: Lecture Hall G

Chair: Vasileios Ntinias

Improved Memristor Control Using Device Physics and Deep Reinforcement Learning

Ming-Jay Yang¹, Zhenming Yu^{1,2}, Giacomo Pedretti³, Emre Neftci^{1,2}, John Paul Strachan^{1,2}

¹Peter Grünberg Institute, Forschungszentrum Jülich GmbH, Germany; ²Fakultät für Elektrotechnik und Informationstechnik, Germany; ³Artificial Intelligence Research Lab, Hewlett Packard Labs, United States

Page: 339 : [Click here to get to the paper](#)

Memory-Efficient Spiking Neural Networks Using MLC FeFETs and Mixed Quantization

Alptekin Vardar, Franz Müller, Gonzalo Federico Cuñarro Podestá, Nellie Laleni, Nandakishor Yadav, Thomas Kämpfe

Fraunhofer Institute for Photonic Microsystems IPMS, Germany

Page: 344 : [Click here to get to the paper](#)

Drift-Aware Regularization for Long-Term Stability in Phase-Change Memory Based Neural Network Implementations

Adnan Haidar, Vasileios Ntinias, Ronald Tetzlaff

Dresden University of Technology, Germany

Page: 349 : [Click here to get to the paper](#)

Semi-Recurrent Physics Informed Neural Networks for Modelling Resistive Memories

Georgios Kleitsiotis¹, Athanasios Passias¹, Evangelos Tsipas¹, Karolos-Alexandros Tsakalos¹, Iosif-Angelos Fyrigos¹, Panagiotis Bousoulas², Stavros Kitsios³, Charalampos Tsiouostas³, Ioannis Vourkas¹, Panagiotis Dimitrakis³, Dimitris Tsoukalas², Georgios Ch. Sirakoulis¹

¹Democritus University of Thrace, Greece; ²National Technical University of Athens, Greece; ³NCSR Demokritos, Greece

Page: 354 : [Click here to get to the paper](#)

C1L-B

Circuits & Systems for AI Accelerators I

Date: Wednesday Apr 30th, 2025

Time: 09:00-10:20

Room: Lecture Hall H

Chair: Rodney Martinez Alonso

Efficient MFCC Computation for Keyword Spotting on Embedded Systems: Optimization Techniques and Performance Analysis

Youssef Kandil, Mohammed Shalan

American University in Cairo, Egypt

Page: 359 : [Click here to get to the paper](#)

An Analog Neural Network ASIC for In-Sensor Processing of Radiation Detector Signals

Susanna Di Giacomo, Michele Ronchi, Mattia Amadori, Giacomo Borghi, Marco Carminati, Carlo Fiorini

Politecnico di Milano / Istituto Nazionale di Fisica Nucleare (INFN), Italy

Page: 364 : [Click here to get to the paper](#)

Energy-Efficient, Nonvolatile 8T-2MTJ Ternary CAM Cell for Low Search Delay

Mohit Gupta^{1,2}, Ravi S. Siddanath^{1,2}, Mudit Gupta^{1,3}, Surya Prakash¹, Manish Goswami¹, Kavindra Kandpal¹

¹*ECE Dept. IIIT Allahabad, India;* ²*Broadcom Inc.;* ³*GlobalFoundries*

Page: 369 : [Click here to get to the paper](#)

Tru-Sign: Lossless Truncated Sign Based Energy Efficient DNN Accelerator

Kyungchul Lee, Dongwoo Lew, Jongsun Park

Korea University Seoul, Korea

Page: 374 : [Click here to get to the paper](#)

C2L-A

Architecture for AI Computing

Date: Wednesday Apr 30th, 2025

Time: 10:40-12:00

Room: Lecture Hall G

Chair: Amir Yousefzadeh

Reconfigurable Dataflow and Tile Shape Optimization in SystemC-Based NoC of 1D PE Array Simulators Integrated with TVM

Yu-Jen Chang, Ching-Te Chiu, Tzu-Chun Huang

National Tsing Hua University, Taiwan

Page: 379 : [Click here to get to the paper](#)

SWAP-GNG: A Scalable HW Architecture for Multi Growing Neural Gas Graph Learning for Continual Learning

Florent Derue, Slavisa Jovanovic, Hassan Rabah, Serge Weber

Université de Lorraine, CNRS, IJL, France

Page: 384 : [Click here to get to the paper](#)

Activation Swapping for Feed Forward Networks in Batched LLM Inference on NPU

Seongho Jeong, Minseok Seo, Jungi Hyun, Hyuk-Jae Lee, Xuan Truong Nguyen

Inter-University Semiconductor Research Center (ISRC), Seoul National University, Korea

Page: 389 : [Click here to get to the paper](#)

An Adaptive Scheme of Threshold Adjustment for Dynamic Sparsity Extraction of Self-Attention Network

Yong-Lun Xiao¹, Chia-Wei Chang¹, Cheng-Ting Shih¹, Jing-Jia Liou¹, Chih-Tsun Huang¹, Yao-Hua Chen¹, Juin-Min Lu²

¹*National Tsing Hua University, Taiwan*; ²*Industrial Technology Research Institute, Taiwan*

Page: 394 : [Click here to get to the paper](#)

C2L-B

Circuits & Systems for AI Accelerators II

Date: Wednesday Apr 30th, 2025

Time: 10:40-12:00

Room: Lecture Hall H

Chair: Kaori Sekine

Codeword Decomposition and Recoding Based Lossless Model Compression Algorithm and its VLSI Decompressor Design

Lyu-Ming Ho, Chih-Yao Liang, Juinn-Dar Huang
National Yang Ming Chiao Tung University, Taiwan
Page: 399 : [Click here to get to the paper](#)

KV Cache Compression Based on Token-Level Redundancy Elimination and Bit-Level Encoding

Tianze Li², Chenjia Xie², Jiayuan Chen¹, Yuan Du², Li Du²
¹*China Mobile Research Institute, China*; ²*Nanjing University, China*
Page: 404 : [Click here to get to the paper](#)

A Multicast-Capable AXI Crossbar for Many-Core Machine Learning Accelerators

Luca Colagrande, Luca Benini
ETH Zürich, Switzerland
Page: 409 : [Click here to get to the paper](#)

DPP-ViT: Dynamic Patch Pruning for Low Complexity Vision Transformer Accelerator

Han Cho¹, Joongho Jo², Seung-Eon Hwang², Jongsun Park²
¹*Georgia Institute of Technology, United States*; ²*Korea University Seoul, Korea*
Page: 414 : [Click here to get to the paper](#)