DOI: 10.4218/etr2.12735

GUEST EDITORIAL

ETRI Journal WILEY

Special issue on next-gen AI and quantum technology

Artificial intelligence (AI) and quantum technology are two key fields that drive the development of modern science and technology, and their developments have had tremendous impacts on academia and industry. AI is a technology that can solve complex problems through data-based learning and inference and is already driving innovation in various industries such as healthcare, finance, and manufacturing. In particular, the development of AI has enabled practical applications in autonomous driving, natural language processing, and image recognition, greatly improving the quality of human life.

Quantum technology utilizes the principles of quantum mechanics to provide new computational capabilities beyond the scope of classical computing methods. Quantum computing has the potential to perform multiple calculations simultaneously using quantum bits (qubits), which is expected to lead to innovative results in complex optimization problems and the analysis of large datasets. Additionally, quantum technology plays an important role in secure communication, with technologies such as quantum key distribution (QKD) providing security surpassing that of existing encryption methods.

The Electronics and Telecommunications Research Institute (ETRI) Journal is a peer-reviewed open-access journal that launched in 1993 and is published bimonthly by ETRI of the Republic of Korea, aiming to promote worldwide academic exchange in information, telecommunications, and electronics. This special issue of the ETRI Journal focuses on exploring the latest research on these cutting-edge technologies and highlighting the challenges and opportunities that each technology presents. The research included in this special issue clearly demonstrates the significant impact that each of the advancements in both AI and quantum technologies have on academia and industry. AI is already driving change in many fields and is focused on creating more efficient and intelligent systems. In contrast, quantum technologies are introducing a novel computing paradigm, revealing groundbreaking possibilities for computational power and secure communication.

The papers selected for this special issue cover various aspects of AI and quantum technologies. In the AI field, the latest hardware architectures, energy-efficient AI systems such as spiking neural networks (SNNs), and AI application technologies such as anomaly detection are introduced. In the field of quantum technology, theoretical developments in quantum computing, quantum photonic systems, and secure communication technologies such as QKD are discussed.

The first paper [1], titled "Trends in quantum reinforcement learning: State-of-the-arts and the road ahead by Park and Kim," is an invited paper. This paper presents the foundational quantum reinforcement learning theory and explores quantum-neural-network-based reinforcement learning models with advantages such as fast training and scalability. It also discusses multi-agent applications, including quantum-centralized critics and multiple-actor networks. Future research directions include federated learning, autonomous control, and quantum deep learning software testing.

In the second paper [2], titled "Optimal execution of logical hadamard with low-space overhead in rotated surface code," Lee et al. propose a novel method for executing the logical Hadamard operation in rotated surface codes with minimal space requirements. Using boundary deformation, this method rotates the logical qubit affected by the transversal Hadamard and restores its original encoding through logical flip-and-shift operations. The space–time cost for this approach is $5d^3 + 3d^2$, offering an efficiency that is approximately four times greater than that of previous methods. It requires only two patches for implementation, in contrast to the traditional seven patches, and maintains parallelism in quantum circuits by avoiding interference between adjacent logical qubits.

The third paper [3], titled "Quantum electrodynamical formulation of photochemical acid generation and its implications on optical lithography" by Lee, refines photochemical acid generation using quantum electrodynamics principles, providing a probabilistic description of acid generation and deprotection density in photoresists.

It combines quantum mechanical acid generation with deprotection mechanisms to analyze stochastic feature formation, thereby offering key insights into the randomness of the deprotection process.

The fourth paper [4], titled "Fabrication of low-loss symmetrical rib waveguides based on x-cut lithium niobate on insulator for integrated quantum photonics" by Kim et al., presents the fabrication of a low-loss lithium niobate on insulator (LNOI) rib waveguide with an optical propagation loss of 0.16 dB/cm, which was achieved by optimizing etching parameters. The shallow etching process improves waveguide symmetry and smoothness on x-cut LNOI, supporting advancements in on-chip quantum photonic devices.

The fifth paper [5], titled "Metaheuristic optimization scheme for quantum kernel classifiers using entanglement-directed graphs" by Tjandra and Sugiarto, presents a novel meta-heuristic approach using a genetic algorithm to optimize quantum kernel classifiers by incorporating entanglement-directed graphs. This method effectively enhances classification performance by designing quantum circuits that leverage entanglement, outperforming classical and other quantum baselines across various datasets. The results demonstrate that the proposed approach successfully identifies optimal entanglement structures for specific datasets, leading to significant improvements in classification accuracy and F1 scores.

In the sixth paper [6], titled "Free-space quantum key distribution transmitter system using wavelength-division multiplexing (WDM) filter for channel integration," Kim et al. introduce a free-space QKD transmitter system using the BB84 protocol that eliminates the need for internal alignment. It uses a custom WDM filter and polarization-encoding module to integrate the quantum and synchronization channels. This integration avoids the complex alignment processes required for conventional systems with bulk optics. The WDM filter efficiently multiplexes 785 and 1550 nm signals, with insertion losses of 1.8 and 0.7 dB, respectively. The system achieved a sifted key rate of 1.6 Mbps and qubit error rate of 0.62% at 100 MHz, exhibiting performance comparable to that of traditional bulk-optic devices.

The seventh paper [7], titled "PF-GEMV: Utilization maximizing architecture in fast matrix-vector multiplication for GPT-2 inference" by Kim et al., presents solutions for overcoming the challenges of processing matrix–vector multiplications (GEMVs). It examines the challenges faced by AI processors in light of the rapid advancement of transformer-based artificial neural networks, particularly the need to perform matrix–vector multiplication efficiently alongside traditional matrix–matrix multiplication (GEMM). The authors noted that existing AI processor architectures are primarily optimized for GEMMs, leading to considerable throughput degradation when handling GEMV.

To address this issue, their paper introduces a port-folding GEMV scheme that incorporates multi-format and low-precision techniques while leveraging an outer-product-based processor designed for conventional GEMM operations. This innovative approach achieved an impressive 93.7% utilization on GEMV tasks with an eight-bit format on an 8×8 processor, resulting in a $7.5\times$ throughput increase compared with the original design. Additionally, when applied to the matrix operations of the GPT-2 large model, the proposed scheme demonstrated a remarkable $7\times$ speedup on single-batch inferences. The eighth paper [8], titled "SNN eXpress: Streamlining low-power AI-SoC development with unsigned weight accumulation spiking neural network" by Jang et al., presents solutions for overcoming the challenges of developing low-power AI-SoCs using analog-circuit-based unsigned weight accumulating spiking neural networks (UWA-SNNs). It introduces the SNN eXpress tool, which automates the design process, enabling the rapid development and verification of UWA-SNN-based AI-SoCs, as demonstrated by the creation of two AI-SoCs.

In the ninth paper [9], titled "XEM: Tensor accelerator for AB21 supercomputing artificial intelligence processor," Jeon et al. introduce the XEM accelerator, which is designed to enhance the AB21 supercomputing AI processor's efficiency when performing the tensor-based linear-algebraic operations that are crucial for hyperscale AI and high-performance computing applications. The XEM architecture, with its outer-product-based parallel floating-point units, is detailed along with new instructions and hardware characteristics, with future verification planned alongside the AB21 processor chip.

The tenth paper [10], titled "NEST-C: A deep learning compiler framework for heterogeneous computing systems with AI accelerators" by Park et al., introduces NEST-C, a deep learning compiler framework designed to optimize the deployment and performance of deep learning models across various AI accelerators. This framework achieves significant computational efficiency by incorporating profiling-based quantization, dynamic graph partitioning, and multilevel intermediate representation integration. The experimental results demonstrate that NEST-C enhances throughput and reduces latency across different hardware platforms, making it a versatile tool for modern AI applications.

The eleventh paper [11], titled "Mixed-mode SNN crossbar array with embedded dummy switch and mid-node pre-charge scheme" by Park et al., presents solutions for overcoming the challenges of processing

GEMVs. This paper introduces a membrane computation error-minimized mixed-mode SNN crossbar array. The authors implemented an embedded dummy switch scheme along with a mid-node pre-charge technique to create a high-precision current-mode synapse. This innovative approach effectively mitigates charge sharing between membrane capacitors and the parasitic capacitance of synapses, thereby reducing computational errors. A prototype chip featuring a 400 × 20 SNN crossbar was fabricated using a 28 nm fully depleted silicon on insulator (FDSOI) complementary metal–oxide–semiconductor (CMOS) process, successfully recognizing 20 MNIST patterns reduced to 20 × 20 pixels with a power consumption of 411 μW. Notably, the peak-to-peak deviation of the normalized output spike count from the 21 fabricated SNN prototype chips remained within 16.5% of the ideal value, accounting for random sample-wise variations.

The twelfth paper [12], titled "Asynchronous interface circuit for nonlinear connectivity in multi-core spiking neural networks" by Oh et al., presents solutions for overcoming the challenges of processing GEMVs. This paper addresses the need for an interface circuit that supports multiple SNN cores to facilitate SNN expansion. The proposed circuit employs an asynchronous design approach to mimic the operational characteristics of the human brain; however, the lack of a global clock introduces timing challenges during implementation. To address this issue, the authors proposed an intermediate latching template that establishes asynchronous nonlinear connectivity and enables multi-pipeline processing utilizing multiple SNN cores.

This design incorporates arbitration and distribution blocks based on the proposed template and is fabricated as a fully custom interface circuit supporting four SNN cores in a 28 nm CMOS FDSOI process. The results indicate that this innovative template can enhance the throughput of the interface circuit by up to 53% compared with conventional asynchronous designs. Additionally, the interface circuit can transmit a spike with an energy consumption of 1.7 pJ at a supply voltage of 0.9 V, supporting 606 Mevent/s for intra-chip communication, and 3.7 pJ at the same voltage for 59 Mevent/s inter-chip communication.

The thirteenth paper [13], titled "AONet: Attention network with optional activation for unsupervised video anomaly detection" by Rakhmonov et al., proposes AONet, a novel attention-based neural network designed for unsupervised video anomaly detection, which incorporates a unique activation function (OptAF) combining the benefits of the rectified linear unit (ReLU), leaky ReLU, and sigmoid functions. This method efficiently captures spatiotemporal features using a temporal shift module and residual autoencoder, achieving superior performance on benchmark datasets compared with state-of-the-art methods. The model was evaluated on three datasets, demonstrating competitive accuracy and speed with an area under the curve (AUC) score of 97.0% on the UCSD Ped2 dataset.

As guest editors, we are pleased to explore the future of these cutting-edge technologies in this special issue. We express our deepest gratitude to the authors for their outstanding contributions, reviewers for their thorough review, and editorial team for their assistance in publishing this special issue. We hope that this special issue will contribute to a broader understanding of the developments in AI and quantum technology and foster further research and innovation.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest.

Ji-Hoon Kim [1]
Ho-Young Cha [2]
Daewoong Kwon [3]
Gyu Sang Choi [4]
HeeSeok Kim [5]
Yousung Kang [6]

[1] *Department of Electronic and Electrical Engineering, Ewha Womans University, Seoul, Republic of Korea*
[2] *School of Electronic and Electrical Engineering, Hongik University, Seoul, Republic of Korea*
[3] *Department of Electronic Engineering, Hanyang University, Seoul, Republic of Korea*
[4] *School of Computer Science and Engineering, Yeungnam University, Gyeongsan, Republic of Korea*
[5] *Department of Cyber Security, College of Science and Technology, Korea University, Sejong, Republic of Korea*
[6] *Cryptography Engineering Research Section, Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea*

Correspondence
Ji-Hoon Kim, Department of Electronic and Electrical Engineering, Ewha Womans University, Seoul, Republic of Korea.
Email: jihoonkim@ewha.ac.kr

REFERENCES

1. S. Park and J. Kim, *Trends in quantum reinforcement learning: state-of-the-arts and the road ahead*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0153
2. S.-M. Lee, K.-S. Jin, S.-C. Oh, J.-H. On, and G.-I. Cha, *Optimal execution of logical hadamard with low-space overhead in rotated surface code*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0129

3. S. Lee, *Quantum electrodynamical formulation of photochemical acid generation and its implications on optical lithography*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0127
4. H.-S. Kim, G. Kim, T. Slusar, J. Kim, J. Park, J. Park, H. Hwang, W. Noh, H. Lee, M.-K. Seo, K. Moon, and J. J. Ju, *Fabrication of low-loss symmetrical rib waveguides based on x-cut lithium niobate on insulator for integrated quantum photonics*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0137
5. Y. Tjandra and H. S. Sugiarto, *Metaheuristic optimization scheme for quantum kernel classifiers using entanglement-directed graphs*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0144
6. M. Kim, K. Lim, J.-S. Choe, B.-S. Choi, K.-J. Kim, J. H. Baek, and C. J. Youn, *Free-space quantum key distribution transmitter system using WDM filter for channel integration*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0142
7. H. Kim, Y. Lee, and C.-G. Lyuh, *PF-GEMV: utilization maximizing architecture in fast matrix-vector multiplication for GPT-2 inference*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0111
8. H. Jang, K. Han, K.-I. Oh, S. Lee, J.-J. Lee, and W. Lee, *SNN eXpress: streamlining low-power AI-SoC development with unsigned weight accumulation spiking neural network*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0114
9. W. Jeon, M. Y. Lee, J. H. Lee, and C.-G. Lyuh, *XEM: tensor accelerator for AB21 supercomputing artificial intelligence processor*, ETRI j **46** (2024), no. 5, DOI 10.4218/etrij.2024-0141
10. J. Park, M. Yu, J. Kwon, J. Park, J. Lee, and Y. Kwon, *NEST-C: a deep learning compiler framework for heterogeneous computing systems with AI accelerators*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0139
11. K.-I. Oh, H. Kim, T. Kang, S.-E. Kim, J.-J. Lee, and B.-D. Yang, *Mixed-mode SNN crossbar array with embedded dummy switch and mid-node pre-charge scheme*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0120
12. S.-E. Kim, K.-I. Oh, T. Kang, S. Lee, H. Kim, M.-J. Park, and J.-J. Lee, *Asynchronous interface circuit for nonlinear connectivity in multi-core spiking neural networks*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0135
13. A. A. U. Rakhmonov, B. Subramanian, B. A. Varnousefaderani, and J. Kim, *AONet: attention network with optional start for unsupervised video anomaly detection*, ETRI J **46** (2024), no. 5, DOI 10.4218/etrij.2024-0115

AUTHOR BIOGRAPHIES

Ji-Hoon Kim received his BS degree (summa cum laude) and PhD in electrical engineering and computer science from KAIST, Daejeon, South Korea, in 2004 and 2009, respectively. In 2009, he joined Samsung Electronics, Suwon, South Korea, as a Senior Engineer and worked on the next-generation architecture for a 4G communication modem SoC. He was an associate professor with the Department of Electronics Engineering, Chungnam National University, Daejeon, Korea, from 2010 to 2016. In 2018, he joined the Department of Electronic and Electrical Engineering, Ewha Womans University, Seoul, South Korea, as a Faculty Member, where he is currently a professor. His current interests include CPU microarchitectures, domain-specific SoCs, and deep neural-network accelerators. Dr. Kim served on the Technical Program Committee and Organizing Committee for various international conferences, including the IEEE International Conference on Computer Design (ICCD), IEEE Asian Solid-State Circuits Conference (A-SSCC), and IEEE International Solid-State Circuits Conference (ISSCC). He was a co-recipient of the Distinguished Design Award at the 2019 IEEE A-SSCC, Best Design Award at the 2007 Dongbu HiTek IP Design Contest, First Place Award at the 2008 International SoC Design Conference (ISOCC) Chip Design Contest, and IEEE/IEIE Joint Award for Young Scientists and Engineers.

Ho-Young Cha received his BS and MS degrees in electrical engineering from Seoul National University in 1996 and 1999, respectively, and PhD in electrical and computer engineering from Cornell University in 2004. He is a professor of electronic and electrical engineering at Hongik University. He was a postdoctoral research associate at Cornell University from 2004 to 2005 and a research engineer at GE Global Research from 2005 to 2007. His main research interests include wide bandgap semiconductor power and RF devices. He serves as a section editor for ETRI Journal.

Daewoong Kwon received his BS degree in electronic engineering from Kwangwoon University in 2005, and his MS degree and PhD in electrical engineering and computer science from Seoul National University in 2009 and 2017, respectively. He was a senior engineer at Samsung electronics on the NAND Flash development team from 2009 to 2014. He worked at UC Berkeley and Intel as a research scientist and senior engineer from 2017 to 2019. He is an assistant professor of electronic engineering at Hanyang University. His main research interests include next-generation low-power logic and memory devices.

Gyu Sang Choi received his PhD from the Department of Computer Science and Engineering, The Pennsylvania State University, University Park, PA, USA, in 2005. From 2006 to 2009, he was a Research Staff Member of the Samsung Advanced Institute of Technology (SAIT), Samsung Electronics. Since 2009, he has been a faculty member with the School of Software Convergence, Yeungnam University, South Korea. His research interests include data mining, natural language processing, and reinforcement learning.

HeeSeok Kim received his BS degree in mathematics from Yonsei University, Seoul, South Korea, in 2006, and his MS degree and PhD in engineering and information security from Korea University, Seoul, in 2008 and 2011, respectively. He was a postdoctoral researcher with the University of Bristol, UK, from 2011 to 2012. From 2013 to 2016, he was a senior researcher with the Korea Institute of Science and Technology Information (KISTI). Since 2016, he has worked at Korea University. His research interests include side-channel attacks, cryptography, and network security.

Yousung Kang received BS and MS degrees in electronics engineering from Chonnam National University, Gwangju, South Korea, in 1997 and 1999, respectively, and a PhD in electrical and electronic engineering from KAIST, Daejeon, South Korea, in 2015. He has been a principal researcher with the Electronics and Telecommunications Research Institute, Daejeon, since 1999. Since 2004, he has been an IT International Standard Expert of the Telecommunications Technology Association in Seoul, South Korea. He was a visiting researcher at Queens University Belfast, Belfast, UK, from 2011 to 2012. His current research interests include quantum security, drone security, side-channel analysis, and key-hiding technologies.