SCHOOL OF COMPUTATION, INFORMATION AND TECHNOLOGY — CHAIR OF AI PROCESSOR DESIGN

TECHNISCHE UNIVERSITÄT MÜNCHEN

Seminar AI Processor Design (WS 2024/25)

Seminar Report

The Next Generation of Brain-Inspired Computing: Hybrid Neural Networks

Author:	Marcos Ibáñez Fandos
Matrikelnummer:	03801227
Examiner:	Prof. Dr.-Ing. Hussam Amrouch
Supervisor:	Paul R. Genssler, Simon Thomann
Submission Date:	June 29, 2025

The Next Generation of Brain-Inspired Computing: Hybrid Neural Networks

Marcos Ibáñez Fandos
Electrical Engineering and Information Technology
TUM School of Computation, Information and Technology
Technical University of Munich
Munich, Germany
Email: marcos.ibanez@tum.de

***Abstract*—The advancement towards Artificial General Intelligence (AGI) favors neural network research and, in particular, the development of Hybrid Neural Networks (HNNs)—architectures that integrate Artificial Neural Networks (ANNs) and Spiking Neural Networks (SNNs). This paper reviews several prominent approaches to hybrid neural computation, including Hybrid Units (HUs), the Tianjic chip architecture, first-principles-based models, and Optical Neural Networks (ONNs). Each methodology is analyzed in terms of its design philosophy, technical advantages, and specific trade-offs, with the ultimate objective of assessing how these hybrid strategies can advance the field toward robust, scalable brain-inspired computation.**

***Index Terms*—Hybrid Neural Networks, Brain-Inspired Computing, Artificial Neural Networks, Spiking Neural Networks.**

I. Introduction

The pursuit of AGI—the capability for machines to understand, learn, and reason across a wide range of tasks and environments—remains a leading challenge in computational neuroscience and artificial intelligence. Meeting this challenge requires systems that can process complex, dynamic, and uncertain information in a manner reminiscent of biological brains. Achieving this level of intelligence is widely regarded as one of the foremost scientific and engineering problems of our era.

Although significant advances have been made in machine learning, conventional paradigms are still far from achieving the flexibility, generalization, and adaptivity exhibited by biological intelligence. A key bottleneck has been the separation between approaches prioritizing computational efficiency and those inspired by neurobiological realism. While many modern AI systems have excelled in narrow, well-defined tasks, their capabilities often do not generalize well to new domains or situations without extensive retraining or manual intervention. This challenge has motivated an intense search for neural architectures that combine the strengths of different computational models, pushing the boundaries of what artificial systems can achieve.

Among the diverse neural paradigms developed to address these goals, ANNs and SNNs stand out for their distinctive strengths and weaknesses. ANNs have demonstrated remarkable success in a variety of domains—such as computer vision, natural language processing, and strategic gameplay—thanks to their ability to leverage large datasets and highly parallelized hardware. However, they tend to lack energy efficiency, temporal precision, and biological plausibility. By contrast, SNNs mimic the event-driven and asynchronous processing of real neurons, achieving ultra-low power consumption and excelling in applications involving temporal or spatiotemporal data streams. Nevertheless, SNNs are currently limited by the complexity of their training algorithms and the difficulty of scaling to large problem sizes with competitive performance.

The limitations of each architecture have motivated the development of HNNs, which seek to merge their complementary properties. Rather than adhering strictly to one computational paradigm, HNNs aim to integrate the learning flexibility and high-dimensional capacity of ANNs with the temporal dynamics and energy efficiency of SNNs. This synthesis opens up new possibilities for applications that demand both robust pattern recognition and real-time adaptation, such as robotics, brain-machine interfaces, sensor fusion, and autonomous decision-making under uncertainty.

The push towards hybridization is further reinforced by rapid advances in neuromorphic hardware, photonic computing, and novel learning algorithms. Hardware platforms capable of supporting heterogeneous neural computations are emerging as a critical enabler for scalable and efficient HNNs. These innovations not only enhance computational performance, but also facilitate the practical deployment of hybrid neural models in resource-constrained environments, including edge devices and embedded systems.

From a research perspective, the hybrid neural paradigm also invites new questions: What are the best ways to combine different neural models within a single framework? How can we ensure interpretability, robustness, and scalability in such heterogeneous systems? What trade-offs arise between performance, generality, and physical realizability? Addressing these questions is vital for charting a course towards next-generation brain-inspired AI and for ultimately achieving the vision of AGI.

In this context, a growing body of work has proposed a range of hybrid neural architectures and platforms. Notable recent contributions include the HUs framework by Zhao et al. [1], which provides a modular approach for seamless integration of ANNs and SNNs; the Tianjic chip by Pei et al. [2], representing a hardware-centric solution for hybrid neural computation; first-principles-based hybrid networks as pioneered by Psichogios and Ungar [3], which merge mechanistic domain knowledge with data-driven learning; and ONNs as explored by Jutamulia and Yu [4], which exploit the unique computational properties of photonics for neural processing.

This paper reviews these state-of-the-art HNNs, examining their underlying concepts, representative implementations, and application domains. Our focus is on providing a panoramic view of the current landscape, critically comparing the different methodologies, and identifying the opportunities and challenges that will shape future research in brain-inspired hybrid computation.

II. Background

A wide range of artificial intelligence research has centered on the development of neural paradigms that capture different aspects of learning, representation, and computation. Among these, ANNs and SNNs represent two foundational yet distinct approaches, each with characteristic advantages and constraints.

ANNs are inspired by the interconnected structure of biological neurons but have been developed primarily for mathematical tractability, efficient parallelism, and powerful data-driven learning. Architecturally, these models consist of layers of nodes ("neurons") that aggregate weighted inputs and apply nonlinear activation functions. Global learning algorithms, such as backpropagation, enable ANNs to excel at complex pattern recognition, function approximation, and feature extraction [5]. However, their operation is largely synchronous and real-valued, requiring significant computational resources and large training datasets, while lacking the event-driven, energy-efficient nature observed in biological neural systems.

By contrast, SNNs more closely emulate the temporal dynamics of real neurons, transmitting information via spikes or action potentials in an asynchronous manner. These networks support temporal coding and extremely low power consumption, particularly when implemented on neuromorphic hardware [6]. Local learning rules, such as Spike-Timing-Dependent Plasticity (STDP), further differentiate their training mechanisms. Despite these advantages, SNNs face major challenges: scalable training, global optimization, and competitive accuracy on static, high-dimensional tasks.

The inherent trade-off between the learning flexibility of ANNs and the biological realism and efficiency of SNNs has led to the emergence of HNNs. These architectures aim to synthesize the complementary strengths of both paradigms, targeting robust and efficient processing of both spatial and temporal data, and enabling greater scalability and interpretability. Hybridization strategies also benefit from recent advances in neuromorphic and hybrid hardware platforms, as well as the development of learning algorithms that combine global and local adaptation. As highlighted by recent surveys [7], [8], HNNs offer a promising direction for bridging the gap between current neural models and the requirements of AGI.

A central challenge in advancing HNNs is the design of architectures and training paradigms that harness the strengths of each neural model while mitigating their respective weaknesses. Strategies for coupling real-valued, synchronous computations with event-driven, asynchronous processing are central to the operation of effective hybrids. Likewise, scalable implementations depend on both algorithmic innovation and hardware support for heterogeneous computing, including cross-modal signal conversion, joint optimization, and real-time adaptation.

The remainder of this article will analyze several representative HNNs frameworks—such as HUs [1], the Tianjic chip [2], first-principles-based architectures [3], and ONNs [4]—evaluating their methodologies, technical properties, and application domains.

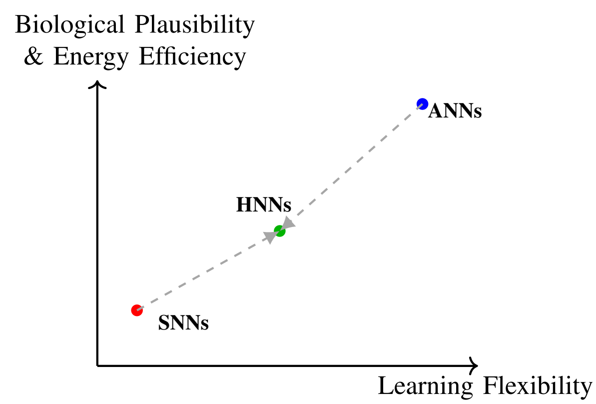

Fig. 1. Conceptual evolution of neural network paradigms. SNNs offer high biological plausibility and energy efficiency, but limited learning flexibility. ANNs provide high learning flexibility at the cost of biological realism and efficiency. HNNs aim to balance both dimensions.

III. HNNs: Methodologies and Implementations

A. HUs Framework

The HU framework, introduced by Zhao et al. [1], offers a novel and systematic method for integrating ANNs with SNNs. Central to this framework is the HU itself—a modular component purpose-built to mediate between these two distinct paradigms. Each HU serves as an interface that can convert asynchronous, event-driven signals from SNNs into the synchronous, real-valued formats required by ANNs, and vice versa. This bidirectional interoperability is essential for managing the complex, heterogeneous data flows encountered in hybrid neural architectures.

Structurally, HUs are composed of several configurable modules, such as windowing mechanisms for temporal segmentation, kernel-based transformations for extracting spatiotemporal features, nonlinear activation functions to preserve approximation capabilities, and discretization or thresholding units. The windowing modules enable the synchronization of continuous data streams by dividing them into discrete segments amenable to ANNs processing. Kernel transformations are particularly important for converting the sparse, temporally coded information from SNNs into dense, real-valued representations. Together, these mechanisms allow HUs to effectively approximate complex mappings between the two network paradigms, thereby facilitating seamless integration within heterogeneous neural network systems.

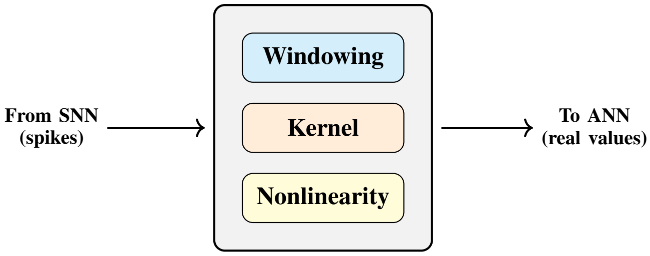

Fig. 2. Modular structure of a HU, which bridges SNNs and ANNs through three main stages: windowing, kernel transformation, and nonlinearity.

One standout feature of the HU framework is its exceptional flexibility and adaptability. Zhao and colleagues outline two main configuration approaches: manual adjustment and autonomous learning. In scenarios where domain expertise is strong and transformations are relatively straightforward, parameters can be set directly, leveraging prior knowledge for deterministic tasks. On the other hand, for more complex or less understood relationships, the framework's automatic learning capabilities come into play—allowing HUs to optimize their own parameters in response to messy, unpredictable data.

This versatility extends to support for hierarchical and multi-domain hybrid models, affording researchers considerable freedom in designing intricate neural architectures. The authors back up this claim with practical examples, such as hybrid sensing networks that seamlessly integrate spatial and temporal data for improved tracking accuracy and energy efficiency. Hierarchical modulation networks further demonstrate the system's ability to support meta-continual learning across diverse tasks, while hybrid reasoning networks offer robust, interpretable solutions for multimodal reasoning challenges.

Compared to other hybrid approaches—like the Tianjic chip, First Principles-HNNs, or optical implementations—the HU framework stands out for its modularity and ease of configuration. Its design strikes a careful balance between flexibility and complexity, giving both researchers and developers a customizable, scalable toolset for implementing advanced hybrid neural systems.

In summary, the HU framework represents a significant advancement in hybrid neural computing. By blending computational flexibility with adaptive learning strategies, the HU framework broadens the practical applicability and scalability of hybrid models for real-world applications.

While the HU framework emphasizes modular software-based integration, other methodologies aim to realize hybridization directly at the hardware level. One prominent example is the Tianjic chip architecture, which will be discussed next.

B. Tianjic Chip Architecture

The Tianjic chip, introduced by Pei et al. [2], marks a significant development in the field of AI hardware. Unlike most existing platforms, which typically focus on either ANNs or SNNs, Tianjic integrates both paradigms within a single, highly reconfigurable architecture. This allows for simultaneous operation and direct interaction between ANNs and SNNs, effectively overcoming the limitations of earlier chips restricted to just one computational approach.

Central to the Tianjic design is a scalable mesh of processing units known as FCores. Each FCore is programmable to function either in ANN mode—supporting multi-bit weighted-sum operations and nonlinear activations—or in SNN mode, which enables spike-based, event-driven computation and local learning rules such as STDP. This dual functionality is achieved through a unified hardware design featuring flexible control logic, shared memory, and configurable communication pathways. The FCores are interconnected using a hybrid routing network, supporting both synchronous and asynchronous data transfers, which facilitates the construction of arbitrarily complex hybrid topologies.

A particularly notable feature of the Tianjic chip is its capacity to execute ANN and SNN workloads concurrently on the same silicon. For instance, in autonomous vehicle applications, the chip can process camera data through an ANN pipeline while simultaneously handling event-based sensor data with SNN pipelines, ultimately integrating both modalities to enable robust decision-making. This hybrid operation has been validated in real-time robotics and control scenarios, highlighting the practical benefits of cross-paradigm neural computing.

In terms of performance, Tianjic demonstrates marked improvements in energy efficiency compared to conventional ANN- or SNN-only architectures. According to reports from both Pei et al. [2] and Zhang et al. [8], Tianjic achieves energy efficiency figures of approximately 9.9 TOPS/W (trillions of operations per second per watt). In contrast, leading ANN chips typically reach about 2.5 TOPS/W, while state-of-the-art SNN chips, such as Intel's Loihi or IBM's TrueNorth, range from 2.0 to 3.0 TOPS/W. This efficiency advantage is illustrated in Fig. 3, underscoring the value of Tianjic's hybridized hardware for energy-constrained edge applications and large-scale AI workloads.

The Tianjic chip, while offering notable advancements, has a set of substantial design hurdles. Integrating two fundamentally distinct computational paradigms on a single hardware platform inevitably increases both design complexity and the physical area required. Efficiently translating high-level hybrid models into something the chip can efficiently process is still a work in progress—current compilation and scheduling tools are under constant development. Additionally, although the chip is theoretically highly scalable, deploying the Tianjic chip at larger scales introduces real-world issues like bandwidth constraints and synchronization bottlenecks throughout the hybrid routing network.

In essence, the Tianjic chip stands as a cutting-edge example

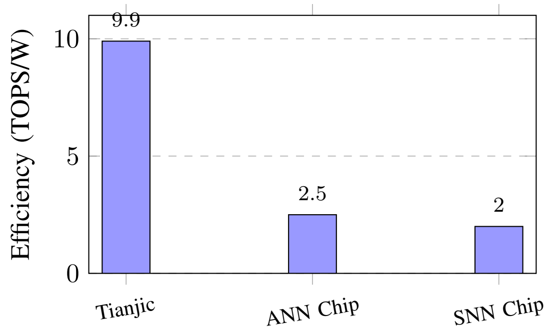

Fig. 3. Energy efficiency comparison: Tianjic chip versus representative ANN and SNN chips (values from Pei et al. 2019 and Zhang et al. 2020). Higher is better.

of brain-inspired, hybrid neural computing implemented at the hardware level. Its adaptive architecture successfully bridges the gap between ANNs and SNNs, resulting in measurable gains in both performance and efficiency. As demand grows for AI accelerators that are both flexible and energy-efficient, the Tianjic chip sets a high bar for what future hybrid and neuromorphic processors could achieve.

While the Tianjic chip exemplifies hardware-based integration, an alternative strategy is to combine data-driven learning with domain knowledge, as in first-principles-HNNs.

C. Hybrid ONNs

Following, we consider hybrid optical neural networks (ONNs), as discussed by Jutamulia and Yu [4]. These represent an early effort to harness the unique strengths of optical hardware for neural computation. The core idea is to merge the parallelism and high bandwidth of optical systems with the flexibility of electronics, aiming for neural computations at speeds that conventional electronics can't match.

In a typical hybrid ONN, computational tasks are assigned according to what each system does best. The optical subsystem handles massively parallel operations—for example, matrix-vector multiplications, which are central in both feedforward and recurrent ANNs. Here, inputs (like pixel intensities or feature values) are encoded optically, using spatial modulators or laser arrays. The optical signals are then processed through lenses, gratings, or holographic elements that serve as the network's weight matrices. Due to light's wave nature, millions of computations can be performed in parallel, limited mainly by optical resolution and coherence.

After this linear and parallel processing, electronic photodetectors convert the optical outputs back to electrical signals. Nonlinear activation functions, which are crucial for neural networks, are typically implemented electronically, since optical nonlinearity remains a significant technical hurdle. The outputs can be re-encoded optically for further processing or used directly for inference.

To summarize, while hybrid ONNs are still evolving, they offer an intriguing approach to accelerating neural computation by strategically combining optical and electronic domains.

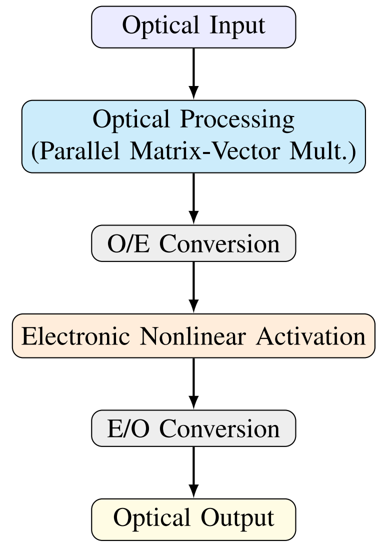

Fig. 4. Schematic of an ONN layer: input data are encoded as optical signals, processed in parallel by optical elements (matrix-vector multiplication), converted to electronic signals for nonlinear activation, and optionally re-encoded optically for subsequent layers.

Jutamulia and Yu [4] presented a range of case studies on hybrid ONNs, demonstrating that these systems can perform tasks such as pattern recognition, associative memory, and image processing at speeds far surpassing those of conventional electronic-only approaches. The use of optics not only accelerates computation but also enables highly parallel storage and retrieval of large datasets through holographic memory techniques.

Nonetheless, there are notable limitations. Optical systems are highly sensitive to factors like alignment, environmental disturbances (such as vibrations and temperature fluctuations), and optical noise, all of which can degrade system performance in practical scenarios. The integration of optical subsystems with standard electronic components introduces significant interface challenges; signal conversion between optical and electrical domains (O/E and E/O) results in increased latency and energy consumption. Additionally, scalability is limited by the physical footprint of optical components and the complexity involved in dynamically reconfiguring optical weight matrices—unlike electronic memories, which can be reprogrammed with relative ease and speed.

With recent advancements in silicon photonics and integrated optics, some of these longstanding obstacles are being addressed, sparking renewed interest in hybrid optical-electronic neural hardware for future applications. However, compared to architectures such as the Tianjic chip [2] or HUs [1], which emphasize flexibility, reconfigurability, and integration, hybrid ONNs remain largely confined to specialized, high-throughput domains—such as large-scale pattern matching, signal processing, or as experimental platforms for new hardware accelerators.

In summary, hybrid ONNs represent a significant milestone in the evolution of brain-inspired hybrid computing, illustrating both the potential and the inherent difficulties of cross-domain neural processing. These systems exemplify how leveraging distinct physical principles—using light for parallelism and electronics for logic and memory—can expand the capabilities of neural computation. Nevertheless, widespread adoption remains limited, with contemporary research focusing primarily on all-electronic or mixed-signal hardware for mainstream use, while optical technologies continue to play a vital role in the development of specialized accelerators and emerging quantum-inspired neural systems.

Taken together, these four methodologies illustrate the breadth of innovation in hybrid neural computation, each offering distinct advantages and facing specific challenges as the field moves toward more powerful and versatile brain-inspired systems.

IV. Discussion, Comparative Assessment, and Future Directions

This section synthesizes the central ideas from the various HNN methodologies discussed throughout the paper. This section addresses technical clarity, scalability, energy demands, and whether these approaches actually perform in real-world settings. Each method's core advantages and drawbacks are outlined, while persistent challenges are flagged, and some specific directions for future research in hybrid brain-inspired computing are suggested.

Each method offers advantages, but also presents considerable limitations. HUs offer solid modularity and are easy to reconfigure—which is a plus for dynamic applications. The Tianjic chip stands out for integrating ANNs and SNNs directly at the hardware level. Real-time, low-latency processing is possible, but design and implementation complexity is increased.

First Principles-HNNs deliver high interpretability and robust generalization—provided there are reliable equations to anchor the solution. When applied to a domain lacking clear physical models, their adaptability decreases substantially. Hybrid optical ONNs are impressive in terms of parallelism and energy efficiency—at least, in theory. In practice, most optical processing remains experimental, challenged by issues with alignment and scalability.

Overall, integration hurdles are significant. It is difficult to coordinate heterogeneous computational models, especially at the hardware-software boundary. There is also no universal framework or toolchain, so portability and reproducibility suffer. Scalability is another major challenge, whether you are running into chip interconnect bottlenecks, optical path limits, or algorithmic challenges with hybrid learning strategies. Interpretability, meanwhile, is still mostly limited to first-principles hybrids. Most current HNNs are heavily customized for narrow domains and nowhere near ready for broad, general-purpose AI deployment.

Despite these challenges, several overarching themes emerge. One is that the landscape is evolving rapidly: the boundaries between approaches are blurring, with hardware, algorithms, and applications influencing one another. For instance, while the Tianjic chip is a hardware-centric solution, its impact on the design of software frameworks and training algorithms is increasingly evident. Similarly, as optical and neuromorphic hardware advances, methods for efficient mapping of high-level hybrid architectures onto physical substrates are becoming a central research focus.

The trade-offs highlighted in Table I underscore the need for holistic evaluation metrics that go beyond traditional accuracy or throughput. Energy efficiency, adaptability, transparency, and hardware-compatibility are all critical in assessing the practical relevance of a given HNN methodology. This is especially true for applications at the edge, in embedded systems, and in mission-critical domains where resource constraints or interpretability requirements can outweigh raw performance.

Another emerging trend is the move towards end-to-end co-design. This involves simultaneous development of learning rules, neural architectures, and hardware—ensuring that each layer is optimized not in isolation, but as part of an integrated whole. Such co-design approaches have already demonstrated substantial improvements in both energy efficiency and system robustness, and are likely to shape future hybrid neural platforms.

Additionally, the integration of domain knowledge and explainability is growing in importance, especially for deployment in regulated or safety-critical sectors such as healthcare, finance, and autonomous vehicles. The adoption of explainable AI (XAI) tools, alongside formal methods for model verification and validation, will be essential for building trust in hybrid neural solutions.

A summary of these challenges and limitations for each HNN methodology can be found in Table I.

TABLE I

Visual summary of the severity of key research challenges for each HNN methodology. Darker color and higher number indicate a more critical issue for the corresponding approach.

	HUs	Tianjic Chip	First-Principles-HNN	Hybrid ONN
Integration Complexity	5	4	2	5
Scalability	3	5	2	2
Energy Efficiency	4	5	3	4
Interpretability	2	2	5	3
Application Generality	3	3	2	3

A. Perspectives and Future Research Directions

Despite persistent challenges, recent advances in neuromorphic hardware, algorithmic innovation, and interdisciplinary collaboration are accelerating the progress of HNNs. Building on both this review and recent analyses in the literature [7], several key research directions have emerged.

A primary objective for the community is the development of unified software and hardware platforms that enable seamless integration, training, and deployment of heterogeneous neural components. Achieving such unification will demand new interfaces, standardized protocols, and modular architectures capable of bridging the distinct requirements of ANNs and SNNs, as well as the domain-specific models and photonic devices found in the broader family of HNNs. In parallel, advances in hardware-algorithm co-design are needed to optimize not only for computational throughput, but also for flexibility, upgradability, and support of new learning paradigms, including online, continual, and meta-learning.

Another promising direction is the development of self-adaptive, context-aware, and task-driven HNNs, capable of autonomously modifying their computational structure in response to changing input statistics, environmental demands, or application-specific constraints. Such adaptive hybrids would be especially valuable in scenarios where uncertainty, novelty, or resource limitations require systems to dynamically trade off between energy use, precision, and real-time responsiveness. The realization of these capabilities will likely require joint innovation in network architecture, algorithmic plasticity, and learning rules that go beyond current supervised or reinforcement learning frameworks.

Improving the transparency, interpretability, and reliability of HNNs is also essential, especially in high-stakes, safety-critical, or regulated domains. Integration of explainable AI methods, as well as formal verification techniques, will be crucial for building trust and providing insight into how heterogeneous neural modules interact and make decisions. The incorporation of domain knowledge, physics-based constraints, or prior structural information can serve as valuable regularizers, potentially leading to more generalizable and robust hybrid architectures.

A further research priority is the explicit minimization of energy consumption and physical footprint, both at the algorithmic and hardware levels. This is particularly pressing for edge computing, IoT, and mobile robotics applications, where energy efficiency often determines the viability of real-world deployment. Efforts to develop training algorithms and hardware primitives that can operate in ultra-low-power regimes, possibly inspired by advances in event-driven, asynchronous, or spike-based computation, are likely to drive further breakthroughs in resource-constrained environments.

Equally important is the creation and dissemination of shared datasets, benchmarking protocols, and open-source toolkits tailored to HNNs. Such resources will enable fair and reproducible comparisons, foster community-driven innovation, and lower the barrier to entry for new researchers. Collaborative research consortia and interdisciplinary networks can play a pivotal role in standardizing best practices, distributing computational resources, and coordinating large-scale empirical studies.

Looking to the future, the scope of HNNs is likely to expand far beyond software and digital hardware. Emerging fields such as neuromorphic materials, memristive and photonic devices, and quantum-inspired computation will offer new substrates and operating principles for hybrid neural architectures. As the boundaries between computer engineering, neuroscience, physics, and materials science blur, cross-disciplinary research will become not only desirable but indispensable. It will also be critical to consider the societal, ethical, and economic implications of deploying advanced HNNs in domains ranging from healthcare and transportation to security and the environment.

Overall, this work provides a unified and critical perspective on the current state of hybrid neural computation, rigorously evaluating leading methodologies with respect to architecture, efficiency, interpretability, and practical viability. By systematically categorizing strengths and weaknesses and identifying focused research priorities, this review seeks to inform and catalyze future advances in scalable, robust, and transparent brain-inspired AI. HNNs hold the promise of bridging the gap between the speed and adaptability of artificial intelligence and the robustness and efficiency of biological systems. Achieving this vision will require continued technological innovation, the convergence of computational and physical sciences, and sustained dialogue between fundamental research and application-driven development. Ultimately, integrating insights from brain science, computer engineering, and emerging device technologies may enable the next leap forward in the realization of trustworthy, adaptive, and intelligent artificial agents.

V. Summary

This paper has systematically reviewed recent advances in HNNs, with particular emphasis on four key methodologies: HUs, the Tianjic chip architecture, first-principles-driven HNNs, and hybrid optical ONNs. Rather than merely cataloguing each strategy, the analysis has focused on technical transparency, scalability, energy efficiency, interpretability, and practical viability. The comparative evaluation highlights the substantial advantages gained by integrating elements from both artificial and biologically inspired neural paradigms, reinforcing the idea that hybridization is essential for overcoming the limitations inherent in single-method systems.

Collectively, HNNs distinguish themselves through their adaptability, energy efficiency, and robustness in the context of brain-inspired computing. By blending the principal features of ANNs, SNNs, domain-specific physical models, and optical computation, HNNs are equipped to address a wide spectrum of applications. These include not only traditional tasks such as robotics, process automation, signal processing, and pattern recognition, but also emerging domains requiring adaptability to changing environments and efficient resource utilization. This hybrid paradigm supports high performance under diverse operational demands, extending the reach of neural computation into scenarios that previously remained inaccessible due to constraints in efficiency or flexibility.

Despite these advances, several significant challenges remain unresolved. Integration complexity, particularly when coordinating heterogeneous computational models and bridging hardware-software boundaries, is an ongoing concern. The absence of standardized development frameworks, as well as a lack of widely accepted benchmarking protocols, continues to hamper reproducibility and fair comparison. Furthermore, scalability remains limited by both algorithmic bottlenecks and physical constraints at the hardware level, whether in chip interconnects or optical pathways. The interpretability of HNNs, while improved in models incorporating first-principles or explainable AI techniques, still lags behind the requirements for critical and regulated applications.

To address these challenges, the field will require sustained and coordinated progress across multiple fronts. Advancements in hardware-software co-design are essential for seamless integration of heterogeneous components. The creation of unified development environments and toolchains will foster portability and accelerate innovation. Research on energy-efficient learning algorithms, particularly those tailored for neuromorphic and edge devices, is likely to drive further improvements in real-world applicability. Equally important is the continued development of strategies for system validation, benchmarking, and interpretability, enabling HNNs to gain trust and regulatory approval in high-stakes environments.

Looking forward, the consolidation of strengths from multiple computational paradigms positions HNNs as a cornerstone for the next generation of intelligent systems. Their capacity to bridge the gap between artificial computation and biological efficiency will be instrumental in realizing advanced, scalable, and trustworthy artificial agents. Continued interdisciplinary progress—combining insights from computational neuroscience, machine learning, hardware engineering, and domain-specific sciences—will further expand the reach and impact of HNNs in both foundational research and practical technologies. As the landscape of brain-inspired AI continues to evolve, HNNs are set to play an increasingly central role in shaping the trajectory of artificial intelligence.

References

[1] R. Zhao, Z. Yang, H. Zheng, Y. Wu, F. Liu, Z. Wu, L. Li, F. Chen, S. Song, J. Zhu *et al.*, "A framework for the general design and computation of hybrid neural networks," *Nature communications*, vol. 13, no. 1, p. 3427, 2022.

[2] J. Pei, L. Deng, S. Song, M. Zhao, Y. Zhang, S. Wu, G. Wang, Z. Zou, Z. Wu, W. He *et al.*, "Towards artificial general intelligence with hybrid tianjic chip architecture," *Nature*, vol. 572, no. 7767, pp. 106–111, 2019.

[3] D. C. Psichogios and L. H. Ungar, "A hybrid neural network-first principles approach to process modeling," *AIChE Journal*, vol. 38, no. 10, pp. 1499–1511, 1992.

[4] S. Jutamulia and F. Yu, "Overview of hybrid optical neural networks," *Optics & Laser Technology*, vol. 28, no. 2, pp. 59–72, 1996.

[5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[6] M. Davies, N. Srinivasa, T.-H. Lin, G. Chinya, Y. Cao, S. Choday, G. Dimou, P. Joshi, N. Imam, S. Jain, S. Liao, H. Lin, A. Lines, R. Liu, D. Mathaikutty, S. McCoy, A. Paul, C.-K. Tse, G. Venkataramanan, Y. Weng, A. Wild, Y. Yang, and N. Yung, "Advancing neuromorphic computing with loihi: A survey of results and outlook," *Nature*, vol. 589, no. 7840, pp. 54–62, 2021.

[7] F. Liu, H. Zheng, S. Ma, W. Zhang, X. Liu, Y. Chua, L. Shi, and R. Zhao, "Advancing brain-inspired computing with hybrid neural networks," *National Science Review*, vol. 11, no. 5, p. nwae066, 2024.

[8] W. Zhang, B. Gao, J. Tang, P. Yao, S. Yu, M.-F. Chang, H.-J. Yoo, H. Qian, and H. Wu, "Neuro-inspired computing chips," *Nature electronics*, vol. 3, no. 7, pp. 371–382, 2020.