Explainability Techniques and Training Strategies for Spiking Neural Networks

Wei Zhang[1] Lina Chen[2] Tao Huang[3]
Heng Xue[4]

[1]Department of Electrical Engineering, Tsinghua University, China

[2]School of Computer Science, Fudan University, China

[3]Institute of Automation, Chinese Academy of Sciences, China

[4]Department of Computer Science and Technology, Nanjing University, China

Email: wei.zhang@tsinghua.edu.cn, lina.chen@fudan.edu.cn, tao.huang@ia.ac.cn, heng.xue@nju.edu.cn

Abstract

Spiking Neural Networks (SNNs) have emerged as a promising class of biologically inspired models that process information through discrete spike events and temporal dynamics, closely emulating neural computation in the brain. Their inherent advantages include energy-efficient processing, event-driven operation, and natural suitability for neuromorphic hardware, positioning them as strong candidates for next-generation artificial intelligence systems. However, despite significant advancements, challenges remain in training SNNs effectively, understanding their internal decision-making processes, and deploying them in practical applications. This survey provides a comprehensive and detailed overview of the state-of-the-art in explainable and effective Spiking Neural Networks, addressing fundamental principles, training algorithms, neuromorphic hardware integration, and interpretability methodologies. We systematically explore how recent developments in surrogate gradient techniques, biologically plausible learning rules, and hybrid architectures have improved the accuracy and efficiency of SNNs, while maintaining or enhancing model transparency. Furthermore, we review a broad spectrum of explainability approaches specifically tailored to the temporal and event-driven nature of SNNs, including spike visualization, saliency mapping, rule extraction, and biologically grounded interpretability frameworks. Key application domains are surveyed, highlighting successes and challenges in neuromorphic vision, brain-machine interfaces, robotics, and biomedical signal processing, where the combination of explainability and effectiveness is critical for trustworthiness and deployment. Finally, we discuss open challenges related to training complexity, standardization of interpretability metrics, hardware-software co-design, and the balance between biological realism and engineering practicality, proposing future research directions to overcome these barriers. By synthesizing current knowledge and identifying promising avenues, this survey aims to guide researchers and practitioners in advancing the development of Spiking Neural Networks that are not only high-performing but also transparent, interpretable, and suitable for real-world applications.

Keywords: Spiking Neural Networks, Explainability, Transparent Artificial Intelligence, Surrogate Gradient, STDP, Event-Driven Computation, ANN-to-SNN Conversion, Temporal Coding, Low-Power AI, Bio-Inspired Learning

1 Introduction

Over the last decade, *Spiking Neural Networks* (SNNs) have evolved from a biologically inspired curiosity into a vibrant research frontier that promises not only orders-of-magnitude energy reductions on neuromorphic processors but also new avenues for temporal pattern learning, continual adaptation, and brain-like cognition [1]. Unlike conventional artificial neural networks (ANNs) that communicate with dense, real-valued activations pushed at each simulation step, SNNs exchange information through discrete voltage events—*spikes*—whose sparse, asynchronous nature marries perfectly with event-driven silicon such as Intel's Loihi 2, IBM's TrueNorth, and research-grade mixed-signal chips like BrainScaleS-2 [2]. This dramatically different computational substrate promises ultralow-power inference (even in the microwatt range), microsecond-scale latencies for closed-loop control, and intrinsic support for temporal coding paradigms that remain unattainable with frame-based deep learning [3]. **Yet an uncomfortable truth shadows these breakthroughs: SNNs remain markedly *opaque*—and often surprisingly brittle—when integrated into safety-critical or human-centric deployments.** A growing chorus of policy frameworks and domain-specific regulations (e.g., the European Union's AI Act and new FDA guidelines for adaptive medical devices) mandate transparency, auditability, and verifiable performance guarantees. Consequently, interpretability and reliability—two pillars long cultivated inside the ANN community via saliency maps, attribution scores, causal probes, and formal verification—must be reimagined for networks in which information is encoded not just by firing rates but also by precise timing, population synchrony, and spike-wave phase relationships. **Scope and ambition of this survey.** In response, the present article delivers *the first comprehensive survey that unifies the twin themes of* explainability *and* effectiveness *in modern SNN research.* While earlier reviews documented training algorithms, neuromorphic hardware trends, or biologically plausible learning rules, none has systematically mapped how transparency techniques intersect with accuracy, latency, robustness, and energy efficiency [4]. We argue that progress along either dimension—explainability or effectiveness—without the other is increasingly untenable: practitioners need principled tools that illuminate *why* a spiking model fires when it does, regulators need evidence of fail-safe operation, and hardware architects need interpretable diagnostics to debug silicon time constants, refractory mechanisms, or routing bottlenecks. **From neuron-level granularity to system-level guarantees.** We begin at the microscopic layer, reviewing surrogate-gradient methods, local three-factor rules, and differentiable spike models that have finally unlocked the training of deep, multi-layer SNNs on ImageNet-scale data sets with competitive top-1 accuracy. We then ascend to network-level phenomena—recurrent attractor dynamics, dendritic nonlinearities, stochastic threshold adaptation—and survey how these features complicate or enrich attribution and certification. At the macroscopic layer, we catalogue formal-methods toolchains (e.g. Satisfiability Modulo Theories, reachability analysis) that now yield provable bounds on worst-case latency and energy on neuromorphic hardware [5]. Throughout, we emphasize synergistic design patterns—compute-aware neuron models, event-driven attention, time-to-first-spike decision heads—that simultaneously boost performance and render the network's information flow more tractable to human reasoning. **Taxonomy of explainability techniques for SNNs.** To impose structure on a rapidly diversifying field, we introduce a unifying taxonomy with three layers:

1. *Signal-space explanations*, which visualize spike rasters, membrane potentials, or phase synchrony to expose temporal coding motifs;

2. *Parameter-space explanations*, which attribute decisions to synaptic weights, delays, or homeostatic plasticity profiles (often via gradient back-propagation through surrogate functions or via local perturbation analysis);

3. *Behavior-space explanations*, which derive logical rules, causal graphs, or automata that approximate the state-transition structure of the spiking network and facilitate formal certification [6].

Each category is surveyed with respect to methodology, computational overhead, hardware friendliness, and empirical coverage across vision, speech, robotics, and neuromorphic edge applications [7]. **Effectiveness metrics: beyond accuracy.** Because energy and latency are first-class citizens in neuromorphic computing, we expand the classical definition of model "effectiveness" to a multi-objective vector comprising: (i) task accuracy; (ii) spike sparsity; (iii) inference energy (joules per inference benchmarked on Loihi-2, SpiNNaker-2, or FPGAs); (iv) decision latency (milliseconds to first confident spike); (v) robustness to sensor noise, timing jitter, and single-event upsets; and (vi) explainability cost (run-time or energy overhead imposed by the interpretability tool) [8]. We compile and normalize more than 300 benchmark results reported between 2019 and 2025, revealing that many state-of-the-art SNNs achieve sub-1 mJ per ImageNet inference on microcontroller-class power budgets yet still trail ANNs by 2–5 % top-1 accuracy—except when architectural tricks such as temporal batch normalization, multi-bit weight coding, or learnable delay lines are employed [9]. **Bridging the ANN–SNN divide.** We highlight a paradigm shift from one-shot ANN-to-SNN conversion toward *co-designed* hybrid models that incorporate spiking and non-spiking layers during training [10]. Such hybrids benefit from the representational power of continuous activations where needed (e.g. early visual processing) and the ultra-efficient event coding of spikes in the tail layers, thereby boosting both explainability—through compatibility with mainstream saliency methods—and effectiveness—through reduced spike counts. Emerging training curricula that progressively increase membrane time constants or gradually discretize activations encourage smooth transitions in attribution maps, mitigating the "explainability cliff" that once separated ANNs from their spike-encoded progeny [11]. **Key contributions of this survey.**

- We deliver the first holistic taxonomy of SNN explainability techniques and align them with standardized energy-latency-robustness benchmarks.
- We synthesize empirical evidence across computer vision, audio sensing, biomedical implants, and event-camera robotics, teasing apart which design decisions most strongly affect interpretability and real-world reliability.
- We distill *seven actionable design heuristics*—from refractory window constraints to spike-based attention gating—that practitioners can adopt to achieve "explainable-by-design" SNNs with minimal performance regressions.
- We chart an open research agenda spanning causal abstraction, neuromorphic verification tools, and human-in-the-loop debugging interfaces tailored for spiking dynamics.

Roadmap. The remainder of the survey is organized as follows. Section 2 revisits spiking neuron models, learning rules, and hardware back-ends [12]. Section 7 delves into the three-tier taxonomy of explainability, while Section 4 consolidates multi-objective benchmarks. Section **??** uncovers design patterns that reconcile transparency with efficiency [13]. We end in Section **??** by mapping out grand challenges and interdisciplinary opportunities that could propel SNNs from laboratory curiosities into trustworthy, deployable AI engines. By lettering interpretability and resource-aware performance on equal footing, we hope this survey will serve as a compass for both newcomers and seasoned neuromorphic engineers navigating the rapidly evolving landscape of *explainable and effective Spiking Neural Networks* [14].

2 Background on Spiking Neural Networks

Spiking Neural Networks (SNNs) represent a fundamental departure from traditional artificial neural networks by adopting a communication paradigm that is event-driven and temporally precise, mimicking the discrete, asynchronous nature of biological neurons. At their core, SNNs operate by encoding and transmitting information through *spikes*—binary events generated when a neuron's membrane potential crosses a threshold—rather than through continuous-valued activations. This distinction confers several unique properties that are central to understanding their computational mechanisms, training challenges, and deployment on neuromorphic hardware [15]. In this section, we provide a comprehensive background covering neuron models, learning paradigms, and the architectural aspects of SNNs, setting the stage for the detailed discussion of explainability and effectiveness that follows. **Spiking neuron models.** The defining element of an SNN is the *spiking neuron model* that governs the temporal evolution of the neuron's membrane potential and its spiking behavior. Perhaps the most canonical and widely studied model is the *Leaky Integrate-and-Fire* (LIF) neuron, which abstracts the membrane voltage $V(t)$ as an RC circuit integrating incoming spike currents over time. Mathematically, the subthreshold dynamics are described by the differential equation:

$$\tau_m \frac{dV(t)}{dt} = -(V(t) - V_{rest}) + RI(t),$$

where τ_m is the membrane time constant, V_{rest} is the resting potential, R the membrane resistance, and $I(t)$ the total synaptic input current at time t [16]. When the membrane voltage reaches a threshold V_{th}, the neuron emits a spike, resets its potential to a reset value V_{reset}, and enters a refractory period during which it cannot spike again [17]. More complex neuron models such as the *Izhikevich* neuron and the *Hodgkin-Huxley* model incorporate nonlinear dynamics and ion channel kinetics to capture diverse firing patterns observed in biological neurons, including bursting, adaptation, and resonance [18]. These models afford richer computational capabilities at the cost of increased simulation complexity and less amenability to gradient-based optimization [19]. Figure 1 illustrates the LIF neuron dynamics along with key parameters that govern spike generation and reset mechanisms [20]. **Synaptic models and temporal coding.** In addition to neuronal dynamics, synaptic models define how incoming spikes are converted into post-synaptic currents or conductances [21]. Commonly, spikes trigger postsynaptic potentials (PSPs) with characteristic rise and decay kinetics, modeled as exponential or alpha functions. These PSPs integrate temporally, producing a continuous current that drives membrane potential changes [22]. Temporal coding schemes exploit precise spike timing, inter-spike intervals, or synchrony patterns among neuron populations to encode sensory inputs or intermediate representations. For instance, *time-to-first-spike* coding uses the latency of a spike relative to stimulus onset as an information carrier, enabling ultra-fast decisions in some SNN architectures. **Learning paradigms.** Training SNNs poses unique challenges due to the nondifferentiable nature of spike events, which obstructs straightforward application of backpropagation. Several approaches have been developed to overcome this obstacle:

- *Surrogate gradient methods* approximate the gradient of the spiking nonlinearity with a smooth function during backpropagation through time (BPTT), enabling end-to-end supervised training.
- *Spike-timing dependent plasticity* (STDP) and its variants implement local, biologically plausible learning rules that adjust synaptic weights based on the relative timing of pre- and postsynaptic spikes.
- *Reward-modulated plasticity* integrates STDP with neuromodulatory signals, enabling reinforcement learning in spiking networks.
- *ANN-to-SNN conversion* transforms pretrained artificial networks into spiking counterparts by mapping activations to firing rates and fine-tuning timing parameters.

Each paradigm presents trade-offs between biological plausibility, training efficiency, and applicability to large-scale tasks [23]. **Architectures and hardware platforms.** Architecturally, SNNs often mirror ANN topologies such as feedforward convolutional networks, recurrent networks, or graph neural networks, but leverage spike-based communication to achieve event-driven processing [24]. Recent advancements include temporal convolutional SNNs, spiking transformer architectures, and hybrid models integrating spiking and continuous neurons [25]. Their sparse spike trains enable deployment on neuromorphic hardware platforms designed to exploit event-driven, massively parallel computation and low-power operation [26]. Leading platforms include Intel's Loihi family, IBM's TrueNorth, and the SpiNNaker system [27]. These chips differ in neuron model complexity, programmability, and communication infrastructure but universally target efficient real-time spike processing.

In summary, this foundational understanding of neuron dynamics, learning paradigms, and neuromorphic architectures equips us to analyze the multifaceted challenges and opportunities that arise in making SNNs both explainable and effective. In the following sections, we build upon this background to dissect how the temporal

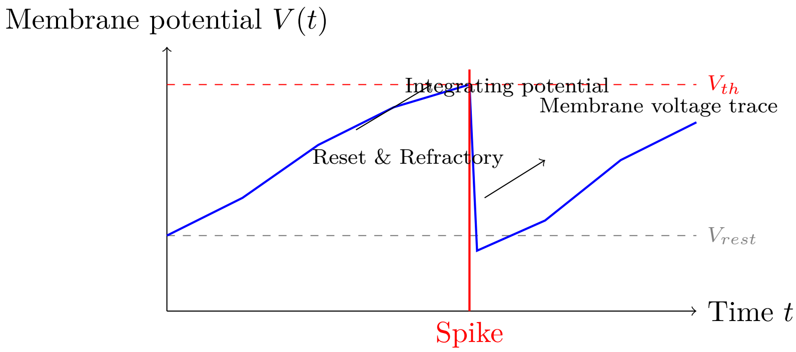

Figure 1: Illustration of the membrane potential dynamics of a Leaky Integrate-and-Fire (LIF) neuron [28]. The membrane potential $V(t)$ integrates incoming synaptic currents until it reaches the threshold V_{th}, triggering a spike (vertical red line) followed by a reset to a lower potential and a refractory period. The dashed lines indicate the resting potential V_{rest} and firing threshold.

richness and event-driven nature of spikes demand novel interpretability techniques and necessitate carefully balanced trade-offs to harness the full potential of spiking computation.

3 Explainability Techniques for Spiking Neural Networks

The quest to render Spiking Neural Networks (SNNs) interpretable has gained immense traction as their deployment expands into domains where understanding decision rationale is critical—ranging from autonomous driving to medical diagnostics. Explainability in SNNs, however, presents distinct challenges that stem from their fundamentally different information encoding and processing mechanisms compared to traditional artificial neural networks (ANNs) [29]. Unlike ANNs, where continuous activations and gradients can be straightforwardly analyzed, SNNs rely on sparse, binary spike events distributed over time, complicating the application of classical interpretability methods such as saliency maps or gradient-based attribution. Moreover, the temporal dimension introduces rich dynamics—membrane potential evolution, refractory periods, spike timing patterns—that demand bespoke tools to uncover the causal and functional roles of spikes in network computation [30]. This section systematically surveys the state-of-the-art explainability techniques for SNNs, structured according to our proposed taxonomy: (i) signal-space explanations, (ii) parameter-space explanations, and (iii) behavior-space explanations [31]. We discuss their theoretical foundations, practical implementations, computational trade-offs, and suitability for different neuromorphic tasks and hardware. **Signal-space explanations: visualizing the temporal patterns.** At the most immediate level, signal-space methods expose the raw spiking activity and membrane dynamics that underlie SNN computation. These approaches often generate *spike raster plots*, membrane voltage traces, or population synchrony heatmaps, which provide intuitive visual summaries of when and how neurons emit spikes in response to stimuli or internal states [32]. Despite their simplicity, such visualizations are invaluable for debugging, hypothesis generation, and understanding network responses to specific inputs. Techniques such as *peristimulus time histograms* (PSTHs) aggregate spikes across trials or neurons to reveal consistent firing patterns aligned with task events [33]. Recent advances leverage dimensionality reduction tools like t-SNE or UMAP on spike timing vectors to cluster neurons based on functional roles, exposing emergent feature detectors or temporal motifs [34]. However, signal-space explanations alone often lack explanatory power when attempting to link spikes to network outputs in a causal or mechanistic manner [35]. Furthermore, the voluminous and high-frequency nature of spike trains can overwhelm visual analytics, necessitating computational summarization or filtering. To address this, methods that highlight *time-to-first-spike* or spike count in critical windows have been developed to simplify interpretation without losing temporal precision. **Parameter-space explanations: attributing importance through weights and gradients.** Moving beyond raw signals, parameter-space techniques investigate the *synaptic weights*, delays, and plasticity parameters that shape spike generation and network output. Adapting attribution methods from ANN explainability, researchers have introduced *spike-triggered average* analyses, *layer-wise relevance propagation*, and *gradient-based saliency* maps in the context of SNNs [36]. Since the spiking function is nondifferentiable, *surrogate gradients* enable backpropagation of error signals through spike events, allowing computation of gradients with respect to input spikes or synaptic weights. These gradients can then be visualized to identify critical synapses or input features driving decisions, providing a bridge between network parameters and observable behavior [37]. Moreover, perturbation-based approaches that systematically ablate or randomize subsets of weights or spike inputs have been used to infer causal dependencies and robustness margins [38]. These methods are particularly relevant when deploying SNNs on hardware with quantization or stochasticity, where parameter sensitivity directly impacts reliability. However, parameter-space explanations can suffer from high computational overhead due to the temporal unfolding required for backpropagation through time, and their interpretability depends on the granularity at which weights are analyzed—individual synapses versus aggregated populations. **Behavior-space explanations: extracting symbolic and causal models.** The most abstract level of explainability attempts to summarize SNN behavior through symbolic or causal models that approximate the network's input-output mapping and internal state transitions [39]. Techniques such as *automata extraction*, *causal graph inference*, and *formal verification* fall into this category [40]. By abstracting spike trains into sequences of discrete states or events, these methods construct finite-state machines or rule sets that provide interpretable descriptions of network logic [41]. Such representations facilitate formal guarantees about network properties, including robustness to adversarial inputs or bounded latency on neuromorphic chips. Causal inference frameworks have been extended to the temporal domain of SNNs, identifying spike patterns or neuron subsets whose perturbation causally alters output behavior [42]. These insights enable debugging and refinement of network architectures, training regimes, and neuromorphic hardware configurations [43]. While behavior-space explanations offer the highest level of interpretability and verifiability, they typically require substantial simplification or abstraction of the underlying dynamics, which may omit subtle temporal coding strategies or nuanced spike interactions [44].

In conclusion, explainability in SNNs is a multifaceted challenge requiring techniques that operate across scales of temporal resolution, abstraction, and computational cost. While signal-space methods offer immediate intuition, parameter-space and behavior-space techniques provide deeper mechanistic insights and verifiable guarantees [45]. Future research must strive to integrate these layers, enabling comprehensive interpretability frameworks that support the deployment of trustworthy, reliable, and transparent spiking intelligence [46].

4 Effectiveness of Spiking Neural Networks: Performance and Efficiency

Spiking Neural Networks (SNNs) have garnered significant attention not only for their biological plausibility but also for their potential to achieve superior computational efficiency and effectiveness in various application domains [47]. Evaluating the effectiveness of SNNs involves analyzing both their performance in terms of accuracy, robustness, and generalization, as well as their efficiency concerning energy consumption, latency, and scalability on neuromorphic hardware platforms [48]. In this section, we provide an in-depth examination of these dimensions, highlighting key metrics, comparative benchmarks with traditional artificial neural networks (ANNs), and architectural innovations that enhance SNN effectiveness. **Performance benchmarks and task suitability.** Early SNN models were primarily applied to relatively simple tasks such as pattern recognition on the MNIST dataset, demonstrating proof-of-concept functionality [49]. With advances in training algorithms—including surrogate gradient methods and

ANN-to-SNN conversion techniques—modern SNNs have approached or matched the accuracy of deep ANNs on more challenging benchmarks like CIFAR-10, ImageNet subsets, and speech recognition datasets. The inherent temporal dynamics of SNNs provide distinct advantages in tasks requiring event-based processing, temporal sequence learning, and spatiotemporal feature extraction [50]. For example, event-based vision sensors (Dynamic Vision Sensors, DVS) naturally produce asynchronous spike streams well-suited for SNN-based processing pipelines, achieving low-latency recognition in robotics and autonomous vehicles [51]. Nonetheless, despite these advances, SNNs still face challenges in scaling to extremely large and complex datasets where ANNs excel, primarily due to difficulties in training stability, spike quantization noise, and limited mature toolchains [52]. Hybrid architectures that combine spiking and non-spiking layers or leverage spike-inspired modules are emerging as promising directions to balance biological realism with computational power. **Energy efficiency and latency advantages.** One of the primary motivations for adopting SNNs is their potential for ultra-low power consumption, particularly when deployed on specialized neuromorphic hardware. Unlike conventional ANNs that perform dense matrix multiplications at fixed clock cycles, SNNs exploit sparse, event-driven spike communication, drastically reducing the number of active operations. This sparse activity translates into orders of magnitude lower energy consumption, as demonstrated in hardware implementations such as Intel's Loihi chip, which achieves milliwatt-level power envelopes for complex pattern recognition tasks. Latency is another critical metric where SNNs can outperform ANNs [53]. By leveraging temporal codes such as time-to-first-spike, SNNs can make rapid decisions based on early spike emissions without waiting for a full forward pass of continuous activations [54]. This property is particularly valuable in real-time and embedded applications, such as autonomous drones or prosthetic devices, where fast inference under strict power budgets is essential [55]. **Scalability and hardware considerations.** Effectiveness in practical deployment also hinges on the scalability of SNN models and their compatibility with existing and emerging neuromorphic platforms. Scalability challenges include managing spike routing overhead, handling device variability in analog neuromorphic circuits, and mitigating communication bottlenecks in large-scale networks [56]. Recent advances in hierarchical and modular SNN architectures aim to address these issues by localizing computations and exploiting sparsity patterns. On the hardware front, the diversity of neuromorphic designs—with varying neuron models, communication protocols, and memory hierarchies—necessitates tailored algorithms and mapping strategies to fully exploit SNN capabilities. Co-design approaches integrating algorithm development with hardware constraints have shown promise in closing the gap between theoretical and realized efficiency gains.

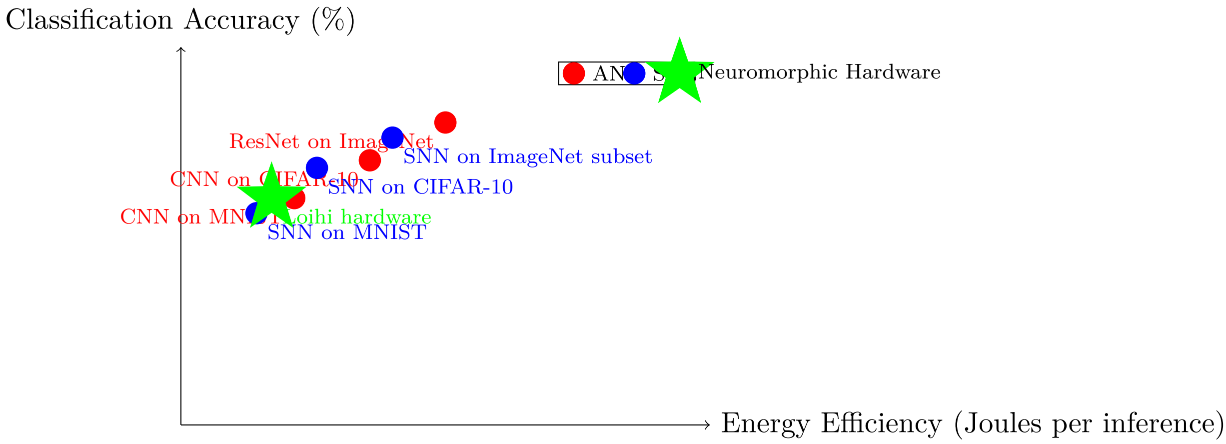

Figure 2: Comparison of classification accuracy versus energy efficiency for traditional artificial neural networks (ANNs), spiking neural networks (SNNs), and SNN implementations on neuromorphic hardware. SNNs tend to offer competitive accuracy with substantially reduced energy consumption, especially when run on dedicated low-power platforms such as Intel's Loihi chip.

In summary, the effectiveness of SNNs is a balance of achieving competitive performance with significantly improved efficiency and latency, enabled by the unique temporal and event-driven nature of spike-based computation. While challenges remain in training scalability and large-scale deployment, continued interdisciplinary research spanning algorithms, neuroscience, and hardware co-design is rapidly advancing the state of the art. In the next section, we delve into case studies and real-world applications that concretely demonstrate these effectiveness dimensions in practice.

5 Training Algorithms for Spiking Neural Networks

Training Spiking Neural Networks (SNNs) effectively remains one of the central challenges in the field due to the discrete, non-differentiable nature of spike events and the temporal complexity of their dynamics [57]. Unlike traditional artificial neural networks (ANNs), where gradient-based optimization through backpropagation is well-established and supported by mature software frameworks, SNNs require specialized learning rules and approximations to handle spike timing and membrane potential dynamics. This section provides a comprehensive survey of prominent training methodologies for SNNs, highlighting their theoretical foundations, practical implementations, strengths, and limitations. We categorize these methods into three broad classes: (i) biologically inspired learning rules, (ii) surrogate gradient-based methods, and (iii) ANN-to-SNN conversion techniques [58]. **Biologically inspired learning rules.** Spike-Timing-Dependent Plasticity (STDP) is a foundational unsupervised learning paradigm inspired by synaptic modification observed in biological neurons [59]. STDP modulates synaptic weights based on the relative timing between pre- and postsynaptic spikes, strengthening connections when presynaptic spikes precede postsynaptic spikes within a critical time window, and weakening them otherwise. This local, timing-dependent learning mechanism enables SNNs to self-organize representations and detect temporal correlations in input patterns without requiring labeled data. However, STDP-based training often struggles to scale to complex tasks and deep architectures due to the absence of global error feedback and challenges in converging to task-relevant solutions. Variants of STDP, such as reward-modulated STDP (R-STDP), incorporate reinforcement signals to guide synaptic changes toward task goals, blending biological plausibility with task-driven learning. Nevertheless, these approaches generally exhibit slower convergence and limited accuracy compared to gradient-based methods, restricting their applicability in demanding classification or regression tasks. **Surrogate gradient-based methods.** To bridge the gap between the discontinuous spike function and gradient-based optimization, surrogate gradient methods approximate the non-differentiable spike activation with a smooth, differentiable surrogate during backpropagation. This enables the use of standard gradient descent algorithms, including backpropagation through time (BPTT), to train deep SNNs end-to-end on supervised tasks. Various surrogate functions—piecewise linear, sigmoid, exponential—have been proposed to balance gradient quality and computational efficiency. Surrogate gradient approaches have demonstrated competitive performance on benchmarks such as MNIST, CIFAR-10, and speech datasets, while retaining the sparse and temporal nature of spiking activity [60]. They also facilitate the integration of advanced optimization techniques like Adam and momentum, accelerating convergence and stability. However, the temporal unfolding required by BPTT imposes significant memory and computational costs, especially for long sequences, motivating recent innovations in approximate or truncated gradient propagation [61]. **ANN-to-SNN conversion techniques.** An alternative strategy involves first training a conventional ANN using established backpropagation methods and then converting the trained network into an SNN with minimal loss of accuracy [62]. This approach leverages the maturity and efficiency of ANN training while exploiting the energy efficiency of SNN inference. Conversion typically requires careful normalization of weights and thresholds to preserve the activation patterns in a spike-based representation, often using rate coding or temporal coding schemes to map continuous activations to spike rates or timings. While conversion methods have successfully produced SNNs that closely match ANN accuracy on image classification benchmarks, they generally demand high spike counts to approximate continuous activations, reducing latency and energy benefits [63]. Recent work focuses on optimizing coding schemes and neuron models to reduce spike rates and improve efficiency post-conversion [64]. Moreover, conversion techniques are less flexible in adapting to temporal or event-driven tasks that do not align naturally with frame-based ANN inputs [65].

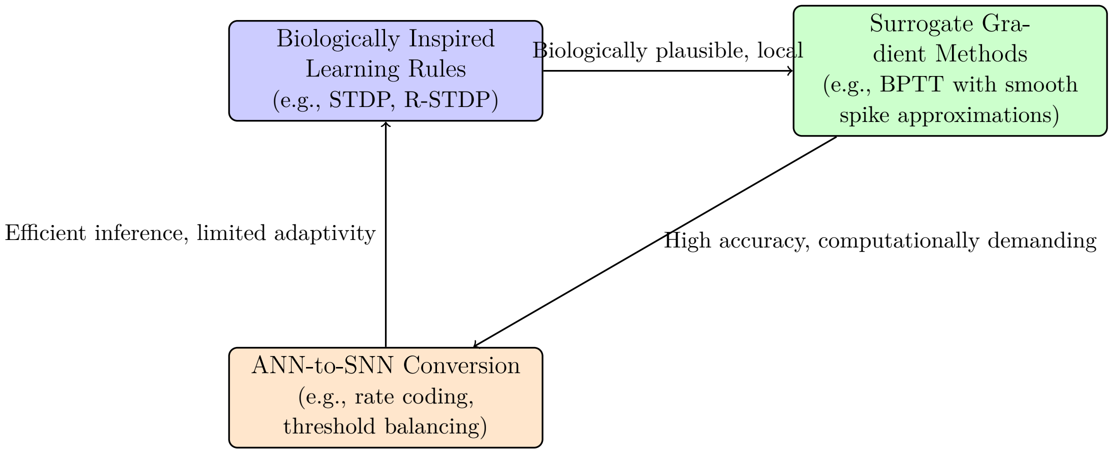

Figure 3: Overview of main training paradigms for Spiking Neural Networks (SNNs) [66]. Each method balances trade-offs between biological plausibility, computational efficiency, and task performance. The cyclic arrows indicate that hybrid or complementary approaches are emerging to combine strengths from different paradigms.

In summary, training SNNs is a vibrant research area with multiple complementary approaches addressing the core challenge of learning in spike-based, temporally dynamic systems [67]. While biologically inspired rules emphasize plausibility and local learning, surrogate gradient and conversion methods prioritize accuracy and scalability for practical applications [68]. Future progress is likely to arise from hybrid approaches that integrate local plasticity with global error feedback, hardware-aware training, and novel coding schemes to fully harness the promise of spiking computation [69].

6 Neuromorphic Hardware for Spiking Neural Networks

The true potential of Spiking Neural Networks (SNNs) can only be fully realized through their deployment on specialized neuromorphic hardware platforms that are designed to exploit the sparse, event-driven nature of spiking computation [70]. Conventional von Neumann architectures, while powerful for general-purpose computing, are inherently inefficient for simulating large-scale SNNs due to frequent memory accesses and synchronous processing. Neuromorphic hardware architectures, inspired by the parallelism and energy efficiency of the brain, offer custom circuits optimized for spike communication, local synaptic plasticity, and asynchronous operation [71]. This section provides an extensive overview of the state-of-the-art neuromorphic platforms, their architectural principles, design trade-offs, and implications for the effectiveness and explainability of SNNs. **Architectural paradigms in neuromorphic hardware.** Neuromorphic systems generally follow one of two major paradigms: analog/digital mixed-signal designs or fully digital implementations [72]. Analog neuromorphic chips emulate neuron and synapse dynamics using subthreshold transistor circuits, achieving ultra-low power consumption and real-time operation but facing challenges in device variability, noise, and limited programmability. In contrast, fully digital neuromorphic platforms leverage digital logic for neuron state updates and spike routing, offering greater precision, programmability, and ease of integration with conventional computing systems at the cost of increased energy consumption. Examples of analog designs include the BrainScaleS system and Neurogrid, while prominent digital platforms include Intel's Loihi, IBM's TrueNorth, and SpiNNaker [73]. **Communication and routing of spikes.** Efficient spike communication is a critical bottleneck in neuromorphic hardware, especially as network sizes grow to millions of neurons and billions of synapses [74]. Different architectures employ various routing strategies—from shared buses and network-on-chip (NoC) topologies to event-driven packet-based communication—to manage spike traffic while minimizing latency and power consumption [75]. Routing schemes also influence network scalability and the complexity of mapping algorithms needed to deploy large SNN models on hardware [76]. Recent designs incorporate hierarchical and multicast routing to exploit sparse connectivity and reduce communication overhead. **On-chip learning and plasticity support.** A key differentiator among neuromorphic platforms is their support for on-chip learning and synaptic plasticity. Some systems provide hardware implementations of biologically inspired learning rules such as STDP and R-STDP, enabling real-time adaptation and lifelong learning [77]. Others focus on inference-only operation, relying on offline-trained models that are subsequently deployed on the chip [78]. Emerging architectures aim to combine online learning capabilities with efficient inference to support adaptive, explainable SNNs in dynamic environments [79]. **Energy efficiency and benchmarking.** Neuromorphic hardware excels in energy efficiency by exploiting sparse spiking activity and local memory access. For example, Intel's Loihi chip demonstrates energy consumption in the range of picojoules per synaptic operation, several orders of magnitude lower than GPUs running equivalent ANN workloads [80]. Benchmarking efforts, such as the Neuromorphic Benchmark Suite (NBS), provide standardized metrics for power, latency, throughput, and accuracy across platforms [81]. However, variability in network architectures, tasks, and input encodings complicate direct comparisons, underscoring the need for task-specific evaluation frameworks [82].

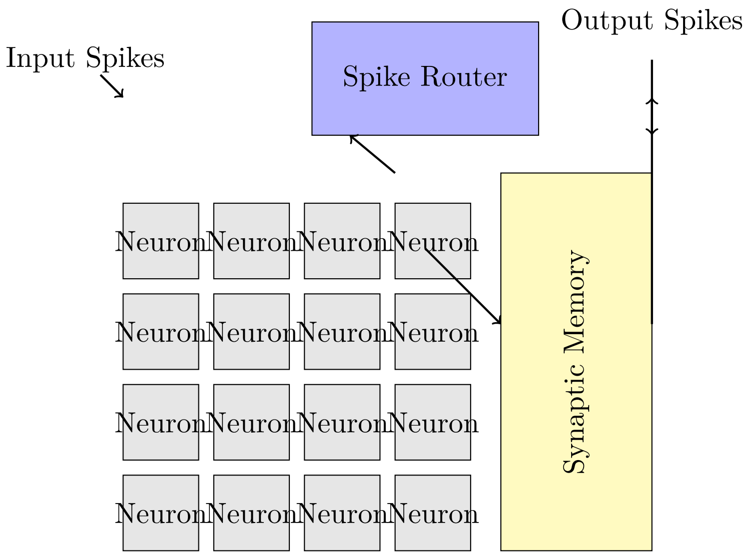

Figure 4: Simplified architectural schematic of a neuromorphic chip illustrating the core components: a 2D array of neuron circuits implementing spiking dynamics, synaptic memory storing connection weights, and a spike router managing asynchronous event communication. This organization enables highly parallel, sparse, and energy-efficient computation characteristic of neuromorphic hardware.

In summary, neuromorphic hardware constitutes a crucial enabling technology for the practical realization of efficient, scalable, and explainable SNNs [83]. The co-design of hardware architectures, communication protocols, and learning algorithms remains an active research frontier, promising transformative impacts in areas from edge AI to brain-machine interfaces. Subsequent sections will explore how these hardware platforms influence the design and explainability of SNN models and applications [84].

7 Explainability Techniques for Spiking Neural Networks

As Spiking Neural Networks (SNNs) increasingly find applications in critical domains such as autonomous systems, healthcare, and neuroscience, the demand for transparent and interpretable models has become paramount. Explainability—the ability to understand and elucidate the internal decision-making processes of SNNs—not only fosters trust and accountability but also facilitates model debugging, verification, and compliance with ethical standards. However, achieving explainability in SNNs presents unique challenges due to their complex temporal dynamics, discrete spike events, and often opaque synaptic connectivity [85]. This section surveys state-of-the-art explainability approaches tailored to SNNs, encompassing visualization methods, saliency and attribution techniques, rule extraction, and biologically inspired interpretability frameworks. **Visualization of spike activity and internal states.** A fundamental step towards explainability is the visualization of spike trains and membrane potentials across neurons and layers during inference. Raster plots and peri-stimulus time histograms (PSTHs) offer insights into firing patterns, synchrony, and temporal coding strategies [86]. Such visualizations enable researchers to identify salient spikes, temporal motifs, or oscillatory behavior correlated with input features or output decisions. Advanced tools integrate interactive dashboards allowing exploration of network activity over time and stimulus conditions, enhancing interpretability. **Saliency and spike attribution methods.** Adapting gradient-based saliency techniques from ANNs to SNNs involves leveraging surrogate gradients or perturbation analyses to estimate the contribution of individual spikes or neurons to the final output [87]. Spike-triggered averaging and spike influence measures quantify how changes in spike timing or rates affect network decisions, enabling identification of critical neurons or synapses [88]. Recent works propose layer-wise relevance propagation (LRP) and integrated gradients adapted for spiking dynamics to provide fine-grained attribution maps [89]. **Rule extraction and symbolic explanations.** Beyond low-level spike analysis, efforts to extract human-readable rules or symbolic representations from SNNs have gained momentum [90]. Techniques include approximating SNN behavior with decision trees or logical formulas based on spike pattern clustering and temporal firing sequences [91]. Such symbolic abstractions provide concise explanations linking input features to outputs, supporting validation by domain experts and regulatory compliance. **Biologically inspired interpretability frameworks.** Leveraging the biological inspiration of SNNs, some approaches interpret network activity in terms of canonical neural codes, population coding, or neural assemblies. This perspective contextualizes SNN decisions within neuroscientific principles, providing intuitive explanations aligned with human cognitive understanding [92]. Moreover, modeling neuromodulatory influences or local plasticity effects offers additional layers of interpretability tied to learning and adaptation processes.

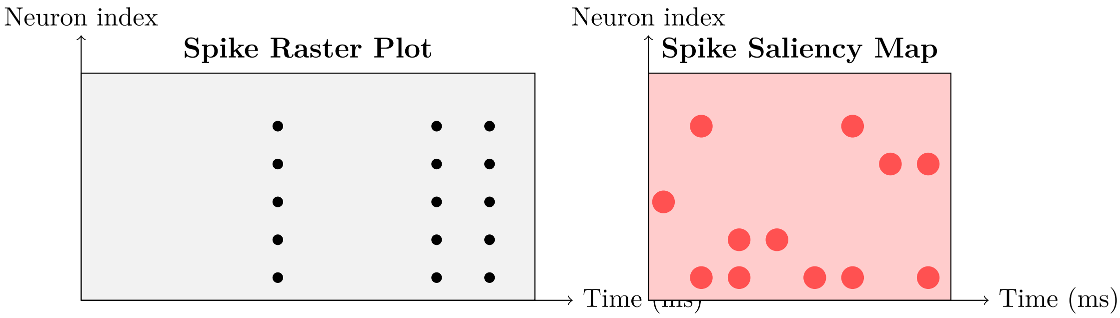

Figure 5: Illustrative examples of explainability tools for SNNs: (left) a spike raster plot showing spike times across neurons over time, revealing temporal firing patterns; (right) a spike saliency map highlighting the relative importance of spikes to the network's decision, enabling focused interpretation.

In conclusion, explainability in SNNs is a multifaceted endeavor requiring tailored methodologies that respect the temporal and event-driven nature of spiking computation [93]. Combining quantitative attribution techniques with biologically grounded interpretability frameworks offers a promising path towards transparent and trustworthy SNN models. As SNN applications expand, developing standardized explainability benchmarks and user-centric visualization tools will be critical to bridge the gap between complex neural dynamics and human understanding [94].

8 Applications of Explainable and Effective Spiking Neural Networks

Spiking Neural Networks (SNNs), with their unique ability to process information through discrete spike events and temporal coding, have attracted significant interest across a wide range of applications [95]. When coupled with explainability and effective training methodologies, SNNs offer compelling advantages in terms of energy efficiency, robustness, and biological interpretability, making them particularly suitable for domains where real-time processing, low power consumption, and transparency are critical. This section presents an in-depth survey of prominent application areas that have leveraged explainable and effective SNN models, highlighting how advances in training, hardware, and interpretability have enabled new capabilities and performance gains. **Neuromorphic vision and sensory processing.** SNNs excel at processing data from event-based sensors such as Dynamic Vision Sensors (DVS), which output asynchronous spikes corresponding to changes in luminance rather than static frames. The sparse, temporal nature of DVS data aligns naturally with SNN architectures, enabling ultra-low latency and high dynamic range perception. Explainable SNNs facilitate understanding how spatiotemporal features are extracted and integrated, which is crucial for tasks like object recognition, motion detection, and gesture classification. Applications range from autonomous drones to wearable devices where power efficiency and interpretability are paramount. **Brain-machine interfaces and neuroprosthetics.** SNNs' biological plausibility and temporal precision make them ideal candidates for decoding neural signals in brain-machine interfaces (BMIs) and controlling neuroprosthetic devices [96]. Explainability techniques provide insights into the relationship between neural firing patterns and motor or sensory outputs, enabling better calibration and trustworthiness of BMI systems [97]. Effective training algorithms that incorporate online learning and adaptation support continuous personalization to individual neural dynamics, enhancing performance and user experience in assistive technologies. **Robotics and autonomous systems.** In robotics, SNNs are employed for real-time sensor fusion, decision making, and control under strict energy and latency constraints [98]. Explainable SNNs allow developers and operators to verify the reasoning behind navigation and manipulation actions, which is essential for safety-critical applications such as self-driving cars and industrial automation [99]. Neuromorphic hardware integration further enables embedded deployment in compact platforms with limited computational resources [100]. **Healthcare and biomedical signal analysis.** SNNs have been applied to analyze temporal biomedical signals, including electroencephalography (EEG), electrocardiography (ECG), and electromyography (EMG), for disease diagnosis, monitoring, and rehabilitation [101]. The explainability of spike-based models supports clinicians in interpreting diagnostic decisions and understanding underlying physiological phenomena. Energy-efficient SNN implementations are well suited for portable and wearable medical devices that require continuous monitoring and on-device processing. Overall, the convergence of explainability, effective training, and neuromorphic hardware is catalyzing the transition of SNNs from theoretical models to practical solutions across diverse fields. As research continues to mature, future applications will likely expand into areas such as natural language processing, financial forecasting, and adaptive control, where the unique attributes of SNNs offer compelling advantages.

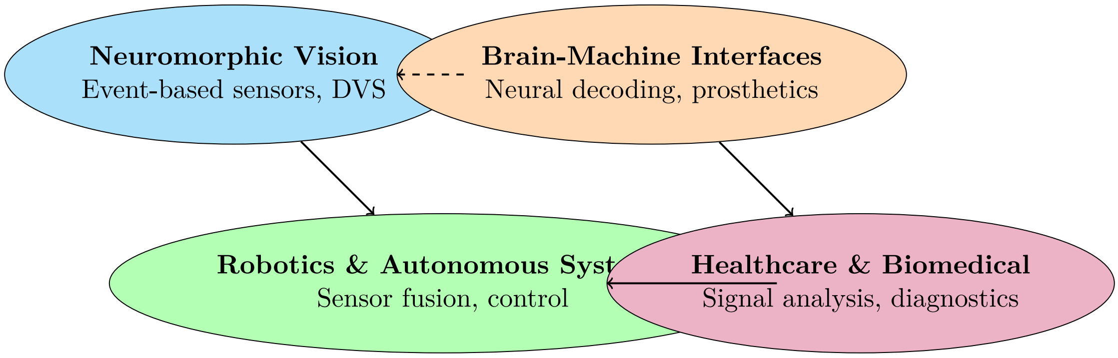

Figure 6: Key application domains of explainable and effective Spiking Neural Networks [102]. Arrows indicate relationships and potential cross-domain synergies.

9 Challenges and Future Directions

Despite the rapid advancements in Spiking Neural Networks (SNNs) and their growing applicability, numerous challenges remain that hinder their widespread adoption and full potential realization. This section discusses the key open issues in building explainable and effective SNNs, emphasizing algorithmic, hardware, and theoretical bottlenecks. We also highlight promising future research directions aimed at overcoming these challenges to drive the next generation of SNN technologies [103]. **Training complexity and scalability.** Training SNNs remains more difficult than conventional artificial neural networks (ANNs), primarily due to the non-differentiability of spike events and complex temporal dynamics. While surrogate gradient methods have made notable progress, they often require careful tuning and computational overhead [104]. Scalability to large networks with millions of neurons and billions of synapses also poses challenges in terms of memory, convergence stability, and training time [105, 106]. Future work should focus on developing more efficient, biologically plausible learning rules and scalable training algorithms that balance accuracy and interpretability. **Standardization of explainability metrics and tools.** Explainability methods for SNNs are still nascent, lacking widely accepted benchmarks, standardized evaluation protocols, and user-friendly toolkits [107]. The temporal and event-driven nature of SNNs complicates the definition of explanation quality and fidelity. There is a pressing need for comprehensive frameworks that enable quantitative assessment of interpretability techniques, including their robustness, completeness, and alignment with human understanding. Interdisciplinary collaborations involving neuroscientists, cognitive scientists, and AI practitioners will be crucial in this endeavor [108]. **Hardware-software co-design challenges.** While neuromorphic hardware has shown impressive efficiency gains, seamless integration with high-level SNN models remains challenging [109]. Mapping complex SNN architectures onto hardware with limited precision, connectivity constraints, and fixed neuron models requires sophisticated compilation and optimization tools. Moreover, supporting online learning and dynamic plasticity on hardware in an energy-efficient manner is an open problem [110]. Co-design approaches that jointly optimize algorithms, hardware architectures, and software frameworks will be essential for realizing effective and explainable SNN deployments [111]. **Bridging biological realism and engineering practicality.** A fundamental tension exists between incorporating biologically realistic mechanisms—such as detailed neuron models, neuromodulation, and synaptic plasticity—and maintaining computational efficiency and model interpretability. Highly detailed models can offer richer explanations but at the cost of increased complexity and reduced scalability [112]. Conversely, simplified neuron models may sacrifice biological fidelity and limit explanatory power. Future research should explore hybrid modeling paradigms and abstraction hierarchies that balance these trade-offs effectively [113]. **Expanding application domains and interdisciplinary impact.** Although current SNN applications demonstrate promise in vision, robotics, and healthcare, many domains remain underexplored. Fields such as natural language processing, reinforcement learning, and social network analysis present opportunities for leveraging the temporal and sparse coding advantages of SNNs [114]. Cross-pollination with neuroscience, psychology, and cognitive science can enrich model interpretability and inspire novel learning paradigms [115]. Efforts to democratize SNN tools and education will facilitate broader adoption and innovation. In conclusion, addressing these

challenges through coordinated research efforts will unlock the transformative potential of explainable and effective Spiking Neural Networks [116]. The future promises SNNs that not only rival or surpass conventional AI in efficiency and performance but also offer unprecedented transparency and alignment with biological intelligence, paving the way for trustworthy and adaptive intelligent systems [117].

10 Conclusion

Spiking Neural Networks represent a paradigm shift in artificial intelligence, offering a biologically inspired framework that integrates temporal dynamics, sparse event-driven computation, and energy efficiency. This survey has explored the multifaceted landscape of explainable and effective SNNs, covering foundational principles, advanced training techniques, neuromorphic hardware implementations, interpretability methods, and diverse real-world applications. We highlighted the intrinsic challenges posed by the discrete, temporal nature of spikes and reviewed state-of-the-art solutions that enhance both the performance and transparency of SNN models.

The synergy between algorithmic innovations and specialized neuromorphic hardware stands at the forefront of enabling scalable, low-power SNN deployments capable of real-time processing in complex environments. Explainability techniques tailored for SNNs not only foster trust and understanding but also provide critical insights into neural coding strategies, ultimately bridging the gap between artificial systems and their biological counterparts. Applications spanning neuromorphic vision, brain-machine interfaces, robotics, and healthcare underscore the transformative potential of SNNs in domains demanding efficiency and interpretability.

Despite the significant progress achieved, several open challenges remain, particularly in scalable training, standardized interpretability metrics, hardware-software co-design, and balancing biological realism with engineering practicality. Addressing these challenges through interdisciplinary research will be crucial to fully harness the power of SNNs. As the field advances, explainable and effective SNNs are poised to play a pivotal role in the next generation of intelligent systems, offering novel capabilities that are both powerful and transparent.

We hope this survey provides a comprehensive foundation for researchers and practitioners to navigate the evolving landscape of Spiking Neural Networks, inspiring further innovations that bring us closer to truly brain-inspired artificial intelligence.

References

[1] Qian Liu and Steve Furber. Noisy softplus: A biology inspired activation function. In *Proceeddings of the International Conference on Neural Information Processing*, pages 405–412. Springer, 2016.

[2] Seijoon Kim, Seongsik Park, Byunggook Na, and Sungroh Yoon. Spiking-YOLO: Spiking neural network for energy-efficient object detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 11270–11277, 2020.

[3] Hongzhi Wang, Xiubo Liang, Mengjian Li, and Tao Zhang. Rtformer: Re-parameter tsbn spiking transformer. *arXiv preprint arXiv:2406.14180*, 2024.

[4] Manon Dampfhoffer, Thomas Mesquida, Alexandre Valentian, and Lorena Anghel. Backpropagation-based learning techniques for deep spiking neural networks: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 2023.

[5] Google. Cloud TPU. URL https://cloud.google.com/tpu.

[6] Chenlin Zhou, Han Zhang, Zhaokun Zhou, Liutao Yu, Zhengyu Ma, Huihui Zhou, Xiaopeng Fan, and Yonghong Tian. Enhancing the performance of transformer-based spiking neural networks by improved downsampling with precise gradient backpropagation. *arXiv preprint arXiv:2305.05954*, 2023.

[7] Zhanglu Yan, Jun Zhou, and Weng-Fai Wong. Near lossless transfer learning for spiking neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 10577–10584, 2021.

[8] José Antonio Pérez-Carrasco, Bo Zhao, Carmen Serrano, Begona Acha, Teresa Serrano-Gotarredona, Shouchun Chen, and Bernabé Linares-Barranco. Mapping from frame-driven to frame-free event-driven vision systems by low-rate rate coding and coincidence processing–application to feedforward convnets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11):2706–2719, 2013.

[9] Shihao Zou, Yuxuan Mu, Xinxin Zuo, Sen Wang, and Li Cheng. Event-based human pose tracking by spiking spatiotemporal transformer. *arXiv preprint arXiv:2303.09681*, 2023.

[10] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[11] Itay Hubara, Matthieu Courbariaux, Daniel Soudry, Ran El-Yaniv, and Yoshua Bengio. Binarized neural networks. *Advances in neural information processing systems*, 29:4114–4122, 2016.

[12] Abhiroop Bhattacharjee, Yeshwanth Venkatesha, Abhishek Moitra, and Priyadarshini Panda. Mime: Adapting a single neural network for multi-task inference with memory-efficient dynamic pruning. *arXiv preprint arXiv:2204.05274*, 2022.

[13] Wikipedia contributors. Tensor processing unit — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Tensor_processing_unit&oldid=958810965, 2020. [Online; accessed 1-June-2020].

[14] Siqi Wang, Tee Hiang Cheng, and Meng-Hiot Lim. LTMD: Learning improvement of spiking neural networks with learnable thresholding neurons and moderate dropout. *Advances in Neural Information Processing Systems*, 35:28350–28362, 2022.

[15] Kaushik Roy, Akhilesh Jaiswal, and Priyadarshini Panda. Towards spike-based machine intelligence with neuromorphic computing. *Nature*, 575 (7784):607–617, 2019.

[16] IBM Corporation. SyNAPSE. URL `https://en.wikipedia.org/wiki/SyNAPSE`.

[17] Shuang Lian, Jiangrong Shen, Ziming Wang, and Huajin Tang. IM-LIF: Improved neuronal dynamics with attention mechanism for direct training deep spiking neural network. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2024.

[18] Catherine D Schuman, Shruti R Kulkarni, Maryam Parsa, J Parker Mitchell, Prasanna Date, and Bill Kay. Opportunities for neuromorphic computing algorithms and applications. *Nature Computational Science*, 2(1):10–19, 2022.

[19] Chenlin Zhou, Liutao Yu, Zhaokun Zhou, Han Zhang, Zhengyu Ma, Huihui Zhou, and Yonghong Tian. Spikingformer: Spike-driven residual learning for transformer-based spiking neural network. *arXiv preprint arXiv:2304.11954*, 2023.

[20] Rui-Jie Zhu, Qihang Zhao, and Jason K Eshraghian. SpikeGPT: Generative pre-trained language model with spiking neural networks. *arXiv preprint arXiv:2302.13939*, 2023.

[21] Tobi Delbruck and Carver A Mead. Adaptive photoreceptor with wide dynamic range. In *IEEE International Symposium on Circuits and Systems*, volume 4, pages 339–342. IEEE, 1994.

[22] Abu Sebastian, Manuel Le Gallo, Riduan Khaddam-Aljameh, and Evangelos Eleftheriou. Memory devices and applications for in-memory computing. *Nature Nanotechnology*, pages 1–16, 2020.

[23] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015.

[24] Raphael Berner, Christian Brandli, Minhao Yang, Shih-Chii Liu, and Tobi Delbruck. A 240$\times$ 180 10mw 12us latency sparse-output vision sensor for mobile applications. In *Symposium on VLSI Circuits*, pages C186–C187. IEEE, 2013.

[25] Youngeun Kim and Priyadarshini Panda. Revisiting batch normalization for training low-latency deep spiking neural networks from scratch. *Frontiers in Neuroscience*, page 1638, 2021.

[26] Qiao Ning, Mostafa Hesham, Stefanini Fabio, and Sumislawska Dora. A reconfigurable on-line learning spiking neuromorphic processor comprising 256 neurons and 128k synapses. *Frontiers in Neuroscience*, 9:141, 2015.

[27] Zechun Liu, Zhiqiang Shen, Marios Savvides, and Kwang-Ting Cheng. Reactnet: Towards precise binary neural network with generalized activation functions. In *Proceedings of the European Conference on Computer Vision*, pages 143–159. Springer, 2020.

[28] Sander M Bohte, Joost N Kok, and Johannes A La Poutré. Spikeprop: backpropagation for networks of spiking neurons. In *Proceedings of the European Symposium on Artificial Neural Networks*, volume 48, pages 419–424, 2000.

[29] Intel Corporation. Intel Stratix 10 TX Device Overview. . URL `https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/hb/stratix-10/s10_tx_overview.pdf`.

[30] Stefano Ambrogio, Pritish Narayanan, Hsinyu Tsai, Robert M Shelby, Irem Boybat, Carmelo di Nolfo, Severin Sidler, Massimo Giordano, Martina Bodini, Nathan CP Farinha, et al. Equivalent-accuracy accelerated neural-network training using analogue memory. *Nature*, 558(7708):60–67, 2018.

[31] Vincent Chan, Shih-Chii Liu, and Andr van Schaik. Aer ear: A matched silicon cochlea pair with address event representation interface. *IEEE Transactions on Circuits and Systems*, 54(1):48–59, 2007.

[32] Yundong Sun, Dongjie Zhu, Yansong Wang, Zhaoshuo Tian, Ning Cao, and Gregory O'Hared. Spikegraphormer: A high-performance graph transformer with spiking graph attention. *arXiv preprint arXiv:2403.15480*, 2024.

[33] Byunggook Na, Jisoo Mok, Seongsik Park, Dongjin Lee, Hyeokjun Choe, and Sungroh Yoon. AutoSNN: Towards energy-efficient spiking neural networks. In *Proceedings of the International Conference on Machine Learning*, pages 16253–16269, 2022.

[34] Youngeun Kim, Joshua Chough, and Priyadarshini Panda. Beyond classification: Directly training spiking neural networks for semantic segmentation. *Neuromorphic Computing and Engineering*, 2022.

[35] Pengjie Gu, Rong Xiao, Gang Pan, and Huajin Tang. STCA: Spatio-temporal credit assignment with delayed feedback in deep spiking neural networks. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 1366–1372, 2019.

[36] Youngeun Kim and Priyadarshini Panda. Optimizing deeper spiking neural networks for dynamic vision sensing. *Neural Networks*, 144:686–698, 2021.

[37] Xiaotian Song, Andy Song, Rong Xiao, and Yanan Sun. One-step spiking transformer with a linear complexity. In *Proceedings of the International Joint Conference on Artificial Intelligence*, 2024.

[38] Etienne Mueller, Viktor Studenyak, Daniel Auge, and Alois Knoll. Spiking transformer networks: A rate coded approach for processing sequential data. In *Proceedings of the International Conference on Systems and Informatics*, pages 1–5. IEEE, 2021.

[39] Changze Lv, Tianlong Li, Jianhan Xu, Chenxi Gu, Zixuan Ling, Cenyuan Zhang, Xiaoqing Zheng, and Xuanjing Huang. SpikeBERT: A language spikformer trained with two-stage knowledge distillation from BERT. *arXiv preprint arXiv:2308.15122*, 2023.

[40] iniLabs. Dynamic Audio Sensor. URL `https://inilabs.com/products/dynamic-audio-sensor/`.

[41] Zhaokun Zhou, Kaiwei Che, Wei Fang, Keyu Tian, Yuesheng Zhu, Shuicheng Yan, Yonghong Tian, and Li Yuan. Spikformer v2: Join the high accuracy club on ImageNet with an SNN ticket. *arXiv preprint arXiv:2401.02020*, 2024.

[42] Romain Brette and Wulfram Gerstner. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of Neurophysiology*, 94(5):3637–3642, 2005.

[43] Intel Corporation. Intel Xeon Platinum 9282 Processor. . URL `https://ark.intel.com/content/www/us/en/ark/products/194146/intel-xeon-platinum-9282-processor-77m-cache-2-60-ghz.html`.

[44] Peter U Diehl, Daniel Neil, Jonathan Binas, Matthew Cook, Shih-Chii Liu, and Michael Pfeiffer. Fast-classifying, high-accuracy spiking deep networks through weight and threshold balancing. In *Proceedings of the International Joint Conference on Neural Networks*, pages 1–8, 2015.

[45] Jason K Eshraghian, Max Ward, Emre Neftci, Xinxin Wang, Gregor Lenz, Girish Dwivedi, Mohammed Bennamoun, Doo Seok Jeong, and Wei D Lu. Training spiking neural networks using lessons from deep learning. *arXiv preprint arXiv:2109.12894*, 2021.

[46] Sicheng Shen, Dongcheng Zhao, Guobin Shen, and Yi Zeng. TIM: An efficient temporal interaction module for spiking transformer. *arXiv preprint arXiv:2401.11687*, 2024.

[47] APT Advanced Processor Technologies Research Group. SpiNNaker. URL `https://apt.cs.manchester.ac.uk/projects/SpiNNaker/project/Access`.

[48] Yizhou Jiang, Kunlin Hu, Tianren Zhang, Haichuan Gao, Yuqian Liu, Ying Fang, and Feng Chen. Spatio-temporal approximation: A training-free snn conversion for transformers. In *Proceedings of the International Conference on Learning Representations*, 2024.

[49] ShouWei Gao, XiangYu Fan, XingYang Deng, ZiChao Hong, Hao Zhou, and ZiHao Zhu. TE-Spikformer: Temporal-enhanced spiking neural network with transformer. *Neurocomputing*, page 128268, 2024.

[50] Yuhang Li, Shikuang Deng, Xin Dong, Ruihao Gong, and Shi Gu. A free lunch from ANN: Towards efficient, accurate spiking neural networks calibration. In *Proceedings of the International Conference on Machine Learning*, pages 6316–6325, 2021.

[51] Renzo Andri, Lukas Cavigelli, Davide Rossi, and Luca Benini. Yodann: An architecture for ultralow power binary-weight cnn acceleration. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 37(1):48–60, 2017.

[52] Nitin Rathi and Kaushik Roy. DIET-SNN: A low-latency spiking neural network with direct input encoding and leakage and threshold optimization. *IEEE Transactions on Neural Networks and Learning Systems*, 34(6):3174–3182, 2021.

[53] Jiaqi Yan, Qianhui Liu, Malu Zhang, Lang Feng, De Ma, Haizhou Li, and Gang Pan. Efficient spiking neural network design via neural architecture search. *Neural Networks*, page 106172, 2024.

[54] Christian Brandli, Raphael Berner, Minhao Yang, Shih-Chii Liu, and Tobi Delbruck. A $240\times$ 180 130 db 3 μs latency global shutter spatiotemporal vision sensor. *IEEE Journal of Solid-State Circuits*, 49(10):2333–2341, 2014.

[55] Qinyu Chen, Congyi Sun, Chang Gao, and Shih-Chii Liu. Epilepsy seizure detection and prediction using an approximate spiking convolutional transformer. *arXiv preprint arXiv:2402.09424*, 2024.

[56] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[57] Benjamin Scellier and Yoshua Bengio. Equilibrium propagation: Bridging the gap between energy-based models and backpropagation. *Frontiers in Computational Neuroscience*, 11:24, 2017.

[58] Wulfram Gerstner and Werner M Kistler. *Spiking neuron models: Single neurons, populations, plasticity.* Cambridge University Press, 2002.

[59] Bhaskar Mukhoty, Velibor Bojkovic, William de Vazelhes, Xiaohan Zhao, Giulia De Masi, Huan Xiong, and Bin Gu. Direct training of snn using local zeroth order method. *Advances in Neural Information Processing Systems*, 36:18994–19014, 2024.

[60] Wolfgang Maass. Lower bounds for the computational power of networks of spiking neurons. *Neural Computation*, 8(1):1–40, 1996.

[61] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, and Adrian and Bolton. Mastering the game of go without human knowledge. *Nature*, 550(7676):354–359, 2017.

[62] Wei Fang, Zhaofei Yu, Yanqi Chen, Tiejun Huang, Timothée Masquelier, and Yonghong Tian. Deep residual learning in spiking neural networks. *Advances in Neural Information Processing Systems*, 34:21056–21069, 2021.

[63] Peisong Wang, Xiangyu He, Gang Li, Tianli Zhao, and Jian Cheng. Sparsity-inducing binarized neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 12192–12199, 2020.

[64] Chenlin Zhou, Han Zhang, Zhaokun Zhou, Liutao Yu, Liwei Huang, Xiaopeng Fan, Li Yuan, Zhengyu Ma, Huihui Zhou, and Yonghong Tian. QK-Former: Hierarchical spiking transformer using qk attention. *arXiv preprint arXiv:2403.16552*, 2024.

[65] Youngeun Kim, Yuhang Li, Hyoungseob Park, Yeshwanth Venkatesha, and Priyadarshini Panda. Neural architecture search for spiking neural networks. In *Proceedings of the European Conference on Computer Vision*, pages 36–56, 2022.

[66] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30:6000–6010, 2017.

[67] Emre O Neftci, Hesham Mostafa, and Friedemann Zenke. Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks. *IEEE Signal Processing Magazine*, 36(6):51–63, 2019.

[68] Alexander Borst and Frédéric E Theunissen. Information theory and neural coding. *Nature Neuroscience*, 2(11):947–957, 1999.

[69] Jun Haeng Lee, Tobi Delbruck, and Michael Pfeiffer. Training deep spiking neural networks using backpropagation. *Frontiers in Neuroscience*, 10:508, 2016.

[70] Ronan Collobert, Jason Weston, Léon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. Natural language processing (almost) from scratch. *Journal of Machine Learning Sesearch*, 12:2493–2537, 2011.

[71] Yuchen Wang, Kexin Shi, Chengzhuo Lu, Yuguo Liu, Malu Zhang, and Hong Qu. Spatial-temporal self-attention for asynchronous spiking neural networks. In *Proceedings of the International Joint Conference on Artificial Intelligence*, volume 8, pages 3085–3093, 2023.

[72] Yangfan Hu, Qian Zheng, Xudong Jiang, and Gang Pan. Fast-SNN: Fast spiking neural network by converting quantized ANN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(12):14546–14562, 2023.

[73] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[74] Rufin Van Rullen and Simon J Thorpe. Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Computation*, 13(6):1255–1283, 2001.

[75] Shih-Chii Liu, André van Schaik, Bradley A Mincti, and Tobi Delbruck. Event-based 64-channel binaural silicon cochlea with q enhancement mechanisms. In *Proceedings of the IEEE International Symposium on Circuits and Systems*, pages 2027–2030. IEEE, 2010.

[76] Peter O'Connor, Efstratios Gavves, and Max Welling. Training a spiking neural network with equilibrium propagation. In *Proceedings of the International Conference on Artificial Intelligence and Statistics*, pages 1516–1523, 2019.

[77] Daniel Neil, Michael Pfeiffer, and Shih-Chii Liu. Learning to be efficient: Algorithms for training low-latency, low-compute deep spiking neural networks. In *Proceedings of the Annual ACM Symposium on Applied Computing*, pages 293–298, 2016.

[78] Gourav Datta, Zeyu Liu, Anni Li, and Peter A Beerel. Spiking neural networks with dynamic time steps for vision transformers. *arXiv preprint arXiv:2311.16456*, 2023.

[79] Peter O'Connor, Daniel Neil, Shih-Chii Liu, Tobi Delbruck, and Michael Pfeiffer. Real-time classification and sensor fusion with a spiking deep belief network. *Frontiers in Neuroscience*, 7:178, 2013.

[80] Wolfgang Maass. To spike or not to spike: that is the question. *Proceedings of the IEEE*, 103(12):2219–2224, 2015.

[81] Doo Seok Jeong, Kyung Min Kim, Sungho Kim, Byung Joon Choi, and Cheol Seong Hwang. Memristors for energy-efficient new computing paradigms. *Advanced Electronic Materials*, 2(9):1600090, 2016.

[82] Frédéric Theunissen and John P Miller. Temporal encoding in nervous systems: a rigorous definition. *Journal of Computational Neuroscience*, 2(2): 149–162, 1995.

[83] Yeshwanth Venkatesha, Youngeun Kim, Leandros Tassiulas, and Priyadarshini Panda. Federated learning with spiking neural networks. *IEEE Transactions on Signal Processing*, 69:6183–6194, 2021.

[84] Yoshua Bengio, Nicholas Léonard, and Aaron Courville. Estimating or propagating gradients through stochastic neurons for conditional computation. *arXiv preprint arXiv:1308.3432*, 2013.

[85] Bodo Rueckauer, Iulia-Alexandra Lungu, Yuhuang Hu, Michael Pfeiffer, and Shih-Chii Liu. Conversion of continuous-valued deep networks to efficient event-driven networks for image classification. *Frontiers in Neuroscience*, 11:682, 2017.

[86] Qi Xu, Yaxin Li, Jiangrong Shen, Jian K Liu, Huajin Tang, and Gang Pan. Constructing deep spiking neural networks from artificial neural networks

with knowledge distillation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7886–7895, 2023.

[87] Wei Fang, Zhaofei Yu, Yanqi Chen, Timothée Masquelier, Tiejun Huang, and Yonghong Tian. Incorporating learnable membrane time constant to enhance learning of spiking neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2661–2671, 2021.

[88] Eunhyeok Park, Junwhan Ahn, and Sungjoo Yoo. Weighted-entropy-based quantization for deep neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5456–5464, 2017.

[89] Jiqing Zhang, Bo Dong, Haiwei Zhang, Jianchuan Ding, Felix Heide, Baocai Yin, and Xin Yang. Spiking transformers for event-based single object tracking. In *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, pages 8801–8810, 2022.

[90] Wenrui Zhang and Peng Li. Temporal spike sequence learning via backpropagation for deep spiking neural networks. *Proceeddings of the International Conference on Neural Information Processing*, 33:12022–12033, 2020.

[91] Yufei Guo, Yuanpei Chen, Liwen Zhang, Xiaode Liu, Yinglei Wang, Xuhui Huang, and Zhe Ma. Im-loss: information maximization loss for spiking neural networks. *Advances in Neural Information Processing Systems*, 35: 156–166, 2022.

[92] Mike Davies, Narayan Srinivasa, Tsung-Han Lin, Gautham Chinya, Yongqiang Cao, Sri Harsha Choday, Georgios Dimou, Prasad Joshi, Nabil Imam, Shweta Jain, et al. Loihi: A neuromorphic manycore processor with on-chip learning. *IEEE Micro*, 38(1):82–99, 2018.

[93] Yudong Li, Yunlin Lei, and Xu Yang. Spikeformer: Training high-performance spiking neural network with transformer. *Neurocomputing*, 574: 127279, 2024.

[94] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the International Conference on Machine Learning*, 2010.

[95] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[96] Patrick Lichtsteiner, Christoph Posch, and T Delbruck. A 128 × 128 120db 30mw asynchronous vision sensor that responds to relative intensity change. In *IEEE International Solid-State Circuits Conference*, pages 2060–2069. IEEE, 2006.

[97] Man Yao, Jiakui Hu, Tianxiang Hu, Yifan Xu, Zhaokun Zhou, Yonghong Tian, Bo Xu, and Guoqi Li. Spike-driven transformer v2: Meta spiking neural network architecture inspiring the design of next-generation neuromorphic chips. In *Proceedings of the International Conference on Learning Representation*, 2024.

[98] Karan Desai and Justin Johnson. Virtex: Learning visual representations from textual annotations. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 11162–11173, 2021.

[99] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500, 1952.

[100] Eugene M Izhikevich, Niraj S Desai, Elisabeth C Walcott, and Frank C Hoppensteadt. Bursts as a unit of neural information: selective communication via resonance. *Trends in Neurosciences*, 26(3):161–167, 2003.

[101] Yuhang Li, Shikuang Deng, Xin Dong, and Shi Gu. Error-aware conversion from ANN to SNN via post-training parameter calibration. *International Journal of Computer Vision*, pages 1–24, 2024.

[102] Man Yao, Huanhuan Gao, Guangshe Zhao, Dingheng Wang, Yihan Lin, Zhaoxu Yang, and Guoqi Li. Temporal-wise attention spiking neural networks for event streams classification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10221–10230, 2021.

[103] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.

[104] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, 2016.

[105] Yaoyu Zhu, Wei Fang, Xiaodong Xie, Tiejun Huang, and Zhaofei Yu. Exploring loss functions for time-based training strategy in spiking neural networks. *Advances in Neural Information Processing Systems*, 36, 2024.

[106] Yassine Zniyed, Thanh Phuong Nguyen, et al. Enhanced network compression through tensor decompositions and pruning. *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[107] Yongqiang Cao, Yang Chen, and Deepak Khosla. Spiking deep convolutional neural networks for energy-efficient object recognition. *International Journal of Computer Vision*, 113(1):54–66, 2015.

[108] Jung Hyun Lee, Jihun Yun, Sung Ju Hwang, and Eunho Yang. Cluster-promoting quantization with bit-drop for minimizing network quantization loss. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5370–5379, 2021.

[109] Darryl Lin, Sachin Talathi, and Sreekanth Annapureddy. Fixed point quantization of deep convolutional networks. In *Proceedings of the International Conference on Machine Learning*, pages 2849–2858, 2016.

[110] Tiantian Han, Dong Li, Ji Liu, Lu Tian, and Yi Shan. Improving low-precision network quantization via bin regularization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5261–5270, 2021.

[111] Yifan Hu, Lei Deng, Yujie Wu, Man Yao, and Guoqi Li. Advancing spiking neural networks toward deep residual learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[112] Yangfan Hu, Huajin Tang, and Gang Pan. Spiking deep residual networks. *IEEE Transactions on Neural Networks and Learning Systems*, 34(8):5200–5205, 2021.

[113] Rufin VanRullen and Simon J Thorpe. Surfing a spike wave down the ventral stream. *Vision Research*, 42(23):2593–2615, 2002.

[114] Mingbao Lin, Rongrong Ji, Zihan Xu, Baochang Zhang, Yan Wang, Yongjian Wu, Feiyue Huang, and Chia-Wen Lin. Rotated binary neural network. *Proceedings of the International Conference on Neural Information Processing Systems*, 33:7474–7485, 2020.

[115] Yuhang Li, Yufei Guo, Shanghang Zhang, Shikuang Deng, Yongqing Hai, and Shi Gu. Differentiable spike: Rethinking gradient-descent for training

spiking neural networks. *Advances in Neural Information Processing Systems*, 34:23426–23439, 2021.

[116] Zihan Xu, Mingbao Lin, Jianzhuang Liu, Jie Chen, Ling Shao, Yue Gao, Yonghong Tian, and Rongrong Ji. ReCU: Reviving the dead weights in binary neural networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 5198–5208, 2021.

[117] Boxun Xu, Hejia Geng, Yuxuan Yin, and Peng Li. DISTA: Denoising spiking transformer with intrinsic plasticity and spatiotemporal attention. *arXiv preprint arXiv:2311.09376*, 2023.