Computação Neuromórfica

Instituto Superior de Engenharia do Porto Arquiteturas de Sistemas Críticos de Computadores - 2024/2025 Reinaldo Reis - 1201560, Nuno Castro - 1240160

Resumo

A computação neuromórfica está a emergir como uma área inovadora com grande potencial em Inteligência Artificial(IA) e outras tecnologias avançadas. Este relatório aborda a computação neuromórfica, uma abordagem inspirada no cérebro humano, que visa otimizar o processamento de informações por meio de arquiteturas e memórias avançadas, como ReRAM, MRAM e memristores. Focada na redução da latência e do consumo energético, em comparação com sistemas tradicionais, a computação neuromórfica utiliza dispositivos como sinapses artificiais e redes neurais de spiking (SNNs).

O artigo explora os desafios associados a esta tecnologia, incluindo questões de escalabilidade, precisão, confiabilidade, e a adaptação dos algoritmos de Deep Learning, além das dificuldades na integração com tecnologias convencionais. Apesar de todos os desafios, esta tecnologia apresenta algumas vantagens, como a maior eficiência energética e o seu potencial em veículos autónomos, cibersegurança, Edge AI, robótica e reconhecimento de padrões.

A computação neuromórfica continua a evoluir, com tendências como o desenvolvimento de novos dispositivos, a integração com a computação quântica e a IA, e o aprimoramento contínuo dos algoritmos. O futuro da computação neuromórfica promete revolucionar a IA, oferecendo sistemas mais rápidos, eficientes e adaptáveis, com um vasto campo de aplicação em diversas áreas.

Introdução

Com o passar dos anos, a procura por capacidade computacional tem aumentado significativamente, impulsionada por aplicações como a Inteligência Artificial (IA) e a Internet das Coisas (IoT). Devido a este aumento, surgiram algumas limitações nas arquiteturas tradicionais, tal como na arquitetura de Von Neumann, especialmente na forma como se processa e gere o fluxo constante de dados entre a memória e o processador. Para superar estas limitações, surge a computação neuromórfica, destacando-se como uma das áreas mais promissoras no campo da ciência e engenharia computacional. Inspirada na estrutura e funcionamento do cérebro humano, esta abordagem tem como objetivo criar dispositivos capazes de aprender, reter informações e realizar deduções lógicas como um ser humano, além de integrar memória e processamento de forma mais eficiente.[1]

Contudo a computação neuromórfica ainda enfrenta muitos desafios, como o desenvolvimento de novos materiais e algoritmos que possam suportar a criação de sistemas eficientes e escaláveis. No entanto, o interesse crescente de investigadores e empresas demonstram o potencial da computação neuromórfica.[1]

Contexto Histórico

A computação neuromórfica tem as suas origens na procura por sistemas que se inspirem na dinâmica dos circuitos neurais biológicos, com um desenvolvimento que remonta ao final da década de 1950. Embora modelos iniciais, como os *perceptrons* de Rosenblatt (1958) e as retinas eletrónicas de Fukushima (1970), tenham estabelecido bases importantes, foi nos anos 1980 que o campo se impulsionou com os avanços da tecnologia *Very-large-scale integration*(VLSI).[2]

Em 1981, Carver Mead introduziu dispositivos analógicos inspirados no funcionamento neural, como retinas e cócleas de silício. O seu trabalho, influenciado por figuras como John Hopfield, Max Delbrück e Richard Feynman, utilizou as propriedades analógicas dos transístores para criar circuitos neuromórficos eficientes, capazes de simular a atividade neural com uma menor necessidade de componentes em comparação com abordagens digitais.[2]

A área consolidou-se com iniciativas como o curso de Física da Computação e o livro *Analog VLSI and Neural Systems* (1989), de Mead, que influenciaram uma nova geração de investigadores. Entre os acontecimentos mais importantes destacam-se o Projeto Cérebro Humano, lançado em 2013, e o desenvolvimento de chips neuromórficos, como o TrueNorth da IBM (2014) e o Loihi da Intel (2018), aplicados em áreas como a robótica, o reconhecimento de padrões e a IA.[2]

Estado da Arte

A computação neuromórfica, inspirada na arquitetura e funcionamento do cérebro humano, tem despertado grande interesse devido ao seu potencial para superar limitações das arquiteturas tradicionais, como a arquitetura de Von Neumann. Este campo combina a neurociência, ciência dos materiais, engenharia de sistemas e IA, promovendo avanços significativos na eficiência energética, processamento paralelo e adaptabilidade [3, 4].

Fundamentos de Computação Neuromórfica

A computação neuromórfica procura replicar funcionalidades cerebrais, utilizando princípios bioinspirados. Este paradigma permite criar sistemas que combinam eficiência energética e capacidade de processamento paralelo, integrando memória e lógica num único dispositivo. Baseados em neurónios e sinapses biológicos, os sistemas neuromórficos focamse em:

- Processamento em paralelo: Inspirado na capacidade do cérebro de processar múltiplos fluxos de dados em simultâneo, reduzindo a latência e aumentando a eficiência global.
- Eficiência energética: Sistemas como o IBM TrueNorth e o Intel Loihi atingem consumos energéticos muito baixos, aproximando-se do desempenho do cérebro humano [3].
- Aprendizagem e adaptação: Mecanismos de plasticidade sináptica permitem ajustamentos dinâmicos, fundamentais para a aprendizagem e adaptação em tempo real [3].
- Integração de sensores e processamento: Sistemas neuromórficos podem integrar sensores bio-inspirados que percebem e processam dados de forma semelhante ao sistema nervoso, proporcionando respostas mais rápidas e otimizadas [5].
- Robustez contra ruído: Inspirando-se na resiliência do cérebro a sinais ruidosos, os sistemas neuromórficos são projetados para operar de forma eficaz mesmo em ambientes com interferências, o que é essencial para aplicações no mundo real [6].
- Implementação em hardware emergente: Tecnologias como memristores, transístores sinápticos e dispositivos spintrónicos estão a impulsionar a computação neuromórfica, proporcionando novas oportunidades para o desenvolvimento de hardware eficiente e compacto [3].

Fundamentos do Cérebro

Para criar um sistema neuromórfico capaz de emular as capacidades do cérebro, é essencial compreender como funciona o cérebro e, de um modo geral, como os sistemas nervosos operam e processam informação.

O neurónio é uma célula especializada reconhecida como o bloco fundamental do cérebro (Figura 1). Os neurónios apresentam características únicas de excitabilidade elétrica, permitindo-lhes processar e trocar sinais em larga escala e a uma velocidade relativamente rápida. Para compreender este processo, é necessário examinar de perto o funcionamento de um neurónio biológico [7].

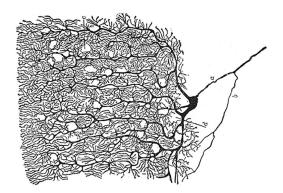


Figure 1: Ilustração de uma célula de Purkinje criada por Ramón y Cajal [7].

Funcionamento de um Neurónio Biológico

O neurónio biológico é a unidade fundamental do sistema nervoso, responsável pela receção, processamento e transmissão de sinais elétricos entre as diferentes partes do corpo e o cérebro. O funcionamento dos neurónios envolve uma série de processos biofísicos altamente complexos que permitem a comunicação rápida e eficiente entre células nervosas. A seguir, apresentamos os principais processos envolvidos no funcionamento de um neurónio biológico:

Potencial de Repouso: Normalmente, os neurónios mantêm um potencial elétrico de aproximadamente -70 mV através da camada bilipídica que separa o interior da célula do exterior (Figura 2). Este potencial, chamado potencial de repouso, é mantido por estruturas na membrana conhecidas como *bombas iónicas*. Estas bombas utilizam energia para expelir 3 iões de sódio (Na⁺) para fora da célula e introduzir 2 iões de potássio (K⁺) no seu interior, criando uma diferença de carga na membrana. [7]

Geração de Potenciais de Ação: Estruturas adicionais na membrana, conhecidas como *canais iónicos*, permitem a passagem seletiva de iões como Na⁺, K⁺ e Ca²⁺. Quando o potencial da membrana é perturbado acima de um limiar, canais de Na⁺ dependentes de voltagem abrem-se, permitindo que estes iões entrem na célula. Este movimento gera um aumento

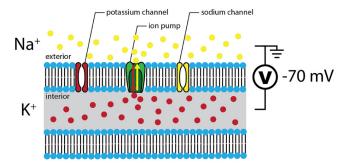


Figure 2: Ilustração da membrana de um neurónio [7].

rápido no potencial da membrana, seguido da ativação de canais de K⁺, que restauram o potencial ao valor inicial. Este ciclo de subida e descida rápida cria um *potencial de ação* (Figura 3) [7].

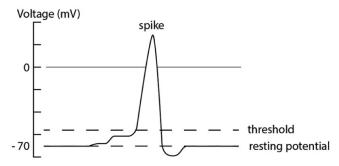


Figure 3: Geração de um potencial de ação. A subida do potencial acima do limiar provoca a abertura de canais de Na^+ , seguida da ativação de canais de K^+ , que restauram o potencial [7].

Transmissão Sináptica: Os dendritos de um neurónio contêm locais onde as mensagens de outros neurónios são recebidas (Figura 6). Os terminais do axónio de outros neurónios aproximam-se dos espinhos dendríticos, formando uma *fenda sináptica* de cerca de 20 nm. Quando um neurónio pré-sináptico dispara, liberta neurotransmissores que se difundem pela fenda e se ligam a recetores na membrana do neurónio pós-sináptico [7].

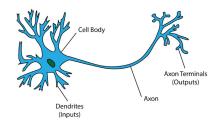


Figure 4: Estrutura genérica de um neurónio [7].

Plasticidade Sináptica: Os *pesos sinápticos* determinam a influência de uma sinapse na atividade do neurónio pós-sináptico. Estes pesos não são estáticos,

sendo ajustados através de processos como a plasticidade dependente do tempo de disparo (*STDP*, Figura 5). Este mecanismo aumenta o peso das sinapses ativas antes do disparo de um neurónio e diminui o peso das sinapses ativas após o disparo [7].

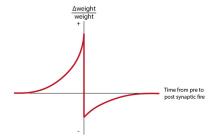


Figure 5: Curva típica de alteração do peso sináptico segundo a teoria STDP. Sinapses ativas antes do disparo são reforçadas, enquanto sinapses ativas após o disparo são enfraquecidas [7].

Sistemas e Arquiteturas Neuromórficas

Tendo descoberto o potencial do que as redes biológicas de neurônios podem realizar, muitos pesquisadores tentaram recriar artificialmente elementos semelhantes a neurônios que capturam alguns aspectos da sua capacidade biológica.

Neurônios Artificiais e Spiking Neural Networks (SNNs): Uma das descrições mais simples de um neurônio é o modelo de *Integrate-and-Fire* (IF) [7]. Neste modelo, o neurônio é controlado por sua tensão interna, começando no potencial de repouso e acumulando carga dos picos recebidos. Quando a voltagem ultrapassa um limite, o neurônio dispara, reiniciando seu potencial após o pico.

O modelo pode incluir uma corrente de fuga, que faz com que a tensão retorne ao nível de repouso, resultando no modelo de *leaky-Integrate-and-Fire* (LIF). Esse modelo tem várias implementações eletrônicas e é útil em redes neurais complexas ou SNNs [8].

Outra modificação é a inclusão de ruído, que simula o comportamento estocástico dos neurônios biológicos. O modelo pode então ser ajustado para um modelo ruidoso, onde a tensão flutua, mas tende a retornar ao repouso.

Embora o modelo IF seja amplamente utilizado, existem outras variações como Hodgkin-Huxley, Izhikevich, FitzHugh-Nagumo e multicompartimentos, que buscam uma representação mais realista do comportamento neuronal.

Hardware Neuromórfico para SNNs: Embora qualquer computador padrão possa simular uma SNN, obter as vantagens potenciais da computação baseada em SNNs depende fundamentalmente de sua execução em hardware neuromórfico especializado. Isso ocorre porque a computação neuronal é intrinsecamente paralela, ou seja, cada neurônio realiza continuamente o cálculo de seu estado interno. Quando representado por uma série de equações diferenciais, um computador von Neumann padrão deve calcular sequencialmente o novo estado de cada neurônio na rede. Esse processo resulta em tempos de computação muito maiores em hardware não neuromórfico, especialmente à medida que a escala da rede aumenta [7].

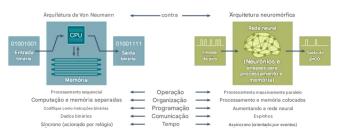


Figure 6: Comparação da arquitetura von Neumann com a arquitetura neuromórfica

Para facilitar a pesquisa eficiente sobre SNNs, muitas arquiteturas neuromórficas foram propostas e desenvolvidas por pesquisadores acadêmicos e pela indústria. Entre as arquiteturas acadêmicas estão o SpiNNaker e o BrainScaleS, ambos apoiados pelo projeto europeu *Human Brain Project*.

- SpiNNaker: Arquitetura digital altamente paralela, permite que processadores independentes resolvam modelos neuronais arbitrários e transmitam mensagens que representam picos [7]. O chip SpiNNaker contém 18 núcleos ARM, podendo atingir até 1 milhão de núcleos em um único sistema. A plataforma pode emular 1 bilhão de neurônios em tempo real, e a versão atualizada do chip é 10 vezes mais eficiente. [3]
- BrainScaleS: Sistema misto analógico-digital que utiliza diretamente a dinâmica elétrica de transistores para criar modelos neuronais, com parâmetros armazenados em *local floating-gate transistors* [7]. Cada wafer pode conter 200.000 neurônios e 44 milhões de sinapses, e os cálculos são acelerados em até 10.000 vezes em relação ao tempo real. [3]

Arquiteturas experimentais também foram criadas por empresas como IBM e Intel:

 TrueNorth (IBM): Explora aplicações práticas de SNNs com neurônios e sinapses baseados no modelo LIF, priorizando eficiência computacional e espacial [7]. O sistema hospeda até 1 milhão de neurônios simples, consumindo cerca de 63 mW por núcleo, com 4096 núcleos idênticos. [3] • Loihi (Intel): Projetado para investigar como Artificial Neural Networks (ANNs) podem ser adaptadas para arquiteturas de spiking, potencialmente permitindo treinamento de ANNs nesses sistemas em vez de GPUs/TPUs mais caras [7]. A versão Loihi 2 contém até 1 milhão de neurônios por chip, fabricado em tecnologia Intel 4 (7 nm). Um sistema com 768 chips Loihi consome menos de 500 watts. [3]

As escolhas de design dessas arquiteturas criam limitações importantes para os utilizadores finais. Por exemplo, o BrainScaleS utiliza componentes analógicos que sofrem de variabilidade, o que pode levar a discrepâncias entre simulações ideais e resultados reais. Já o TrueNorth possui limites estritos na conectividade de cada neurônio, o que pode restringir aplicações específicas [7].

Tecnologias de Memórias Resistivas e Memristores

As tecnologias de memórias resistivas e memristores são fundamentais na computação neuromórfica, pois oferecem uma abordagem inovadora para integrar memória e processamento no mesmo dispositivo, imitando características dos neurônios e sinapses biológicas. Ao contrário das arquiteturas tradicionais, onde a memória e o processamento são separados, estas tecnologias permitem a realização de operações computacionais diretamente onde os dados estão armazenados. Algumas dessas tecnologias são:

- RAM resistiva (ReRAM): Combina baixa latência com baixo consumo energético, simulando sinapses artificiais. Quando uma tensão é aplicada, a resistência do material muda, permitindo a gravação de estados binários ou até mesmo estados analógicos.[4]
- RAM magnetorresistiva (MRAM): A MRAM utiliza mudanças na resistência magnética para armazenar dados. É uma tecnologia não volátil que mantém os dados mesmo quando a energia é desligada. [9]
- Memória de Mudança de Fase (PCM): Utiliza alterações de fase para armazenar dados, suportando operações analógicas em redes neuromórficas [3].
- Memristores: Ajustam a resistência em resposta a estímulos, permitindo operações como aprendizagem supervisionada e não supervisionada [3].

Integração de Arquiteturas Neuromórficas e Tecnologias de Hardware

A integração de arquiteturas neuromórficas e tecnologias de hardware avançado é fundamental para explorar todo o potencial da computação inspirada no cérebro. Os modelos computacionais teóricos, como redes neurais de spiking (SNNs), ganham eficiência prática quando combinados com plataformas especializadas e dispositivos eletrônicos otimizados. A seguir, destacam-se os principais elos entre as arquiteturas neuromórficas e as tecnologias de hardware.

Sinergia entre Arquiteturas e Memórias Resistivas. A combinação de memórias resistivas, como ReRAM e memristores, com arquiteturas neuromórficas, como SpiNNaker e Loihi, permite uma aproximação mais eficiente da computação sináptica. Por exemplo:

- Memórias Resistivas como Sinapses: Dispositivos como memristores replicam o comportamento de sinapses biológicas ao ajustar sua resistência com base em estímulos. Isso complementa arquiteturas digitais como Loihi, onde cada sinapse programável pode utilizar dispositivos de baixo consumo energético para armazenar estados [3, 4].
- Integração com Processamento Analógico: Arquiteturas mistas como o BrainScaleS exploram a natureza analógica de dispositivos como PCM para armazenar e processar dados diretamente, evitando a latência associada à leitura e escrita sequencial [7].

Processamento Neuromórfico

O processamento neuromórfico codifica informações de sinais sob a forma de sequências de eventos, reduzindo a quantidade de dados transmitidos e processados para incluir apenas as informações relevantes, como alterações no sistema. Este princípio inspira-se diretamente nos sistemas nervosos biológicos, onde a transmissão de informações (sob a forma de picos neuronais) e o consequente processamento têm impacto no consumo de energia. Assim, os sistemas biológicos tendem a minimizar o número de picos transmitidos e processados para realizar uma tarefa computacional. Este é o princípio que a computação neuromórfica procura seguir.[3]

A **Figura** 7 ilustra um exemplo deste método de codificação eficiente, orientado a eventos e sem enquadramento.

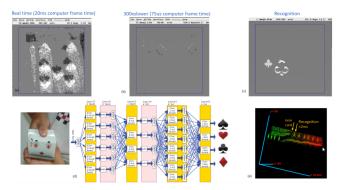


Figure 7: Exemplo de processamento neuromórfico. (a) Visualização rápida de um baralho de cartas, com eventos capturados a cada 20 ms para construir um fotograma. (b) Reprodução em baixa velocidade com fotogramas reconstruídos a 77 µs. (c) Símbolo de carta seguido e reconhecido em tempo real. (d) Estrutura CNN orientada a eventos para classificação de quatro símbolos de cartas. (e) Representação de eventos {x, y, t} durante 20 ms, exibindo eventos da câmara e do reconhecimento.[3]

Exemplo Prático de Processamento Neuromórfico Na Figura 7, vemos as seguintes etapas:

- (a) Um baralho de cartas é percorrido rapidamente, capturando eventos a cada 20 ms para construir um fotograma.[3]
- (b) Os mesmos eventos são reproduzidos em baixa velocidade, com fotogramas reconstruídos a 77 μs.[3]
- (c) Um símbolo de carta é rastreado e o resultado do reconhecimento é exibido em tempo real.[3]
- (d) Uma rede neural convolucional orientada a eventos (CNN) classifica os símbolos das cartas.[3]
- (e) Durante a transição de uma carta para outra, o sistema reconhece o novo símbolo em menos de 2 ms.[3]

Note-se que o sistema, composto por um sensor e a rede de reconhecimento, trabalha em conjunto para atingir este desempenho. Em contraste, os sistemas artificiais convencionais precisam de adquirir duas imagens consecutivas (tipicamente 25 ms por imagem) para detetar alterações.[3]

Desafios na Computação Neuromórfica

A computação neuromórfica, embora promissora, enfrenta uma série de desafios técnicos, materiais e algorítmicos que precisam ser superados para alcançar seu pleno potencial [3, 7]. A seguir, são abordados os principais obstáculos que ainda limitam a adoção e evolução desta tecnologia.

Desafios de Escalabilidade

A escalabilidade é um dos maiores desafios da computação neuromórfica, exigindo dispositivos eficientes, de baixa latência e baixo consumo de energia. Protótipos como SpiNNaker e Loihi mostram potencial, mas a replicação em larga escala ainda é limitada por recursos computacionais e físicos [4]. A integração de memória e processamento é essencial, mas a escalabilidade e eficiência continuam sendo obstáculos, com tecnologias como memristores e ReRAM enfrentando problemas de confiabilidade [3]. Além disso, simular grandes redes de neurónios e sinapses exige arquiteturas avançadas e software otimizado [4].

Desafios de Precisão e Confiabilidade

Apesar dos benefícios em eficiência energética, a precisão das operações em dispositivos de memória não volátil, como memristores, é limitada devido à variabilidade e não linearidade desses dispositivos, o que compromete o desempenho de redes neuromórficas, especialmente em tarefas que exigem alta precisão, além de levantar preocupações sobre a estabilidade a longo prazo e durabilidade, especialmente em aplicações de aprendizagem contínua ou em tempo real [4].

Desafios de Algoritmos e Aprendizagem

Adaptar algoritmos de *Deep Learning* para redes neuromórficas permanece um desafio significativo. As SNNs têm requisitos específicos de aprendizagem que diferem dos métodos convencionais utilizados em redes neurais artificiais.

- Treinamento de Redes Neuronais de Spiking: As SNNs dependem de mecanismos como a plasticidade sináptica, que não são compatíveis com retropropagação. Métodos como *Spike-Timing-Dependent Plasticity* (STDP) estão sendo explorados, mas carecem de técnicas escaláveis [3].
- Aprendizagem Online e Adaptativa: A aprendizagem contínua, uma característica atrativa das SNNs, enfrenta desafios na adaptação eficiente a novas informações sem a necessidade de re-treinamento completo [4].

Desafios de Integração com Tecnologias Existentes

A integração de sistemas neuromórficos com arquiteturas de computação convencionais é crucial para sua adoção em aplicações reais. Isso requer novas formas de interconexão e comunicação para garantir compatibilidade com plataformas existentes. Para

o sucesso comercial, é necessário integrar a computação neuromórfica aos fluxos de trabalho atuais, sem exigir mudanças significativas no hardware ou no desenvolvimento de software [3].

Desafios de Sustentabilidade e Custos

Apesar da eficiência energética prometida, os custos de produção e o impacto ambiental da computação neuromórfica continuam sendo barreiras significativas.

- Custo de Produção: A fabricação de dispositivos emergentes, como memristores, exige processos especializados que elevam os custos [4].
- Sustentabilidade: A sustentabilidade ambiental dos novos materiais e processos precisa ser considerada, especialmente em um mundo que prioriza a eficiência energética [3].

Casos de Uso

Embora as aplicações reais de sistemas neuromórficos ainda sejam limitadas, o paradigma computacional pode ser aplicado nas seguintes áreas:

Veículos Autónomos

A computação neuromórfica, devido ao seu alto desempenho e eficiência energética, pode revolucionar a navegação de veículos autónomos. Esta tecnologia permite correções de trajetória mais rápidas e melhora a capacidade de evitar colisões, enquanto reduz as emissões energéticas. [10].

Cibersegurança

Os sistemas neuromórficos podem ajudar a detectar padrões ou atividades incomuns que possam sinalizar ataques cibernéticos ou violação de dados. Esses riscos podem ser rapidamente neutralizados graças à baixa latência e à rápida computação dos dispositivos neuromórficos [10].

Edge AI

As características da arquitetura neuromórfica tornam-na adequada para Edge AI. O baixo consumo energético pode ajudar a lidar com a curta duração da bateria de dispositivos como smartphones e Wearables, enquanto a sua adaptabilidade e natureza orientada a eventos são adequadas para os métodos de processamento de informações de sensores remotos, drones e outros dispositivos de IoT [10].

Reconhecimento de Padrões

Devido à sua vasta capacidade de processamento paralelo, a computação neuromórfica pode ser utilizada em aplicações de aprendizagem automática para reconhecer padrões em linguagem natural, analisar imagens médicas e processar sinais de imagem de exames de ressonância magnética funcional (fMRI) e de eletroencefalogramas (EEG), que medem a atividade elétrica no cérebro [10].

Robótica

Como uma tecnologia adaptável, a computação neuromórfica pode ser usada para aprimorar as habilidades de aprendizagem em tempo real e tomada de decisões de um robô, ajudando-o a reconhecer melhor objetos e operar de forma mais rápida em linhas de montagem [10].

Tendências Futuras

A computação neuromórfica está em constante evolução, com desenvolvimentos significativos tanto em termos de hardware quanto de algoritmos. As tendências futuras neste campo não apenas visam superar os desafios atuais, mas também explorar novas fronteiras que permitirão que a computação neuromórfica desempenhe um papel fundamental em várias aplicações tecnológicas emergentes. As principais tendências incluem inovações em dispositivos, integração com outras tecnologias, avanços em algoritmos de aprendizagem e uma maior adoção em setores industriais e comerciais.

Desenvolvimento de Dispositivos e Materiais Emergentes

Os dispositivos de memórias resistivas, como os memristores, continuam a ser o principal foco de pesquisa em computação neuromórfica. No entanto, novas tecnologias de materiais emergentes estão a abrir novas possibilidades para melhorar o desempenho e a escalabilidade dos sistemas neuromórficos. A seguir, estão algumas das principais tendências futuras na área de dispositivos e materiais:

 Memristores de nova geração: A pesquisa está a explorar novos tipos de memristores com propriedades melhoradas, como maior estabilidade e maior precisão de programação. O desenvolvimento de memristores baseados em materiais como óxidos metálicos, materiais 2D e materiais ferroelétricos pode oferecer um desempenho superior em termos de resistência e durabilidade [3].

- Tecnologias fotônicas e spintrônicas: A integração de dispositivos fotônicos e spintrônicos com memristores pode levar a novos paradigmas em computação neuromórfica, melhorando a paralelização e a eficiência energética, além de reduzir a latência. O uso de memristores fotônicos, por exemplo, pode permitir uma maior capacidade de processamento em redes neurais de spiking [11].
- Computação com materiais 2D: Materiais bidimensionais, como o grafeno e o disseleneto de tungstênio (WSe₂), estão a ser explorados como possíveis candidatos para dispositivos neuromórficos devido à sua capacidade de integrar memórias e lógica de processamento em um único chip [3].

Integração de Computação Neuromórfica com Tecnologias Emergentes

Com o avanço da computação quântica, IA e IoT, há um forte impulso para integrar a computação neuromórfica com essas tecnologias emergentes. As futuras arquiteturas neuromórficas provavelmente envolverão uma combinação dessas tecnologias, levando a um novo nível de processamento computacional eficiente. Algumas áreas chave incluem:

- Computação Quântica e Neuromórfica: A fusão entre computação quântica e neuromórfica pode revolucionar a forma como processamos informações, permitindo o desenvolvimento de sistemas de *Deep Learning* mais rápidos e eficientes, além de melhorar a capacidade de processamento de dados temporais ou sequenciais. A computação quântica oferece a capacidade de explorar vastos espaços de estados simultaneamente, o que pode ser benéfico para redes neurais de spiking [4].
- Redes Neurais Híbridas: O uso de redes neurais híbridas, que combinam redes neurais profundas com redes de spiking, pode ser uma tendência importante no futuro. Essas redes podem aproveitar as vantagens dos modelos de aprendizagem profundo enquanto mantêm a eficiência dos sistemas neuromórficos, permitindo a implementação de algoritmos de aprendizagem mais poderosos e rápidos [3].

Conclusão

As tendências futuras da computação neuromórfica são vastas e empolgantes, abrangendo desde avanços em dispositivos e materiais até a integração com tecnologias emergentes e novos algoritmos de aprendizagem. O futuro da computação neuromórfica promete uma revolução no campo da IA, levando a sistemas mais rápidos, eficientes e adaptáveis, com aplicações em diversos setores. A pesquisa e desenvolvimento contínuos serão fundamentais para realizar essas possibilidades e garantir que a computação neuromórfica se torne uma tecnologia amplamente adotada.

Referências

- [1] Marcello Massimini et al. "A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior". In: Frontiers in Neuroscience 5 (2011). DOI: 10.3389/fnins. 2011.00118. URL: https://www.frontiersin.org/articles/10.3389/fnins.2011.00118/full.
- [2] Quail Technologies. Day 18: Neuromorphic Computing. https://medium.com/quail-technologies/day-18-neuromorphic-computing-752e6020332d. Accessed: 2024-12-06. 2023.
- [3] Adnan Mehonic et al. "Roadmap to Neuromorphic Computing with Emerging Technologies". In: arXiv preprint arXiv:2407.02353v2 (2024). URL: https://arxiv.org/abs/2407.02353v2.
- [4] Teresa Serrano-Gotarredona and Bernabe Linares-Barranco. "Neuromorphic Engineering: Past, Present, and Future". In: Frontiers in Neuroscience 18 (2024). DOI: 10.3389/fnins.2024. 1511987. URL: https://doi.org/10.3389/fnins.2024.1511987.
- [5] Johannes Schemmel et al. "BrainScaleS A wafer-scale neuromorphic hardware system for large-scale neural modeling". In: Proceedings of the International Joint Conference on Neural Networks (IJCNN) (2010). URL: https://ieeexplore.ieee.org/document/5536970.
- [6] Kwabena Boahen. *A Neuromorphic Treatise*. MIT Press, 2022. ISBN: 9780262046775.
- [7] William Olin-Ammentorp and Nathan Cady. "Biologically-Inspired Neuromorphic Computing". In: *Science Progress* 102.3 (2019). DOI: 10. 1177/0036850419850394.
- [8] S. B. Furber et al. "The SpiNNaker project". In: *Proceedings of the IEEE* 107.4 (2019). DOI: 10. 1109/JPROC.2018.2874565.
- [9] Sumio Ikegawa et al. "Magnetoresistive Random Access Memory: Present and Future". In: *IEEE Transactions on Electron Devices* 67.4 (2020). DOI: 10.1109/TED.2020.2965403.

- [10] IBM. Neuromorphic Computing: The Next Frontier in AI. Acessado em 7 de dezembro de 2024. 2024. URL: https://www.ibm.com/think/topics/neuromorphic-computing.
- [11] Renjie Li et al. *Photonics for Neuromorphic Computing: Fundamentals, Devices, and Opportunities.* LAdv. Mater. n/a, 2312825 (2024). 2023. DOI: 10.1002/adma.202312825. eprint: arXiv: 2311.09767.