



Title	Energy-Efficient Architectural Design of High-Dimensional Computing Paradigm for Edge Devices
Author(s)	Liang, Dehua
Citation	大阪大学, 2024, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/98690
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

Abstract of Thesis

Name (Dehua Liang)	
Title	Energy-Efficient Architectural Design of High-Dimensional Computing Paradigm for Edge Devices (エッジデバイスのための高次元コンピューティングパラダイムの省エネルギーーアーキテクチャ設計)

Abstract of Thesis

With the emergence of the IoT, devices are generating massive data streams. Running big data processing algorithms, e.g., machine learning, on edge devices poses substantial technical challenges due to limited device resources. Compared to the sophisticated machine learning method, the high-dimensional computing paradigm are considered as promising alternative in terms of energy efficiency and robustness. As a novel computing paradigm, how to reduce the hardware cost while maintaining sufficient accuracy performance is still an open problems. The goal of this thesis is thus to provide the strategies to design an energy efficient high-dimensional computing paradigm on edge devices.

In Chapter 3, we discussed the reservoir computing (RC) systems, which is one of the typical high-dimensional computing paradigms. Suffered from the huge memory usage and expensive arithmetic operations, there was still gap between the RC concept to practical implementation. Therefore, this thesis propose a novel RC architecture EnsembleBloomCA, which utilizes cellular automata (CA) and an ensemble Bloom filter to organize an RC system. By adopting CA as the reservoir, it can be implemented using only binary operations and is thus energy efficient. The rich pattern dynamics created by CA can provide more features for the classifier. Applying the ensemble Bloom filters as the classifier, the features provided by the reservoir can be effectively memorized. The novel RC architecture successfully eliminates all floating-point calculation and integer multiplication. Our experiment result demonstrated that 43x and 8.5x reduction is achieved in terms of memory usage and power consumption, while the accuracy performance is maintained.

Although the extreme energy efficiency is achieved with novel RC architecture, the high-dimensional computing are also expected to achieve competitive accuracy. In Chapter 4, we found similar issues should be tackled in hyper-dimensional computing (HDC), which are also considered high-dimensional computing paradigm. To alleviate the huge memory cost during encoding, we propose a novel HDC architecture StrideHD that utilizes the window striding in image classification. It encodes data points to distributed binary hypervectors and eliminates the expensive CiM and iM, which significantly reduces the hardware cost. For the accuracy improvement, we provide the iterative learning mode for the proposed architecture. It also enables HDC systems to be trained and tested using binary hypervectors and achieves very competitive accuracy. The experiment shows that the proposed HDC model achieves extreme memory efficiency (27.6x) with acceptable accuracy under the single-pass mode, while its iterative mode provides competitive performance (11.33% accuracy improvement) and keeping the memory efficiency (8.7x improvement). The iterative retraining can be accomplished within fewer iterations compared to the baseline HDC works.

Besides the trade-off between accuracy and hardware resource, the robustness issue are also arousing the attention of circuit designers. Scaling down the supply voltage is a promising way to reduce the energy consumption, while the aggressive voltage scaling will pose designers several problems such as the performance variation and functional failures. It is a potential solution to tolerate the functional failures by the superior robustness of the high-dimensional computing. Therefore, we introduces the concept of margin enhancement for model retraining and utilizes noise injection to improve the robustness in the proposed HDC framework DependableHD. We additionally propose the dimension-swapping technique, which aims at handling the stuck-at errors induced by aggressive voltage scaling. The experiment shows that under 8% memory stuck-at error, the proposed method exhibits a 2.42% accuracy loss on average, which achieves a 14.1x robustness improvement compared to the baseline HDC. The work also supports the systems to reduce the supply voltage from 430mV to 340mV for both iM and AM, which provides a 41.8% energy consumption reduction while maintaining competitive accuracy.

論文審査の結果の要旨及び担当者

氏 名 (Dehua Liang)		
	(職)	氏 名
論文審査担当者	主査	准教授 塩見 準
	副査	教授 三浦 典之
	副査	准教授 粟野 皓光 (京都大学大学院情報学研究科)
	副査	教授 櫻井 保志

論文審査の結果の要旨

本論文は、エッジデバイスのための高次元コンピューティングパラダイムの省エネルギーーアーキテクチャ設計に関する研究成果をまとめたものである。Deep Neural Network (DNN)等の現在主流の機械学習技術は非常に高い分類精度・汎用性を有する一方、ハードウェアコスト・エネルギー消費が膨大である。本論文は、ハードウェア資源・エネルギー供給源が乏しいエッジ環境に機械学習技術を取り入れることを目指し、ハードウェアコスト・エネルギー効率がDNNと比べて優れる高次元コンピューティング、特にReservoir Computing (RC)とHyper-Dimensional Computing (HDC)の高効率なハードウェア実装手法を提案した。以下に主要な成果を示す。

1. EnsembleBloomCA: Cellular Automata (CA) と Bloom Filter (BF) を用いた省メモリRCシステム

本論文は、CAとBFを併用したRCシステムであるEnsembleBloomCAを提案した。CAをリザバ層として用いることで、2値データのみが演算対象となり省エネルギー化・省コスト化を実現できる。リザバ層が生成する特徴量を効率良く記憶するBFを分類器に用いることで、計算コストの高い浮動小数点演算・整数乗算を排除できた。従来手法と比べて、分類精度を保ちながらメモリ容量を43分の1に、消費電力を8.5分の1に削減できた。

2. StrideHD: Window stridingを用いた2値形式HDCシステム

HDCはRCより高い分類能力を持つ高次元コンピューティング手法であるが、エンコーダに膨大なメモリが必要である。本論文はWindow stridingを活用したStrideHDを提案した。StrideHDは2値形式の高次元特徴量を生成し、従来型HDCシステムのエンコーダに必要であった膨大なメモリを削減できる。分類精度を向上するため、反復学習手法も提案した。従来型のHDCシステムと比較して、単一学習では分類精度を約13%向上しつつ27.6倍のメモリ実装効率を実現した。反復学習では8.7倍のメモリ実装効率を実現し、分類精度がさらに約11%改善した。

3. DependableHD: 低電圧動作のためのディペンダブルHDCシステム

集積回路の低電圧動作は省エネルギー動作のための強力な手法の1つであるが、極端な低電圧化は回路の部分的な誤動作を引き起こし、高次元特徴量に誤りが生じる。本研究は、低電圧動作により生じる高次元特徴量の誤りが誤分類を引き起こさないよう、回路誤動作を想定したHDCシステムの事前学習方式、および実行時回路内検査システムを取り入れたDependableHDを提案した。従来HDCシステムと比べて、分類精度を保ったまま電源電圧を430mVから340mVに削減でき、結果として約42%の消費エネルギー削減を実現した。

EnsembleBloomCAはハードウェアコストの制約が著しく厳しい環境での単純な分類用途に、StrideHDはやや複雑な分類用途に、DependableHDはノイズの多い信頼性のない環境での分類用途に適している。

エッジコンピューティングは今後の情報化社会の持続的発展を支えるキー技術である。以上のように、本論文はエッジ環境の様々な応用先に機械学習技術を実装することを可能にし、情報化社会の持続的発展に寄与するものと期待できる。よって、本論文は博士(情報科学)の学位論文として価値のあるものと認める。