Development of a Hyperdimensional Computing System with Dynamic Hypervector Generation for Resource-Limited Device

Name

Haoyu Tang

Laboratory's name

Computing Architecture Laboratory

Supervisor's name

Professor Yasuhiko Nakashima

Abstract

Hyperdimensional Computing (HDC) is an emerging paradigm for energy-efficient, noise-tolerant learning, inspired by the brain's use of high-dimensional representations. HDC encodes data into dense binary hypervectors and performs inference via simple algebraic operations. However, conventional HDC systems rely on pre-stored level hypervectors and complex encoding pipelines, resulting in high memory overhead and limited scalability, making them unsuitable for edge devices.

This dissertation proposes DystoHD, an area and memory-efficient HDC system tailored for resource-constrained platforms. To address the bottlenecks of conventional encoders, DystoHD integrates Stochastic Computing (SC) to perform dynamic encoding. Real-valued input features are quantized into probabilistic bitstreams and bound with sparse index vectors to generate hypervectors on-the-fly, removing the need for stored templates. A dynamic encoding mechanism further adapts to input statistics, allowing flexible adjustment across data distributions and quantization levels. Additionally, DystoHD reuses position hypervectors within a systolic array, enabling highly parallel execution and reducing memory access overhead.

The architecture is modular and reconfigurable, supporting deployment across a range of classification tasks with minimal tuning. Its general-purpose nature and low hardware cost make it suitable for both energy-constrained and task-adaptive edge systems. Implemented on FPGA, DystoHD achieves a 97.5–99.9% reduction in encoding and training memory usage, a 64% improvement in area-delay product, and maintains classification accuracy within 1–2% of conventional models. These results demonstrate that combining dynamic stochastic encoding with systolic computation enables scalable and efficient HDC deployment. This work contributes a practical solution for real-time and adaptive learning in embedded environments and lays the foundation for future low-power HDC-based intelligent systems.