Dataflow-Aware Memory-Centric Computing: From Synaptic Devices to Practical AI Systems

Bokyung Kim

North Carolina State University

Abstract:

Neuromorphic computing is often motivated by synaptic devices that co-locate neural state and computation. Yet their energy advantage is not determined by device physics alone; it also depends on which data remain local, how intermediate results move, and how effectively the architecture exposes device-level parallelism. My research treats this as a cross-layer dataflow problem and pursues two complementary paths.

For long-term neuromorphic systems, memristive devices support bio-inspired functions such as artificial nociception, while 3D integration increases synaptic density. A 3D memristive convolution processor and an input-stationary ReRAM architecture demonstrate how vertical organization and operand placement can translate these device capabilities into scalable neural acceleration. In parallel, conventional memories provide a near-term path for pressing needs in biomedical intelligence and privacy-preserving learning while emerging devices continue to mature. An SRAM computing-in-memory chip co-designs model and dataflow for multi-class patient-independent seizure classification, while a DRAM processing-in-memory architecture reorganizes differentially private training around its per-sample computations.

Together, these efforts establish dataflow co-design as the bridge between neuromorphic device innovation and practical memory-centric systems for efficient and trustworthy AI.