Topic:

We introduce Neuro.ZERO—a co-processor architecture consisting of a main microcontroller (MCU) that executes scaled-down versions of a deep neural network1 (DNN) inference task, and an

accelerator microcontroller that is powered by harvested energy

and follows the intermittent computing paradigm [76]

Goal:

The goal of the accelerator is to enhance the inference performance of the DNN that is running on the main microcontroller.

Approach

To enable these modes, we propose two sets of algorithms: (1) energy and

intermittence-aware DNN inference and training algorithms, and

(2) a fast and high-precision adaptive fixed-point arithmetic that

beats existing floating-point and fixed-point arithmetic in terms

of speed and precision, respectively, and achieves the best of both.

Result:

To evaluate Neuro.ZERO, we implement low-power image and audio recognition applications and demonstrate that their inference speedup increases by 1.6× and 1.7×, respectively, and the inference accuracy increases by 10% and 16%, respectively, when compared to battery-powered single-MCU systems.

Good:

Figure 2: Neuro.ZERO’s four modes are display clearly by figures

Use a variation of LeNet architecture [70] as the baseline DNN

against two baseline solutions (skip-out and no-skip-out)

Bad:

No val dataset