

## **A scalable, configurable neural network accelerator supporting on-device training**

Jaehwa Kwak, Senior Staff Engineer, LG Electronics, USA

With great success of deep learning in the past years, many manufacturers are trying to apply a machine learning technique to consumer electronics, including refrigerators, washers, and TVs. However, due to the limited computing power of these systems, they often require Internet connectivity of high bandwidth, by which they can offload machine learning tasks to cloud servers. This cloud-backed operation model is of uncontrollable latency, according to dynamic routing paths and server availability.

To address this issue, LG has designed the Neuro-Morphic Processor (NMP), which is a neural network accelerator that can be configured suitable for both low-end and high-end embedded systems. The NMP features (1) the scalable architecture that can adjust the number of cores depending on application requirements, and (2) the on-device training capability that can update the parameters of a neural network at a running device alone using dedicated hardware acceleration logics without being connected to a cloud server. In more detail, the NMP architecture can have up to 128 cores that operate individually or in group, and the number of cores is configurable at the design phase based on the requirement. Each core contains inference and training accelerators, and a RISC-V processor that deals with any non-predefined operations.

The NMP's architecture and products have been verified and proven through both simulation and silicon fabrication with 28nm CMOS technology. The processor can support most of popular neural network models in various fields such as image classification and object detection. The overall processing performance is by orders of magnitude faster compared to those of typical microcontrollers. Currently, the NMP has been adopted in many LG home appliances such as robot vacuum cleaners, washers, and UHD TVs.