Reconfigurable 2D Devices for Artificial Neural Network Applications

Ji-Ung Lee [1]

[1]*SUNY Polytechnic Institute, U.S.A.*

We describe a reconfigurable device that is ideally suited for machine learning hardware. The resulting device architecture is expected to reduce the computational power of advanced neural network algorithms compared those using Si CMOS.Cognitive computing models such as Deep Neural Networks (DNNs) are successful in applications such as computer vision and speech recognition. Within DNN, the most accurate implementations are based on convolutional neural networks (CNNs). To achieve these accuracy levels, CNN-based recognition systems require power hungry GPUs and large memory, which are not suitable for IoT and mobile applications. Recently, XNOR-Nets have attracted a lot of attention as a simpler approximation to CNN with far greater efficiency without the loss of accuracy. Here, we describe a new device based on 2D transition metal dichalcogenide (TMD) semiconductors that can implement logic gates, including XNOR gates, more efficiently compared to Si CMOS. The device relies on its ability to dynamically reconfigure. For example, a single device can implement the PN diode, the MOSFET, and the BJT functions. Furthermore, TMD films allow fine-grained 3D monolithic integration to allow higher device integration. We will describe the property of the reconfigurable logic device and its potential for 3D monolithic integration to address a critical need in machine learning hardware.