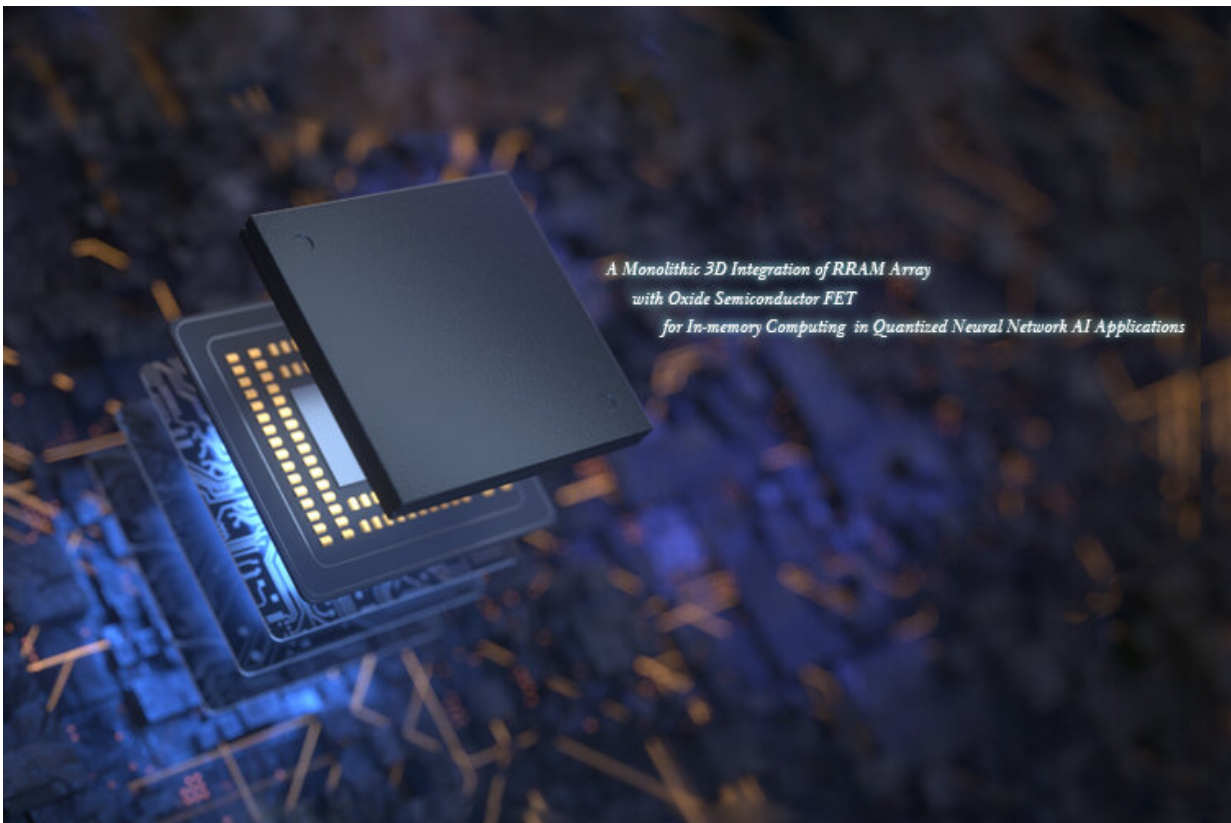


Circular reasoning: Spiraling circuits for more efficient AI

June 15 2020



Researchers from The University of Tokyo create a new integrated 3D-circuit architecture for AI applications with spiraling stacks of memory modules, which may help lead to specialized machine-learning hardware that uses much less electricity. Credit: Institute of Industrial Science, The University of Tokyo

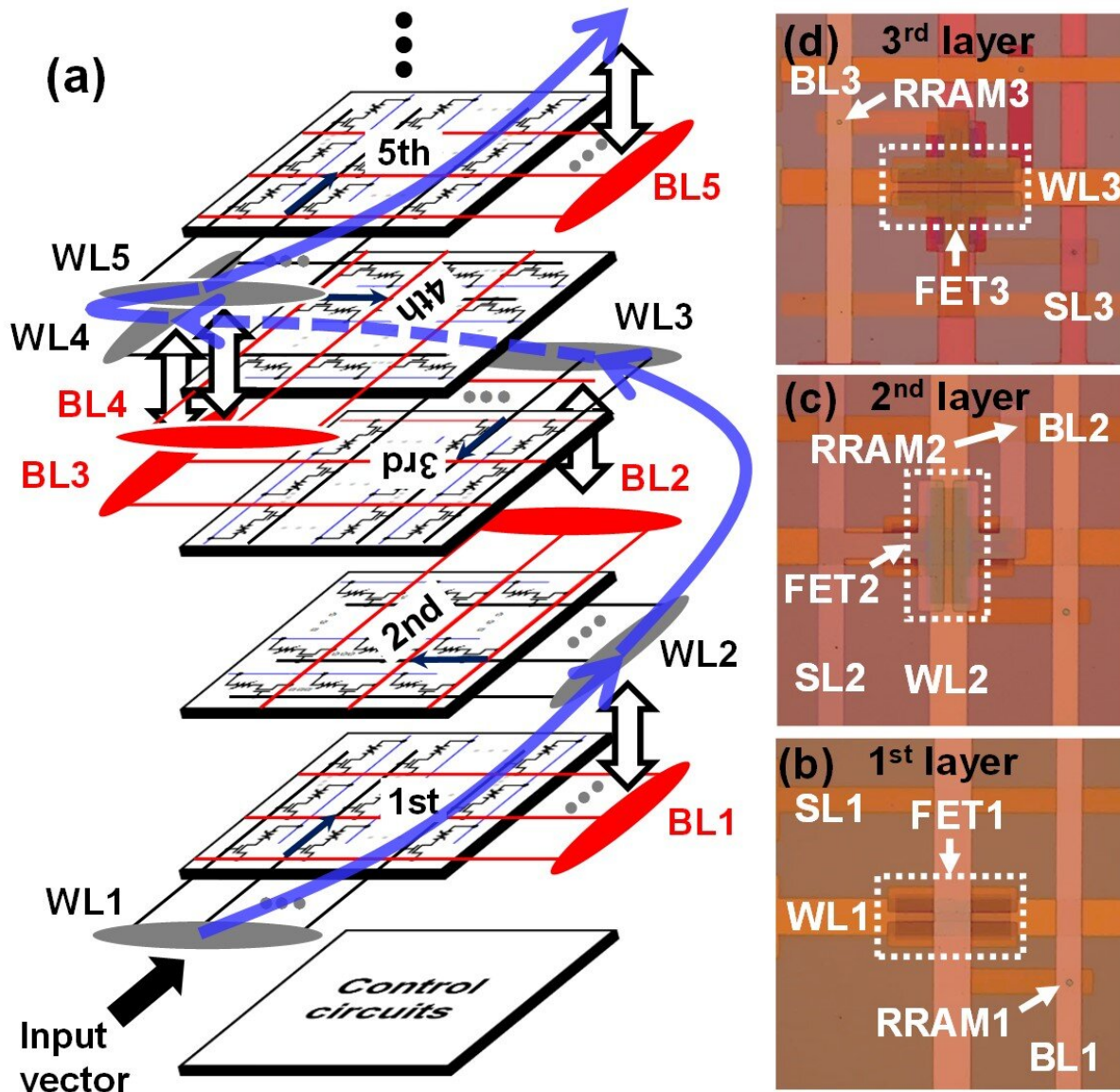
Researchers from the Institute of Industrial Science at the University of

Tokyo designed and built specialized computer hardware consisting of stacks of memory modules arranged in a 3-D spiral for artificial intelligence (AI) applications. This research may open the way for the next generation of energy-efficient AI devices.

Machine learning is a type of AI by which computers are trained with sample data to make predictions for new instances. For example, a smart speaker algorithm like Alexa can learn to understand your voice commands, so it can understand you even when you ask for something for the first time. However, AI tends to require a great deal of electrical energy to train, which raises concerns about adding to climate change.

Now, scientists from the Institute of Industrial Science at The University of Tokyo have developed a [novel design](#) for stacking resistive random-access memory modules with oxide semiconductor (IGZO) access transistor in a three-dimensional spiral. Having on-chip nonvolatile memory placed close to the processors makes the [machine learning](#) training process much faster and more energy-efficient. This is because [electrical signals](#) have a much shorter distance to travel compared with conventional computer [hardware](#). Stacking multiple layers of circuits is a natural step, since training the algorithm often requires many operations to be run in parallel at the same time.

"For these applications, each layer's output is typically connected to the next layer's input. Our architecture greatly reduces the need for interconnecting wiring," says first author Jixuan Wu.



Schematic of the proposed spiral stacking of RRAM array; (b)~(d) Top down microscope image of fabricated IGZO FETs on 1st, 2nd,3rd layer, respectively. Provided by University of Tokyo, 2020 Symposia on VLSI Technology and Circuits

The team was able to make the device even more energy efficient by implementing a system of binarized neural networks. Instead of allowing

the parameters to be any number, they are restricted to be either +1 or -1. This both greatly simplifies the hardware used, as well as compressing the amount of data that must be stored. They tested the device using a common task in AI, interpreting a database of handwritten digits. The scientists showed that increasing the size of each circuit layer could enhance the accuracy of the algorithm, up to a maximum of around 90%.

"In order to keep energy consumption low as AI becomes increasingly integrated into [daily life](#), we need more specialized hardware to handle these tasks efficiently," explains Senior author Masaharu Kobayashi.

This work is an important step toward the "Internet of Things," in which many small AI-enabled appliances communicate as part of an integrated "smart-home."

The study has been presented at the VLSI Technology Symposium 2020.

More information: J. Wu et al. "A Monolithic 3D Integration of RRAM Array with Oxide Semiconductor FET for In-memory Computing in Quantized Neural Network AI Applications." 2020 Symposia on VLSI Technology and Circuits.

Provided by University of Tokyo

Citation: Circular reasoning: Spiraling circuits for more efficient AI (2020, June 15) retrieved 9 April 2024 from <https://techxplore.com/news/2020-06-circular-spiraling-circuits-efficient-ai.html>

| |
|---|
| This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only. |
|---|