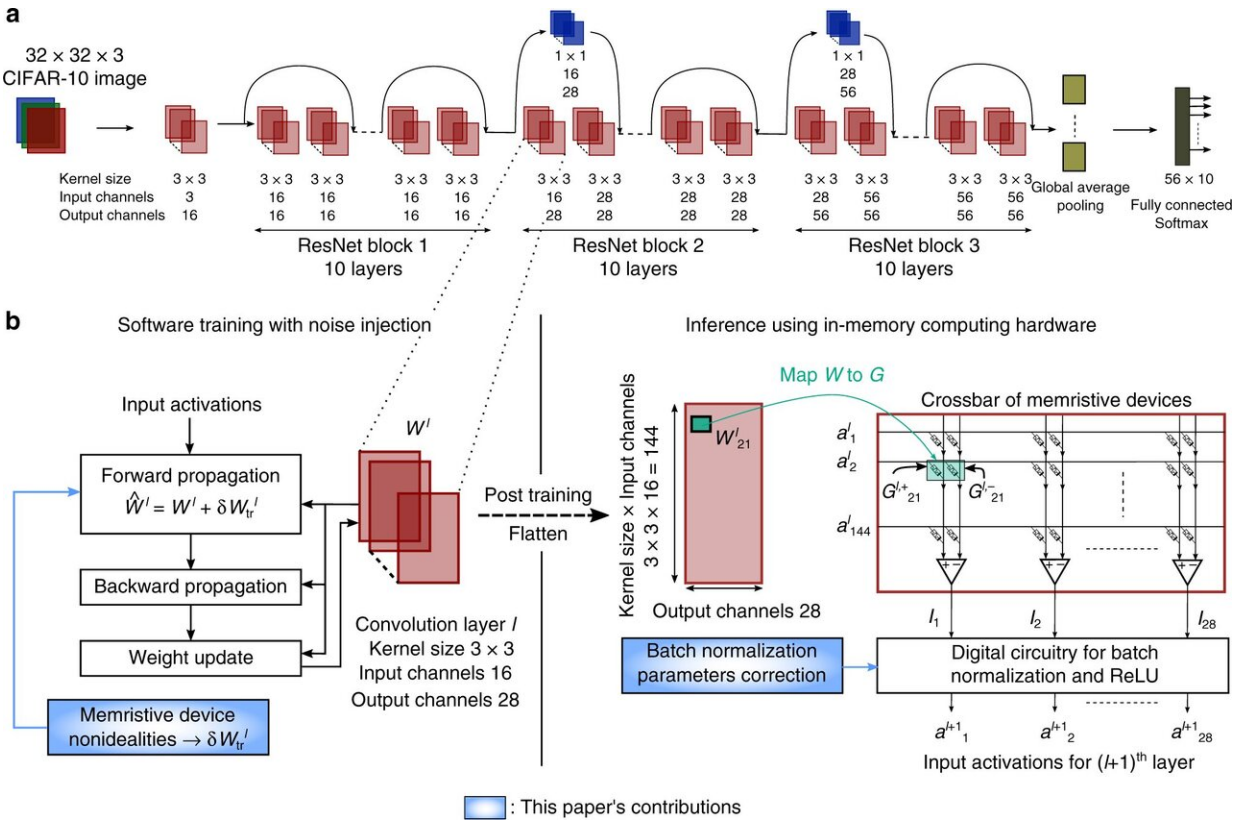


IBM sees AI benefits in phase-change memory

May 19 2020, by Peter Grad



Training and inference methodology. Credit: *Nature Communications* (2020). DOI: 10.1038/s41467-020-16108-9

In a development that holds promise of more sophisticated programming of mobile devices, drones and robots that rely on artificial intelligence, IBM researchers say they have devised a programming approach that

achieves greater accuracy and reduced energy consumption.

AI systems generally employ procedures that divide memory and processing units. This practice means time is consumed transferring data between the two waypoints. The volume of data transfer is massive enough to accrue costly energy tabs.

Nature Communications reported this week that IBM devised an approach that relies on [phase-change memory](#) to execute code faster and cheaper. This is a type of random access memory containing elements that can rapidly change between amorphous and crystalline states, offering performance superior to the more commonly used Flash memory modules. It is also known as P-RAM or PCM. Some refer to it as "perfect RAM" because of its extraordinary performance capabilities.

PCM relies on chalcogenide glass, which has a unique capacity to alter its state when a current passes through. A key advantage of phase change technology, first explored by Hewlett Packard, is that the memory state does not require continuous power to remain stable. The addition of data in PCM does not require an erase cycle, typical of other types of memory storage. Also, since code may be executed directly from memory rather than being copied into RAM, PCM operates faster.

IBM recognized that the growing requirements of operations relying on [deep neural networks](#) in the fields of image and speech recognition, gaming and robotics demand greater efficiencies.

"As [deep learning](#) continues to evolve and demand greater processing power," an IBM team studying solutions posted on a company blog, "companies with large data centers will quickly realize that building more [power plants](#) to support an additional one million times the operations needed to run categorizations of a single image, for example, is just not economical, nor sustainable."

"Clearly, we need to take the efficiency route going forward by optimizing microchips and hardware to get such devices running on fewer watts," the report states.

IBM compared PCM to the human brain, noting that it "has no separate compartments to store and compute data, and therefore consumes significantly less energy."

One drawback with PCMs is the introduction of computational inaccuracies due to read and write conductance noise. IBM addressed that problem by introducing such noise during AI [training sessions](#).

"Our assumption was that injecting noise comparable to the device noise during the training of DNNs would improve the robustness of the models," the IBM report states.

Their assumption was correct. Their model achieved an accuracy of 93.7 percent, which IBM researchers say is the highest accuracy rating achieved by comparable [memory](#) hardware.

IBM says more work needs to be done to obtain even higher degrees of accuracy. They are pursuing studies using small-scale convolutional neural networks and generative adversarial networks, and recently reported on their progress in *Frontiers in Neuroscience*.

"In an era transitioning more and more towards AI-based technologies, including internet-of-things battery-powered devices and autonomous vehicles, such technologies would highly benefit from fast, low-powered, and reliably accurate DNN inference engines," the IBM report says.

More information: Vinay Joshi et al. Accurate deep neural network inference using computational phase-change memory, *Nature Communications* (2020). [DOI: 10.1038/s41467-020-16108-9](https://doi.org/10.1038/s41467-020-16108-9)

S. R. Nandakumar et al. Mixed-Precision Deep Learning Based on Computational Memory, *Frontiers in Neuroscience* (2020). [DOI: 10.3389/fnins.2020.00406](https://doi.org/10.3389/fnins.2020.00406)

IBM Blog: [www.ibm.com/blogs/research/2020 ... in-memory-computing/](https://www.ibm.com/blogs/research/2020-05-19-in-memory-computing/)

© 2020 Science X Network

Citation: IBM sees AI benefits in phase-change memory (2020, May 19) retrieved 20 April 2024 from <https://techxplore.com/news/2020-05-ibm-ai-benefits-phase-change-memory.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.