

Forschungskolloquium Neuromorphe Elektronik

Freitag 20. Mai 2022 -- 14:00 Uhr

Building efficient AI hardware with imprecise nanoscale devices

Bipin Rajendran

Department of Engineering, King's College London, London, UK

AI models based on deep learning have demonstrated super-human performance for a wide variety of tasks - ranging from language translation to protein folding. However, the energy required to optimise some of the state-of-the-art deep learning models on traditional computational systems can be as high as the equivalent of driving a car to the moon and back. Traditional von Neumann systems are not optimised for implementing these data-intensive workloads as the constant shuttling of data between memory and processor units consumes huge amounts of energy and time.

To address this challenge, in-memory computing architectures based on beyond-CMOS nanoscale devices are being explored as matrix computations can be performed efficiently using Kirchhoff's laws, and avoiding data movement. In this talk, I will present our results demonstrating the feasibility of building AI accelerators using nanoscale memristive devices based on chalcogenide alloys that suffer from various nanoscale non-idealities while achieving near software-equivalent accuracies and over 100x improvement in energy-efficiency compared to digital CMOS designs.

Bio: Bipin Rajendran is a Reader in Engineering at King's College London (KCL). He received a B. Tech degree from I.I.T. Kharagpur in 2000, and M.S. and Ph.D. degrees in Electrical Engineering from Stanford University in 2003 and 2006, respectively. He was a Master Inventor and Research Staff Member at IBM T. J. Watson Research Center in New York during 2006-'12 and has held faculty positions in India and the US. His research focuses on building algorithms, devices, and systems for brain-inspired computing. He has co-authored 89 papers in peer-reviewed journals and conferences, one monograph, one edited book, and 59 issued U.S. patents. He is a recipient of the IBM Faculty Award (2019), IBM Research Division Award (2012), and IBM Technical Accomplishment Award (2010). He was elected a senior member of the US National Academy of Inventors in 2019.

Meeting link:

<https://tu-ilmenau.webex.com/tu-ilmenau-en/j.php?MTID=mcb0d4f6659750cf4cd50f3c04d39e8ae>
