EUROPEAN AI FOR FUNDAMENTAL PHYSICS CONFERENCE
EuCAIFCon 2024

Contribution ID: **214** Type: **Flashtalk with Poster**

Deep support vector data description models on an analog in-memory computing platform for real-time unsupervised anomaly detection.

Tuesday, 30 April 2024 18:05 (3 minutes)

One of the most important challenges in High Energy Physics today is to find rare new physics signals among an abundance of Standard Model proton-proton collisions, also known as anomaly detection. Deep Learning (DL) based techniques for this anomaly detection problem are increasing in popularity [1]. One such DL technique is the Deep SVDD model [2], which shows great results when applied to the anomaly detection problem in particle physics [3]. These Deep SVDD models are relatively computational inexpensive, which is promising for real-time implementation in particle detectors such as the ATLAS detector at LHC. For the Atlas detector specifically the incoming event rate is 40MHz and is brought down to a final collection rate of 300Hz using the three staged trigger system. Therefore, any model deployed within this trigger system, especially the first level trigger, requires a high throughput.

The progress of these DL techniques could be further improved by utilizing special neuromorphic hardware, i.e. hardware specifically designed to accelerate machine learning tasks. A promising candidate is the Analog In-Memory Computing (AIMC) platform, such as the HERMES core [4]. In this hardware the Matrix Vector Multiplication (MVM), a prominent part of inference in machine learning tasks, is done in-memory. This mitigates the von Neuman bottle neck giving access to potential faster and more energy efficient computations.

In this work we investigate an implementation of Deep SVDD models on an AIMC platform for unsupervised new physics detection at 40 MHz, using a dataset specifically generated for unsupervised new physics detection at 40 MHz at the LHC [5]. First, we investigated the energy consumption and throughput of these Deep SVDD on CPUs and GPUs, which we then compared to the estimated performance of an AIMC platform [4]. We predict that the state-of-the-art AIMC is up to 1000x more energy efficient than CPUs and GPUs at a throughput that is 10x faster than CPUs and GPUs [6]. This suggest high potential for faster and more sustainable anomaly detection in fundamental physics and beyond.

[1] G. Kasieczka et al., The LHC Olympics 2020 a community challenge for anomaly detection in high energy physics, Rep. Prog. Phys. 84, 124201 (2021), doi:10.1088/13616633/ac36b9.
[2] Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S.A., Binder, A., Müller, E. Kloft, M.. (2018). Deep One-Class Classification. Proceedings of the 35th International Conference on Machine Learning, in Proceedings of Machine Learning Research 80:4393-4402 Available from https://proceedings.mlr.press/v80/ruff18a.html.
[3] S. Caron, L. Hendriks, and R. Verheyen, "Rare and Different: Anomaly Scores from a combination of likelihood and out-of-distribution models to detect new physics at the LHC," SciPost Phys. 12, 77 (2022).
[4] R. Khaddam-Aljameh, M. Stanisavljevic, J. Fornt Mas, G. Karunaratne, M. Brändli, F. Liu, A. Singh, S. M. Müller, U. Egger, A. Petropoulos, T. Antonakopoulos, K. Brew, S. Choi, I. Ok, F. L. Lie, N. Saulnier, V. Chan, I. Ahsan, V. Narayanan, S. R. Nandakumar, M. Le Gallo, P. A. Francese, A. Sebastian, and E. Eleftheriou, "HERMES-core—a 1.59-TOPS/mm2 PCM on 14-nm CMOS in-memory compute core using 300-ps/LSB linearized CCO-based ADCs," IEEE Journal of Solid-State Circuits 57, 1027–1038 (2022).
[5] E. Govorkova, E. Puljak, T. Aarrestad, M. Pierini, K. A. Wo´zniak, and J. Ngadiuba, "LHC physics dataset for unsupervised New Physics detection at 40 MHz," Scientific Data 9, 1–7 (2022).
[6] Dominique J. Kösters, Bryan A. Kortman, Irem Boybat, Elena Ferro, Sagar Dolas, Roberto Ruiz de Austri, Johan Kwisthout, Hans Hilgenkamp, Theo Rasing, Heike Riel, Abu Sebastian, Sascha Caron, Johan H. Mentink; Benchmarking energy consumption and latency for neuromorphic computing in condensed matter and particle physics. APL Mach. Learn. 1 March 2023; 1 (1): 016101. https://doi.org/10.1063/5.0116699

Primary authors: SEBASTIAN, Abu (IBM Research Europe - Zürich); KORTMAN, Bryan (Nikhef); KOSTERS, Dominique (Radboud Universiteit, IBM, IMM,); FERRO, Elena (ETH Department of Information Technology and Electrical Engineering); HILGENKAMP, Hans (University of Twente, Faculty of Science and Technology); RIEL, Heike (IBM Research Europe - Zürich); BOYBAT, Irem (IBM Research Europe - Zürich); KWISTHOUT, Johan (Radboud University, Donders Institute for Brain, Cognition and Behaviour); MENTINK, Johan (Radboud University, Institute for Molecules and Materials); RUIZ DE AUSTRI, Roberto; DOLAS, Sagar (Surf Cooperation, Innovation Team); CARON, Sascha (Nikhef); RASING, Theo (Radboud University, Institute for Molecules and Materials)

Presenter: KOSTERS, Dominique (Radboud Universiteit, IBM, IMM,)

Session Classification: 3.3 Hardware acceleration, FPGAs & Uncertainty quantification

Track Classification: Session B