

Efficient U-Net Model on FPGAs: A VitisAI Approach with Xilinx ZCU104

Thursday 10 July 2025 14:40 (20 minutes)

Field programmable gate arrays (FPGA) are reconfigurable and serve as prototyping platforms for application-specific integrated circuits (ASIC). FPGA are composed of configurable logic blocks and programmable interconnects and allow rapid testing of digital designs. Many challenges in medical field, in agriculture and in security are solved by the implementation of artificial intelligence techniques in high performance computing units with an inherently higher quantitative accuracy than any user. The implementation of these techniques on portable computing units (including FPGA) has gained great attention in the new field of tinyML. Here, we present an efficient FPGA implementation of UNET architecture segmentation of unstained brightfield cell images achieving a 97% accuracy. Quantization and pruning are used by effectively reducing the model size without a significant decrease in the accuracy. This work indicates the importance of combining algorithmic based compression techniques with hardware-aware optimizations.

Keywords: FPGA, tinyML, deep learning, Medical image segmentation.

References:

- [1] Ng, W. S., Goh, W. L., & Gao, Y. (2024, May). High Accuracy and Low Latency Mixed Precision Neural Network Acceleration for TinyML Applications on Resource-Constrained FPGAs. In 2024 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 1-5). IEEE.
- [2] Uka, A., Topalli, G., Hoxha, J., & Vrana, N. E. (2021). FPGA Implementation of Filters in Medical Imaging. In BIODEVICES (pp. 195-200).
- [3] Prakash, S., Callahan, T., Bushagour, J., Banbury, C., Green, A. V., Warden, P., ... & Reddi, V. J. (2023, April). Cfu playground: Full-stack open-source framework for tiny machine learning (tinymml) acceleration on fpgas. In 2023 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS) (pp. 157-167). IEEE.

Primary authors: UKA, Arban (Epoka University); Prof. KARLIK, Bekir (Epoka University); ALUSHLLARI, Griselda (Epoka University); Mr STERGU, Igli (Epoka University)

Presenter: UKA, Arban (Epoka University)

Session Classification: Poster Session 3

Track Classification: S06 –Interdisciplinary Physics, Mathematical and Computational Methods