

INTISARI

PERANCANGAN AKSELERATOR *ARTIFICIAL NEURAL NETWORK* (ANN) SEDERHANA BERBASIS *SYSTOLIC ARRAY* PADA FPGA

Oleh

Jihan Aanisa Zulfaani

21/480765/PA/20904

Perkembangan teknologi mendorong penerapan *Artificial Neural Network* (ANN) secara luas pada berbagai aplikasi kecerdasan buatan. Namun, proses inferensi ANN memerlukan operasi perkalian dan penjumlahan dalam jumlah besar yang membutuhkan sumber daya komputasi tinggi, sehingga menjadi tantangan ketika diterapkan pada perangkat dengan keterbatasan sumber daya. Oleh karena itu, diperlukan akselerator perangkat keras yang mampu mempercepat proses inferensi ANN secara efisien.

Pada penelitian ini, diimplementasikan akselerator ANN berbasis *systolic array* 6×6 pada FPGA PYNQ-Z2 dengan model ANN berarsitektur 5–5–2, fungsi aktivasi *sigmoid*, dan representasi *fixed-point* Q6.10. Hasil pengujian menunjukkan bahwa hasil inferensi akselerator konsisten dengan model referensi berbasis Python. Waktu eksekusi inferensi pada FPGA mencapai sekitar 0,0003 detik per data, lebih cepat dibandingkan eksekusi CPU sebesar 0,0006 detik per data. Meskipun waktu total meningkat pada jumlah data besar akibat *overhead* DMA, performa komputasi inti FPGA tetap lebih cepat dan stabil, sehingga akselerator ANN berbasis *systolic array* berpotensi digunakan pada sistem *edge-device*.

Kata Kunci: *systolic array*, akselerator ANN, FPGA, PYNQ Z2

ABSTRACT

**ARTIFICIAL NEURAL NETWORK (ANN) BASED ON SYSTOLIC ARRAY
ON FPGA**

By

Jihan Aanisa Zulfaani

21/480765/PA/20904

The development of technology has encouraged the widespread application of Artificial Neural Networks (ANNs) in various artificial intelligence applications. However, the ANN inference process requires a large number of multiplication and addition operations that demand high computational resources, which poses challenges when implemented on resource-constrained devices. Therefore, hardware architectures that can accelerate the ANN inference process efficiently and reliably are required.

In this research, a 6×6 systolic-array-based ANN accelerator is implemented on the PYNQ-Z2 FPGA using a 5–5–2 ANN architecture with sigmoid activation and fixed-point Q6.10 representation. The system is integrated into a System-on-Chip (SoC) design using AXI-Stream and AXI-Lite interfaces with data transfer handled through Direct Memory Access (DMA). Experimental results show that the inference outputs produced by the accelerator are consistent with those of the Python-based reference model. The FPGA inference execution time reaches approximately 0.0003 seconds per data sample, which is faster than CPU execution at around 0.0006 seconds per data sample. Although the total execution time increases for larger datasets due to DMA overhead, the core computation performance of the FPGA remains faster and stable.

Keywords: *systolic array, ANN accelerator, FPGA, PYNQ Z2*