Oral S14

HW Acceleration for Biomedical and BioInformatics Applications

Date/Time	8/3(四) 11:30-12:30
Chair(s)	魏一勤／長庚大學電機工程學系 陳元賀／長庚大學電子工程學系

S14.1 | 11:30—11:45

A High-Throughput FPGA Accelerator of FM-index Based Paired-end Short-read Mapping for Next-Generation Sequencing

Chung-Hsuan Yang[1], Yi-Chung Wu[1], Yen-Lung Chen[1], Chao-Hsi Lee[2], Jui-Hung Hung[2,3], Chia-Hsiang Yang[1,2]

[1] Graduate Institute of Electronics Engineering, National Taiwan University

[2] GeneASIC Technologies Corp.

[3] Graduate Institute of Computer Science and Engineering, National Yang Ming Chiao Tung University

This paper presents an Ferragina-Manzini index (FM-index) based paired-end short-read mapping hardware accelerator. Four techniques are proposed to significantly reduce the number of memory accesses and operations. First, an interleaved data structure is proposed to reduce the processing time by 51.8% by leveraging the data locality. Second, the boundaries of possible location candidates can be retrieved from a lookup table within only one memory access, reducing the number of DRAM accesses by 60%. Third, an additional filtering step is added to skip the time-consuming repetitive location candidates conditionally. Lastly, an early termination method is proposed to terminate the process if a pair of high enough scores is detected. Overall, the computation time is reduced by 92.6%. Compared to prior arts, realized on a Xilinx Alveo U250 FPGA, this work delivers a 1.7-to-18.6x higher throughput in a memory-efficient way. Paired-end short-read mapping is exploited to achieve the highest 99.3% accuracy on true human DNA dataset.

S14.2 | 11:45—12:00

A Trainable Deep Learning Accelerator for Seizure Detection Application

Yen-Hsing Tsai, Ming-Yueh Ku, Shuenn-Yuh Lee, and Chou-Ching Lin
Department of Electrical Engineering, National Cheng-Kung University; National Cheng-Kung University Hospital

This paper proposes a seizure detection algorithm and a Trainable Deep Learning Accelerator (TDLA). The algorithm includes a Discrete Wavelet Transform (DWT) and a z-score normalization for data pre-processing and an 8-layer Convolution Neural Network (CNN) model for classification. The algorithm provides 99.49% accuracy, 99.58% sensitivity, and 99.37% specificity on floating-point operations while 99.31% accuracy, 99.53% sensitivity, and 99.01% specificity on 16-bit fixed-point operations with fusing part of layers to reduce 26.14% operations. The algorithm can be implemented on the TDLA through instructions and configurations generated by the self-defined compiler. Moreover, the CNN model can also be fine-tuned on the TDLA through the Stochastic Gradient Descent with Momentum (SGDM) optimizer to make the model personalized. More importantly, the TDLA is not only trainable but also programmable, so it can execute various CNN models for different wearable biomedical applications.

S14.3 | 12:00—12:15

An AAMI Standard Arrhythmia Classification with a High-Performance CNN Based AI Accelerator

Wei-Cheng Tseng, Shuenn-Yuh Lee, and Ju-Yi Chen
Department of Electronic Engineering, National Cheng Kung University

This paper presents an arrhythmia classification system based on the Association for the Advancement of Medical Instrumentation (AAMI) standard. The algorithm includes a general electrocardiography (ECG) data preprocessing flow that can be applied to various ECG databases. This paper proposes a Convolution Neural Network (CNN) based model to achieve arrhythmia classification. In addition, the prediction accuracy of the model is improved by incorporating R-peak interval features. The proposed model is trained and tested using the MIT-BIH database, achieving an accuracy of 98.5%. Furthermore, this paper provides a custom artificial intelligence (AI) accelerator for hardware implementation, which utilizes hybrid stationary techniques to achieve high-performance computing. The accelerator design is implemented by the Xilinx PYNQ-Z2. The implementation consumes 6458 lookup tables (LUTs), 2656 flip-flops, 5KB block random access memory (BRAM), and 24 digital signal processors (DSPs). The power consumption consists of a static power of 0.106 W and a dynamic power of 0.002 W in 10-MHz operation frequency. And the accelerator can execute inference in only 0.683 ms.

S14.4 | 12:15—12:30

High Resolution Time-to-Digital Converter Implemented by FPGA for Time-of-Flight Positron Emission Tomography

Yeuk-Ho Lai, Don-Gey Liu, and Ching-Hwa Cheng
Department of Electronic Engineering, Feng Chia University

In this paper, a proposed design for the Vernier Delay Line (VDL) is used to make the difference in the propagation times of the two delay paths very small so that the resolution of the related Time-to-Digital Converter (TDC) can be improved. By this method, the slightly different path delay times can be generated automatically by our algorithm rather than by manual tuning. In this study, a Zed board FPGA was employed as the development platform. And the smallest timing resolution of our design can reach down to 3ps which corresponds to a spatial resolution of 0.9 mm for tomography images.