FPGA-based Cloudification of ECG Signal Diagnosis Acceleration

Dongkyu Lee[1], SeungMin Lee[1], and Daejin Park[1*]
[1]School of Electronic and Electrical Engineering, Kyungpook National University
*Correspondence to: boltanut@knu.ac.kr

***Abstract*—Recently, studies to analyze heart disease using ECG signals are emerging. The proposed platform generates multiple reference signals trained for individuals in real time by reducing the learning time. The data in the cluster is compressed by linear approximation to speed up diagnosis and reduce memory usage, allowing more diagnosis to be performed with limited resources. Platforms using FPGA can accelerate ECG signal diagnosis by adding hardware. As a result of diagnosing ECG signals of 10 people using the processor and accelerator, the execution time when using the accelerator was 71% lower than that when using the processor.**

***Keywords*—FPGA acceleration; co-design; cloudification; electrocardiogram; linear approximation;**

I. Introduction

Healthcare, one of the fields where IoT is applied, is being studied as people are more interested in health due to an increase in life expectancy. Among the vital signals handled in healthcare, the electrocardiogram (ECG) signal, which is the most important signal, can detect arrhythmia, which is the most fatal disease for humans. Because the arrhythmia signal to be measured using ECG signal is rare signal, the ECG signal must be measured for several hours or longer. The embedded device, which collects data from a sensor and performs simple calculations, has a small memory size and low performance because the supply power is limited. Therefore, it is difficult to perform ECG signal diagnosis in which a large amount of data is accumulated over a long period of time at the embedded device. There are studies to reduce execution time by improving the ECG signal diagnosis algorithm and to perform diagnosis using an FPGA accelerator in an embedded device [1].

There are studies on ECG diagnosis using cloud computing to compensate for the low performance of edge devices [2]. Cloud computing is a system that transmits data collected by sensors to a server so that a server with many resources can calculate data at a high speed. In this paper, we propose a cloudification platform that accelerates ECG diagnosis based on FPGA. We adopted the cloud computing method to diagnose ECG signals from multiple people at the same time [3]. The proposed platform uses an improved ECG reference signal learning algorithm to reduce the memory usage required for diagnosis. The embedded device compresses before transmitting data according to the trained reference signal. We diagnosed ECG signal from 10 individuals using FPGA in the MIT-BIH arrhythmia database on the platform and measured the power consumption and execution time [4].

This study was supported by the BK21 FOUR project funded by the Ministry of Education, Korea (4199990113966, 10%), Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2019R1A2C2005099, 10%), and Ministry of Education (NRF-2018R1A6A1A03025109, 10%, NRF-2020R1I1A1A01072343, 20%). This work was partly supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2021-0-00944, Metamorphic approach of unstructured validation/verification for analyzing binary code, 50%)

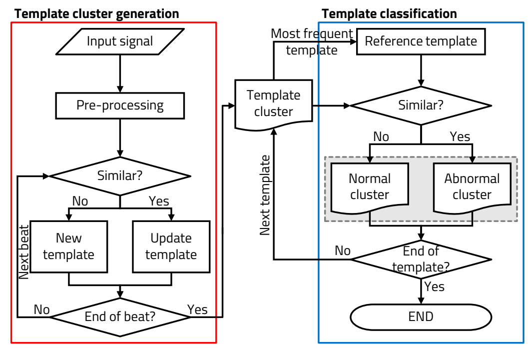

Fig. 1. The flow chart of the template cluster generation algorithm

II. Proposed Architecture and Experimentation

Wearable devices that measure ECG signals have a small memory size, low performance, and small battery capacity. To perform a diagnosis that requires lots of computation, the devices transmit the measured ECG signal to the cloud server. Existing ECG diagnosis algorithms classify an input ECG signal using a reference signal learned in advance that has different detection rate for each person. Learning ECG signals which is different for each person to make single reference signal takes a lot of memory and a long time. In the proposed platform, the time required for learning the reference signal is reduced by decreasing the fidelity using weighted average learning. Furthermore, the accuracy of diagnosis is improved by using template clusters that collects the normal and abnormal reference signals. Fig. 1 shows the template cluster learning algorithm used in the proposed platform [5].

The input ECG raw data is pre-processed into beat data centered on R-peak through filtering and QRS complex detection [6]. The first learning beat entered for template cluster initialization becomes a template. From the second beat, the entire wave and the P wave are compared with the entire template using the Pearson similarity shown in Eq. 1. When the similarity exceeds the update threshold, the template with the highest similarity is updated using Eq. 2. C_T is the data of the template, C_w is the weight of the template, and S is the data of the beat to be updated. After updating the data in the template, the weight is increased by 1. If the similarity does not exceed the threshold, a new template is created and the weight is set to 1.

$$\rho(X,Y) = \frac{1}{N-1}\sum_{i=1}^{N}\frac{(X_i-\mu_X)}{\sigma_X}\cdot\frac{(Y_i-\mu_Y)}{\sigma_Y} \tag{1}$$

$$C_T = \frac{C_w \times C_T + S}{C_w + 1} \tag{2}$$

When the template cluster is generated, the signal with the highest weight is selected as the reference normal template using ECG's characteristic. The remaining templates are calculated for Pearson similarity with the reference normal template and stored in a normal cluster when the similarity exceeds separate threshold, and in an abnormal cluster when the similarity does not exceed separate threshold. Using linear approximation (LA), stored normal/abnormal clusters are compressed.

Fig. 2 (a) shows the operation of LA with minimal memory usage. When performing LA, N_{bit} was set as the maximum distance of vertices to prevent the distance between vertices from being too far apart. The base matrix $C_0(i,j)$ stores the error value between the line that connects i to j and the beat from i to j. Each row of the cost matrix stores the minimum error value when there are v vertices between 1 and j. The X array is a one-dimensional array with N_{bit} rows, and the Y array is a one-dimensional array with N_{bit} columns.

Fig. 3 shows the diagnosis algorithm using FPGA. The cluster created at the learning process is stored in the FPGA. The ECG beat measured at the sensor is compressed using the index of each cluster and transmitted to the FPGA. For the received data, the average of the similarity with the normal cluster and the average of the similarity with the abnormal cluster are calculated. The diagnosis of the ECG signal is determined by the higher similarity among the normal/abnormal cluster's similarity.

In this paper, we cloudificated the FPGA-based ECG signal diagnosis acceleration to provide an optimized diagnosis service for individuals, shown in Fig. 4. The platform is divided into software running on the processor and hardware that is synthesized in real time on the FPGA to accelerate computation. The clusters that created for the individual in the learning stage is transferred to the FPGA. The FPGA is synthesized including the received clusters and diagnostic algorithm. In the diagnosis process, personal data is preprocessed in the processor and transmitted to the FPGA in the form of an ECG matrix.

For the experiment of the proposed platform, a Xilinx Alveo U200 FPGA and two Intel Xeon Bronze 3204 processors are

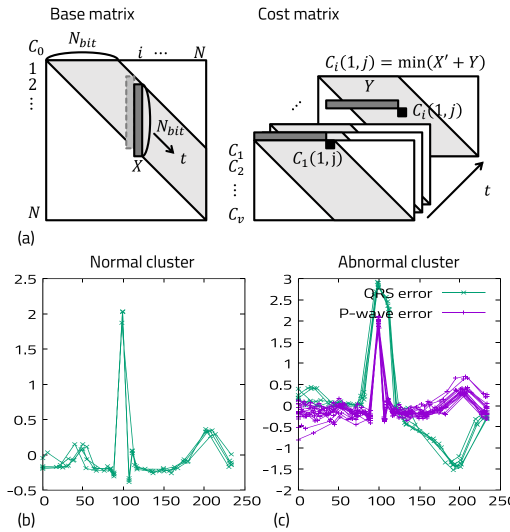

Fig. 2. (a) Linear approximation with minimal memory usage, and result of linear approximation for (b) the normal and (c) the abnormal cluster

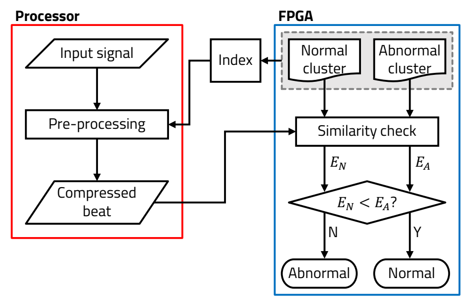

Fig. 3. The flow chart of the diagnosing algorithm

configured in the server. The Xeon processor has a maximum clock speed of 1.9 GHz. The Alveo U200 card based on Xilinx's UltraScale architecture includes 892000 number of look-up tables (LUTs) and 100 MHz clock sources. We diagnosed ECG signals in 10 patients with 1987 beats on the MIT-BIH arrhythmia database. Fig. 5 shows the execution time of diagnostics using software and hardware accelerators. Using only processor without accelerator, represented as SW, diagnosis takes 11.39 seconds. Accelerators are divided into HW 1 to HW 10 depending on how many diagnosis can be performed at the same time. The execution time when using HW 1 is reduced to 3.33 seconds, which is 0.29% of the software execution time. The more work the accelerator performs, the less execution time takes. Fig. 6 shows the number of LUTs required when implementing an accelerator.

We estimates the total power consumption using Eq. 3. The Alveo U200 has 892000 LUTs, that represented by N, and it consumes 100 W of power on average (P_{HW}). The processor consumes 85 W on average (P_{SW}). We assumed that the average power consumption of the accelerator is proportional to the number of used LUTs, represented by

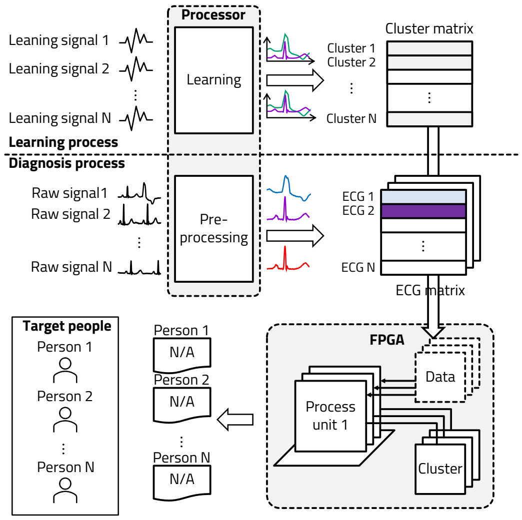

Fig. 4. Execution of the proposed ECG diagnosis platform

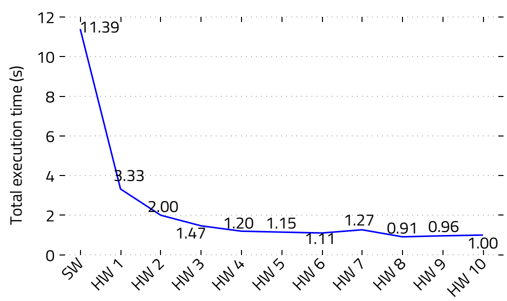

Fig. 5. The total execution time according to the number of hardware running concurrently

n_{LUTs}. Fig. 7 shows the power consumption calculated by Eq. 3 using the execution time and the number of LUTs. The accelerators consume more instantaneous power than software, but their execution time is shorter, reducing overall power consumption. The larger the accelerator, the shorter execution time. However, when the platform uses the accelerator that is larger than HW 9, the instantaneous power consumption is larger than the reduced execution time, resulting in increased total power consumption. In this experiment, HW 8, which can diagnose 8 beats in parallel, consumes the least power.

$$E = (P_{HW} \times \frac{n_{LUTs}}{N} + P_{SW}) \times t \tag{3}$$

III. Conclusion

In this paper, a hardware acceleration system is cloudificated to energy-efficiently diagnose ECG signals in IoT environments where large-scale data is generated. The proposed platform uses simple weighted learning to reduce the training time, and uses linear approximation to decrease memory usage. Because the learning time and the memory required by the template is small, the platform can serve multiple patients at the same time, optimizing reference signals for individual patients. The ECG signal detected by the sensor is compressed based on the learned template cluster and transmitted to the accelerator. Accelerator that implemented in real time controls the number of diagnoses that can be executed at the same time according to the number of patients, the precision of diagnostics required, and the state of server, thereby obtaining flexibility in software execution and high performance of the hardware. As a result of diagnosing 10 people's ECG signal, using an accelerator takes 3.33 seconds, which is 71% less than the 11.39 seconds, which is the time taken using software.

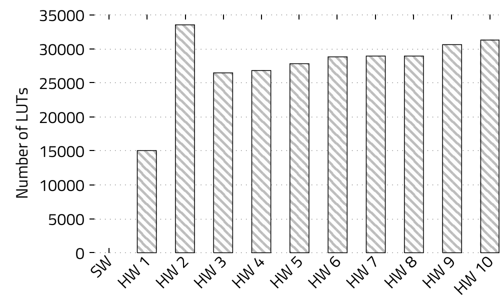

Fig. 6. The number of used LUTs according to the number of hardware

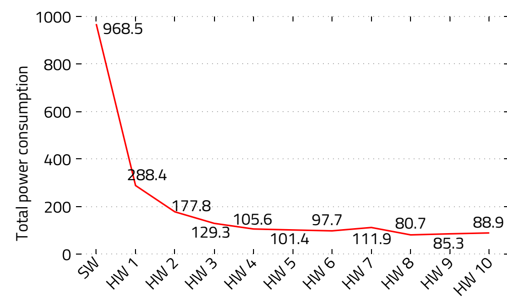

Fig. 7. The total power consumption according to the number of hardware

References

[1] D. Azariadi, V. Tsoutsouras, S. Xydis, and D. Soudris, "Ecg signal analysis and arrhythmia detection on iot wearable medical devices," in *2016 5th International conference on modern circuits and systems technologies (MOCAST)*. IEEE, 2016, pp. 1–4.

[2] X. Gu, Y. Zhu, S. Zhou, C. Wang, M. Qiu, and G. Wang, "A real-time fpga-based accelerator for ecg analysis and diagnosis using association-rule mining," *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 15, no. 2, pp. 1–23, 2016.

[3] D. Lee, H. Moon, S. Oh, and D. Park, "miot: Metamorphic iot platform for on-demand hardware replacement in large-scaled iot applications," *Sensors*, vol. 20, no. 12, p. 3337, 2020.

[4] G. B. Moody and R. G. Mark, "The mit-bih arrhythmia database on cd-rom and software for use with it," in *[1990] Proceedings Computers in Cardiology*. IEEE, 1990, pp. 185–188.

[5] S. Lee and D. Park, "Advanced real-time dynamic programming in the polygonal approximation of ecg signals for a lightweight embedded device (in press)," *IEEE Access*, 2021.

[6] S. Lee, Y. Jeong, D. Park, B.-J. Yun, and K. H. Park, "Efficient fiducial point detection of ecg qrs complex based on polygonal approximation," *Sensors*, vol. 18, no. 12, 2018. [Online]. Available: https://www.mdpi.com/1424-8220/18/12/4502