

ASIC 기반 저면적 저전력 Multiply-Accumulate (MAC) 구현

이유진, 임지환, 유호영*
충남대학교 전자공학과

ASIC-Based Low-Area Low-Power Multiply-Accumulate (MAC) Implementation

Yujin Lee, Jihwan Lim, Hoyoung Yoo*

Electronics Engineering Department

Chungnam National University

E-mail: yjlee01.cas@gmail.com, jhlim.cas@gmail.com, *hyyoo@cnu.ac.kr

Abstract

In this paper, we design an optimal structure by comparing four types of MAC units in terms of area and power. To calculate the CNN algorithm, convolution operation which is calculated by accumulating multiplication and addition is required. It can be composed of MAC units. RCA, CLA were used as additives used in the implementation, AM and MBM were used as multipliers, and registers were used as accumulators. The experiment was conducted on 4-bit and 8-bit inputs. Using Synopsis Design Compiler, it was synthesized with an operating frequency of 140 MHz and a CMOS 180 nm process. For both 4-bit and 8-bit inputs, the MAC unit of the AM+RCA+Register is suitable for the low area, and the MAC unit of the MBM+RCA+Register is suitable for low power.

I. 서론

Internet of Thing(IoT)와 같은 저성능 프로세서에서 주로 사용되는 인공지능 알고리즘인 Convolution Neural Network (CNN) 알고리즘을 동작 시키기 위해서는 저면적과 저전력의 하드웨어 구성이 필요하다[2]. CNN 알고

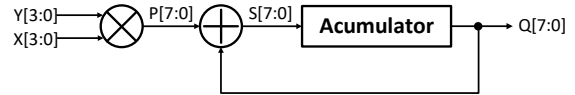


그림 1. MAC unit

리즘은 곱셈과 덧셈 연산을 순차적으로 진행하는 Convolution 연산이 추가 되어 처리된다. Convolution 연산은 하드웨어로 Multiply-Accumulate (MAC) unit 을 통해 구현할 수 있다 [1]. 따라서, 본 논문에서는 2 종류의 덧셈기와 2 종류의 곱셈기 그리고 레지스터의 조합으로 MAC unit 을 설계하고 면적과 전력, 소요 시간 측면에서 비교하여 저면적과 저전력에서 가장 효율적인 구조를 설계한다.

II. 배경 지식

2.1 MAC unit

MAC unit 은 곱셈과 덧셈을 누적하여 수행하는 연산 장치로 그림 1 과 같이 곱셈기와 덧셈기 그리고 누산기로 구성된다[1].

2.2 Ripple Carry Adder (RCA)

RCA 는 그림 2 와 같이 직관적이고 간단한 덧셈기의 구조이다. 첫번째 Full Adder(FA)에서 생성된 캐리가

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2022-001170), National R&D Program through the National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT (2020M3H2A1078119). The EDA tool was supported by the IC Design Education Center (IDEC), Korea.

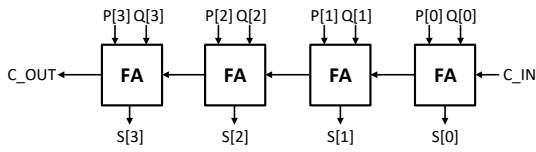


그림 2. Ripple Carry Adder

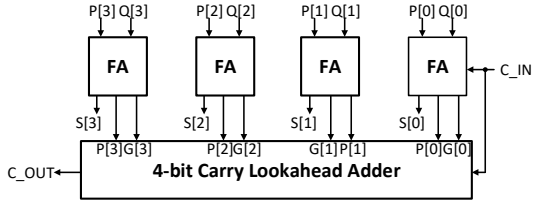


그림 3. Carry Lookahead Adder

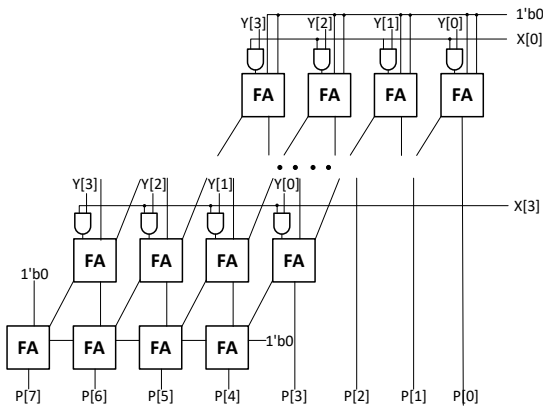


그림 4. Array Multiplier

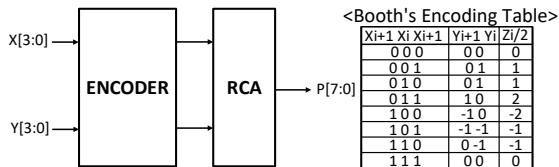


그림 5. Modified Booth's Multiplier

다음 단 FA 의 입력으로 전파되며 덧셈이 진행된다. 이 때 각 FA 에서 출력된 S 로 최종 합을 출력한다.

2.3 Carry Lookahead Adder (CLA)

CLA 는 그림 3 과 같이 각 비트에 대한 캐리를 즉시 계산하지 않고, 최종 캐리를 미리 계산하여 출력하는 덧셈기다. 따라서 캐리 전파에 소요되는 시간을 단축할 수 있다.

2.4 Array Multiplier (AM)

AM 은 그림 4 와 같이 피연산자의 각 비트에 대해 계산하는 AND Gate 와 이를 덧셈하는 FA 로 구성된다. AM 은 N-bit 입력에 대해 n^2 의 FA 가 추가되기 때문에 입력 비트가 커질 수록 면적이 증가한다.

표 1. MAC 구조 별 면적과 전력, 소요시간 비교

Design	4-bit			8-bit		
	EGC	Power [mW]	Delay [ns]	EGC	Power [mW]	Delay [ns]
AM+RCA+Register	320.09	0.167	6.05	1543.66	1.137	6.41
MBM+RCA+Register	504.80	0.106	5.98	2682.91	0.594	6.41
AM+CLA+Register	347.77	0.183	5.95	1665.71	1.201	6.40
MBM+CLA+Register	528.14	0.119	5.82	2840.65	0.609	6.41

2.5 Modified Booth's Multiplier (MBM)

MBM 은 그림 5 와 같이 Booth's Encoder, 승수와 피승수로 구성된다[3]. 부분합은 Booth's Encoding 방식에 의해 생성되며 최종단에서 모든 부분합을 더하여 최종 곱셈 결과가 출력된다. MBM 을 사용하면, 부분합의 수가 줄어들어 전체적인 지연이 줄어들지만, Booth's Encoder 의 복잡성으로 면적이 증가한다.

III. 실험결과 및 결론

MAC unit 구현 결과를 ASIC 으로 확인하기 위해 140MHz 의 동작 주파수로 합성하였다. Synopsis 사의 Design Compiler 를 사용하여 CMOS 180nm 공정으로 합성하였다. 표 1 의 Equivalent Gate Count (EGC) 와 Power 의 결과값을 통해 4 종류의 MAC unit 을 면적과 전력 측면에서 비교하였다. 4-bit 와 8-bit 입력 모두 면적 측면에서는 AM+RCA+Register 의 MAC unit 이 가장 작은 면적을 차지하여 저면적에 적합하고, 전력 측면에서는 MBM+RCA+Register 의 MAC unit 이 가장 적은 전력 소모를 보여 저전력에 적합하다.

참고문헌

- [1] M. Sai Kumar, D. A. Kumar and P. Samundiswary, "Design and performance analysis of Multiply-Accumulate (MAC) unit," *2014 International Conference on Circuits, Power and Computing Technologies*, 2014, pp. 1084-1089
- [2] J. Castro-Godínez, D. Hernández-Araya, M. Shafique and J. Henkel, "Approximate Acceleration for CNN-based Applications on IoT Edge Devices," *2020 IEEE 11th Latin American Symposium on Circuits & Systems*, pp. 1-4
- [3] S. -R. Kuang, J. -P. Wang and C. -Y. Guo, "Modified Booth Multipliers with a Regular Partial Product Array," *IEEE Transactions on Circuits and Systems*, pp. 404-408