Approximate Signed Carry Disregard Multiplier

MRS.VADLAMUDI SREE LAKSHMI 1, KOMATIGUNTA GOURI LAKSHMI 2, JONNAKUTI SHARATH CHANDRA3, KOLLIBOINA PAVAN GANESH 4,MADDALA PRAVEEN BABU 5

[1]Assistant Professor in Department of Electronics and Communication Engineering, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru

[2,3,4,5] UG Students with Electronics and Communication Engineering in Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru.

Abstarct_ This paper describes the design and development of the Signed Carry Disregard Multiplier (SCDM), a family of signed approximate multipliers designed for use with Convolutional Neural Networks (CNNs). Our study studied the resiliency of CNN models to approximate multipliers, and found that the present system had the best resiliency with an average difference in accuracy, whereas CNN had the lowest resiliency. These findings help to select energy-efficient approximate multipliers for CNN-based systems, as well as contribute to the creation of energy-efficient deep learning systems by providing an effective approximation technique for CNN multipliers. The proposed SCDM family of approximate multipliers opens up new possibilities for efficient deep learning applications, allowing for large energy savings while maintaining near-perfect accuracy.

Index Terms: Energy efficiency, convolutional neural network, approximate multiplier, image classification.

1.INTRODUCTION

Multiplication is the most common arithmetic operation. Optimising the multiplication units should enhance system performance significantly. Using approximation computation in multiplier unit design and applications can reduce critical path time, power consumption, and area. Designers employ approximation in three computing phases for multipliers. The operands truncation method uses the reality that not all operand bits are equally important.A smaller core multiplier results from picking only a segment of operand bits. Another way is lowering PPs to approximate. PP reduction in approximate Wallace and Dadda multipliers is done via the approximation compressor. CNN uses pruning, quantisation, and weight sharing to balance power efficiency and accuracy. Deep neural network approximation multiplication effects. They found important factors in CNN layers that permitted accurate predictions despite approximate multiplication errors. CNN-optimized energy-efficient approximate multipliers using Mitchell's log multiplication. Various design methods and operand truncation reduced energy consumption. CNN-optimized energy-efficient approximate multipliers using Mitchell's log multiplication.

Various design methods and operand truncation reduced energy consumption.

2.LITERATURE SURVEY

2.1 "Development of Convolutional Neural Network and Its Application in Image Classification: A Survey" by W. Wang, Y. Yang, X. Wang, W. Wang, and J. Li:

In recent years, convolutional neural networks (CNNs) have been widely used in various computer visual recognition tasks and have achieved good results compared with traditional methods. Image classification is one of the basic and important tasks of visual recognition, and the CNN architecture applied to other visual recognition tasks (such as object detection, object localization, and semantic segmentation) is generally derived from the network architecture in image classification. We first summarize the development history of CNNs and then analyze the architecture of various deep CNNs in image classification.

2.2 "Convolutional Neural Network: A Review of Models, Methodologies and Applications to Object Detection" by A. Dhillon and G. K. Verma:

This paper presents a comprehensive review of convolutional neural networks (CNNs), focusing on their models, methodologies, and applications in object detection. The authors discuss various CNN architectures, training techniques, and the evolution of object detection frameworks, highlighting the advancements and challenges in the field. "On Semantics and Deep Learning for Event Detection in Crisis Situations" by G. Burel, H. Saif, M. Fernandez, and H. Alani: In this paper, we introduce Dual-CNN, a semantically-enhanced deep learning model to target the problem of event detection in crisis situations from social media data. A layer of semantics is added to a traditional Convolutional Neural Network (CNN) model to capture the contextual information that is generally scarce in short, ill-formed social media messages.

2.3 "Convolutional Neural Network: A Review of Models, Methodologies and Applications to Object Detection" by A. Dhillon and G. K. Verma:

This paper presents a comprehensive review of convolutional neural networks (CNNs), focusing on their models, methodologies, and applications in object detection. The authors discuss various CNN architectures, training techniques, and the evolution of object detection frameworks, highlighting the advancements and challenges in the field.

3.PROPOSED SYSTEM

proposed approximate 8 × 4 units for Group A and Group B, which are based on disregarding the carries.

In each of these units, the hexadecimal parameter x determines that the carries are disregarded from the first column to the x-th column.

Digital signal processing (DSP) systems include discrete cosine transform (DCT) [1] and rapid Fourier transform [2] as well as finite impulse response filters [3] make extensive use of multipliers. For many years, people have debated the need of a basic but accurate fixed width multiplier for usage in DSP systems. Two of the most often used fixed-width

multipliers are the Booth multiplier [7]-[16] and the Baugh-Wooley (BW) array multiplier [4]-[6]. The Booth encoder lowers the number of truncated partial products, so Booth multipliers have more accuracy than BW multipliers [12]. By means of a simulation and statistical probability of a low-error fixed-width Booth multiplier (FWBM), the compensating bias of truncated partial products is approximated. Although its circuit design is time-consuming, simulation based error compensation is an accurate technique for application in FWBMs. From a simulation, Jou et al. [7] obtained the statistical characteristics and applied linear regression to attain constant compensation in FWBMs. Song et al. [8] achieved adaptive compensation and lowered the truncation error by means of a curve-fitting method. Similarly, an exhaustive simulation was used in the circuit used in [9] to establish a comparison algorithm; Wang et al. [10] used additional product information to improve the accuracy; although exhaustive simulation achieves accurate compensation, it is time-consuming especially for long-width multiplication.

Presenting probability techniques for shortening the simulation duration, these techniques determine the likelihood of an element belonging to the truncated partial products of multipliers [11]-[16]. From theoretical computation in [11], a constant compensation value produced by a circuit was obtained known as probabilistic estimation bias (PEB). Using information from more columns (w) in the truncation section, the generalised PEB (GPEB) technique was applied in [12] to generate a more accurate estimate of the area penalty; the GPEB circuit offers a suitable tradeoff between accuracy and area/power. Presented in [13] for use in compensated circuit design, an adaptive conditional-probability estimator (ACPE) was proven to increase multiplier accuracy. In [14] the method of compensation—called probability and computer simulation (PACS)—combined statistical methods with a tiny area/ Delay penalty to achieve great accuracy with a limited area. Still, a lot of time is devoted configuring some of the hybrid circuits. Adopted in [15] the multilayer conditional-probability (MLCP) approach, in which conditional probability is more complicated. MLCP helps one to reach a great accuracy, but the area cost rises. On the basis of the column data, this approach allows one to acquire a signal-to---noise ratio (SNR) comparable to the ideal SNR value of a post-truncated (P-T) FWBM.

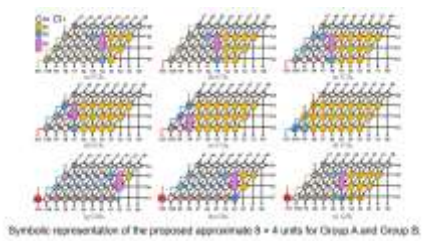

Fig 1:Proposed System

3.1 Exact Signed Multipliers

Among the most often used and basic computer processes in many different fields is signed multiplication. Negative numbers can be shown in several ways, the most often occurring being two's complement. Multiplying two's complement integers can be difficult, though, since the output is not just the sum of PPs unlike in unsigned multiplication. a Baugh-Wooley optimised method for computing the complement of signed two's numbers. Still, by considering the two's complement of the negative numbers, subtraction can be turned into addition.

Thus, just two kinds of partial product units are needed to build an 8-bit signed multiplier. shows a precise 8-bit signed array multiplier including many π 0s, each comprising one AND Gate (^) and an FA utilised for both single-bit multiplication and sum of PPs. Apart from π 0, this multiplier has other partial product units that vary solely in the usage of a NAND Gate rather than one ^ (marked by "N" in Figure 3.1). Apart from these partial product units, the first row of the multiplier utilised just for producing PPs has several ^s and one NAND Gate. The HA of the last row decides the most important portion of the output. With a long critical path shown in red in 1, this accurate signed multiplier The delay of the 8-bit multiplier can be lowered, though, by cutting the multiplier horizontally in the centre and separating it into two smaller 8×4 groups.

4.RESULTS AND DISCUSSION

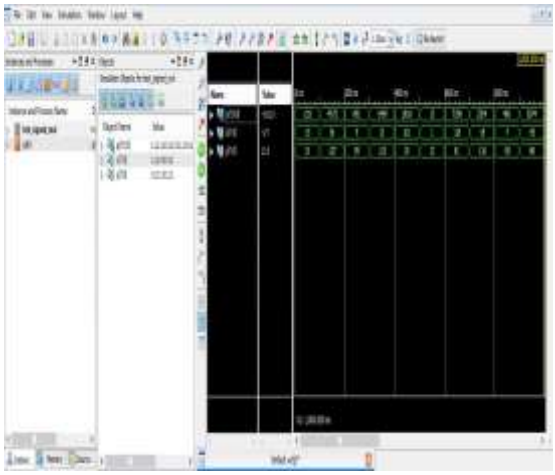

Figure 2: Simulation result of the 8 bit signed multiplier

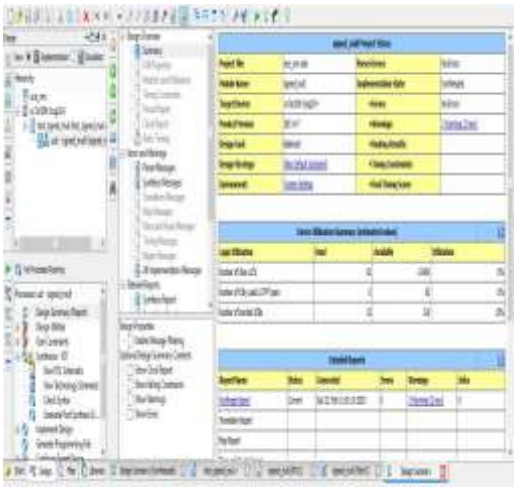

Figure 3: Summary report of the 8 bit signed multiplier

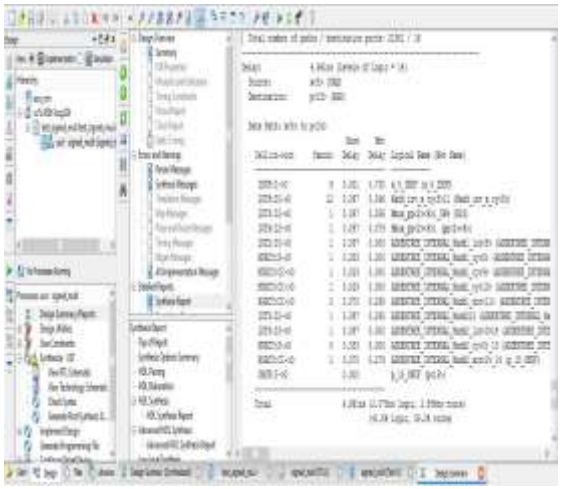

Figure 4: Delay report of the 8 bit signed multiplier

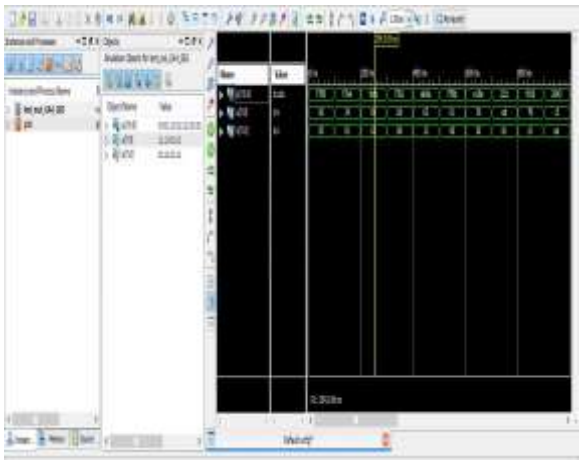

Figure 5: Simulation result of the 8 bit proposed signed multiplier

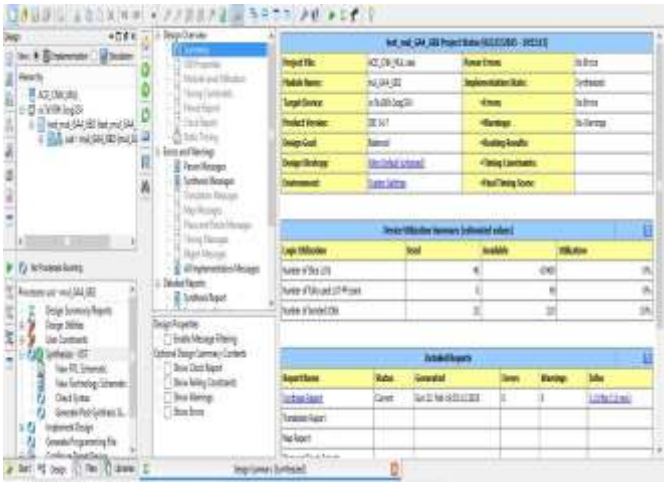

Figure 6: Summary report of the 8 bit proposed signed multiplier

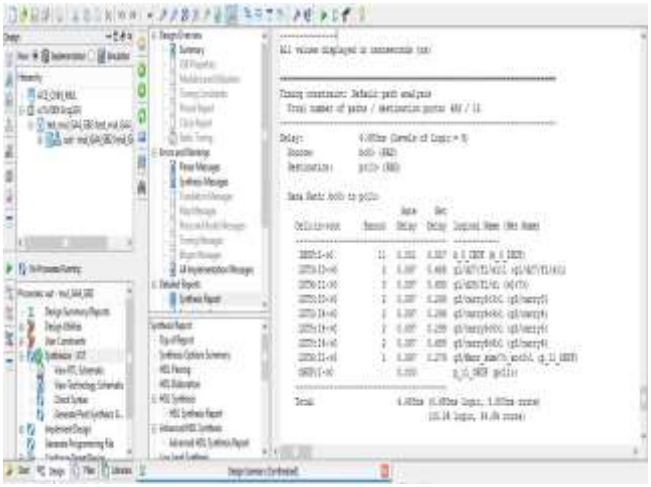

Figure 7: Delay report of the 8 bit proposed signed multiplier

5.CONCLUSION

In this study, we introduced the Signed Carry Disregard Multiplier (SCDM), a novel family of signed approximate multipliers designed to enhance the energy efficiency of CNN-based systems. Our findings indicate that CNN models exhibit varying levels of resiliency to approximation, with the proposed system demonstrating the highest tolerance, maintaining a minimal accuracy difference. This highlights the potential of SCDM in selecting energy-efficient approximate multipliers while preserving classification performance.

The proposed SCDM multipliers pave the way for more power-efficient deep learning applications by significantly reducing computational energy consumption without compromising accuracy. These insights contribute to the broader goal of developing energy-aware CNN accelerators, making deep learning more feasible for resource-constrained environments. Future work will focus on further optimizing SCDM designs and evaluating their

impact on diverse deep learning architectures and real-world applications.

REFERENCES

[1] W. Wang, Y. Yang, X. Wang, W. Wang, and J. Li, "Development of convolutional neural network and its application in image classification: A survey," Opt. Eng., vol. 58, no. 4, 2019, Art. no. 040901. [2] A. Dhillon and G. K. Verma, "Convolutional neural network: A review of models, methodologies and applications to object detection," Prog. Artif. Intell., vol. 9, no. 2, pp. 85–112, Jun. 2020.

[3] G. Burel, H. Saif, M. Fernandez, and H. Alani, "On semantics and deep learning for event detection in crisis situations," in Proc. Work shop Semantic Deep Learn. (SemDeep). Portoroz, Slovenia: ESWC, May 2017.

[4] T.-J. Yang, Y.-H. Chen, and V. Sze, "Designing energy-efficient con volutional neural networks using energy-aware pruning," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jul. 2017, pp. 5687–5695.

[5] A. Bouguettaya, A. Kechida, and A. M. Taberkit, "A survey on lightweight CNN-based object detection algorithms for platforms with limited computational resources," Int. J. Inform. Appl. Math., vol. 2, no. 2, pp. 28–44, 2019.

[6] N. TaheriNejad and S. Shakibhamedan, "Energy-aware adaptive approx imate computing for deep learning applications," in Proc. IEEE Comput. Soc. Annu. Symp. VLSI (ISVLSI), Jul. 2022, p. 328.

[7] Md. M. H. Shuvo, S. K. Islam, J. Cheng, and B. I. Morshed, "Efficient acceleration of deep learning inference on resource-constrained edge devices: A review," Proc. IEEE, vol. 111, no. 1, pp. 42–91, Jan. 2023.

[8] R. Tang, W. Wang, Z. Tu, and J. Lin, "An experimental analysis of the power consumption of convolutional neural networks for keyword spotting," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Apr. 2018, pp. 5479–5483.

[9] V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," Proc. IEEE, vol. 105, no. 12, pp. 2295–2329, Dec. 2017.

[10] H. Jiang, F. J. H. Santiago, H. Mo, L. Liu, and J. Han, "Approximate arithmetic circuits: A survey, characterization, and recent applications," Proc. IEEE, vol. 108, no. 12, pp. 2108–2135, Dec. 2020.

Author's Profiles

Mrs.V. Sree Lakshmi is presently working as an Assistant Professor in the Department of Electronics and Communication Engineering at Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, specializing in Electronics and Communication Engineering. With over 20 years of teaching experience, she has been instrumental in shaping the careers of numerous students in ECE.

She holds a strong academic background and has contributed significantly to the field through her expertise and mentorship. Her dedication to research and teaching in the field of electronics has made a lasting impacting academia, inspiring students to excel in the ever-evolving world of VLSI.

KOMATIGUNTA GOURI LAKSHMI is a B.Tech student specializing in Electronics and Communication Engineering at Seshadri Rao Gudlavalleru Engineering College. She is passionate about exploring emerging technologies and continuously enhancing her skills through hands-on challenges. She has completed internships in VLSI and Artificial Intelligence and Data Visualization under AIMER Society, gaining practical experience. Additionally, she has earned NPTEL certification in Introduction to Internet of Things, Digital VLSI Testing and Sensors and Actuators. She is skilled in PCB(Printed Circuit Board) Making.

JONNAKUTI SHARATH CHANDRA is a B.Tech Student specialising in Electronics and Communication Engineering at Seshadri Rao Gudlavalleru Engineering College. He has completed internships in embedded systems, IoT, and VLSI, gaining expertise in both hardware and software. Certified by NPTEL in IoT, he is skilled in Power BI, Verilog HDL, and Python, with a strong foundation in digital design and functional verification. He has experience in PCB design, object detection using Roboflow, and visual question-answering with Hugging Face. His expertise includes sensors, actuators, and embedded system interfacing.

KOLLIBOINA PAVAN GANESH is a B. Tech student specializing in Electronics and Communication Engineering at Seshadri Rao Gudlavalleru Engineering College. He is passionate about exploring emerging technologies and continuously enhancing his skills through hands-on challenges. He has completed internships in Artificial Intelligence and in Code tantra. Additionally, he has earned NPTEL certification in Introduction to Internet of Things 2.0 . He is skilled in PCB (Printed Circuit Board) Making.

MADDALA PRAVEEN BABU is a B.Tech student specializing in Electronics and Communication Engineering at Seshadri Rao Gudlavalleru Engineering College. He is passionate about emerging technologies and gains hands-on experience through internships in embedded systems, IoT, and data science. He is NPTEL-certified in Cloud Computing and skilled in data visualization using Power BI. His projects include PCB design, object detection with Roboflow, and a visual question-answering model using Hugging Face. He also has expertise in VLSI, sensors, and actuators, enhancing his technical foundation.