

Atmospheric neutron single event effects for multiple convolutional neural networks based on 28-nm and 16-nm SoC

Xu Zhao(赵旭), Xuecheng Du(杜雪成), Chao Ma(马超), Zhiliang Hu(胡志良), Weitao Yang(杨卫涛), and Bo Zheng(郑波)

Citation: Chin. Phys. B, 2025, 34 (**1**): 018501. DOI: 10.1088/1674-1056/ad8b38

What follows is a list of articles you may be interested in

Single event effects evaluation on convolution neural network in Xilinx 28 nm system on chip

Xu Zhao(赵旭), Xuecheng Du(杜雪成), Xu Xiong(熊旭), Chao Ma(马超), Weitao Yang(杨卫涛), Bo Zheng(郑波), and Chao Zhou(周超)

Chin. Phys. B, 2024, 33 (**7**): 078501. DOI: 10.1088/1674-1056/ad3b82

Experiment and simulation on degradation and burnout mechanisms of SiC MOSFET under heavy ion irradiation

Hong Zhang(张鸿), Hongxia Guo(郭红霞), Zhifeng Lei(雷志锋), Chao Peng(彭超), Zhangang Zhang(张战刚), Ziwen Chen(陈资文), Changhao Sun(孙常皓), Yujuan He(何玉娟), Fengqi Zhang(张凤祁), Xiaoyu Pan(潘霄宇), Xiangli Zhong(钟向丽), and Xiaoping Ouyang(欧阳晓平)

Chin. Phys. B, 2023, 32 (**2**): 028504. DOI: 10.1088/1674-1056/ac8cda

Investigation of single event effect in 28-nm system-on-chip with multi patterns

Wei-Tao Yang(杨卫涛), Yong-Hong Li(李永宏)†, Ya-Xin Guo(郭亚鑫), Hao-Yu Zhao(赵浩昱), Yang Li(李洋), Pei Li(李培), Chao-Hui He(贺朝会), Gang Guo(郭刚), Jie Liu(刘杰), Sheng-Sheng Yang(杨生胜), and Heng An(安恒)

Chin. Phys. B, 2020, 29 (**10**): 108504. DOI: 10.1088/1674-1056/ab99b8

Synergistic effect of total ionizing dose on single event effect induced by pulsed laser microbeam on SiGe heterojunction bipolar transistor

Jin-Xin Zhang(张晋新), Hong-Xia Guo(郭红霞), Xiao-Yu Pan(潘霄宇), Qi Guo(郭旗), Feng-Qi Zhang(张凤祁), Juan Feng(冯娟), Xin Wang(王信), Yin Wei(魏莹), Xian-Xiang Wu(吴宪祥)

Chin. Phys. B, 2018, 27 (**10**): 108501. DOI: 10.1088/1674-1056/27/10/108501

Three-dimensional simulation of fabrication process-dependent effects on single event effects of SiGe heterojunction bipolar transistor

Jin-Xin Zhang(张晋新), Chao-Hui He(贺朝会), Hong-Xia Guo(郭红霞), Pei Li(李培), Bao-Long Guo(郭宝龙), Xian-Xiang Wu(吴宪祥)

Chin. Phys. B, 2017, 26 (**8**): 088502. DOI: 10.1088/1674-1056/26/8/088502

Atmospheric neutron single event effects for multiple convolutional neural networks based on 28-nm and 16-nm SoC

Xu Zhao(赵旭)[1], Xuecheng Du(杜雪成)[1,†], Chao Ma(马超)[1], Zhiliang Hu(胡志良)[2,3], Weitao Yang(杨卫涛)[4], and Bo Zheng(郑波)[1]

[1] *School of Nuclear Science and Technology, University of South China, Hengyang* 421001, *China*
[2] *Spallation Neutron Source Science Center, Dongguan* 523000, *China*
[3] *Institute of High Energy Physics, Chinese Academy of Sciences (CAS), Beijing* 100049, *China*
[4] *School of Microelectronics, Xidian University, Xi'an* 710071, *China*

(Received 22 September 2024; revised manuscript received 21 October 2024; accepted manuscript online 25 October 2024)

The single event effects (SEEs) evaluations caused by atmospheric neutrons were conducted on three different convolutional neural network (CNN) models (Yolov3, MNIST, and ResNet50) in the atmospheric neutron irradiation spectrometer (ANIS) at the China Spallation Neutron Source (CSNS). The Yolov3 and MNIST models were implemented on the XILINX 28-nm system-on-chip (SoC). Meanwhile, the Yolov3 and ResNet50 models were deployed on the XILINX 16-nm FinFET UltraScale+MPSoC. The atmospheric neutron SEEs on the tested CNN systems were comprehensively evaluated from six aspects, including chip type, network architecture, deployment methods, inference time, datasets, and the position of the anchor boxes. The various types of SEE soft errors, SEE cross-sections, and their distribution were analyzed to explore the radiation sensitivities and rules of 28-nm and 16-nm SoC. The current research can provide the technology support of radiation-resistant design of CNN system for developing and applying high-reliability, long-lifespan domestic artificial intelligence chips.

Keywords: single event effects, atmospheric neutron, system on chip, convolutional neural network

PACS: 85.30.De, 61.82.Fk, 85.35.–p, 61.80.Hg

DOI: 10.1088/1674-1056/ad8b38 **CSTR:** 32038.14.CPB.ad8b38

1. Introduction

CNNs have demonstrated outstanding performance across various fields such as image recognition, and object detection.[1,2] Meanwhile, they show utter promising applications in satellite remote sensing, deep space exploration, and so on.[3–5] The CNN system based on SoC benefits from these advantages, which are supporting large-scale parallelism, high performance-to-power ratio, and variable precision. Moreover, it can greatly improve the reliability and performance, reduce the weight and power consumption of space electronic system. Thereby, the CNN system based on SoC has the tremendous application potential for spacecraft edge computing and terminal.[6] However, the high-energy particles in the space radiation environment, such as protons and heavy ions, pose one significant threat: SEEs.[7–9] The atmospheric neutrons interacting with atomic nuclei produce high-energy secondary particles, capable of causing SEEs in various aerospace electronic systems around terrestrial.[10,11] And such atmospheric radiation environments exist both on the earth's surface and in near-earth space. The resulting SEEs can lead to some issues in numerous safety and reliability artificial intelligence applications in terrestrial environments.[12,13]

The recent research has only focused on the impact of SEEs on simple CNN models and specific single SoC,[14–18] especially a little research on radiation experiment and theoretical analysis. Given the massive computational requirements of CNNs and the complexity of hardware, the impact of SEEs on CNNs is exceptionally intricate. Therefore, the various types of CNN models with artificial intelligence application scenarios on different SoCs and implementations are necessary to be studied. The atmospheric neutrons irradiation accelerated environment can be produced by ANIS at CSNS, making it feasible to assess the SEEs of CNN induced by atmospheric neutron.

In this paper, three different CNN models (Yolov3, MNIST CNN, and ResNet50) have been implemented on the 28-nm SoC and 16-nm SoC using Vitis-AI and Xilinx Vitis high level synthesis (HLS) tools. The SEE susceptibilities of the involved CNNs have been evaluated by the ANIS at CSNS. The complexity and practical applications of CNN models have preliminarily revealed the phenomenon and patterns of atmospheric neutron SEEs. The structure of the paper is as follows. Section 2 gives description of the tested SoCs and CNN network models. Section 3 introduces the CNN system design, atmospheric neutrons experiment platform and process. Section 4 analyzes the experiment results. Finally, Section 5 is a summary of this article.

†Corresponding author. E-mail: duxuecheng@usc.edu.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

2. Tested devices and networks

2.1. Device under test

The two devices under test are shown in Table 1, one is the 28-nm SoC and the other is 16-nm SoC.[19,20] They are composed of a processing system (PS) coupled with programmable logic (PL) components. For the 28-nm SoC, the PS part consists of dual Cortex-A9 Cores. When it comes to the 16-nm SoC, the PS section contains dual Cortex-A53 and Cortex-R5 cores. On the other hand, their PL side is implemented in the form of an FPGA, having flexible and configurable. And it is noteworthy that the 16-nm SoC has more abundant resources and powerful computational capabilities for executing CNNs.

Table 1. Comparison between 28-nm SoC and 16-nm MPSoC specifications.

	28-nm SoC	16-nm MPSoC
Processor core	Dual Cortex-A9	Dual Cortex-A53+Cortex-R5
System logic cells (K)	85	103
DSP slices	220	240
On-chip memory	DDR3	DDR4
Block RAMs (Mb)	4.9	5.3

2.2. Convolution neural network models and dataset

Three CNN models were selected for their advantages such as real-time image recognition and object detection applications. The architecture of these CNN models comprises several major components: including convolution layer (Conv), maximum pooling player (Max-pooling), rectified linear unit (ReLU), fully connected layer (FC), residual blocks (RB) + batch normalization (BN), global average pooling (GAP), and feature pyramid networks (FPN), as summarized in Table 2.

For the 28-nm SoC, the Yolov3 model is trained on the Common Objects in Context (COCO) 2017 dataset, which contains a large-scale object detection images categorized into 80 classes.[21] The second model is based on the MNIST CNN model, trained on the MNIST dataset consisting of 28×28 pixels images divided into 10 different categories.[22] The third model keeps the same architecture as the MNIST CNN, while it is implemented using Vitis HLS.[23]

For the 16-nm SoC, the Yolov3 model is trained on the PASCAL Visual Object Classes (PASCAL VOC) 2012 dataset, which consists of object detection images divided into 20 categories.[24] The second model is the ResNet-50, trained on MNIST, a dataset consisting of 32×32 pixels images divided into 10 categories.[25]

It is noteworthy that these CNNs based on 28-nm SoC and 16-nm SoC are all implemented by Xilinx Vitis-AI tools, except for the third one in the 28-nm SoC.

Table 2. The convolution neural network architectures.

Network	Dataset	Output types	Conv	Max-pooling	ReLU	FC	RD+BN	GAP	FPN
Yolov3@28-nm SoC	COCO2017	80	√		√(Leaky)		√		√
MNIST CNN@28-nm SoC Vitis-AI	MNIST	10	√	√	√	√			
MNIST CNN@28-nm SoC Vitis HLS	MNIST	10	√	√	√	√			
ResNet50@16-nm SoC	MNIST	10	√	√	√	√	√	√	
Yolov3@16-nm SoC	VOC2012	20	√		√(Leaky)		√		√

3. Convolution neural networks system design and atmospheric neutrons experiment

3.1. Vitis-AI and Vitis HLS tools

In this paper, several CNN models have been developed based on the 28-nm SoC and 16-nm SoC, leveraging the Vitis-AI tools and Vits HLS tools.[26,27] In additions, the model's inference time, accuracy, the position of the anchor boxes (height/width), and classification result, are individually evaluated, as shown in Table 3.

(i) Vitis-AI tools

This framework includes quantization, optimization, compilation, as well as libraries to run inference and the acceleration runtime (XRT) for deploying quantized neural networks on SoCs, as depicted in Fig. 1. It provides minimal effort for migrating an application from one edge platform to another. And the N2Cube, based on the XRT, provides the application programming interfaces to implement CNN application in Python or other language. In additions, some DPU un-supported operations like softmax and pooling are implemented on the CPU side.

Table 3. The CNN systems output detection results.

Experiment SoC	Inference support	Network	Inference time	Accuracy	Classification result	Height/width
28-nm SoC	Vitis-AI tools	Yolov3	√	√	√	
	Vitis-AI tools	MNIST CNN			√	
	Vitis HLS tools	MNIST CNN			√	
16-nm SoC	Vitis-AI tools	ResNet50	√		√	
	Vitis-AI tools	Yolov3	√	√	√	√

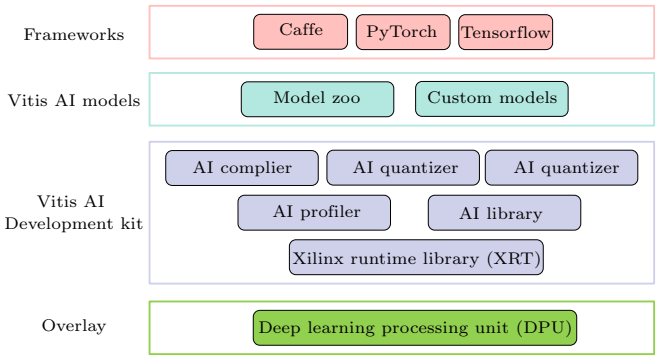


Fig. 1. The Xilinx Vitis-AI framework.

(ii) Vitis HLS tools

The Xilinx HLS tool is applied to simplify the CNN's complex design flows. It converts the high-level specification languages to low-level hardware languages, and automates the compilation of the high-level specification functional language into a register-transfer level (RTL) acceleration IP core. Consequently, then the acceleration IP core can be used by Vivado software suite. The CNN models can be rapidly implemented by this approach.

3.2. CNNs on devices

3.2.1. Design process of CNN systems-based Vitis-AI

The Vitis-AI utilizes the INT8 data type for quantizing the CNN models, facilitating the model compression to reduce area, memory, and other resources. The implementation flow of Vitis-AI is illustrated taking an example of the 16-nm SoC CNN system. This process includes training CNN models on personal computer (PC), quantizing, optimizing, and compiling models on Vitis-AI tools, and subsequent hardware development, as drawn in Fig. 2.

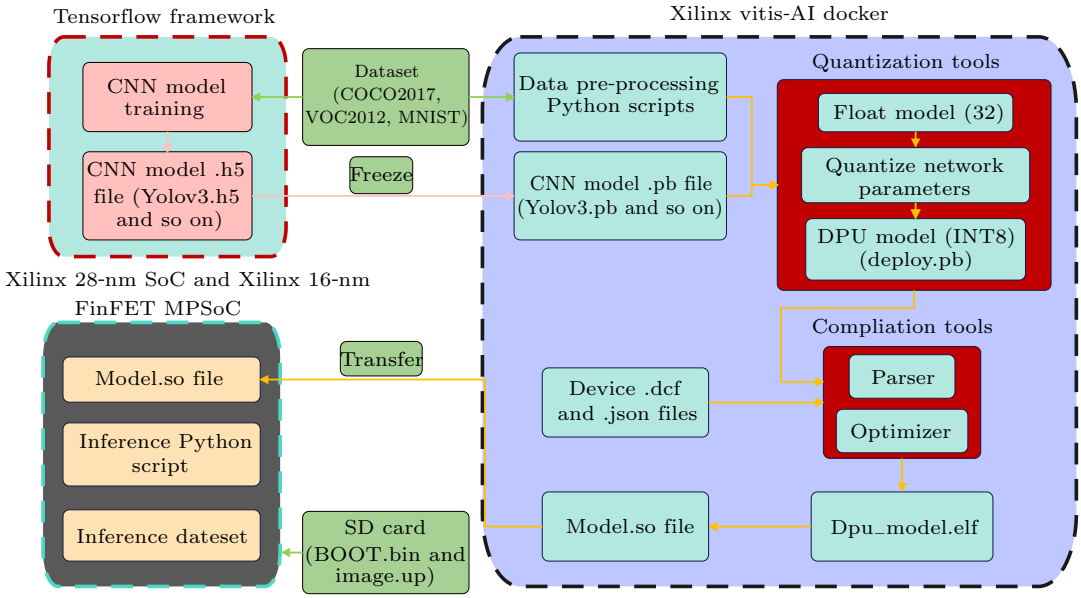


Fig. 2. The Xilinx Vitis-AI design flow.

The Yolov3 network is trained on the VOC2012 dataset. The model is trained using the standard 9 anchor boxes. Input images size is resized to 416×416 by Yolov3, after which objects are detected. The bounding box coordinates (height/width) are then output with relative to the original dataset size.

Firstly, the Yolov3 is trained by TensorFlow framework on the PC. The training consists of two stages. During the first stage, most layers are frozen and only the last few output layers are trained; in the second stage, all layers are unfrozen and the model is fine-tuned. The pretraining process includes 50 epochs in each stage. Then the light-weight model graph Yolov3.pb will be generated.

Secondly, the deploy model (deploy.pb) is generated according to the frozen network graph Yolov3.pb serves as the input, with the input and output nodes, input shape and dataset preprocessing script.

Thirdly, according to device's data processing unit (DPU) types, hardware architecture description files (.json and .dcf), the quantized network is compiled by Vitis AI deep neural network development kit (DNNDK) tool. This step compile layer information for DPU, summarizing operations required for both DPU and CPU sides. Then, the dynamic link library file (.so) is generated using Xilinx runtime docker.

Finally, the generated .so file, dataset and inference Python script are transferred into the SD card for deployment. This implements the Yolov3 model on the MPSoC system.

The ResNet50 network is trained on the MNIST dataset. Input images are converted into 32×32 pixels by ResNet50 for number classification. The ResNet50 network follows the same Vitis-AI flow as Yolov3. Additionally, the Yolov3 network for the 28-nm SoC is trained on the COCO2017 dataset.

This network design flow mirrors the Vitis-AI flow as Yolov3 in MPSoC. Regarding another two MNIST dataset CNN networks, one follows the Vitis-AI flow, and the other follows the Vits HLS flow. The input images size is 28×28 pixels, and they classify number. The former adheres to Vitis-AI flow, similar to Yolov3 and ResNet50.

3.2.2. The design process of CNN system-based Vitis HLS

SoC-based inference accelerators can be implemented using Vitis HLS. The design flow is shown in Fig. 3. For the MNIST CNN, the network accelerator IP core can be designed by Vitis HLS based on the network framework,

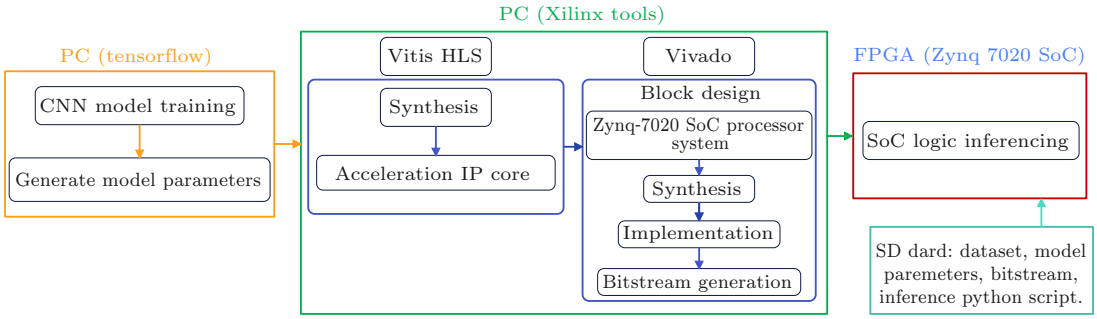
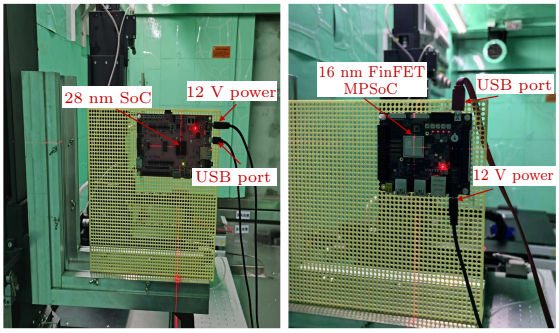


Fig. 3. The design flow of Vitis HLS tools.

weights, and biases. The accelerator IP can then be integrated into the block design. The Vivado software suite is employed to connect the accelerator IP core with Xilinx 28-nm Zynq-7020 SoC in the block design. In this manner, the same architecture MNIST CNN can be implemented.

3.3. Atmospheric neutrons experiment

In this study, simulated atmospheric neutron irradiation experiments were conducted at the ANIS facility of CSNS. During the irradiation tests, the neutron intensity exceeded 10^6 n/cm^2·s. The atmospheric neutron irradiation experiments fluence was measured by ANIS neutron detector, then the fluence data acquired by CSNS-ANIS host computer software. The JEDEC is the terrestrial neutron spectrum data.[28] The ANIS neutron spectrum is highly consistent with the JEDEC standards. During experiments, the test chip was positioned about 20 m from the target center, utilizing the accelerator beam with the power of 140 kW. A beam spot size of 2 cm×2 cm was employed with the uniformity of ±10%, ensuring coverage of the test chips. Figure 4 shows the SoC and MPSoC CNN systems of neutron irradiation.



(a) 28-nm SoC (b) 16-nm MPSoC

Fig. 4. The 28-nm SoC and 16-nm MPSoC CNN systems of atmospheric neutron irradiation experiment.

During the irradiation test, the Yolov3, ResNet50, and MNIST CNN were executed in real-time, and various metrics including inference time, accuracy, height/width, and classification result were recorded. The test process of atmospheric neutron irradiation experiment is drawn in Fig. 5.

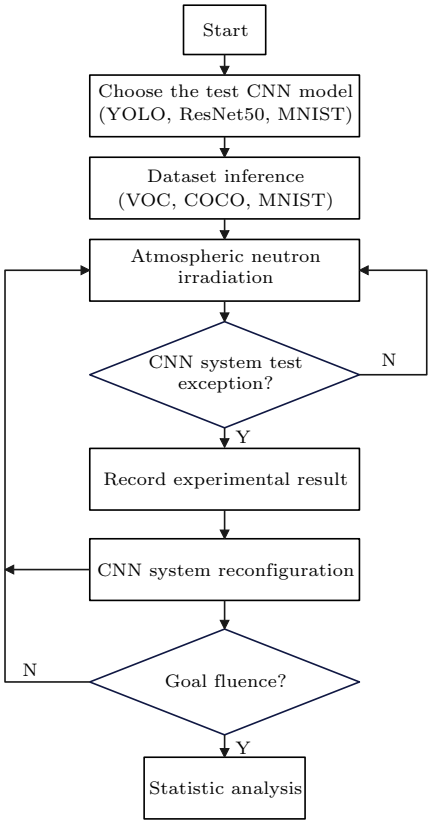


Fig. 5. The test procedure of atmospheric neutron irradiation experiment.

The 28-nm SoC and 16-nm SoC-based CNN systems have been implemented. The process of the irradiation experiment is as follows. Firstly, the SoC system and CNN model

are selected. Secondly, the CNN model performs inference process in real-time. Thirdly, the atmospheric neutron irradiation experiment was launched at the ANIS and the system outputs are monitored in real-time. Finally, the CNN system is reconfigured in case of system exception. Experimental results are recorded, and the irradiation is halted when achieving the goal fluence. The process of neutron irradiation is repeated until all CNN models have been tested.

4. Atmospheric neutrons experiment results and analysis

4.1. Xilinx 28-nm Zynq-7020 SoC experiment results and analysis

The observed SEE soft errors can be categorized into eight main types: false negative (FN), false positive (FP), classification error (CE), accuracy decrease (AD), accuracy increase (AI), system crash (SC), system halt (SH), duplicate recognition (DR).

FN: the number of detected objects is different from the golden prediction and is lower.

FP: the number of detected objects differs from the golden prediction and is higher.

CE: the number of detected objects is same as the golden prediction, but the objects classification results are different.

AD: the object detection accuracy is lower than the golden prediction.

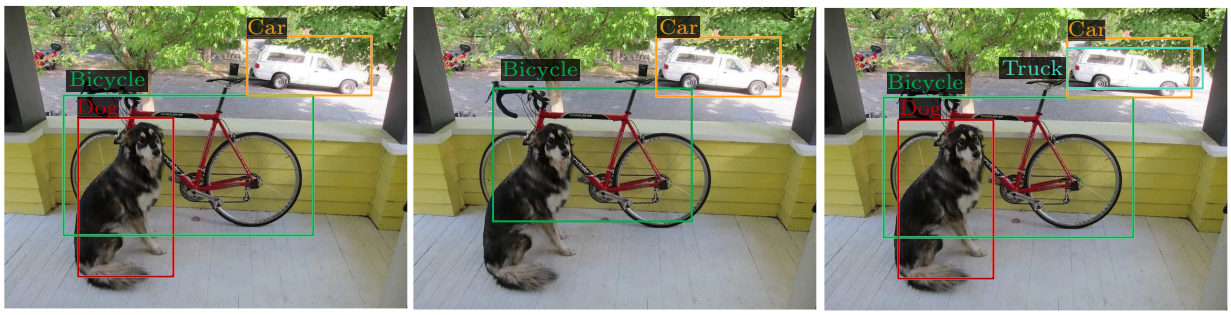
AI: the object detection accuracy is higher than the golden prediction.

SC: the SoC CNN system fails to perform dataset inference.

SH: the SoC CNN system stops outputting messages.

DR: the SoC CNN system repeatedly outputs the same object detection results.

The definition of the FN and FP can better be depicted in Fig. 6.[29] In a typical environment, the Yolov3 CNN system is able to detect the three objects. However, in an irradiation environment, the Yolov3 CNN system may detect either two objects or four objects.



(a) Golden prediction (b) FN: lower recognition (c) FP: higher recognition

Fig. 6. The Yolov3 objects detection example.

Table 4 provides a summary of the SEE soft errors detected during the atmospheric neutron irradiation. The SEEs soft errors are categorized into eight types for Yolov3, while four types for Vitis-AI MNIST CNN and two types for Vitis HLS MNIST CNN. And the SEE cross-section were computed.

The SEE cross-section can be calculated by[30]

$$\sigma = \frac{N}{F}, \tag{1}$$

where σ is the cross section of SEE soft error in unit cm^2; N is the number of SEE soft error events; F is the neutron fluence in particles per cm^2. The atmosphere neutron experiment fluence is 2.358×10^9 n/cm^2, 1.211×10^{10} n/cm^2, 1.342×10^{10} n/cm^2, respectively.

In the atmospheric neutron irradiation environment, the SoC-based Yolov3 CNN system was sensitive to SEE. The FP, FN, CE, AD, and AI severely impacted the inference performance of the Yolov3 CNN system. During irradiation, the Yolov3 CNN system was more prone to missing objects detection, leading to decreased accuracy. In addition, it is worth noting that some object recognition accuracy improved as well. The objects detection associated errors (FP, FN, and CE) accounts for 19.91%, while the recognition accuracy decrease and increase accounted for 75.29% of all errors.

For the Vitis-AI MNIST CNN system, CE was the predominant error type.

In contrast, for the Vitis HLS MNIST CNN system, the SC was the main error type. The CNN system stops inference when the SC error occurs. And it requires to be reconfigured, severely impacting the normal operation of the CNN system. This indicates that the implement CNN systems using Vitis HLS tools are more sensitive in processor system.

It is important to note that regardless of the deployment method, the 28-nm SoC-based CNN system all occurred SC and SH errors. The SC manifested as several types as follows: firstly, the first stage boot loader (FSBL) exception, secondly, the CNN system inference stage outputs system exception codes. For the same architecture MNIST CNN, the Vitis HLS implementation of network is more vulnerable to SEE.

Table 4. The SEE soft errors neutron experiment results of Xilinx 28-nm SoC.

Experiment network	Error types	Error count	Fluence (n·cm^{-2})	SEE cross-section (cm^2)
Yolov3	false negative	93	2.358×10^{9}	3.944×10^{-8}
	false positive	20		8.482×10^{-9}
	classification error	24		1.018×10^{-8}
	accuracy decrease	460		1.951×10^{-7}
	accuracy increase	58		2.460×10^{-8}
	system crash	22		9.330×10^{-9}
	system hang	5		2.120×10^{-9}
	duplicate recognition	6		2.545×10^{-9}
Vitis-AI MNIST CNN	classification error	6455	1.211×10^{10}	5.330×10^{-7}
	system crash	75		6.193×10^{-9}
	system hang	12		9.909×10^{-10}
	duplicate recognition	1		8.258×10^{-11}
Vitis HLS MNIST CNN	system crash	68	1.342×10^{10}	5.067×10^{-9}
	system hang	7		5.216×10^{-10}

4.2. Xilinx 16-nm MPSoC experiment results and analysis

The detected SEE soft errors can be categorized into nine main types: FN, FP, CE, AD, AI, system interrupt (SI), height/width (HW), DR, auto reboot (AR).

SI: the MPSoC CNN system inference interrupt during irradiation.

HW: the bounding box coordinates of objects is different between the golden prediction. This difference includes height, width, or both.

AR: the MPSoC CNN system auto reboots during irradiation.

Table 5 lists the SEE soft errors detected during the atmospheric neutron irradiation. The SEEs soft errors were categorized into seven types for Yolov3, and four types in ResNet50. The SEE cross section can be calculated by Eq. (1). The atmospheric neutron experiment fluence is 1.626×10^{10} n/cm^2 and 1.516×10^{10} n/cm^2, respectively.

Table 5. The SEE soft errors neutron experiment results of Xilinx 16-nm MPSoC.

Experiment network	Error types	Error count	Fluence (n·cm^{-2})	SEE cross-section (cm^2)
ResNet50	classification error	23601	1.626×10^{10}	1.451×10^{-6}
	duplicate recognition	2		1.230×10^{-10}
	interrupt	2		1.230×10^{-10}
	auto reboot	2		1.230×10^{-10}
Yolov3	false negative	323	1.516×10^{10}	2.131×10^{-8}
	false positive	188		1.240×10^{-8}
	classification error	45		2.968×10^{-9}
	accuracy decrease	3579		2.361×10^{-7}
	accuracy increase	752		4.960×10^{-8}
	interrupt	6		3.958×10^{-10}
	width and height	2069		1.365×10^{-7}

In the atmospheric neutron irradiation environment, the MPSoC-based Yolov3 CNN system underwent testing. The FP, FN, CE, AD, and AI significantly impact the reliability of the Yolov3 CNN system. Similar to the 28-nm Yolov3 CNN system, the neutron irradiation results of the MPSoC Yolov3 CNN system indicate that the number and accuracy of object detection are affected, with both increases and decreases observed. The object detection-associated errors accounts for 11.36%, while the recognition accuracy decrease and increase accounts for 88.51% of all error types. In addition, the HW errors were detected during irradiation. These errors resulted in changes in the object detection (numbers and categories of objects, prediction accuracy) and different with the golden data.

For the MPSoC-based ResNet50 CNN system, the CE was also the predominant error type. This is because the CNN system only outputs the single number category.

4.3. Comparing 28-nm SoC with 16-nm SoC SEE results and discussion

According to the irradiation experiment, the SEE soft errors observed as shown in Fig. 7. Based on the distribution results, the accuracy decrease dominates the errors in Yolov3

networks, accounting for more than 50%. And the FP, FN, and CE also severely affected the reliability of Yolov3 CNN system. In addition, the object detection accuracy might also increase, which causes a threat to the reliability of the CNN system. For the ResNet50 and Viti-AI MNIST CNN, CE was the predominant error, accounting for more than 98%. For Viti HLS MNIST CNN, the SC error accounts for more than 90%. The SoC-based CNN system encounters significant SC errors, which should be noticed.

The SEE cross section for the five CNN models are calculated and shown in Table 6. Based on these SEE cross-section data, the 16-nm SoC is much more sensitive than 28-nm SoC.

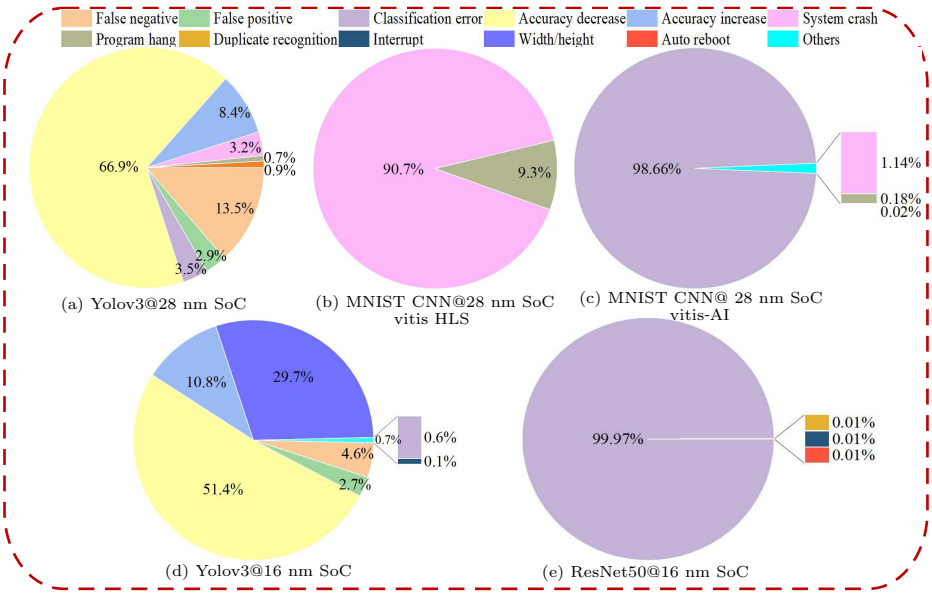


Fig. 7. The five CNN models distribution of SEE soft errors.

Table 6. The SEE cross section of different CNN systems.

Experiment network	SEE cross-section (cm^2)
Yolov3@28-nm SoC	2.833×10^{-8}
MNIST CNN@28-nm SoC Vitis-AI	5.403×10^{-7}
MNIST CNN@28-nm SoC Vitis HLS	5.589×10^{-9}
ResNet50@16-nm SoC	1.452×10^{-6}
Yolov3@16-nm SoC	4.592×10^{-7}

The results of the neutron irradiation experiment for different CNN models are presented in Table 7. Whether based on 28-nm SoC or 16-nm MPSoC, the CE, FP, FN, AD, and AI were identified as the primary same types of errors for Yolov3 network. Additionally, the MPSoC has a higher error count and error rate.

The same MNIST dataset was used for the 28-nm Vitis-AI and 16-nm ResNet50 networks. Since these CNN networks only output the number category, when the SEE occurs, it only indicates in the classification results and system performance. Therefore, the majority of the SEEs result is misclassification, and the 16-nm SoC has higher classification errors. In addition, the 28-nm SoC CNN system suffers from more system exception, which may be related to the packaging of the SoC itself.

Table 7. The atmospheric neutron irradiation experiment results of different CNN systems.

Experiment network	Test objects	SEE error types	CE	FN	FP	AD	AI	Error (CE+FN+FP) rate (%)	Total errors
Yolov3@28-nm SoC	2415	8	24	92	20	460	58	5.67	688
MNIST CNN@28-nm SoC Vitis-AI	122058	4	6455					5.29	6543
ResNet50@16-nm SoC	380716	4	23601					6.20	23607
Yolov3@16-nm SoC	7243	6	45	323	188	3579	752	7.68	5029

In terms of the inference time, which include the convolution time, decode time, prob time, and total time, tests were conducted based on the Yolov3 in both the SoC and MPSoC, as well as the ResNet50 network. For the 28-nm SoC Yolov3 CNN system, the impact of neutron irradiation did not significant impact the inference time. However, for the 16-nm SoC CNN system, neutron irradiation increased the prob time by approximately 100 ms for Yolov3. For example, without irradiation: (prob time: 321.77 ms); whereas with the irradiation, the prob time increased to 419.14 ms. For 16-nm SoC ResNet50 CNN system, the impact of neutron irradiation on the inference time of the 16-nm SoC ResNet50 CNN system

was negligible, with only some slight slowdown observed in object detection inference. As the semiconductor technology nodes scaling from 28 nm to 16 nm, the device dimensions and operating voltage are reduced. Thus, the CNN system gets more sensitive to SEEs. The 16-nm SoC CNN system, with more complex system integration, functional modules, and memory units, is particularly susceptible.

5. Conclusion

The impact of SEEs induced by atmospheric neutron on CNN system based on 28-nm and 16-nm SoCs have been comprehensively evaluated by utilizing ANIS at CSNS. During the atmospheric neutron experiment, eight soft error types were detected in the 28-nm CNN system, nine soft error types were observed in the 16-nm CNN system. The SEE cross section and distribution of these CNN models were calculated, revealing the CNN system based on Xilinx 16-nm MPSoC is more sensitive than the 28-nm SoC. The experiment results illustrate that the SEEs severely impact the operation of CNN system through atmospheric neutron irradiation and analysis of experimental results, a comprehensive evaluation of the impact mechanism of SEE on CNN systems. This research provides critical radiation-resistant design technologies support and theoretical basis to enhance the reliability of domestic artificial intelligence chips, supporting the development and application of high-reliability, long-lifespan domestic artificial intelligence chips.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant No. 12305303), the Natural Science Foundation of Hunan Province of China (Grant Nos. 2023JJ40520, 2024JJ2044, and 2021JJ40444), the Science and Technology Innovation Program of Hunan Province, China (Grant No. 2020RC3054), the Postgraduate Scientific Research Innovation Project of Hunan Province, China (Grant No. CX20240831), the Natural Science Basic Research Plan in the Shaanxi Province of China (Grant No. 2023-JC-QN-0015), and the Doctoral Research Fund of University of South China (Grant No. 200XQD033).

References

[1] Cosmas K and Kenichi A 2020 *Aerospace.* **7** 159
[2] Gao H B, Cheng B, Wang J Q, Li K Q, Zhao J H and Li D Y 2018 *IEEE Trans. Industr. Inform.* **14** 4224
[3] Jin X M, Chen W, Li J L, Qi C, Guo X Q, Li R B and Liu Y 2019 *Chin. Phys. B* **28** 104212
[4] Guo X T, Pimentel A D and Stefanov T 2023 *IEEE Internet Things J.* **10** 5843
[5] Cheng T, Zhao R S, Wang S, Wang R and Ma H Y 2024 *Chin. Phys. B* **33** 040303
[6] Meloni P, Capotondi A, Deriu G, Brian M, Conti F, Rossi D, Raffo L and Benini L 2018 *ACM T Reconfig. Techn.* **11** 1
[7] Yang S H, Zhang Z G, Lei Z F, Huang Y, Xi K, Wang S L, Liang T J, Tong T, Liu X H, Peng C, Wu F G and Li B 2022 *Chin. Phys. B* **31** 126103
[8] Yang W T, Li Y H, Guo Y X, Zhao H Y, Li Y, Li P, He C H, Guo G, Liu J, Yang S S and An H 2020 *Chin. Phys. B* **29** 108504
[9] Wei J N, Li Y, Liao W L, Liu F, Li Y H, Liu J C, He C H and Guo G 2022 *Chin. Phys. B* **31** 086106
[10] Hu Z L, Yang W T, Zhou B, Liu Y N, He C H, Wang S L, Yu Q Z and Liang T J 2022 *J. Nucl. Sci. Technol.* **60** 473
[11] Mo L H, Yu Q Z, Hu Z L, Zhou B, Yi T C, Yuan L B, Shen F and Liang T J 2023 *Microelectron. Reliab.* **146** 114997
[12] Lbrahim Y, Wang H B, Bai M, Liu Z, Wang J N, Yang Z M and Chen Z M 2020 *IEEE Access* **8** 19490
[13] Abich G, Garibotti R, Reis R and Ost L 2022 *IEEE T. Circuits-II* **69** 679
[14] Zhao X, Du X C, Xiong X, Ma C, Yang W T, Zheng B and Zhou C 2024 *Chin. Phys. B* **33** 078501
[15] Wang H B, Wang Y S, Xiao J H, Wang S L and Liang T J 2021 *IEEE Trans. Nucl. Sci.* **68** 394
[16] Agiakatsikas D, Foutris N, Sari A, Vlagkoulis V, Souvatzoglou L, Psarakis M, Ye R Q, Goodacre J, Luján M, Kastriotou M, Cazzaniga C and Frost C 2024 *IEEE Trans. Reliab.* **73** 771
[17] Ruospo A and Sanchez E 2021 *Appl. Sci.* **11** 6455
[18] Rech P 2024 *IEEE Trans. Nucl. Sci.* **71** 377
[19] Xilinx, Inc. 2018 Zynq-7000 SoC Data Sheet: Overview, DS190 (v1.11.1)
[20] Xilinx, Inc. 2024 Zynq UltraScale+ MPSoC Data Sheet: Overview, DS891 (v1.11).
[21] Liang S Y, Wu H, Zhen L, Hua Q Z, Garg S, Kaddoum G, Hassan M M and Yu K P 2022 *IEEE Trans. Intell. Transp. Syst.* **23** 25345
[22] Cheng Q, Huang M Q, Man C H, Shen A, Dai L Y, Yu H and Hashimoto M 2023 *IEEE T. Circuits-I* **70** 3978
[23] Traiola M, dos Santos F F, Rech P, Cazzaniga C, Sentieys O and Kritikakou A 2024 *IEEE Trans. Nucl. Sci.* **71** 845
[24] Liu B K, He F J, Du S Q, Li J C and Liu W J 2023 *J. Intell. Fuzzy Syst.* **45** 5807
[25] Mascarenhas S and Agarwal M 2021 *International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications* (CENTCON), November 19–21, 2021, Bengaluru, India, p. 96
[26] Xilinx, Inc. 2023 Vitis High-Level Synthesis User Guide, UG1399
[27] Xilinx, Inc. 2021 Vitis AI User Guide, UG1414 (v1.2)
[28] Measurement and Reporting of Alpha Particle and Terrestrial Cosmic Ray-induced Soft Errors in Semiconductor Devices
[29] Bosio A, Bernardi P, Ruospo A and Sanchez E 2019 *IEEE Latin American Test Symposium (LATS)*, March 11–13, 2019, Santiago, Chile, p. 1
[30] Howard J W, Carts M A, Stattel R, Rogers C E, Lrwin T L, Dunsmore C, Sciarini J A and LaBel K A 2001 *IEEE Radiation Effects Data Workshop (REDW) and IEEE Nuclear and Space Radiation Effects Conference (NSREC)*, July 16–20, 2001, Vancouver, BC, Canada, p. 38