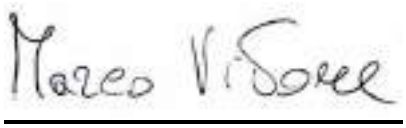
PhD in Information Technology and Electrical Engineering
Università degli Studi di Napoli Federico II

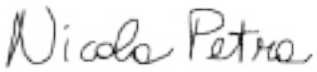
PhD Student: Marco Vitone

Cycle: XXXVI

Training and Research Activities Report

Year: Second



Tutor: Prof. Prof. Nicola Petra 

Co-Tutor: Ing. Claudio Giaccio (Micron Semiconductor Italia)

Date: October 30, 2022

Training and Research Activities Report

PhD in Information Technology and Electrical Engineering

Cycle: XXXVI **Author: Marco Vitone**

1. Information:

- **PhD student: Marco Vitone**
- **DR number: DR995147**
- **Date of birth: 11/10/1994**
- **Master Science degree: Electrical Engineering University: Università degli studi di Napoli Federico II**
- **Doctoral Cycle: XXXVI**
- **Scholarship type: *Funded by Micron Semiconductor Italia S.R.L.***
- **Tutor: Prof. Nicola Petra**
- **Co-tutor: Ing. Claudio Giaccio**

2. Study and training activities:

Activity	Type[1]	Hours	Credits	Dates	Organizer	Certificate[2]
FPGA per l'elaborazione dei segnali	M.Sc. Course	72	9	Semester II	Prof. Nicola Petra	Y
Cambridge English: Preliminary (PET)	Ah hoc Course	40	6	25/03/22 24/06/22	Centro Linguistico di Ateneo (CLA)	Y
Cyber security in Akka Technologies	Seminar	2	0.4	3/11/2021	Prof. D. Cotroneo Prof. S.P. Romano Prof. R. Natella	Y
Vehicular Hacking in Akka Technologies	Seminar	1.5	0.3	3/11/2021	Prof. D. Cotroneo Prof. S.P. Romano Prof. R. Natella	Y
Exploring the early Universe through the cosmic microwave	Seminar	1.5	0.3	4/11/2021	Scuola Superiore Meridionale	Y
Evolution by curvature of networks in the plane	Seminar	1	0.2	11/11/2021	Scuola Superiore Meridionale	Y
Turbolent dynamics in viscous fluids: a complex phenomenon ubiquitous in nature	Seminar	1	0.2	18/11/2021	Scuola Superiore Meridionale	Y
Graphons: A tool for the analysis of systems on large networks	Seminar	1	0.2	25/11/2021	Scuola Superiore Meridionale	Y
GDPR basics for computer scientists	Seminar	2	0.4	14/12/2021	Prof. P. Bonatti	Y
All roads leads to	Seminar	2	0.4	16/12/2021	Prof.	Y

WebRTC					S.P. Romano	
Designing Quantum Algorithms	Seminar	2	0.4	16/12/2021	Prof. A.S. Cacciapuoti	Y
Social network dynamics leading to community formation and residential segregation	Seminar	1.5	0.3	16/12/2021	Scuola Superiore Meridionale	Y
The challenge of gravitational wave detectors of the 3rd generation. Cultural and technological aspects	Seminar	1.5	0.3	20/1/2022	Scuola Superiore Meridionale	Y
Space Weather: Science or Application?	Seminar	1.5	0.3	27/2/2022	Scuola Superiore Meridionale	Y
The needle in the haystack: the search for rare processes and fundamental laws of Nature	Seminar	1.5	0.3	17/2/2022	Scuola Superiore Meridionale	Y
An overview of the transient sky at high-energies	Seminar	1.5	0.3	3/3/2022	Scuola Superiore Meridionale	Y
Global and cluster synchronization in complex networks and beyond	Seminar	1.5	0.3	10/3/2022	Scuola Superiore Meridionale	Y
High energy X-ray Astrophysics from Space: revealing the backbones of the Universe	Seminar	1.5	0.3	17/3/2022	Scuola Superiore Meridionale	Y
From basic principles in spintronics to some recent developments toward spin-orbinics	Seminar	1.5	0.3	31/3/2022	Scuola Superiore Meridionale	Y
Capillary Surfaces and a Model of Nanowire Growth	Seminar	1.5	0.3	7/4/2022	Scuola Superiore Meridionale	Y
Towards a Political Philosophy of AI	Seminar	1	0.2	11/4/2022	Prof. Giuseppe Luongo Prof. Flora Amato	N
Power-Law Gels, Scott Blair and fractional Calculus of Soft Multi-scale Materials	Seminar	1.5	0.3	5/5/2022	Scuola Superiore Meridionale	Y
An informational	Seminar	1.5	0.3	12/5/2022	Scuola	Y

Discussion around stochastic control and free boundary problems					**Superiore Meridionale**	
A day in the life of a Chief Data Officer	**Seminar**	**1**	**0.2**	**9/5/2022**	**Prof. Giuseppe Luongo Prof. Flora Amato**	**N**
Quantum fluids of atoms and of light as analog models of gravity: a fruitful synergy of gravitational physics and quantum optics	**Seminar**	**1.5**	**0.3**	**9/6/2022**	**Scuola Superiore Meridionale**	**Y**
From resilience assessment to design for resilience: what is missing?	**Seminar**	**1.5**	**0.3**	**14/7/2022**	**Scuola Superiore Meridionale**	**Y**

1) Courses, Seminar, Doctoral School, Research, Tutorship
2) Choose: Y or N

2.1. Study and training activities - credits earned

	Courses	Seminars	Research	Tutorship	Total
Bimonth 1	**0**	**3.1**	**4**	**0**	**7.1**
Bimonth 2	**0**	**0.9**	**8**	**0**	**8.9**
Bimonth 3	**0**	**1.7**	**8**	**0**	**9.7**
Bimonth 4	**9**	**1.1**	**2**	**0**	**12.1**
Bimonth 5	**6**	**0.3**	**4**	**0**	**10.3**
Bimonth 6	**0**	**0**	**11,9**	**0**	**11,9**
Total	**15**	**7.1**	**37.9**	**0**	**60**
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	

3. Research activity:

System on Chip is a complex Integrated Circuits that includes several components as CPU, Programmable Logic (FPGA), on-chip memory, storage interfaces and so on.
The design of SoC foresees many aspects to be solved, form the implementation of the hardware accelerator to the validation and debug of the SoC architecture. Increasing the complexity of SoC, the design of the latter becomes challenging.
My research activity focuses on design of System on Chip, dealing with both hardware implementation and validation of SoC.
For the sake of clarity, I divided my research activity report into two sub-sections:

1. Development of hardware accelerator for SoC
2. Development of methodologies for validation and debug of SoC

1- Development of hardware accelerator for SoC

Convolution is the major is a major operation performed in many applications. Mono-dimensional convolution is used in typical digital filter signal processing [1-2]. Mono- and bi-dimensional convolutions are used in Neural Networks for image recognition and detection, speech recognition, text classification, etc. [3-6]. As an example, in image classification a k by k kernel is convoluted with the image. Typical values of k range from 3 to 11. Multidimensional convolution is used in Neural Networks for medical applications [7-8], industrial applications [9], video classification [10-11].

Some applications, such as multimedia [10-11], IoT[12] and industrial [9], require the convolution to be implemented through dedicated hardware accelerators. The multiplication operation is the bottleneck for these kinds of circuits impacting the power dissipation and the speed of the hardware accelerator. For this reason, many results have been presented in Literature that address the efficient hardware implementation of convolution.
During my second year of PhD course, I worked on developing an Fast FIR Algorithm to implement the convolution with a reduced number of multiplications by extending and generalizing the algorithm proposed in [13]. The aim of this work is to optimize the mono-dimensional algorithm and then to find a general procedure that allows us to apply the same procedure to the implementation of bi- and multi- dimensional application.
The mono-dimensional convolution can be expressed as

$$y(n) = \sum_{i=0}^{k-1} x(n-i)h(i) \quad (1)$$

where $x(n)$ is the infinite length sequence, $h(n)$ is the finite length sequence, and k is the length of the sequence $h(n)$.
In the developed algorithm the sequence $y(n)$ in (1) can be computed as a linear combination of two kinds of partial values as defined in (2)-(3):

$$P_1(\delta,\gamma,i) = \sum_{j1=\delta}^{\delta+\gamma-1} x((k\cdot i-1)-j1) * \sum_{j2=k-(\delta+\gamma)}^{k-\delta-1} h(j2) \quad (2)$$

$$P_2(\delta,\gamma,i) = [x\big((k\cdot i-1)-\delta\big) + x\big((k\cdot i-1)-(\delta+\gamma)-1\big)] * [h\big(k-(\delta+\gamma)\big) + h(k-\delta-1)] \quad (3)$$

where i indicates the iteration of the algorithm and

$$\delta \in \{0, k-1\} \quad (4)$$

$$\gamma \in \{0, k-1\} \quad (5)$$

Each partial values requires a single multiplication. We have found that a number of partial values lower than k can be used to compute (1) thus reducing the overall number of multiplications to be performed by the accelerator.
A similar approach can be used in multi-dimensional convolution.
In the following Table I show a comparison between the developed algorithm and previous art for bi-dimensional convolution.

Kernel size	Convolution method	Computed Outputs	Multiplications	Additions	Multiplications/ Outputs
3x3	Standard [3] Proposed	1 3 3	9 18 18	8 36 36	9 6 6
5x5	Standard [3] Proposed	1 5 5	25 75 70	24 175 185	25 15 14
7x7	Standard [3] Proposed	1 7 7	49 196 175	48 448 511	49 28 25
11x11	Standard [3] Proposed	11 N/A 11	121 N/A 649	120 N/A 1097	121 N/A 59

Table I – Comparison of FFA algorithms

As can be seen, the proposed algorithm allows for great reduction on the number of multiplications thus improving latency and throughput of the hardware accelerator.
An integrated circuit (IC) implementing the proposed algorithm is in development. The IC will be fabricated in a the 28nm TSMC CMOS technology and contains 42 multipliers and 128 adders. The circuit can be used to accelerate in hardware the operations of the AlexNet neural network with the following performance.

Layer	Kernel	Total multiplications required from standard convolution	Total multiplications required from proposed algorithm	Proposed Datapath Latency (clock cycles)
Conv1	11x11	105×10^6	76×10^6	0.8×10^6
Conv2	5x5	223×10^6	143×10^6	1.75×10^6
Conv3	3x3	149×10^6	115×10^6	1.35×10^6
Conv4	3x3	112×10^6	86×10^6	1×10^6
Conv5	3x3	74×10^6	57×10^6	0.7×10^6

Table II- IC performance related to AlexNet computations, the IC clock period is equal to 2 ns

The design of the IC and of the test environment will be completed in the final year of the Ph.D. course.

2-Development of methodologies for validation and debug of SoC

The validation and debug of SoC requires a great effort and time consumption due to the complexity of these architectures. Thereby, a systematic and methodological approach is needed to address this issue.

Accellera Systems Initiative developed the Universal Verification Methodology (UVM) [14] which was accepted by IEEE as a standard. The UVM is a methodology for the functional verification of digital hardware; thus, this methodology allows developing of reusable and robust simulation environments for SoC validation purposes.

My research activity focuses on :

- development of the UVM simulation environments for state of the art serial interfaces
- implementation of the emulation environment where the UVM verification is accelerated in hardware using SoC devices

The implementation of a verification environment for embedded systems requires a high-level abstraction modeling of both environment and design under test (DUT). This latter might handle packets flowing back and forth, or process instructions, or performs other types of functionalities; therefore, the first level of the abstraction of the DUT is based on the enumeration of all the interfaces or bus signals that the simulation environment must control.

The UVM provides a set of transaction level communication interfaces and channels that can be used to connect components at the system level.

The UVM uses the transaction level modelling (TLM) which allows the verification engineer to control the interface overcoming the need to look at the single signal and at the clock-cycle.

A typical UVM simulation environment architecture is shown in Fig.2. A UVM sequencer is the verification object that handles the transition-level data. The driver converts the transaction-level stimulus into pin-level stimulus while the UVM monitor performs the opposite operation. This latter is a passive component that sends transaction-level information to other components for checking or coverage purposes.

The UVM creates modular, robust environments for the software simulations, however for complex systems this approach could became a bottleneck in terms of time consumptions. As a results, the emulation of the UVM architecture by using a SoC board could drastically decrease the testing-time of a digital design.

The goal of the emulation is to:

- Synthesize the DUT on the programmable logic of the SoC
- Introduce on FPGA additional circuits that emulates the behavior of drivers, sequencers and monitors.
- Create C/C++ routines for the CPU of SoC in order to translate
  - the UVM test and the system level behavior
  - the stimuli randomization
  - the UVM Scoreboard and all the checks needed to validate the DUT
- Reproduce on SoC a test procedure that comes from the UVM Software simulator

Over the last year, I collaborated Micron Hardware Validation Tool Team to implement a UVM-based simulation environment for a custom Micron SoC (All rights are reserved). The SoC architecture includes several hardware accelerators, bus interfaces (such as AXI bus, JTAG bus), and storage interfaces (Universal Flash Storage). By using the UVM approach, the simulation environment is improved in terms of reusability, robustness and coverage analysis. The proposed environment for the SoC validation achieves high speed performance thus reducing the simulation time compared to previous no-UVM simulation environments, developed by Micron.

I have started the development of the emulation environment which will be completed in the last year of the Ph.D. course.

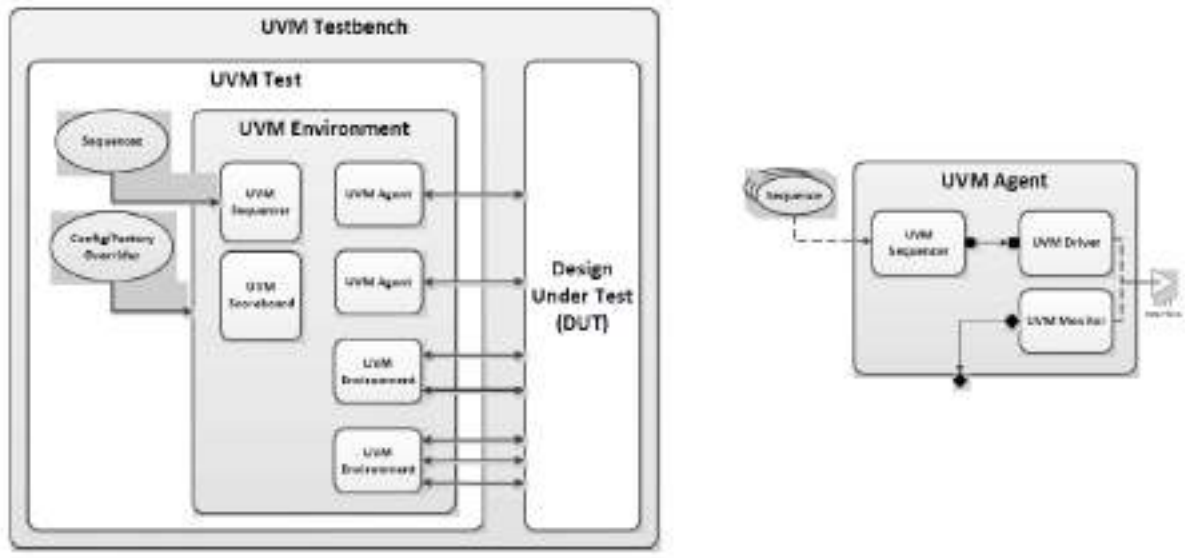


Fig.2 UVM Architecture

References

[1] Chao Cheng and Keshab K. Parhi, "Hardware Efficient Fast Parallel FIR Filter Structures Based on Iterated Short Convolution," Ieee Transactions On Circuits And Systems—I: Regular Papers, vol. 51, no. 8, august 2004, pp., doi
[2] Yu-chi Tsao. Ken Choi, "Area-Efficient Vlsi Implementation For Parallel Linear-Phase Fir Digital Filters of Odd Length Based On Fast Fir Algorithm," IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS, VOL. 59, NO. 6, JUNE 2012
[3] Jichen Wang , Jun Lin, and Zhongfeng Wang, "Efficient Hardware Architectures for Deep Convolutional Neural Network," IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I: REGULAR PAPERS, VOL. 65, NO. 6, JUNE 2018
[4] Chao Cheng and Keshab K. Parhi, "Fast 2D Convolution Algorithms for Convolutional Neural Networks," IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I: REGULAR PAPERS, VOL. 67, NO. 5, MAY 2020
[5] Yizhi Wang , Jun Lin, Senior Member, IEEE, and Zhongfeng Wang, "FPAP: A Folded Architecture for Energy-Quality Scalable Convolutional Neural Networks," IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS–I: REGULAR PAPERS, VOL. 66, NO. 1, JANUARY 2019
[6] Huizheng Wang, Weihong Xu, Zaichen Zhang, Xiaohu You, Chuan Zhang, "An Efficient Stochastic Convolution Architecture Based on Fast FIR Algorithm," IEEE TRANSACTION ON CIRCUITS AND SYSTEMS II: EXPRESS BRIEFS, 2021
[7] Pengbo Zhang , Xue Wang, Senior Member, IEEE , Weihang Zhang, and Junfeng Chen, "Learning Spatial–Spectral–Temporal EEG Features With Recurrent 3D Convolutional Neural Networks for Cross-Task Mental Workload Assessment," IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 27, NO. 1, JANUARY 2019
[8] Ahmet Remzi Ozcan and Sarp Erturk, "Seizure Prediction in Scalp EEG Using 3D Convolutional Neural Networks With an Image-Based Approach," IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 27, NO. 11, NOVEMBER 2019
[9] Xiaoyu Jiang , and Zhiqiang Ge, "Augmented Multidimensional Convolutional Neural Network for Industrial Soft Sensing," IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 70, 2021
[10] Geranun Boonyuu, and Sumek Wisayataksin, "Configurable Hardware Architecture of Multidimensional Convolution Coprocessor," 2021 Second International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP)
[11] Yixing Li,Zichuan Liu,Wenye Liu,Yu Jiang,Wang Ling Goh,Hao Yu and Fengbo Ren, "A 32-FPS 698-GOP/s/W Binarized Deep Neural Network-Based Natural Scene Text Interpretation Accelerator for Mobile Edge Computing", IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS,VOL.66,NO.9,SEMPTEMBER 2019
[12] Li Du, Yuan Du, Yilei Li, Junjie Su, Yen-Cheng Kuan, Chun-Chen Liu, Mau-Chung Frank Chang, "A Reconfigurable Streaming Deep Convolutional Neural Network Accelerator for Internet of Things", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-I:REGULAR PAPERS,VOL.65,NO.1,JANUARY 2018
[13] Marco Vitone, Nicola Petra, "Reconfigurable Datapath for Hardware Acceleration of Convolutional Neural Network",SIE-2021,52nd Annual Meeting of Associazione Società Italiana di Elettronica
[14] Accellera,"Universal Verification Methodology User's Guide",2012

4. Research products:

5. Conferences and seminars attended

- Seminar :
  - Title: "Introduction of Universal Verification Methodology"
  - Lecturer: Marco Vitone
  - Place : Online Microsoft Teams, channel of Ms.C. Course of System on Chip
  - Date: 9/11/2021

6. Periods abroad and/or in international research institutions

7. Tutorship

7. Plan for year three

The activities planned for the third PhD year could be divided in two macro activities:

- The first one regards to section 3.1 and the implementation of the hardware accelerator for Neural Networks, I will realize the dedicated ASIC test chip and test its performance
- The second one regards to section 3.2, I will continue the study of the hardware emulation and I will explore different solutions for the implementation of an automated tool for both software simulation and hardware emulation.