europass

Marco Angioli

ISTRUZIONE E FORMAZIONE

01/11/2022 – ATTUALE Roma, Italia

Dottorato in Tecnologie dell'Informazione e delle Comunicazion
Ingegneria Elettronica Sapienza University Of Rome

Sito Internet https://phd.uniroma1.it/web/MARCO-ANGIOLI_nP1773057_IT.aspx | **Campo**
Machine Learning

09/2016 – 10/12/2019 Italia

Laurea Triennale in Ingegneria Elettronica Sapienza Università d

Voto finale 100/110 | **Tesi** Stato dell'arte nella computazione neuromorfica (neuromorph

09/2019 – 15/07/2022

Laurea Magistrale in Ingegneria Elettronica Sapienza Università

Voto finale 110/110L | **Tesi** Analisi ed implementazione di algoritmi Contextual Bandits s
vettoriale riconfigurabile

COMPETENZE LINGUISTICHE

LINGUA MADRE: Italiano

Altre lingue:

Inglese

Ascolto B1	
Produzione orale B1	
Lettura B1	
Interazione orale B1	
Scrittura B2	

Livelli: A1 e A2: Livello elementare B1 e B2: Livello intermedio C1 e C2: Livello avanzato

COMPETENZE DIGITALI

LINGUAGGI DI PROGRAMMAZIONE
Programmazione VHDL | Programmazione C++ | Programmazione Bash | Matlab
mazione Python | System Verilog

SOFTWARE
Padronanza del Pacchetto Office (Word Excel PowerPoint ecc) | ModelSim | Vivado
nopsis Verdi | Synopsis Fusion Compiler

ALTRE COMPETENZE
Machine Learning | AI and Deep Neural Networks | DIGILENT ZYBO | Embedded L
(FPGA) | Programmazione FPGA | Reinforcement Learning | Conoscenza del Data
Classificazione, Clustering o Regressione | Risoluzione dei problemi | Elaborazione
broswer

SISTEMI OPERATIVI
Windows | Unix Linux

PROGETTI

Design VHDL di un'unità FMA (Fused Multiply Add) ad 8 bit

Progetto sviluppato durante il corso Digital Integrated System Architectures della Laurea Magistrale in Ingegneria Elettronica. Design dell'unità in VHDL, sintesi su Vivado, valutazione delle performance ed ottimizzazioni.

Implementazione della funzione di Motion Detection sulla board Zybo Z7-20 ed ottimizzazione tramite accelerazione hardware

Progetto svolto durante il corso Digital System Programming della Laurea Magistrale in Ingegneria Elettronica. Utilizzo dell'evaluation board Zybo Z7-20 e della telecamera esterna Pcam 5C per la realizzazione di una funzione di motion detection ad alte performance grazie all'accelerazione hardware.

Manutenzione predittiva tramite reti neurali ricorrenti

Progetto svolto durante il corso di Machine Learning for Signal Processing della Laurea Magistrale in Ingegneria Elettronica. Utilizzo di reti neurali ricorrenti su Python per la predizione di possibili malfunzionamenti delle turbine degli aerei. Utilizzo di LSTM, Bi-LSTM e GRU.

Manutenzione predittiva sui motori dei treni

Progetto sviluppato durante il corso di Computational Intelligence della Laurea Magistrale in Ingegneria Elettronica. Realizzazione di una rete che, a partire dall'analisi di campioni di pressione raccolti dai compressori dei treni, permette di prevedere guasti o possibili condizioni fallimentari del motore stesso. Utilizzo di un classificatore SVM con kernel RBF ed algoritmo di ottimizzazione genetico.

Valutazione del grado rotazionale di cifre scritte a mano tramite rete neurale convolutiva

Progetto svolto durante il corso di Circuiti e algoritmi per il Machine Learning della Laurea Magistrale in Ingegneria Elettronica. Utilizzo di una rete neurale convolutiva per la risoluzione di un problema di regressione mirato alla valutazione del grado rotazionale delle cifre scritte a mano contenute nel dataset MNIST.

PUBBLICAZIONI

2022

Contextual Bandits Algorithms for Reconfigurable Hardware Accelerators

Authors: *Marco Angioli, Marcello Barbirotta, Abdallah Cheikh, Antonio Mastrandrea, Francesco Menichelli, Saeid Jamili, and Mauro Olivieri*

Abstract: *Reconfigurable processing cores for IoT and edge computing applications are emerging topics to calibrate costs, energy consumption and area occupation with performance and reliability on Commercial Off the Shelf (COTS) devices. This work analyzes how to take advantage of Machine Learning to potentially automate the reconfiguration process of an Hardware accelerator inside the Klessydra Vector Cooprocessor Unit (VCU) [1][2][3], choosing the best configuration according to the workload. The problem is modeled with a contextual bandits approach using the Linear UCB algorithms and validated with offline Python simulations.*

Accepted at ApplePies,Genova, Italy, 26-27 september 2022, published by Springer, indexed in Scopus

2022

Implementation of Dynamic Acceleration Unit Exchange on a RISC-V Soft-Processor

Authors: *Saeid Jamili, Abdallah Cheikh, Antonio Mastrandrea, Marcello Barbirotta, Francesco Menichelli, Marco Angioli, Mauro Olivieri*

Abstract: Reconfigurable computing, also known as adaptive computing, exploits the reconfigurability of reprogrammable logic devices like FPGAs to perform runtime hardware reconfigurations, enabling the system to better adapt to the underlying application. By using reconfigurable computing, parts of the logic implemented on an FPGA can be dynamically changed according to the task demands during runtime. The underlying hardware can be changed to trade off performance/power or can be modified to perform functional reconfiguration, reprogramming the behavior of a functional unit. By exploiting this flexibility, we can significantly scale the performance and power efficiency of a system. We present a dynamic acceleration unit exchange on a RISC-V soft-processor, based on the open-source Klessydra-T13 RISC-V core. We demonstrate reconfiguration for functional versatility or for improving the hardware accelerator performance, providing, as a case study, an example of how a deep neural network like VGG16 can be accelerated by using runtime reconfiguration techniques.

Accepted at ApplePies,Genova, Italy, 26-27 september 2022, published by Springer, indexed in Scopus

2023

Automatic Hardware Accelerators Reconfiguration through LinearUCB Algorithms on a RISC-V Processor

Authors: Marco Angioli, Marcello Barbirotta, Antonio Mastrandrea, Saeid Jamili, Mauro Olivieri

Abstract: Reconfigurable processors are hardware architectures that allow for the dynamic configuration of processing resources to optimize performance and power consumption, using partial reconfiguration to modify a portion of the design or update it without affecting the entire system. In this work, we present an automatic reconfiguration technique that leverages machine learning (ML) algorithms to automatically select the optimal configuration of a general-purpose hardware accelerator according to the workload and reconFigure the architecture at run-time. The problem is formulated as a Contextual Bandit (CB) case using the Linear Upper Confidence Bound (LinearUCB) algorithms and verified using the RISC-V Klessydra family cores as a case of study.

2023 18th Conference on Ph. D Research in Microelectronics and Electronics (PRIME)

2023

Fault-Tolerant Hardware Acceleration for High-Performance Edge-Computing Nodes

Authors: Marcello Barbirotta, Abdallah Cheikh, Antonio Mastrandrea, Francesco Menichelli, Marco Angioli, Saeid Jamili, Mauro Olivieri

Abstract: High-performance embedded systems with powerful processors, specialized hardware accelerators, and advanced software techniques are all key technologies driving the growth of the IoT. By combining hardware and software techniques, it is possible to increase the overall reliability and safety of these systems by designing embedded architectures that can continue to function correctly in the event of a failure or malfunction. In this work, we fully investigate the integration of a configurable hardware vector acceleration unit in the fault-tolerant RISC-V Klessydra-fT03 soft core, introducing two different redundant vector co-processors coupled with the Interleaved-Multi-Threading paradigm on which the microprocessor is based. We then illustrate the pros and cons of both approaches, comparing their impacts on performance and hardware utilization with their vulnerability, presenting a quantitative large-fault-injection simulation analysis on typical vector computing benchmarks, and comparing and classifying the obtained results. The results demonstrate, under specific conditions, that it is possible to add a hardware co-processor to a fault-tolerant microprocessor, improving performance without degrading safety and reliability.

Electronics 2023

2024

Design, Implementation and Evaluation of a New Variable Latency Integer Division Scheme

Authors: Marco Angioli, Marcello Barbirotta, Abdallah Cheikh, Antonio Mastrandrea, Francesco Menichelli, Saeid Jamili, Mauro Olivieri

Abstract: Integer division is key for various applications and often represents the performance bottleneck due to its inherent mathematical properties that limit its parallelization. This paper presents a new data-dependent variable latency division algorithm derived from the classic non-performing restoring method. The proposed technique exploits the relationship between the number of leading zeros in the divisor and in the partial remainder to dynamically detect and skip those iterations that result in a simple left shift. While a similar principle has been exploited in previous works, the proposed approach outperforms existing variable latency divider schemes in average latency and power consumption. We detail the algorithm and its implementation in four variants, offering versatility for the specific application requirements. For each variant, we report the average latency evaluated with different benchmarks, and we analyze the synthesis results for both FPGA and ASIC deployment, reporting clock speed, average execution time, hardware resources, and energy consumption, compared with existing fixed and variable latency dividers.

IEEE Transactions on Computers 2024

2024

AeneasHDC: An Automatic Framework for Deploying Hyperdimensional Computing Models on FPGAs

Authors: Marco Angioli, Saeid Jamili, Marcello Barbirotta, Abdallah Cheikh, Antonio Mastrandrea, Francesco Menichelli, Antonello Rosato, Mauro Olivieri

Abstract: Hyperdimensional Computing (HDC) is a bioinspired learning paradigm, that models neural pattern activities using high-dimensional distributed representations. HDC leverages parallel and simple vector arithmetic operations to combine and compare different concepts, enabling cognitive and reasoning tasks. The computational efficiency and parallelism of this approach make it particularly suited for hardware implementations, especially as a lightweight, energy-efficient solution for performing learning tasks on resource-constrained edge devices. The HDC pipeline, including encoding, training, and comparison stages, has been extensively explored with various approaches in the literature. However, while these techniques are mainly oriented to improve the model accuracy, their influence on hardware parameters remains largely unexplored. This work presents AeneasHDC, an automatic and open-source platform for the streamlined deployment of HDC models in both software and hardware for classification, regression and clustering tasks. AeneasHDC supports an extensive range of techniques commonly adopted in literature, automates the design of flexible hardware

accelerators for HDC, and empowers users to easily assess the impact of different design choices on model accuracy, memory usage, execution time, power consumption, and area requirements.

IEEE WCCI 2024 - The IEEE World Congress on Computational Intelligence

2024

Exploring Variable Latency Dividers in Vector Hardware Accelerators

Authors: Marco Angioli, Marcello Barbirotta, Abdallah Cheikh, Antonio Mastrandrea, Mauro Olivieri

Abstract: Efficient hardware implementation of integer division remains one of the most significant challenges in the design of vector hardware accelerators, particularly in achieving a balance between performance and hardware overhead. This work uses the RISC-V Klessydra-T13 Vector Coprocessor Unit as a case study to explore the impact of variable latency division architectures for vector hardware accelerators. We analyze existing designs in the literature to identify the most effective approach, detailing the whole integration process with new instructions for supporting vector divisions in the RISC-V custom instruction set and optimizing the implementation for the target accelerator by exploiting hardware reuse. We also provide real-world computation kernels and Monte Carlo simulations to demonstrate how a variable latency divider can significantly improve the division time over traditional methods with negligible hardware overhead.

IEEE Proceedings - 19th Conference on Ph.D Research in Microelectronics and Electronics (PRIME)
