POLITECNICO DI MILANO

PhD in INGEGNERIA DELL'INFORMAZIONE / INFORMATION TECHNOLOGY - 38th cycle

Research Area n. 1 - Computer Science and Engineering

PNRR_352 Research Field: DESIGN METHODOLOGIES FOR THE CUSTOMIZATION AND PROGRAMMABILITY OF IN-MEMORY COMPUTING DEVICES FOR NEURAL NETWORKS

Monthly net income of PhDscholarship (max 36 months)
€ 1400.0 In case of a change of the welfare rates during the three-year period, the amount could be modified.

Context of the research activity	
Motivation and objectives of the research in this field	In-memory computing (IMC) based on emerging memory technologies such as Phase Change Memories is an attractive solution to sustain the significant amount of arithmetic operations needed by a deep neural network. IMC accelerators can enable even more sophisticated inference workloads on low-power edge and IoT devices. From an industrial perspective, however, the costs associated with such application-specific accelerators must be justified by the extent of genericity of workload types that they can sustain, e.g. how many data types, deep learning operators and compression types they do support, whether they enable different types of neural models (e.g. Spiking neural networks) and/or if they can accelerate different data-parallel computation models. . In this scenario, the main goal of the PhD research is to define an innovative design methodology for IMC accelerators to support the design-time customizability and programmability.
Methods and techniques that will be developed and used to carry out the research	- Techniques for introducing support, in IMC, for different levels of activation compression, sparsity of weights and/or activations and (vector/scalar) deep learning operators both at design-time as well as at later stages. - Techniques for mapping and parallelizing NNs onto multi -tile architectures and validation on real world use cases

	(imaging, audio, control, automotive, etc.) - Techniques for the development of neural accelerators that can support non-conventional (e.g., Spiking) neural networks and/or emerging memories (e.g. Phase Change Memory) - Techniques and dedicated instruction set extensions to traditional micro-controllers to enhance the use of IMC tiles (e.g, bitwise vector operations).
Educational objectives	Acquire and/or consolidate knowledge and/or practical skills around: Computer Architectures Architectural exploration Design of Hardware Accelerators Compression, data representation Parallelization, mapping, compilers
Job opportunities	The research proposal addresses an output profile to respond to the needs of the research and design sector of edge and IoT systems, which requires technical experts with highly interdisciplinary skills and in-depth knowledge in the following areas: (i) mathematical / statistical modeling, (ii) design of high-performance parallel computing and data analysis applications for edge and IoT systems (iii) design of accelerators for the Deep Learning sector.
Composition of the research group	1 Full Professors 2 Associated Professors 2 Assistant Professors 2 PhD Students
Name of the research directors	Prof. Cristina Silvano

Contacts
cristina.silvano@polimi.it; +39 02 2399 3692; https://silvano.faculty.polimi.it

Additional support - Financial aid per PhD student per year (gross amount)	
Housing - Foreign Students	--

Housing - Out-of-town residents (more than 80Km out of Milano)	--

Scholarship Increase for a period abroad	
Amount monthly	700.0 €
By number of months	6

National Operational Program for Research and Innovation	
Company where the candidate will attend the stage (name and brief description)	STMicroelectronics Italia
By number of months at the company	6
Institution or company where the candidate will spend the period abroad (name and brief description)	ETH Zurich
By number of months abroad	6

Additional information: educational activity, teaching assistantship, computer availability, desk availability, any other information

Attinenza alla tematiche, alle missioni/componenti prescelte del bando PNRR v. D.M. 352, art.6

The research activities addressed by this proposal are perfectly in line with the "vision" of the National Recovery and Resilience Plan (PNRR) as part of Mission 1: "Digitalization, innovation, competitiveness, culture and tourism" with emphasis on the component M1C2: "Digitalization, innovation and competitiveness in the production system". In particular, we refer to Transition 4.0, aimed at empowering the digital transition of businesses and to the rate of innovation of the industrial and entrepreneurial fabric of the country. The design methodologies techniques to be developed for the acceleration of Deep Learning might represent a significant contribution to several application domains, such as automotive, aeronautics and aerospace and more in general microelectronic devices for mobile and real-time systems. Moreover, design methodologies techniques to the Deep Learning accelerators will reinforce a synergistic research method between university and industry and it will help to sustain the processes for innovation and technology trasfer as well as the competences to support innovation.

Impresa, presso cui si svolgerà l'attività esterna

STMicroelectronics Italia

Ente, università, azienda, centro di ricerca presso cui si svolgerà il periodo di studio e ricerca all'estero.

Research group of prof. Onur Mutlu at ETH Zurich, expert in In-memory computing architectures

All information regarding educational activities, personal funding, regulations and

obligations of Ph.D. candidates are available on the web site https://dottoratoit.deib.polimi.it/