europass

Manuele Rusci

[redacted]

EDUCATION AND TRAINING

[01/11/2014 – 27/04/2018] **Ph.D. in Electronics, Telecommunications, and Information Technologies Engineering**

Università di Bologna

[01/10/2011 – 02/04/2014] **M.Sc. in Electronics Engineering**

Università di Firenze

[01/10/2008 – 30/09/2011] **B.Sc. Information Engineering**

Università di Siena

WORK EXPERIENCE

KU Leuven

City: Leuven | **Country:** Belgium

[01/09/2022 – 31/08/2024] **Marie-Curie Post-Doctoral Fellow**

Greenwaves Technologies

City: Grenoble | **Country:** France

[01/06/2023 – 31/08/2023] **Research Secondment**

Università di Bologna

City: Bologna | **Country:** Italy

[01/11/2017 – 31/08/2022] **Post-Doctoral Reasearcher**

Greenwaves Technologies

City: Bologna | **Country:** Italy

[01/05/2019 – 31/08/2022] **Embedded Machine Learning Consultant**

NXP Semiconductor

City: Eindhoven | **Country:** Netherlands

[01/09/2016 – 28/02/2017] **Internship**

SAITEC

City: Figline Valdarno | **Country:** Italy

[01/05/2014 – 31/10/2014] **FPGA Firmware Engineer**

RESEARCH ACTIVITY

Research Profile

I am specialized in **Embedded Machine Learning** and Deep Learning (aka TinyML) for ultra-low power IoT sensor devices, and multi-core RISC-V Microcontroller Units (MCUs) in particular. My contributions have concerned low-bitwidth quantization methods and optimized software frameworks for inference on MCUs. I have also extensively worked on full-system integration of smart audio or video sensor solutions with a power consumption as low as few mWs (with GreenWaves Technologies).

Recently, my research focus has moved towards **On-Device Learning** to improve at runtime the prediction capacity of IoT smart sensors. In this space, I am investigating the integration of training algorithms on MCUs and efficient (continual) learning strategies for resource-constrained platforms. Thanks to my EU-funded Marie-Curie grant, I have studied novel adaptation mechanisms for ultra-low power speech recognition sensors (SEA2Learn project).

Previously, I was at Universita' di Bologna (Prof. Benini's group), where I was involved in the PULP project - targeting new RISC-V based IoT systems - and several projects in the **Di gital Design Domain**, e.g., FPGA-based equipment for railway applications or, during my PhD, the design of smart camera sensors with event-based cameras.

Main Achievements (recent)

- [2020] Mixed Low Precision Quantization (@ UNIBO), I was the first to show how to bring a Imagenet-scale deep learning model on an MCU using mixed-precision quantization [3][5], thus enabling real-world applications on ultra-low power edge devices.
- [2020] Accelerating Low Precision Inference (@ UNIBO), I contributed the fastest software library for NN inference on multi-core RISC-V MCU, namely PULP-NN [2] (open-source: https://github.com/pulp-platform/pulp-nn). The project was led by a PhD student starting from my initial concept (demo at IROS2018).
- [2021-2023] I demonstrated the first real-world scale audio and visual smart sensor applications with multi-core MCUs (@ GWT). Among the others: automatic license plate detection and recognition [7], object detection on nano-robots [11], audio denoising for hearables [8].
- [2023] Continual Learning on MCUs (@ UNIBO), I led the work with 2 PhD students towards the release of the fastest software library for on-device training, PULPTrainLib [6] (https://github.com/pulp-platform/pulp-trainlib) for continual learning applications [4][9].
- [2024] On-Device Customization and label-free incremental learning (@ KUL). I proposed the first MCU-based solution for on-device customization of the target classification classes (keyword spotting use case), preventing application-specific retraining [10], and a solution for unsupervised incremental learning via pseudo-labeling [12] (https://github.com/mrusci/ondevice-learning-kws).

Technical Experitise

- **RTL Design and FPGA programming**. Design and test of RTL circuits (System Verilog or VHDL) and development on FPGAs (Xilinx SoCs, low-power Microsemi), following an IP-based modular design flow. Validation and testing of application-specific circuits on FPGA boards.
- **HW Architectures of ultra-low power RISC-V-based processors**. Analysis of the system micro-architectures and hardware-software trade-offs for digital signal processing.

- **Software Optimization for Digital Signal Processing on Microcontrollers**. Design and hardware-aware code optimization for accelerating computation-intensive kernels on MCUs. *C* programming for embedded targets.
- **Deep Learning for ultra-low power devices**. Design of memory and computation-efficient Deep Learning algorithms for sensor analysis (using *pytorch*). Data collection and fine-tuning of custom models and integration in the system application code.
- **Deployment frameworks for DNN on MCUs & Quantization**. Methods for porting Deep Learning models on MCUs. This includes quantization and code generation for performance-oriented developments.
- **Audio and Image smart MCU-based sensor integration**. System level deployment with sensor and radio interconnections and debug. System level power cost analysis and board-level design (schematic and components).

PROJECTS

[01/09/2022 – 31/08/2024] **SEA2Learn: SElf-Adaptive and Automated LEARNing Framework for Smart Sensors**

Type of project: Research Project at KUL funded by the Horizon Europe MSCA-PF program (Pillar I, excellent science). Budget: 175k€, individual project. https://cordis.europa.eu/project/id/101067475

Description: This project investigates *continual learning strategies* for a network of ultra-low power speech (non-wearable) sensors, with the overall aim of gaining a system solution that can adapt in the field and achieve high-accurate recognition within the target environment.

My role: I ideated, wrote the proposal, and executed the technical activity of the project.

[2021 – 2024] **On-Device Learning for MicroControllers**

Type of project: Internal Research Project at UNIBO.

Description: This project aimed at evaluating Continual Learning strategies on RISC-V multi-core Microcontrollers and developing an energy-efficient software framework for accelerating training primitives and automatic tools for deployment of the training pipeline.

My role: I have coordinated the activity of 2 PhD students (Leonardo Ravaglia and Davide Nadalini), ideating the roadmap and reviewing the technical activity.

[2021 – 2022] **On-device Audio Denoising System**

Type of project: Internal Industry project at GWT.

Description: I developed and ideated end-to-end a solution for audio denoising on an ultra-low power RISC-V multicore processor [8]. The task involved the selection, design and training of a RNN model for audio denoising and the deployment and system-level integration in the final system. This was the first system demonstration of this kind of application, which was used by the company for marketing.

My role: I led and executed the technical work.

[2019 – 2022] **Efficient Deployment of Inference DNN models on GAP systems**

Type of project: Internal Industry project at GWT.

Description: I have been responsible for benchmarking state-of-the-art Deep Learning models (image and audio) on the processors developed by GWT (GAP8 and GAP9), analyzing the interactions between tools, and algorithms (e.g. quantization, graph optimization) and software runtime. The task involved the measurement of algorithm effectiveness and chip performance, as well as giving feedback to the hardware and

software team. I have been the interface with the marketing and business team to release performance metrics and technical contents for general public or specialized audiences. Among the notable results, the teamwork led to the best rank on the MLPerfTiny benchmark.

My role: I led and executed the technical work and coordinated with the rest of the team.

[2020 – 2022] **Self-localization of Nano-drone Swarm**

Type of project: Industry project at UNIBO, committed by TII (UAE).

Description: Task of a multi-partner industry project concerning the design and realization of a UWB-based localization system for a swarm of nano-drones.

My role: I was the task leader @UNIBO student (Prof. Benini was the PI), coordinating the technical activity of 1 PhD (Mahyar Pourjabar) + reporting to the client.

[2018 – 2023] **Efficient TinyML inference on microcontrollers**

Type of project: Internal research project at UNIBO.

Description: I have investigated methods aiming at enabling deep learning inference on resource-constrained but software-programmable embedded devices, such as Cortex-M class or RISCV multi-core MCUs. The activity focused in particular, on efficient software libraries for quantized neural networks and and mixed-precision (quantization-aware training) methodologies. We developed new software packages (PULP-NN, CMix-NN) and new applications for nano-drones, precision agriculture monitoring, ...

My role: I started the project and developed the technical activity. Over time, I coordinated the activity of various PhD and master students involved in the project (Angelo Garofalo, Luca Bompani, Lorenzo Lamberti, Marco Fariselli, etc..)

[2018 – 2021] **FPGA-based devices for railways control applications**

Type of project: Industry project at UNIBO, committed by RFI (Rete Ferroviaria Italiana).

Description: The project concerned the design and realization of two devices based on FPGA SoC for two railway modules. The design resolved around an FSK demodulator for communication and security.

My role: I initially completed the technical activity of the first devices. The successful work led to the contract extension and more budget (+30k€) to hire 1 young researcher (Guido Bettini), who I trained and I reviewed the work.

[2017 – 2017] **Workload-Aware Dynamic Voltage and Frequency Scaling**

Type of project: Internship during My PhD @ NXP.

Description: I ideated and demonstrated a module for workload-aware DVFS for a novel MCU platform. A patent has been granted on this idea.

My role: I led and executed the project.

[2014 – 2018] **Ultra-low power event-driven smart cameras**

Type of project: My PhD project @ FBK and UNIBO.

Description: I realized a smart camera system based on an event-driven ultra-power image sensor and a multi-core MCU [1]. The system achieved a 10-100x lower power consumption wrt common frame-based systems for monitoring applications.

My role: I led and executed the project.

OTHER ACTIVITIES

Teaching

- Teaching tutor of the master course HARDWARE-SOFTWARE DESIGN (Teacher Prof. Luca Benini). University of Bologna - 2017-2022. The teaching activity concerns HDL programming languages, digital circuit design theory, FPGA design flow and software for embedded systems. During this time, I also developed a lab course on RISC-V MCU programming and DNN acceleration on MCUs (material available on GitHub and still used at UNIBO).
- I have co-advised the activities of 4 Ph.D. students (main supervisor is Prof. Luca Benini) and tutored several master (8) and bachelor (12) theses, in addition to numerous semester student projects (also at ETH Zurich).

Invited Talks

- 2024, M. Rusci, "Adaptive on-device deep-learning for audio and visual sensors at the extreme-edge.", ITEM Workshop, Vilnius (Keynote Speaker).
- 2024, M. Rusci, "On-device Customization of Tiny Deep Learning Models for Keyword Spotting with Few Examples", DATE24 conference, Valencia, (Invited Talk).
- 2023, M. Rusci, "From Deep Inference to On Device Learning with Multi-Core RISC-V MicroControllers at the Extreme Edge", MSCA Apropos Winter School, Zurich, (Invited Talk).
- 2023, M.Rusci, "From Deep Inference to On Device Learning with Multi-Core RISC-V MicroControllers at the Extreme Edge", WIC Mid-winter meeting 2023, Eindhoven (Invited Talk)
- 2023, M. Rusci "TinyDenoiser: RNN-based Speech Enhancement on a Multi-Core MCU with Mixed FP16-INT8 Post-Training Quantization", TinyMLTalk, Online, (Talk).
- 2020, M. Rusci "A Parallel, Ultra-low-power and flexible RISC-V based IoT Application Processor for the TinyML ecosystem", TinyML, Online (Talk).
- 2018, M.Rusci, L. Benini. "An always-on visual node based on a μW imager with contrast-extraction and motion-detection capabilities", IROS, Madrid (Invited Talk).

Tutorials at international venues

- 2024, M. Rusci, D.Nadalini, C. Cioflan "On-device Continual Learning meets ultra low-power processing", DATE24 conference, Valencia.
- 2020, M.Rusci, "Deep Inference on a multi-core RISCV processor at the very edge: Challenges and Tools", AICAS, Online (Tutorial).
- 2020, M. Rusci. "Enabling Embedded AI at the edge of a Network", HIPEAC, Bologna (Tutorial).

Awards

- **2021 Horizon Europe MSCA Postdoctoral Fellowship** (Project: SEA2Learn). This is the most prestigious individual research grant in Europe at my career level.
- Best Paper Award at the 20th Embedded Vision Workshop (CVPR Workshop 2024)
- Best Paper Candidate at IEEE Conference on AgriFood Electronics (CAFÉ) 2023
- Prize sponsored by the company "Austrade per l'Italia S.p.A", reserved to the best 3 students in the first year of a Master Course at the Engineering Faculty of the University of Florence, 2012.

Patents

- *Event-Based Power Manager* registered on 15/11/2017 to USPTO (United States Patent and Trademark Office). The patent is in collaboration with NXP. The patent is about power management policies in ultra-low-power architectures performing digital signal processing. More in details, it envisions the adaptation of functional

characteristics of microcontrollers (e.g. voltage supply, operating frequency), based on information collected from performance counters available within the architecture, easing the power management to the end-user making them pro-active, hence more effective.

Service to the Scientific Community

- TPC Member: ITEM Workshop 2023-2024, DATE 2024, CODAI Workshop 2023-
- I regularly act as a reviewer for the main journal publications in the embedded system domain: ACM Transactions on Embedded Computing Systems, IEEE Transactions on Instrumentation & Measurement, IEEE Journal on Emerging and Selected Topics in Circuits and Systems, IEEE Transactions on Circuits and Systems I: Regular, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

SELECTED PUBLICATION

Full list at: https://scholar.google.it/citations?user=18H0d0gAAAAJ&hl=en

1. **M. Rusci**, D. Rossi, E. Farella and L. Benini, "*A Sub-mW IoT-Endnode for Always-On Visual Monitoring and Smart Triggering*," in IEEE Internet of Things Journal, vol. 4, pp. 1284-1295, 2017.
2. A. Garofalo, **M. Rusci**, F. Conti, D. Rossi and L. Benini. "*PULP-NN: accelerating quantized neural networks on parallel ultra-low-power RISC-V processors*". Philosophical Transactions of the Royal Society A, 378(2164), 2020.
3. A. Capotondi, **M. Rusci**, M. Fariselli, L. Benini. "*CMix-NN: Mixed Low-Precision CNN Library for Memory-Constrained Edge Devices*". IEEE Transactions on Circuits and Systems II: Express Briefs, 67(5), pp.871-875, 2020.
4. L. Ravaglia, **M. Rusci**, D. Nadalini, A. Capotondi, F. Conti and L. Benini. "*A TinyML Platform for On-Device Continual Learning with Quantized Latent Replays*." IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2021.
5. **M. Rusci**, A. Capotondi, and L. Benini. "*Memory-driven mixed low precision quantization for enabling deep network inference on microcontrollers*." Proceedings of Machine Learning and Systems 2, 2020.
6. D. Nadalini, **M. Rusci**, L. Benini and F. Conti. "*Reduced precision floating-point optimization for Deep Neural Network On-Device Learning on MCUs*", in Future Generation Computer Systems, 2023.
7. L. Lamberti, **M. Rusci**, M. Fariselli, F. Paci, and L. Benini. "*Low-Power License Plate Detection and Recognition on a RISC-V Multi-Core MCU-based Vision System*." IEEE International Symposium on Circuits and Systems (ISCAS), 2021.
8. **M. Rusci**, M. Fariselli, M. Croome, F. Paci, and E. Flamand. "*Accelerating RNN-Based Speech Enhancement on a Multi-core MCU with Mixed FP16-INT8 Post-training Quantization*." In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 606-617, 2022.
9. C. Cioflan, L. Cavigelli, **M. Rusci**, M. De Prado, and L. Benini. "*Towards on-device domain adaptation for noise-robust keyword spotting*." In IEEE 4th Int. Conf. on AICAS, 2022.
10. **M. Rusci** and T. Tuytelaars. "*On-device Customization of Tiny Deep Learning Models for Keyword Spotting with Few Examples*." IEEE Micro, 2023.
11. L. Lamberti, L. Bompani, V. Kartsch, **M. Rusci**, D. Palossi, and L. Benini. "*Bio-inspired autonomous exploration policies with cnn-based object detection on nano-drones*". In DATE 2023.
12. **M. Rusci**, F. Paci, M. Fariselli, E. Flamand, and T. Tuytelaars. "*Self-Learning for Personalized Keyword Spotting on Ultra-Low-Power Audio Sensors.*" IEEE Internet of Things Journal, 2024.

Bologna, 15/06/2025

In fede, Dr. Manuele R[redacted]