Fellow	Ramesh Sambangi

Host Organisation	Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

Scientific coordinator	Markus Eppel

I – SCIENTIFIC ACTIVITY DURING YOUR FELLOWSHIP

As part of my **ERCIM Fellowship**, I had contributed to the advancement of hardware accelerator design and exploration for deep neural networks (DNNs), with a particular focus on energy-efficient solutions for Edge AI applications. In modern sensor-based systems, especially those requiring mobility and compact form factors energy efficiency remains a major challenge, particularly for battery-powered or energy-harvesting devices. These systems often struggle to support the local data processing required for real-time inference tasks, typically resorting to external cloud computing resources. This approach, however, incurs significant energy overhead due to wireless transmission and raises concerns about data privacy and latency.

To address these limitations, AI accelerators are increasingly being integrated into sensor nodes, enabling on-chip inference and reducing reliance on external computation. This shift not only minimizes energy consumption but also enhances data security by retaining raw data within the local device. This paradigm, known as Edge AI, supports fast and intelligent data analysis at the network's edge, and is rapidly becoming central to next-generation embedded systems. Among these accelerators, inference engines are critical for executing AI workloads, with digital accelerators currently being mainstream. However, mixed-signal architectures, which combine analog and digital processing through in-memory computing, are emerging as promising alternatives due to their potential for superior energy efficiency.

At the Fraunhofer IIS Erlangen, significant research is underway on the development of mixed-signal AI accelerators that leverage in-memory computing principles. These architectures inherently reduce data movement and energy usage by performing analog computation directly within memory arrays. My specific role involves the design and implementation of a mapper module, a critical component that identifies the optimal configuration settings for mapping DNN workloads onto the underlying accelerator hardware. This task is non-trivial due to the vast design space and the absence of formal methods for accelerator workload mapping.

To address this, I have implemented a Particle Swarm Optimization (PSO)-based mapping algorithm that efficiently explores the configuration space. The algorithm optimizes cost function that evaluates potential mappings based on factors such as energy consumption, computational load distribution, and data movement efficiency. This method allows the mapper to intelligently assign DNN layer weights across multiple computational cores of the mixed-signal accelerator, ensuring balanced workload distribution and minimized inter-core communication. By aligning architectural exploration with optimal mapping strategies, this work directly contributes to the broader goal of realizing high-performance, energy-efficient AI accelerators tailored for edge computing environments.

II – PUBLICATION(S) DURING YOUR FELLOWSHIP

NA

III – ATTENDED SEMINARS, WORKHOPS, CONFERENCES

NA

IV – RESEARCH EXCHANGE PROGRAMME (REP)

As part of my Research Exchange Program, I had the opportunity to visit the **Department of Electronic Systems** at **NTNU, Trondheim campus, Norway**, from **October 7 to October 12, 2024**. I was hosted by **Prof. Per Gunnar Kjeldsberg**, who facilitated my engagement with the department. During my visit, I presented my research work to a diverse audience comprising faculty members, master's students, and PhD scholars. I also had the opportunity to interact with researchers and faculty members to explore potential avenues for future collaboration. Additionally, I visited several research laboratories, which provided valuable insights into the ongoing projects and cutting-edge research being conducted at NTNU. Overall, it was a highly enriching experience to visit one of the leading technical institutes in Europe and engage in meaningful academic exchange with its vibrant research community.

local scientific coordinator Details:

Per Gunnar Kjeldsberg
Professor, acting head of department
Department of Electronic Systems
Faculty of Information Technology and Electrical Engineering
Email: pgk@ntnu.no
+4773594405 +4793459550
B309, Elektro B Gløshaugen, Trondheim