INTERNATIONAL CONFERENCE ON RESEARCHES IN ENGINEERING, SCIENCE, TECHNOLOGY, MANAGEMENT AND HUMANITIES (ICRESTMH – 2024)

25TH AUGUST, 2024

CERTIFICATE NO : ICRESTMH /2024/C0824819

A Study of Neural Network Training for Embedded Systems

Ravinder Gaja

Research Scholar, Ph. D. in Electronics & Communication
Sri Satya Sai University of Technology and Medical Sciences, Sehore, M.P., India.

ABSTRACT

Neural network training for embedded systems is an emerging research domain aimed at integrating the computational power of artificial intelligence into low-power, resource-constrained devices. As embedded systems become increasingly central to applications such as IoT, autonomous vehicles, wearable technology, and industrial automation, there is a growing need for efficient neural network models that can be trained and deployed directly on these platforms. Traditional neural network training is computationally intensive and typically carried out on high-performance servers or cloud-based infrastructures. However, recent advancements in lightweight neural architectures, hardware accelerators, and optimization algorithms have enabled partial or full training of models on embedded systems. Techniques such as quantization, pruning, and transfer learning have been instrumental in minimizing memory footprint and energy consumption without significantly compromising model accuracy. Additionally, on-device training capabilities enhance data privacy and reduce latency, which is critical for real-time and sensitive applications. Despite these advancements, challenges persist in balancing computational efficiency with performance, particularly in dynamic and real-world environments. This paper explores current methods, tools, and trends in neural network training for embedded systems, highlighting the trade-offs and innovations that are shaping this rapidly evolving field. The study also outlines future directions, emphasizing the need for collaborative hardware-software co-design and adaptive learning mechanisms.