



Politecnico
di Torino
SmartData@PoliTO



May 29th 2023, 17:30 PM CEST

SmartTalk: Covivio

<https://smartdata.polito.it/category/smarttalks/>

Philippe Bich

Philippe Bich obtained the B.Sc. in Computer Engineering in 2018 and the M.Sc. in Mechatronics at Politecnico di Torino in 2021. Currently, he is a PhD student of the Department of Electronics and Telecommunications (DET) at Politecnico di Torino (PoliTO), actively collaborating with researchers from the University of Bologna (UniBo), Boston University (BU) and the King Abdullah University of Science and Technology (KAUST) in Saudi Arabia. He is also a member of the SmartData@Polito research center. His research focuses on the optimization of Deep Neural Networks to allow their implementation on edge devices with limited resources.



MAM: a Multiply-And-Max/min Neuron Paradigm for Aggressively Prunable Deep Neural Networks

ABSTRACT

The growing interest in Internet of Things (IoT) and mobile Artificial Intelligence (AI) applications is driving research into Deep Neural Networks (DNNs) that can operate efficiently on low-resource/energy devices at the edge. However, the limited computational resources typically available in edge devices pose a significant challenge for accommodating complex neural networks. Traditional techniques, such as pruning and quantization, have emerged as popular strategies for reducing the size of DNNs.

In this talk, we explore an alternative approach to address the resource constraints on edge devices. While we do not present a novel pruning technique, we introduce a new neuron architecture called Multiply-And-Max/Min, which inherently exhibits a higher propensity for pruning. We discuss the advantages and characteristics of these neurons and showcase their potential to optimize DNNs for edge deployment.



Season 3