

Avis de Soutenance

Monsieur Mateus VIDALETTI COSTA
Electronique, Micro et Nanoélectronique, optique et laser

Soutiendra publiquement ses travaux de thèse intitulés
Accélérateurs de convolution photonique pour les CNNs à grande vitesse

Travaux dirigés par Monsieur Ian O'CONNOR et Monsieur Arnan MITCHELL
Cotutelle avec l'université "Institut royal de technologie de Melbourne" (Australie)

Soutenance prévue le **vendredi 17 juillet 2026** à 9h00
Lieu : 36 Av. Guy de Collongue, bâtiment W1, 69130 Écully
Salle : Amphithéâtre 203

Composition du jury proposé

M. Ian O'CONNOR	Professeur des universités	Ecole Centrale de Lyon	Directeur de thèse
M. Jean-Michel PORTAL	Professeur des universités	AMU-IM2NP - Aix-Marseille Université	Examineur
M. Arnan MITCHELL	Professeur des universités	COMBS - RMIT University	Co-directeur de thèse
Mme Ioana VATAJELU	Chargée de recherche HDR	CNRS-TIMA - Université Grenoble Alpes	Rapporteure
M. Daniel CHILLET	Professeur des universités	ENSSAT-IRISA - Université de Rennes	Rapporteur

Mots-clés : Photonique, Convolution, Kernels, Noise,
Quantization, Multiply-and-Accumulate

Résumé :

La croissance rapide de l'intelligence artificielle (IA) a considérablement augmenté les besoins de calcul des systèmes modernes d'apprentissage automatique. Les réseaux de neurones convolutifs (CNN), largement utilisés en vision par ordinateur, reposent sur un grand nombre d'opérations matricielles et de transferts de données, entraînant une consommation énergétique élevée et une complexité matérielle croissante. Bien que les accélérateurs électroniques spécialisés, tels que les GPU et les TPU, aient soutenu les récents progrès de l'IA, les architectures électroniques conventionnelles restent limitées par la consommation d'énergie, la fréquence d'horloge et le ralentissement de la miniaturisation des transistors. Dans ce contexte, le calcul photonique apparaît comme une alternative prometteuse grâce à sa très large bande passante, son parallélisme intrinsèque et son potentiel d'efficacité énergétique. Cette thèse étudie les réseaux de neurones convolutifs photoniques (PCNN), en se concentrant sur l'impact des non-idéalités matérielles sur la qualité des convolutions et les performances d'inférence. Un cadre de simulation est développé afin d'évaluer le comportement des PCNN dans des conditions de fonctionnement réalistes, en tenant compte des effets de quantification et du bruit. L'étude analyse également l'influence des propriétés des noyaux de convolution sur la robustesse du système face aux contraintes matérielles. Par ailleurs, la thèse explore les compromis entre la précision de quantification, l'exactitude de classification et l'efficacité énergétique. Les résultats montrent qu'une représentation à faible nombre de bits permet de réduire significativement la consommation des convertisseurs numérique-analogique (DAC) tout en maintenant des performances comparables. Des stratégies sont également proposées pour exploiter les ressources de calcul inutilisées de l'architecture, permettant de générer des cartes de caractéristiques supplémentaires sans accroître la complexité matérielle. Dans l'ensemble, ce travail apporte une meilleure compréhension des limites pratiques et des opportunités offertes par les accélérateurs neuronaux photoniques. Il met en évidence leur potentiel en tant que solution évolutive et économe en énergie pour les futurs systèmes matériels dédiés à l'intelligence artificielle.

Summary:

The rapid growth of Artificial Intelligence (AI) has greatly increased the computational requirements of modern machine learning systems. Convolutional Neural Networks (CNNs), which are widely used in computer vision, rely on extensive matrix operations and data movement, leading to high energy consumption and increasing hardware complexity. While electronic accelerators such as GPUs and TPUs have driven recent AI advancements, traditional computing architectures face limitations related to power consumption, clock frequency, and the slowing progress of transistor scaling. To address these challenges, photonic computing has emerged as a promising alternative due to its high bandwidth, inherent parallelism, and potential for energy-efficient processing. This thesis investigates Photonic Convolutional Neural Networks (PCNNs), with a particular focus on how hardware non-idealities affect convolution accuracy and overall inference performance. A simulation framework is developed to evaluate PCNN behavior under realistic operating conditions, including the effects of quantization and noise. The study also examines how different kernel characteristics influence system robustness when hardware constraints are present. Furthermore, the thesis explores the trade-offs between quantization precision, classification accuracy, and energy efficiency. The results show that low-bit operation can significantly reduce Digital-to-Analog Converter (DAC) power consumption while maintaining comparable performance. The work also proposes methods to utilize otherwise unused computational resources within the architecture, enabling the generation of additional feature maps without increasing hardware complexity. Overall, the research provides insights into both the limitations and opportunities of photonic neural network accelerators, demonstrating their potential as scalable and energy-efficient hardware solutions for future AI systems.