

Proposition de thèse

Réseaux de neurones impulsionnels adaptatifs et sécurisés pour des objets connectés autonomes

Adaptive and secure Spiking Neural Networks for autonomous connected objects.

Laboratoire d'accueil : LEAT, campus Sophia Antipolis, Université de Côte d'Azur

Période : octobre 2025 – septembre 2028

Encadrants : Alain PEGATOQUET (Professeur à Université Côte d'Azur), Benoit Miramond (Professeur à Université Côte d'Azur) Ghattas AKKAD (Maitre de Conférences à Université Côte d'Azur) et Matthieu GAUTIER (Professeur à l'IRISA Lannion, Université de Rennes)

Mots clés : Apprentissage automatique (MLP, CNN), Réseaux de neurones impulsionnels (SNN), Electronique numérique, Architecture matérielle, Langage VHDL ou Verilog,

Contacts : alain.pegatoquet@univ-cotedazur.fr, benoit.miramond@univ-cotedazur.fr

Dossier de candidature : CV, dernier diplôme, relevés de notes post-bac, rapport de stage de fin de master ou de fin d'étude, 2 lettres de recommandation ou contact de deux référents.

1. Introduction et contexte

Les réseaux de neurones impulsionnels (SNN) font partie des modèles neuronaux bio-inspirés qui imitent le cerveau humain et la manière dont l'information se propage le long des axones via des impulsions asynchrones codant les événements se produisant en entrée [1]. Les SNN sont ainsi théoriquement plus économes en énergie que les réseaux de neurones classiques (ANN), car les calculs sont rares et ne sont effectués que lorsque les neurones reçoivent des impulsions [2]. Les réseaux de neurones impulsionnels (SNN) représentent ainsi une solution particulièrement prometteuse pour les nœuds IoT autonomes en énergie. Leur comportement dynamique peut par ailleurs être exploité pour s'adapter à l'évolution de l'énergie disponible dans des nœuds autonomes.

Or, le déploiement de nœuds autonomes facilite leur accès physique et offre ainsi à des personnes malveillantes l'opportunité de mener des attaques physiques contre ces systèmes. Ces menaces de sécurité (déni de service, extraction d'informations, ...) représentent un risque important pour les concepteurs de systèmes, les fournisseurs de services et les utilisateurs finaux, et affectent à la fois les architectures ANN et SNN [3][4]. Les problématiques de sécurité des SNN ainsi que leur intégration dans des nœuds autonomes n'ont pas été étudiées de manière approfondie. Le projet SpiNNAch, financé par l'ANR, a pour objectif d'établir une étude comparative de la robustesse des SNN par rapport aux ANN classiques, ceci pour différents modèles tels que les réseaux denses et convolutionnels. L'architecture SPLEAT [5] et son environnement de programmation open-source QUALIA [6], tous deux développés au LEAT, serviront d'environnement de développement et de test pour étudier la sécurité des SNN et mesurer efficacement leurs performances, en termes de latence, de précision et de robustesse, directement sur des cibles embarquées (FPGA). L'objectif de ce projet

de thèse est donc d'étudier la robustesse des SNN et de développer de nouvelles méthodes de contre-mesures contre les attaques et autres failles de sécurité.

2. Objectifs de la thèse

L'objectif de cette thèse est de développer des **modèles SNN adaptatifs pour des nœuds IoT** à récupération d'énergie. Les SNNs proposés offriront des compromis dynamiques entre consommation énergétique et qualité de service (QoS). Les efforts porteront sur le développement de nouvelles techniques exploitant les caractéristiques intrinsèques des SNN, ainsi que sur l'adaptation des techniques existantes issues des ANN vers les SNN.

Dans cette thèse, nous allons proposer d'aborder les problématiques de recherche suivantes :

- Nous proposons tout d'abord d'explorer des **techniques spécifiques aux SNN** pour adapter dynamiquement le nombre de *timestep* [7], le seuil d'activation des neurones [8] ou la sparsité [9], ainsi que différents modèles de quantification, pour les couches convolutives et denses. Nous étudierons, par exemple, la possibilité d'entraîner un modèle SNN partageant ses poids pour différents nombres de *timestep*. Afin de réduire la latence, une autre solution consistera à stopper dynamiquement l'inférence du SNN après un certain nombre limité de *timestep*. Cette méthode d'« **early-stopping** » nécessite un mécanisme de contrôle mais présente l'avantage que le SNN soit entraîné pour un nombre unique de *timestep*.
- Différents **schémas de quantification** (pour les poids, biais et potentiels membranaires) offrant divers compromis entre latence et performance seront étudiés. Un entraînement Quantization-Aware (QAT) des SNN, pris en charge par QUALIA, sera très certainement indispensable pour maintenir les performances.
- Une méthode pour **paralléliser un modèle SNN quantifié donné sur un nombre variable d'unités de traitement neuronal (NPU)** dans l'architecture SPLEAT sera étudié.
- Afin de développer des **SNN capables de supporter différentes configurations architecturales**, nous proposons enfin d'adapter des approches déjà proposées pour les réseaux neuronaux classiques, telles que les **Dynamic Neural Networks** [10] ou encore **SuperNet** [11]. L'objectif est cette fois de faire varier la profondeur du réseau (i.e. le nombre de couches), ainsi que le nombre de canaux par couche, et de proposer typiquement des solutions de type Big/Little en faisant cohabiter deux modèles. Des techniques issues des DNN, telles que l'« **early-exiting** », seront exploitées pour éviter l'exécution des couches les plus profondes lorsque cela est possible.

Les techniques et méthodes proposées aideront à constituer une nouvelle méthodologie d'exécution adaptative des SNN pour améliorer leur robustesse et leur sécurité. Cette méthodologie intégrera également les contre-mesures proposées par l'un des partenaires pour pallier les vulnérabilités et failles de sécurité identifiées. Les études seront principalement menées en simulation, et pourront aboutir à une implémentation matérielle sur FPGA.

Un profiling (en latence et consommation) et des modèles énergétiques de ces solutions SNN adaptatives seront également proposés.

3. Profil recherché

- Nous recherchons un(e) candidat(e) motivé(e), passionné(e) par la recherche, les sciences et les nouvelles technologies.
- Le (la) candidat(e) devra posséder une solide formation en intelligence artificielle et en électronique numérique. Une maîtrise des techniques récentes d'apprentissage automatique (Deep learning) et d'un langage de conception matériel tel que VHDL est indispensable. Des connaissances dans le domaine de l'Internet des Objets (IoT) seraient un plus.
- Cette thèse sera co-supervisée par le LEAT de l'Université Côte d'Azur (UniCA) et l'IRISA de Lannion (Université de Rennes). Elle s'effectuera dans le cadre du projet collaboratif SpiNNach, le candidat sera ainsi amené à interagir avec les différents partenaires du projet et à participer aux réunions d'avancement.
- Le doctorant(e) retenu(e) devra également publier et présenter son travail en anglais à des conférences internationales et dans des revues scientifiques à comité de lecture : une bonne maîtrise de l'anglais oral et écrit est requise.
- Lieu de travail : Laboratoire LEAT, Bât. Forum, Campus SophiaTech, 930 Route des Colles, 06903 Sophia Antipolis.

4. Le laboratoire de recherche (LEAT)

Le **LEAT** est un laboratoire de recherche de l'Université Cote Azur (UniCA) – labélisée Initiative d'Excellence (IDEX), dans le domaine des STIC. Les champs scientifiques portent sur les thèmes antennes, électromagnétisme et micro-ondes et les thèmes objets communicants, réseaux sans fil, intelligence artificielle, systèmes embarqués et Systems on Chip (SoC). Les activités de recherche sont menées dans le domaine des télécommunications, du radar, de l'e-santé, de la sécurité, des bâtiments et villes intelligents, de l'observation de la terre, du développement durable, etc. Le LEAT est organisé en trois équipes : EDGE (Edge Computing and Digital Systems), CMA (Conception et Modélisation d'Antennes) et ISA (Imagerie microonde et Systèmes d'Antennes associés). Compte tenu des objectifs de la thèse, c'est l'équipe EDGE du LEAT qui est logiquement impliquée dans ce projet de recherche.

L'équipe EDGE s'appuie historiquement sur les disciplines scientifiques des systèmes embarqués, de la modélisation et de l'optimisation des réseaux de capteurs sans fil, des Systèmes sur puce (SoC ou System on Chip) et des objets communicants. Une forte activité de recherche dans le domaine de l'intelligence artificielle et plus particulièrement des algorithmes bio-inspirés et des réseaux de neurones dédiés à l'électronique embarquée s'est développée ces dernières années. L'ensemble des travaux de recherche de l'équipe EDGE est structuré selon 3 axes : l'IA bio-inspirée et réseaux de neurones (eBrain), les réseaux sans-fils économes en énergie (eWise) et l'optimisation de la consommation d'énergie dans les Soc (eSoC).

5. Références bibliographiques

- [1] M. Bouvier et al., "Spiking Neural Networks Hardware Implementations and Challenges: A Survey," ACM J. Emerg. Technol. Comput. Syst., vol. 15, no. 2, Apr. 2019.
- [2] E. Lemaire et al., "An Analytical Estimation of Spiking Neural Networks Energy Efficiency," in *Neural Information Processing*, Cham: Springer International Publishing, 2023.
- [3] Y. Hu et al., "Artificial Intelligence Security: Threats and Countermeasures," ACM Comput. Surv., vol. 55, no. 1, Nov. 2021.
- [4] N. Nazari et al., "Securing On-Chip Learning: Navigating Vulnerabilities and Potential Safeguards in Spiking Neural Network.
- [5] N. Abderrahmane et al., "SPLEAT: SPiking Low-power Event-based ArchiTecture for in-orbit processing of satellite imagery," in 2022 International Joint Conference on Neural Networks (IJCNN), Jul. 2022.
- [6] P.-E. Novac et al., "Quantization and Deployment of Deep Neural Networks on Microcontrollers," Sensors, vol. 21, no. 9, Jan. 2021.
- [7] Y. Li et al., "Input-Aware Dynamic Timestep Spiking Neural Networks for Efficient In-Memory Computing," in 2023 60th ACM/IEEE Design Automation Conference (DAC), Jul. 2023.
- [8] A. Castagnetti et al., "SPIDEN: deep Spiking Neural Networks for efficient image denoising," Front. Neurosci., vol. 17, Aug. 2023.
- [9] X. Shi et al., "Towards Energy Efficient Spiking Neural Networks: An Unstructured Pruning Framework," 2024.
- [10] Y. Han et al., "Dynamic Neural Networks: A Survey," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 11, Nov. 2022.
- [11] H. Cai et al., "Once-for-All: Train One Network and Specialize it for Efficient Deployment." arXiv, Apr. 29, 2020.