

LABO - Axe ou Equipe : Institut Pascal, Equipe Comsee

Directeur de thèse : Frederic chausse, email frederic.chausse@uca.fr

Co-encadrant :

Achref Elouni, achref.elouni@uca.fr

Adelaide albouy-kissi, Adelaide.KISSI@uca.fr

Oussama habachi, oussama.habachi@uca.fr

Titre du sujet de thèse : Apprentissage fédéré auto-supervisé avec parallélisme intra-couche et orchestration dynamique des ressources

Résumé du sujet de thèse : L'apprentissage auto-supervisé s'impose aujourd'hui comme un moteur essentiel de l'IA moderne [2][3], permettant d'apprendre à partir de données non annotées. Malgré son efficacité, il nécessite des infrastructures de calcul massives et coûteuses [6][7], difficiles à utiliser dans des environnements distribués. À l'inverse, l'apprentissage fédéré offre un cadre décentralisé où chaque appareil contribue à un modèle global tout en conservant ses données localement [8], ce qui correspond bien aux contextes IoT, industriels ou embarqués. Pourtant, l'hétérogénéité des ressources et des données rend les approches existantes difficiles à stabiliser et souvent inefficaces.

Cette thèse explore une nouvelle voie : l'exploitation d'un parallélisme intra-couche, permettant à plusieurs machines de coopérer pour entraîner les neurones d'une même couche, contrairement aux approches traditionnelles basées sur des couches entières [9][10][11]. Combinée à une orchestration dynamique des ressources, capable d'adapter automatiquement la répartition du calcul, cette approche pourrait rendre l'apprentissage fédéré auto-supervisé plus rapide, plus flexible et plus robuste.

Les progrès récents de l'intelligence artificielle reposent sur des modèles profonds nécessitant des volumes considérables de données [4]. L'apprentissage auto-supervisé, qui apprend sans annotations, a permis des avancées majeures dans la vision et le langage [2][3]. Toutefois, il reste fortement dépendant de centres de calcul puissants [6], ce qui limite son utilisation dans des environnements distribués. À l'inverse, l'apprentissage fédéré permet à chaque appareil de contribuer à un modèle global sans partager ses données [8], une approche particulièrement adaptée aux réseaux IoT, aux systèmes embarqués ou aux infrastructures industrielles.

Cependant, ces environnements sont extrêmement variés : certains appareils disposent d'une grande puissance de calcul, d'autres très peu ; certaines connexions sont stables, d'autres intermittentes [13]. Les données sont hétérogènes et les échanges de paramètres coûtent cher, ce qui ralentit l'entraînement. Les méthodes actuelles tentent d'atténuer ces limites : SplitNN propose un découpage client-serveur rigide [10], FedLAMA adapte la synchronisation entre couches [9], et Lw-FedSSL réduit la mémoire en entraînant couche par couche [11]. Néanmoins, elles partagent toutes la même contrainte structurelle : une couche est toujours considérée comme un bloc indivisible.

Cette thèse propose un changement conceptuel majeur : exploiter un parallélisme intra-couche, où les neurones d'une même couche sont répartis entre plusieurs machines. Une telle granularité fine, encore inexplorée dans le cadre fédéré, introduit de nouveaux enjeux méthodologiques, notamment pour garantir la stabilité et la convergence d'un modèle dont les neurones sont distribués entre plusieurs nœuds. Elle nécessite également de repenser l'organisation des échanges afin de limiter les communications tout en maintenant un niveau de performance élevé. Les travaux existants ne considèrent pas ces problématiques, car ils se limitent à une granularité par couche [9][10][11].

Le second axe majeur de la thèse consiste à développer une orchestration dynamique, inspirée de travaux en optimisation distribuée et en personnalisation fédérée [12][14]. Cet orchestrateur devra adapter en temps réel la répartition des neurones, la fréquence des synchronisations et l'allocation des

ressources. Dans un système fédéré, certaines machines sont surchargées, d'autres inactives, la bande passante fluctue et la disponibilité varie continuellement. En tenant compte de ces contraintes, l'orchestrateur permettra d'assurer une coopération neuronale stable, réactive et efficace.

Les approches actuelles restent trop rigides pour exploiter pleinement les environnements distribués modernes. En introduisant une granularité nouvelle centrée sur les neurones et en l'associant à une orchestration intelligente, la thèse propose un cadre inédit pour l'apprentissage fédéré auto-supervisé. Ce travail se positionne au croisement de l'auto-supervision [2][3], de l'apprentissage distribué [8][9], et des architectures neuronales avancées [5], tout en ouvrant de nouvelles perspectives pour les systèmes embarqués et les réseaux distribués.

Les méthodes proposées dans cette thèse seront évaluées sur une base de données issue du domaine agro-environnemental, avec un accent particulier sur l'analyse des images de sols pour la quantification des micro-organismes (bactéries, champignons, nématodes, ...), caractérisant la santé des sols.

Planning

La première année débutera par un premier semestre entièrement consacré à l'état de l'art portant sur l'apprentissage auto-supervisé, l'apprentissage fédéré et les différentes formes de parallélisme de modèles. Cette revue approfondie permettra de comprendre précisément les limitations actuelles et d'identifier les pistes pouvant bénéficier d'un parallélisme intra-couche. Au second semestre, le doctorant définira le cadre conceptuel du projet en formalisant les principes de base de cette nouvelle forme de coopération neuronale. Il développera également les premiers prototypes afin de tester la faisabilité pratique du modèle et d'identifier les premiers défis liés à la convergence, à la communication et à la répartition des calculs.

La deuxième année sera consacrée au développement de l'orchestrateur dynamique, qui jouera un rôle central dans l'adaptation en temps réel de la répartition des neurones et de l'utilisation des ressources. Le doctorant explorera différentes approches d'apprentissage par renforcement ou d'optimisation adaptative pour concevoir un système capable de prendre des décisions éclairées en fonction de l'état des machines. Il mettra en place des environnements expérimentaux reproduisant des conditions réalistes d'hétérogénéité matérielle, de variabilité réseau et de fluctuations de charge. Une fois opérationnel, l'orchestrateur sera intégré au pipeline auto-supervisé afin d'évaluer ses bénéfices en termes de vitesse d'entraînement, d'efficacité globale et de réduction des coûts.

La troisième année commencera par l'intégration complète des modules développés au cours des deux années précédentes, suivie d'expérimentations à grande échelle pour valider le système dans des environnements variés. Le doctorant évaluera la solution sur des plateformes distribuées réelles afin de mesurer l'adaptabilité, la robustesse et les gains de performance apportés par l'approche. Il ajustera ensuite les paramètres du système et consolidera les résultats afin de garantir des contributions solides et scientifiquement valides.

Enfin, les six derniers mois seront exclusivement consacrés à la rédaction du manuscrit de thèse, à la finalisation des publications et à la préparation de la soutenance.

[1] J. Xu, Z. Chen, S. Yang, J. Li, H. Wang, and E. C.-H. Ngai, "Mentor: Multi-level Self-Supervised Learning for Multimodal Recommendation," Proc. AAAI Conf. Artif. Intell., vol. 39, pp. 12908–12917, 2025.

[2] J.-B. Grill et al., "Bootstrap Your Own Latent: A new approach to self-supervised learning," Adv. Neural Inf. Process. Syst., 2020.

[3] M. Caron et al., "Emerging properties in self-supervised vision transformers," IEEE Int. Conf. Comput. Vis. (ICCV), 2021.

- [4] T. B. Brown et al., “Language Models Are Few-Shot Learners,” Adv. Neural Inf. Process. Syst., 2020.
- [5] A. Ouni, C. Samir, Y. Bouaziz, and A. Fradi, “ConvKAN: Towards robust, high-performance and interpretable image classification,” Proc. VISSAP, 2025.
- [6] D. Patterson et al., “Carbon emissions and large neural network training,” arXiv preprint arXiv:2104.10350, 2021.
- [7] Deepak et al., “Efficient large-scale language model training on GPU,” USENIX OSDI, 2021.
- [8] B. McMahan, E. Moore, D. Ramage, B. Hampson, and B. A. Arcas, “Communication-efficient learning of deep networks from decentralized data,” Proc. AISTATS, 2017.
- [9] Y. Chen, T. Huang, J. Wang, C. Xu, and D. Tao, “FedLAMA: Layer-wise adaptive model aggregation for scalable federated learning,” IEEE Int. Conf. Comput. Vis. (ICCV), 2022.
- [10] P. Vepakomma, O. Gupta, A. Swedish, and R. Raskar, “Split learning for health: Distributed deep learning without sharing raw patient data,” Proc. AAI, 2020.
- [11] Anonymous Author, “Lw-FedSSL: Resource-efficient layer-wise federated self-supervised learning,” arXiv preprint, 2024.
- [12] C. T. Dinh, N. Tran, and T. Nguyen, “Personalized federated learning with Moreau envelopes,” Adv. Neural Inf. Process. Syst., 2020.
- [13] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated optimization in heterogeneous networks,” arXiv preprint arXiv:1812.06127, 2020.
- [14] S. P. Karimireddy et al., “SCAFFOLD: Stochastic controlled averaging for on-device federated learning,” Proc. Int. Conf. Mach. Learn. (ICML), 2020.

Publications /références associées au sujet :

A. Ouni, C. Samir, Y. Bouaziz, and A. Fradi, “ConvKAN: Towards robust, high-performance and interpretable image classification,” Proc. VISSAP, 2025.

CAUNES, Andrew, CHATEAU, Thierry, et FRÉMONT, Vincent. 3d can be explored in 2d: Pseudo-label generation for lidar point clouds using sensor-intensity-based 2d semantic segmentation. In : 2024 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2024. p. 2192-2197.

Leveraging semantic segmentation for hybrid image retrieval methods A Ouni, E Royer, M Chevaldonné, M Dhome, Neural Computing and Applications 34 (24), 21519-21537

Deep model compression and architecture optimization for embedded systems: A survey; A Berthelie, T Chateau, S Duffner, C Garcia, C Blanc, Journal of Signal Processing Systems 93 (8), 863-878

Toward industrial use of continual learning: new metrics proposal for class incremental learning, MA Konaté, AF Yao, T Chateau, P Bouges, 2023 International Joint Conference on Neural Networks (IJCNN), 01-07

