

**Institut Pascal – ISPR / DREAM team**

**Thesis subject: Federated AI on a wireless smart camera network: a unified architecture based on autoencoders and distributed ledger**

**Thesis supervisor: François BERRY (PR) - Institut Pascal, [francois.berry@uca.fr](mailto:francois.berry@uca.fr)**

**Co-supervisors: François DELOBEL (MCF)-LIMOS, [francois.delobel@uca.fr](mailto:francois.delobel@uca.fr), Christophe BLANC (MCF)-Institut Pascal), [christophe.blanc@uca.fr](mailto:christophe.blanc@uca.fr)**

**Funding: This thesis is funded by the MIAI Embed-AI research chair, whose goal is to propose hardware frugality solutions for AI.**

With the proliferation of intelligent visual sensors, wireless camera networks are becoming a key research area for numerous applications, such as surveillance, object recognition and real-time scene analysis. However, optimizing data processing and transmission remains a major challenge due to bandwidth constraints, energy consumption and the limited computational capacity of embedded sensors. In a spirit of frugality, this thesis proposes an original distributed architecture based on federated artificial intelligence and a distributed ledger mechanism built on latent representations, in order to optimize information exchange within the network while ensuring data confidentiality and traceability.

To preserve the privacy of individuals, the cameras used are of very low resolution (32×32 pixels), thereby reducing the risk of identification while ensuring the collection of relevant information for the targeted applications. These cameras already exist and consist of a low-resolution sensor, an analog PIR sensor, a small FPGA (Max10 – low cost and low power) and an ESP32-C6 for managing network communication via Bluetooth.

#### *Research problem*

The use of wireless cameras in a dynamic environment imposes contradictory requirements. On the one hand, it is necessary to capture and analyze data to ensure optimal accuracy. On the other hand, energy consumption and transmission latency must be minimized. The existing hardware architecture, based on a small FPGA (Max10) and an ESP32 for Bluetooth communication, imposes strong constraints in terms of embedded processing optimization and exchange protocol efficiency.

To address these constraints, we propose a unified and original approach: each network node integrates an autoencoder whose latent space serves a dual purpose. On the one hand, it acts as a compressed representation of visual information, significantly reducing the volume of data to be transmitted over the Bluetooth network. On the other hand, each timestamped latent vector constitutes an entry in the distributed ledger, replicated across all nodes, allowing the reconstruction of the observation history by querying any node in the network. Federated AI then intervenes to align the latent spaces across different nodes, ensuring the interoperability of representations without requiring centralization of raw data.

#### *Research objectives*

The main objective is to design a unified architecture where the autoencoder embedded on each node serves as the cornerstone: it simultaneously ensures information compression, confidentiality preservation (latent representations are not visually interpretable) and the feeding of the distributed ledger.

This objective is broken down into four axes: designing quantized autoencoders (8 bits or less) capable of running on a Max10 FPGA while optimizing the trade-off between latent size, reconstruction quality and hardware resources; building a distributed ledger where each block is a timestamped latent, cryptographically chained and replicated across all nodes, with a potentially hybrid consensus mechanism (cryptographic combined with semantic coherence); implementing federated learning to align latent spaces across nodes and ensure their interoperability without ever exchanging raw images; and optimizing the Bluetooth protocol for exchanging latent vectors and federated update parameters under bandwidth and energy constraints.

### *Methodological approach*

The methodology relies on a hybrid approach combining theory, simulation and experimentation on existing hardware, organized in four phases. The first phase is devoted to the design and implementation on the Max10 FPGA of quantized autoencoders (8-bit, 4-bit, binary) for 32×32 images, exploring efficient architectures and evaluating the trade-off between latent size, reconstruction quality and hardware resources. The second phase focuses on the distributed ledger: cryptographic chaining of latent vectors, Bluetooth replication protocol, management of the ESP32 memory constraints, history pruning mechanisms and formalization of the hybrid crypto-semantic consensus. The third phase is dedicated to federated learning for inter-node autoencoder alignment, with the evaluation of aggregation strategies adapted to network constraints and the study of latent space drift due to heterogeneous conditions across sensors. The fourth phase consists of integration on a real prototype followed by a global evaluation covering energy consumption, latency, AI accuracy, reconstruction quality from latent vectors and distributed ledger integrity.

### *Expected impact and contributions*

The main contribution lies in the proposal of a unified architecture where the autoencoder's latent space serves as the native support for the distributed ledger, merging compression, privacy protection and data traceability into a single operation. This is particularly suited to ultra-constrained nodes where each operation must serve multiple objectives. The thesis will also contribute to pushing the limits of federated learning on low-cost FPGAs and microcontrollers, a still largely unexplored area since the literature focuses on more powerful nodes such as smartphones or edge servers, with contributions on model quantization and heterogeneous latent space alignment. The custom-designed distributed ledger mechanism, with its hybrid cryptographic and semantic consensus, will open avenues applicable to other IoT sensor networks facing the same limitations. Finally, confidentiality is ensured at three complementary levels – low sensor resolution, exclusive exchange of visually non-interpretable latent vectors, and no centralization of raw data thanks to federated learning – which constitutes a coherent multi-layer privacy-by-design approach, paving the way for varied applications ranging from surveillance to autonomous object recognition in a context of frugality.

### *Conclusion*

The originality of this thesis lies in the deep integration between embedded autoencoder, distributed ledger and federated learning within a wireless camera network with ultra-constrained resources. By making the latent space the vehicle for compression, confidentiality and distributed traceability alike, this architecture proposes a coherent and frugal solution to the challenges posed by intelligent visual sensor networks. The use of very low-resolution cameras, combined with decentralized collaborative learning, will reconcile performance, privacy protection and energy sobriety, while opening new perspectives for embedded technologies.

**Institut Pascal – ISPR / équipe DREAM**

**Directeur de thèse : François BERRY (PR)- Institut Pascal, [francois.berry@uca.fr](mailto:francois.berry@uca.fr)**

**Co-encadrant : François DELOBEL (MCF)- LIMOS, [francois.delobel@uca.fr](mailto:francois.delobel@uca.fr), Christophe BLANC (MCF) - Institut Pascal, [christophe.blanc@uca.fr](mailto:christophe.blanc@uca.fr)**

**Titre du sujet de thèse : IA fédérative sur un réseau de caméras intelligentes sans fil**

**Financement : Cette thèse est financée par la chaire de recherche MIAI Embed-AI dont le but est de proposer des solutions de frugalités matérielles pour l'IA.**

Avec la prolifération des capteurs visuels intelligents, les réseaux de caméras sans fil deviennent un domaine de recherche essentiel pour de nombreuses applications, telles que la surveillance, la reconnaissance d'objets et l'analyse de scènes en temps réel. Cependant, l'optimisation du traitement des données et de leur transmission reste un défi majeur en raison des contraintes de bande passante, de consommation d'énergie et de capacité de calcul des capteurs embarqués. Dans une logique de frugalité, cette thèse propose une architecture distribuée originale reposant sur l'intelligence artificielle fédérative et un mécanisme de registre distribué fondé sur des représentations latentes, afin d'optimiser les échanges d'informations dans le réseau tout en garantissant la confidentialité et la traçabilité des données.

Pour préserver la vie privée des individus, les caméras utilisées sont de très basse résolution (32×32 pixels), réduisant ainsi le risque d'identification tout en assurant une collecte d'informations pertinentes pour les applications visées. Ces caméras sont déjà existantes et sont composées d'un capteur basse résolution, d'un capteur analogique PIR, d'un petit FPGA (Max10 – Low cost et low power) et d'un ESP32-C6 pour la gestion de la communication en réseau via Bluetooth.

### *Problématique*

L'utilisation de caméras sans fil dans un environnement dynamique impose des exigences contradictoires. D'une part, il est nécessaire de capturer et d'analyser des données pour garantir une précision optimale. D'autre part, il faut minimiser la consommation énergétique et la latence de transmission. L'architecture matérielle existante, reposant sur un petit FPGA (Max10) et un ESP32 pour la communication en Bluetooth, impose des contraintes fortes en termes d'optimisation des traitements embarqués et d'efficacité du protocole d'échange.

Face à ces contraintes, nous proposons une approche unifiée et originale : chaque nœud du réseau intègre un auto encodeur dont l'espace latent remplit une double fonction. D'une part, il sert de représentation compressée de l'information visuelle, réduisant considérablement la quantité de données à transmettre sur le réseau Bluetooth. D'autre part, chaque vecteur latent horodaté constitue une entrée du registre distribué, répliqué sur l'ensemble des nœuds, permettant de reconstituer l'historique des observations en interrogeant n'importe quel nœud du réseau. L'IA fédérative intervient alors pour aligner les espaces latents entre les différents nœuds, garantissant l'interopérabilité des représentations sans nécessiter de centralisation des données brutes.

### *Objectifs de la recherche*

L'objectif principal est de concevoir une architecture unifiée où l'auto encodeur embarqué sur chaque nœud est la pierre angulaire : il assure à lui seul la compression de l'information, la préservation de la confidentialité (les latents ne sont pas interprétables visuellement) et l'alimentation du registre distribué.

Cet objectif se décline en quatre axes : concevoir des auto encodeurs quantifiés (8 bits ou moins) capables de tourner sur un FPGA Max10 en optimisant le compromis taille du latent / qualité de reconstruction / ressources matérielles ; construire un registre distribué où chaque bloc est un latent horodaté, chaîné cryptographiquement et répliqué sur tous les nœuds, avec un consensus potentiellement hybride (crypto + cohérence sémantique) ; mettre en œuvre un apprentissage fédéré qui aligne les espaces latents entre nœuds pour garantir leur interopérabilité sans jamais échanger d'images brutes ; et optimiser le protocole Bluetooth pour l'échange des latents et des paramètres fédérés sous contraintes de bande passante et d'énergie.

### *Approche méthodologique*

La méthodologie s'appuie sur une approche hybride combinant théorie, simulation et expérimentation sur le matériel existant, organisée en quatre phases. La première est consacrée à la conception et l'implémentation sur FPGA Max10 d'auto encodeurs quantifiés (8 bits, 4 bits, binaire) pour les images 32×32, en explorant des architectures efficaces et en évaluant le compromis taille du latent / qualité de reconstruction / ressources matérielles. La deuxième phase porte sur le registre distribué : chaînage cryptographique des latents, protocole de réplication Bluetooth, gestion des contraintes mémoire de l'ESP32, mécanismes d'élagage de l'historique et formalisation du consensus hybride cryptosémantique. La troisième phase est dédiée à l'apprentissage fédéré pour l'alignement inter-nœuds des auto encodeurs, avec l'évaluation de stratégies d'agrégation adaptées aux contraintes du réseau et l'étude de la dérive des espaces latents due aux conditions hétérogènes entre capteurs. La quatrième phase consiste en l'intégration sur un prototype réel suivie d'une évaluation globale portant sur la consommation énergétique, la latence, la précision IA, la qualité de reconstruction depuis les latents et l'intégrité du registre distribué.

### *Impact et contributions attendues*

La contribution principale réside dans la proposition d'une architecture unifiée où l'espace latent de l'auto encodeur sert de support natif au registre distribué, fusionnant en une seule opération la compression, la protection de la vie privée et la traçabilité des données, ce qui est particulièrement adapté aux nœuds ultra-contraints où chaque opération doit servir plusieurs objectifs. La thèse contribuera également à repousser les limites de l'apprentissage fédéré sur des FPGA low-cost et des microcontrôleurs, un terrain encore peu exploré puisque la littérature se concentre sur des nœuds plus puissants comme les smartphones ou les serveurs « edge », avec des apports sur la quantification des modèles et l'alignement d'espaces latents hétérogènes. Le mécanisme de registre distribué conçu sur mesure, avec son consensus hybride cryptographique et sémantique, ouvrira des pistes transposables à d'autres réseaux de capteurs IoT confrontés aux mêmes limitations. Enfin, la confidentialité est assurée à trois niveaux complémentaires — basse résolution des capteurs, échange exclusif de latents non interprétables visuellement, et absence de centralisation des données brutes grâce au fédéré — ce qui constitue une approche de « privacy by design » multicouche cohérente, ouvrant la voie à des applications variées allant de la surveillance à la reconnaissance d'objets autonomes dans un contexte de frugalité.

### *Conclusion*

L'originalité de cette thèse réside dans l'intégration profonde entre auto encodeur embarqué, registre distribué et apprentissage fédéré au sein d'un réseau de caméras sans fil à ressources ultra-contraintes. En faisant de l'espace latent le vecteur à la fois de la compression, de la confidentialité et de la traçabilité distribuée, cette architecture propose une solution cohérente et frugale aux défis posés par les réseaux de capteurs visuels intelligents. L'utilisation de caméras de très basse résolution, couplée à l'apprentissage collaboratif décentralisé, permettra de concilier performance, respect de la vie privée et sobriété énergétique, tout en ouvrant de nouvelles perspectives pour les technologies embarquées.