

Sentinelles – IA embarquée

Traitements décentralisés et optimisation des échanges

Philippe VOISIN¹, Christophe HENNEQUIN¹, Pierre RAYMOND¹,
Nicolas HUEBER¹ et Alexander PICHLER¹

¹ Institut franco-allemand de recherches de Saint-Louis, Saint-Louis, France

Résumé

L'objectif de nos travaux de recherche est de contribuer à la constitution d'une situation tactique locale partagée entre plusieurs entités d'un même théâtre d'opération. Le contexte hostile et des ressources réseaux extrêmement limitées imposent l'optimisation du couple décentralisation des traitements et gestion optimale des flux échangés. Nos travaux ISL portent sur l'intégration au niveau du capteur de traitements et concepts d'IA pour le rendre plus autonome, plus simple à mettre en œuvre et plus robuste à son environnement. Les capteurs ISL intègrent une capacité à détecter, classifier et archiver localement les événements placés dans leur champ de vision. Ils apportent une solution 24h/24 avec un fonctionnement dans le domaine visible ou infrarouge. Le fil conducteur de ces travaux est une maîtrise totale des aspects matériels et logiciels des capteurs conçus et réalisés. Des mécanismes de Machine Learning et de Deep Learning embarqués au niveau composant permettent l'analyse et la compréhension de l'environnement pour une vision tactique opérationnelle locale et une aide à la prise de décision.

Mots clés : Réseaux de neurones, Deep Learning, Détection, Reconnaissance, Identification, Système Embarqué, Infrarouge, SWaPT.

1 Observation et renseignement en milieu hostile

Le défi est d'obtenir le plus rapidement et efficacement possible, par des UGS¹, une situation tactique opérationnelle locale pour la surveillance de FOB², de sites, d'aires de stockage, d'ouverture de convoi, ... Le milieu est hostile et les médias de communication y sont limités voire inexistant. En conséquence, les capteurs déposés sont autonomes et surveillent discrètement une grande zone. Pour ne pas augmenter la charge de travail humaine, ils intègrent des fonctions de DRI³ et d'archivage des informations. Une communication temps réel des événements vers un poste de contrôle est possible par une liaison radio ou filaire.

¹ Unattended Ground Sensor

² Forward Operational Base

³ Détection, de Reconnaissance et d'Identification

1.1 Contexte et contraintes de terrain

La couverture de vastes zones ainsi que l'analyse d'environnements complexes imprédictibles sont génératrices de grands flots de données. Transmettre des informations brutes, en milieu hostile, est aujourd'hui inconcevable. L'idée directrice est de minimiser les échanges d'information entre le capteur positionné au plus près du danger et l'utilisateur en position de sécurité. De plus, afin de conserver une certaine discrétion, minimiser la détectabilité et augmenter l'autonomie énergétique, le capteur ne doit transmettre au moment opportun qu'une information de très haut niveau.

Pour atteindre cet objectif, il faut imaginer et adapter des outils de DRI innovants et réalistes pouvant s'accommoder des ressources calculatoires limitées disponibles au niveau des capteurs. Pour traiter les cas complexes de l'analyse de l'environnement et du mouvement, l'IA embarquée permet de dupliquer l'expertise humaine au niveau du capteur. Le capteur de terrain devient alors un expert électronique pouvant être dupliqué et positionné en zone sensible.

Cette étude du domaine défense et sécurité a relevé plusieurs défis technologiques et scientifiques aux niveaux logiciels et matériels. Concernant le défi technologique, les capteurs doivent répondre à des critères de robustesse, de taille, de poids, d'énergie et de temps de calcul. Le défi scientifique consiste à réaliser localement le maximum de reconnaissances par seconde et par Watt avec un taux de fausses alarmes le plus faible possible.

L'approche retenue pour la fonction Détection (DRI) est un filtre Sigma-Delta [1, 2,10]. Les événements les plus pertinents sont sélectionnés et labellisés sur des critères de taille, de vitesse et de temps de présence. La fonction analyse est réalisée par une d'IA embarquée et entraînée sur ordinateur à partir d'une base de données. Pour la fonction de Reconnaissance (DRI), seule l'architecture du réseau est portée au niveau capteur. Le principal challenge a consisté à adapter et coder ces fonctions complexes pour être implémentés sur un composant de type SoC⁴ tout en maîtrisant la consommation électrique. Cette consommation ne devra pas excéder 6W.

2 L'IA au niveau du capteur

Notre objectif est d'être en mesure de suivre une cible à bord d'UxV⁵ et de réaliser des missions de surveillance et de renseignement par des UGS. Dans les deux cas, le milieu est non-prédictif. Nous avons évalué et porté sur le même composant FPGA deux approches complémentaires de l'IA :

1. le classifieur TEAM⁶_ISL pour des fonctions de poursuite incrémentale de cible
2. les réseaux DNN-SSD⁷ [11] pour des fonctions de Reconnaissance-Identification (DRI)

⁴ System on Chip

⁵ Unmanned Vehicle

⁶ Target Evolutive Associative Memory

⁷ Deep Neural Network, Single Shot Detector

2.1 Les moteurs d'IA câblés de Sentinelle

Approche TEAM. Le processeur TEAM [3], se compose de trois classifieurs hautement parallélisés sur un composant FPGA Xilinx : kNN⁸, PNN⁹ et RBF¹⁰. Cette approche permet de gérer le nombre de neurones à implémenter en fonction l'application. La dimension du vecteur d'entrée est paramétrable [256-2048] ainsi que le nombre de neurones [256-4096].

- Le réseau KNN recherche les k plus proches voisins entre un échantillon d'entrée et un ensemble de prototypes enregistrés. Il utilise la norme L1. A un neurone correspond un prototype.
- Pour éviter tout biais lié au nombre d'échantillons par classe, le réseau PNN fournit une densité de probabilité de classification qui tient compte de la population de chaque classe.
- Le réseau RBF apporte une capacité de reconnaissance et de généralisation. Dans ce cas, un neurone code pour un ensemble de prototypes d'entrée appartenant au même espace de décision. Par rapport aux approches classiques, nous avons automatisé au niveau composant les processus chronophages d'attribution d'un neurone et de mise à jour de l'ensemble des poids des connections pendant la phase d'apprentissage. Les zones d'influence (AIF¹¹) de l'ensemble des neurones (voir Fig. 1) s'ajustent automatiquement et de manière transparente pour l'utilisateur, pour permettre la prise en compte d'un nouvel échantillon.

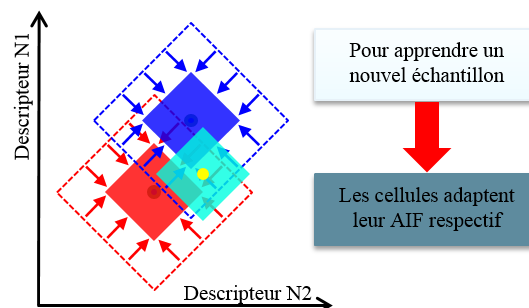


Fig. 1. Espace de décision 2D d'un réseau RBF : les zones d'influences des neurones (rouge, bleu) s'ajustent automatiquement pour permettre au réseau d'engager un nouveau neurone (point jaune) avec son espace de décision associé (bleu clair).

Rapides et à configurations dynamiques, le réseau RBF est capable d'apprendre à tout moment de nouveaux échantillons, même en phase de reconnaissance. Cette fonc-

⁸ k Nearest Neighbors

⁹ Probabilistic Neural Network

¹⁰ Radial Basis Function

¹¹ Active Influence Field

tionnalité peut être utile pour la poursuite d'une cible pour laquelle il faut pouvoir ré-actualiser le modèle de la cible par apprentissage incrémental. Parallélisé sur SoC, l'avantage de cette approche est un temps de reconnaissance constant indépendant de la taille de la base de données avec une efficacité très largement supérieure à une approche linéaire classique sur un ordinateur (voir Fig. 2).

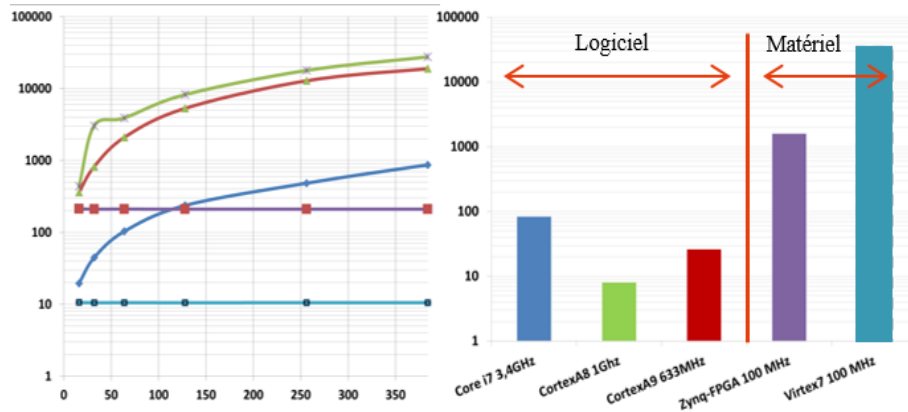


Fig. 2. Performances du RBF évaluées sur plusieurs plateformes avec la base de caractères MNIST (2009). A gauche le temps de reconnaissance en μs en fonction du nombre de cellule de décision. A droite l'efficacité avec 384 cellules de décision ($\text{Reco.s}^{-1}.\text{W}^{-1}$)

L'efficacité de cette approche est dépendante de la pertinence et de la définition « manuelle » des éléments caractéristiques retenus pour les évènements à reconnaître : un facteur limitant de cette méthode.

Approche DNN-SSD. Le second moteur d'IA retenu est le réseau SSD. Il présente l'avantage de définir lui-même les éléments caractéristiques des objets à reconnaître. Pour entrainer un tel réseau, il faut une très grande base de données labellisées et fiables, et un ordinateur performant. L'apprentissage hors-ligne aboutit à un fichier de configuration décrivant les couches, leurs fonctions et les coefficients des filtres de convolutions. Ce fichier a été porté sur la clé USB NCS¹² MOVIDIUS-INTEL Myriad 2 [4] qui fait office de coprocesseur du SoC. Programmé avec un réseau SSD MobileNet pré entraîné avec une base VOC 0712¹³ [13], il permet la reconnaissance de plusieurs évènements dans une image RGB de résolution 300x300 pixels. Le temps de reconnaissance est de $\sim 100\text{ms}$ par image pour une consommation de 1W et passe à $\sim 30\text{ms}$ avec la clé Myriad X [5]. Ce temps de reconnaissance est compatible avec nos applications de DRI embarquées. Mais, l'évaluation sur le terrain a montré un échauffement trop important de la clé auquel s'ajoute un problème de robustesse. Nous avons remédié à

¹² Neural Compute Stick

¹³ 20 classes (classification d'images)

ce problème d'intégration, de consommation et d'échauffement en intégrant l'IP¹⁴ Xilinx-DPU¹⁵ [6] directement au niveau du SoC (Xilinx-XC7Z020) équipant nos prototypes [7]. Les performances avec le réseau MobileNet V2 SSD et des images RGB de 480x360 pixels sont moindres (base BDD100K¹⁶[14]). Le temps de reconnaissance est de ~230ms par image. Ce temps de calcul est compatible d'une application DRI embarquée sur un UGS. Une nouvelle implémentation et évaluation de laboratoire sur une plateforme ZynqMP Ultrascale MPSoC¹⁷ (Xilinx- XCZ07EV) [8,9] avec le réseau MobileNet V2 SSD a atteint un taux de reconnaissance de ~33ms/image mais au prix d'une plus forte consommation électrique (~10W avec la carte de développement utilisée ZCU104 Xilinx). Une approche MPSoC optimisée (~ 8W) permettra d'augmenter la capacité de traitement DRI de nos concepts d'UGS.

3 Validation et expérimentation

3.1 Unité de calcul

Le défi est d'atteindre une souveraineté aux niveaux électronique, matériel et logiciel. Pour les capteurs ISL nous avons développé une carte mère (55x55mm²) polyvalente (voir Fig. 3), programmable en langage C sous un OS-Linux. Sa structure permet de porter sur son SoC les algorithmes d'analyse et de traitement décrits précédemment. Les prochains développements se feront sur MPSoC.

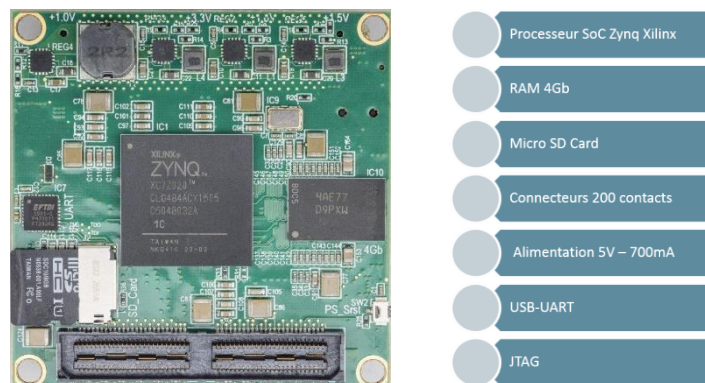


Fig. 3. Carte mère ISL équipée d'un SoC Xilinx-XC7Z020 (55x55mm²)

Pour des applications de surveillances multidirectionnelles et infrarouge, la même carte mère peut gérer jusqu'à 4 flux vidéo RGB WXVGA en parallèle [12].

¹⁴ Intellectual Property

¹⁵ Deep Processing Unit

¹⁶ 11 classes du domaine de la conduite : piétons, véhicules, feu de signalisation, ...

¹⁷ Multi Processor System on Chip

3.2 Les capteurs Sentinelles

Le défi de Sentinelle est d'apporter à l'homme de terrain une aide à la prise locale de décision. Sentinelle est adaptable et reconfigurable en fonction de la mission. L'effort de conception a été porté sur les aspects mobilité, discrétion de communication et d'intégration dans le milieu naturel. Parfaitement autonome, elle ne requière aucun paramétrage sur site. Elle archive localement 24h/7 les mouvements ou changements significatifs détectés tout en précisant leur nature : véhicule ou piéton. Cette fonction de reconnaissance est réalisée avec l'IP Xilinx DPU portée sur le SoC de la carte mère (voir Fig. 3). Cette approche hautement intégrée a servi à établir une preuve de concept. Les performances obtenues à partir de bases civiles : BDD100K et VOC 0712¹⁸ sont respectivement de ~230 ms avec un réseau MobilenetV2 SSD et de ~600 ms avec un réseau VGG16 SSD.

Deux générations de Sentinelles ont été développées pour répondre aux contraintes de terrain : Sentinelles visible (EOV) (voir Fig. 4) et infrarouge (EOIR) (voir Fig. 5). En mode visible, le capteur de 1/3" utilisé a une résolution de 752x480 pixels et une fréquence de 10fps en analyse. Une monture S (M12x0.5) permet d'utiliser des optiques de faible encombrement. Pour une optique de 50 mm, le champ de vue est de 5°, le volume occupé est de 43x25x25 mm³ pour un poids de 36 g (©LENSATION). Le poids de Sentinelle EOV est de ~400g et son volume sans les optiques est de 70x70x45mm³ et sa consommation électrique est de 0,80A sous 5V.



Fig. 4. Changement d'optique sur Sentinelle EOV. La distance de détection des événements, de 10 à 1000m, est fonction de la focale de l'objectif interchangeable.

Dans sa version infra rouge (EOIR) (voir Fig.5), les caractéristiques de Sentinelle restent sensiblement les mêmes. Sentinelle EOIR s'appuie sur un capteur 1" Thermo-technix de 640x480 pixels de taille 17μm et une fréquence de 10fps en analyse. Il opère en bande III non refroidi 8μm-14μm avec une résolution de 14bits. Pour une optique de 60mm, le champ de vue est de 10°, le volume occupé de 62x62x50mm³ pour un poids de 218g (© UMICORE). Le poids de Sentinelle EOIR est de ~600g et sa consommation électrique est de 1,14A sous 5V.

¹⁸ 20 classes VOC 0712 ou 2 classes : véhicules et personnes VOC 0712 complétées avec 2 classes ISL.



Fig. 5. Sentinelle EOIR avec une optique de 60mm.

3.3 Sentinelle sur le terrain

Capacité de Détection. Sentinelle EOIR a été testée avec succès lors d'une campagne à Mourmelon pour valider les distances de détection et de reconnaissance (voir Fig. 6). Avec une optique de 50 mm, un piéton est détecté à 400m et un véhicule à 1300m.



Fig. 6. Détection (DRI) d'un soldat (cercle orange) en mouvement en lisière de bois par Sentinelle EOIR distante de 400m avec une focale de 50mm.

Pour les applications de terrain, l'IHM sur tablette tactile géo localise et visualise la position angulaire de Sentinelle sur un fond de carte référencé UTM¹⁹(voir Fig. 7).

¹⁹ Universal Transverse Mercator



Fig. 7. L'IHM permet de visualiser sur une carte la position au format UTM de Sentinelle (point rouge) le secteur angulaire observé (cône bleu) ainsi que les angles de la direction des évènements (trait jaune).

Capacité de communication. En fonction de la situation, la communication avec la Sentinelle se fait via une communication Gigabit Ethernet (IHM²⁰) (voir Fig.8) ou par clé USB (milieu hostile).



Fig. 8. L'IHM permet de consulter la base des évènements archivée dans Sentinelle via une connexion IP filaire ou IP UHF.

Capacité de reconnaissance. Sentinelle gère localement une base des évènements détectés au format SQL, consultable depuis l'IHM (voir Fig.9 et Fig. 10). A chaque évènement sont associés un temps d'occurrence, une position angulaire, une imagerie, un index de classe avec son taux de vraisemblance (DRI) et une séquence vidéo de $\pm 2s$.

²⁰ Interface Homme Machine

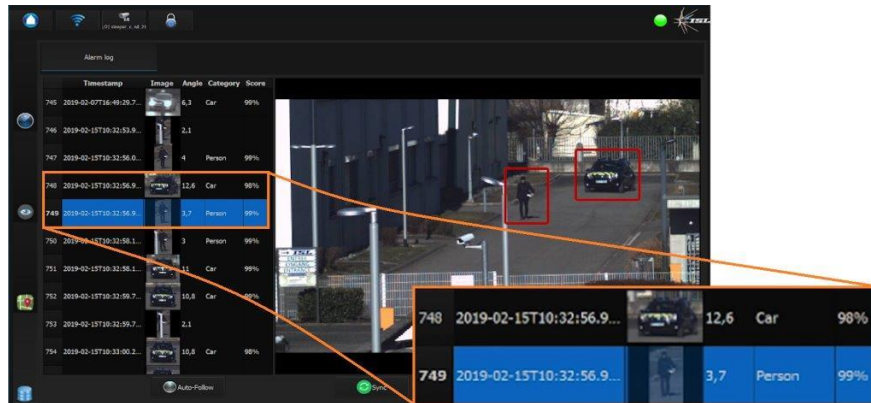


Fig. 9. Consultation de la base des évènements de Sentinelle EOY depuis l'IHM.

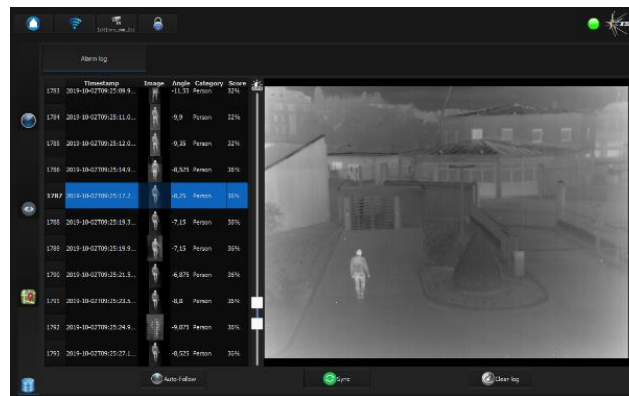


Fig. 10. Consultation et création de la base des évènements de Sentinelle EOIR depuis l'IHM.

4 Conclusion et perspectives

Dans le domaine de la Défense et de la Sécurité, le traitement local de l'information est un point clé des capteurs déposés pour établir une vision tactique opérationnelle temps réel en milieu hostile. L'approche IA embarquée permet des actions de DRI, de contribuer à limiter les besoins en énergie et en bande passante sur le terrain. Les progrès académiques réalisés dans le domaine de l'IA, en parallèle avec l'évolution technologique, ont permis de rendre réalistes des applications embarquées ainsi que le démontre le prototype Sentinelle, validé sur le terrain.

Dans cette étude, plusieurs approches complémentaires d'IA ont été adaptées et portées au niveau composant. Les classifieurs parallélisés dans TEAM sont performants mais supposent une définition préalable des éléments caractéristiques définissant les évènements à reconnaître. Sur cet aspect, l'approche SSD est plus efficace mais suppose une grande base de données significatives et labellisées. Mais, contrairement aux classifieurs (TEAM), le SSD ne permet pas l'apprentissage d'un nouvel évènement en

phase de reconnaissance. Une étude en cours associe ces deux approches en vue de permettre au SSD un apprentissage à la volée. Le réseau SSD configuré sera utilisé comme extracteur de caractéristiques formatées sous la forme d'un vecteur exploité par TEAM avec un temps de réponse constant.

Le point clé de ces concepts reste l'accès à des bases de données fiables et pertinentes : drones, véhicules, piétons ... Des travaux ISL en cours ambitionnent de générer des bases spécifiques. La première étude sera centrée sur des véhicules militaires, dans le spectre visible. Les images seront générées à partir de maquettes à l'échelle 1/48^{ème}, dans des conditions d'éclairage contrôlées, avec différentes orientations et sous différents angles de vue. La seconde étude sera focalisée sur la création d'une base de données véhicules et piétons en bande III acquises en milieu naturel par Sentinelle EOIR. Dans les deux cas, il s'agit de créer les quelques milliers d'images labellisées, nécessaires à l'entraînement d'un réseau DNN embarqué, puis d'étudier la pertinence de telles bases, spécifiquement construites pour répondre à une mission.

Références

1. Manzanera, A., Richefeu, J.: A new motion detection algorithm based on Σ - Δ background estimation, *Pattern Recognition Letters* 28, 320-328 (2007).
2. N. Hueber, C. Hennequin, al.: Real-time movement detection and analysis for video surveillance applications, *SPIE Proceedings*, vol. 9079, June (2014).
3. P. Raymond: Modélisation, simulation d'un réseau neuronal RBF. Amélioration du modèle RBF vers une approche câblée de l'intelligence artificielle appliquée à la perception de l'environnement, Rapport ISL, RV-212_2008.
4. Intel, Intel neural compute stick, <https://software.intel.com/en-us/movidius-ncs>, last accessed 2019/09/25.
5. Intel, Intel neural compute stick 2, <https://software.intel.com/en-us/neural-compute-stick>, last accessed 2019/09/25.
6. Xilinx, Dpu for convolutional neural network, <https://www.xilinx.com/products/intellectual-property/dpu.html>, last accessed 2019/09/25.
7. Xilinx, Zynq-7000 SoC, <https://www.xilinx.com/products/silicon-devices/soc/zynq-7000.html>, last accessed 2019/09/25.
8. Xilinx, Zynq-7000 SoC, <https://www.xilinx.com/products/silicon-devices/soc/zynq-ultrascale-mpsoc.html>, last accessed 2019/09/25.
9. Xilinx, Xilinx zynq ultrascale + MPSoC zcu104 evaluation kit, <https://www.xilinx.com/products/boards-and-kits/zcu104.html>, last accessed 2019/09/25.
10. Manzanera, A. and Richefeu, J. C., A robust and computationally efficient motion detection algorithm based on sigma-delta background estimation, *Indian Conference on Computer Vision, Graphics and Image Processing in ICVGIP'04* (2004).
11. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S. E., Fu, C.-Y., and Berg, A. C., SSD: Single shot multibox detector, in *ECCV* (2016).
12. Hueber, N., Raymond, P., Hennequin, C., Pichler, A., Perrot, M., Voisin, P., and Moeglin, J.-P., "Bio-inspired approach for intelligent unattended ground sensors," *Proceedings SPIE 9494, Next-Generation Robotics II; and Machine Intelligence and Bio-inspired Computation: Theory and Applications IX* (2015).
13. PASCAL VOC 0712, <http://host.robots.ox.ac.uk/pascal/VOC>, last accessed 2019/09/25.
14. BDD100K, <https://bdd-data.berkeley.edu/index.html>, last accessed 2019/09/25.