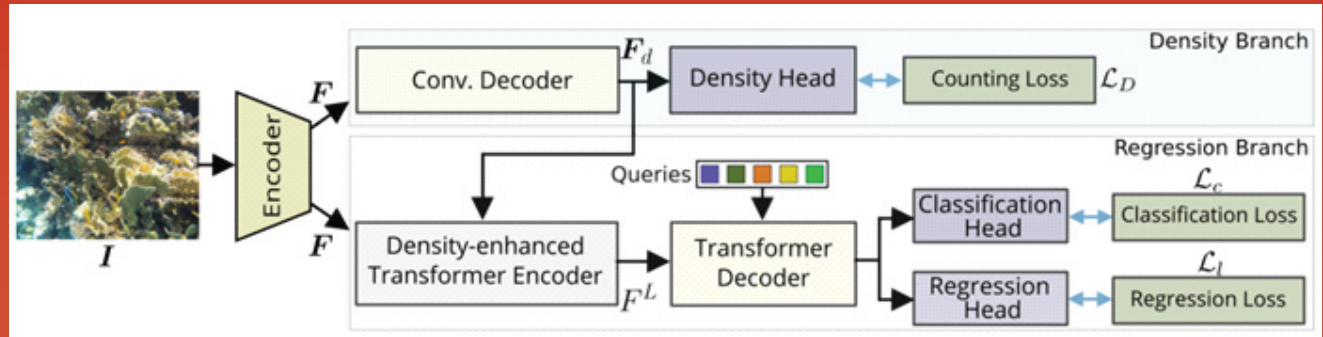


La Revue 3E.I



Ressource publiée sur Culture Sciences de l'Ingénieur : <https://sti.eduscol.education.fr/si-ens-paris-saclay>

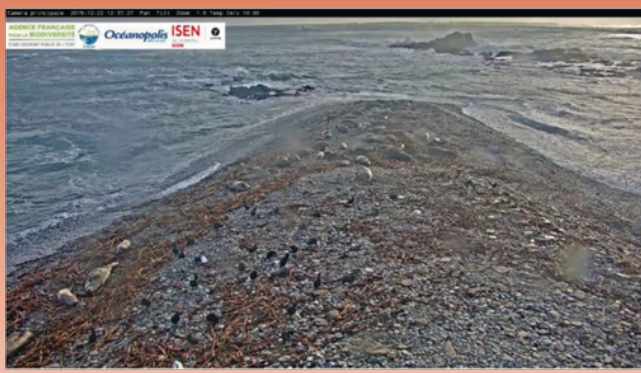


Prompt Input
"superhero style, ironman, iron, light, dangerous"

Image Input	Canny	Image Output



Vision et traitement d'images



Publication trimestrielle du Cercle Thématique 13.01 de la SEE
ENSEIGNER L'ÉLECTROTECHNIQUE ET L'ÉLECTRONIQUE INDUSTRIELLE



Société de l'Électricité, de l'Électronique
et des Technologies de l'Information
et de la Communication

N°115
2^{ème} trimestre 2025

Abonnez-vous à la

La REE est une publication trimestrielle de la SEE



Choisissez votre formule d'abonnement :



Version papier

1 an - 4 numéros**

(Parution : mars, juin, octobre, décembre). Distribution postale

Livraison France	Livraison UE	Livraison Hors UE
☐ 135 € TTC	☐ 150 € TTC (146,91 € HT*)	☐ 165 € TTC (161,61 € HT*)



Version numérique

Accès aux publications numériques
ouvert pendant un an à compter de la date de paiement

France - UE - Hors UE
☐ 90 € TTC (88,15 € HT*)



Version duo

Version imprimée + version numérique

Livraison France	Livraison UE	Livraison Hors UE
☐ 165 € TTC	☐ 180 € TTC (173,30 € HT*)	☐ 195 € TTC (190,99 € HT*)

* Prix HT valide si le pays de facturation est hors UE, ou si la TVA Intracommunautaire est fournie pour un pays de l'UE.

** Abonnement papier servi à partir de la date de paiement

Votre adhésion à la SEE* pour 2025

☐ Standard	☐ Retraité ☐ Enseignant ☐ Jeune actif (< 35 ans)	☐ Etudiant ☐ En recherche d'emploi
130 €	70 €	15 €

* Adhésion d'un an à compter de la date de paiement.

+ Votre abonnement REE** (Tarif réservé aux adhérents, version papier)

☐ Livraison France : 68 € TTC	☐ Livraison UE : 78 € TTC (76,40 € HT*)	☐ Livraison Hors UE : 83 € TTC (81,70 € HT*)
--	--	---

TVA de la revue REE : 2,1 %. Adhésion collective possible via des conventions de partenariat - Contactez-nous à : sg@see.asso.fr

* Prix HT valide si le pays de facturation est hors UE, ou si la TVA Intracommunautaire est fournie pour un pays de l'UE.

** Abonnement papier servi à partir de la date de paiement.

Adresse de livraison

☐ Mr ☐ Mme Prénom* : _____

Nom* : _____

Adresse* : _____

Code postal* : _____ Pays* : _____

Ville* : _____

Tél.* : _____

e-mail* : _____

*Obligatoire

Adresse de facturation (Si différente)

☐ Je joins le bon de commande administratif N° _____
et je désire recevoir une facture au nom de mon employeur pour paiement à réception

Raison sociale de l'employeur : _____

Service : _____ Activité (facultatif) : _____

Adresse : _____

Code postal : _____ Ville : _____

Pays : _____ N° TVA : _____

N° TVA intracommunautaire : obligatoire pour règlement HT en UE hors de France

Votre règlement

Je règle la somme de _____ €

par ☐ Chèque à l'ordre de la SEE

☐ Virement après réception de la facture

☐ Carte bancaire, lien de paiement sécurisé (visa, etc)

e-mail* : _____

Date* _____ Signature* et cachet si il y a lieu : _____

*Obligatoire

BULLETIN À COMPLÉTER ET RENVOYER À : SEE - 17 rue de l'Amiral Hamelin - 75116 Paris - France
Tél. +33(0)1 56 90 37 17 - abo@see.asso.fr

ABONNEMENT PLUS RAPIDE EN LIGNE : www.see.asso.fr

☐ Je consens à recevoir les autres diffusions de la SEE & de ses activités (congrès, soirées débats, revues, etc.) qui sont extérieures aux diffusions liées à mon abonnement.

Conformément aux dispositions légales et réglementaires en matière de données personnelles, les informations recueillies sur ce formulaire sont enregistrées dans un fichier informatisé par la SEE (Société de l'électricité, de l'électronique et des technologies de l'information et de la communication) pour la mise en place et le suivi de l'abonnement souscrit ainsi que pour l'envoi de courriers, e-mails de réabonnements. Elles sont conservées et sont destinées à être utilisées par la SEE et les prestataires techniques de la SEE afin de permettre la bonne réception du magazine et d'assurer le service client. Vous pouvez exercer votre droit d'accès aux données vous concernant par courrier : SEE - Service abonnements 17 rue de l'Amiral Hamelin 75116 Paris ou par le formulaire de contact du site web : www.see.asso.fr. Offre d'abonnement, valable du 01/12/2023 au 25/11/2024 inclus, dans la limite des quantités disponibles.





La Revue 3E.I
Publication trimestrielle
de la SEE

**SOCIÉTÉ de l'ELECTRICITE, de l'ELECTRONIQUE
et des TECHNOLOGIES de l'INFORMATION
et de la COMMUNICATION.**

17, rue de l'Amiral Hamelin, 75116 PARIS
Tél : 01 56 90 37 17
www.see.asso.fr

SEE, association reconnue d'utilité publique par le décret du 7 décembre 1886
Siret 785 393 232 00042, APE 9412 Z, n° d'identification FR 44 785 393 232

Hébergé par :

**Culture Sciences
de l'Ingénieur**

4 avenue des Sciences, 91190 Gif sur Yvette
tel : 01 81 87 55 22
<https://eduscol.education.fr/sti/si-ens-paris-saclay>

La Revue 3E.I

**3EI : Enseigner l'Électrotechnique et l'Électronique
Industrielle**

La Revue 3EI, Édition SEE,
17 rue de l'Amiral Hamelin
75116 PARIS

Directeur de la publication
François GERIN
Président de la SEE

Rédacteur en Chef
Franck LE GALL

Adresser les propositions d'article à :
revue3ei@gmail.com

Communication :
Mme. Mélisande DE LASSENCE
Communication1@see.asso.fr
01 56 90 37 17

Dépôt Légal : janvier 2024
Commission Paritaire 1222 G 78028
ISSN 1252-770X

Comité de publication

Morgan ALMANZA (ENS Paris-Saclay)

Hamid BEN AHMED (ENS Rennes)

Afef BEN ABDELGHANI BENNANI (ENIT, Tunisie)

Arnaud BRUGIER (IUT GIM Saint-Denis)

François COSTA (SATIE UMR 8029, UPEC)

Hervé DISCOURS (IUT GEII Cachan)

Jean-Michel GAY (Retraité STI2D-BTS ET Versailles)

Hélène HORSIN-MOLINARO (Culture Science de
l'Ingénieur)

Jean-Philippe ILARY (IUT GEII Ville-d'Avray)

Anthony JUTON (ENS Paris-Saclay)

Franck LE GALL (ISEN Brest)

Ingrid MININGER (BTS CIEL ER Cachan)

Emmanuel MONNOT (STI2D Versailles)

Abir REZGUI (ESIEE Paris)

Magali SAUVERGEAT (BTS CIEL IR Arpajon)

Jean-François SERGENT (Retraité Univ Lille)

Sommaire du n° 115

p. 3 *Éditorial*

Thème : Vision et traitement d'images

- p. 4 *Thibault Napoléon - Ayoub Karine - Hamza Masmoudi, « Méthodes d'intelligence artificielle pour la surveillance d'une colonie de phoques gris »*
- p. 15 *Simon Playe, « Computer Vision : Introduction au calcul du volume des objets »*
- p. 21 *Simon Playe, « IA Générative : Apprenez à contrôler la génération d'images avec Stable Diffusion »*

Hors Thème :

- p. 32 *Rachida Saroui et Anthony Juton, « Localisation en intérieur avec Ultra Wide Band »*
- p. 44 *Joss Brunet, Pascal Champaney, Hervé Discours, « Conception de systèmes radiofréquences. Entretien avec la PME française « Always Wireless »*

Editorial

La vision par ordinateur et le traitement d'images s'imposent aujourd'hui comme des domaines incontournables dans de nombreuses applications du génie électrique. Que ce soit dans le contexte des véhicules autonomes, de la robotique ou, plus largement, de l'exploitation des données issues de capteurs optiques, ces technologies ouvrent des perspectives nouvelles souvent encore peu familières à nombre d'entre nous.

Dans ce numéro du second trimestre 2025, nous vous proposons d'explorer cet univers à la croisée des capteurs et du traitement informatique des données. Un domaine passionnant, riche en innovations, dont nous pourrons, espérons-le, tirer le meilleur pour nos projets à venir !

« Dossier : Vision et traitement des images »

L'article introductif nous transporte au large de Brest, sur l'îlot de Morgol, au cœur de la mer d'Iroise, où une colonie de phoques gris a élu domicile. Thibault Napoléon et ses co-auteurs y présentent une méthode de surveillance basée sur l'intelligence artificielle. Cette approche, non intrusive pour les espèces observées, permet une estimation bien plus précise de la population de phoques, contribuant ainsi à une meilleure protection de leur habitat naturel. Une belle illustration du potentiel de l'IA appliquée à l'environnement.

Avez-vous déjà envisagé de calculer le volume d'objets à partir de vidéos en exploitant l'intelligence artificielle ? C'est précisément la problématique abordée dans l'article de Simon Playe, qui détaille les différentes étapes du calcul et les leviers d'optimisation possibles – un enjeu majeur pour des applications telles que la gestion de stocks ou l'aménagement d'espaces.

Dans un second article, Simon Playe revient sur un autre sujet d'actualité : l'amélioration de la génération d'images à l'aide de l'outil Stable Diffusion. Il nous guide dans l'utilisation de cet outil, conçu pour produire des visuels via une API, et nous montre comment en tirer le meilleur parti dans un cadre professionnel.

« Hors Thème »

Comment se localiser efficacement à l'intérieur d'un bâtiment alors que la géolocalisation par satellite, si performante en extérieur, devient peu fiable une fois en intérieur ? Rachida Saroui et Anthony Juton nous font découvrir la technologie Ultra Wide Band (UWB), qui offre une nouvelle approche prometteuse : la localisation d'objets en environnement clos.

Ce numéro se conclut par un entretien mené par Hervé Discours avec J. Brunet et P. Champaney, de la société Always Wireless. Spécialisée dans les sous-ensembles pour les communications radio sans fil, l'entreprise nous ouvre les coulisses de ses projets – de la conception à la qualification produit. Parmi leurs réalisations : les systèmes de communication pour les arbitres UEFA (VOGO), les modules de géolocalisation GNSS des vélos électriques DECATHLON, les solutions d'éclairage urbain intelligentes de LACROIX et BH Technologies, ou encore les antivols connectés utilisés par DELIVEROO et LA POSTE.

Dans les prochains numéros de la revue 3EI, nous aborderons les thématiques de la compatibilité électromagnétique (CEM), de la robotique et de RISC-V. N'hésitez pas à nous faire parvenir vos propositions d'articles, nous les attendons avec enthousiasme !

Méthodes d'intelligence artificielle pour la surveillance d'une colonie de phoques gris

Thibault NAPOLÉON¹ - Ayoub KARINE² - Hamza MASMOUDI¹

Édité le
17/03/2025

école
normale
supérieure
paris-saclay

¹ ISEN Ouest, LabISEN, Vision-AD, 20 rue Cuirassé Bretagne, 29200 Brest

² Université Paris Cité, LIPADE, F-75006 Paris, France

Cette ressource fait partie du N°115 de La Revue 3EI du deuxième trimestre 2025.

Dans le cadre de la conservation de la biodiversité marine, les technologies de surveillance et d'analyse jouent un rôle crucial. En effet, la gestion efficace des habitats naturels nécessite une compréhension approfondie des écosystèmes et de leur dynamique. Cela est particulièrement vrai dans des zones sensibles comme l'îlot de Morgol, situé dans la réserve naturelle nationale d'Iroise, qui abrite une population importante de phoques gris. Ces dernières années, les méthodes traditionnelles de suivi ont été remises en question, notamment à cause de leur impact potentiellement perturbateur sur la faune.

Cet article se concentre sur le développement d'une méthode de surveillance qui minimise les interférences avec les espèces surveillées tout en fournissant des données précises et exploitables. En particulier, nous proposons ici une approche basée sur la vision par ordinateur et les méthodes modernes d'intelligence artificielle pour : 1. Mesurer la fréquence des dérangements liés au passage de bateau ou au débarquement d'humains. 2. Estimer la densité de phoque gris sur l'îlot. La méthodologie employée s'appuie d'une part sur l'utilisation de YOLO v8 pour la détection des perturbateurs et sur l'approche IOCFomer ré-entraînée pour la phase de comptage d'autre part. Comparés à la méthode précédemment utilisée, les résultats obtenus sont plus riches et montre une nette amélioration de l'estimation de la colonie pour des coûts de calcul acceptable.

1 - Introduction

La vision par ordinateur présente aujourd'hui un fort intérêt pour l'observation des espaces naturels au travers de tâches allant de la détection d'objets à la segmentation sémantique en passant par le comptage d'individus [1]. Cet attrait pour ce champ de recherche est poussé par le coût faible des capteurs de vision ainsi que par la croissance des approches d'intelligence artificielle traitant les données issues des capteurs optiques. En particulier, les réseaux de neurones profonds ont permis un développement rapide de nouvelles méthodes grâce à l'essor de deux éléments complémentaires. D'une part, les grandes bases de données d'images, telles qu'ImageNet [2], nécessaires à l'apprentissage des modèles d'intelligence artificielle. D'autre part, les puissances de calcul disponible au travers des processeurs graphiques (c.-à-d. GPU) qui se sont avérés particulièrement efficaces pour exécuter les approches neuronales. Ainsi, il est désormais possible d'employer ce type d'approche d'intelligence artificielle pour des projets concrets alliant robustesse des traitements et rapidité des calculs dans des applications diverses telles que la reconnaissance de visages, le déploiement de voitures autonomes ou encore la préservation de l'environnement.

1.1 - Présentation du projet

L'approche présentée ici s'inclut dans le projet de remplacement de l'observatoire qui permettait avant 2024 d'effectuer le comptage des individus d'une colonie de phoques gris après sa destruction par les intempéries. Ce dispositif, installé sur l'îlot de Morgol dans l'archipel de Molène au large des côtes Finistériennes, a pour rôle de faciliter le suivi de la faune sauvage qui s'y trouve avec comme objectif de minimiser les dérangements. L'observatoire mis en place est équipé d'une caméra, de panneaux solaires offrant l'autonomie énergétique, ainsi qu'un lien radio permettant la retransmission des images à terre, voir figure 1. La caméra disponible sur l'observatoire possède deux capteurs, dans les domaines visible et infrarouge, mais seule la première modalité est utilisée ici. Les images captées par l'observatoire sont traitées à terre à l'aide d'une carte de traitement *Jetson Nano Orin* embarquant les algorithmes mis en œuvre dans le projet. En particulier, deux réseaux de neurones ont été déployés pour le suivi de la colonie dans le temps :

1. Un réseau dédié à la détection de bateaux naviguant aux abords de l'îlot ou de personnes y débarquant.
2. Un réseau dédié au comptage s'appuyant sur l'estimation de la densité de phoques gris sur l'îlot.

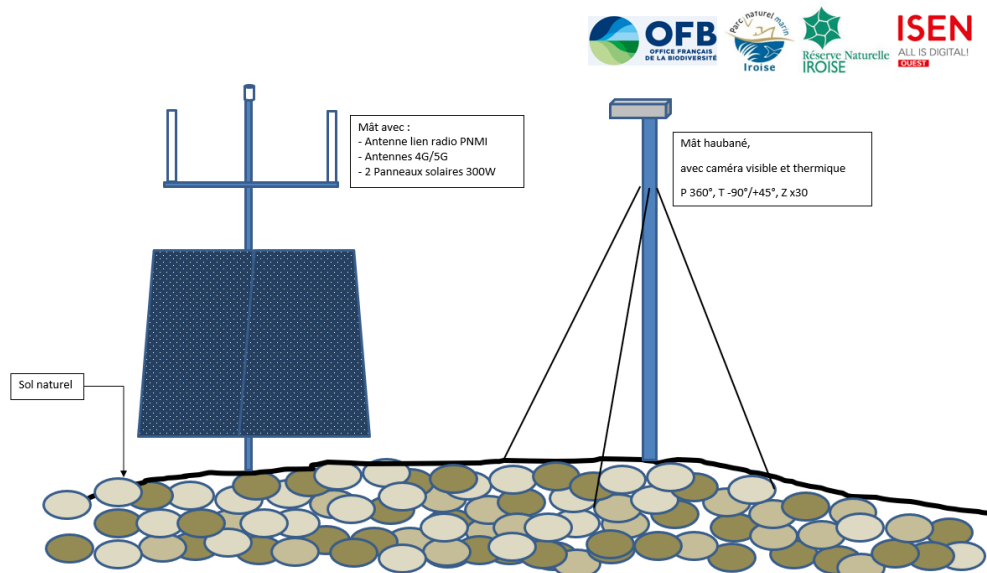


Figure 1 : Schéma de l'observatoire qui sera installé sur l'îlot de Morgol. On y trouve une caméra bispectrale, deux panneaux solaires ainsi qu'une antenne pour le lien radio.

1.2 - Travaux existants

Une première version de l'observatoire sur l'îlot de Morgol avait été réalisée en 2019 pour répondre à la problématique du suivi de la colonie de phoque. La méthode proposée par Karine et al [3] s'appuie sur une approche de classification des vidéos de phoques en utilisant l'apprentissage par transfert sur un réseau de neurones convolutifs (CNN). Grâce au transfert de connaissance apporté par cette technique, qui consiste à réutiliser les hyperparamètres d'un modèle pré-entraîné sur un large ensemble de données, il est possible de limiter la quantité d'annotations nécessaire à la maturation (c.-à-d. apprentissage) du réseau de neurone employé. La mise en œuvre de méthode proposée repose sur deux phases distinctes présentes dans la majorité des approches d'apprentissage supervisé : 1. Une phase « hors ligne » permettant l'entraînement du modèle à partir de vidéos annotées. 2. Une phase « en ligne » utilisant le modèle entraîné pour classer automatiquement les images issues des nouvelles vidéos pour en extraire les phoques et ainsi estimer l'évolution de leur nombre au cours du temps.

Phase « hors ligne »

Cette phase est appelée hors ligne, car elle est réalisée une seule fois pour apprendre au réseau de neurones profond à reconnaître les phoques sur une image. Dans celle-ci, des imagerie de taille 70×100 pixels sont utilisés pour entraîner le modèle de classification. En particulier, une annotation de 353 imagerie de la classe « phoque » ont été utilisées conjointement à 357 imagerie de la classe « non-phoque » comme vérité terrain pour la phase d'apprentissage, voir figure 2. En effet, étant donné le cadre applicatif de la méthode, aucun jeu de données n'était disponible pour entraîner le modèle. Le réseau utilisé est un CNN basé (a) sur l'architecture VGG-16 [4] pré-entraîné sur la base de données *ImageNet*. Ce réseau contient 13 couches convolutives suivies de couches de *pooling* et de couches complètement connectées (c.-à-d. des couches denses). Une couche de 256 neurones ayant une fonction d'activation de type *ReLU* et avec un dropout fixé à 0.5 a été ajoutée pour éviter le surapprentissage. Enfin, la dernière couche de sortie est une classification binaire avec fonction d'activation de type *Softmax* pour permettre une prédiction probabiliste. L'apprentissage par transfert du réseau a été réalisé avec un optimiseur Adam, une taille de lot de 8 sur 50 époques. Pour cette phase, la base d'annotation a été séparée en trois parties, entraînement (60 %), validation (20 %) et test (20 %). La base de validation permet de limiter le risque du surapprentissage tandis que la base de test permet d'évaluer la pertinence du modèle sur des données annotées, mais inconnues du système d'apprentissage. Les résultats obtenus montrent une précision globale de 87.24 % sur la base de test. Les erreurs de classification sont en partie dues à la ressemblance entre les phoques et les galets présents dans les images.

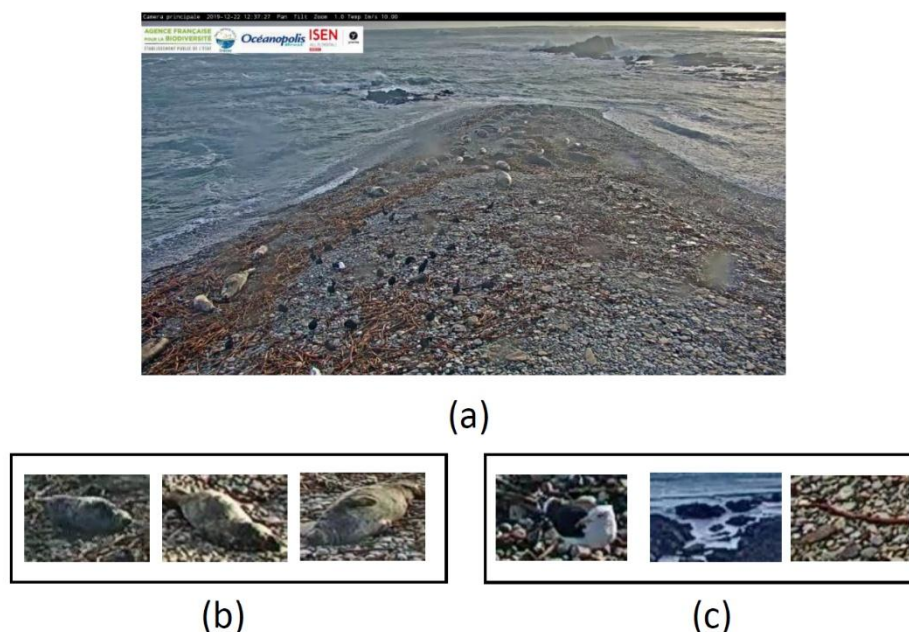


Figure 2 : (a) Exemple d'image issue de la caméra installée sur l'îlot de Morgol. (b) Exemple de vignettes de la classe « phoque ». (c) Exemple de vignettes de la classe « non-phoque ».

Phase « en ligne »

Une fois l'entraînement du modèle réalisé, le système peut alors être utilisé pour estimer le nombre de phoques sur l'îlot. Pour cette phase, une stratégie de fenêtre glissante, avec un décalage de 100 pixels, est utilisée pour classer chaque vignette de l'image entre les classes « phoque » et « non-phoque ». Ici, la taille des vignettes est identique à celle de la phase d'apprentissage, à savoir 70×100 pixels. Dans cette approche, pour toutes les images captées sur l'îlot et retransmises sur le continent, le système d'intelligence artificielle va parcourir l'image pour y détecter les vignettes présentant des phoques. Conjointement à cette détection, des informations temporelles sont extraites de la vidéo pour horodater les présences/absences de phoque. En particulier, les données temporelles sont composées de la date et de l'heure d'acquisition de la vidéo ainsi que de l'instant

précis de la détection en heure, minute et seconde. Ainsi, à la suite de la phase de classification, une agrégation des résultats permet d'estimer le nombre phoque sur chaque image et de suivre ainsi l'évolution de la colonie au cours du temps.

Ce dispositif est désormais inopérant après 5 années en conditions difficiles, incluant la tempête Ciarán qui a eu lieu à la fin de l'année 2023. De ce fait, nous proposons une nouvelle implémentation de l'approche de suivi de la colonie de phoque au travers de nouvelles méthodes qui suivent l'évolution de l'état de l'art dans le domaine.

2 - Approche proposée

Dans le nouveau système, deux approches disjointes, mais complémentaires, ont été mise en place. La première est basée sur *YOLO v8* pour détecter les bateaux et les humains qui sont des perturbateurs pour l'écosystème marin tandis que la seconde s'appuie sur l'approche *IOCFomer* ré-entraînée pour estimer la densité de phoque gris.

2.1 - Mesure des perturbations : *YOLO v8*

Afin d'étudier précisément le comportement de la colonie de phoque gris, l'évaluation de l'impact des événements extérieurs est nécessaire. Pour cela, un outil d'intelligence artificielle permettant la détection d'événements tels que le passage d'un bateau, l'arrivée d'un kayak ou le débarquement de personnes sur l'île est mise en place. Ainsi, associé aux informations temporelles, il est possible d'évaluer le degré de perturbation de la colonie et le temps de retour à l'état normal du reposoir. L'outil développé est basé sur un réseau de neurones convolutifs « léger » permettant le traitement des vidéos en temps réel. En effet, l'une des contraintes du projet est la quantité de données à analyser, à savoir, un flux vidéo Full HD permanent. Le réseau de neurones profonds utilisé est la version 8 de la famille *YOLO* [5] (You Only Look Once) développé par Ultralytics qui améliore la précision et la vitesse par rapport aux versions précédentes grâce à des optimisations architecturales et de nouvelles techniques d'apprentissage. Ce modèle modulaire et polyvalent permet de réaliser plusieurs tâches en vision par ordinateur, dont la détection d'objets.

L'architecture de *YOLO v8* est composée de trois parties principales : *backbone*, *neck* et *head*. Le réseau de base (c.-à-d. Le *backbone*) permet d'extraire les caractéristiques visuelles de l'image telles que les contours, les textures ainsi que les motifs colorimétriques. Pour cela, il s'appuie sur des optimisations de l'architecture *CSPDarknet* permettant une meilleure efficacité. La partie *neck* de l'architecture apporte une capacité multi-échelle à la détection en fusionnant les informations issues de différentes échelles de l'image. Il est basé sur une optimisation de la méthode *PANet*. Enfin, la partie *head* du réseau génère les prédictions sous forme de boîtes englobantes associées, pour chacune, aux classes prédites avec leurs scores de confiances.

Ce réseau pré-entraîné sur la base de données *COCO* [6] permet de détecter et de reconnaître un ensemble de concept, dont les bateaux et les humains. Appliqué sur chaque image de la vidéo, il permet de détecter les événements extérieurs pouvant perturber la colonie. Afin de rendre plus robuste le système dans les conditions difficiles, une étape de post-traitements a été mise en place. Étant donné qu'une perturbation de type « humain » ou « bateau » dure dans le temps, un lissage des détections est ajouté au système en suivant deux stratégies : 1. Une fusion des détections similaires proches dans le temps (perte de détection). 2. Une suppression des détections isolées (fausse détection). La première stratégie permet d'agréger deux détections similaires qui apparaissent à moins de 10 secondes d'intervalle tandis que la seconde élimine ensuite les détections de moins de 5 secondes.

2.2 - Comptage s'appuyant sur l'estimation de la densité : *IOCFormer*

Le modèle utilisé dans la version précédente [3] présente des lacunes significatives lorsqu'il s'agit de résoudre les ambiguïtés créées par des phoques qui sont partiellement occultés par d'autres. Cette limitation découle de la nature des données d'entraînement et de la structure même du réseau de neurones, qui ne permet pas de détecter plus d'un phoque dans une fenêtre de taille 70×100 pixels. Cette incapacité peut avoir des conséquences importantes pour l'estimation de la taille de la colonie, dégradant de ce fait les données fournies au parc marin pour lequel la qualité du comptage est importante pour évaluer correctement l'état de la population de phoque dans le but de conserver leur habitat. Pour pallier ces imprécisions, un changement de paradigme a été opéré dans la nouvelle version du projet. En effet, l'algorithme proposé ne repose plus sur une classification par fenêtre glissante, mais sur un comptage par estimation de densité.

Cette technique vise à convertir le problème de comptage en une tâche de prédiction de densité. Au lieu de compter directement les objets, cette méthode génère une carte de densité où chaque pixel représente la densité probable d'objets à cet emplacement. Cette technique est particulièrement utile dans des scènes où les objets sont groupés ou se chevauchent, car elle permet de modéliser la densité d'objets même dans les zones où ils ne sont pas clairement séparés. En particulier, le modèle *IOCFormer* [7] a été utilisé, car il présente une approche performante pour traiter le problème du comptage des objets dits indiscernables. Après une extraction de caractéristiques au moyen d'un encodeur (*ResNet-50* [8]), ce modèle s'appuie sur deux branches complémentaires, à savoir : 1. Une branche de densité qui estime la position des objets (c.-à-d. des phoques dans notre cas) de manière floue. 2. Une branche de régression qui prédit la position exacte des objets précédemment détectés.

Précisément, la branche de densité utilise un ensemble de convolutions pour produire une carte de densité approximative de la position des objets à partir des caractéristiques disponibles en sortie de l'encodeur. Cette carte est supervisée par une fonction de perte pour aligner la densité estimée avec le nombre réel d'objets dans l'image. La seconde branche, quant à elle, prend en entrée les caractéristiques extraites par l'encodeur ainsi que la carte de densité produite par la première branche pour affiner les prédictions, notamment en améliorant la détection des objets qui se chevauchent ou se fondent dans l'environnement. Ceci est mis en œuvre à l'aide d'un *transformer*, nommé *DETE*, supervisé par une fonction de perte qui compare les prédictions aux annotations. L'innovation clé de l'approche est la mise au point du *DETE* (c.-à-d. *Density-Enhanced Transformer Encoder*) qui joue un rôle central dans l'amélioration des performances, voir figure 3.

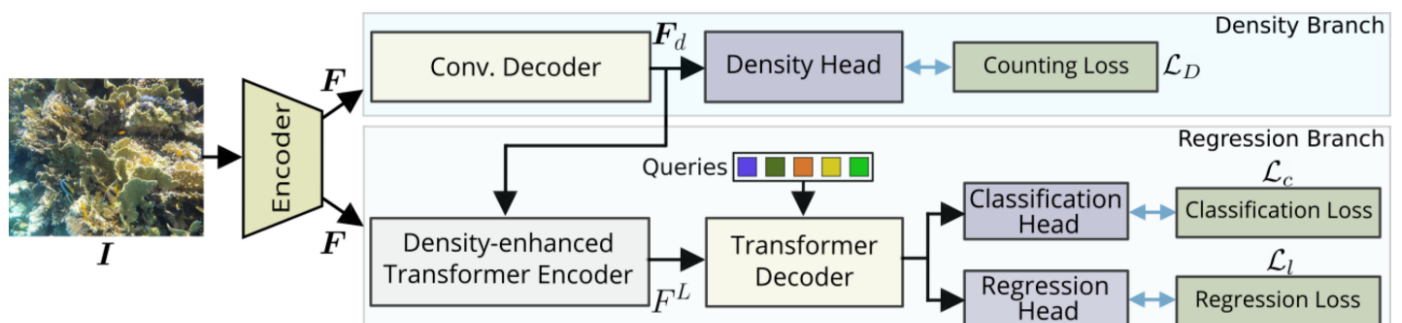


Figure 3 : Architecture de la méthode *IOCFormer* avec, à gauche, l'entrée est l'extraction des caractéristiques par l'encodeur *ResNet-50*, en haut à droite la branche de densité et en bas à droite la branche de régression (image issue de l'article original [7]).

3 - Expérimentations et résultats

Avant une utilisation en conditions réelles, les deux outils proposés doivent être amenés à maturité, quand c'est nécessaire, et validés sur un jeu de données préétabli. Pour cela, nous présentons dans cette partie, les bases de données qui ont permis l'évaluation quantitative des modèles, les détails liés à l'apprentissage de l'architecture *IOCFomer* ainsi que les résultats obtenus.

3.1 - Validation des mesures de perturbations

Pour la détection des perturbations humaines, le réseau de neurones *YOLO v8 m* (c.-à-d. la version de taille moyenne) a été utilisé tel qu'il est fourni, pré-entraîné sur la base de données d'images *COCO*. Cette version offre un bon compromis entre performance et rapidité en adéquation avec la puissance de calcul disponible à terre pour analyser les images. Afin de détecter les perturbations, seules les détections relatives aux classes « humain » et « bateau » sont conservées. Aussi, toutes les détections dont la confiance est inférieure à 0.75 (c.-à-d. 75 % de confiance) sont éliminées pour éviter les fausses alarmes. Enfin, les deux stratégies présentées dans la partie 2.1 sont appliquées pour consolider les détections. La première permet de fusionner les détections similaires proches dans le temps pour corriger les pertes de détection. La seconde limite les fausses alarmes en supprimant les détections isolées.

Base de données

Afin de valider les performances de l'approche mise en œuvre, une base de données a été constituée. Elle comprend 200 images réparties équitablement en deux jeux de données : perturbation par un humain et/ou un bateau (possiblement plusieurs), pas de perturbation. Pour chacune des images de la première catégorie, les boîtes minimales englobant les perturbations ont été manuellement annotées et utilisées comme vérité terrain, voir figure 4. Ces images sont issues de la précédente campagne d'acquisition qui embarquait une caméra similaire avec une résolution de 1920×1080 pixels filmant à 10 images par seconde.

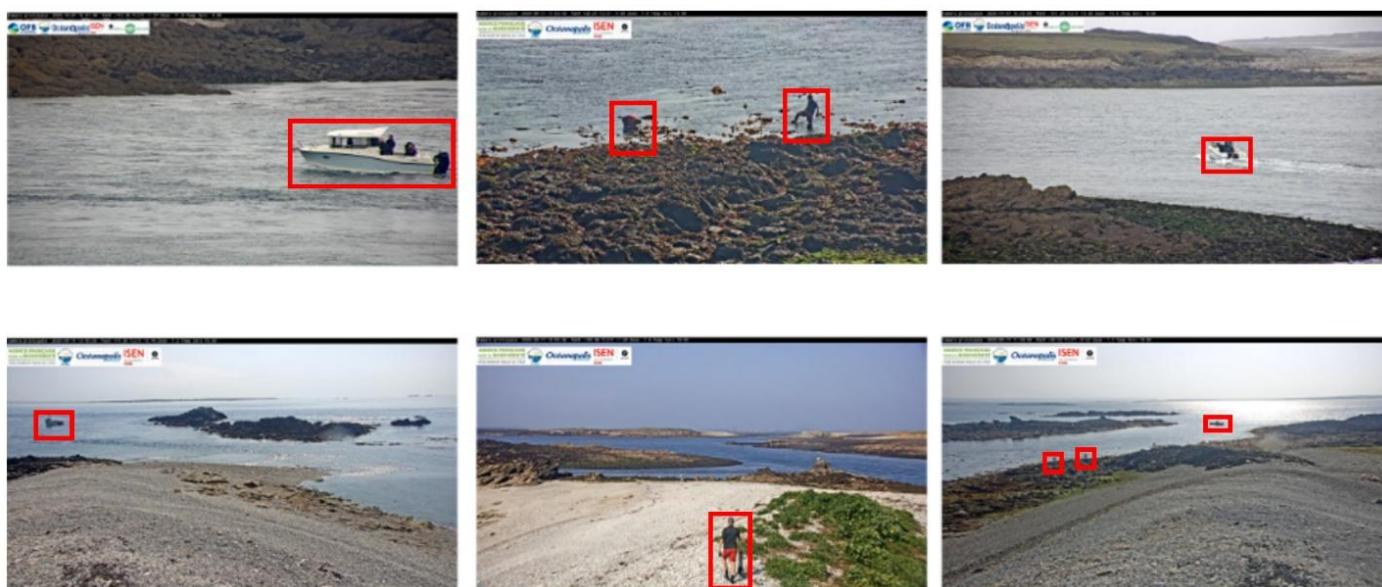


Figure 4 : Exemple de 6 images de la base de données appartenant à la catégorie « perturbation ». En rouge, les boîtes minimales englobantes des objets (humain ou bateau) utilisées comme vérité terrain.

Résultats

Sur la base de données de test, les résultats obtenus montrent une performance de 97 % avec la répartition suivante entre les deux jeux de données : 94 % de bonnes détections pour la catégorie « perturbation », 100 % de non-détections pour la catégorie « pas de perturbation ». En particulier,

1 erreur a été commise sur la non-détection d'un bateau et 5 sur la non-détection d'humains. Ces erreurs sont dues à une difficulté importante à différencier les pêcheurs à pied de certains rochers lorsque les individus sont très éloignés de la caméra, voir figure 4. Concernant les bateaux, la détection est plus aisée, même s'ils restent parfois difficiles à détecter lorsqu'ils sont au loin, parfois masqués à intervalles réguliers par les vagues au large de l'îlot. Pour permettre une exploitation future par les conservateurs du parc marin, les résultats obtenus sont exportés dans un fichier CSV afin d'archiver les perturbations. Un exemple d'extraction est visible dans le tableau 1. On y retrouve la date et l'heure de la perturbation, son type et sa position dans l'image. Ainsi, il est aisé de croiser les résultats des perturbations et de comptage pour estimer l'impact humain sur la colonie. À titre d'exemple, la figure 4 montre que lorsqu'il y a une perturbation, les phoques quittent généralement l'îlot.

Date et heures	Types d'objets	Position
21/09/2020 13:32:15	Bateau	[10, 12, 512, 312]
21/09/2020 14:52:10	Bateau	[150, 565, 220, 622]
21/09/2020 18:02:10	Humain	[155, 154, 223, 312]
21/09/2020 18:02:10	Humain	[256, 189, 353, 359]
22/09/2020 10:17:15	Bateau	[45, 652, 125, 732]
22/09/2020 12:22:48	Bateau	[456, 785, 589, 936]
22/09/2020 15:02:51	Bateau	[221, 375, 452, 427]

Tableau 1 : Exemple d'export au format CSV des perturbations détectées.

3.2 - Validation du comptage

Le comptage des phoques gris a été réalisé avec l'approche *IOCFomer*, initialement utilisée et évaluée sur une base de données de poissons. Étant donné les différences entre les deux contextes expérimentaux, poissons et phoques, une nouvelle base de données annotée a dû être mise en place pour permettre un réapprentissage du réseau de neurones.

Base de données

La base de données établie a été collectée à partir de 5 vidéos, chacune d'une durée de 60 minutes, capturant divers scénarios. Cette extraction vise à capturer des scènes variées afin d'assurer que le modèle puisse généraliser à différentes conditions de prise de vue ou de météo par exemple. Un échantillon de 1000 images a été sélectionné pour servir de base à l'entraînement, à la validation et au test du modèle. Afin d'établir une vérité terrain nécessaire à mesurer quantitativement la robustesse de la méthode, chacune des 1000 images a été annotée à l'aide des outils *LabelImg* [9] et *RectLabel*. En particulier, les phoques présents dans les images ont été annotés par un point situé au centre de l'individu, comme le requiert la méthode *IOCFomer*. Cette phase a permis d'obtenir un total de 6084 annotations. Afin d'entraîner le modèle, de le valider et finalement de le tester, la base de données a été divisée en 3 jeux de données répartis comme suit : 800 images pour l'entraînement, 100 images pour la validation et 100 images pour le test.

Apprentissage du modèle

L'implémentation de la méthode *IOCFomer* s'appuie sur le code officiel [11] développé avec *PyTorch* pour des GPU NVIDIA. Les spécificités de l'architecture sont les mêmes que pour la méthode originale, à savoir 4 blocs *transformer* pour la partie *DETE* avec 700 *queries*. Les caractéristiques sont extraites avec un encodeur ResNet-50 pré-entraîné sur *ImageNet*. Pour les augmentations de données, nous utilisons un redimensionnement aléatoire et un retournement horizontal. Les images sont recadrées aléatoirement pour obtenir des images d'entrées de taille 256 x 256 pixels. L'entraînement est réalisé sur 2 GPU NVIDIA Quadro RTX 8000 avec des lots de 4 images sur une

durée de 1500 *epochs* avec l'optimiseur Adam [12]. Lors de l'inférence, les images sont divisées en vignettes de taille identique à celles utilisées pendant l'entraînement. Enfin, un seuil de 0.35 est utilisé pour filtrer les prédictions [13].

Les métriques utilisées pour vérifier la qualité de l'apprentissage sont le MAE (*Mean Absolute Error*) et le MSE (*Mean Squared Error*). L'Erreur Absolue Moyenne (MAE) est la moyenne des différences absolues entre les valeurs prédites et la vérité terrain. Elle est également connue sous le nom de norme L1 ou distance de Manhattan :

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

Où N est le nombre d'observation, y_i la valeur prédite pour la i^e observation et $\hat{y}_i$ la vérité terrain associée à cette même observation.

Le MSE (Erreur Quadratique Moyenne), quant à lui, est la moyenne des carrés des différences entre les valeurs prédites et la vérité terrain. Elle est également connue sous le nom de norme L2 ou distance euclidienne :

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

La figure 5, montre l'évolution des métriques au cours de l'apprentissage. On remarque des résultats cohérents pour les deux métriques observées avec des erreurs qui se stabilisent à partir de 1200 *epochs*. La sélection du meilleur modèle a été réalisée en étudiant les performances du MSE sur le jeu de données de validation à chaque *epoch*.

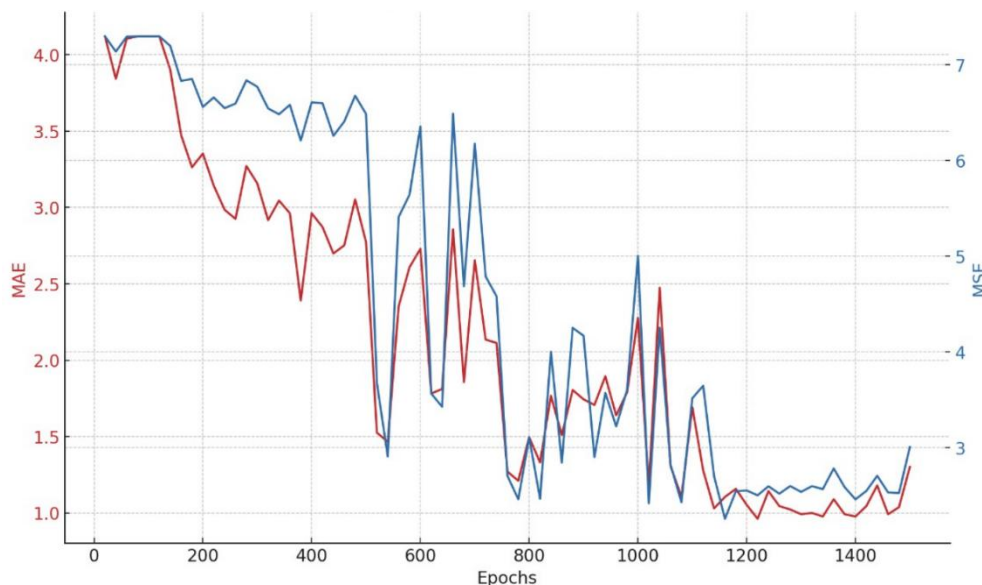


Figure 5 : Visualisation de l'évolution du MAE et du MSE au cours de l'apprentissage. On remarque une stabilisation des métriques après 1200 epochs environ.

Résultats quantitatifs

Les résultats obtenus sur la base de données de test montrent de bons résultats qualitatifs, même sur des images présentant un grand nombre de phoques, voir figure 6. La performance globale du système sur l'ensemble de test est de 91.2 %, en progression par rapport à la méthode précédente qui avait une performance de 87.24 %. On note cependant que les résultats ne sont pas

comparables, car les bases de données utilisées pour l'évaluation sont différentes. Cependant, l'approche proposée montre des perspectives intéressantes pour le comptage des phoques en situation difficile. Finalement, les résultats du comptage sont exportés dans un fichier CSV regroupant les informations visibles dans le tableau 2. On y retrouve l'horodatage du comptage ainsi que le nombre d'individus présent sur l'image.

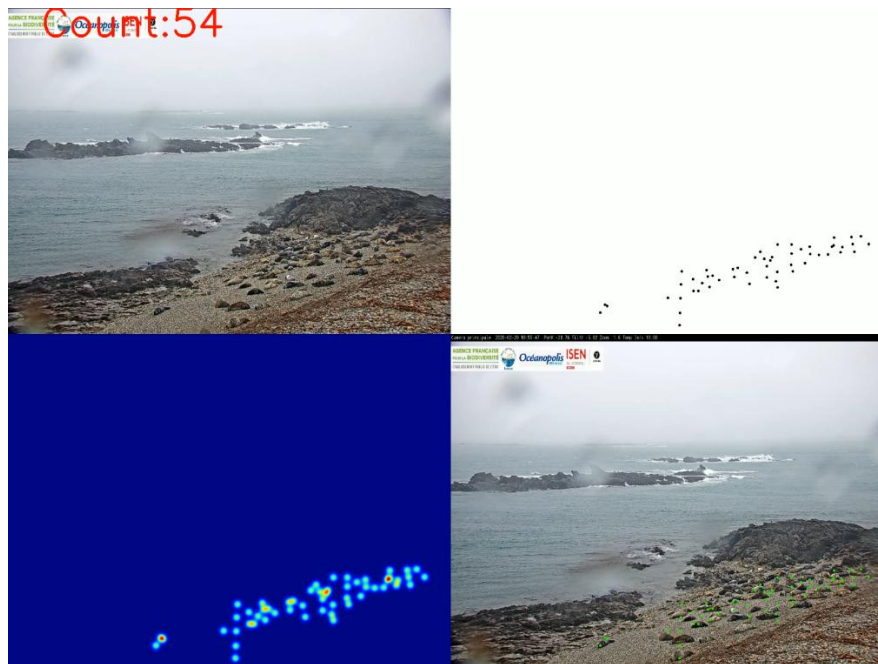


Figure 6 : Résultats de la méthode de comptage proposée avec, l'image originale avec le nombre de phoques (en haut à gauche), l'estimation de la densité (en bas à gauche), l'extraction des individus à partir de la densité (en haut à droite) et la prédiction finale (en bas à droite).

Date et heures	Nombre
22/12/2019 11:00	58
22/12/2019 11:15	55
22/12/2019 11:30	59
22/12/2019 11:45	62
22/12/2019 12:00	58
22/12/2019 12:15	46
22/12/2019 12:30	48

Tableau 2 : Exemple d'export au format CSV du nombre de phoques détectés.

4 - Conclusion

Les résultats obtenus avec les deux approches proposées montrent que l'étude des perturbations ainsi que le comptage sont possibles en utilisant la vision par ordinateur. D'une part, l'architecture YOLO v8 a permis d'identifier avec robustesse les perturbations humaines qui peuvent survenir autour de l'îlot. D'autre part, l'approche IOCFormer a montré sa pertinence pour l'estimation de la densité de la colonie dans des conditions parfois difficiles. Cependant, même si les résultats obtenus répondent bien aux attentes du projet, certains points d'améliorations restent possibles. Le premier élément qui n'a pas été pris en compte dans le projet, malgré sa disponibilité, est l'information issue du capteur infrarouge de la caméra. En effet, cette nouvelle modalité pourrait permettre de robustifier l'estimation de la densité de phoques gris en réalisant une fusion des informations issues des deux types d'images. Cette fusion serait alors réalisée, soit tardivement en fusionnant les estimations de densité, soit précocement lors de l'apprentissage du réseau de neurones qui prendrait alors en entrée les deux modalités d'image plutôt que seulement l'image

dans le domaine visible. D'autre part, il faut noter que derrière les réussites liées aux grandes tâches de vision telles que la reconnaissance de visage, le développement des voitures autonomes ou la génération d'images synthétiques se cache des besoins importants en bases de données annotées, nécessaires aux phases d'entraînement, ainsi qu'en puissance de calcul. Afin de pallier ces contraintes, qui pourraient avoir un impact significatif si le système de décision était amené à devenir autonome sur l'îlot, une approche de frugalité pourrait être envisagée. En particulier, une approche de réduction de la taille, et donc du coût calculatoire, des réseaux de neurones pourrait être envisagée en s'appuyant sur les techniques de distillation de connaissances par exemple [14]. Finalement, la réussite de ce projet montre que les techniques de vision par ordinateur récentes peuvent permettre de répondre, avec une certaine simplicité, à des projets tels que la préservation de la biodiversité, en limitant les opérations humaines qui peuvent perturber les écosystèmes au travers de leurs observations sur le terrain. Ainsi, les données qui seront recueillies en 2025 et 2026 pourront permettre de valider la méthodologie mise en œuvre pour mieux comprendre les impacts humains sur la colonie de phoques gris installée sur l'îlot de Morgol dans l'archipel de Molène au large des côtes Finistérienne.

Remerciement :

Nous remercions Jean-Yves MULOT, Philippe FORJONEL et Léo-Paul PELLETIER du LabISEN pour leur aide tout au long du projet ainsi que l'Office Français de la Biodiversité et le Parc Naturel Marin d'Iroise pour le financement du projet ainsi que pour leur soutien.

Références :

- [1] Zizhu FAN, Hong ZHANG, Zheng ZHANG et al. A survey of crowd counting and density estimation based on convolutional neural network. *Neurocomputing*. 2022. vol. 472, p. 224-251.
- [2] Jia DENG, Wei DONG, Richard SOCHER et al. Imagenet: A large-scale hierarchical image database. *Conference on Computer Vision and Pattern Recognition (CVPR)*. 2009. p. 248-255.
- [3] Ayoub KARINE, Thibault NAPOLÉON, Jean-Yves MULOT et al. Video seals recognition using transfer learning of convolutional neural network. *International Conference on Image Processing Theory, Tools and Applications (IPTA)*. 2020. p. 1-4.
- [4] Karen SIMONYAN and Andrew ZISSERMAN. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- [5] Glenn JOCHER, Ayush CHAURASIA and Jing QIU. Ultralytics YOLOv8. <https://github.com/ultralytics/ultralytics>. 2023.
- [6] Tsung-Yi LIN, Michael MAIRE, Serge BELONGIE et al. Microsoft coco: Common objects in context. *European Conference on Computer Vision (ECCV)*. 2014. p. 740-755.
- [7] Guolei SUN, Zhaochong AN, Yun LIU et al. Indiscernible object counting in underwater scenes. *Computer Vision and Pattern Recognition (CVPR)*. 2023. p. 13791-13801.
- [8] Kaiming HE, Xiangyu ZHANG, Shaoqing REN et al. Deep residual learning for image recognition. *Computer Vision and Pattern Recognition (CVPR)*. 2016. p. 770-778.
- [9] <https://github.com/HumanSignal/labellmg>
- [10] <https://rectlabel.com/>
- [11] <https://github.com/GuoleiSun/Indiscernible-Object-Counting/tree/main>

[12] Diederik P. KINGMA and Jimmy BA. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.

[13] Dingkan LIANG, Wei XU and Xiang BAI. An end-to-end transformer model for crowd localization. European Conference on Computer Vision (ECCV). 2022. p. 38-54.

[14] Ayoub KARINE, Thibault NAPOLÉON and Maher JRIDI. Channel-spatial knowledge distillation for efficient semantic segmentation. Pattern Recognition Letters, 2024, vol. 180, p. 48-54.

Computer Vision : Introduction au calcul du volume des objets

Simon PLAYE¹

Édité le
12/03/2025

école
normale
supérieure
paris-saclay

¹ Data Scientist chez Sicara (Theodo Data & AI)

Cette ressource fait partie du N° 115 de La Revue 3EI du deuxième trimestre 2025.

Avez-vous déjà travaillé sur le calcul du volume des objets à partir de vidéos en utilisant l'IA ? Cela s'est-il avéré soit peu pertinent, soit extrêmement difficile à réaliser ? Si vous avez répondu « oui » à ces deux questions et que vous êtes passionné par le calcul du volume ou l'IA, cet article est fait pour vous. En utilisant deux vidéos, je vais vous montrer comment calculer le volume des objets.

Les cas d'utilisation concernent principalement la gestion des stocks. Alors que de plus en plus d'entreprises cherchent à réduire leur impact environnemental, un tel outil peut permettre d'améliorer l'efficacité des chaînes d'approvisionnement et optimiser la gestion des stocks. Cet outil peut également être utilisé pour le placement de meubles : design d'intérieur, planification d'événements, construction ou rénovation...

Le calcul du volume des objets soulève plusieurs défis :

- Localiser les objets dans la vidéo
- Définir leurs limites
- Mettre à l'échelle la vidéo pour convertir une distance en pixels, en mètres

Cet article passera en revue toutes les étapes ci-dessous pour expliquer comment effectuer le calcul du volume des objets :

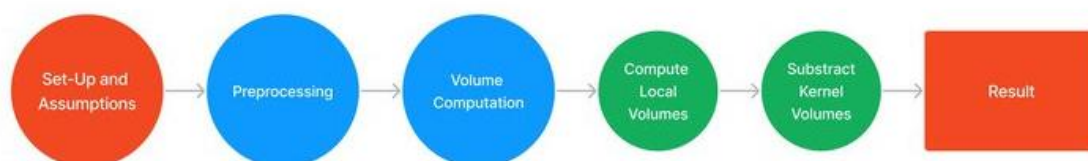


Figure 1 : Étapes du processus de calcul du volume.

1 - Configuration sous-jacente et hypothèses

Pour effectuer le calcul du volume des objets, je m'appuie sur quelques hypothèses :

- L'espace filmé est une pièce fermée.
- Cette pièce est filmée deux fois depuis le même point de vue : une fois sans les objets et une fois avec eux.
- Les vidéos sont identiques, à l'exception des objets.
- Les objets sont placés directement contre un mur, sans aucun espace.

Par souci de simplicité, dans le reste de cet article, la pièce filmée est vide, et les objets sont des boîtes rectangulaires. Par conséquent, le calcul du volume sera effectué sur ces boîtes. Voici un exemple de séquence vidéo respectant ces hypothèses :



Figure 2 : Un exemple de séquence vidéo respectant ces hypothèses. Ici, les objets étudiés sont les deux boîtes.

L'outil le plus important pour effectuer le calcul du volume est la caméra. En effet, une caméra 2D normale n'affiche que des valeurs de couleur pour tous les pixels filmés. Dans cet article, la configuration consiste en une caméra spéciale nommée Intel RealSense D435. Cette caméra est composée de deux caméras infrarouges qui offrent une représentation 3D de ce qui est filmé. Une API Python, `pyrealsense2`, permet de récupérer, pour chaque pixel, une coordonnée 3D. Ainsi, il est possible d'obtenir la distance en mètres par rapport à la caméra sur l'axe des x (de droite à gauche), sur l'axe des y (de haut en bas), et sur l'axe des z (de près à loin) pour tous les pixels. N'importe quelle autre caméra peut être utilisée tant qu'elle fournit ces coordonnées.

1.1 - Quelques mots sur `pyrealsense2`

`Pyrealsense2` est l'enveloppe Python pour le SDK Intel RealSense 2.0, qui est une bibliothèque pour les caméras Intel RealSense. Le SDK peut normalement être utilisé avec `librealsense`, un package C++. Selon la page Github de `librealsense` : « Le SDK permet la diffusion de profondeur et de couleur, et fournit des informations sur la calibration intrinsèque et extrinsèque. La bibliothèque offre également des flux synthétiques (nuage de points, profondeur alignée avec la couleur et vice-versa), ainsi qu'un support intégré pour l'enregistrement et la lecture des sessions de diffusion. » Pour fournir ces services, le SDK utilise un modèle déterministe qui repose sur les entrées des deux caméras infrarouges.

1.2 - Prétraitement

Entrons maintenant dans le processus de calcul du volume. Voici la profondeur des images de ces deux vidéos :



Figure 3a : Image brute avec boîte



Figure 3b : Image brute sans boîte

Ces images illustrent un problème statistique classique : les valeurs aberrantes (outliers) :

- Les zones noires sur les images des vidéos correspondent à des pixels sans coordonnées (la profondeur est définie à 0).
- Les zones colorées correspondent à des coordonnées erronées (la profondeur est de 10 mètres, alors que le mur est à 1,5 mètre de la caméra).

1.3 - Traiter les valeurs aberrantes : bonnes pratiques

Tout d'abord, définissons plus précisément ce qu'est une valeur aberrante. Ici, une valeur aberrante est un pixel dont la valeur de la coordonnée z (correspondant à la profondeur) est égale à 0 ou supérieure à 1,6 mètre. Puisque le mur est à environ 1,5 mètre de la caméra, aucun pixel ne peut avoir une valeur z supérieure à 1,6 mètre.

Cette méthode permet également de sélectionner les pixels ayant des valeurs aberrantes pour les coordonnées x et y . En effet, un contrôle rapide montre que les pixels avec une valeur de coordonnée z égale à 0 ou supérieure à 1,6 ont aussi des valeurs aberrantes pour les coordonnées x et y . En revanche, aucun pixel n'a été trouvé avec des coordonnées x et y aberrantes, mais une valeur de coordonnée z comprise entre 0 et 1,6. Par conséquent, seule la valeur de la coordonnée z définit si un pixel est une valeur aberrante.

Prenons un pixel spécifique d'une image d'une vidéo et imaginons que ce pixel soit une valeur aberrante.

La meilleure manière de remplacer la valeur de la coordonnée z est d'utiliser la valeur de la coordonnée z du pixel non aberrant le plus proche. Ici, le « pixel non aberrant le plus proche » est le pixel ayant la distance euclidienne la plus faible par rapport à notre pixel.

La méthode utilisée pour récupérer les coordonnées x et y est plus compliquée. En nous concentrant d'abord sur x , utilisons la grille ci-dessous pour visualiser la méthode employée. Cette grille donne les coordonnées x des pixels situés sur une portion 7x7 de l'image. Les cellules rouges représentent des valeurs aberrantes, les cellules vertes représentent des pixels non aberrants. La valeur de la coordonnée x est écrite à l'intérieur des pixels non aberrants.

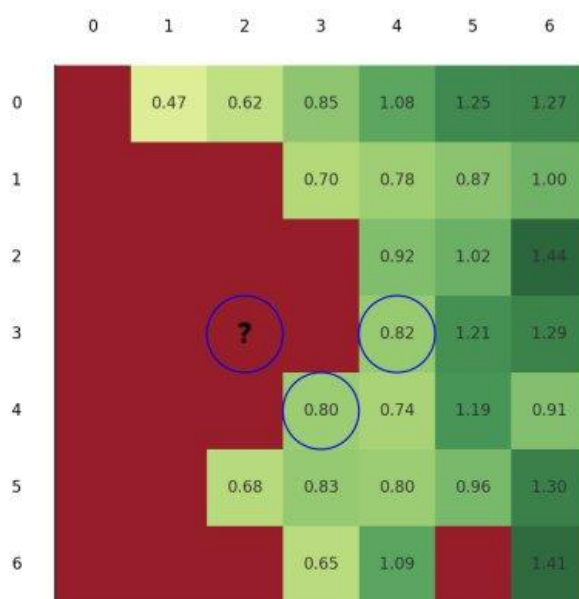


Figure 4 : Sous-ensemble de l'image : les zones rouges sont des valeurs aberrantes, les zones vertes sont des pixels avec des valeurs

Ici, je souhaite calculer la coordonnée x pour l'outlier entouré « ? » situé en (3,2). Pour ce faire, j'ai d'abord sélectionné les deux points les plus proches de cet outlier (entourés dans la grille) qui ne se trouvent pas dans la même colonne. La différence de coordonnée x entre ces deux pixels est $0.82 - 0.80 = 0.02$. Ces deux pixels sont distants d'une colonne. Par conséquent, on peut estimer qu'en se déplaçant d'une colonne, la valeur de x change de 0.02. Ainsi, puisque l'outlier est situé à deux colonnes de ce pixel entouré (0.82), sa valeur estimée est $0.78 (= 0.82 - 0.02 * 2)$.

La même méthode est utilisée pour remplacer les valeurs des coordonnées y pour les outliers.

Ces méthodes de prétraitement permettent d'obtenir une bien meilleure représentation 3D de ce que la caméra a filmé (ici pour la profondeur) :

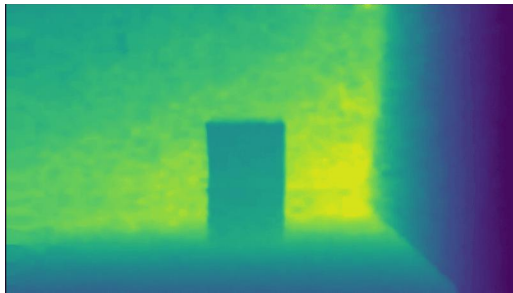


Figure 5a : Image brute avec boîte

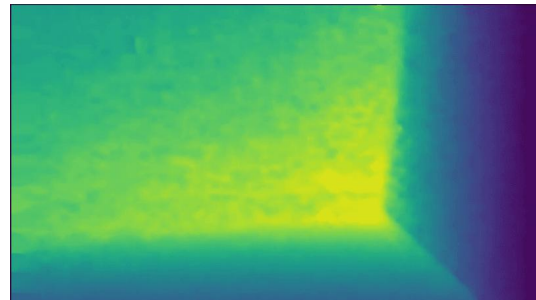


Figure 5b : Image brute sans boîte

2 - Processus de calcul du volume

2.1 - Technique pour effectuer le calcul du volume à l'aide de deux vidéos

En disposant de vidéos avec des coordonnées 3D cohérentes, je peux effectuer le calcul du volume. En lisant cet article, vous vous êtes peut-être demandé : « Pourquoi insister sur l'enregistrement d'une vidéo sans boîtes alors que seul le volume des boîtes est nécessaire ? » L'astuce pour effectuer le calcul du volume des boîtes est de calculer le volume d'une image de la vidéo avec les boîtes, puis de soustraire ce volume du volume de la même image sans boîtes. Cette différence correspond au volume calculé des boîtes.

2.2 - Calcul des volumes locaux à l'aide de noyaux

Comment effectuer le « calcul du volume » sur une seule image ? En regardant l'image ci-dessous, on peut la diviser en petits carrés, appelés noyaux :

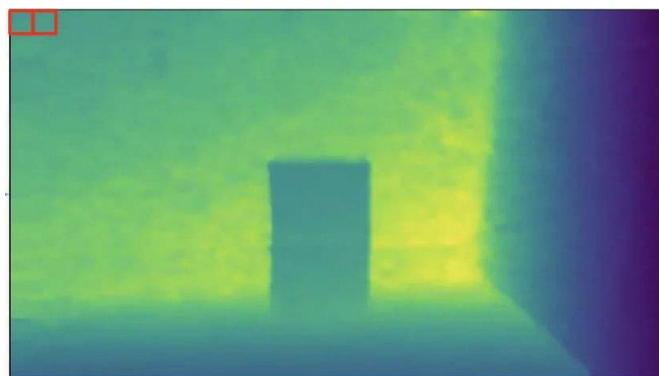


Figure 6 : Image traitée avec boîtes et noyaux

Simplifions ce problème en commençant par essayer de calculer les aires des noyaux, puis leurs volumes correspondants.

Ayant les coordonnées 3D des quatre pixels qui composent un noyau, on peut calculer l'aire du noyau de différentes manières. Pour être plus précis, nous avons approximé un noyau à un parallélogramme. Ensuite, j'ai éliminé les coordonnées z des points et utilisé cette formule :

L'aire d'un parallélogramme peut être calculée en utilisant la formule suivante :

$$A = \|\overrightarrow{AB} \wedge \overrightarrow{AD}\|$$

Formule vectorielle de l'aire d'un parallélogramme

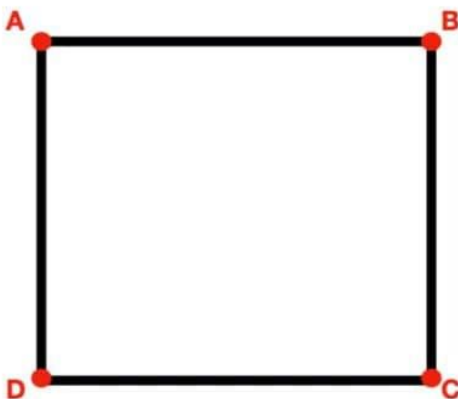


Figure 7 : Un noyau

Cette formule ignore un point (ici C), et les quatre points sont considérés comme étant dans le même plan 2D, car les coordonnées z de chaque point sont ignorées.

En se concentrant sur un seul noyau, on peut considérer ce noyau comme la base d'un prisme rectangulaire, l'autre base étant située aux coordonnées z=0 et étant identique. Une manière de l'imaginer est de penser que la caméra est située sur un plan perpendiculaire à l'axe des z. Par conséquent, le prisme rectangulaire aura une base sur ce plan et l'autre sera le noyau sur votre image. La hauteur de ce prisme rectangulaire sera la moyenne des coordonnées z de tous les points à l'intérieur du noyau, comme les points A, B, C, D, E, F et G dans l'image ci-dessous :

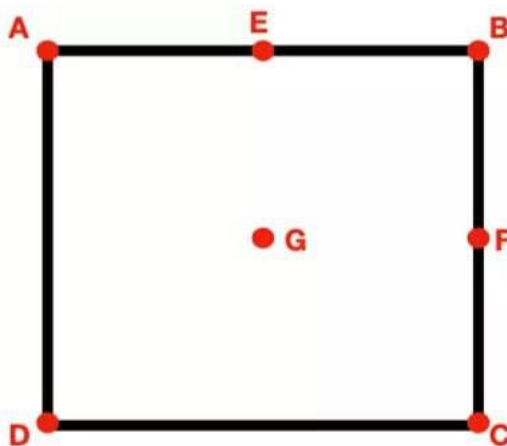


Figure 8 : E, F et G sont également utilisés pour calculer la hauteur moyenne

Ainsi, en calculant l'aire d'un noyau et en obtenant la hauteur du prisme rectangulaire correspondant, on peut effectuer le calcul du volume pour un noyau. En couvrant une image avec des noyaux et en additionnant leurs volumes correspondants, il est possible d'obtenir le volume total de l'image. Voici une représentation visuelle du volume : un pixel de chaque image représente le volume calculé avec un noyau de 2x2 pixels :

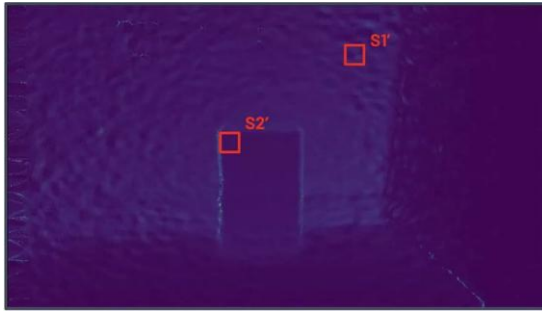


Figure 9a : Calcul du volume- Image brute avec boîte

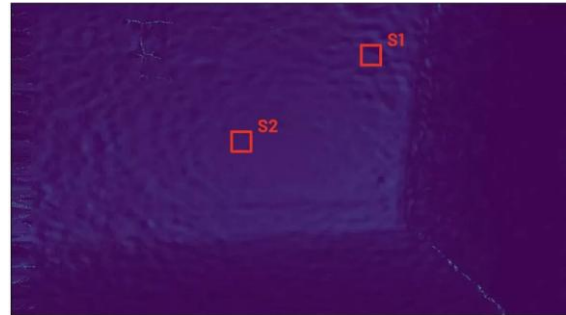


Figure 9b : Calcul du volume - Image brute sans boîte

Dans les images ci-dessus, j'ai ajouté deux carrés : S1 et S2 sur l'image sans boîtes, et S1' et S2' sur l'image avec boîtes. En comparant S1 et S1', les volumes devraient être approximativement les mêmes, car ils couvrent la même surface et ont la même profondeur. Ainsi, la différence entre S1 et S1' devrait être proche de 0. Cependant, en comparant S2 et S2', le volume de S2 devrait être plus grand que celui de S2' car S2' se trouve sur une boîte. Ainsi, même si S2 et S2' ont la même surface, la profondeur moyenne des pixels dans S2' est inférieure à celle de S2. La différence entre S2 et S2' correspond au volume occupé par la boîte. Par conséquent, en recouvrant les deux images de carrés et en calculant la différence de volume pour chaque carré correspondant, on peut obtenir le volume des boîtes.

3 - Conclusion

Pour conclure, le calcul du volume en utilisant cette technique donne de très bons résultats pour estimer le volume des objets. Pour deux boîtes, l'erreur absolue entre le volume calculé et le volume réel est inférieure à 1 %. Pour une et trois boîtes, cette erreur est inférieure à 3 %. Malgré les nombreuses hypothèses sous-jacentes concernant la pièce et les objets, cet algorithme se révèle efficace pour le calcul du volume des objets. De plus, il fournit une base pour des applications intéressantes en vision par ordinateur, telles que le calcul du volume de différents types d'objets en utilisant des algorithmes de détection d'objets.

4 - Références :

[1]: <https://data-ai.theodo.com/en/technical-blog/mastering-volume-computation-of-objects-from-videos>

IA Générative : Apprenez à contrôler la génération d'images avec Stable Diffusion

Simon PLAYE¹

Édité le
11/03/2025

école —————
normale —————
supérieure —————
paris — saclay —————

¹ Data Scientist chez Sicara (Theodo Data & AI)

Cette ressource fait partie du N° 115 de La Revue 3EI du deuxième trimestre 2025.

1 - Introduction

Avez-vous déjà utilisé ChatGPT pour créer des images et avez-vous été déçu par les résultats ? Si vous avez ressenti cette frustration ou si vous êtes simplement intéressé par des outils de génération d'images plus efficaces, Stable Diffusion pourrait être la solution idéale.

Avec la popularité de ChatGPT, nombreux sont ceux qui sont curieux de découvrir ce que l'IA générative peut accomplir. Elle ne se limite pas à la création de texte ; cette technologie peut également générer des images, des vidéos et même de la musique. Pourtant, il est parfois difficile d'obtenir les images que vous souhaitez simplement à partir d'une invite textuelle.

Cet article se penche sur Stable Diffusion, un outil conçu spécifiquement pour générer des images. Ici, je vous montrerai comment utiliser une API pour contrôler la création d'images avec Stable Diffusion, en couvrant des options tant pour les non-développeurs que pour les développeurs. Les méthodes abordées dans cet article incluent :

- Diriger le processus de diffusion vers une image spécifique (IP Adapter).
- Maintenir certaines caractéristiques dans une image via le processus de diffusion (ControlNet).
- Extraire des caractéristiques à partir d'une image initiale (Image-to-Image).
- Former un modèle spécifique avec un ensemble limité d'images pour des ajustements ciblés (LORA, Model Fine-Tuning).
- Ajouter de nouveaux poids aux couches de croisement d'attention pour modifier les caractéristiques de l'image (LORA).

2 - Qu'est-ce que Stable Diffusion ?

Stable Diffusion est un modèle de diffusion spécifiquement conçu pour la génération d'images. Il commence avec un motif initial de bruit aléatoire et affine systématiquement ce bruit ou le "débruite" pour produire des images qui ressemblent de près à des photos réelles. Le modèle guide cette transformation en appliquant certaines conditions (par exemple une invite textuelle) durant le processus de débruitage.

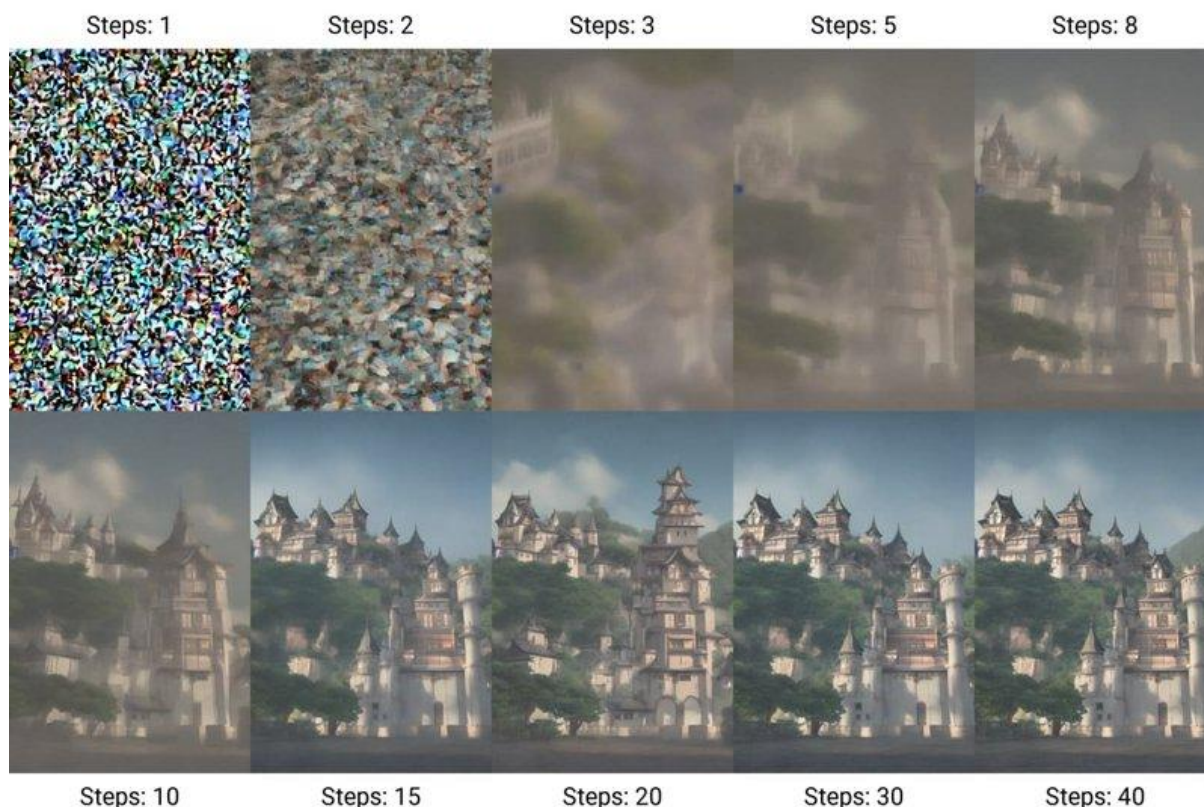


Figure 1 : Les étapes de débruitage

source: https://en.wikipedia.org/wiki/Diffusion_model#/media/File:X-Y_plot_of_algorithmically-generated_AI_art_of_European-style_castle_in_Japan_demonstrating_DDIM_diffusion_steps.png

Contrairement aux méthodes traditionnelles qui se contentent de prompts textuels, Stable Diffusion offre des contrôles plus flexibles et sophistiqués :

- Conditionnement basé sur l'image : Grâce à des fonctionnalités comme l'IP-Adapter et ControlNet, Stable Diffusion peut utiliser une image existante pour orienter la génération, permettant des modifications ou des améliorations basées sur cette image.
- Transformation Image-à-Image : Cette fonctionnalité permet au modèle de partir non pas de zéro mais d'une image existante, qu'il transforme ensuite en une nouvelle création.
- Contrôle de style : Les utilisateurs peuvent choisir des modèles générant des images dans des styles artistiques spécifiques, ou même entraîner leurs modèles pour produire des styles personnalisés.
- Conditionnement hybride : Il permet de combiner des invites textuelles, des images et différents modèles pour conditionner la génération, offrant un contrôle sans précédent sur la sortie.

Pour expérimenter avec ces diverses techniques de contrôle de la génération d'images, ModelsLab propose une API Stable Diffusion pratique. Elle est disponible sous deux formats :

- **Version Playground** : Une interface conviviale conçue pour l'expérimentation et l'exploration sans nécessiter de connaissances techniques étendues.
- **API pour Développeurs** : Offre un contrôle et une personnalisation plus détaillés, adaptés aux développeurs souhaitant intégrer ces capacités dans leurs applications.

Dans les sections suivantes, je vais approfondir chaque technique, en montrant comment utiliser la plateforme ModelsLab pour réaliser votre vision créative.

2.1 - Stable Diffusion text-to-image

Avant de plonger dans les divers outils permettant de contrôler la génération d'images, commençons par une introduction à l'API Stable Diffusion text-to-image. Cette API, semblable à Dall-E, facilite la génération d'images à partir de prompts textuels.

Pour les développeurs, il existe deux points de terminaison principaux : **text2img** et **realtime-stable-diffusion**. Le point de terminaison **text2img** est le choix principal pour la génération d'images, accessible via le playground et l'API pour développeurs. Il permet aux utilisateurs de créer des images avec des modèles de diffusion entraînés par la communauté, disponibles sous les versions **Stable Diffusion** et **Stable Diffusion XL**. La version XL offre des images plus précises mais nécessite plus de temps pour la génération.

Dans le playground, les utilisateurs peuvent personnaliser leur génération d'images en définissant différents paramètres tels que :

- **Negative Prompt** : Spécifiez ce que vous ne voulez pas voir apparaître dans l'image.
- **Guidance Scale** : Définissez l'importance de l'influence du prompt sur le débruitage.
- **Steps** : Déterminez le nombre d'étapes de génération, ce qui affecte le détail de l'image.

L'API développeur offre également des options supplémentaires, telles que :

- **enhance_style** : Choisir un style spécifique pour l'image.
- **highres_fix** : Créer des images en haute résolution.

De plus, le point de terminaison **realtime-stable-diffusion** offre moins d'options de personnalisation et ne permet pas de choisir un modèle de diffusion, mais il est plus rapide pour générer des images.

Voici un exemple rapide de fonctionnement de **text2img**.

Prompt Input

"A black cat with a Christmas hat dancing on a table"

Image Output



Figure 2 : Un chat dansant sur une table avec un chapeau de Noël

3 - Contrôle de la génération d'images avec Stable Diffusion

Dans cette section, nous explorerons trois outils avancés fournis par Stable Diffusion qui offrent des façons alternatives de guider la génération d'images, au-delà du prompt textuel conventionnel :

- **IP-Adapter**
- **ControlNet**
- **Image-to-Image**

3.1 - IP-Adapter

Imaginez que vous puissiez guider la génération d'images non pas en donnant des instructions textuelles, mais en utilisant une image cible à la place. C'est exactement ce que fait **IP-Adapter**. Il permet de diriger le processus de génération avec une image, tout comme un prompt textuel. Cependant, l'IP-Adapter doit être utilisé avec un prompt textuel (ou une image, comme nous le verrons ci-dessous).

Cependant, l'IP-Adapter est uniquement disponible via l'API développeur sur le point de terminaison **img2img**. Dans ce paramètre, vous pouvez ajuster des paramètres tels que :

- **ip_adapter_id**, qui définit comment l'image est encodée,
- **ip_adapter_scale**, qui affecte l'impact de l'image IP-Adapter sur la réduction du bruit,
- **ip_adapter_image**, qui est l'URL de l'image IP-Adapter.

Adapter avec un prompt textuel

Prompt Input

"pink, girly, castle, disney, animation, love, flower"

IP-Adapter Image



Output Image



Figure 3 : Create a rose and cute castle with IP-Adapter (Créez un château rose et mignon avec IP-Adapter)

3.2 - ControlNet

ControlNet fonctionne de manière similaire à l'IP-Adapter, mais avec une approche plus nuancée. Contrairement à l'IP-Adapter, qui influence la génération d'images en fonction de l'image entière, ControlNet cible des caractéristiques spécifiques dans l'image pour guider le processus de génération vers ces détails.

Il existe plusieurs types de ControlNet qui permettent d'extraire des caractéristiques spécifiques d'une image :

- **softedge** : trouve les contours dans les images
- **canny** : délimite avec précision les frontières dans des environnements contrôlés
- **openpose** : détecte les points clés du corps humain, des mains, du visage et des pieds
- etc.

Vous pouvez également intégrer plusieurs modèles ControlNet en énumérant plusieurs paramètres `controlnet_model`, séparés par des virgules, comme "canny, softedge".

L'adaptabilité de ControlNet signifie qu'il peut fonctionner en parallèle avec des prompts textuels et des images IP-Adapter. Il est facile à utiliser pour les non-développeurs dans l'interface, tandis que les développeurs peuvent ajuster des paramètres supplémentaires, tels que :

- **controlnet_type** : Cela définit l'un des différents types acceptés de modèles ControlNet, comme canny, depth, hed, etc. Vous pouvez trouver la liste complète des types de modèles disponibles ici.
- **controlnet_model** : Cela spécifie le modèle ControlNet spécifique utilisé, qui peut être un modèle par défaut ou un modèle communautaire. Dans le cas d'un modèle par défaut, `controlnet_model` correspond directement à `controlnet_type`.
- **controlnet_conditioning_scale** : Cela détermine dans quelle mesure ControlNet influence le processus de réduction du bruit.
- **control_image (optionnel)** : Il s'agit de l'image à partir de laquelle les caractéristiques seront extraites. Si ce paramètre n'est pas spécifié et qu'une image initiale est fournie, cette image sera utilisée.

ControlNet sans prompt textuel et sans image IP-Adapter.

Image Input



Softedge



Image Output



Figure 4 : Transforming a bird picture with a softedge ControlNet (Transformation d'une image d'oiseau avec un softedge ControlNet)

Prompt Input

"superhero style, ironman, iron, light, dangerous"

Image Input



Canny

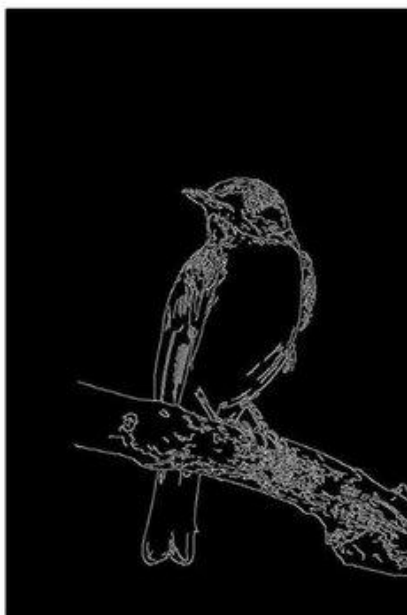


Image Output



Figure 5 : Transforming a bird picture with a prompt and a canny ControlNet (Transformer une photo d'oiseau grâce à un prompt et un canny ControlNet)

3.3 - Génération d'Image-à-Image

Enfin, au lieu d'orienter la génération d'image en conditionnant la sortie, pourquoi ne pas commencer la génération à partir d'une image qui partage des caractéristiques avec votre image finale ? C'est ainsi que fonctionne l'API Stable Diffusion img2img. Elle ne part pas du bruit aléatoire complet, mais ajoute du bruit à l'image initiale. Elle est conçue pour capturer les caractéristiques générales de l'image initiale, comme sa couleur et sa composition.

La génération Image-à-Image peut être combinée avec un prompt textuel, ControlNet et/ou IP-Adapter.

Parce que la génération img2img fonctionne de manière similaire à la génération texte-à-image, les deux API partagent presque les mêmes fonctionnalités. Les modèles communautaires et realtime-stable-diffusion peuvent être appliqués à l'un ou l'autre type de génération.

Génération d'image avec un prompt textuel et sans IP-Adapter

Prompt Input

"plane, flying, blue, sky, sun"

Image Input



Image Output



Figure 6 : A plane inspired from a bird picture (Un avion inspiré d'une photo d'oiseau)

Génération d'image sans commande textuelle et sans IP-Adapter

Image Input



Image Output



Figure 7 : Image d'un paysage de « fantasy » inspiré d'un dessin, source : <https://tinyurl.com/2eyf5ky5>

3.4 - Conclusion sur le Contrôle de la Génération d'Images

En conclusion, les générateurs d'images traditionnels comme Dall-E s'appuient uniquement sur des prompts textuels pour le contrôle, mais **Stable Diffusion** introduit des outils puissants tels que **IP-Adapter**, **ControlNet** et **Image-à-Image** pour diversifier et améliorer la génération d'images. Les utilisateurs peuvent utiliser ces outils indépendamment ou en combinaison, offrant une grande flexibilité et créativité pour générer des images, avec ou sans prompts textuels.

Sélection et entraînement de modèles spécifiques

Jusqu'à présent, nous avons parlé de l'ajout de conditions pour contrôler la génération d'images. Mais pourquoi ne pas modifier totalement le processus de génération ? Derrière la génération d'images se cache un modèle de diffusion. Cependant, ces modèles peuvent être modifiés pour générer des images plus spécifiques. Ici, je vais présenter deux façons d'améliorer le processus de génération : en sélectionnant ou en affinant un modèle de diffusion ou un **LoRA**. Ces deux approches peuvent être combinées avec les outils présentés ci-dessus.

Modèles de Diffusion

Comme mentionné précédemment, **Stable Diffusion** fonctionne avec un modèle de diffusion. Un modèle de diffusion est un modèle formé pour générer des images. Pour ce faire, ce modèle est entraîné sur un ensemble d'images qu'il cherche à reproduire. Par exemple, si vous voulez un modèle qui génère des images de chiens, vous devrez entraîner un modèle de diffusion en utilisant de nombreuses images de chiens.

La puissance de **Stable Diffusion** réside dans le fait qu'il permet un accès facile aux modèles affinés par la communauté. L'**affinage** est un processus de réentraînement partiel du modèle standard de **Stable Diffusion** sur votre propre ensemble d'images. Au lieu d'utiliser le modèle standard qui génère des images classiques, vous pouvez choisir un modèle qui génère des images de pixels, par exemple.

Image Input



Image output



Figure 8 : A pixelated fish using model Chibi Pixel Art Style (Un poisson pixelisé utilisant le modèle Chibi Pixel Art Style)

Ou pour générer des images de cartoons :

Image Input



Image Output



Figure 9 : Cartoon Image using model Cartoon Backgrounds (Image de bande dessinée utilisant un modèle Arrière-plans de bande dessinée)

La sélection de modèles est disponible à la fois pour les utilisateurs non-développeurs et via l'API développeur.

Enfin, si aucun modèle ne correspond à vos attentes, vous pouvez également affiner votre propre modèle. Selon la documentation, l'affinage nécessite seulement 7 à 8 images. Cependant, cela n'est disponible que via l'API développeur.

4 - LoRA

Enfin, **LoRA** est le dernier outil que je vais vous présenter pour avoir plus de contrôle sur la génération d'images. Les **LoRA** sont des versions compactes des modèles **Stable Diffusion**, généralement 10 à 100 fois plus petites. Les LoRAs ajoutent des modifications par-dessus le modèle de diffusion. Elles affinent les modèles standards, modifiant subtilement des styles spécifiques ou l'apparence générale des images générées. L'avantage des LoRAs réside dans leurs faibles besoins en calcul et leurs temps d'entraînement rapides.

Tout comme pour les modèles de diffusion, vous pouvez facilement utiliser des LoRAs affinées par la communauté. Mais, contrairement aux modèles de diffusion et de manière similaire à **ControlNet**, vous pouvez utiliser plusieurs LoRAs à la fois en les séparant par des virgules. La sélection de LoRAs est disponible tant pour les utilisateurs non-développeurs que via l'API développeur.

De plus, entraîner votre propre LoRA est simple à l'aide de l'API développeur, et cela nécessite typiquement seulement 7 à 8 images.

Vous trouverez deux nouveaux paramètres pour les LoRAs dans l'interface utilisateur et l'API développeur :

- **lora_model** : Spécifie quel modèle LoRA utiliser.
- **lora_strength** : Détermine l'ampleur de l'impact du LoRA pendant le processus de réduction du bruit.

Voici un exemple d'un LoRA intitulé "Princess" :

Image Input



Image Output

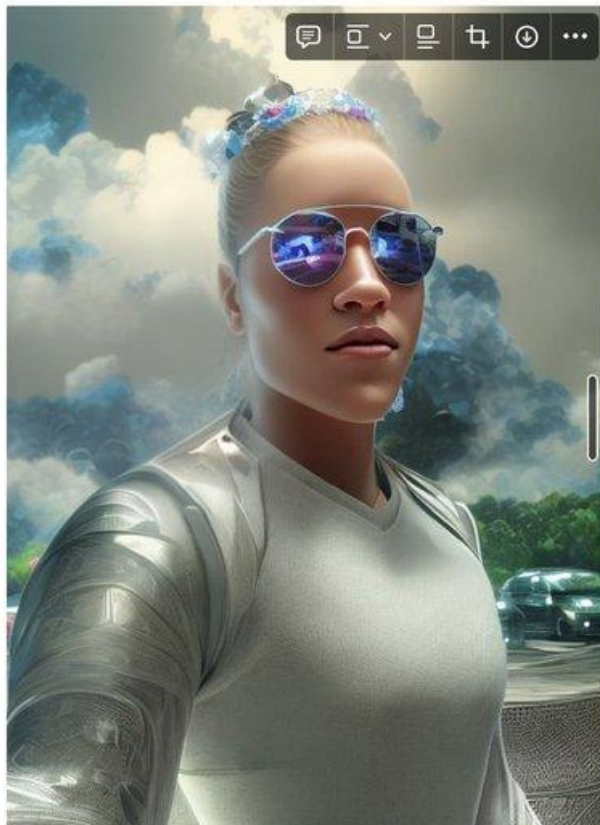


Figure 10 : LoRA : Princess

Et ici une illustration pour enfant :

Image Input



Image Output



Figure 11 : LoRA : Illustration pour enfant

Localisation en intérieur avec Ultra Wide Band

Rachida SAROUI^{1,2} - Anthony JUTON¹

Édité le
11/04/2025

¹Ecole Normale Supérieure Paris-Saclay, ² TU Graz

Cette ressource fait partie du N°115 de La Revue 3EI du deuxième trimestre 2025.

Cette ressource présente l'Ultra Wide Band (UWB ou ultra large bande) avec le principe de la localisation en intérieur, un exemple de mise en œuvre à l'aide de modules accessibles Qorvo DWM1001-DEV, ainsi que le problème lié à la présence d'obstacles.

1 - Introduction

1.1 - Problématique de la localisation en intérieur

Alors que la localisation en extérieur repose sur les satellites, le rapport signal/bruit (SNR) rend souvent difficile le décodage de ces signaux en l'intérieur. Cette limitation a stimulé la recherche sur la localisation en intérieur. Pour permettre un Internet des objets géolocalisés, des systèmes de positionnement en intérieur dédiés doivent compléter la technologie GNSS (Système Mondial de Navigation par Satellite), ouvrant ainsi la voie à une nouvelle dimension connue sous le nom de la Localisation des Objets.

1.2 - Applications

L'Ultra Wide Band (UWB) offre des applications variées et innovantes. Dans l'industrie automobile, elle permet des solutions comme la clé numérique, qui verrouille ou déverrouille un véhicule sans contact. Dans le domaine de la santé, UWB optimise le suivi des équipements médicaux et améliore la coordination des soins. Elle est également utilisée pour le suivi des matériaux et outils en fabrication, facilitant la gestion des stocks et augmentant l'efficacité de production [3]. Enfin, les dispositifs grand public, comme l'AirTag d'Apple [1], exploitent l'UWB pour une localisation précise des objets perdus.

Le tableau ci-dessous répertorie les entreprises, les téléphones et produits IoT (Internet of Object ou Internet des Objets) qu'elles ont lancés depuis 2020, ainsi que le module UWB intégré à ces appareils [2].

Company (products released between 2020 and 2024)	UWB Controller
Apple (Smartphones, Smartwatches, IoT devices (AirPods, HomePod, AirTag))	Apple U1 / U2
Google (Smartphones)	Qorvo DW5720
Samsung and Xiaomi (Smartphones)	NXP SR100T

2 - Principe de la localisation UWB

2.1 - Définition

Selon la définition de la Commission fédérale des communications (FCC) et du Secteur des radiocommunications de l'Union internationale des télécommunications (UIT-R), les signaux UWB ont une bande passante absolue $B \geq 500$ MHz ou une bande passante fractionnaire (relative) B_{frac} supérieure à 0,2 [6], où B_{frac} est définie comme :

$$B_{frac} = \frac{B}{f_c} = \frac{f_s - f_i}{f_c}$$

Où f_s et f_i désignent respectivement les fréquences supérieure et inférieure auxquelles la densité spectrale de puissance (PSD) est inférieure de 10 dB à son maximum, et f_c est la fréquence centrale.

Le concepteur de circuits intégrés UWB Spark microsystems propose sur son site [16] deux figures présentant l'occupation temporelle et spectrale des signaux :

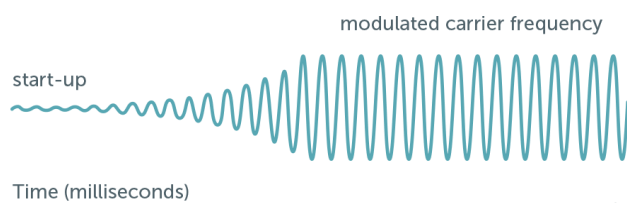


Figure 1 : Solution sans-fil narrow band (Wifi, BT, 4G...)



Figure 2 : Solution UWB source [16]

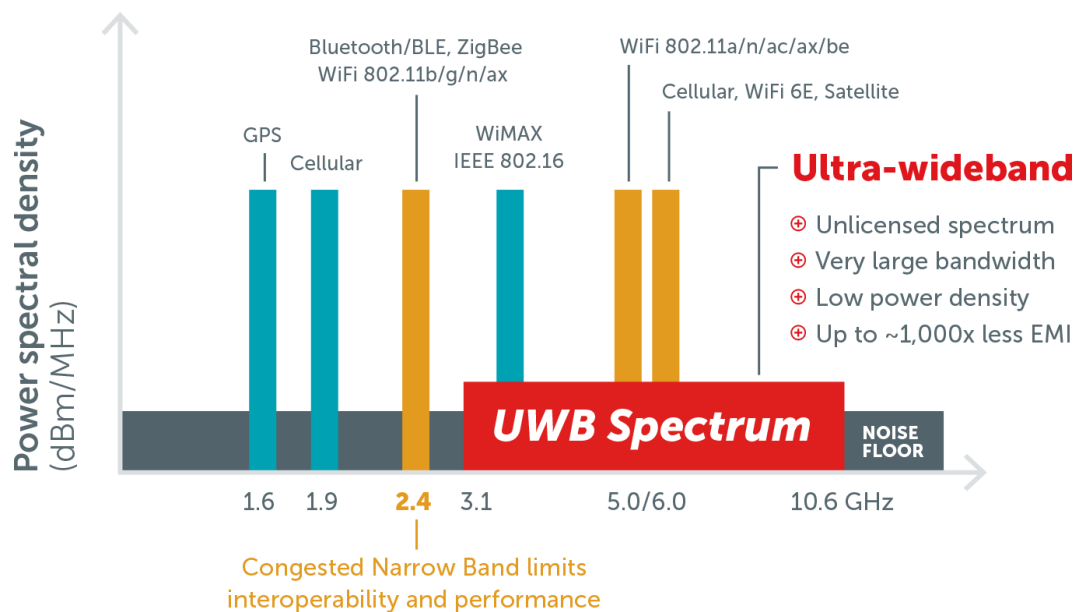


Figure 3 : Bande spectrale autorisée pour les signaux UWB

2.2 - L'Ultra Wide Band vs la bande étroite (narrowband)

L'exploitation d'une bande passante de plusieurs centaines de mégahertz correspond à une résolution temporelle de l'ordre de quelques nanosecondes. Dans le domaine spatial, cela permet de distinguer les composantes de trajets multiples (MPCs) provenant de surfaces réfléchissantes séparées par seulement quelques décimètres. Cela permet de dissocier le premier trajet du signal ou la composante de ligne de vue (LOS) des trajets multiples, ce qui permet de déterminer avec précision et exactitude son temps d'arrivée (ToA) et, par conséquent, le temps de vol du signal et

donc la distance à la source. À l'inverse, lorsque des systèmes à bande étroite sont utilisés, les MPCs interfèrent de manière constructive et destructive avec la composante de ligne de vue (LOS), ce qui introduit de l'affaiblissement et des distorsions d'impulsions, augmentant ainsi l'incertitude de l'estimation du ToA.

Cependant, cela peut potentiellement causer des interférences avec d'autres systèmes radio à proximité. C'est pourquoi des règles complexes régissent l'utilisation des fréquences, les niveaux de puissance autorisés et le rythme des transmissions.

Frequency Range (MHz)	960-1610	1610-1990	1990-3100	3100-10600	Above 10600	1164-1240 1559-1610
Ground penetrating radar, wall imaging	-65.3	-53.3	-51.3	-41.3	-51.3	-75.3
Through D-wall imaging systems:						
UWB bandwidth below 960 MHz	-65.3	-53.3	-51.3	-51.3	-51.3	-75.3
UWB bandwidth 1990-10600 MHz	-46.3	-41.3	-41.3	-41.3	-51.3	-56.3
Surveillance systems	-53.3	-51.3	-41.3	-41.3	-51.3	-63.3
Medical imaging systems	-65.3	-53.3	-51.3	-41.3	-51.3	-75.3/-53.3
Indoor UWB systems	-75.3	-53.3	-51.3	-41.3	-51.3	-85.3
Hand held UWB systems	-75.3	-63.3	-61.3	-41.3	-61.3	-85.3
Frequency Range (MHz)		1610-22000	22000-29000	29000-31000	Above 31000	
Vehicular radar systems	-75.3	-61.3	-41.3	-51.3	-61.3	-85.3
Emissions at 1164-1240 and 1559-1610 MHz to be measured using a resolution bandwidth no less than 1 kHz. All other frequencies to be measured using a resolution bandwidth of 1 MHz.						

Figure 4 : Limites d'émission (EIRP en dBm) pour les différents types de systèmes UWB. En dessous de 960 MHz, tous les systèmes doivent se conformer aux limites d'émission existantes qui s'appliquent à tous les émetteurs non intentionnels (Partie 15.209) [9]

2.3 - Modèle OSI

Le modèle OSI ci-dessous montre que des efforts de normalisation sont faits dans le but de garantir l'interopérabilité des composants, notamment pour la partie communication. Pour la localisation, on reste essentiellement sur des solutions propriétaires.

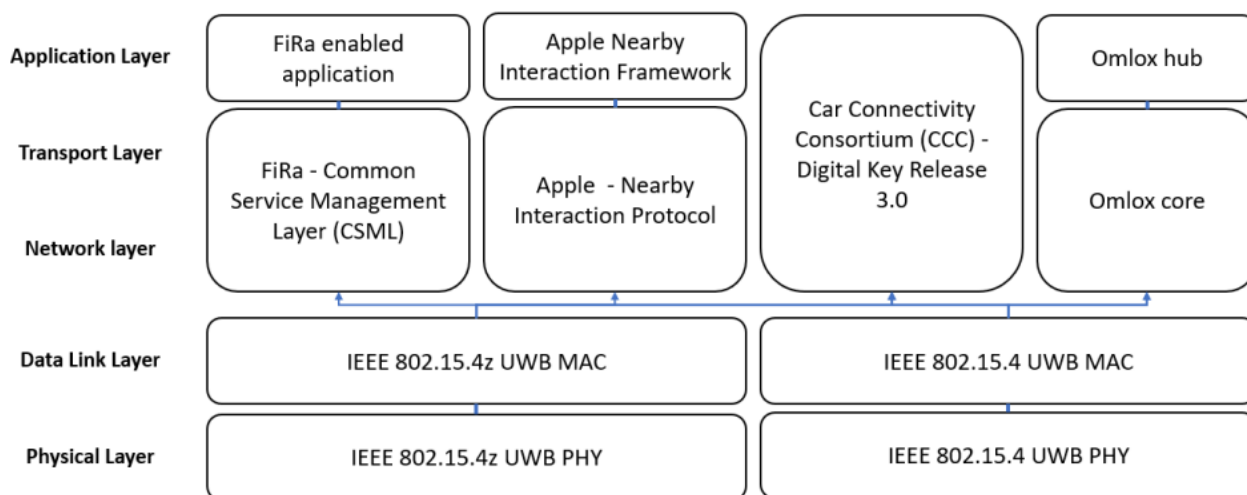


Figure 5 : Les normes UWB et leur intégration dans le modèle OSI [7]

2.4 - Localisation UWB

L'Ultra Wide Band (UWB) est conçue pour l'échange de données à faible consommation d'énergie et la mesure sécurisée de distance. L'UWB peut être utilisée pour construire un système de localisation en temps réel (RTL) tout en promettant une précision au centimètre près. À cette fin, il est

nécessaire d'utiliser au moins trois nœuds fixes pour estimer la position spatiale d'un quatrième nœud (algorithme de trilatération).

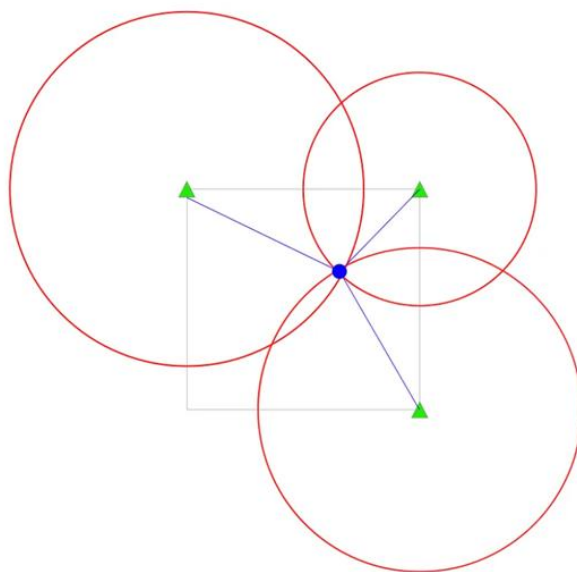


Figure 6 : Illustration de la trilatération [13]

2.5 - Mesure de Distance Bidirectionnelle

La figure 7 illustre le principe de la mesure de distance entre le tag et l'ancre ou balise. Il s'agit du procédé de "two-way ranging" [10]. Le tag envoie un message à la balise (anchor) qui lui renvoie un message contenant le temps entre la réception du signal tag et l'envoi de son signal balise Treply. Le tag, en mesurant la différence entre la date de réception du message balise et la date d'envoi du signal tag, obtient le double du temps de trajet entre le tag et la balise. Cette valeur peut bien sûr être transmise à la balise.

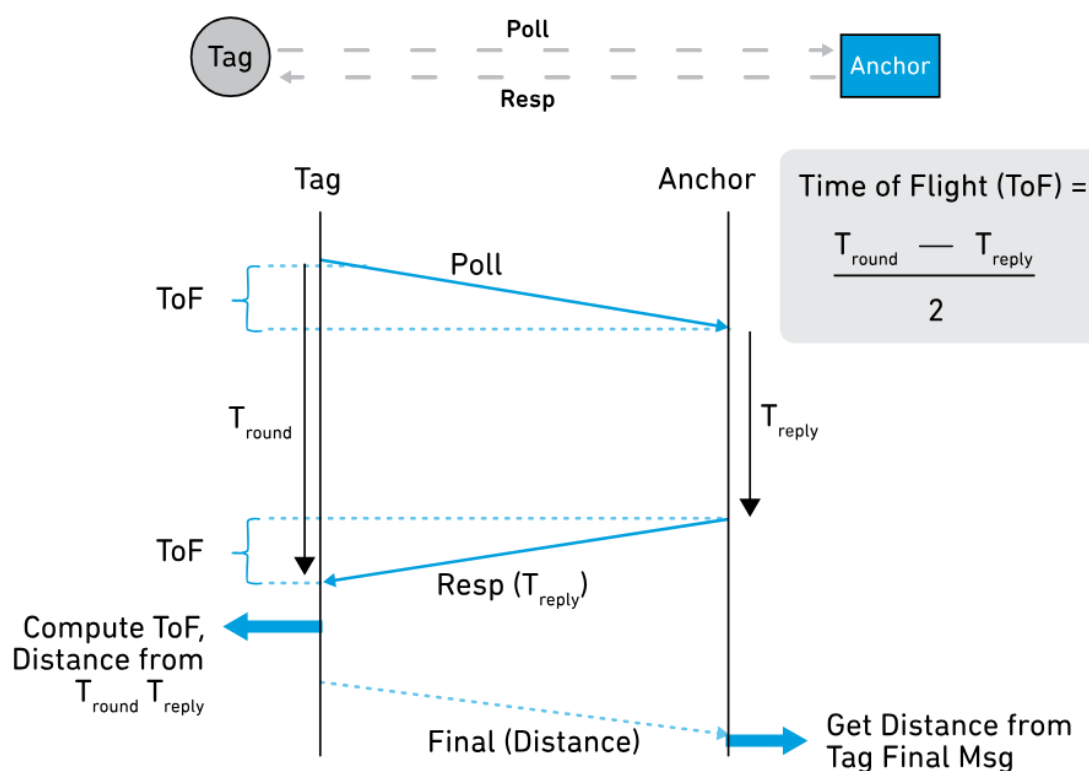


Figure 7 : Principe de la mesure de distance entre le tag et l'ancre - Procédé « two-way ranging »

3 - Un exemple de mise en œuvre à l'aide des kits DWM1001-DEV

La mise en œuvre présentée ici a mené à la mesure de précision de localisation d'un kit mobile (le tag) dans une pièce équipée de 4 autres kits fixes (les ancres).

Il est aussi possible, avec deux kits, de simplement effectuer du *ranging* entre deux cartes et observer la précision.

3.1 - Présentation du module DWM1001

- Précision de localisation jusqu'à 10 cm
- Antenne PCB imprimée UWB Canal 5 (6,5 GHz)
- Débit de données de 6,8 Mbps, conforme à IEEE 802.15.4-2011 UWB
- Connectivité Bluetooth® avec puce Nordic Semiconductor nRF52832
- Antenne Bluetooth® intégrée
- Capteur de mouvement : accéléromètre 3 axes
- Consommation optimisée pour le mode veille à faible consommation : <15µA
- Tension d'alimentation : 2,8 V à 3,6 V
- Interfaces SPI, UART, I2C et Bluetooth® pour connecter le DWM1001 à un appareil hôte
- Firmware intégré prêt à l'emploi pour des applications RTLS (Two Way Ranging)

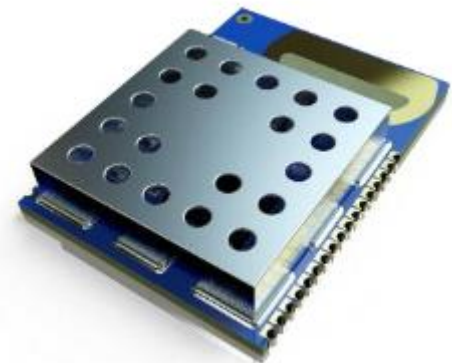


Figure 8 : Module DWM1001

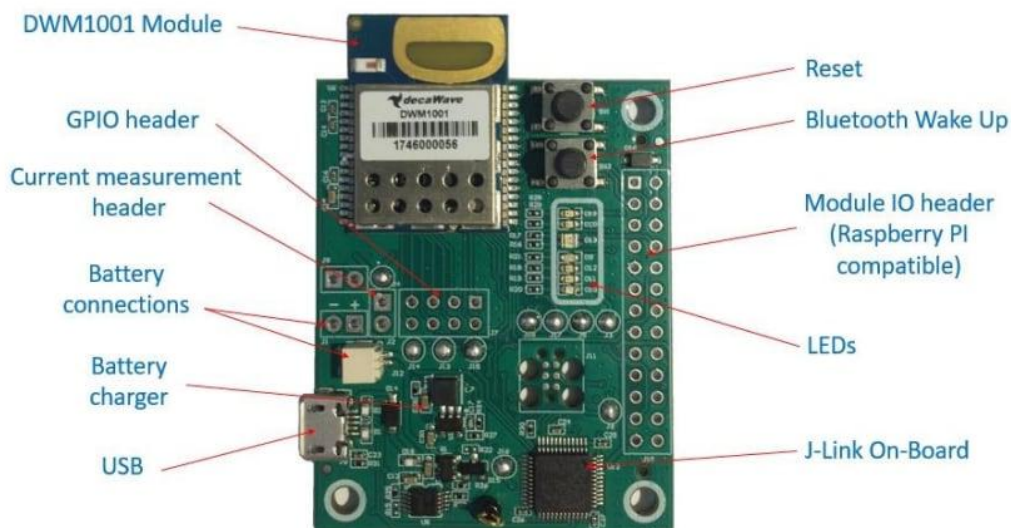


Figure 9 : Kit Qorvo DWM1001-dev source Qorvo

3.2 - Installation

Télécharger le package logiciel Decawave

Ce package comprend la documentation et le logiciel nécessaire pour la carte de développement DWM1001. Il contient un grand nombre de fichiers, et doit être téléchargé pour configurer correctement les modules UWB.

Flasher les capteurs UWB avec J-Flash Lite

J-Flash Lite est une interface graphique simple et gratuite qui permet de télécharger le micrologiciel embarqué (firmware) dans la mémoire flash des modules UWB.

Après l'installation, suivre les étapes suivantes pour flasher les quatre modules :

1. Connecter le module à son ordinateur à l'aide d'un câble de données micro USB.
2. Flasher l'image du module DWM1001 :
 - a. Ouvrir J-Flash Lite.
 - b. Choisir nrf52832_XXAA comme périphérique et SWD comme interface, utiliser la vitesse par défaut de 1000. Cliquer sur "OK".
3. Cliquer sur "Erase Chip" pour effectuer un effacement complet de la puce.
4. Dans "Data File", cliquer pour parcourir le package logiciel Decawave téléchargé (/DWM1001/Factory_Firmware_Image/DWM1001_PANS_R2.0.hex) à flasher, puis cliquer sur "Program Device".

3.3 - Configuration via BLE

La manière la plus simple de commencer à utiliser les modules UWB est via l'application Android. Ouvrir le lien avec un smartphone, télécharger le SDK Android et l'installer.

Une fois l'application correctement installée, alimenter chaque module en USB (il est recommandé d'utiliser une batterie externe). Si les LEDs des modules clignotent en vert et rouge, il est possible d'établir la connexion en Bluetooth. Rechercher les modules avec l'application Decawave. Une fois les quatre modules détectés, ils doivent être connectés au même réseau que celui créé auparavant (dans les illustrations ci-dessous, le réseau est nommé RTLS Demo).

Il est alors possible de définir trois ancres et un tag :

- Sélectionner les modules dans le menu. Le PANID (identifiant de réseau sur 16 bits) doit être défini par l'utilisateur. Un groupe de nœuds formant un réseau UWB doit avoir le même PANID. L'ancre initiatrice initialise et contrôle le réseau.
- Définir les coordonnées de chaque ancre : X Y Z

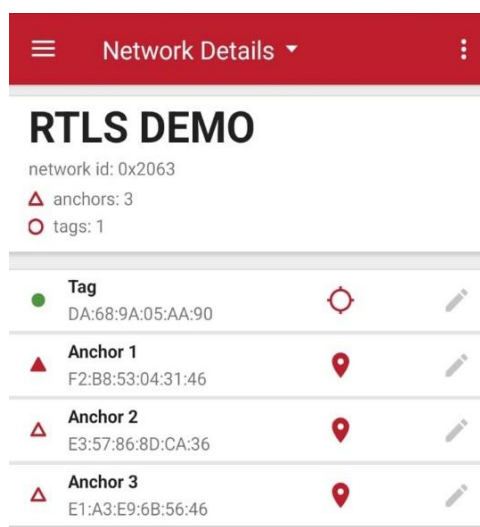


Figure 10 : Aperçu de la page de configuration de l'application Android Decawave



Figure 11 : Page de localisation de l'application Decawave

3.4 - Configuration via UART

Il est aussi possible de configurer les modules sans passer par l'application Android, en utilisant la liaison série UART :

- | | |
|---|---|
| <ul style="list-style-type: none"> • Anchor : <ol style="list-style-type: none"> 1. nis 0x1234: PAN ID 2. aps X Y Z 3. nma / nmi si initiator | <ul style="list-style-type: none"> • tag <ol style="list-style-type: none"> 1.nis 0x1234: PAN ID 2.nmt |
|---|---|

```
DWM1001 TWR Real Time Location System
Copyright : 2016-2019 LEAPS and Decawave
License : Please visit https://decawave.com/dwm1001\_license
Compiled : Mar 27 2019 03:35:59

Help : ? or help

dwm> nis 0x1234
nis: ok
dwm> aps 0 0 100
aps: ok
dwm> nmi

DWM1001 TWR Real Time Location System
Copyright : 2016-2019 LEAPS and Decawave
License : Please visit https://decawave.com/dwm1001\_license
Compiled : Mar 27 2019 03:35:59

Help : ? or help

dwm>
```

Figure 12 : Configuration d'une ancre via la liaison série

Lorsqu'un nœud est configuré en tant que **anchor** (balise fixe), il renvoie via l'**UART** les distances aux trois ancrs ainsi que leurs identifiants respectifs.

Quand un nœud est configuré en tant que **tag**, il renvoie via l'**UART** les données suivantes concernant les balises environnantes dans le même réseau :

- Identifiant du nœud (ID MAC).
- Coordonnées de position prédéfinies en X, Y, Z en millimètres.
- RSSI (Indicateur de force du signal radio).
- Numéro dans le réseau et réseau voisin.

3.5 - Récupération des données par un nano-ordinateur Raspberry Pi

Pour transmettre les informations de localisation au nano-ordinateur (nommé directement Raspberry Pi par la suite), il est possible d'utiliser un module série-USB ou de connecter directement le kit DWM1001-DEV sur les broches UART du connecteur GPIO de la raspberry pi.

Il est possible d'utiliser le code disponible sur le dépôt de Patrick Hallek :

<https://gist.github.com/PatrickHallek/3f800d0d39e740a3f955733a549011c3#file-dwm1001-localization-py>

Ce code fait l'importation des bibliothèques JSON et Redis, ce qui permet de stocker les données reçues du module dans une base de données Redis. Si tout fonctionne comme prévu, la chaîne suivante est renvoyée :

```
DIST,3,AN0,8182,2.18,0.00,0.00,0.83,AN1,0F8C,2.12,1.04,0.00,0.18,AN2,5C2F,0.00,0.00,0.00,0.59,POS,0.83,0.52,0.00
```

Les données des ancrs suivent cette structure :

DIST,Nombre d'ancres,[AN+Numéro de l'ancre,ID de l'ancre,Position X,Position Y,Position Z,Distance au module],POS,Position X du module,Position Y du module,Position Z du module

Les informations entre crochets sont répétées pour chaque ancre.

Le script Python divise cette chaîne en plusieurs éléments, puis stocke les informations dans un tableau. Les données des ancrs et la position du module sont ensuite sauvegardées séparément dans la base de données.

Les données de position sont maintenant stockées en temps réel sur la Raspberry Pi.

3.6 - Pour aller plus loin

Pour approfondir et comprendre les différents types de ranging (double-sided, single-sided two-way ranging), ainsi que pour explorer toutes les informations disponibles, comme les diagnostics RX, une liste d'exemples à tester est fournie. Il est recommandé d'exécuter ces exemples sur Segger Embedded Studio ou, sinon, de créer un makefile adapté à son environnement.

Dans [11], se trouvent toutes les fonctions de l'API du DW1001 utilisables pour ses projets. Les exemples sont également expliqués dans le document.

4 - Enjeu de la connaissance de la présence d'un obstacle

4.1 - Problématique

Les environnements intérieurs posent souvent des défis aux systèmes de positionnement par radiofréquence (RF). Les objets présents peuvent perturber la propagation du signal en absorbant, réfléchissant ou dispersant les ondes, ce qui crée des erreurs dans l'estimation du temps d'arrivée (TOA) et donc de la position. Récemment, des approches classiques et basées sur l'apprentissage automatique (ML) ont permis de détecter ces perturbations en analysant les réponses impulsionnelles des canaux (CIR) des systèmes de positionnement Ultra Wide Band.

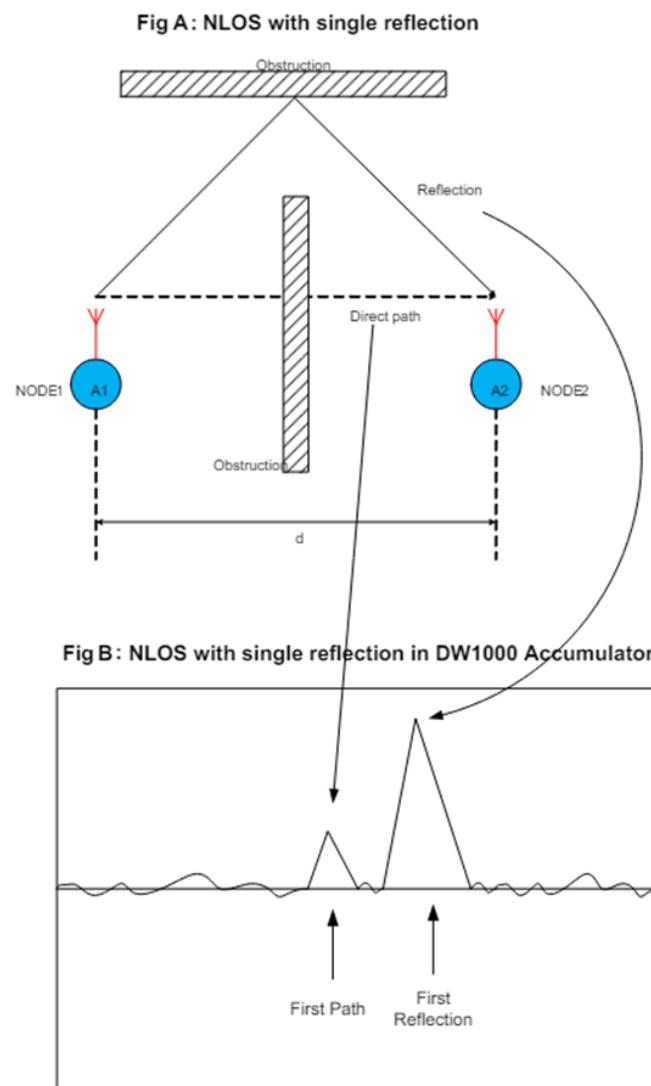


Figure 13 : La figure A montre le schéma d'un émetteur et d'un récepteur où la ligne de vue directe est obstruée. Une réflexion est également représentée.

La figure B montre les échantillons accumulés dans l'accumulateur du DW1000. Le premier chemin correspond à la voie directe, tandis que la première réflexion correspond à la réflexion illustrée dans la figure A.

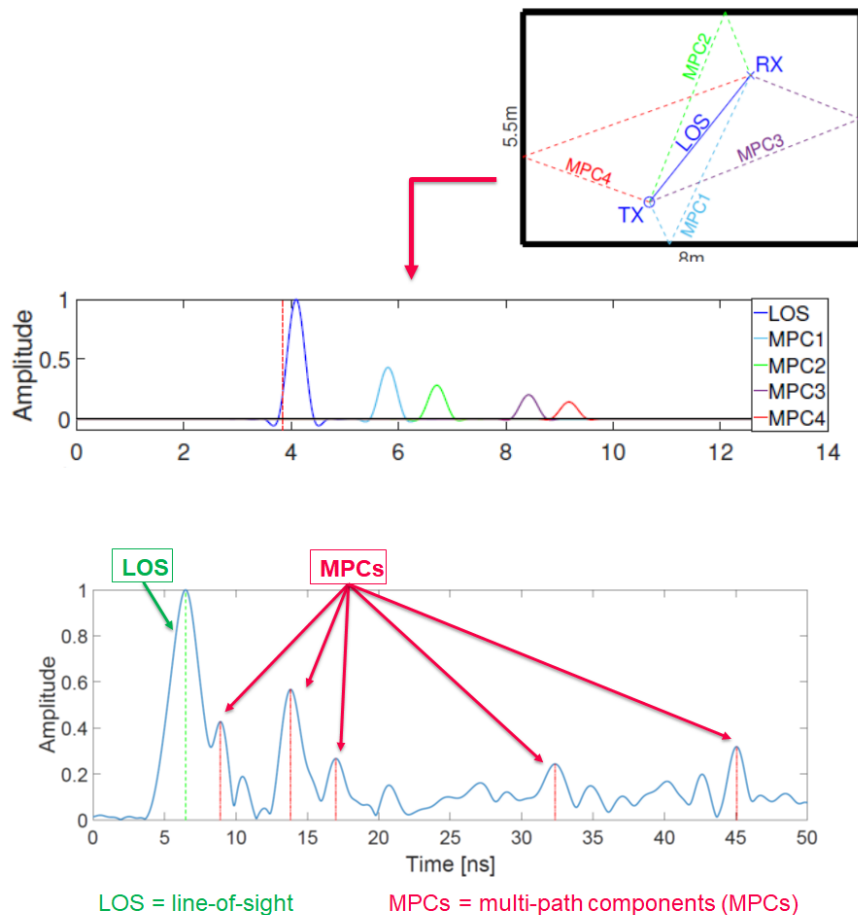


Figure 14 : Amplitudes des signaux reçus lors d'une connexion multitrajet

4.2 - L'utilisation de l'apprentissage automatique pour la classification LOS/NLOS et la correction d'erreurs

Les CIR peuvent être utilisés directement comme entrées pour des modèles d'apprentissage automatique pour la classification LOS/NLOS et la correction d'erreurs. Cependant, beaucoup d'approches se concentrent surtout sur les performances de classification et la correction des erreurs, sans se soucier des préoccupations pratiques comme la mémoire ou le temps d'exécution. De nombreux modèles, surtout les réseaux de neurones convolutifs (CNN), demandent beaucoup de puissance de calcul. C'est pourquoi certains chercheurs recommandent de les déployer sur des dispositifs edge, comme des Raspberry Pi avec des CPU/GPUs puissants et suffisamment de mémoire [14].

Dans [15], les auteurs ont développé un framework InSight qui facilite le déploiement des méthodes de machine learning sur des dispositifs Ultra Wide Band intégrés, tout en faisant un compromis entre la précision et les contraintes de ressources, comme un temps d'exécution minimal pour des taux de mises à jour rapides ou une empreinte mémoire réduite pour les dispositifs à mémoire limitée. L'intégration avec le système existant est fluide. Le générateur de modèles prend en charge plusieurs algorithmes de machine learning, notamment les machines à vecteurs de support (SVM), les arbres XGBoost et les réseaux de neurones. Le moteur NLOS intégré est responsable de la classification et de la correction d'erreurs sur le dispositif, conçu pour s'intégrer facilement à diverses applications.

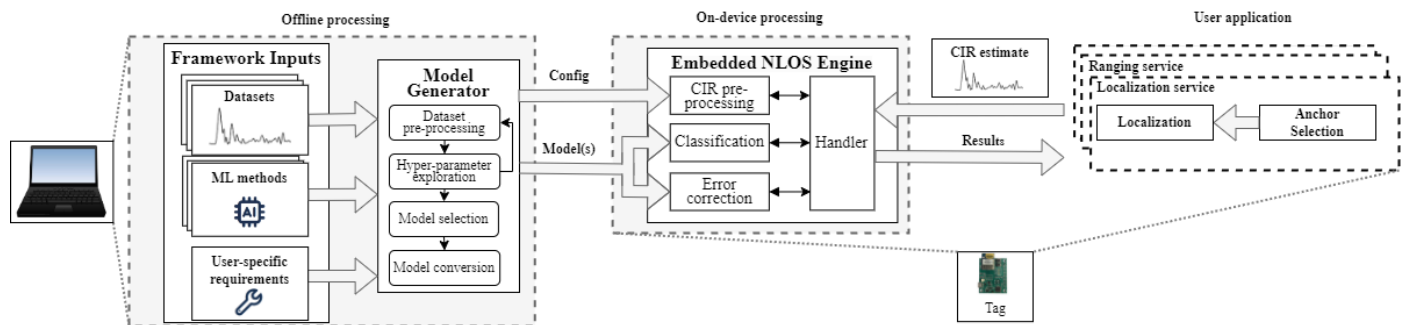


Figure 15 : Présentation de la framework InSight, comprenant le générateur de modèles hors ligne et le moteur NLOS intégré sur l'appareil. Le générateur de modèles hors ligne choisit le modèle le plus approprié en fonction de la configuration d'entrée, tandis que le moteur NLOS intégré effectue la classification et/ou la correction d'erreurs en utilisant les données de réponse impulsionnelle du canal (CIR) fournies par un service de localisation ou de mesure de distance [15].

5 - Présentation du module DWM3000 et comparaison avec DWM1001

- Conformes au BPRF IEEE 802.15.4-2015 et IEEE 802.15.4z
- Entièrement alignés avec les spécifications FiRa™,
- PHY, MAC et le développement de la certification
- Prend en charge les canaux 5 (6,5 GHz) et 9 (8 GHz)
- SoC nRF52833 Nordic avec émetteur-récepteur BLE
- Tension d'alimentation VCC de 2,5 V à 3,6 V
- Antenne puce BLUETOOTH®
- Bloc de sécurité AES 128/256



Figure 16 : Module Qorvo DWM3000

Caractéristiques nouvelles :

- Angle d'arrivée (QM33120W)
- STS (Scrambled Timestamp Sequence) : une séquence d'impulsions pseudo-aléatoires générées à l'aide d'un générateur de bits aléatoires déterministe (DRBG) et incluse dans le paquet UWB)

Comparaison de Performances :

En conditions LOS, le DW3000 est plus précis avec une réduction de l'écart type de 5,5 %, et plus fiable, avec une diminution de 71 % du RMSE. En moyenne, il garantit une précision de distance constamment inférieure à 10 cm en conditions LOS et avec une orientation optimale de l'antenne. Le DW3000 présente également une consommation d'énergie réduite par un facteur de 2× par rapport à son prédécesseur, et un pic de courant diminué de 45 % un atout majeur pour tout dispositif IoT à faible consommation [12].

Références :

- [1] Apple, "AirTag", Accessed: July 26, 2023. Available : <https://www.apple.com/airtag/>
- [2] https://en.wikipedia.org/wiki/List_of_UWB-enabled_mobile_devices
- [3] [PhD Thesis \(grosswindhager.com\)](#)
- [6] International Telecommunication Union Radiocommunication Sector (ITU-R), "Recommendation ITU-R SM.1755: Characteristics of UWB Technology," 2006 (International Telecommunication Union Radiocommunication Sector (ITU-R))
- [7] Coppens, D., Shahid, A., Lemey, S., Van Herbruggen, B., Marshall, C., & De Poorter, E. (2023). "An Overview of UWB Standards and Organizations (IEEE 802.15.4, FiRa, Apple): Interoperability Aspects and Future Research Directions".
- [9] A Summary of FCC Rules for Ultra Wideband Communications By Gary Breed, High Frequency Design.
- [10] Ultra-Wideband For Dummies®, Qorvo Special Edition.
- [11] DW1000 DEVICE DRIVER APPLICATION PROGRAMMING INTERFACE (API) GUIDE 2016
- [12] Polonelli, T., Schläpfer, S., & Magno, M. "Performance Comparison between Decawave DW1000 and DW3000 in Low-Power Double Side Ranging Applications." D-ITET, ETH Zurich.
- [13] <https://medium.com/@newforestberlin/precise-realtime-indoor-localization-with-raspberry-pi-and-ultra-wideband-technology-decawave-191e4e2daa8c>
- [14] Stefano MARANO et al. « NLOS Identification and Mitigation for Localization Based on UWB Experimental Data ». In : IEEE Journal on Selected Areas in Communications 28.7 (sept. 2010). Stefano Marano and Wesley M. Gifford are Student Members, IEEE ; Henk Wymeersch is a Member, IEEE ; Moe Z. Win is a Fellow, IEEE, p. 1157-1166.
- [15] Markus Gallacher et al. "InSight: Enabling NLOS Classification, Error Correction, and Anchor Selection on Resource-Constrained UWB Devices." EWSN 2023, Rende, Italy, ACM.
- [16] Le site de Spark Microsystems : <https://www.sparkmicro.com/ultra-wideband-technology/>

Conception de systèmes radiofréquences. Entretien avec la PME française « Always Wireless »

Joss BRUNET¹, Pascal CHAMPANEY¹, Hervé DISCOURS²

Édité le
10/03/2025

école _____
normale _____
supérieure _____
paris-saclay _____

¹ Société « Always Wireless », ² IUT Cachan

Cette ressource fait partie du N° 115 de La Revue 3EI du deuxième trimestre 2025.

Cet article est un entretien avec J. Brunet, et P. Champaney qui développent des sous-ensembles dans le domaine des communications radioélectriques sans fil. Ils nous font part de leurs activités de projet depuis la conception, la réalisation, jusqu'à la qualification d'un produit.

Avec plus de 30 ans d'expérience, la Société « Always Wireless » est spécialisée dans les produits radio-électroniques grand public ou professionnel, principalement dédiés aux objets connectés d'aujourd'hui.

Située en région grenobloise, et composée uniquement de deux personnes, Always Wireless traite en moyenne 40 à 50 projets par an. Cela va des systèmes de communication pour arbitres de football UEFA « VOGO », en passant par des systèmes de localisation GNSS des vélos électriques DECATHLON, ou des gestionnaires de luminaires urbains LACROIX et BH TECHNOLOGIES, ou encore les antivols des vélos DELIVEROO et LA POSTE...

1 - Métier et domaine d'ingénierie

La Revue 3EI : Pouvez-vous décrire votre métier en quelques mots ?

J. Brunet et P. Champaney : Notre métier consiste à étudier et mettre au point les sous-ensembles nécessaires à établir des communications sans fil par voie radioélectrique entre des produits communicants : émetteurs, récepteurs et antennes radio (WiFi, Bluetooth, 4/5G et autres...). Les grandeurs électroniques manipulées sont très proches de la physique électronique basique, avec les limites associées. C'est un domaine peu représenté en termes de compétences, dont l'enseignement est limité à une théorie assez lourde et indigeste, alors qu'il requiert énormément d'expérience pratique.

La Revue 3EI : Quel est votre domaine d'ingénierie et en quoi consiste-t-il concrètement ?

J. Brunet et P. Champaney : Nous assurons les Conception, Développement et Qualification des chaînes radio issues de nos réflexions : Choix des composants radio stratégiques, proposition des éléments d'antennes, compléments et suivi des schématiques et CAO PCB, optimisation des prototypes et qualification fonctionnelle et pré-réglementaire.

De façon générale, nous travaillons sur des équipements tels que : modulateurs & démodulateurs analogiques, numériques ou logiciels, amplificateurs faible bruit pour la réception, amplificateurs de puissance pour l'émission, et antennes afin de convertir de manière bidirectionnelle, les signaux électriques en ondes électromagnétiques.

La Revue 3EI : Quels sont les projets ou missions les plus courants dans votre activité ?

J. Brunet et P. Champaney : Le client nous fait part de son besoin en termes de communication sans fil : nature et volume des données transmises, portées souhaitées, milieu d'exploitation des équipements, territoire géographique de l'usage des produits. A partir ces données nous proposons des solutions techniques et nous en assurons les intégrations et optimisations.

On peut assurer la conception de « front-end » RF en bandes 433 et 868/915 MHz, ainsi que la conception d'antennes intégrées. L'ajout d'un « front-end » RF permet d'augmenter significativement les performances et de répondre à des besoins clients spécifiques, initialement non couvert par les circuits intégrés traditionnels, disponibles sur le marché.

La Revue 3EI : Avez-vous des interactions fréquentes avec d'autres corps de métier (designers, techniciens, commerciaux, etc.) ?

J. Brunet et P. Champaney : Comme nous apportons seulement un complément à un produit électronique, nous sommes obligatoirement en échange permanent avec le bureau d'étude électronique classique du client. Les fonctions sont aujourd'hui très imbriquées et interdépendantes, tant en termes de matériel que des logiciels. Il nous faut également travailler avec les designers mécaniques car l'intégration des antennes apporte un lot de contraintes dont le respect est essentiel pour assurer un rayonnement radio le plus efficace possible.

La Revue 3EI : Avez-vous des interactions avec des laboratoires de recherche ?

J. Brunet et P. Champaney : Très peu. Nous travaillons une électronique très appliquée avec des fortes contraintes technico-économiques, donc avec peu de place pour les risques de l'innovation de pointe. De surcroît, nos quelques expériences passées avec des laboratoires de recherche ont été assez décevantes du fait d'objectifs et fonctionnements trop distincts.

La Revue 3EI : Quelle formation est nécessaire pour votre métier ? Quelles sont les compétences indispensables ?

J. Brunet et P. Champaney : Il faut une formation de technicien supérieur ou ingénieur en électronique, avec, dans la mesure du possible, une spécialisation vers les domaines de l'analogique ou des télécommunications. Cependant, et très au-delà de la formation, il faut une sensibilité à la physique électromagnétique (sans forcément s'abrutir avec sa théorie), être capable de « ressentir » les phénomènes et bien en comprendre les mécanismes et limites. Nous sommes à des lieux des numérisations et virtualisations actuellement très à la mode !

2 - L'activité de projet

La Revue 3EI : Quelles sont les principales étapes lors de la conception d'un produit RF ?

J. Brunet et P. Champaney : Définir correctement le besoin exact : Nature et volume des données transmises, portées souhaitées, milieu d'exploitation des équipements, territoire géographique de l'usage des produits.

- Choix des composants radios stratégiques, que ce soient des circuits intégrés ou des modules et leurs composants périphériques (quartz, filtres, amplis transistorisés, réseaux d'adaptation d'impédance...)
- Proposition des éléments d'antennes, selon les nécessités et possibilités d'intégration mécanique et les objectifs économiques.

- Fournir les compléments schématiques du sous-ensemble radio et assurer sans bonne intégration électrique au produit complet : cohérence des signaux et alimentations.
- Suivre la CAO des circuits imprimés pour s'assurer du respect des règles essentielles de la radio électricité : chemin courts, respect des équipotentiels, réduction des pollutions électriques.
- Optimisation des prototypes afin de trouver les réglages et les valeurs de composants offrant des performances maximales.
- Qualification fonctionnelle et pré-réglementaire pour obtenir des liens fiables, opérant aux distances attendues et conformes aux législations en vigueur.

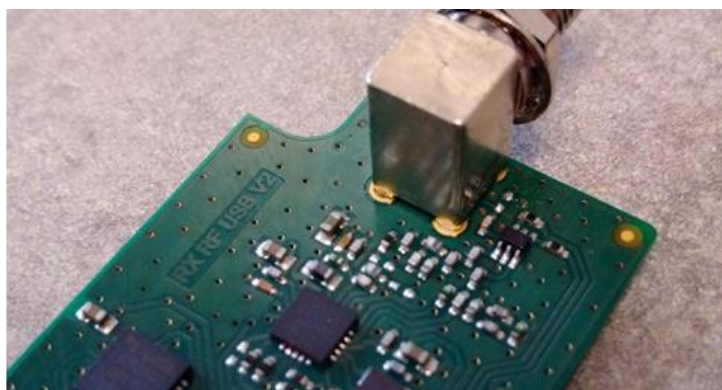
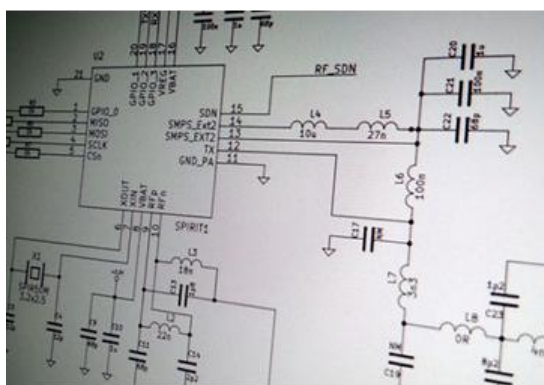


Figure 1 : De la conception à la réalisation

La Revue 3EI : Quelles sont les principales gammes de fréquence avec lesquelles vous travaillez ?

J. Brunet et P. Champaney : Nous travaillons de 30 kHz à 8 GHz. Au-delà, les activités changent considérablement vers les faisceaux télécoms et les radars, domaines que nous n'adressons pas. Nous exerçons néanmoins majoritairement dans les bandes sans licence d'exploitation : 125 kHz, 13.56 MHz, 169 MHz, 433 MHz, 868 MHz, 915MHz (US-CAN-AUS) 2450 MHz, 5300 MHz pour les produits monobandes et de 650 à 3800 MHz pour le cellulaire 4/5G.

La Revue 3EI : Quels sont les principaux protocoles de communication que vous utilisez actuellement ?

J. Brunet et P. Champaney : Il y a clairement un repositionnement du marché vers les standards très classiques, type Bluetooth, WiFi, LoRaWAN, 802.15.4 et consorts et les 4/5G et leur « descendance » LTE-M et NB-IoT. Aujourd'hui développer son propre protocole coûte cher en temps de design logiciel alors qu'il existe pléthore de solutions où le protocole est embarqué. Seuls les domaines très sécuritaires tels que les systèmes d'alarmes et contrôles d'accès persistent à exploiter des protocoles propriétaire, mais de moins en moins.

La Revue 3EI : Vous parlez parfois de « Always Design to cost », de quoi s'agit-il ?

J. Brunet et P. Champaney : Nos solutions techniques et designs sont toujours orientés afin de présenter un coût minimal : choix de composants peu coûteux et très standards, composants discrets proposés en multisources, solution d'antennes composants céramiques « low cost » ou 100% imprimée. Toute « subtilité » est toujours examinée avec un œil économique critique pour éviter les exotismes.

La Revue 3EI : Quels sont les principaux logiciels que vous utilisez ?

J. Brunet et P. Champaney : Nous utilisons beaucoup de logiciels freewares type Simsmith, AppCAD et RFSim qui ne nécessitent pas de « permis de conduire » pour être exploités efficacement. Notre schématique est produite sous KiCAD, gratuit également. Nos seules licences sont sur Sonnet Suite, un simulateur électromagnétique pour les antennes 2D et Keysight VEE pour l'automatisation des mesures pré-réglementaires.

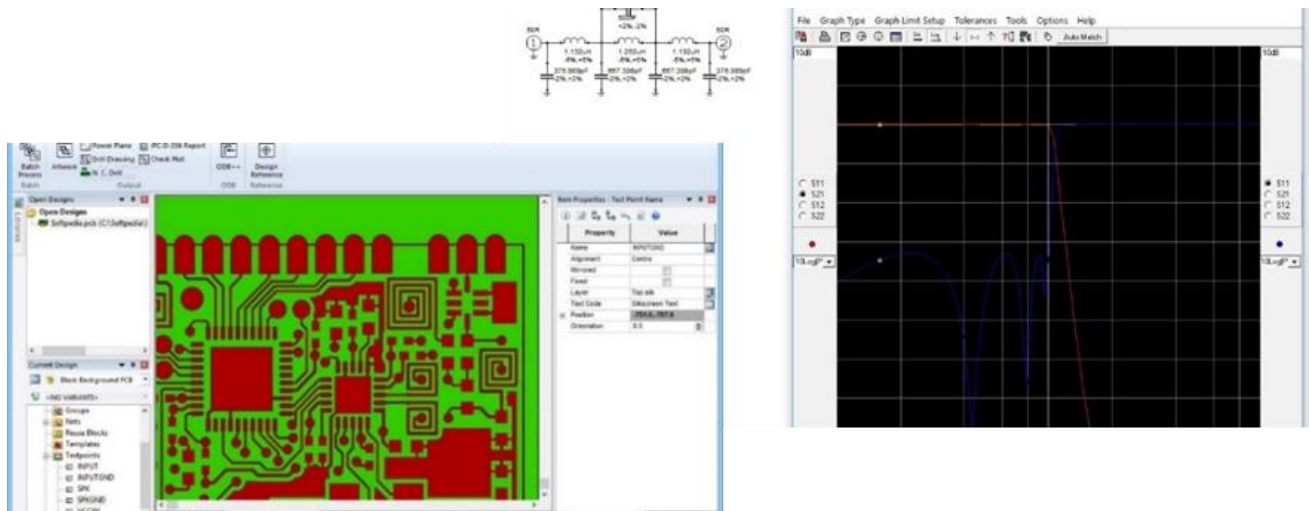


Figure 2 : Logiciels CAO

La Revue 3EI : Avez-vous quelques exemples de réalisation qui illustrent votre savoir-faire ?

J. Brunet et P. Champaney : Nous avons conçus les sous-ensembles radio-électroniques des équipements « connus » suivant :

- Systèmes de communications de arbitres de football UEFA et « world rugby » VOGO
- Traceurs de localisation GNSS des vélos électriques DECATHLON
- Antivol des vélos DELIVEROO et LA POSTE
- Supervision des voies ferroviaires SNCF
- Gestionnaires de luminaires urbains LACROIX et BH TECHNOLOGIES
- Répéteurs de radiocommunications en tunnel TDF et SEE
- Gestionnaire de chauffage domestique connecté TADO
- Calculateurs de véhicules professionnels ACTIA

... et bien d'autres car Always Wireless traite en moyenne 40 à 50 sujets / an !



Figure 3 : Mise au point de prototypes

La Revue 3EI : Dans le développement de sous-ensembles intégrés et des réseaux passifs associés (Filtres, PA, LNA) quelles sont les principales normes à respecter ?

J. Brunet et P. Champaney : Les normes radio les plus classiques dont les points critiques sont puissances et rayonnements harmoniques à l'émission sont :

- EN 300-220 Europe / LoRa et autres 434-868 MHz
- EN 300-330 Europe / NFC & RFID 13.56 MHz
- EN 300-328 Europe / Bluetooth & WiFi
- EN 301-908 Europe / 4/5G et dérivés
- EN 301-489 Europe / CEM des produits radio
- FCC Part 15 / Toutes radio dans les Amériques

... et le puzzle asiatique où rien n'est harmonisé !

Il est à noter, qu'il existe des modules radio RF pré-certifié, donc des composants électroniques intégrant un émetteur et/ou un récepteur radiofréquence, qui a déjà été testé et certifié pour répondre aux réglementations de certains pays ou régions. Cela simplifie le développement de produits, car le fabricant du produit final peut s'appuyer sur la certification existante du module.

La Revue 3EI : Utilisez-vous LoRa® ou SigFox et que pensez-vous de leurs performances ?

J. Brunet et P. Champaney : Nous utilisons LoRa énormément car aujourd'hui 95% de nos projets sont en 868-915 MHz. Sa force inégalée est d'offrir des sensibilités de réception énormes tout en conservant une bonne immunité aux perturbateurs. Sa seule contrepartie est un débit binaire limité à 30 kbps max, ce qui interdit quasiment les mises à jour logiciel OTA ("Over The Air").

Sigfox a quasiment disparu de nos projets alors qu'il était leader voilà 10 ans. Comme quoi la chute va vite dès que la pertinence technique est noyée par du marketing ...

La Revue 3EI : Comment se déroule l'étape de qualification, notamment en termes de performances rayonnées ?

J. Brunet et P. Champaney : Nous relevons majoritairement quatre points essentiels : la puissance rayonnée fonctionnelle, les rayonnements indésirables et harmoniques à l'émission, et les remontés parasites en réception, pour la conformité réglementaire.

Pour ces opérations, le produit est complet en boîtier et il développe une émission permanente. La mesure du champ électrique à quelques mètres permet d'en déduire la puissance émise.

Always Wireless est complètement indépendant des laboratoires d'essais sur ces relevés.



Figure 4 : Mesures de rayonnement et spectre

Les mesures sont généralement effectuées en utilisant une antenne de référence, placée à une distance fixe de l'équipement à mesurer et connectée à un analyseur de spectre. L'utilisation d'une antenne de référence permet de disposer d'une table de gain en fréquence. Cette table, une fois renseignée dans l'analyseur de spectre, permet d'effectuer des mesures précises de 500 MHz à 13,6 GHz.

L'équipement sous tests peut être mis en rotation à distance sur les plans X, Y & Z afin de mesurer pleinement son rayonnement dans l'espace.

La Revue 3EI : Comment s'effectue des tests de robustesse et d'immunité radio ?

J. Brunet et P. Champaney : Le récepteur sous test reçoit un signal « utile » périodique produit par un générateur RF et dont il indique la bonne réception. Un second générateur développe un signal « perturbateur » qui est additionné à l'utile. Son niveau est relevé jusqu'à perdre la réception.

Ce test peut être dans le canal de réception (cocal) ou les canaux adjacents et/ou bandes latérales (blocking).

Pour ces tests, Always Wireless est capable d'enregistrer des trames de test clients pour les reproduire à l'identique mais de niveau maîtrisé, via un générateur RF.

3 - Intégration d'antennes

La Revue 3EI : Comment bien prendre en compte l'environnement et le contexte d'une antenne, avant sa conception ?

J. Brunet et P. Champaney : Le point premier consiste à travailler le dégagement de l'antenne. Le champ électrique RF ne peut être produit que si celle-ci est éloignée de tout matériau conducteur. Il lui faut souvent une surface conductrice en son pied pour assurer un contrepoids électrique mais son extrémité doit toujours être la plus dégagée possible.

Heureusement, les fabricants d'antennes « composants » ont fait de gros efforts de documentation de leurs solutions. A respecter à la lettre et éviter, ici aussi, tout exotisme !

De façon générale, une antenne est composée de métal, de forme géométriquement adaptée à une ou plusieurs fréquences de résonance. Peu importe sa technologie mise en avant par son marketing, sa principale propriété physique réside dans le fait que ses dimensions sont liées à sa fréquence de résonance, il n'est pas possible de faire autrement. Attention donc aux datasheets « magiques » !

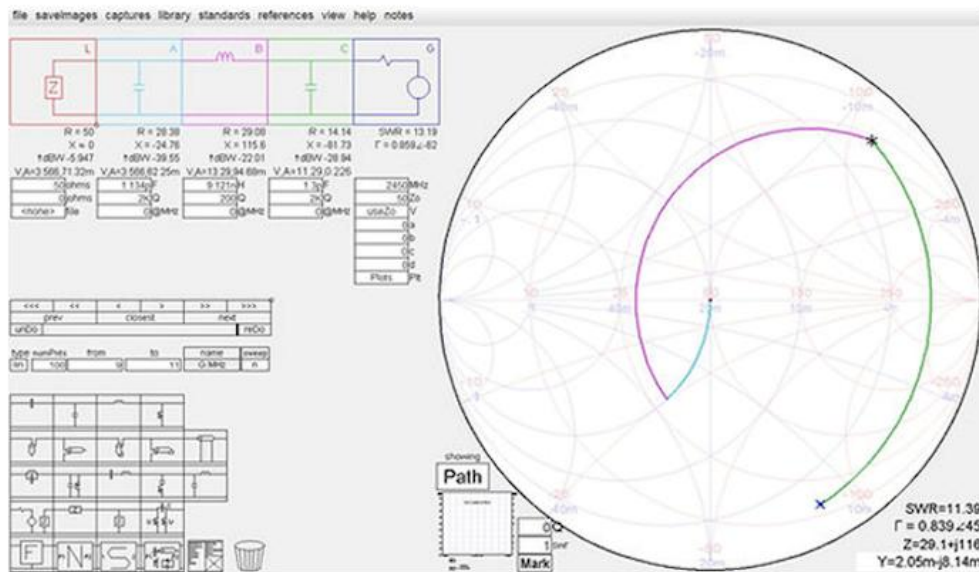
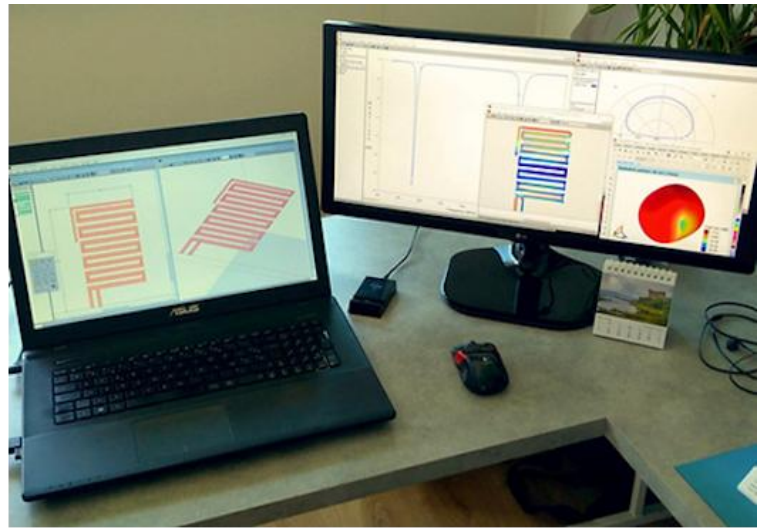
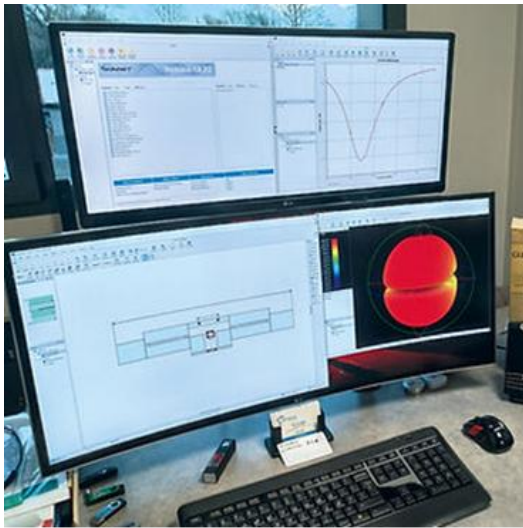


Figure 5 : Conception d'antennes et inductances imprimées

La Revue 3EI : Quelles sont les principales étapes lors de la conception de solutions personnalisées ?

J. Brunet et P. Champaney : Always Wireless ne développe que des antennes dédiées qu'en forme « 2D imprimée ». La société ne propose pas de design 3D car trop coûteux en licence logicielle et temps de modélisation. Impossible à amortir et assumer sauf à ne faire que cela, ce qui n'est pas la vocation de la société.

La première étape consiste à bien déterminer les encombrements et positions possibles (et faire preuve de pédagogie car les barrières sont ici physiques donc incontournables !). Une fois le contexte établi, il faut placer les formes des résonateurs d'antenne en fréquences croissantes et optimiser leur structure pour maximiser les rayonnements.

Bien comprendre le principe d'un résonateur électromagnétique est ici essentiel et cela requiert des années et des années de pratique. La théorie n'apporte rien.

Nous pouvons simuler électromagnétiquement les designs d'antenne. Ce qui peut prendre plusieurs jours pour une antenne LTE multibande.

Après avoir fourni un plan d'implémentation et du support à l'implémentation, nous demandons un prototype complet, mécanique incluse (car la permittivité du plastique décale la fréquence de

résonnance des antennes) afin de procéder à la mise au point d'un réseau d'adaptation d'impédance.

Enfin, pour terminer, nous procédons à des mesures de rayonnement comme expliqué précédemment afin de juger des performances réelles.

La Revue 3EI : Quels outils ou savoir-faire sont nécessaires pour la phase d'optimisation ?

J. Brunet et P. Champaney : Les outils indispensables sont l'analyseur de réseau vectoriel puis l'analyseur de spectre.

Le premier permet de travailler les accords en fréquence des résonateurs de l'antenne et d'assurer un bon transfert de puissance électrique vers ceux-ci. L'observation se fait en conduisant en examinant la réponse de l'impédance au pied de l'antenne et en ajustant les composants associés.

Le second donne un diagramme « Puissance RF / Fréquence » et permet de juger de l'efficacité de rayonnement de l'élément. La structure peut alors être retouchée pour augmenter son efficacité.

Le sens physique du comportement des résonateurs électromagnétiques est ici primordial.

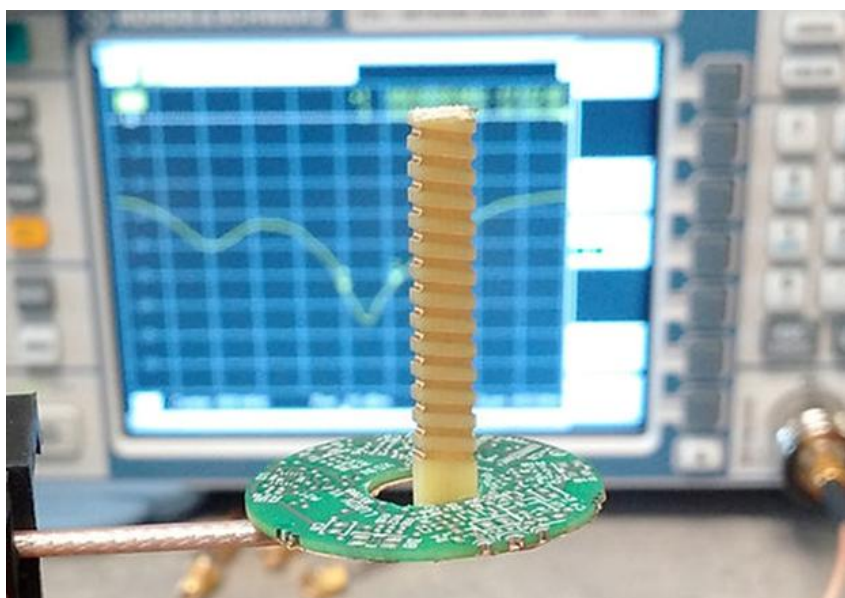


Figure 6 : Relevé d'une réponse en impédance. La résonance à la fréquence d'accord se manifeste par un creux dans la réponse.

La Revue 3EI : Comment se déroule l'étape de qualification, notamment pour l'extraction de diagrammes de rayonnement ?

J. Brunet et P. Champaney : Le produit est complet en boîtier et il développe une émission permanente. Il est à quelques mètres d'une antenne de référence (gain connu) reliée à un analyseur de spectre. La mesure du champ électrique vue par l'antenne de référence permet d'en déduire la puissance émise. Ce rayonnement est examiné par rotation de 360° selon les trois axes.

- Le champ maximum enregistré donne la puissance rayonnée maximale ou le gain maximum d'antenne.
- Un relevé pas par pas (typiquement 15°) donne le diagramme de rayonnement.
- Ces mesures sont faites en extérieur pour s'affranchir de tout obstacle.

4 - Laboratoire et instrumentation

La Revue 3EI : La figure ci-dessous, montre votre poste de travail pour la qualification de produits RF, avec différents instruments. Quels sont-ils exactement ?

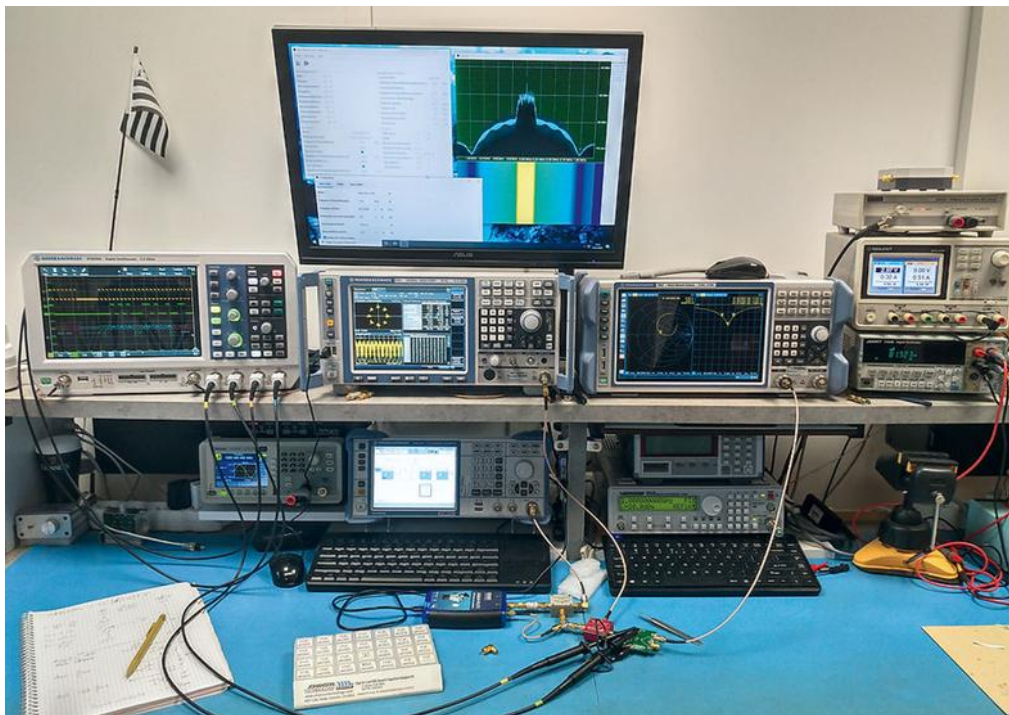


Figure 7 : Appareillage de mesure RF

J. Brunet et P. Champaney : Nous disposons de l'ensemble des instruments nécessaires pour la conception RF analogique et numérique :

- R&S FSV14 Analyseur de signaux et analyseur de spectre : cet instrument permet d'analyser n'importe quelle modulation analogique ou numérique afin de vérifier le contenu et la conformité de la modulation. Sa fonction primaire est l'analyse de spectre, ses très bonnes performances de balayage permettent de visualiser spectralement des modulations numériques large bande.
- R&S SMBV100B Générateur de signaux IQ 6 GHz : cet instrument permet de moduler de manière arbitraire l'ensemble des modulations classiques analogique ainsi que l'ensemble des modulations numériques avancées, jusqu'à une bande passante de 500 MHz. Il est également capable de « rejouer » des signaux RF enregistrés avec le R&S FSV14.
- R&S SMBV100A Générateur de signaux IQ 3 GHz : nous utilisons cet instrument en second du SMBV100B. Sur ce dernier, nous pouvons en plus des caractéristiques citées plus haut, générer une constellation satellite de type GPS afin d'effectuer une mesure de sensibilité de récepteurs GPS. Cela permet de valider de manière répétable, les performances de tout récepteur GPS ayant bénéficié d'une conception customisée.
- R&S ZNL6 Analyseur de réseaux vectoriels & analyseur de spectre 6 GHz : cet instrument permet de mesurer les paramètres S de manière simple ou différentielle de n'importe quels quadripôles. L'obtention des paramètres S est une étape obligatoire lors de la conception d'un circuit d'adaptation d'impédance. Sa fonction secondaire d'analyseur de spectre permet de compléter le R&S FSV14.
- R&S RTB2004 Oscilloscope numérique 2.5 Gs/s : cet instrument à mémoire numérique nous permet d'analyser le signal démodulé en sortie du R&S FSV14. Avec une mémoire allant

jusqu'à 20 millions de points et des fonctions mathématiques intégrées, l'analyse d'une modulation est rendue agréable.

- Divers alimentations & générateur de signaux BF & RF.

Les trois plus gros instruments du bas sont les générateurs RF. Le premier en haut à gauche est un oscilloscope numérique, hélas peu utilisé vue notre activité. En haut, sous le moniteur se trouve l'analyseur de spectre, ici en démodulateur numérique. A sa droite, l'analyseur de réseau vectoriel, où l'on distingue une réponse d'impédance sur abaque de Smith.

La Revue 3EI : Vous disposez d'une cellule GTEM. De quoi s'agit-il et en quoi vous est-elle utile ?

J. Brunet et P. Champaney : La GTEM (cellule Giga Transverse ElectroMagnétique), est une mini chambre anéchoïque qui embarque une antenne de référence sur sa face intérieure arrière. Elle offre une (petite) zone où le champ électrique est parfaitement maîtrisé. Cela permet de faire des mesures de rayonnement relatif et de juger rapidement des améliorations ou non d'un produit à l'émission.

Une fois la cellule close, le produit est isolé de l'extérieur afin d'évaluer des sensibilités de réception.



Figure 8 : Cellule GTEM

Néanmoins, cela reste un instrument peu précis, uniquement exploitable en développement afin de parfaire des mises au point. Des mesures absolues en émission imposent un site extérieur ou une grande chambre anéchoïque.

La Revue 3EI : Vous êtes amenés à réaliser des interventions sur un site de mesure en plein champ. Quelles en sont les raisons ?

J. Brunet et P. Champaney : Il n'est hélas pas possible de faire des relevés en rayonné sur un site quelconque, car non maîtrisé et souvent garni d'obstacles. Notre site extérieur limite au maximum les réflexions et sa réponse de 30 MHz à 15 GHz nous est parfaitement connue. Nos qualifications sont donc limitées à cet emplacement.

De fait, et même s'ils sont encombrants, ce sont nos clients qui nous envoient les produits pour qualification. Heureusement, la majorité d'entre eux sont de petite taille.



Figure 8 : Mesure de rayonnement en plein champ

La Revue 3EI : Merci beaucoup pour toutes ces informations, où se mélangent science, électronique et passion !

Ressource publiée sur Culture Sciences de l'Ingénieur : <https://sti.eduscol.education.fr/si-ens-paris-saclay>

En tant que membre de la SEE, vous bénéficiez de nombreux avantages



1 **Élargir son réseau professionnel par la participation à des structures de réflexion adaptées**

• 6 Clubs techniques

- Cybersécurité et réseaux Intelligents
- Ingénierie des systèmes d'information et de communication
- Radar, sonar et systèmes radioélectriques
- Stockage et nouveaux moyens de production
- Eco-conception en génie électrique
- Systèmes électriques

• 6 Groupes régionaux

2 **Participer et bénéficier de tarifs préférentiels pour les Conférences et Journées d'études SEE**

- Conférences nationales et internationales
- Journées d'études thématiques
- Conférences-débats
- Congrès internationaux, en partenariat ou non avec d'autres sociétés scientifiques

3 **Consulter et télécharger gratuitement les publications (REE et 3EI) en version numérique et s'abonner aux publications papier à tarif préférentiel**

- **La Revue de l'électricité et de l'électronique (REE)** est destinée aux ingénieurs, chercheurs, enseignants, décideurs techniques et économiques intéressés par les secteurs de l'électricité, de l'électronique, de l'information et de la communication. Paraissant cinq fois par an, la revue s'articule autour de dossiers techniques, flash-Infos, articles invités, entretiens avec des personnalités du monde de la recherche et de l'industrie.
- **La revue 3EI** est une publication trimestrielle destinée aux professeurs, universitaires et industriels concernés par l'enseignement de l'électricité et de l'électronique industrielle.

4 **S'engager dans une association qui reconnaît les talents et crée une émulation parmi ses membres**

- Grades senior et émérite SEE
- Remise de prix et trophées pour les professionnels confirmés, jeunes actifs et étudiants (Brillouin-Glavieux, général Ferrié, André Blanc Lapierre...)
- Remise de médailles (Ampère, Blondel...)

Pour rejoindre la SEE, deux modes d'adhésion :



Adhésion individuelle :
via le site www.see.asso.fr
ou le bulletin d'adhésion



Adhésion collective :
partenariat d'entreprise ou
d'école via des conventions

NOUS CONTACTER :

SEE - Service adhésions - 17 rue de l'amiral Hamelin - 75116 Paris - France
+33 (0)1 56 90 37 17 - adhesion@see.asso.fr - www.see.asso.fr

En 2025, ✓ j'adhère à la SEE



La SEE, société savante française fondée en 1883, forte de 2 000 membres, couvre les secteurs de l'électricité, de l'électronique et des technologies de l'information et de la communication.

BULLETIN À COMPLÉTER ET RENVoyer À :

SEE - Service adhésions - 17 rue de l'Amiral Hamelin
75116 Paris - France.
Tél : +33(0)1 56 90 37 17 - adhesion@see.asso.fr

☐ J'adhère à la SEE

☐ Standard	130 €
☐ Retraité	70 €
☐ Enseignant	70 €
☐ Jeune actif < 35 ans	70 €
☐ Etudiant	15 €
☐ Recherche d'emploi	15 €

* Une remise de 10% est accordée aux membres IEEE

Merci d'indiquer votre n° de membre IEEE :

La SEE change son mode d'adhésion et passe à une adhésion d'un an, date à date à partir de la date de paiement. Le 1^{er} numéro servi pour la REE sera le numéro suivant la date de paiement. Plus d'informations sur l'abonnement sur le site web de la SEE.

☐ Je m'abonne à la revue REE de la SEE à un tarif préférentiel !*



REE - La Revue de l'Électricité et de l'Électronique

4 numéros par an (Mars, Mai, Octobre Décembre)

- ☐ Livraison France : **68 € TTC**
- ☐ Livraison UE : **78 € TTC** (76,40 € HT*)
- ☐ Livraison Hors UE : **83 € TTC** (81,70 € HT*)



La Revue 3EI

4 numéros par an

Retrouvez la revue 3EI en accès gratuit sur le site web de la SEE

* Prix HT valide si le pays de facturation est hors UE, ou si la TVA Intracommunautaire est fournie pour un pays de l'UE. ** Prix réservés aux adhérents - Abonnement dans la limite des stocks disponibles.

Mes coordonnées / Adresse de livraison

☐ Mr ☐ Mme Prénom* : _____ Nom* : _____
Adresse* : _____
Code postal* : _____ Ville* : _____ Pays* : _____
Téléphone* : _____ email** : _____
(*Obligatoire) **En adhérant à la SEE, votre mail est le moyen de contact et d'information de l'association & de ses activités (congrès, soirées débats, revues, etc.). Vous acceptez donc de recevoir les diffusions de l'association.

Adresse de facturation (si différente)

Raison sociale de l'employeur : _____ Service : _____ Activité : _____
(Facultatif)
Adresse : _____
Code postal : _____ Ville : _____ Pays : _____
email* : _____
(*Obligatoire)

BON DE COMMANDE :

Mon règlement

- ☐ Virement bancaire :
BNP Paribas, Paris Associations (02837)
IBAN - FR76 3000 4002 7400 0103 3624 258
- ☐ Chèque bancaire ou postal **à l'ordre de la SEE**
- ☐ Carte bancaire, lien de paiement sécurisé (visa, etc)

Signature obligatoire



Adhérez dès aujourd'hui via le site internet de la SEE sur : www.see.asso.fr

DOSSIER

Géolocalisation

P.46



CAHIER AZUR

Journées «Compatibilité radioélectrique»

Rennes 2024 P.84

➤ POINT DE VUE

L'ÉLECTRIFICATION CROISSANTE OFFRE
AUX JEUNES DES MÉTIERS D'AVENIR
PATRICK LECLERC P.112

➤ ARTICLE INVITÉ

LA CONCEPTION DE RÉSEAUX TELECOM
EN FONCTION DE LEUR IMPACT
BERNA SAYRAC & JEAN-MARC KELIF P.38

➤ ENTRETIEN AVEC...

EMMANUEL GRAVIER,
PRÉSIDENT DU CONSUEL
ALAIN BRENAC & JACQUES HORVILLEUR P.122



ETTC 2025

European Test and Telemetry Conference

10-11 June 2025, Toulouse

Upcoming generation of Testing for new mobile
platforms and communication systems

Partnership guide



11 Esplanade Compans Caffarelli, 31000 Toulouse

Contact : Lucas Dominguez | lucas.dominguez@carte-blanche.fr

05 63 72 31 25 – 06 28 15 07 77