

Titre: Détection d'activités humaines par l'utilisation du CSI
Title:

Auteur: Laure Pourcel
Author:

Date: 2023

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Pourcel, L. (2023). Détection d'activités humaines par l'utilisation du CSI
[Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/54374/>

Document en libre accès dans PolyPublie

Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/54374/>
PolyPublie URL:

Directeurs de recherche: Chahe Nerguzian, & Gabriela Nicolescu
Advisors:

Programme: Génie électrique
Program:

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Détection d'activités humaines par l'utilisation du CSI

LAURE POURCEL
Département de Génie électrique

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Génie électrique

Août 2023

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

Détection d'activités humaines par l'utilisation du CSI

présenté par **Laure POURCEL**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

Christian CARDINAL, président

Chahé NERGUIZIAN, membre et directeur de recherche

Gabriela NICOLESCU, membre et codirectrice de recherche

Elham BALADI, membre

DÉDICACE

*À mes parents, Philippe et Marie Reine
Qui m'ont toujours soutenu et aimé,
Pendant les 22 dernières années,
Promis, je rentre dans deux mois à peine.*

REMERCIEMENTS

Je tiens à remercier le professeur Chahé Nerguzian pour ses conseils tout au long de ma maitrise et le temps qu'il a accordé à ce projet. Sa confiance a été le socle d'une relation de travail agréable et fructueuse.

Je remercie également Humanitas pour avoir proposé ce projet, financé la recherche et fournit un appui technique au cours des deux dernières années. Tout particulièrement je souhaite exprimer ma gratitude envers la docteure Maroua Ben Attia pour toutes les idées qu'elle a donné et le docteur Abdo Shabah pour sa bienveillance sans faille.

J'adresse un remerciement tout particulier au docteur et ami Antoine Bernard pour la relecture de ce manuscrit. Il est rare de rencontrer des personnes aussi sincères et gentilles. J'aspire à te ressembler un peu plus.

Je veux aussi adresser ma reconnaissance envers la professeure Gabriela Niculescu, directrice du département de génie informatique et logiciel, et MITACS pour avoir contribué au financement de ce projet de recherche, ainsi que Lizbeth Meneses Flores, secrétaire du département de génie informatique et logiciel, pour ses contributions administratives rapides et efficaces.

Il est aussi très important de remercier Pierre et Clélia pour leur contribution lors des collectes de données.

Enfin un très grand merci à tous les bougs, pour vos contributions mineures commes majeures. À Flokie, Florian, Gabriel, Mathias, Nero, Peno, Samuel, Tanguy et Vi. Même dispersés au quatre coins du monde on pourra compter les uns sur les autres. Merci pour tout.

RÉSUMÉ

La technologie Wireless Fidelity (Wi-Fi) a été développée à la fin des années 1990 et peut être considérée comme arrivée pleinement à maturité. L'une des preuves de cette maturité est l'existence d'usages qui se détournent de l'objectif premier de la technologie, qui est de fournir un accès internet sans fil. Le Wi-Fi est désormais utilisé dans des applications de localisation, de domotique et internet des objets ou de communication vocale. Ces applications utilisent certaines caractéristiques techniques du Wi-Fi de manière créative. Par exemple, la puissance du signal reçu (RSSI) est un indicateur qui a pour but de déterminer la qualité du signal reçu, permettant de moduler dynamiquement la vitesse de transmission ou d'initier un processus de déconnexion quand le signal est trop faible. Certaines applications de localisation utilisent le RSSI pour déterminer la position d'un utilisateur, une manière détournée de mettre à profit des spécifications préexistantes.

L'atout principal d'applications s'appuyant sur le Wi-Fi est le potentiel de les commercialiser à très grande échelle. En effet, le Wi-Fi domine largement les réseaux domestiques et publics. De même, il est prévalant dans les réseaux professionnels. Ainsi, le déploiement de nouvelles fonctionnalités peut s'adresser à un public particulièrement large. On peut citer le protocole VoIP qui permet de transmettre de la voix sur les réseaux IP, développé au cours des années 1980 et 1990 et dont l'utilisation était relativement confidentielle jusqu'à la démocratisation du Wi-Fi. Il est désormais utilisé par des applications populaires de messagerie et d'appels, telles que WhatsApp, Skype, FaceTime, Viber, Zoom, Microsoft Teams, dont le nombre d'utilisateurs est estimé à des centaines de millions.

Nous allons appuyer notre étude sur la connaissance de l'état du canal (CSI). Cet indicateur s'appuie sur les nombreuses sous-porteuses du Wi-Fi. Il nous informe sur la qualité du signal transmis sur chaque sous-porteuse entre plusieurs antennes. Le CSI est impacté par les différents phénomènes physiques qui s'appliquent à un signal qui se propage en espace libre, réflexion, réfraction et diffraction. La variation de la qualité de la transmission des signaux du Wi-Fi a été déterminée empiriquement suffisamment fine pour pouvoir différencier plusieurs activités. Les modèles théoriques sont actuellement trop difficiles à mettre en place, ces phénomènes physiques dépendant de nombreux paramètres, tandis que lors d'un mouvement réalisé par un sujet, l'environnement est modifié à chaque seconde.

La littérature cite de nombreuses utilisations du CSI, qui vont du comptage de foule à la localisation de personnes ou la détection d'activités humaines. Elles présentent des résultats très encourageants, certains articles atteignent une précision supérieure à 90% sur la

reconnaissance de sept activités, tandis que d'autres atteignent 95% sur la reconnaissance de gestes de la main. On note cependant que les techniques utilisées sont dures à mettre en place en dehors du cadre d'un laboratoire ou d'une expérience unique. Cela tient principalement au matériel utilisé, ainsi qu'aux conditions d'une expérience scientifique difficilement reproductibles en dehors d'un laboratoire dont l'environnement est parfaitement contrôlé et statique.

Notre objectif final est de détecter des chutes dans un environnement hospitalier. En nous appuyant sur la littérature, nous cherchons à déterminer au cours de notre étude les limites de la détection par le CSI en utilisant un microcontrôleur ESP32. Ce microcontrôleur relativement peu onéreux présente l'avantage de pouvoir collecter des CSI et les traiter sans avoir besoin d'une connexion filaire à un outil réalisant le traitement comme un ordinateur. Nous mettrons en place des expériences avec différentes méthodes de collecte. La méthode passive connecte un appareil à un ESP32, et est donc relativement simple à mettre en place. Connecter son appareil personnel à un ESP32 qui a été correctement déployé au préalable n'est pas plus compliqué que de se connecter à son réseau Wi-Fi domestique. La méthode de collecte active est plus complexe dans la mise à échelle d'une application de détection d'activités humaines puisqu'elle nécessite de connecter deux ESP32 entre eux, ce qui est une installation plus lourde, mais qui permet un meilleur contrôle dans l'envoi des paquets. À ces paquets, une étape de prétraitement pour modifier leur représentation, mettre en exergue les petites variations ou débruiter peut ou non être appliquée. Nous utilisons ensuite l'apprentissage machine, en anglais Machine Learning (ML) sur les données que nous collectons. Plus précisément, devant le volume significatif de données, nous employons des représentations graphiques analysées avec un réseau de neurones à convolutions (CNN), algorithme de ML particulièrement adapté à la gestion de données de haute dimensionnalité.

Nous avons fait varier graduellement la difficulté perçue des scénarios de classification. Dans un premier temps, nous avons utilisé la méthode de collecte active, qui a déjà fait ses preuves dans la littérature. Nous avons d'abord réalisé la détection d'aucune, une ou deux personnes et obtenu une précision de 100% avec un prétraitement minime. Nous avons ensuite réalisé la classification sur sept activités - marcher, courir, être allongé, se lever, s'asseoir, se pencher et tomber et atteint une précision de 98% dans la classification subséquente. Nous avons enfin réalisé une expérience pour distinguer trois cas avec deux personnes réalisant respectivement les activités marcher toutes les deux, marcher et tomber, marcher et être allongée. Empiriquement, les réflexions des ondes sur deux personnes complexifient beaucoup le travail de classification. Plus précisément, le nombre de trajectoires possibles pour le CSI augmente grandement. Nous avons effectivement obtenu une précision de 82% parmi les trois classes considérées.

Ensuite, nous avons implémenté une connexion passive, nous avons souhaité détecter la présence d'aucune, une ou deux personnes dans trois environnements différents. Cette étape est importante, car ne pas réaliser d'analyse quand une pièce est vide permet une économie d'énergie importante. De même, il n'a pas été démontré qu'un modèle qui détecte les chutes efficacement en la présence d'une personne en est également capable en présence de deux personnes. Les résultats sont encourageants, avec une détection de présence qui avoisine les 100%, tandis que la classification entre zéro, une ou deux personnes s'établit autour des 93%.

ABSTRACT

The Wireless Fidelity (Wi-Fi) technology was developed in the late 1990s and should be considered to have reached full maturity by now. One of the proof of this is the fact that new applications are developed and stray from Wi-Fi's original objective, which is providing a wireless internet access. Wi-Fi may now be used for localization, home automation, Internet of Things (IoT) or voice communications. These applications use creatively technical features of Wi-Fi. For instance, the Received Signal Strength Index (RSSI) is normally used in order to quantify the intensity of a signal, thus allowing to dynamically modify the velocity of transmission or start a reconnection process if the signal is too weak. Some applications now leverage the RSSI for localization, a new way to use a preexisting technology.

The main advantage of Wi-Fi-based application is scalability. We can't ignore the fact that Wi-Fi largely dominates other internet access technology, in both domestic and public networks, and is prevalent as well in professional networks. Therefore, any new application could theoretically reach a particularly wide audience. One notable example is the VoIP protocol, developed in the late 1980s and early 1990s, this protocol allowed voice transmission over IP. While it was relatively confidential before the democratization of Wi-Fi, it is now used by applications like WhatsApp, Skype, FaceTime, Zoom and Microsoft Teams, which have hundreds of millions of users.

In our study, we aim to leverage Channel State Information (CSI). This index is associated to each subcarrier of Wi-Fi in each transmitted packet. It informs us on the quality of the transmitted signal, which like every signal propagating in free space is subject to physical phenomenons such as reflection, refraction and diffraction. It was experimentally proven that the variation of CSI was a stable enough indicator to pursue Human Activity Recognition (HAR). Theoretical models are yet to be developed because the complexity of considering every physical interaction for multiple frequency in fast moving and non-reproducible settings should not be underestimated.

There is a vast diversity of CSI-based applications, ranging from crowd counting to localization or HAR. Some have fairly impressive results, reaching a 90% accuracy on recognizing activities among 7 different possibilities or reaching a 95% on hand gesture recognition. There is however a shortage of real life applications, most do not leave the confined environment of a laboratory. This is explained by the high level of control one has in such a setting, which is an enviable feature as you carry out an experiment.

Our ultimate goal is to detect fall events in a hospital. After our literature review, we decided

to implement our CSI-based HAR solution using an ESP32. This microcontroller is relatively cheap compared to others. It is able to collect CSI, which is the most important criteria we need to take into account, and it may do the operations required to launch a Machine Learning model in order to do classification on board later on. We will do experiments with two different ways of collecting our data. The first one is called the passive way, it connects any device using Wi-Fi to a previously set-up ESP32. It is fairly easy to scale, because this method only requires one piece of hardware, since everyone already using Wi-Fi also owns equipment to connect to it. It's also pretty simple to use because connecting to the Wi-Fi of a correctly installed ESP32 is as simple as connecting to any Wi-Fi which you know the password of. We will then use the active method of collection, which is slightly harder to scale because it requires two ESP32. They send packets to each other, one acts as the transmitter and one acts as the receiver. While this method requires two times more hardware, it allows for a better control over how many packets are sent.

Once the collecting concerns are addressed, we preprocess our data, then use Machine Learning to process and thus classify them. Specifically, since our data have a high dimensionality we use Convolutional Neural Network (CNN) algorithms, which are particularly useful in such cases.

We gradually varied the perceived difficulty of the classification scenarios. Initially, we employed the well-established active data collection method. We first performed detection of no, one, or two persons and achieved a 100% accuracy with minimal preprocessing. Next, we conducted classification on seven activities - walking, running, lying down, standing up, sitting down, bending, and falling - and achieved a subsequent classification accuracy of 98%. Finally, we conducted an experiment to differentiate between three cases involving two persons performing walking and walking, walking and falling, and walking and lying down activities. Empirically, reflections of waves from two individuals greatly complicated the classification task. Specifically, the number of possible trajectories for the Channel State Information (CSI) increased significantly. We indeed achieved an accuracy of 82% among the three considered classes.

Subsequently, we implemented a passive connection to detect the presence of no, one, or two persons in three different environments. This step is important as it allows for significant energy savings by avoiding unnecessary analysis when a room is empty. Similarly, it has not been demonstrated that a model that efficiently detects falls in the presence of one person is also capable of doing so in the presence of two persons. The results are promising, with presence detection reaching approximately 100%, while classification between zero, one, or two persons is at around 93%.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	viii
TABLE DES MATIÈRES	x
LISTE DES TABLEAUX	xiii
LISTE DES FIGURES	xiv
LISTE DES SIGLES ET ABRÉVIATIONS	xvi
LISTE DES ANNEXES	xvii
 CHAPITRE 1 INTRODUCTION	1
1.1 Définitions et concepts de base	1
1.1.1 Le Wi-Fi	1
1.1.2 Le CSI	2
1.1.3 L'apprentissage Machine	4
1.2 Objectifs de recherche	5
1.3 Plan du mémoire	6
 CHAPITRE 2 REVUE DE LITTÉRATURE	7
2.1 Différentes utilisations du CSI	8
2.2 Méthodes de collecte du CSI	10
2.3 Technique de prétraitement	11
2.3.1 Choix de la représentation des données	11
2.3.2 Prétraitement appliqué aux données considérées	12
2.3.3 Réduction de la dimensionnalité	13
2.3.4 Algorithme de traitement	14
2.3.5 Jeux de données	14

CHAPITRE 3 PROTOCOLE DE COLLECTE DE DONNÉES	16
3.1 Composants considérés	16
3.1.1 bcm43455c0	16
3.1.2 ESP32	16
3.1.3 Choix de l'ESP32	17
3.2 Collecte active	17
3.2.1 Set-up expérimental	17
3.2.2 Données	18
3.3 Collecte passive	20
3.3.1 Set-up expérimental	21
3.3.2 Données	22
CHAPITRE 4 PRÉTRAITEMENT	24
4.1 Caractéristiques de nos données	24
4.1.1 Sous-porteuses de garde	24
4.1.2 Tailles des échantillons	24
4.1.3 Valeurs du CSI dans nos échantillons	29
4.2 Prétraitement	31
4.2.1 Filtre Hampel pour la détection d'anomalies	32
4.2.2 Filtre passe bande sur les représentations des amplitudes	33
CHAPITRE 5 RÉSULTATS	37
5.1 Présentation des algorithmes de traitement	39
5.2 Collecte active	43
5.2.1 Détection de la présence d'aucun, un ou deux personnes	43
5.2.2 Détection de l'activité réalisée par une personne	44
5.2.3 Détection de l'activité réalisée par deux personnes	48
5.3 Collecte passive	48
5.3.1 Détection de la présence dans trois environnements	48
5.3.2 Détection de la présence avec entraînement et validation dans des environnements différents	51
5.3.3 Détection de la présence d'aucune, une ou deux personnes dans trois environnements différents	53
CHAPITRE 6 CONCLUSION	55
6.1 Synthèse des travaux	55
6.2 Limitations de la solution proposée	57

6.3 Améliorations futures	58
RÉFÉRENCES	59
ANNEXES	64

LISTE DES TABLEAUX

Tableau 3.1	Table des données collectées dans le cadre de la collecte active	20
Tableau 3.2	Table des données collectées dans le cadre de la collecte passive	23
Tableau 4.1	Valeurs des moyennes et des écarts types du nombre de paquets lors de la première collecte de données passive	26
Tableau 4.2	Valeurs des moyennes et des écarts types du nombre de paquets lors de la seconde collecte de données passive	26
Tableau 4.3	Valeurs des moyennes et des écarts types du nombre de paquets lors de la collecte de données active	28
Tableau 5.1	Cas de prétraitement considérés pour la détection d'activités avec la méthode de collecte active	47
Tableau 5.2	Précisions obtenues	48
Tableau 5.3	Cas de prétraitement considérés	50
Tableau 5.4	Précisions obtenues	51
Tableau 5.5	Différents arrangements lors de l'entraînement et la validation sur différents sets de données	52
Tableau 5.6	Précisions obtenues	52

LISTE DES FIGURES

Figure 1.1	Fréquences orthogonales en OFDM	2
Figure 1.2	Réflexions du CSI sur le corps humain	3
Figure 1.3	Schéma de principe d'un CNN	5
Figure 2.1	Les principales utilisations du CSI	8
Figure 3.1	Chambre d'essai	18
Figure 3.2	Bureau	22
Figure 3.3	Couloir	22
Figure 3.4	Salle des switchs	22
Figure 4.1	Répartition de la taille de tous les échantillons de la première collecte passive	25
Figure 4.2	Diagramme en boîte de la taille des échantillons en fonction de leur classe et environnement après la seconde collecte	27
Figure 4.3	Répartition finale de la taille de tous les échantillons de la collecte passive	28
Figure 4.4	Répartition de la taille de tous les échantillons de la collecte passive	30
Figure 4.5	Répartition de la taille de tous les échantillons de la collecte passive	31
Figure 4.6	Valeurs moyennes des paquets successifs dans un échantillon	32
Figure 4.7	Valeurs moyennes des paquets successifs dans un échantillon après l'usage du filtre Hampel	33
Figure 4.8	Représentation du gradient des amplitudes pour l'échantillon 15 de la classe marcher avant ajout du contraste	35
Figure 4.9	Représentation du gradient des amplitudes pour l'échantillon 15 de la classe marcher avant ajout du contraste	35
Figure 5.1	Différentes expériences considérées	38
Figure 5.2	Architecture du premier CNN	40
Figure 5.3	Architecture du deuxième CNN	41
Figure 5.4	Architecture du troisième CNN	41
Figure 5.5	Architecture du quatrième CNN	42
Figure 5.6	Précision et perte sur les ensembles d'entraînement et de validation pour la classification entre aucune, une ou deux personnes présente dans le cadre de la collecte active	43
Figure 5.7	Échantillon avant et après le retrait des sous-porteuses de garde sans prétraitement	45

Figure 5.8	Précision et perte sur les ensembles d'entraînement et de validation pour une collecte active et la classification entre sept activités	46
Figure 5.9	Précision et perte sur les ensembles d'entraînement et de validation pour une collecte active et la classification des activités de deux personnes	49
Figure 5.10	Précision et perte sur les ensembles d'entraînement et de validation pour une collecte passive et la détection de présence d'aucune, une ou deux personnes	54

LISTE DES SIGLES ET ABRÉVIATIONS

AP	Access Point
CNN	Convolutional Neural Network
CSI	Channel State Information
DNN	Deep Neural Network
ESP	Electronic Stability Program
FDMA	Frequency Division Multiple Access
HAR	Human Activity Recognition
IEEE	Institute of Electrical and Electronics Engineers
ML	Machine Learning
NDC	Nano Data Center
OFDM	Orthogonal Frequency Division Multiplexing
PCA	Principal Component Analysis
SVM	Support Vector Machine
Wi-Fi	Wireless Fidelity

LISTE DES ANNEXES

Annexe A	Exemple d'un paquet CSI	64
Annexe B	Répartition des valeurs totales	65

CHAPITRE 1 INTRODUCTION

Les chutes constituent la première cause de blessures chez les Canadiens âgés, avec des conséquences allant du bénin au fatal [1]. Il existe une corrélation directe entre le temps de prise en charge et les conséquences de la chute [2]. Plusieurs systèmes de détection de chute ont été développés, s'appuyant sur l'analyse d'images [3,4], la détection d'accélération [5], ou une alerte manuelle [6].

Dans cette étude, nous proposons d'utiliser la connaissance de l'état du canal (CSI), qui est un indicateur de qualité des réflexions des sous-porteuses Wi-Fi. Comme le CSI est défini pour les sous-porteuses du Wi-Fi, il ne nécessite pas l'ajout de matériel supplémentaire. Effectivement, l'hôpital où nous souhaitons déployer notre solution est déjà équipé de relais Wi-Fi dans toutes les pièces où l'on veut réaliser la détection de chutes.

Nous voulons détecter dans un premier temps la présence puis les chutes. L'avantage du CSI est qu'il ne nécessite pas de ligne de vue directe, une bonne luminosité ou une action de la personne qui chute. Des systèmes de détection de chute qui s'appuient sur le CSI existent déjà [7]. Dans notre étude, nous utilisons des Nano Data Center (NDC) qui embarquent une carte ESP32, carte qui nous permet de collecter 64 sous-porteuses, soit plus que les 30 habituelles dans la littérature.

1.1 Définitions et concepts de base

1.1.1 Le Wi-Fi

Le Wi-Fi est une technologie de communication sans-fil qui permet de transférer des données entre différents appareils. Cet ensemble de protocoles de communications sans-fil s'est imposé comme dominant dans les réseaux domestique, professionnel et public. En effet, il présente des avantages certains, telle la mobilité, le haut débit ou la simplicité d'utilisation.

L'Institute of Electrical and Electronics Engineers (IEEE) définit les normes de communications de la technologie Wi-Fi pour les versions successives : 802.11 a/b/g/n/ac/ax. Ces normes définissent des spécifications techniques sur différentes couches du réseau, comme la bande de fréquence, la vitesse de transmission, le nombre de sous porteuses, la gestion de la mobilité entre plusieurs points d'accès, etc.

L'accès multiple par répartition en fréquence (FDMA)

Le FDMA est une technique d'accès multiple qui permet à plusieurs appareils de transmettre simultanément en utilisant différentes sous-porteuses. À chaque appareil est attribué une ou plusieurs sous-porteuses pour l'envoi et la réception de données, le nombre de ces sous-porteuses dépend de la norme considérée. La norme 802.11n définit 64 ou 128 sous-porteuses en fonction de la taille du canal, tandis que la norme 802.11ac atteint 512 sous-porteuses. Cette méthode d'accès permet d'optimiser l'utilisation du spectre disponible.

Le FDMA permet à plusieurs appareils de se connecter simultanément, mais n'est pas suffisant pour transmettre à plusieurs appareils en même temps. Pour ce faire, il faut utiliser une modulation appropriée, telle que la modulation Orthogonal Frequency Division Multiplexing (OFDM). En effet, en OFDM, les fréquences sont orthogonales donc elles n'interagissent pas entre elles, le Wi-Fi peut donc transmettre des données simultanément sur plusieurs sous-porteuses sans craindre d'interférence. En OFDM, certaines sous-porteuses sont conservées nulles ou constantes, on les appelle des fréquences de garde, elles servent à bien conserver l'orthogonalité des fréquences. Une représentation graphique est donnée à la figure 1.1.

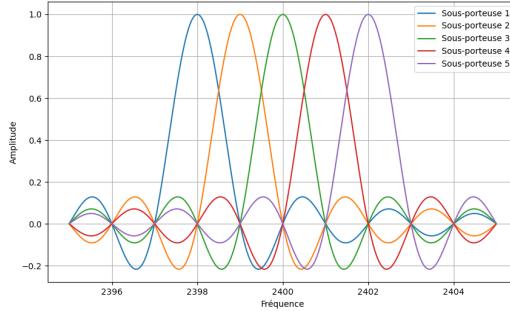


Figure 1.1 Fréquences orthogonales en OFDM

1.1.2 Le CSI

La norme 802.11n [8] de l'IEEE définit le Channel State Information (CSI). Le CSI est un indicateur de qualité des différentes sous-porteuses du Wi-Fi. Il est impacté par les facteurs environnementaux autour des antennes émettrices et réceptrices, qui causent de la diffraction, de la réfraction et de la réflexion.

Le CSI est de plus en plus utilisé dans des applications de détection, telles que la détection de présence, de mouvements ou de chutes. On se sert de l'impact du corps humain en mouvement,

qui fait varier les phénomènes physiques qui s'appliquent au CSI. Un schéma de principe tiré de [9] est donné à la figure 1.2.

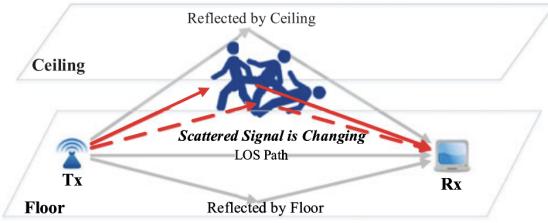


Figure 1.2 Réflexions du CSI sur le corps humain

Le CSI est une valeur complexe, composé d'une phase et une amplitude. Dans le cas où l'on collecte 64 sous-porteuses, on le considère matriciellement, sous la forme d'une matrice ligne \mathbf{H} avec autant de colonnes que de sous-porteuses. Chaque élément est un CSI associé à une sous-porteuse. On peut représenter la matrice \mathbf{H} telle que :

$$\mathbf{H} = [CSI_1 \ CSI_2 \ CSI_3 \ \dots \ CSI_{64}]$$

Avec :

$$CSI_k = A_k e^{j\phi_k}$$

Où CSI_k représente le CSI de la k-ème sous-porteuse, A_k est l'amplitude du CSI de la k-ème sous-porteuse, ϕ_k est la phase du CSI de la k-ème sous-porteuse.

Une nouvelle matrice \mathbf{H} est associée à chaque paquet transmis par le Wi-Fi. Le taux d'envoi des paquets peut varier d'une dizaine à plusieurs centaines par seconde. Nous utiliserons la finesse des variations des amplitudes et des phases au cours du temps dans notre étude.

Les avantages de la détection de chute par le CSI

La détection de chutes par le CSI offre plusieurs avantages comparatifs dans le domaine de la surveillance et de la sécurité des personnes. Pour en citer quelques-uns :

- C'est une technique qui est respectueuse de la vie privée. Au contraire de la détection de chute qui s'appuie sur l'analyse d'image [3, 4] qui peut se révéler intrusive puisque pour être efficace, il faut être filmé en continu, ce qui freine le déploiement des solutions de détection de chute s'appuyant sur de la surveillance vidéo intelligente à domicile.

- Le temps de réaction est rapide. En effet, le principal problème des systèmes qui s'appuient sur une alarme déclenchée manuellement est le temps de réaction [7]. En considérant une chute relativement grave, le patient peut s'évanouir et donc ne pas être en mesure de donner l'alarme. Cela est d'autant plus problématique que ces chutes requièrent une prise en charge immédiate.
- Les conditions d'applicabilité sont relativement larges. Le CSI n'a pas besoins de ligne de vue directe (LOS) ni d'une bonne luminosité au contraire des systèmes d'analyse d'image. De leurs côtés, les systèmes à déclencher manuellement ou basé sur la détection d'accélération doivent être portés en continu. Si le porteur oublie de les mettre, cela annule leur efficacité. L'avantage d'un système de détection de chutes basé sur le CSI est qu'il est actif sans avoir besoin d'actions supplémentaires.
- Il existe déjà des travaux de recherche dans le domaine qui ont fait leurs preuves, nous les détaillerons dans le chapitre 2. Cependant, il reste des défis, sinon une solution commerciale existerait déjà, les seules solutions que nous avons trouvées sont des preuves de concept.
- L'écrasante majorité des domiciles sont couverts par le Wi-Fi. une solution de détection de chutes s'appuyant sur le CSI, donc le Wi-Fi, est potentiellement déployable à très grande échelle. En somme, cela ne nécessiterait pas de matériel supplémentaire.

En résumé, la détection de chutes par le CSI présente de nombreuses qualités. Ces avantages font du CSI un outil prometteur pour améliorer la sécurité des personnes et répondre rapidement aux incidents de chute, contribuant ainsi à prévenir les blessures et à sauver des vies.

1.1.3 L'apprentissage Machine

L'apprentissage automatique, Machine Learning (ML) en anglais, est une branche de l'informatique qui se concentre sur l'utilisation d'algorithmes pour permettre aux ordinateurs de reconnaître des modèles dans des données sans connaître la fonction mathématique réelle qui les décrit. Il tente d'approximer la fonction réelle en associant des poids successifs aux données d'entraînement qui lui sont fournies. On distingue les algorithmes de ML supervisés, dont les données d'entraînement porte un label correspondant à la classification désirée, et non supervisés où les données ne sont pas labellisées.

Les réseaux de neurones convolutionnels

Les réseaux de neurones convolutionnels (CNN) sont un type de modèle de ML utilisé principalement pour l'analyse d'images, le traitement du signal et d'autres tâches qui impliquent

un grand volume de données. Le recours à un CNN est adapté pour notre étude qui nécessite l'analyse de nombreux paquets successifs, contenant eux-mêmes plusieurs CSI.

Un CNN est composé de couches de convolution, qui vont appliquer des filtres sur les valeurs numériques des couleurs rouge, verte et bleue (RGB) qui composent chaque pixel. Ces filtres réalisent des multiplications matricielles avec les valeurs RGB, appelées convolutions. Les valeurs au sein des matrices sont changées au cours du temps pour améliorer les classifications subséquentes.

Les résultats en sortie des convolutions sont ensuite passés en entrée dans une fonction d'activation, qui introduit de la non-linéarité dans notre modèle. La fonction d'activation la plus commune est la fonction ReLU qui associe à tout argument négatif la valeur zéro et garde les arguments positifs inchangés. Les couches dites de pooling vont après réduire le nombre de valeurs numériques considérées en sortie d'une fonction d'activation, en gardant uniquement la plus grande ou la moyenne des valeurs dans une certaine fenêtre prédéfinie. Cela permet de diminuer la dimensionnalité des données tout en conservant de l'information. Ce processus est répété plusieurs fois, puis les sorties des couches de convolution, des fonctions d'activation et des couches de pooling sont données en entrées d'un réseau de neurones classique. On représente à la figure 1.3 un schéma de principe du fonctionnement des CNN, tiré du site jeveuxetredatascientist.com.

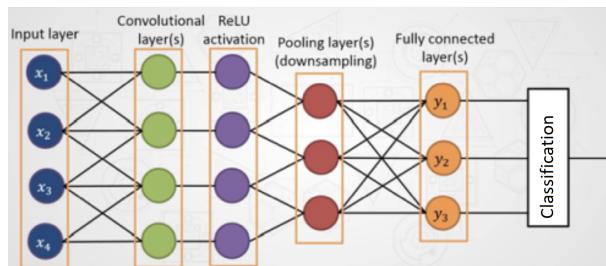


Figure 1.3 Schéma de principe d'un CNN

Les CNN ont déjà largement fait leurs preuves dans des domaines tels à la reconnaissance d'image, détection d'activités, reconnaissance faciale, etc.

1.2 Objectifs de recherche

Nous effectuons notre travail de recherche pour un hôpital dont l'objectif est de détecter les chutes passivement afin de réduire le temps de prise en charge des patients à la suite des événements de chute. En gardant cet objectif final en tête, nous pouvons distinguer plusieurs étapes nécessaires pour mener à bien le projet.

Dans un premier temps, il est nécessaire de collecter des CSI, dans plusieurs cas d'utilisation, représentant la présence ou les activités réalisées par les potentiels patients. Nous allons chercher à classifier les activités réalisées par une personne dans une pièce, mais aussi chercher à dénombrer le nombre de personnes dans une pièce et poserons les bases de l'analyse des activités de deux personnes. Une fois ce riche jeu de données collecté, nous réaliserons une étape de prétraitement afin de pouvoir l'analyser efficacement. Il est à noter que le prétraitement pourra varier d'un cas d'utilisation à un autre. Nous allons ensuite tester des CNN sur nos données prétraitées, dans le double objectif de vérifier qu'elles sont effectivement corrélées avec leurs classes associées et donc classifiables. Le second objectif est de déterminer quel est le meilleur prétraitement parmi ceux que nous considérons. Pour ce faire, on comparera les performances de classifications.

1.3 Plan du mémoire

Le reste de ce mémoire est composé de cinq chapitres. Le prochain chapitre sera la revue de littérature, qui nous permet d'appuyer notre étude sur un socle théorique important. Il est suivi du chapitre 3 qui porte sur la collecte des données, spécifiant le matériel utilisé, le protocole de collecte et donnant un profil des données collectés dans les quatre expériences différentes que nous considérons. Le chapitre 4 détaillera ensuite les différents prétraitements de données que nous allons — ou non — appliquer aux données que nous avons collectées, dans le but d'améliorer les performances de notre modèle. Enfin, le chapitre 5 donnera les résultats de classification pour les données collectées activement puis passivement. Le chapitre 6 nous permettra de conclure sur les avancées de notre travail, la faisabilité d'un déploiement et les améliorations possibles.

Les références bibliographiques et des annexes se trouvent à la fin du mémoire.

CHAPITRE 2 REVUE DE LITTÉRATURE

La détection s'appuyant sur le Wi-Fi est un domaine en plein essor. On discerne la détection qui s'appuie sur des capteurs qui utilisent le Wi-Fi pour communiquer et la détection qui s'appuie sur des indicateurs ou spécifications techniques propres au Wi-Fi.

Parmi les applications qui communiquent via le Wi-Fi, on retrouve un grand nombre d'applications de domotique. On peut citer des capteurs de température [7], des systèmes connectés contrôlant les accès à un domicile [10] ou des systèmes de sécurité utilisant des caméras [11].

Pour la détection intelligente s'appuyant sur les indicateurs du Wi-Fi, il existe principalement des applications qui s'appuient sur le RSSI, le CSI et l'angle d'arrivée (AoA) du signal Wi-Fi. Pour citer quelques utilisations :

- L'AoA peut être utilisé de concert avec une méthode de triangulation s'appuyant sur la différence temporelle entre des paquets reçus successivement ou par des équipements différents, dans l'objectif de localiser en environnement interne [12]. En localisant à intervalles réguliers, on peut également utiliser l'AoA pour déterminer la trajectoire suivie par un appareil qu'on localise [13]. Cela répond au manque de précision du système GPS qui permet une localisation efficace en extérieur, mais n'est pas suffisamment précis en environnement interne. La marge d'erreur sur la précision de la position est de l'ordre du mètre, ce qui est problématique en intérieur puisque quelques mètres de différence peuvent induire en erreur quant à l'étage auquel se situe ce que l'on souhaite localiser, le numéro d'appartement ou encore la pièce [14].
- Le RSSI est également utilisé pour localiser en environnement interne, en s'appuyant sur la force du signal émis par un appareil que l'on souhaite localiser [15]. Si l'on souhaite localiser quelqu'un il doit porter l'appareil qui émet. Cette technologie s'appuie sur des modèles théoriques et empiriques de l'atténuation permettant d'estimer la distance à un point en fonction de la force d'un signal reçu, et de localiser si l'on a 3 points ou plus qui transmettent la force du signal qu'ils reçoivent. Le RSSI peut aussi être utilisé pour détecter la présence d'individus. Cela semble logique que l'on puisse détecter la présence si l'on est déjà en capacité de localiser une personne. Cependant, la différence fondamentale est que la localisation s'appuie sur un équipement qu'on localise, tandis que la détection de présence est réalisée sans aucun capteur supplémentaire. Ce procédé s'appuie simplement sur la variation de la puissance d'une transmission préexistante. Celle-ci est impactée par les mouvements d'un individu qui vient modifier l'environnement où a lieu la transmission et modifie donc la puissance

du signal reçu [16]. La détection de présence passive a des applications dans le domaine de la sécurité.

Nous détaillerons les différentes applications du CSI dans la section suivante.

2.1 Différentes utilisations du CSI

Le CSI est un indicateur suffisamment fin dans ses variations pour être utilisé dans des applications plus complexes que la localisation ou l'identification de présence. On présente ici les usages retrouvés dans la littérature, des plus répandus aux plus ponctuels.



Figure 2.1 Les principales utilisations du CSI

Reconnaissance d'activités : Les possibilités associées au développement d'un module de reconnaissance d'activités humaines (HAR) sans fil ni capteur à porter sont très nombreuses. Au sein de maisons connectées, reconnaître quelqu'un qui se lève permet d'entamer des tâches de manière préemptive. À l'inverse, quelqu'un qui s'allonge dans son lit pourrait être un signal pour éteindre ou réduire les lumières. On retrouve de très nombreux articles qui s'intéressent à la détection d'activités [17–22]. Au sein des applications de HAR on retrouve une branche de la recherche qui s'est concentrée sur la détection de chutes. Effectivement, elles constituent un enjeu de santé public majeur. Il existe déjà des systèmes de détections

de chutes, utilisant l'analyse d'image, ce qui soulève une problématique de vie privée, ou bien utilisant des capteurs d'accélération qui ont encore des performances imparfaites, ou finalement utilisant des systèmes d'alarme à activer manuellement, ce qui n'est pas possible quand la chute est grave et la personne s'évanouit. Ainsi, il existe de nombreuses recherches sur la détection de chutes à l'aide du CSI [7, 9, 21, 23–26]

Détection de présence et localisation : La détection de présence est principalement utilisée pour des utilités à visée sécuritaire. On pourrait également imaginer des usages commerciaux ou domotiques dont le but est d'économiser de l'énergie - par exemple éteindre des enseignes lumineuses, des écrans ou des lumières quand aucune présence n'est détectée. La détection de présence par le CSI obtient d'excellents résultats, s'approchant des 100 % [27, 28]. La localisation en intérieur est l'évolution de la détection de présence, dans le sens qu'une personne est détectée et en plus une position lui est assignée. Elle est intéressante pour des applications commerciales. Le CSI est couramment utilisée pour des applications de géolocalisation [29–32]. Il obtient des résultats encourageants, avec une précision inférieure à l'ordre de grandeur du mètre, et présente l'avantage de ne pas avoir besoins de matériel supplémentaire, contrairement à la localisation par le RSSI. Le principal défaut de la géolocalisation par le CSI est qu'elle ne s'appuie pas sur un modèle théorique, et nécessite donc une prise de données systématique.

Reconnaissance de mouvements : La reconnaissance de mouvement est un domaine particulièrement prometteur, car les débouchés commerciaux sont très nombreux, on peut imaginer des objets intelligents réagissant à certains mouvements — déverrouiller son portable d'un signe de la main, de même éteindre une lumière ou verrouiller une porte d'un simple geste. De même, ses applications pourraient se révéler utiles pour augmenter l'inclusivité pour les personnes handicapées de certaines tâches du quotidien. On retrouve donc des travaux de recherche pour distinguer les mouvements des doigts, des bras voir de tout le corps en restant à une position statique [21, 33–35].

Comptage de foule : Il existe de nombreuses applications qui poussent le principe de la détection de présence afin de compter le nombre de personnes dans un lieu. Cela peut répondre à des contraintes réglementaires sur le nombre d'individus maximales qui peuvent être simultanément dans un endroit [36] ou alors, on peut utiliser ces applications en extérieur pour compter le nombre de participants à un événement, à des fins marketing ou pour des visites touristiques [37].

Application médicale : Le CSI est utilisé dans des applications à visée médicales comme le monitorage de la respiration afin de mieux comprendre et traiter des maladies comme l'apnée du sommeil [38]. On retrouve également un usage du CSI pour surveiller les battements du cœur de patients, dans l'objectif de surveiller des arythmies [39–41].

Usages ponctuels : Il n'est pas possible de citer toutes les manières dont le CSI est utilisé tant elles sont nombreuses. D'autres fonctions plus anecdotiques par leur représentation dans la littérature sont la détection d'incendies [42], le suivi des niveaux de moisissure dans le blé [43] ou la reconnaissance de vibration dans un milieu industriel [44].

2.2 Méthodes de collecte du CSI

L'extraction du CSI est une étape qui peut se révéler technique. Il faut prendre en compte que toutes les cartes réseaux ne le permettent pas. Le CSI a été défini dans la norme 802.11n. Les réseaux et composant qui mettent en œuvre cette norme ou les subséquentes peuvent permettre de collecter le CSI, les précédents ne le permettent pas. Ensuite, si les normes suivantes définissent le CSI, toutes les cartes réseau ne permettent pas d'accéder à cette information, et aucune ne la donne spontanément. Il est toujours nécessaire de coupler un matériel avec un module de collecte.

Dans la littérature, plusieurs méthodes pour collecter sont utilisées. On retrouve de manière prédominante la carte Intel 5300 avec le module d'extraction linux 802.11n CSI tool [45], notamment dans les articles [7, 25, 27, 29, 32, 33, 37, 37, 39, 41–44]. Elle présente l'avantage d'être relativement simple d'utilisation.

Les cartes Atheros [36, 38] permettent également l'extraction du CSI en utilisant le module d'extraction Atheros CSI Tool [46]. Elles présentent l'avantage de collecter les CSI associés à 56 sous-porteuses, contre 30 pour la carte Intel 5300. De plus, l'intégration avec une distribution Linux est facilitée quand on utilise ces cartes.

Il est aussi possible d'utiliser les cartes Broadcom [20, 40] avec le module NEXMON [47]. Les puces Broadcom sont notamment intégrés au sein des Raspberry Pi, des équipements qui peuvent être utilisés au sein de projets complexes et variés.

Enfin, dans notre étude, nous nous concentrerons sur l'utilisation de microcontrôleurs ESP32 [21], qui présentent de nombreux avantages comparatifs. Notre premier prérequis est la possibilité de collecter le CSI, ce qui est permis par l'outil ESP32 CSI TOOL [48]. L'ESP32 est par ailleurs un microcontrôleur relativement facile à se procurer, en opposition avec la carte Intel 5300. Il est programmable à l'aide de l'interface ESP-IDF et à la différence des cartes

Wi-Fi, peut être utilisé indépendamment sans avoir besoin d'autres matériaux dans le cas d'un déploiement industriel.

2.3 Technique de prétraitement

À la collecte, le CSI est représenté sous forme complexe, soit avec une partie imaginaire et une partie réelle. Plus précisément, on récupère x et y tels que :

$$x + iy = A \cdot e^{i\theta}$$

avec x la partie réelle, y la partie imaginaire, A l'amplitude et θ la phase.

Il existe plusieurs manières d'extraire les informations contenues dans le CSI. On s'intéressera dans un premier temps aux différentes représentations considérées, puis aux techniques de prétraitement et de débruitage que l'on retrouve dans la littérature.

2.3.1 Choix de la représentation des données

Amplitude : Comme mentionnée précédemment, à la collecte, on obtient la partie réelle et imaginaire du CSI. Le choix de représentation le plus couramment utilisé est de considérer uniquement l'amplitude du CSI [7, 43]. C'est-à-dire dans l'équation de la partie précédente uniquement A . Ce choix est motivé par la praticité de l'amplitude, qui est obtenue par une simple norme euclidienne depuis les parties réelles et imaginaires. De plus, par définition, elle contient les informations relatives à la force du signal et la qualité de sa réception.

Phase : Il est considéré que la phase du CSI est un indicateur volatil [39]. Pour cette raison, il est rarement utilisé. Par contre, des transformations sur la phase du CSI peuvent rendre l'index plus stable. On peut citer une transformation linéaire appliquée à la phase mesurée pour la transformer en phase réelle [27]. Il est également possible de coupler la transformation linéaire à une seconde étape qui a pour but de faire la différence entre deux phases consécutives [9].

Amplitude et phase : Il est peu courant d'utiliser uniquement la phase du CSI mais on retrouve plusieurs articles qui couplent une représentation de la phase et l'amplitude [38, 41, 44]. Les informations relativement stables que constituent les amplitudes sont légèrement pondérées par les variations plus fines des amplitudes linéarisées.

Représentation statistique : Qualifier le signal peut passer par une représentation des indicateurs principaux qui le caractérisent. C'est l'approche utilisée dans l'article [7], qui utilise sept indicateurs comme l'écart type normalisé ou la rapidité du changement du signal pour réaliser ensuite avec succès une classification d'activités.

Transformée en ondelettes discrètes : La Transformée en ondelettes discrètes est une transformation mathématique qui permet de représenter le signal dans un domaine temporel et fréquentiel. Elle est par exemple utilisée dans les articles suivants [36, 42]. C'est une transformation initiale relativement lourde, dont la plus-value devant des représentations plus simples comme l'amplitude uniquement n'est pas prouvée [21].

2.3.2 Prétraitement appliqué aux données considérées

Une fois le choix des données réalisé vient l'étape de prétraitement. Pour ce faire, les indicateurs considérés sont modifiés, dans le but de mettre en exergue les variations significatives qui permettront d'atteindre l'objectif fixé, que ce soit de la détection de présence, d'activité, d'incendie, ou pour des applications à visée médicale.

Dans un premier temps, il faut noter que de nombreux articles font le choix de n'utiliser aucun prétraitement [24, 33]. Le fondement théorique d'une telle approche est l'hypothèse que toute transformation va faire perdre de la richesse à nos données. Chaque prétraitement peut par inadvertance retirer en effet des variations intéressantes dans l'étape de traitement subséquente. On obtient des résultats de classifications très encourageants sans prétraitement, ce qui vient souligner que cette approche est intéressante, en plus d'être simple à mettre en place.

Interpolation : Une étape de prétraitement très commune est l'interpolation, qui consiste en espacer linéairement les temps d'arrivée des paquets. Cela permet de les passer dans le domaine fréquentiel [9, 34, 36], et simplifie globalement la manipulation des paquets.

Usage de filtres : S'appuyant sur des observations quant à la période des mouvements que l'on souhaite classifier, on retrouve l'utilisation de filtre dans la phase de prétraitement de nombreux articles. Certains articles utilisent un filtre passe-bas afin de couper les fréquences trop hautes [34, 38]. On retrouve également un filtre passe-bande [9] quand l'hypothèse que les variations sur les fréquences trop petites ne sont pas pertinentes. Le filtre Butterworth constitue un passe-bande dont l'atténuation aux fréquences de coupure et au-delà est plus progressive, c'est une autre possibilité également utilisée [42, 43]. Enfin, le filtre dont la com-

plexité est la plus importante lors de l'étape de prétraitement est le filtre Savitsky-Golay, qui consiste en approximer des points successifs en minimisant l'erreur quadratique moyenne avec des polynômes de faibles rangs [41].

Détection d'anomalies et débruitage : Comme tout système physique, le matériel ou le logiciel pour collecter le CSI peut commettre des erreurs lors de la retranscription des valeurs. C'est en partant de ce constat que plusieurs articles décident d'implémenter de la détection de valeurs aberrantes afin de supprimer les points que l'on ne considère pas représentatif de notre signal réel. Le filtre Hampel notamment est utilisé à cette fin, il fixe à la médiane les valeurs qui dépassent un certain seuil, fixé entre la médiane plus ou moins un à trois écarts types [27, 29]. L'une des critiques du filtre Hampel est qu'il retire potentiellement des variations importantes pour la classification subséquentes. En s'appuyant sur une réflexion analogue à celle menée pour les anomalies, une méthode de débruitage s'appuyant sur l'usage de moyenne dans des fenêtres glissantes est également proposée [7, 34].

2.3.3 Réduction de la dimensionnalité

Les données CSI, que l'on considère les amplitudes ou les différences de phases, sont d'une haute dimensionnalité. Il faut considérer à chaque seconde au moins plusieurs dizaines de paquets, et ceux pendant des durées variant de quelques secondes à plusieurs heures. Ainsi, l'un des enjeux du prétraitement peut être de réduire la dimensionnalité des données que nous souhaitons analyser afin de rendre le traitement plus rapide et simplifier la mise à l'échelle de solution en temps réel.

Principal Component Analysis : Le PCA analyse la contribution relative des valeurs propres dans les n premières directions demandées, et représente ensuite nos données dans ces directions. Cette technique est très efficace pour réduire la dimensionnalité des valeurs à analyser tout en conservant les informations importantes [44]. Cependant, la phase de calibration où les données sont projetées peut s'avérer relativement lourde.

Corrélation statistique entre les sous-porteuses : Certains articles mettent en avant des comportements similaires entre des sous-porteuses adjacentes et proposent donc de les représenter par une seule valeur [17, 39]. Il faut noter cependant que le consensus scientifique n'établit pas clairement quelles sous-porteuses sont corrélées ou non, ainsi que le nombre de sous groupes distincts que l'on peut former pour représenter efficacement les corrélations.

2.3.4 Algorithme de traitement

On note l'usage de la plupart des algorithmes classiques de ML pour les applications se basant sur le CSI. Pour citer les principaux et leur fonctionnement :

- Les Support Vector Machine (SVM) sont très courants [9, 27, 29, 36, 44]. Un SVM représente les points d'entraînement qui lui sont passés en arguments dans autant de dimensions que ces points ont de caractéristiques. L'algorithme va ensuite tracer des surfaces de décision qui séparent les points des différentes classes en tentant de minimiser la distance euclidienne globale entre ses surfaces de décision et les points.
- L'algorithme Random Forest (RF) est également relativement fréquent [22, 24]. Un RF est un algorithme de ML qui crée ce que l'on appelle des arbres. Les arbres vont s'entraîner sur une sous-sélection des points d'entraînement de l'algorithme. Ils créent des feuilles, qui sont des critères de décision sur certaines caractéristiques aléatoirement sélectionnées du sous groupe d'échantillons. Le RF crée ainsi plusieurs arbres. Quand l'algorithme rencontre un nouvel échantillon, chaque arbre va lui assigner une classe en fonction de comment le nouveau point passe les différentes feuilles de l'arbre. Enfin, un vote majoritaire déterminera la classification du nouveau point.
- On retrouve également du traitement par des Deep Neural Network (DNN) [33, 38]. Les DNN sont composés de plusieurs couches de neurones. Chaque neurone aura une fonction sous-jacente, généralement affine, pour transformer les valeurs en entrée en nouvelles valeurs numériques. On appelle poids les facteurs de la fonction sous-jacente du neurone. Ils sont modifiés à chaque itération pour améliorer les performances de classification. On retrouve ensuite une fonction d'activation qui introduit de la non-linéarité dans nos modèles. Enfin, chaque neurone d'une couche est connecté à tous ceux de la couche suivante, refaisant la transformation linéaire des valeurs d'entrée à chaque neurone de chaque couche.
- Enfin, l'un des algorithmes le plus représenté est le CNN [22, 24], que nous avons présenté à la section 1.1.3. Il présente l'avantage d'être très adapté aux données de haute dimensionnalité.

2.3.5 Jeux de données

Une corrélation claire est établie entre la fréquence des paquets reçus et la précision des classifications subséquentes [21]. Au-delà du nombre de paquets reçus par seconde, on note également qu'il est important d'avoir un nombre d'échantillons relativement important, dont le nombre varie grandement. Pour citer quelques exemples de jeu de données observés dans la littérature :

- Dans l'objectif de classifier 7 activités, on retrouve 60 échantillons pour chaque activité [24].
- Dans l'objectif de classifier 11 activités, on retrouve 100 échantillons pour chaque activité [20].
- Dans l'objectif de classifier 8 activités, on retrouve 200 échantillons pour chaque activité [19].
- Dans l'objectif de classifier 3 activités, on retrouve 240 échantillons pour chaque activité [18].
- Dans l'objectif de classifier 3 activités, on retrouve 260 échantillons pour chaque activité [17].
- Dans l'objectif de classifier 10 activités, on retrouve 50 échantillons pour chaque activité [26].

Il faut noter que dans chaque cas, le nombre d'échantillons est le même pour chaque activité. Cela fait partie des standards en ML, dans le cas de jeux de données avec des répartitions inéquitables, les algorithmes ont souvent de moins bonnes performances [49].

Il n'existe pas de règle unique pour déterminer la taille optimale d'un jeu de données. Il doit être suffisamment important pour permettre d'observer l'émergence de motifs récurrents. Cependant, entraîner sur un jeu de données trop important est coûteux en termes de ressources, tant pour les calculs que le temps nécessaire pour le constituer.

CHAPITRE 3 PROTOCOLE DE COLLECTE DE DONNÉES

Le premier objectif qu'il faut remplir pour mener à bien notre projet est de collecter des CSI afin de constituer un jeu de données. C'est sur celui-ci que nous appliquerons les pré-traitements retenus avant de mettre en place la dernière étape de classification. Ce chapitre présentera d'abords le matériel que nous avons retenu et avec lesquels nous avons collectés des CSI, pour ensuite présenter les deux modes de collecte que nous avons implémenté et les jeux de données associés.

3.1 Composants considérés

Dans le cadre de notre étude, nous souhaitons utiliser des composants qui s'intègrent dans des systèmes programmables. Il a déjà été prouvé que la détection de chutes avec une carte Intel 5300 est réalisable [7, 9, 23–26]. Le set-up expérimental est cependant difficilement reproducible en dehors du cadre d'un laboratoire. Dans tous ces articles, une connexion est établie entre un ordinateur sur lequel le module d'extraction linux 802.11n csitool [45] est installé et la carte Intel5300. Pour que la détection de chute ait lieu, il faut donc un ordinateur dédié qui réalise les classifications.

Notre objectif en utilisant d'autres composants pour l'extraction est de limiter la complexité du set-up expérimental.

3.1.1 bcm43455c0

Nous avons dans un premier temps réalisé l'extraction des CSI sur une carte bcm43455c0, elle-même intégré au sein d'un Raspberry Pi B3. Pour ce faire, nous avons flashé le projet NEXMON [47] sur une carte Raspberry Pi avec un module Arduino.

3.1.2 ESP32

Les ESP32 mis à notre disposition sont intégrés dans des NDC qui font tourner une distribution linux. On peut se connecter via ssh à ces NDC et exécuter des scripts shell dessus, que ce soit pour installer des projets ou combiner des projets à des instructions bash.

En utilisant le projet ESP32 CSI TOOL [48], il est possible de collecter de deux manières différentes des CSI avec un ESP32 :

- La première méthode est dite passive. L'ESP32 avec le projet installé préalablement

va passivement collecter autour de lui les paquets de données transmis sur la fréquence à laquelle il écoute et fournir les CSI associés à ces paquets.

- La seconde manière de collecter est dite active. Elle consiste en installer des projets distincts sur deux ESP32, qui agiront respectivement comme transmetteur (dit Station) et récepteur (dit Access Point (AP)).

Dans notre étude, nous avons utilisé dans un premier temps la méthode passive, puis la méthode active.

3.1.3 Choix de l'ESP32

La puce bcm43455c0 et l'ESP32 nous ont tous deux permis de collecter des CSI. Nous avons choisi d'utiliser l'ESP32 car ceux mis à notre disposition sont intégrés au sein de NDC qui utilisent une distribution Linux. Cela les rend plus simples à programmer dans le cadre de notre étude, et donc plus simple à déployer potentiellement.

3.2 Collecte active

Nous utilisons la collecte active, car elle permet un meilleur contrôle du nombre de paquets reçus sur une période de temps définie. Nous avions initialement prévu d'utiliser la collecte passive uniquement. Devant le faible taux de paquets collectés et les précisions attenantes, nous avons fait le choix de réaliser une expérience de détection de présence avec le mode de collecte actif également.

Nous avons ensuite mené une seconde expérience, dont l'objectif est de classifier sept activités réalisées dans un environnement. Les activités sont marcher, courir, se lever, s'asseoir, être allongé, tomber et se pencher.

Finalement, nous avons réalisé une dernière expérience de classification pour déterminer quelles activités sont réalisées par deux personnes, parmi trois paires d'activités : marcher et marcher, marcher et tomber, marcher et être allongé.

3.2.1 Set-up expérimental

Nous utilisons deux ESP32, l'un agit comme transmetteur et l'autre comme récepteur. Ils sont connectés en Wi-Fi. Sur chacun est installé le code correspondant : active station et active access point [48]. La pièce est une chambre, qui fait 6 mètres de long et 2,5 de large. On y retrouve un lit, une commode et une table de chevet. La chambre est représentée à la figure 3.1 par deux prises de vue du même environnement.

Les ESP32 sont disposés sur l'un des murs de la chambre, entre le lit et la commode, comme on peut l'observer sur les photographies. On utilise une approche qui se base sur la réflexion des paquets sur le corps humain avant d'atteindre le second ESP32, en opposition avec une approche qui se base sur la transmission entre les deux ESP32.



Figure 3.1 Chambre d'essai

3.2.2 Données

Avec un sujet

Les données sont collectées de deux manières différentes. On sépare en deux catégories les activités, on a d'une part, celles qui sont continues comme marcher ou courir et d'autre part, celles qui ont un début et une fin précis comme tomber ou s'asseoir.

Les données correspondant à des activités qui peuvent être réalisées en continu, marcher, courir ou être allongé sont collectées en réalisant deux longues collectes de 2 minutes. Les quatre premières et dernières secondes de ces échantillons sont coupés. On justifie cela par le fait que pour commencer et terminer la collecte, une action manuelle sur l'ordinateur est requise, qui pourrait être confondu dans nos données avec l'activité que nous souhaitons collecter, comme marcher ou courir. On collecte également de cette manière les échantillons associés à

notre pièce sans aucun sujet à l'intérieur. Ces échantillons serviront à la classification entre aucune, une ou deux personnes.

Les données correspondant à des activités ayant un début et une fin précise sont collectées une par une. Nous avons regroupé dans un même échantillon s'asseoir et se lever puis séparé cet échantillon en deux. Pour chaque échantillon d'activité réalisé au coup par coup, nous coupions les quatre premières secondes et les quatre dernières en utilisant la même logique, voulant que l'activité de se pencher sur l'ordinateur et taper sur le clavier pourrait être apparenté avec celle que nous souhaitons réellement classifier.

Nos échantillons contiennent un nombre variable de paquets CSI. Ils sont chacun composé de la partie réelle et imaginaire du CSI associé à chaque sous-porteuse. Par une transformation, on obtient les phases et les amplitudes du CSI de 64 sous-porteuses. Parmi ces 64 phases et amplitudes, 11 sont nuls ou constantes. On parle de fréquences de garde. En annexe A, nous fournissons un exemple de fichier contenant des paquets CSI.

Avec deux sujets

Dans ce cas, nous collectons trois paires d'activités dans l'objectif de limiter le nombre de collectes. Les paires considérées sont un premier sujet en train de marcher dans chaque cas et un second en train de marcher, de tomber et de rester allongé. Nous avons choisi ces combinaisons car marcher, puis tomber et être allongé constituent la décomposition d'un événement de chute commun, tandis que marcher est une activité probable pour la seconde personne présente.

La collecte s'est découpée entre les deux paires continues, marcher-marcher et marcher-être allongé qui ont été collectés en 2 longs échantillons d'une centaine de secondes, auxquels on enlève les quatre premières et dernières secondes, puisqu'on sépare en 35 échantillons de cinq secondes. La paire marcher-tomber a été collectée au coup par coup, en réalisant 35 fois consécutives une chute tandis que le second sujet marchait. Le début et la fin des échantillons ont également été retirés pour ne garder que la chute et non les mouvements depuis et vers l'ordinateur pour lancer et arrêter la collecte.

Dans un esprit de synthèse, on regroupe dans le tableau 3.1 le nombre d'échantillons et les activités associées dans le cadre de la collecte active.

Modification du protocole de collecte

Nous avions initialement collecté 20 échantillons pour chaque activité, les résultats obtenus n'étaient pas en accord avec les standards de la littérature.

Activité	Nombre de personnes	nombre d'échantillons
Aucune	0	35
Marcher	1	35
Se lever	1	35
S'asseoir	1	35
Tomber	1	35
Être allongé	1	35
Se pencher	1	35
Courir	1	35
Marcher	2	35
Marcher et tomber	2	35
Marcher et être allongé	2	35

Tableau 3.1 Table des données collectées dans le cadre de la collecte active

Devant ce résultat, nous avons collecté de nouveau des échantillons, afin d'arriver à 35 par activités. De plus, nous avons découplé les échantillons des activités s'asseoir et se lever pour être collectés séparément afin d'être sûr qu'il n'y ait aucun chevauchement entre les deux. Nous présentons dans le chapitre 5 les résultats de cette seconde collecte uniquement.

3.3 Collecte passive

L'objectif de notre première expérience utilisant la collecte passive est de réaliser de la détection de présence dans plusieurs environnements. Cela nous permettra d'analyser s'il est possible de détecter la présence d'une ou plusieurs personnes et ensuite de voir s'il est possible de généraliser la détection de présence d'un environnement à un autre.

La détection de présence s'inscrit dans le cadre de notre étude parce qu'elle répond à un double objectif d'économie d'énergie et de performance. En effet, en l'absence d'un sujet dans une pièce, il n'est pas pertinent de réaliser des calculs complexes en continu sur notre microcontrôleur. Il est plus pertinent de le mettre en veille pendant une période définie et recommencer la classification des activités si après cette période de veille, il y a de nouveau un sujet dans la pièce. De plus, la littérature ne fait pas état de la possibilité de généraliser un modèle performant pour la classification d'activités d'une seule personne à deux personnes. En l'absence de preuve allant dans ce sens, nous considérons que le même modèle ne peut être utilisée dans ces deux cas de figure. Ainsi, la classification du nombre de personnes nous permet de sélectionner le modèle à utiliser, dans l'éventualité où l'on en utilise un.

Nous aurions également souhaité utiliser la collecte passive pour réaliser la détection d'activités et plus précisément la détection de chutes. Cependant, nous avons abandonné ce but

devant les résultats obtenus pour la détection de présence, qui seront détaillés dans le chapitre 5.

3.3.1 Set-up expérimental

Pour réaliser la collecte passive, nous avons installé le projet correspondant, ESP32 CSI TOOL passive [48] sur un ESP32. Il agira comme receveur, dit Station.

Nous avons connecté un ordinateur portable en Wi-Fi au NDC. Nous avons utilisé cet ordinateur comme il le serait normalement afin de générer du trafic. Les utilisations réalisées sont le visionnage d'une vidéo et le téléchargement d'un fichier.

La différence fondamentale entre le set-up où un ordinateur est connecté à une carte Intel 5300 et celui que nous utilisons est que l'ordinateur est connecté en Wi-Fi et utilise la connexion à l'ESP32 comme un relai Wi-Fi. Il n'est pas en train de faire tourner un module de collecte dédié, mais est simplement utilisé dans des conditions normales. Nous avons utilisé un ordinateur, il faut noter qu'un téléphone portable peut remplir le même rôle. L'avantage de ces conditions expérimentales est la simplicité pour la mise en place. L'ESP32 agit comme un relai Wi-Fi auquel on se connecte, comme on le fait déjà quand on est chez soi et que l'on se connecte au réseau Wi-Fi. La grosse limite de cette expérience est que l'ordinateur ou le téléphone doivent être en cours d'utilisation dans l'objectif de générer du trafic. Pour réaliser de la détection de chute, cela est une hypothèse contraignante, bien qu'envisageable, pour de la simple détection de présence ce sont des conditions trop complexes pour un objectif atteignable de manières plus simples [16].

Nous avons réalisé la collecte dans trois pièces différentes, de tailles et formes différentes :

- La première pièce est un bureau, dont les mesures sont 3,90x6,90 mètres et meublé de deux tables, un canapé, quatre chaises et quatre ordinateurs. Deux des murs de ce bureau sont vitrés.
- La seconde est un couloir vitré également, attenant au bureau. Les dimensions du couloir sont 1,10x10 mètres.
- Le troisième environnement est une plus petite pièce de 2,20x2,90 mètres, meublé par un large bureau sur toute la longueur du mur. On y retrouve de nombreux câbles.

Les trois environnements, au-delà de l'aspect pratique puisqu'ils se situent au sein d'Humanitas Solutions, ont été choisis pour leur diversité spatiale apparente, tant par leurs tailles respectives que le mobilier. Une photographie du bureau, du couloir et de la salle des switchs est donnée aux figures 3.2, 3.3 et 3.4 respectivement.



Figure 3.2 Bureau



Figure 3.3 Couloir

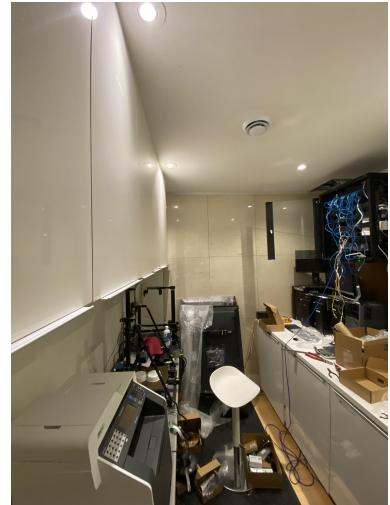


Figure 3.4 Salle des switchs

3.3.2 Données

Chaque échantillon correspond aux CSI collectés passivement pendant quatre secondes. Nous les avons collectés quand il n'y avait aucune personne dans les environnements sélectionnés, une personne puis deux personnes.

Nous avons développé un logiciel capable de collecter en continu pendant 2 minutes, après un délai initial de 5 secondes, et de séparer les paquets dans différents fichiers, correspondant aux 30 échantillons différents. Nous avons mis en place le délai de cinq secondes car le premier paquet était systématiquement plus court, ce que nous attribuons au temps pour que le premier message soit reçu qui peut être plus long, tandis que l'ouverture et la fermeture de socket se fait presque instantanément une fois la connexion établie. De plus, ce temps nous permet de quitter le lieu dans le cas où la pièce doit être vide.

Pour trois classes différentes, correspondant à aucune personne, une personne et deux personnes dans les trois lieux décrits, nous avons un total de 270 échantillons. Nous résumons cela dans le tableau 3.2.

Les échantillons sont tous composés de 64 parties réelles et imaginaires, comme dans la collecte précédente. On obtient finalement 53 amplitudes et phases.

Lieu	Nombre de personnes	nombre d'échantillons
Bureau	0	30
Bureau	1	30
Bureau	2	30
Couloir	0	30
Couloir	1	30
Couloir	2	30
Salle des Switch	0	30
Salle des Switch	1	30
Salle des Switch	2	30

Tableau 3.2 Table des données collectées dans le cadre de la collecte passive

CHAPITRE 4 PRÉTRAITEMENT

Au cours de ce chapitre, nous détaillerons quelques caractéristiques sur nos données comme la taille ou la répartition des valeurs. Nous décrirons ensuite les trois prétraitements que nous avons retenus dans le cadre de notre étude, qui sont le filtre Hampel, l'uniformisation de la taille des échantillons et l'usage de la normalisation lors de la représentation. Nous expliquerons enfin les trois représentations que nous avons testées, la représentation de toutes les amplitudes, la sélection de certains indicateurs pertinents et l'utilisation de la moyenne uniquement.

4.1 Caractéristiques de nos données

Avant d'entamer le prétraitement, nous regardons le profil de nos données. Cela nous permet de tirer des conclusions sur la méthode de collecte la plus fiable et d'entamer le prétraitement en prenant en compte ces considérations.

4.1.1 Sous-porteuses de garde

Parmi les 64 sous-porteuses que nous collectons, on retrouve des fréquences de garde. Dans toute la suite de notre étude, à moins de faire explicitement mention du contraire, nous retirerons les 11 sous-porteuses concernées. En effet, puisqu'elles sont communes à tous les échantillons que nous collectons, elles n'apportent aucune plus-value pour la classification. Au contraire, elles densifient le volume d'information sans présenter des distinctions classifiables. Elles peuvent donc avoir un effet antagoniste pour la classification que nous souhaitons réaliser.

4.1.2 Tailles des échantillons

Après avoir réalisé nos collectes de données, nous avons analysé les tailles des échantillons reçus, pour la collecte passive et la collecte active.

Cela nous permet de déterminer le taux d'échantillonnage dans le cadre de la collecte passive, en effet, nous n'avons pas de contrôle sur le taux d'envois dans ce mode de collecte. Nous souhaitons donc déterminer empiriquement quel est le meilleur taux d'échantillonnage afin de le comparer avec ce que l'on observe dans la littérature. De plus, en analysant la moyenne et l'écart type, cela nous permet d'analyser s'il existe une certaine cohérence entre les échantillons.

Dans le cas de la collecte active, cela nous permet de vérifier que la taille des échantillons reste similaire, ce qui est attendu, car le taux d'envoi devrait être constant puisque le matériel est programmé ainsi.

Collecte passive

On commence par analyser la taille des 270 échantillons, répartis comme décrit dans le tableau 4.1. Nous souhaitons déterminer si une tendance générale est observable avant de rentrer dans le détail des neuf classes. La répartition est donnée dans la figure 4.1.

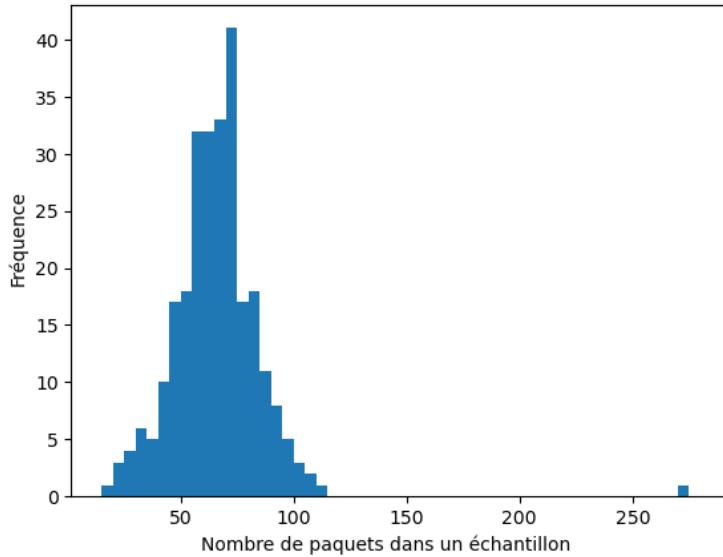


Figure 4.1 Répartition de la taille de tous les échantillons de la première collecte passive

Plusieurs choses sont notables au vu de la répartition de nos données :

- On détermine que l'échantillon contenant 271 paquets n'est pas représentatif et provient d'une erreur lors de la collecte, on ne le considérera pas lors de la suite de notre étude. L'erreur peut être un message d'arrêt qui n'est pas reçu, entraînant l'aggrégation de plusieurs échantillons différents.
- Le taux d'échantillonnage de nos données est très faible. La taille de nos échantillons varie de 15 à 114 paquets, avec une moyenne de 65. Cela constitue en moyenne un taux d'échantillonnage de 16 Hz, puisque chaque échantillon correspond à 4 secondes de collecte. Dans la littérature, le plus faible taux utilisé que nous avons observé est de 30 Hz [19], sinon il est assez commun de s'établir entre 50 et 100 Hz, une fréquence d'échantillonnage plus élevée peut être préférable [17]. La littérature n'indique pas comment traiter un taux d'échantillonnage aussi faible. Ainsi, nous faisons le choix de

supprimer tous les échantillons dont la fréquence d'échantillonage est inférieure à 10 Hz, correspondant à 10 paquets par seconde ou 40 paquets au total.

- l'écart type sur la longueur des échantillons est relativement important, car il s'établit à 21. Cependant, il est fortement impacté par notre échantillon composé de 271 paquets, nous vérifierons s'il se réduit après le retrait de celui-ci.

Nous faisons le choix de poursuivre l'analyse pour déterminer si les tailles sont réparties équitablement entre les différents cas considérés et entre les différentes collectes. Pour ce faire, on représente la moyenne et l'écart type des classes que nous avons collectées lors de la première prise de données que nous avons réalisée.

Lieu	Nombre de personnes	Tailles moyennes	Écart type
Bureau	0	61	9
Bureau	1	63	7
Couloir	0	63	12
Couloir	1	62	13
Salle des Switch	0	28	11
Salle des Switch	1	45	21

Tableau 4.1 Valeurs des moyennes et des écarts types du nombre de paquets lors de la première collecte de données passive

Les données collectées dans la salle des Switch ont une taille moyenne bien inférieure dans le deux cas, pour une pièce vide et une pièce avec une personne dedans. Nous attribuons cela à un défaut momentané de l'équipement, qui peut avoir été mal branché ou mal configuré. Nous décidons alors pour la poursuite de notre étude de réaliser une seconde collecte dans cette pièce seulement pour ces deux classes, afin de vérifier qu'il est possible de collecter des données reproductibles, car nous considérons que si les échantillons sont affectés par une corruption momentanée, il est possible que les amplitudes soient également affectées. Nous rajoutons également la classe avec deux personnes lors de la seconde collecte dans les trois environnements. Dans le tableau 4.2, on donne les valeurs des moyennes et des écarts type.

Lieu	Nombre de personnes	Tailles moyennes	Écart type
Salle des Switch	0	77	11
Salle des Switch	1	69	12
Couloir	2	72	18
Salle des Switch	2	72	19
Bureau	2	67	13

Tableau 4.2 Valeurs des moyennes et des écarts types du nombre de paquets lors de la seconde collecte de données passive

Les tailles moyennes lors de la seconde collecte sont toutes plus importantes que lors de la première, bien que cela ne soit pas statistiquement significatif puisque les intervalles de confiance se chevauchent, comme montré dans la figure 4.2. Dans cette figure, la moyenne est représentée par un point et l'écart type est représenté par les barres autour des points.

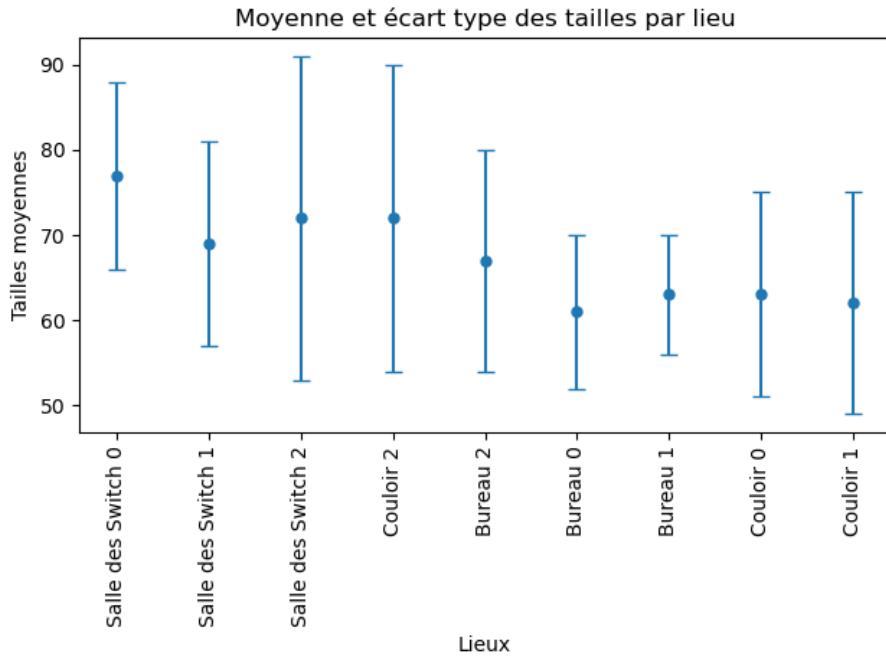


Figure 4.2 Diagramme en boîte de la taille des échantillons en fonction de leur classe et environnement après la seconde collecte

On a pris la décision de retirer les échantillons dont le nombre de paquets est inférieur à 40, car cela représente moins de 10 paquets par secondes. Il est trop aléatoire d'espérer observer des variations dans le CSI quand les paquets sont transmis à une fréquence aussi basse [21]. On détaille la distribution finale de la taille des paquets. Elle a une moyenne de 68 paquets, ce qui correspond à une fréquence d'échantillonnage de 17 Hz. Cela reste très inférieur à ce que l'on retrouve dans la littérature. L'écart type est de 13 paquets, on observe donc bien une diminution. Le jeu de données après la collecte passive initiale et le retrait des échantillons de taille aberrante est composé de 262 échantillons.

On représente la distribution finale dans la figure 4.3.

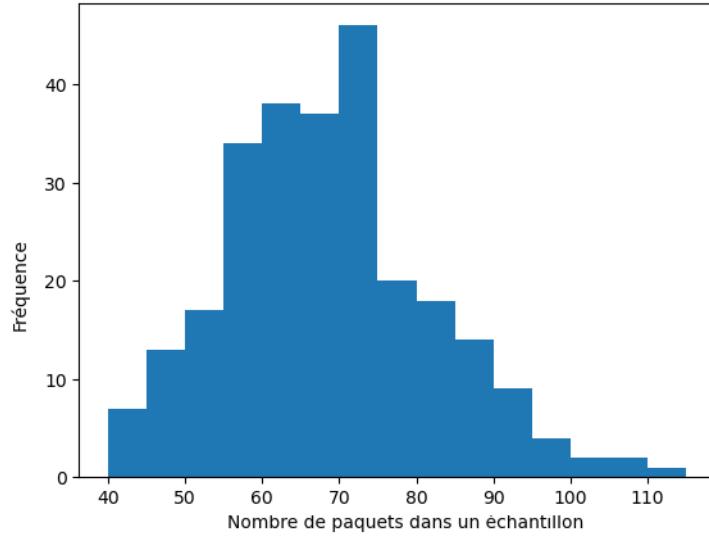


Figure 4.3 Répartition finale de la taille de tous les échantillons de la collecte passive

Collecte active

Dans le cas de la collecte active, la taille des échantillons est légèrement variable en fonction de l'activité. On reporte dans le tableau 4.3 les moyennes et écarts types pour chaque activité :

Activité	Nombre de personnes	Tailles moyennes	Écart type
Se lever	1	254	72
S'asseoir	1	402	37
Se pencher	1	257	84
être allongé	1	325	41
Marcher	1	283	58
Courir	1	306	42
Tomber	1	377	78
Marcher	2	361	22
Marcher et tomber	2	456	68
Marcher et être allongé	2	385	27
Total		341	67

Tableau 4.3 Valeurs des moyennes et des écarts types du nombre de paquets lors de la collecte de données active

On observe que la fréquence d'échantillonnage est plus élevée. Cela est un résultat que nous attendions, et qui s'avèrera sûrement plus prometteur pour les résultats des classifications. De plus, on note une certaine variabilité dans la taille des échantillons contenant des activités qui ont une durée qui n'est pas fixe, mais un écart type relativement faible. Pour toutes les ac-

tivités, la fréquence d'échantillonnage était la même, le nombre de paquets reçus dépend donc des rebonds sur le corps humain et l'environnement. On n'applique pas de transformation dans le cas de la collecte passive et conservons tous les échantillons.

4.1.3 Valeurs du CSI dans nos échantillons

On s'intéresse à la distribution des valeurs de l'amplitude du CSI. En effet, c'est un sujet très peu abordé qui nous semble pourtant pertinent. Il existe une approche fréquentielle pour distinguer quelles variations sont intéressantes [17] à classifier, mais à notre connaissance aucune approche se basant sur la prédominance des valeurs dans un jeu de données. Il est important de comprendre qu'en utilisant le ML toutes les variations n'auront pas un impact sur la classification [50]. De la même manière que les sous-porteuses de garde qui ont des valeurs constantes ne sont pas utiles, les modèles qui apparaissent trop souvent ou au contraire trop rarement ne sont pas significatifs [50]. En effet, un modèle peut trop s'entraîner et apprendre le moindre détail des données d'entraînement, ce que l'on appelle overfit. Un modèle de ML qui overfit n'arrive pas bien à généraliser, car il s'attache trop aux spécificités et ne discerne plus les tendances générales. Il est probable que les détails qu'il a appris dans son set d'entraînement soient trop précis et ne se retrouvent pas dans d'autres données. À l'inverse, sur des données qui sont toutes trop similaires, le modèle n'arrivera pas à discerner les différences suffisamment clairement pour déterminer des tendances, ce que l'on appelle underfit [51].

Nous souhaitons donc déterminer quelles sont les amplitudes les plus représentées, pour effectuer d'éventuels prétraitements par la suite. Nous souhaitons également observer les différences entre la collecte active et passive de cette manière, et éventuellement les différences au sein des classes.

Collecte passive

On s'intéresse dans un premier temps à la distribution des amplitudes en dB du CSI dans le cas de la collecte passive. On représente toutes les amplitudes dans la figure 4.4, un histogramme avec en abscisse les valeurs des amplitudes et en ordonnée la fréquence de leur apparition. On donne les fréquences d'apparition et les pas considérés dans l'annexe B.

Les valeurs sont comprises entre 3.61 et 47.76 dB, avec une moyenne à 21.66 dB et un écart type de 3.48 dB. On détermine les intervalles minimums pour représenter certains pourcentages de nos données :

- Intervalle contenant 99% des données est compris entre 29.22 et 12.88 dB.

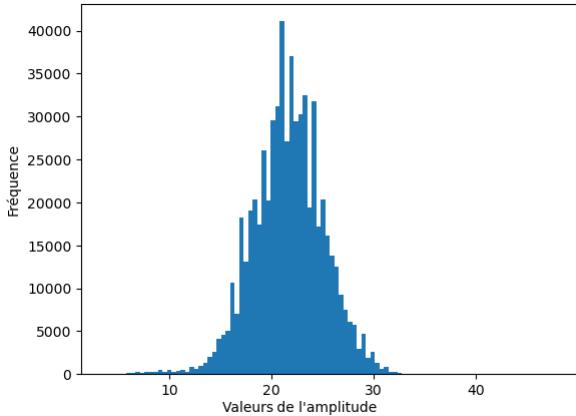


Figure 4.4 Répartition de la taille de tous les échantillons de la collecte passive

- Intervalle contenant 95% des données est compris entre 26.57 et 15.97 dB.
- Intervalle contenant 90% des données est compris entre 25.24 et 16.85 dB.
- Intervalle contenant 75% des données est compris entre 23.47 et 19.06 dB.

Il est intéressant de noter que 95% des données sont comprises dans un intervalle inférieur au quart de l'étendue totale des valeurs que prend l'amplitude des CSI collectés.

On veut déterminer si les valeurs des amplitudes suivent une distribution normale. Ce résultat nous intéresse pour mieux connaître empiriquement la distribution que suivent les amplitudes. Théoriquement nous n'attendons pas une distribution normale comme expliqué ci-après. On effectue donc un test de Shapiro-Wilk afin de mesurer la distance entre la distribution observée et une distribution normale. On réfute l'hypothèse H_0 car la p-value est inférieure à 10^{-13} , pour un critère de rejet tel que la p-value $< 0,05$. Les valeurs des amplitudes du CSI ne suivent pas une distribution normale. Ce qui est un résultat attendu [30] puisque les amplitudes du CSI sont impactées par de nombreux phénomènes physiques. Ces phénomènes peuvent être non centrés comme l'impact des multitrajets, leurs distributions peuvent être asymétriques comme l'atténuation qui a une queue de distribution plus longue. Enfin, la propagation des ondes radio dans un environnement reste toujours complexe et les variations spatiales impliquent des fluctuations importantes.

Collecte active

On représente désormais la distribution des valeurs du CSI lors de la collecte active. On cherche à déterminer si les amplitudes des CSI seront différentes. En effet, on sait que le CSI est très impacté par toutes les modifications spatiales au sein d'un même environnement et que son pouvoir de généralisation est faible [31]. On représente la distribution des valeurs de

l'amplitude du CSI dans le cas de la collecte active dans la figure 4.5. On rappelle qu'il faut distinguer la différence dans les valeurs absolues des amplitudes et les motifs ou tendances de variations, même si cette première information reste intéressante.

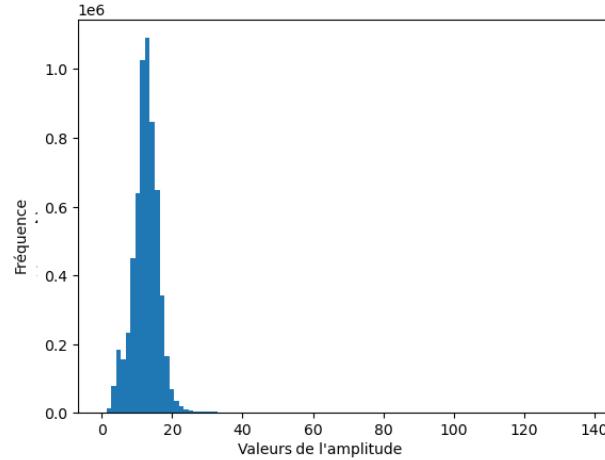


Figure 4.5 Répartition de la taille de tous les échantillons de la collecte passive

Les valeurs sont comprises entre 0.00 et 136.53 dB, avec une moyenne à 12.55 dB et un écart type de 4.00 dB. On détermine les intervalles minimums pour représenter certains pourcentages de nos données :

- Intervalle contenant 99% des données est compris entre 20.48 et 2.73 dB.
- Intervalle contenant 95% des données est compris entre 16.38 et 5.46 dB.
- Intervalle contenant 90% des données est compris entre 15.02 et 6.83 dB.
- Intervalle contenant 75% des données est compris entre 12.29 et 9.56 dB.

Il est intéressant de noter que 95% des données sont comprises dans un intervalle inférieur au dixième de l'étendue totale des valeurs que prend l'amplitude des CSI collectés.

On effectue de nouveau un test de Shapiro-Wilk. On réfute l'hypothèse H_0 car la p-value du test est inférieure à 10^{-18} , pour un critère de rejet fixé à p-value < 0,05. Les valeurs des amplitudes du CSI ne suivent une fois de plus pas une distribution normale.

4.2 Prétraitement

Dans toute la suite de notre étude à moins de faire mention du contraire, nous retirons les 11 sous-porteuses nulles. Elles sont effectivement utiles pour conserver l'orthogonalité des différents canaux, mais n'apportent aucun bénéfice lors de la classification, car elles ne varient pas avec le temps.

4.2.1 Filtre Hampel pour la détection d'anomalies

Au sein d'un même échantillon, les paquets ont des valeurs différentes. Ce comportement est parfaitement normal et témoigne des variations dans l'environnement. C'est grâce à ces variations qu'il est possible de différencier les activités.

On travaille cependant avec un système physique, qui peut être sujet à du bruit et des erreurs d'encodage. Ainsi, on note que certains paquets vont avoir une valeur quand on fait la moyenne de toutes les amplitudes des sous-porteuses bien supérieures aux moyennes des autres paquets au sein d'un même échantillon. On représente ce phénomène dans la figure 4.6.

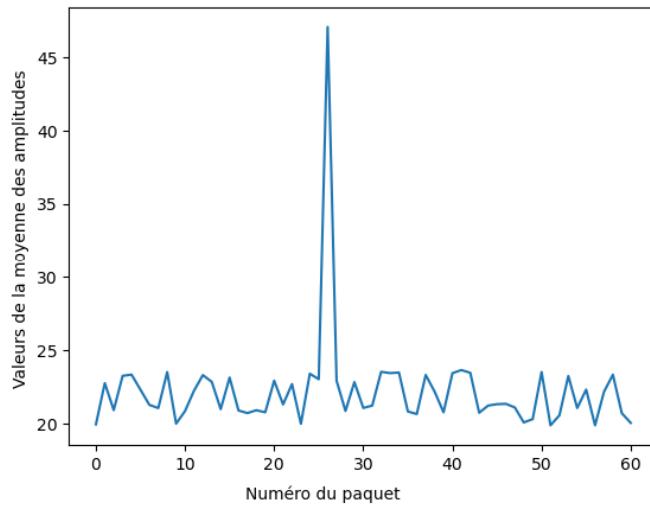


Figure 4.6 Valeurs moyennes des paquets successifs dans un échantillon

Nous proposons de remplacer les paquets aux valeurs moyennes si élevées, car ils ne contiennent pas une variation importante à classifier. Le plus probable est que ces paquets proviennent d'une erreur lors de la collecte. De plus, une telle augmentation peut être représentée de manière disproportionnée dans un algorithme de ML, car les valeurs extrêmes contraignent à restreindre les poids.

Pour détecter ces valeurs anormales, nous utilisons un filtre Hampel [52]. Nous appliquons à chaque paquet la contrainte sur sa moyenne telle que $\mu_i \in [\mu - \epsilon\sigma, \mu + \epsilon\sigma]$, où μ_i est la moyenne du paquet, μ est la médiane des moyennes sur l'échantillon total, σ est l'écart-type des moyennes sur l'échantillon total et ϵ est une constante. Ainsi, si la moyenne de ce paquet est en dehors de l'intervalle créé, on le remplace par le paquet médian. Cela permet de ne pas modifier la médiane et donc éviter une récursivité dans l'algorithme.

On considérera les valeurs $\epsilon=3$ et $\epsilon=1$ au cours de notre étude pour déterminer laquelle est la meilleure. On fera varier la présence du filtre et le niveau de la coupure car certains articles listent que le point faible du filtre Hampel est de couper des variations intéressantes. Ainsi, nous pourrons déterminer quel est le niveau optimal pour ce filtre.

On représente dans la figure 4.7 les valeurs moyennes du même échantillon présenté dans la figure 4.6 afin d'illustrer le fonctionnement du filtre Hampel.

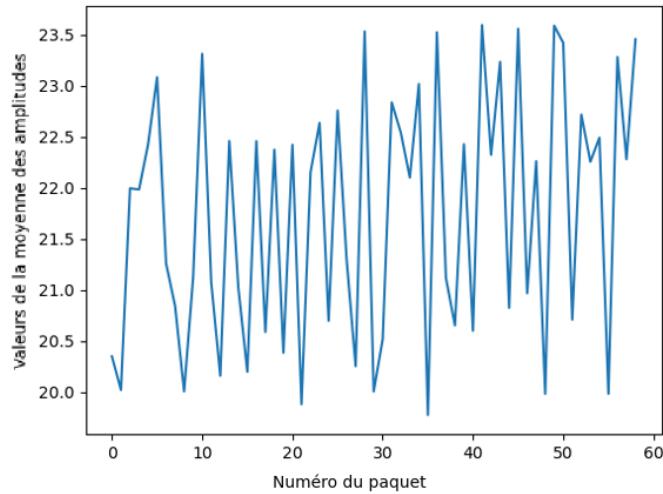


Figure 4.7 Valeurs moyennes des paquets successifs dans un échantillon après l'usage du filtre Hampel

4.2.2 Filtre passe bande sur les représentations des amplitudes

Comme mentionné précédemment, afin d'analyser efficacement le volume de données important que constituent les amplitudes du CSI sur 53 sous-porteuses et dans un nombre de paquets compris entre 40 et 500, nous utilisons des représentations graphiques. Plus précisément, nous représenterons dans toute la suite de notre étude les amplitudes par un gradient de couleur, allant du bleu foncé au rouge foncé. Le bleu foncé représente les amplitudes les plus faibles et le rouge foncé les amplitudes les plus élevées. On a en abscisse le numéro de la sous-porteuse, et en ordonnée les numéros décroissants des paquets. On aura donc un nombre compris entre 40 et 500 de lignes et toujours 53 colonnes, pour les 53 sous-porteuses variables.

À la représentation de nos échantillons, on note que la fenêtre relativement large des valeurs que prennent les amplitudes nuit à la représentation graphique. Pour le reformuler, le contraste est relativement faible, représenter le gradient des valeurs des échantillons sans

autre forme de prétraitement est problématique car :

- Dans le cas de la détection d'activités, l'intervalle des valeurs que nous représentons est théoriquement compris entre 0 et 136 dB. Il faut noter cependant que la proportion des valeurs supérieures à 20 dB représente moins de 1% des valeurs totales. Cependant, il est probable que des couleurs très similaires seront assignées aux valeurs 12 et 9 dB par exemple, puisque numériquement devant la taille de l'intervalle, elles sont proches. Cela est problématique, puisque l'on sait que 75% des données sont comprises entre 9 et 12 dB, il serait donc pertinent de pouvoir les distinguer correctement.
- Les échantillons ont tous des valeurs maximales et minimales différentes. Nous avons mis en évidence des statistiques sur la répartition totale des amplitudes, mais il existe également une variabilité au sein des échantillons. Pour donner un exemple, les valeurs de l'amplitude du CSI pour le premier échantillon de l'activité 'tomber' sont comprises entre $A_{mp} \in [8.06, 44.01]$. Pour le second échantillon, elles appartiendront à $A_{mp} \in [10.0, 27.20]$. Il n'existe donc pas de normalisation dans la représentation des amplitudes. La valeur absolue 10 dB peut correspondre à deux couleurs différentes d'un échantillon à un autre.

Pour illustrer le manque de contraste, nous représentons un échantillon de l'activité 'marcher' à la figure 4.8. On a en abscisse les différentes sous-porteuses et en ordonnée les paquets successifs. Chaque rectangle coloré va représenter une amplitude, avec les plus faibles en bleu foncé et les plus élevées en rouge foncé.

On indique par une flèche sur la figure 4.8 la présence de valeurs extrêmes, représentées par du rouge foncé. On observe très bien que la présence d'une amplitude élevée aplatis le contraste sur tout le reste de l'échantillon. Visuellement, il y a peu de contraste. Ce que nous appelons du contraste et que nous distinguons à l'œil nu se traduit par des valeurs RGB proches pour les pixels. Instinctivement, nous devinons qu'un algorithme de ML n'aura pas de très bonnes performances. En effet, tous les algorithmes de ML effectuent des opérations numériques ou logiques sur les valeurs qu'il a en entrées. Il est préférable que ses valeurs soient donc relativement distinctes.

Pour parer à ce problème, nous proposons une étape de prétraitement supplémentaire, qui consiste en limiter la valeur minimale et maximale associée aux couleurs bleu foncé et rouge foncé. Plus simplement, cela est équivalent à appliquer un filtre passe bande, en dessous d'une certaine amplitude les valeurs seront représentées comme le seuil par la couleur bleu foncé, au-dessus d'une certaine amplitude, elles seront représentées comme la borne maximale par du rouge.

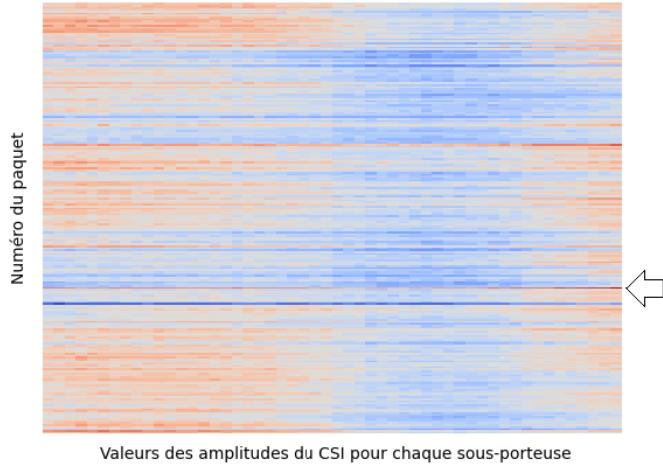


Figure 4.8 Représentation du gradient des amplitudes pour l'échantillon 15 de la classe marcher avant ajout du contraste

Nous représentons dans la figure 4.9 le même échantillon qu'à la figure précédente, après l'application de contrainte dans la représentation des valeurs, avec les bornes $A_{mp} \in [8.55, 18.00]$. On indique par une flèche le même paquet que précédemment.

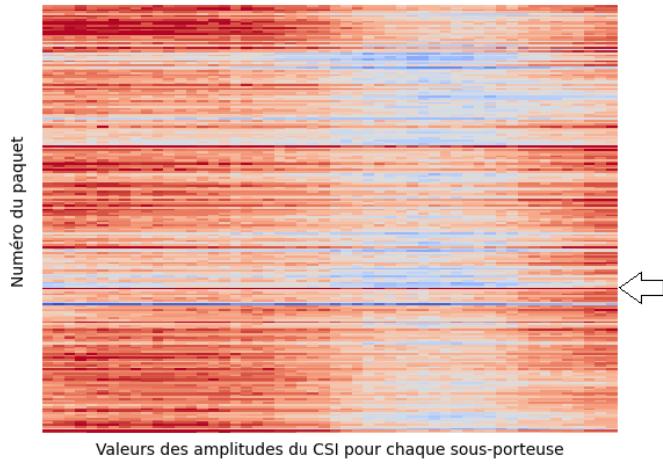


Figure 4.9 Représentation du gradient des amplitudes pour l'échantillon 15 de la classe marcher avant ajout du contraste

Cette étape de prétraitement revient essentiellement à augmenter la saturation. Comme mentionné auparavant, la saturation revient à représenter un plus large éventail de valeurs RGB. On justifie numériquement en relevant les valeurs RGB des amplitudes du CSI avant et après le prétraitement. Dans la première représentation à la figure 4.8, les valeurs RGB des amplitudes du CSI de la première sous-porteuse pour le premier et le deuxième paquet sont

respectivement (235,211,198) et (239, 207, 191). Dans la seconde, les valeurs RGB des amplitudes du CSI sur le premier et le deuxième paquet pour la première sous-porteuse sont respectivement (232,118,92) et (237,131,102). La variation en valeur absolue est de 15 avant le prétraitement, contre 28 après le prétraitement.

Nous faisons l'hypothèse que des valeurs RGB plus différentes au sein d'un même échantillon permettront de mieux distinguer les différences, et accroîtront donc la précision du CNN. Nous vérifierons cette hypothèse en appliquant ou non ce prétraitement.

Par ailleurs, nous souhaitons tester différentes bornes, de manière à déterminer s'il existe des amplitudes plus pertinentes que d'autres. Les premières valeurs du filtre à tester sont les bornes des intervalles contenant 99%, 95% et 75% des données. Nous souhaiterons ensuite effectuer éventuellement d'autres tests. À notre connaissance il n'existe pas de raisonnement sur les valeurs des amplitudes dans la littérature, nous souhaitons vérifier si couper les amplitudes trop faibles ou à l'inverse trop élevées pourrait être intéressant.

CHAPITRE 5 RÉSULTATS

Nous présentons dans ce chapitre les résultats des expériences que nous avons menées. La première section 5.1 sera dédiée à la présentation des algorithmes de traitement que nous utilisons et la répartition des échantillons entre le jeu de données pour l'entraînement et la validation. Nous donnons ensuite les résultats des expériences faites avec la méthode de collecte active. Enfin, nous détaillerons les résultats obtenus pour les expériences dans lesquelles nous avons utilisé une méthode de collecte passive.

Dans le cadre de la collecte de données active, nous avons effectué trois expériences de classification. Les expériences réalisées sont :

1. Détection de la présence d'aucune, une ou deux personnes.
2. Détection de l'activité réalisée par une personne.
3. Détection des activités réalisées par deux personnes.

On attend des résultats qui se rapprochent d'une précision parfaite pour la détection de présence. En effet, les standards de la littérature se rapprochent de cela avec une précision qui s'établit à 97.2% pour la détection de présence dans un environnement [53]. Pour la deuxième expérience, la littérature mentionne en l'absence de prétraitement et pour les sept activités que nous considérons une précision maximale de 92% [24]. Nous réaliserons la comparaison à toutes choses égales, si ce n'est le hardware utilisé, qui une fois de plus opposera la carte Intel5300 à l'ESP32. Une fois cette comparaison effectuée, nous ajouterons une phase de prétraitement pour déterminer si l'ajout de prétraitement améliore nos classifications. Enfin, la dernière expérience a pour objectif de déterminer s'il est possible de classifier les activités réalisées par plusieurs sujets au sein d'un même environnement. À notre connaissance, il n'existe pas de standards pour la détection d'activités de plusieurs personnes simultanément.

Dans le cadre de la collecte passive, nous avons mené les expériences suivantes :

1. Détection de la présence d'aucune ou une personne, avec le jeu de données d'entraînement et de validation collecté dans trois environnements différents.
2. Détection de la présence d'aucune ou une personne, avec le jeu de données d'entraînement constitué de deux environnements et celui de validation composé du troisième.

3. Détection de la présence d'aucune, une personne ou deux personnes, avec le jeu de données d'entraînement et de validation collecté dans trois environnements différents.

Au sein d'un seul environnement, la précision obtenue pour la détection de présence avec le CSI collecté passivement par un ESP32 oscille entre 99% et 100% [54]. Nous souhaitons rajouter à cela le fait d'entraîner et de valider dans plusieurs environnements pour utiliser un modèle uniformisé unique, au sein d'une maison par exemple. Au cours de la première expérience, nous réalisons la recherche du meilleur prétraitement. Nous faisons varier trois paramètres et considérons 13 prétraitements différents. En nous appuyant sur la précision obtenue sur le jeu de données de validation, nous déterminons quel est le meilleur prétraitement. Nous l'utilisons ensuite dans les deux expériences suivantes. Dans la seconde expérience, nous souhaitons déterminer s'il est possible de généraliser l'entraînement d'un lieu à un autre. À notre connaissance, il n'existe pas de réponses définitives à cette question dans la littérature. Enfin, nous complexifions la première expérience en rajoutant une classe avec deux personnes. L'objectif est de vérifier s'il est possible d'utiliser l'ESP32 pour compter sommairement combien de personnes se trouve dans différents lieux. On retrouve dans la littérature une précision de 97.2% pour le comptage de foule avec l'ESP32 en environnement interne [53]. La différence, au même titre que pour la première expérience, est que nous considérons trois lieux distincts.

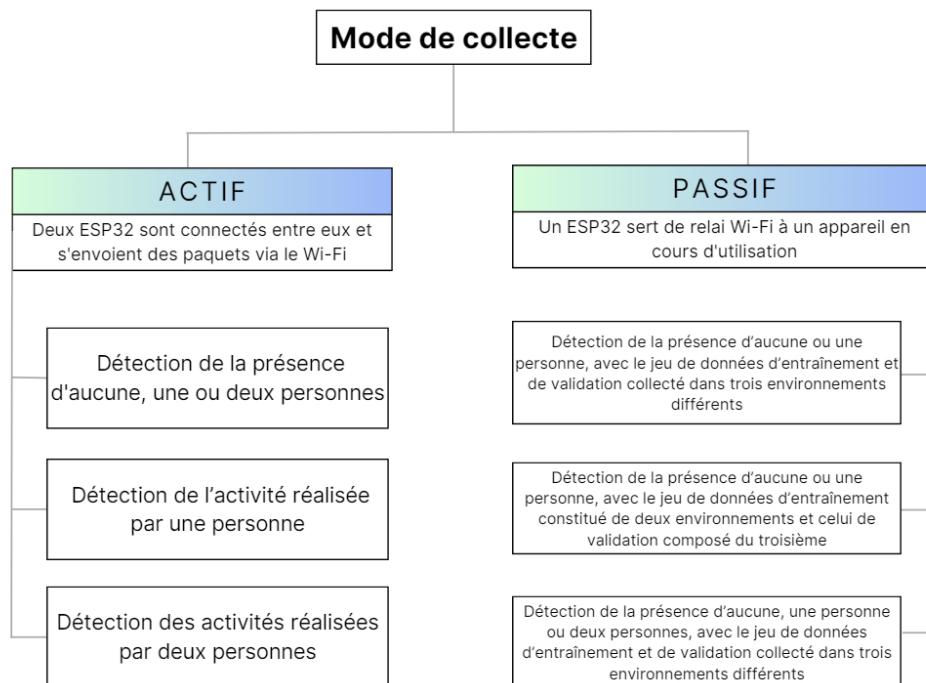


Figure 5.1 Différentes expériences considérées

5.1 Présentation des algorithmes de traitement

Avant de détailler les résultats, nous souhaitons donner quelques précisions sur les jeux de données utilisés et les algorithmes de traitement retenus.

Les données collectées de manière active sont réparties dans les proportions 25/10 (arrondies à 71/29) entre le set d'entraînement et de validation dans les trois expériences exécutées. Cela signifie que 25 images pour chaque classe seront allouées à l'entraînement et 10 à la validation.

Pour les expériences une et deux de la collecte passive, nous séparons notre jeu de données total dans les proportions 70/30. C'est-à-dire que 70% est dédié à l'entraînement et 30% à la validation. On sélectionne aléatoirement les échantillons qui vont servir à l'entraînement et à la validation, tout en conservant des proportions égales pour chaque classe et chaque environnement. Cela signifie que pour chaque classe retenue dans les expériences une et trois, il y a 63 échantillons qui sont dédiés à l'entraînement et 27 à la validation. Ils sont eux-mêmes équitablement répartis entre les trois environnements, ce qui représente 21 échantillons par environnement dans le jeu d'entraînement et 9 dans celui de validation. Dans le cas de la deuxième expérience, les échantillons sont répartis dans les proportions 120/60 (arrondies à 67/33), puisqu'on entraîne sur les 120 échantillons collectés dans deux environnements et nous validons sur les 60 appartenant au dernier.

Nous savons que la précision peut varier en fonction de l'algorithme de traitement retenu. Afin de ne pas être limité par cela, nous avons fait le choix d'utiliser quatre algorithmes de traitements différents. Nous reportons uniquement la meilleure précision obtenue dans les résultats que nous donnons ensuite. Nous précisons que tous les modèles sont construits de manière séquentielle en utilisant la bibliothèque Keras de Python, les tailles de filtre sont 3x3 et les fonctions d'activation pour les couches convolutionnelles sont ReLU. La métrique que nous considérons est la précision. Les quatre algorithmes sont des CNN. La première couche de chaque est une couche appelée "input" qui n'applique aucune opération. Les CNN sont tels que :

- Le premier est composé de 3 couches de convolution, suivies de deux couches entièrement connectées. La première utilise 16 filtres, la seconde 32 et la dernière, 16 de nouveau. Nous utilisons entre la première et la deuxième couche de convolution une couche de max pooling, qui sélectionne les composantes maximum et augmente donc la robustesse de notre CNN devant les petites variations tout en réduisant la dimensionnalité de nos données. Entre la deuxième et la troisième couche, nous utilisons un average pooling, qui permet d'obtenir une représentation plus globale de nos données

en les lissant. Enfin, nous utilisons de nouveau une couche de max pooling en sortie de notre troisième couche de convolution. Les données sont ensuite aplatis (flatten) et passent dans un réseau de neurones composé d'une couche avec 256 neurones et la couche finale contenant autant de neurones qu'il y a de classes dans notre expérience.

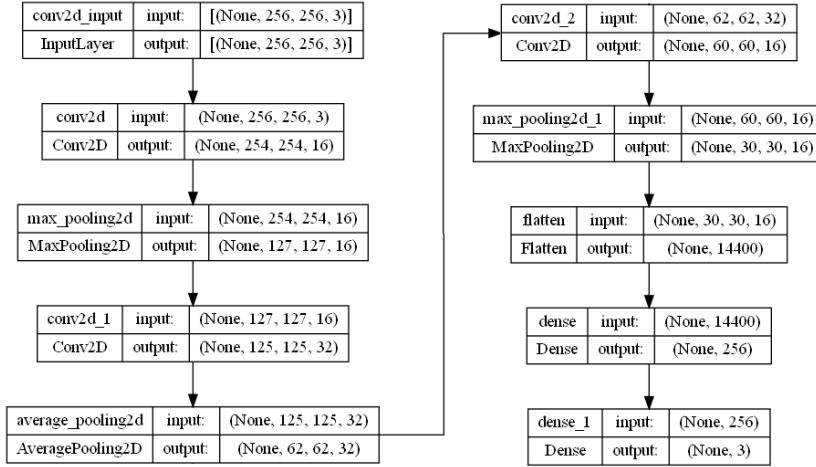


Figure 5.2 Architecture du premier CNN

- Le deuxième est composé de 4 couches de convolution, suivies de deux couches entièrement connectées. La première utilise 16 filtres, la deuxième et la troisième 32 et la dernière 16 de nouveau. Nous utilisons après la première et la dernière couche de convolution des couches de max pooling afin de conserver uniquement les caractéristiques les plus pertinentes pour la classification, et entre les autres couches de convolution un average pooling afin de garder quand même les variations les plus fines. Les valeurs sont ensuite aplatis et passent dans le même réseau de neurones que précédemment, composé d'une couche de 256 neurones et une couche avec autant de neurones qu'il y a de classes dans l'expérience.
- Le troisième est composé de 5 couches de convolution, dont les tailles de filtres sont 16, 32, 64, 32 et finalement 16 de nouveau. Les couches de pooling sont réparties comme précédemment avec d'abord un max pooling, ensuite des pooling qui gardent la moyenne des filtres et finalement avant que les données soient aplatis un max pooling de nouveau. Le réseau de neurones est également composé d'une couche de 256 neurones et la couche finale avec autant de neurones que de classes.
- Le quatrième et dernier CNN que nous utilisons pour obtenir les précisions associées à chaque expérience que nous allons réaliser est composé de 6 couches de convolution. Les tailles des filtres sont successivement 32, 32, 64, 64, 32 et de nouveau 32. Les

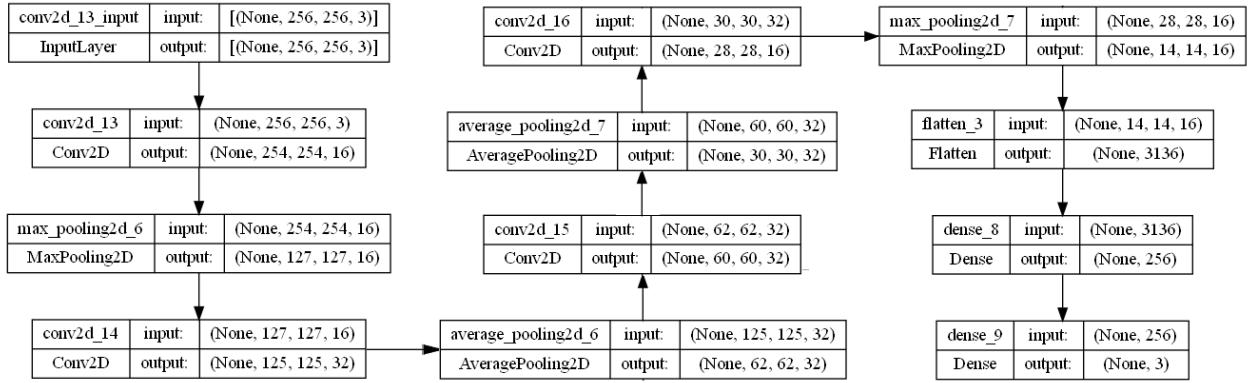


Figure 5.3 Architecture du deuxième CNN

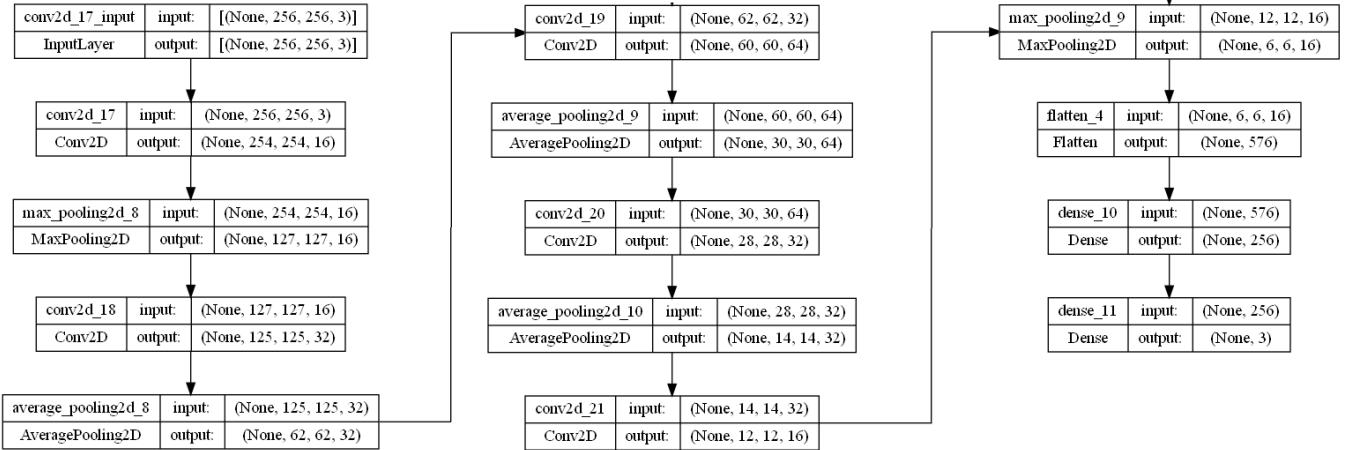


Figure 5.4 Architecture du troisième CNN

couches de pooling sont max pooling, quatre couches d'average pooling successives et finalement une couche de max pooling. On a rajouté après l'aplatissement un réseau de neurones avec 5 couches, constituées de 128, 256, 256, 128 et finalement 2,3 ou 7 neurones en fonction de l'expérience.

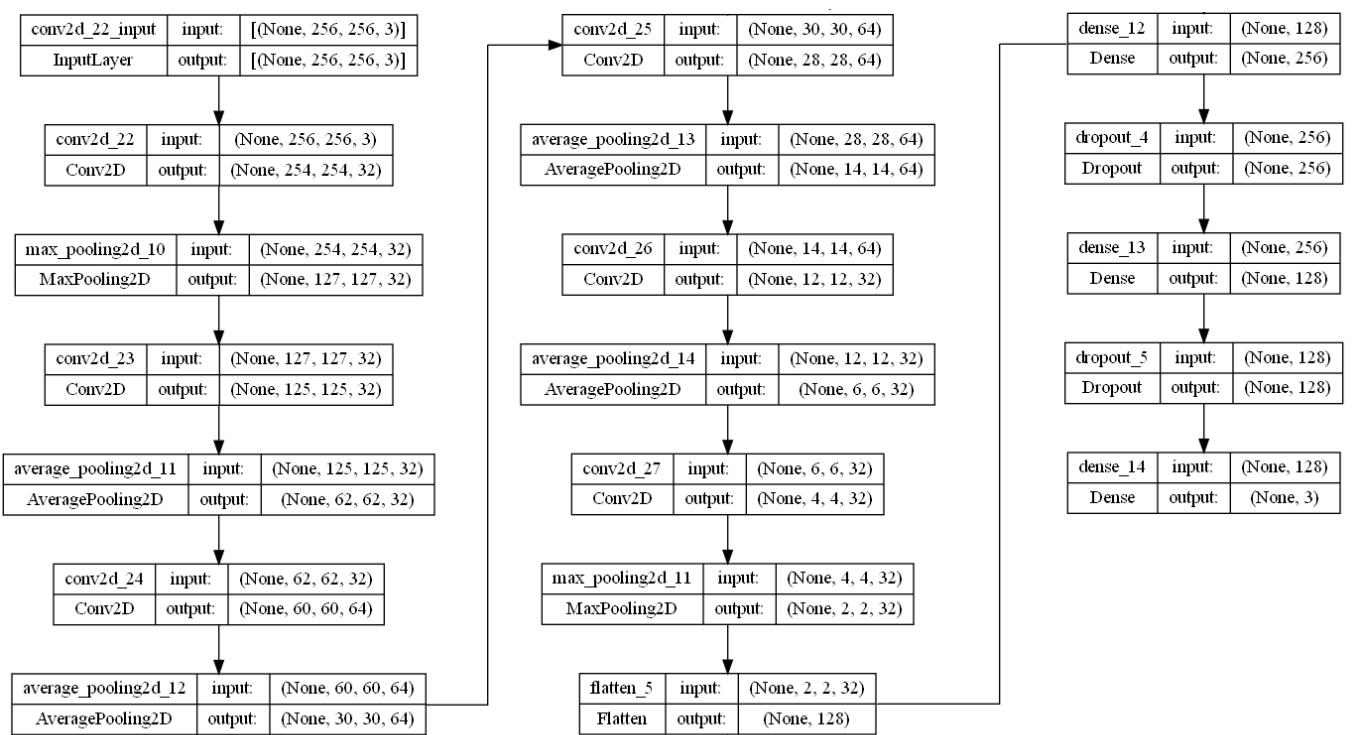


Figure 5.5 Architecture du quatrième CNN

5.2 Collecte active

5.2.1 Détection de la présence d'aucun, un ou deux personnes

Dans le cadre de la détection de présence d'aucune, une ou deux personnes avec un mode de collecte actif, on réalise d'abord la classification sur les 105 échantillons considérés sans prétraitement.

On atteint sur le set d'entraînement et de validation une précision de 100% avec tous les CNNs considérés. Les différents algorithmes ont des convergences plus ou moins rapides vers la précision maximale, s'étalant entre 12 et 20 époches. On représente à la figure 5.6 la courbe d'apprentissage du premier CNN composé de trois couches de convolution. Devant la précision atteinte, nous n'essayons pas d'appliquer de prétraitement à nos données, car nous ne pourrons pas observer leurs contributions relatives si elles sont positives.

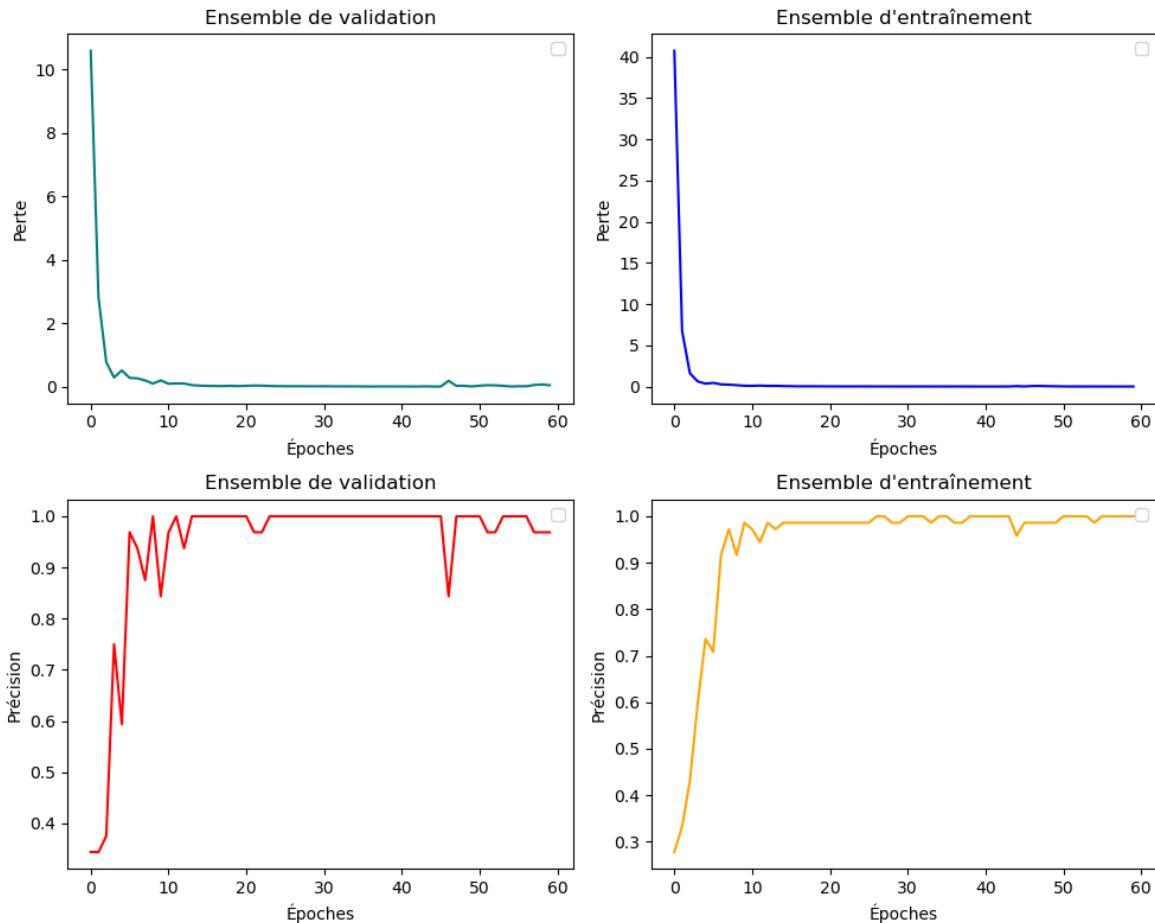


Figure 5.6 Précision et perte sur les ensembles d'entraînement et de validation pour la classification entre aucune, une ou deux personnes présente dans le cadre de la collecte active

Ce résultat est du même ordre de grandeur que ce que nous attendions, puisque nous le comparons avec une précision oscillant autour de 98% quand le comptage de foule est réalisé par deux ESP32 connectés activement [53]. Cela confirme que le CSI est un indicateur avec une granularité suffisante pour effectuer de la détection de présence et la distinction entre une et deux personnes. Par ailleurs, quand on compare la simple détection de présence réalisée avec une carte Intel 5300, on note que la carte ESP32 obtient une précision plus élevée pour une tâche plus complexe. On note tout de même que la différence est minime puisque la précision est systématiquement au-dessus de 95% avec la carte Intel 5300 pour de la détection de présence [27].

Ce résultat pour notre expérience de détection de présence complexifiée est excellent et nous permet d'envisager sereinement la poursuite de notre étude.

5.2.2 Détection de l'activité réalisée par une personne

Avec 20 échantillons par activités

Nous avions d'abord collecté 20 échantillons pour chacune des sept activités : s'asseoir, se lever, marcher, courir, tomber, se pencher, être allongé. En nous inspirant de travaux déjà réalisés dans le domaine de la détection de chutes, nous avons réalisé la classification avec aucun prétraitement.

Plus précisément, nous avons d'abord testé nos CNN sur les représentations des données brutes avant le retrait des sous-porteuses de garde et sans contraste ni filtre Hampel. La précision obtenue lors de cette classification est de 14%. Cela correspond à la random baseline, pour une classification entre sept classes, soit 1/7.

Ainsi, sans aucun prétraitement et avant le retrait des sous-porteuses de garde, nos algorithmes de traitement ne sont pas capables de mettre en évidence des corrélations entre nos données. Ce résultat nous surprend, car bien qu'elles soient plus compliquées à mettre en évidence, car noyées sous le volume d'informations, il existe normalement des corrélations entre les classes.

Cependant, pour remettre en perspective, on note que la première sous-porteuse a une valeur constante de 146 dB et les autres sous-porteuses de garde sont nulles, ce qui aplatis totalement les autres variations du CSI. Afin d'illustrer cela, on représente le même échantillon avant et après le retrait des sous-porteuses de garde 5.7.

La première sous-porteuse, tout à gauche de la représentation, est rouge foncé, puisque c'est la plus grande amplitude. Pour s'en convaincre, on peut revenir à la section 4.1.3, où il est indiqué que 99% des valeurs de l'amplitude du CSI sont comprises entre 20.48 et 2.73 dB.

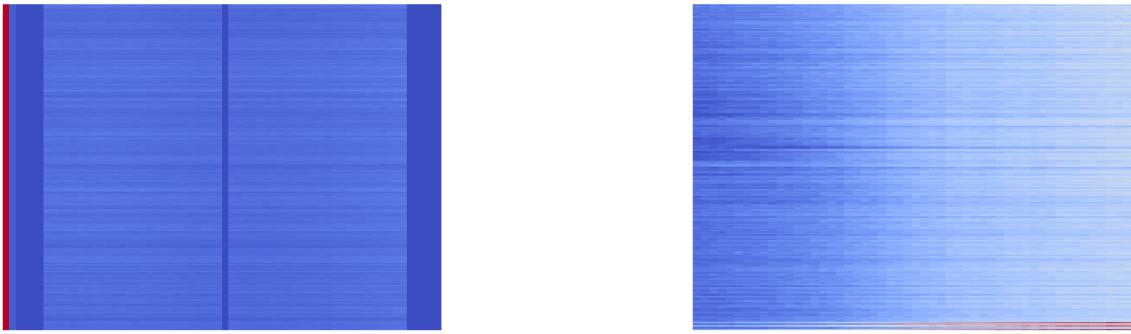


Figure 5.7 Échantillon avant et après le retrait des sous-porteuses de garde sans prétraitement

La valeur de 146 dB qui est constante sur la première sous-porteuse, considérée comme une sous-porteuse de garde, est donc bien supérieure et aplatis complètement le contraste. On représente à droite dans la figure le même échantillon une fois les sous-porteuses de garde retirées.

Nous réalisons désormais la classification en retirant les sous-porteuses de garde. Nous n'appliquons pas d'autre prétraitement. Les conditions sont très similaires à celles de l'article de référence [24], avec deux différences notables : le nombre d'échantillons et le matériel utilisé. Nous avons collecté 20 échantillons pour chaque activité tandis qu'il y en a 60 dans l'article en question. De plus, nous utilisons un ESP32 au lieu d'une carte Intel 5300. On obtient des classifications qui atteignent une précision de l'ordre de 40%.

Ce résultat est très inférieur à la précision obtenue dans l'article cité, qui s'établit à 92%. Nous supposons que le problème ne vient pas de l'ESP32 car il a été prouvé qu'il est possible de réaliser de la HAR avec [21]. Nous augmentons donc le nombre d'échantillons afin de vérifier que la différence de précision vient de ce facteur.

Avec 35 échantillons par activités

Nous avons identifié que la principale différence de méthodologie entre l'expérience que l'on vient de mener et la littérature est le nombre d'échantillons de chaque classe. Nous pensons qu'il est très probable d'obtenir de meilleurs résultats avec plus d'échantillons et faisons le choix de rajouter des échantillons pour chaque classe. Devant la longueur du processus de collecte, nous décidons dans un premier temps de rajouter 15 échantillons.

Nous réalisons la classification avec les 4 CNNs que nous avons décrit plus tôt sur les données non prétraitées. Nous donnons la courbe d'apprentissage à la figure 5.8.

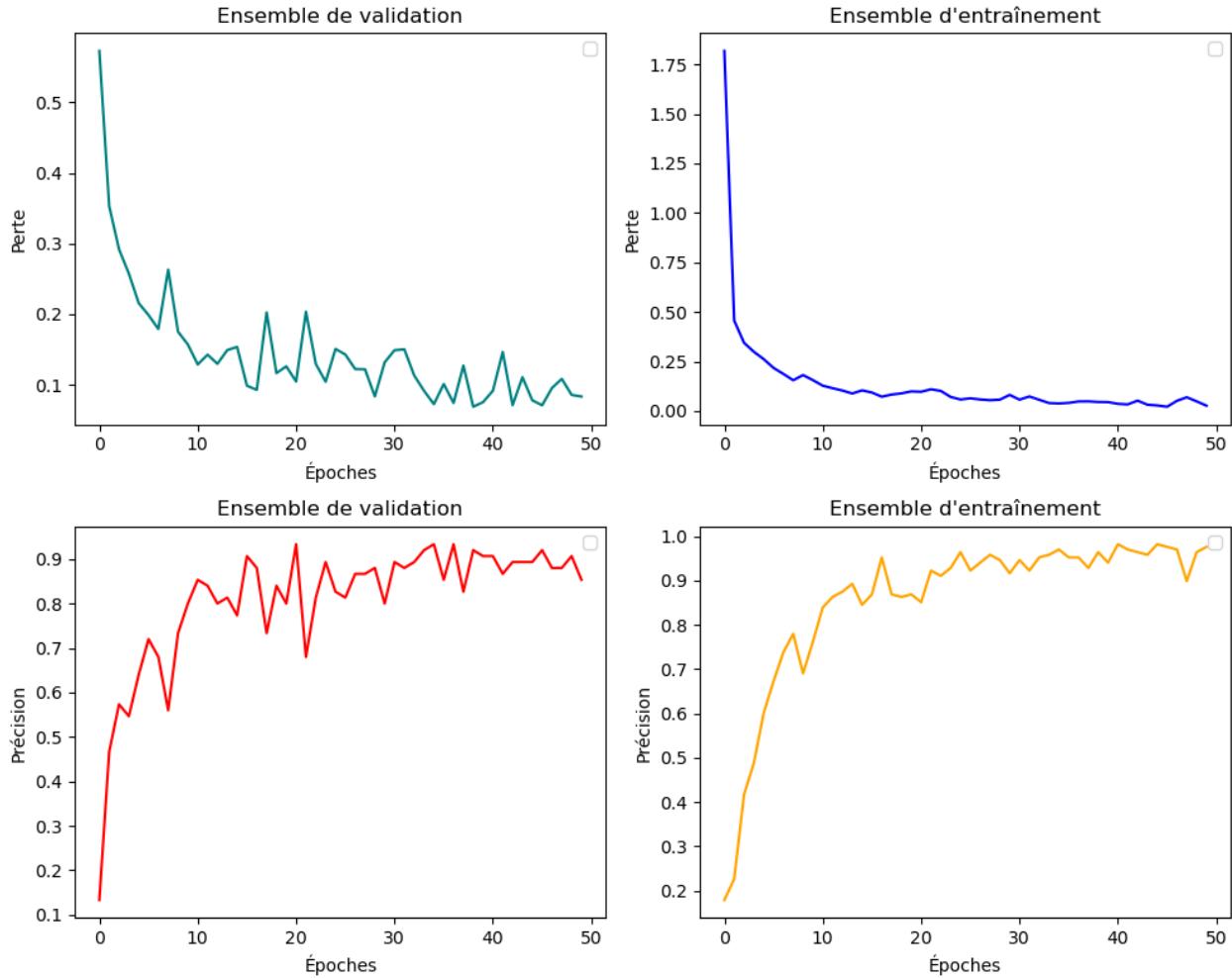


Figure 5.8 Précision et perte sur les ensembles d'entraînement et de validation pour une collecte active et la classification entre sept activités

La précision sur le set de validation atteint 93%. Cela est de l'ordre de grandeur que nous attendions, puisque pour la classification des sept mêmes activités collectées avec une carte Intel 5300 la précision obtenue est de 92%. On ne considère pas que l'augmentation de la précision est significative. En effet, notre jeu de données de validation contient 70 échantillons, une différence d'1% correspond donc à un seul échantillon.

On applique désormais différents prétraitements comme proposé dans le chapitre 4. Ici, appliquer un prétraitement est pertinent, car il y a une marge d'amélioration possible entre une précision de base de 93% et la précision optimale de 100%. On fait varier :

1. le critère de coupure du filtre Hampel, qui est soit absent, soit retire les paquets

aberrants avec un seuil de 1σ ou 3σ

2. La présence ou non d'un filtre passe bande à la représentation des données, augmentant le contraste lors de la représentation. On a considéré un cas sans cette contrainte, un cas avec cette contrainte entre les valeurs 5.38 et 16.46 dB correspondant à 95 % de la distribution, un autre avec cette contrainte entre les valeurs 6.55 et 12.55 dB correspondant aux valeurs d'amplitudes les plus faibles et finalement entre les valeurs 10 et 18 dB correspondant aux plus grandes amplitudes reçues.

Le premier cas, sans filtre Hampel ni contrainte sur les valeurs représentées, correspond à celui dont on a représenté les courbes d'apprentissage précédemment, car il n'y a pas de prétraitement. On résume dans le tableau 5.1 les 12 cas de prétraitement.

	Contrainte sur les valeurs	Filtre Hampel
cas 1	Aucune	
cas 2	6.55 à 12.55 dB	
cas 3	5.38 à 16.46 dB	
cas 4	10 à 18 dB	
cas 5	Aucune	1σ
cas 6	6.55 à 12.55 dB	1σ
cas 7	5.38 à 16.46 dB	1σ
cas 8	10 à 18 dB	1σ
cas 9	Aucune	3σ
cas 10	6.55 à 12.55 dB	3σ
cas 11	5.38 à 16.46 dB	3σ
cas 12	10 à 18 dB	3σ

Tableau 5.1 Cas de prétraitement considérés pour la détection d'activités avec la méthode de collecte active

On consigne ensuite dans le tableau 5.2 les meilleures précisions associées à chaque prétraitement avec les CNN que nous avons décrit précédemment :

Dans un premier temps, on observe qu'il peut y avoir une plus-value à prétraiter nos données. Ensuite, en l'absence de filtre Hampel la représentation la plus pertinente est celle qui constraint les valeurs sur l'intervalle qui contient 95% des valeurs. L'amélioration est notable.

Ensuite, en la présence d'un filtre Hampel la meilleure représentation est celle qui représente plus finement les valeurs d'amplitude les plus élevées. Les deux prétraitements sont complémentaires, car le filtre Hampel retire les valeurs — supérieures comme inférieures — aberrantes, ce qui augmente la robustesse de la classification qui sinon augmenterait le contraste sur ces valeurs. Une précision de 96% pour la classification de sept activités est un résultat qui s'aligne ou dépasse légèrement ce qui est fait dans la littérature.

	Précision
cas 1	93%
cas 2	94%
cas 3	95%
cas 4	92%
cas 5	92%
cas 6	91%
cas 7	95%
cas 8	96%
cas 9	93%
cas 10	95%
cas 11	92%
cas 12	96%

Tableau 5.2 Précisions obtenues

5.2.3 Détection de l'activité réalisée par deux personnes

Enfin, nous réalisons la classification de 35 échantillons de chaque couple d'activités réalisé par deux personnes. Les couples sont marcher-marcher, marcher-être allongé et marcher-tomber.

Nous considérons les deux cas de prétraitement qui ont obtenu les meilleures performances dans la section précédente. Nous représentons à la figure 5.9 les courbes d'apprentissage du cas 12 de prétraitement.

On obtient une précision maximale de 81%. Nous n'avons pas de référence dans la littérature pour comparer cette précision, que nous estimons trop basse pour permettre le développement d'une solution commerciale s'appuyant dessus. Nous attribuons la faiblesse de la précision aux réflexions multiples dues à la présence de deux personnes en mouvement. Cela est un défi intéressant pour le futur.

5.3 Collecte passive

5.3.1 Détection de la présence dans trois environnements

Le but de cette expérience est dans un premier temps de déterminer si la collecte passive permet de généraliser la détection de présence dans plusieurs environnements. En effet, il a été prouvé que l'ESP32 qui collecte des CSI passivement peut réaliser de la détection de présence [54] avec une précision qui varie entre 99 et 100%. Nous ne referons pas cette expérience et chercherons plutôt à rajouter une composante généralisatrice. On détermine qu'il peut être intéressant d'utiliser le même modèle pour une application simple à plusieurs

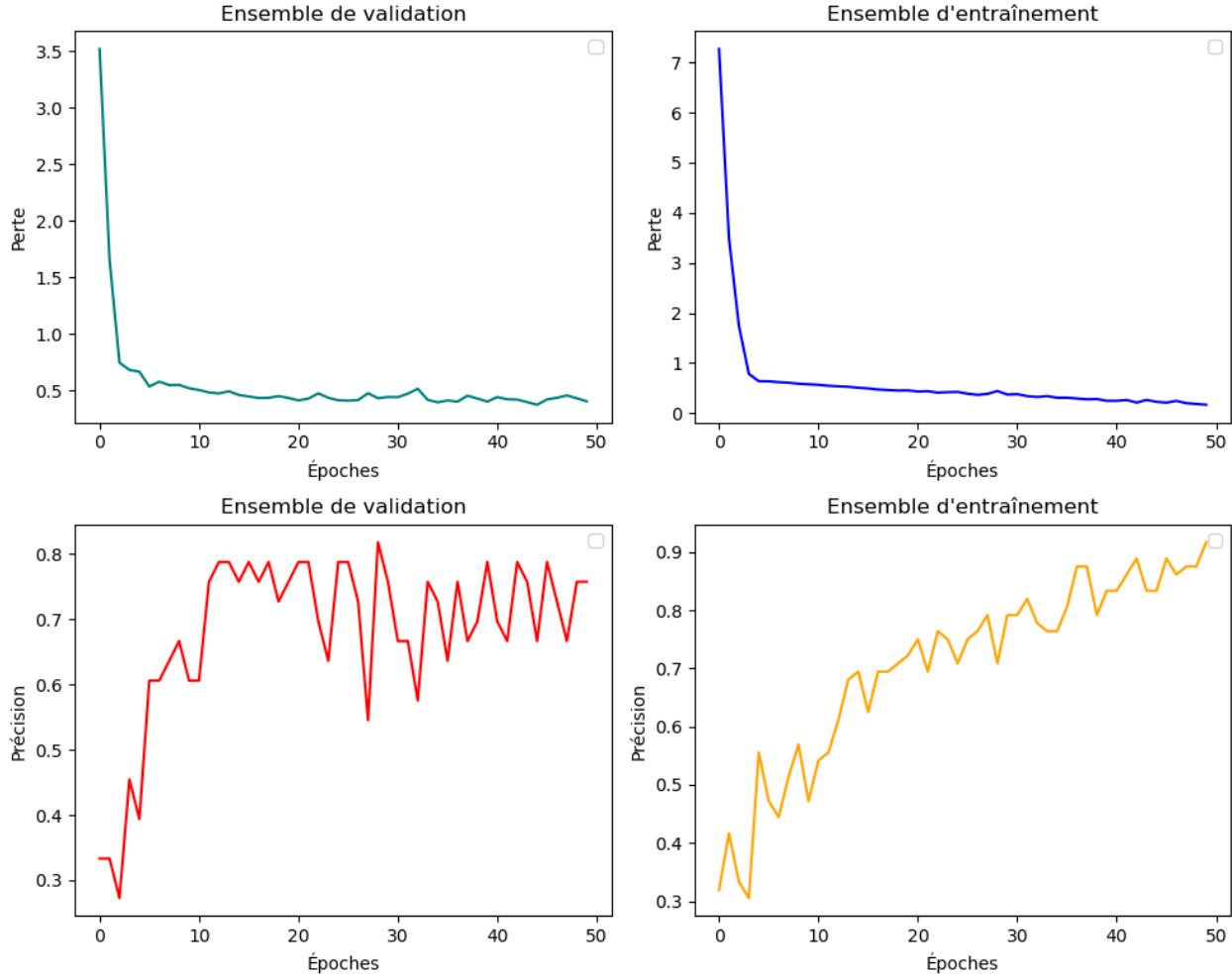


Figure 5.9 Précision et perte sur les ensembles d'entraînement et de validation pour une collecte active et la classification des activités de deux personnes

endroits. Par exemple, utiliser le même modèle de détection de présence dans toutes les pièces équipées de NDC dans l'hôpital. Cela permet lors du déploiement de ne pas avoir à considérer chaque NDC séparément, mais utiliser un modèle commun. À notre connaissance, il n'existe pas d'article qui utilise le même modèle de ML dans plusieurs environnements, car il a toujours été considéré que le pouvoir de généralisation du CSI est trop faible [31].

Ainsi, on entraîne les CNNs sur le jeu de données d'entraînement décrit à la section 5.1 sans prétraitement en dehors du retrait des sous-porteuses de garde. On obtient une précision de 96%. Cela signifie qu'en entraînant notre modèle de ML sur des pièces qui ont des caractéristiques différentes, le modèle peut tout de même déterminer les corrélations et, en l'occurrence, classifier correctement 52 des 54 échantillons de validation.

Puisqu'il existe une possibilité d'amélioration dans la précision de la classification, on considère différents prétraitements dans cette partie. Cela nous permet de comparer les contributions relatives de chaque prétraitement. Nous faisons varier :

1. le critère de coupure du filtre Hampel, qui est soit absent, soit retire les paquets aberrants avec un seuil de 1σ ou 3σ
2. La présence ou non d'un filtre passe bande à la représentation des données, augmentant le contraste lors de la représentation. On a considéré un cas sans cette contrainte, un cas avec cette contrainte entre les valeurs 15.97 et 26.57 dB correspondant à 95 % de la distribution, un autre avec cette contrainte entre les valeurs 12.88 et 21.66 dB correspondant aux valeurs d'amplitudes les plus faibles et finalement entre les valeurs 21.66 et 29.22 dB correspondant aux plus grandes amplitudes reçues.

En faisant varier deux critères qui peuvent prendre respectivement 3 et 4 valeurs, on obtient 12 cas de prétraitement différents. On rajoute également, afin de servir de témoin, les échantillons avec les sous-porteuses de garde, qui constituera le cas 13 de prétraitement. Dans un souci de clarté, nous explicitons les cas de prétraitement dans le tableau 5.3.

	Contrainte sur les valeurs	Filtre Hampel
cas 1	Aucune	
cas 2	15.97 et 26.57 dB	
cas 3	12.88 et 21.66 dB	
cas 4	21.66 et 29.22 dB	
cas 5	Aucune	1σ
cas 6	15.97 et 26.57 dB	1σ
cas 7	12.88 et 21.66 dB	1σ
cas 8	21.66 et 29.22 dB	1σ
cas 9	Aucune	3σ
cas 10	15.97 et 26.57 dB	3σ
cas 11	12.88 et 21.66 dB	3σ
cas 12	21.66 et 29.22 dB	3σ
cas 13	Aucune	

Tableau 5.3 Cas de prétraitement considérés

On prend en compte que notre jeu de données est légèrement biaisé puisque le couloir est attenant au bureau. Cependant, on ne considère pas cela problématique, puisque dans le cadre d'un déploiement à l'hôpital, certaines pièces seront également attenantes, d'autres non, donc notre jeu de données biaisé reproduit décemment les conditions expérimentales qui le seraient également.

On consigne ensuite dans le tableau 5.4 les meilleures précisions associées à chaque prétraitement avec les CNN que nous avons décrit précédemment. Cela nous permettra de comparer l'impact relatif des prétraitements.

	Précision
cas 1	96%
cas 2	96%
cas 3	98%
cas 4	98%
cas 5	98%
cas 6	96%
cas 7	96%
cas 8	96%
cas 9	96%
cas 10	96%
cas 11	96%
cas 12	96%
cas 13	92%

Tableau 5.4 Précisions obtenues

On tire plusieurs conclusions des précisions associées aux différents prétraitements. Déjà, on note qu'il y a effectivement une plus-value à retirer les sous porteuses de garde, ce qui est attendu. C'est un résultat que nous avions déjà relevé dans la section précédente 5.2. Ensuite, on note que les meilleures précisions sont pour les cas 3,4 et 5. Cela ne nous permet pas de tirer des conclusions définitives, car ces trois cas ont tous une représentation différente. De même, parmi ces 3 cas 2 n'utilisent pas de filtre Hampel tandis que le troisième utilise un filtre Hampel avec un critère de coupure à 1σ . Ainsi, on considère qu'il n'existe pas définitivement un prétraitement plus efficace que les autres sur la base des tests réalisés dans cette section.

Pour déterminer si l'un des prétraitements s'avère plus performant, on pourrait augmenter le nombre de points de données dans le set de validation.

5.3.2 Détection de la présence avec entraînement et validation dans des environnements différents

Nous souhaitons ici vérifier s'il est possible de généraliser la détection de présence d'un environnement à un autre. En effet, dans la littérature, un fingerprinting est nécessaire pour chaque nouvel environnement considéré. Cela alourdit considérablement tout déploiement potentiel. Il serait intéressant pour l'industrie de pouvoir limiter le nombre de fingerprinting nécessaire.

Nous tentons donc une nouvelle approche, pour tester s'il est possible de diminuer le nombre de fingerprinting requis au minimum. Nous allons entraîner notre modèle de ML sur deux environnements et valider le modèle sur un troisième disjoint. On détaille la répartition des lieux pour les sets d'entraînement et de validation dans le tableau 5.5.

	Entraînement	Validation
cas 1	Bureau et Couloir	Switch room
cas 2	Bureau et switch room	Couloir
cas 3	Couloir et switch room	Bureau

Tableau 5.5 Différents arrangements lors de l'entraînement et la validation sur différents sets de données

Il faut noter que le couloir est attenant au bureau, et séparé uniquement par un mur vitré. On s'attend donc à obtenir de bons résultats pour l'un quand on entraîne sur l'autre environnement. En effet, on fait l'hypothèse que les profils des CSI seront similaires, non pas identiques car collectés depuis des points différents, mais similaires. Nous ne faisons pas d'hypothèses quant à la précision dans la pièce des switchs quand nous entraînons nos algorithmes de ML sur les autres environnements, car nous n'avons pas de points de comparaisons dans la littérature.

On utilise les cas de prétraitement 3, 4 et 5 présentés dans la section précédente puisque ce sont ceux qui présentaient les meilleures performances. On reporte ensuite les précisions des classifications attenantes dans le tableau 5.6 pour le cas de prétraitement présentant la précision la plus élevée.

	Précision
cas 1	93 %
cas 2	100%
cas 3	100%

Tableau 5.6 Précisions obtenues

Les précisions sont effectivement très bonnes, 100%, quand on entraîne sur le bureau ou le couloir et vérifie sur l'autre. Cela était attendu. Le résultat dans le cas 1 est moins impressionnant, et dénote effectivement du pouvoir de généralisation relativement faible du CSI.

Notre étude est un premier pas dans la direction de la généralisation du CSI d'un environnement à un autre. Nous avons la preuve empirique qu'il est possible d'entraîner et de vérifier dans des lieux différents. Maintenant, il reste d'autres étapes à mettre en place dans des travaux futurs pour pouvoir déterminer combien de lieux différents il est nécessaire de fingerprinter pour réaliser de la détection de présence dans tout un domicile ou un hôpital.

5.3.3 Détection de la présence d'aucune, une ou deux personnes dans trois environnements différents

Nous attendons des performances proches ou légèrement supérieures à celles obtenues précédemment dans la section 5.3.1, car rajouter une troisième classe peut permettre à notre CNN de mieux distinguer les deux classes précédemment considérées.

Dans la littérature, le comptage de foule avec des ESP32 atteint une précision de 97.2% [53]. La différence majeure est que dans l'étude que nous citons, huit ESP32 sont utilisés tandis que nous n'en utilisons qu'un. Cela pourrait expliquer des performances moindres, ou alors justifier de diminuer le nombre d'ESP32 à utiliser dans des travaux futurs. On considère une fois de plus les trois cas de prétraitements qui fonctionnaient le mieux à la section 5.2.1.

Nous représentons les courbes d'apprentissage obtenues pour le meilleur prétraitement à la figure 5.10

On obtient effectivement des performances proches de celles obtenues dans la section 5.3.1, puisqu'elle s'établit à 93%. Cela infirme qu'ici rajouter une classe permet plus aisément de faire la distinction entre les deux classes considérées précédemment. Dans le cadre d'un déploiement industriel de la solution, il serait important d'améliorer ce résultat, car de la détection efficace du nombre de personnes présentes dépend le modèle de ML utilisé pour déterminer les activités. Se tromper dans le nombre de sujets dans un environnement et rater subséquemment la détection d'une chute par exemple pourrait s'avérer très problématique.

Dans un premier temps, on pourrait augmenter le nombre de points de données pour chaque classe afin de vérifier si les CNN sous-performent à cause de cela. Si malgré l'augmentation, nous n'obtenons toujours pas des performances satisfaisantes, il pourrait être intéressant d'envisager d'autres prétraitements. Nous n'envisageons pas d'augmenter le nombre d'ESP32 qui collectent puisque au-delà d'un ESP32, on peut réaliser une connexion active entre eux, qui présentent de très bonnes performances pour ce cas précis, comme montré à la section 5.2.1.

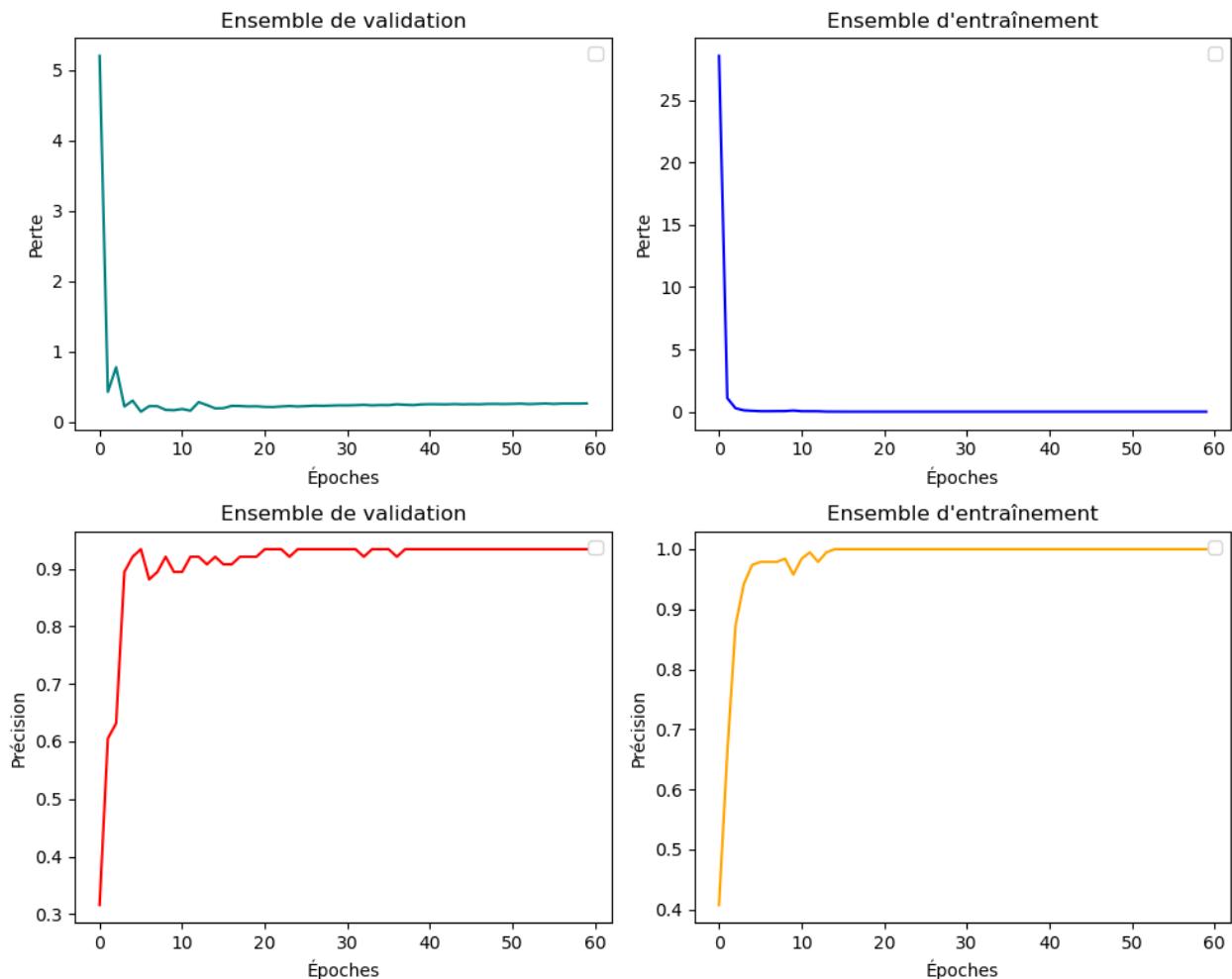


Figure 5.10 Précision et perte sur les ensembles d'entraînement et de validation pour une collecte passive et la détection de présence d'aucune, une ou deux personnes

CHAPITRE 6 CONCLUSION

Dans ce chapitre, nous proposons une synthèse des travaux effectués au long de ce mémoire. Nous mettrons en évidence les limitations de notre méthodologie et les possibilités d'améliorations futures pour approfondir le travail mené.

6.1 Synthèse des travaux

Le but principal de notre recherche était de détecter efficacement les chutes en environnement interne. Pour atteindre cet objectif, nous avons séparé notre travail en plusieurs parties :

- Analyser les composants permettant de collecter des CSI de manière à utiliser celui qui présente le plus d'avantages comparatifs. Nous avons sélectionné l'ESP32 qui permet de collecter 64 sous-porteuses, avec 11 sous-porteuses de garde, ce qui est plus que l'Intel 5300 largement utilisé dans la littérature, qui permet de collecter 30 sous-porteuses par paire RX-TX. L'ESP32 présente également l'avantage d'être un microcontrôleur et non une carte réseau, il peut donc effectuer des calculs directement sans le support d'un ordinateur et s'intègre relativement bien avec une distribution Linux.
- Collecter des CSI. Nous avons collecté des CSI dans 3 objectifs différents, détecter la présence et le nombre de personnes présente dans une pièce, différencier les activités d'un sujet parmi sept possibilités, différencier les activités de deux sujets parmi trois possibilités. Les CSI ont d'abord été collectés dans une configuration active, proche de ce qui existe déjà dans la littérature. Nous avons ensuite collecté des CSI dans une configuration passive qui présente l'avantage d'être très simple à déployer, puisqu'elle ne nécessite qu'un seul ESP32 et non deux. De plus, la connexion passive permet la captation des paquets d'un appareil personnel qui se connecte à l'ESP32 comme on se connecte à un routeur Wi-Fi.
- Prétraiter les CSI collectés. Nous avons d'abord effectué une sélection des échantillons dans le cadre de la collecte passive, puisque certains étaient de tailles soit trop faible, soit trop importante, puis complété ce jeu de données afin d'avoir de nouveau 30 échantillons par classe et par environnement. Nous avons ensuite détaillé un prétraitement communément utilisé dans la littérature, le filtre Hampel, et un prétraitement relativement nouveau à notre connaissance, le filtre passe bande sur les amplitudes,

que nous appliquons à nos données.

- Classifier nos données dans l'objectif de détecter une présence. La détection de présence permet de ne pas utiliser en continu un modèle de HAR quand il n'y a aucune activité à détecter, et répond donc au besoin de sobriété énergétique. Nous avons jugé intéressant de réaliser la détection de présence avec un modèle unique pour différents environnements. En effet, il a déjà été prouvé dans la littérature que la détection de présence par un ESP32 avec un mode de collecte passif fonctionne. Dans le cadre du développement d'une application industrielle, nous avons donc jugé intéressant de simplifier le déploiement en utilisant un modèle unique pour plusieurs lieux. Nous atteignons une précision de 98% quand on entraîne et valide dans les trois environnements considérés.
- Classifier nos données dans l'objectif de réaliser de la détection de présence sans entraînement préalable dans le lieu considéré. Cela constitue une approche relativement nouvelle, puisque pour l'instant peu d'articles existent sur le sujet. Limiter le nombre de collectes de données à réaliser pour détecter la présence dans un environnement interne complexe, composé de plusieurs pièces et couloirs est un objectif premier dans l'industrie actuellement, puisque le nombre de fingerprinting à réaliser est l'une des grosses limites de l'implémentation de solutions utilisant le CSI. Nous avons atteint une précision minimale de 93% en entraînant dans 2 environnements distincts et en validant dans le troisième.
- Classifier nos données dans l'objectif de classifier les activités d'une personne parmi sept activités : tomber, marcher, courir, s'asseoir, se lever, être allongé, se pencher. Nous avons mis en évidence que les meilleurs prétraitements utilisaient un filtre Ham-pel, avec un critère à 1σ ou 3σ , et une contrainte sur l'amplitude des valeurs afin de discerner plus finement les hautes amplitudes.
- Classifier nos données dans l'objectif de classifier les activités de deux personnes parmi trois couples d'activités : marcher et tomber, marcher et marcher, marcher et être allongé. Nous avons obtenu une précision maximale de 81%. Ce résultat était prévisible puisque rajouter une seconde personne en mouvement démultiplie le nombre de trajectoires possibles pour chaque sous-porteuse, compliquant la classification. À notre connaissance, il n'existe pas de standards auquel nous pouvons nous comparer.

Dans nos travaux, nous avons mis en évidence qu'en utilisant la détection active, il est possible de créer une application de détection de chutes relativement efficace, puisque nous sommes en mesure de détecter les activités à 98%. De plus, nous pouvons détecter efficacement la présence, afin de ne pas faire tourner en continu le module de HAR qui est gourmand en termes de calcul. La détection d'activité pour 2 personnes ne fonctionne pas parfaitement. Cependant, on peut argumenter qu'il n'est pas nécessaire de détecter les chutes dans ce cas puisque la seconde personne présente peut donner l'alerte. Nous prouvons également que dans tous les cas de figures pour la collecte active, appliquer une contrainte dans la représentation de nos amplitudes est bénéfique. Cela accélère la convergence dans le cas du comptage, qui converge de toute façon vers 100%, et augmente la précision dans les autres cas. Le filtre Hampel lui n'améliore pas systématiquement les résultats. Nous l'expliquons par deux raisons : il coupe certaines variations qui sont intéressantes à classifier et est redondant avec le filtre passe-bande appliqué sur les représentations qui aplatis déjà les valeurs extrêmes.

6.2 Limitations de la solution proposée

Comme tout travail qui essaye de reproduire des conditions de vie réelles avec la rigueur de la méthode scientifique, on retrouve de nombreuses limitations dans cette étude. Pour la détection d'activités, il existe bien d'autres activités que nous aurions pu rajouter. Les premières qui nous viennent à l'esprit sont être assis, s'allonger, se lever depuis la position allongée, lever un bras pour attraper quelque chose en hauteur, la position statique debout, etc. Le modèle que nous avons développé est relativement limité en ce sens. Par ailleurs, nous n'avons pas rajouté une catégorie 'autre' dans notre modèle, qui aurait pu contenir des échantillons d'autres mouvements.

Une autre limitation dans notre travail dans le cadre de la collecte passive est le fait d'utiliser des pièces attenantes séparées uniquement par une vitre. Cela rajoute nécessairement un biais, qui peut enfiler nos résultats de classification. Nous avons considéré que le biais est proche de celui que nous pouvons retrouver dans un hôpital, contenant également des pièces vitrées. Cependant, en toute rigueur, nous reconnaissions qu'utiliser des données corrélées peut biaiser nos classifications.

Enfin, dans les cas de détection de présence avec la collecte passive et de détection d'une activité, il faut noter que les jeux de données utilisés correspondent à la présence et aux activités réalisées par un seul et même sujet. En effet, dans la pratique, il était plus simple de collecter les activités d'une seule personne, plutôt que de se coordonner avec d'autres. Cela peut être problématique, car nous ne savons pas si les modèles que nous avons entraînés sont en mesure de réaliser des classifications, tant sur la présence que pour les activités si l'on

change de sujet.

6.3 Améliorations futures

Les pistes d'améliorations futures sont nombreuses. L'amélioration la plus évidente et la plus intéressante pour l'industrie est de réaliser le prétraitement et la classification sur l'ESP32 directement. Nous savons que le matériel est capable de réaliser cela [21] et pensons que c'est la prochaine étape de ce projet. Il reste de nombreux travaux dans le cadre de l'utilisation d'un modèle unique commun à plusieurs lieux pour la détection de présence. Le domaine est porteur, car utiliser un seul modèle facilite la mise en place de l'application puisqu'il n'est pas nécessaire de vérifier que chaque ESP32 utilise le modèle de la pièce dans laquelle il est. Cela peut sembler trivial, mais dans un hôpital comptant plusieurs centaines de chambres, c'est une opération fastidieuse. De même, nous avons mis en évidence qu'il est possible d'entraîner sur un nombre limité d'endroits pour en couvrir d'autres. Il reste des travaux à réaliser ici, puisque nous n'avons considéré que trois lieux. Cela limitait les nombres d'arrangements possibles. Dans des travaux futurs, nous pourrions considérer plusieurs chambres et plusieurs couloirs, par exemple, pour déterminer quel est le seuil du nombre de pièces dans lesquelles nous devons entraîner un algorithme réalisant de la détection de présence qui fonctionnerait ensuite partout.

RÉFÉRENCES

- [1] S. Alshammari *et al.*, “Falls among elderly and its relation with their health problems and surrounding environmental factors in riyadh,” *Journal of Family & Community Medicine*, vol. 25, p. 29–34, 01 2018.
- [2] L. S. Mubashir et L. Seed, “A survey on fall detection : Principles and approaches,” *Neurocomputing*, vol. 100, p. 144–152, janv. 2013.
- [3] L. Yang, Y. Ren et W. Zhang, “3d depth image analysis for indoor fall detection of elderly people,” *Digital Communications and Networks*, vol. 2, n°. 1, p. 24–34, 2016. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S2352864815000681>
- [4] K. Chouhan *et al.*, “Human fall detection analysis with image recognition using convolutional neural network approach,” dans *Proceedings of Trends in Electronics and Health Informatics*, M. S. Kaiser *et al.*, édit. Singapore : Springer Nature Singapore, 2022, p. 95–106.
- [5] J. R. Villar *et al.*, “Autonomous on-wrist acceleration-based fall detection systems : unsolved challenges,” *Neurocomputing*, vol. 452, p. 404–413, 2021. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0925231220317549>
- [6] F. J. S. Thilo *et al.*, “Usability of a wearable fall detection prototype from the perspective of older people—a real field testing approach,” *Journal of Clinical Nursing*, vol. 28, n°. 1-2, p. 310–320, 2019. [En ligne]. Disponible : <https://onlinelibrary.wiley.com/doi/abs/10.1111/jocn.14599>
- [7] Y. Wang, K. Wu et L. M. Ni, “Wifall : Device-free fall detection by wireless networks,” *IEEE Transactions on Mobile Computing*, vol. 16, n°. 2, p. 581–594, 2017.
- [8] “Eee standard for information technology— local and metropolitan area networks— specific requirements— part 11 : Wireless lan medium access control (mac)and physical layer (phy) specifications amendment 5 : Enhancements for higher throughput,” oct. 2009.
- [9] H. Wang *et al.*, “Rt-fall : A real-time and contactless fall detection system with commodity wifi devices,” *IEEE Transactions on Mobile Computing*, vol. 16, n°. 2, p. 511–526, 2017.
- [10] R. Reza, A. Alam et M. Islam, “Iot and wi-fi based door access control system using mobile application,” 11 2019, p. 21–24.
- [11] M. Albany *et al.*, “A review : Secure internet of thing system for smart houses,” *Procedia computer science*, p. 437–444, 03 2022.

- [12] F. Zafari, A. Gkelias et K. Leung, “A survey of indoor localization systems and technologies,” *IEEE Communications Surveys & Tutorials*, vol. PP, 09 2017.
- [13] Z. Tian *et al.*, “Rtil : A real-time indoor localization system by using angle of arrival of commodity wifi signal,” dans *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2019, p. 1–6.
- [14] P. Sirish Kumar et V. Srilatha Indira Dutt, “The global positioning system : Popular accuracy measures,” *Materials Today : Proceedings*, vol. 33, p. 4797–4801, 2020, international Conference on Nanotechnology : Ideas, Innovation and Industries. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S2214785320362520>
- [15] S. Sadowski et P. Spachos, “Rssi-based indoor localization with the internet of things,” *IEEE Access*, vol. 6, p. 30149–30161, 2018.
- [16] M. Habaebi et R. Rosli, “Rssi-based human presence detection system for energy saving automation,” *Indonesian Journal of Electrical Engineering and Informatics*, vol. 5, p. 339–350, 12 2017.
- [17] W. Wang *et al.*, “Understanding and modeling of wifi signal based human activity recognition,” dans *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, ser. MobiCom ’15. New York, NY, USA : Association for Computing Machinery, 2015, p. 65–76. [En ligne]. Disponible : <https://doi.org/10.1145/2789168.2790093>
- [18] S. Arshad *et al.*, “Wi-chase : A wifi based human activity recognition system for sensorless environments,” dans *2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2017, p. 1–6.
- [19] J. Yang *et al.*, “A framework for human activity recognition based on wifi csi signal enhancement,” *International Journal of Antennas and Propagation*, vol. 2021, p. 1–18, 02 2021.
- [20] G. Forbes, S. Massie et S. Craw, “Wifi-based human activity recognition using raspberry pi,” dans *2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI)*, 2020, p. 722–730.
- [21] S. M. Hernandez et E. Bulut, “Wifi sensing on the edge : Signal processing techniques and challenges for real-world systems,” *IEEE Communications Surveys & Tutorials*, vol. 25, n°. 1, p. 46–76, 2023.
- [22] I. A. Showmik, T. F. Sanam et H. Imtiaz, “Human activity recognition from wi-fi csi data using principal component-based wavelet cnn,” *Digital*

- Signal Processing*, vol. 138, p. 104056, 2023. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S1051200423001513>
- [23] S. Palipana *et al.*, “Falldefi : Ubiquitous fall detection using commodity wi-fi devices,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, n°. 4, jan 2018. [En ligne]. Disponible : <https://doi.org/10.1145/3161183>
- [24] P. Fard Moshiri *et al.*, “A csi-based human activity recognition using deep learning,” *Sensors*, vol. 21, 10 2021.
- [25] Y. Zhou *et al.*, “Human fall recognition based on wifi csi with dynamic subcarrier extraction of interference index,” *Journal of Physics : Conference Series*, vol. 1861, n°. 1, p. 012072, mar 2021. [En ligne]. Disponible : <https://dx.doi.org/10.1088/1742-6596/1861/1/012072>
- [26] Y. Zhang *et al.*, “Human activity recognition across scenes and categories based on csi,” *IEEE Transactions on Mobile Computing*, vol. 21, n°. 7, p. 2411–2420, 2022.
- [27] K. Qian *et al.*, “Pads : Passive detection of moving targets with dynamic speed using phy layer information,” vol. 2015, 12 2014.
- [28] F.-Y. Chu *et al.*, “Wifi csi-based device-free multi-room presence detection using conditional recurrent network,” dans *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, 2021, p. 1–5.
- [29] X. Cai *et al.*, “Identification and mitigation of nlos based on channel state information for indoor wifi localization,” 10 2015, p. 1–5.
- [30] K. Wu *et al.*, “Csi-based indoor localization,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, n°. 7, p. 1300–1309, 2013.
- [31] A. Sobehy, E. Renault et P. Muhlethaler, “Generalization aspect of accurate machine learning models for csi-based localization,” *Annals of Telecommunications*, vol. 77, 06 2021.
- [32] J. M. Rocamora *et al.*, “Survey of csi fingerprinting-based indoor positioning and mobility tracking systems,” *IET Signal Processing*, vol. 14, n°. 7, p. 407–419, 2020. [En ligne]. Disponible : <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-spr.2020.0028>
- [33] B. Wei *et al.*, “No need of data pre-processing : A general framework for radio-based device-free context awareness,” *ACM Trans. Internet Things*, vol. 2, n°. 4, aug 2021. [En ligne]. Disponible : <https://doi.org/10.1145/3467980>
- [34] R. Nandakumar, B. Kellogg et S. Gollakota, “Wi-fi gesture recognition on existing devices,” 2014.

- [35] H. F. Thariq Ahmed, H. Ahmad et A. C.V., “Device free human gesture recognition using wi-fi csi : A survey,” *Engineering Applications of Artificial Intelligence*, vol. 87, p. 103281, 2020. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0952197619302441>
- [36] H. Zou *et al.*, “Freecount : Device-free crowd counting with commodity wifi,” dans *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, 2017, p. 1–6.
- [37] W. Xi *et al.*, “Electronic frog eye : Counting crowd using wifi,” dans *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, 2014, p. 361–369.
- [38] S. Mosleh *et al.*, “Monitoring respiratory motion with wi-fi csi : Characterizing performance and the breathesmart algorithm,” *IEEE Access*, vol. 10, p. 131 932–131 951, 2022.
- [39] X. Wang, C. Yang et S. Mao, “On csi-based vital sign monitoring using commodity wifi,” *ACM Trans. Comput. Healthcare*, vol. 1, n°. 3, may 2020. [En ligne]. Disponible : <https://doi.org/10.1145/3377165>
- [40] J. C. Soto *et al.*, “A survey on vital signs monitoring based on wi-fi csi data,” *Computer Communications*, vol. 195, p. 99–110, 2022. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0140366422003127>
- [41] Y. Zeng *et al.*, “Fullbreath : Full human respiration detection exploiting complementarity of csi phase and amplitude of wifi signals,” *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, n°. 3, sep 2018. [En ligne]. Disponible : <https://doi.org/10.1145/3264958>
- [42] S. Zhong *et al.*, “Wi-fire : Device-free fire detection using wifi networks,” 05 2017, p. 1–6.
- [43] S. Jian, S. Ishida et Y. Arakawa, “Initial attempt on wi-fi csi based vibration sensing for factory equipment fault detection,” dans *Adjunct Proceedings of the 2021 International Conference on Distributed Computing and Networking*, ser. ICDCN ’21. New York, NY, USA : Association for Computing Machinery, 2021, p. 163–168. [En ligne]. Disponible : <https://doi.org/10.1145/3427477.3429462>
- [44] W. Yang *et al.*, “Wi-wheat : Contact-free wheat moisture detection with commodity wifi,” dans *2018 IEEE International Conference on Communications (ICC)*, 2018, p. 1–6.
- [45] D. Halperin. (2010) Github. [En ligne]. Disponible : <https://github.com/dhalperi/linux-80211n-csitool/>
- [46] Y. Xie, Z. Li et M. Li, “Precise power delay profiling with commodity wifi,” dans *Proceedings of the 21st Annual International Conference on Mobile Computing and*

- Networking*, ser. MobiCom '15. New York, NY, USA : ACM, 2015, p. 53–64. [En ligne]. Disponible : <http://doi.acm.org/10.1145/2789168.2790124>
- [47] M. Schulz, D. Wegemer et M. Hollick. (2017) Nexmon : The c-based firmware patching framework. [En ligne]. Disponible : <https://nexmon.org>
- [48] S. M. Hernandez et E. Bulut, “Lightweight and Standalone IoT Based WiFi Sensing for Active Repositioning and Mobility,” dans *21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)* (WoWMoM 2020), Cork, Ireland, juin 2020.
- [49] J. M. Johnson, “Survey on deep learning with class imbalance,” *Journal of Big Data*, vol. 6, p. 27, 03 2019.
- [50] M. Tamir et S. Elay, “Machine understanding and deep learning representation,” *Synthese*, January 2023. [En ligne]. Disponible : <http://philsci-archive.pitt.edu/21482/>
- [51] X. Guyon et J. feng Yao, “On the underfitting and overfitting sets of models chosen by order selection criteria,” *Journal of Multivariate Analysis*, vol. 70, n°. 2, p. 221–249, 1999. [En ligne]. Disponible : <https://www.sciencedirect.com/science/article/pii/S0047259X99918286>
- [52] R. Pearson *et al.*, “Generalized hampel filters,” *EURASIP Journal on Advances in Signal Processing*, vol. 2016, 08 2016.
- [53] H. Choi *et al.*, “Wi-cal : Wifi sensing and machine learning based device-free crowd counting and localization,” *IEEE Access*, vol. 10, p. 24 395–24 410, 2022.
- [54] A. Natarajan, V. Krishnasamy et M. Singh, “A machine learning approach to passive human motion detection using wifi measurements from commodity iot devices,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, p. 1–10, 2023.

ANNEXE A EXEMPLE D'UN PAQUET CSI

```
CSI_DATA,CO:EE:40:80:3B:1A,[67.0522, 1, 12.0416, 14, 12, 12.0416, 12.0416, 13,  
12.1655, 11.0454, 12.1655, 11.4018, 11.1803, 11.4018, 11.4018, 10.198, 11.1803,  
11.1803, 10.198, 10.0499, 10.0499, 9.05539, 9, 8, 8.06226, 8.06226, 8.24621, 0,  
0, 0, 0, 0, 0, 0, 0, 9.05539, 9, 9.05539, 10.198, 9.48683, 10.4403, 9.48683,  
10.7703, 10.7703, 10.7703, 10.7703, 11.1803, 12.083, 11.1803, 11.1803, 11.7047,  
12.6491, 12.6491, 12.6491, 11.7047, 12.3693, 12.3693, 12.3693, 13.1529, 13.1529,  
13.0384],[0.302885, 1.5708, 3.05845, 3.14159, 3.14159, 3.05845, -3.05845, 3.14159,  
-2.97644, -3.05093, -2.97644, -2.87534, -2.96174, -2.87534, -2.87534, -2.9442,  
-2.96174, -2.96174, -2.9442, -3.04192, -3.04192, -3.03094, 3.14159, 3.14159, 3.01724,  
3.01724, 2.89661, 0, 0, 0, 0, 0, 0, 0, 0, 0, -3.03094, 3.14159, 3.03094,  
2.9442, 2.81984, 2.85014, 2.81984, 2.76109, 2.76109, 2.76109, 2.76109, 2.67795,  
2.71497, 2.67795, 2.67795, 2.79282, 2.81984, 2.81984, 2.81984, 2.79282, 2.89661,  
2.89661, 2.89661, 2.98894, 2.98894, 3.06482]
```

ANNEXE B RÉPARTITION DES VALEURS TOTALES

Collecte passive : (counts([8.0000e+00, 2.6000e+01, 0.0000e+00, 7.5000e+01, 4.7000e+01, 1.6300e+02, 1.6600e+02, 2.8800e+02, 2.0100e+02, 2.2500e+02, 3.1000e+02, 3.0100e+02, 5.2600e+02, 2.3100e+02, 4.8200e+02, 3.0300e+02, 3.2500e+02, 4.9800e+02, 2.1300e+02, 8.5000e+02, 6.1100e+02, 9.1000e+02, 1.3290e+03, 2.0740e+03, 2.5590e+03, 4.1560e+03, 4.5780e+03, 5.0330e+03, 1.0683e+04, 7.0150e+03, 1.8201e+04, 1.3088e+04, 1.9073e+04, 2.0299e+04, 1.7442e+04, 2.6039e+04, 2.0239e+04, 2.9540e+04, 3.1160e+04, 4.1089e+04, 2.7115e+04, 3.7010e+04, 2.9438e+04, 3.0249e+04, 3.2532e+04, 1.9458e+04, 3.1770e+04, 1.7141e+04, 2.0390e+04, 1.6100e+04, 1.3820e+04, 1.2465e+04, 9.1970e+03, 7.5080e+03, 6.0670e+03, 5.7050e+03, 2.9080e+03, 4.7150e+03, 1.9050e+03, 2.5940e+03, 1.2740e+03, 6.2400e+02, 8.2900e+02, 2.0800e+02, 2.4300e+02, 1.1000e+02, 6.7000e+01, 7.1000e+01, 3.9000e+01, 2.7000e+01, 1.8000e+01, 2.1000e+01, 1.4000e+01, 8.0000e+00, 1.4000e+01, 1.3000e+01, 1.2000e+01, 1.5000e+01, 1.0000e+01, 5.0000e+00, 8.0000e+00, 1.1000e+01, 8.0000e+00, 8.0000e+00, 1.2000e+01, 5.0000e+00, 4.0000e+00, 2.0000e+00, 7.0000e+00, 6.0000e+00, 0.0000e+00, 5.0000e+00, 3.0000e+00, 1.0000e+00, 3.0000e+00, 1.0000e+00, 2.0000e+00, 0.0000e+00, 0.0000e+00, 2.0000e+00]), bins([3.60555 , 4.0470925 , 4.488635 , 4.9301775 , 5.37172 , 5.8132625 , 6.254805 , 6.6963475 , 7.13789 , 7.5794325 , 8.020975 , 8.4625175 , 8.90406 , 9.3456025 , 9.787145 , 10.2286875 , 10.67023 , 11.1117725 , 11.553315 , 11.9948575 , 12.4364 , 12.8779425 , 13.319485 , 13.7610275 , 14.20257 , 14.6441125 , 15.085655 , 15.5271975 , 15.96874 , 16.4102825 , 16.851825 , 17.2933675 , 17.73491 , 18.1764525 , 18.617995 , 19.0595375 , 19.50108 , 19.9426225 , 20.384165 , 20.8257075 , 21.26725 , 21.7087925 , 22.150335 , 22.5918775 , 23.03342 , 23.4749625 , 23.916505 , 24.3580475 , 24.79959 , 25.2411325 , 25.682675 , 26.1242175 , 26.56576 , 27.0073025 , 27.448845 , 27.8903875 , 28.33193 , 28.7734725 , 29.215015 , 29.6565575 , 30.0981 , 30.5396425 , 30.981185 , 31.4227275 , 31.86427 , 32.3058125 , 32.747355 , 33.1888975 , 33.63044 , 34.0719825 , 34.513525 , 34.9550675 , 35.39661 , 35.8381525 , 36.279695 , 36.7212375 , 37.16278 , 37.6043225 , 38.045865 , 38.4874075 , 38.92895 , 39.3704925 , 39.812035 , 40.2535775 , 40.69512 , 41.1366625 , 41.578205 , 42.0197475 , 42.46129 , 42.9028325 , 43.344375 , 43.7859175 , 44.22746 , 44.6690025 , 45.110545 , 45.5520875 , 45.99363 , 46.4351725 , 46.876715 , 47.3182575 , 47.7598])),

Collecte active : (counts([1.948000e+03, 1.258000e+04, 7.901900e+04, 1.834860e+05, 1.552200e+05, 2.317950e+05, 4.497090e+05, 6.401380e+05, 1.024874e+06, 1.090481e+06, 8.452000e+05, 6.488890e+05, 3.427440e+05, 1.664800e+05, 6.774800e+04, 3.482400e+04, 1.881300e+04, 9.764000e+03, 5.590000e+03, 5.114000e+03, 3.507000e+03, 2.936000e+03, 3.011000e+03,

2.442000e+03, 2.293000e+03, 1.927000e+03, 1.255000e+03, 9.760000e+02, 5.670000e+02, 3.470000e+02, 3.100000e+02, 2.230000e+02, 1.910000e+02, 1.860000e+02, 1.690000e+02, 1.450000e+02, 1.380000e+02, 1.280000e+02, 9.900000e+01, 9.200000e+01, 8.400000e+01, 7.300000e+01, 7.900000e+01, 6.300000e+01, 6.900000e+01, 6.700000e+01, 6.000000e+01, 6.800000e+01, 6.400000e+01, 4.700000e+01, 6.000000e+01, 6.700000e+01, 4.900000e+01, 3.700000e+01, 6.400000e+01, 5.300000e+01, 5.100000e+01, 5.200000e+01, 5.700000e+01, 4.100000e+01, 4.600000e+01, 4.300000e+01, 4.700000e+01, 4.300000e+01, 3.400000e+01, 5.300000e+01, 5.900000e+01, 5.500000e+01, 5.900000e+01, 6.800000e+01, 4.800000e+01, 5.000000e+01, 6.000000e+01, 4.400000e+01, 4.200000e+01, 3.400000e+01, 3.600000e+01, 2.700000e+01, 2.700000e+01, 2.600000e+01, 2.900000e+01, 2.600000e+01, 2.400000e+01, 1.700000e+01, 2.100000e+01, 2.300000e+01, 1.700000e+01, 1.300000e+01, 1.900000e+01, 1.400000e+01, 3.000000e+00, 1.000000e+01, 8.000000e+00, 5.000000e+00, 7.000000e+00, 7.000000e+00, 3.000000e+00, 6.000000e+00, 2.000000e+00, 2.000000e+00]), bins([0., 1.36528385, 2.73056771, 4.09585156, 5.46113541, 6.82641927, 8.19170312, 9.55698697, 10.92227083, 12.28755468, 13.65283853, 15.01812239, 16.38340624, 17.74869009, 19.11397395, 20.4792578, 21.84454165, 23.20982551, 24.57510936, 25.94039321, 27.30567707, 28.67096092, 30.03624477, 31.40152863, 32.76681248, 34.13209633, 35.49738019, 36.86266404, 38.22794789, 39.59323174, 40.9585156, 42.32379945, 43.6890833, 45.05436716, 46.41965101, 47.78493486, 49.15021872, 50.51550257, 51.88078642, 53.24607028, 54.61135413, 55.97663798, 57.34192184, 58.70720569, 60.07248954, 61.4377734, 62.80305725, 64.1683411, 65.53362496, 66.89890881, 68.26419266, 69.62947652, 70.99476037, 72.36004422, 73.72532808, 75.09061193, 76.45589578, 77.82117964, 79.18646349, 80.55174734, 81.9170312, 83.28231505, 84.6475989, 86.01288276, 87.37816661, 88.74345046, 90.10873432, 91.47401817, 92.83930202, 94.20458588, 95.56986973, 96.93515358, 98.30043744, 99.66572129, 101.03100514, 102.396289, 103.76157285, 105.1268567, 106.49214056, 107.85742441, 109.22270826, 110.58799211, 111.95327597, 113.31855982, 114.68384367, 116.04912753, 117.41441138, 118.77969523, 120.14497909, 121.51026294, 122.87554679, 124.24083065, 125.6061145, 126.97139835, 128.33668221, 129.70196606, 131.06724991, 132.43253377, 133.79781762, 135.16310147, 136.52838533]),