

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA
RECHERCHE SCIENTIFIQUE
UNIVERSITÉ ECHAHID HAMMA LAKHDAR D'EL-OUED



Faculté des Sciences Exactes
Département d'Informatique

THÈSE

Pour l'obtention du grade de Docteur Science En
Informatique

Titre

Conception d'une technologie ambiante d'assistance
domiciliaire pour un patient âgé

Présenté par : BERDJOUH Chafik

Jury:

Dr. Messaoud Abbas	Président	Université d'El Oued
Pr . Mohammed Charaf Eddine Meftah	Rapporteur	Université d'El Oued
Dr . Toumi Lyazid	Examineur	Université de Setif 1
Pr. Rezeg Khaled	Examineur	Université de Biskra
Pr. Bennoui Hammadi	Examineur	Université de Biskra
Dr. Drid Hamza	Examineur	Université de Batna2

2023/2024

Abstract

As people age, their need for care increases. With the world's elderly population growing, it is important that this care is delivered effectively. The rising cost of nursing care and the lack of caregivers has led to an increased need for assisted living at home. Home healthcare for the elderly has become an active area of research, particularly focusing on the recognition of human activities and the detection of anomalies in these activities.

Automatic health monitoring is an inexpensive way to obtain the information needed to provide effective and accurate care for the elderly. Activities of Daily Living (ADL) are a common method for assessing the cognitive and physical well-being of older people. These activities, such as walking, running, and sleeping, are performed daily.

This PhD thesis examines human activity recognition (HAR) and anomaly detection in elderly people living in smart homes, using wearable sensors and advanced deep learning techniques. Data is collected using accelerometers, gyroscopes, and magnetometers, then processed using convolutional neural networks (CNNs), long short-term memory networks (LSTMs), and hybrid CNN-LSTM models. These models classify daily activities and detect anomalies in real time, thereby improving the safety and independence of the elderly. The experimental results show high accuracy in activity recognition and effective detection of anomalies, offering significant advances for health management and emergency response in smart home environments.

My thesis comprises two contributions. Firstly, I developed an intelligent human activity recognition system based on wearable sensors and deep learning. The proposed models were implemented on a real dataset, and the hybrid CNN Multi-heads-LSTM model performed best, achieving an accuracy of 98.69%. Secondly, I proposed a new hybrid algorithm inspired by the PGTO nature to improve the performance of DFFNN, and tested it on the problem of detecting anomalies in

human activities in the elderly. The results show that the proposed approach has enhanced detection performance.

Keywords: Ambient Assisted Living (AAL); Human Activity Recognition (HAR); Activities of Daily Living (ADL); Deep Learning; PGTO; Anomaly detection.



Résumé

À mesure que les personnes vieillissent, leur besoin de soins augmente. Avec la croissance de la population âgée mondiale, il est important que ces soins soient dispensés efficacement. L'augmentation des coûts des soins infirmiers et le manque de soignants ont conduit à un besoin accru de vie assistée à domicile. Les soins de santé à domicile pour les personnes âgées sont devenus un domaine de recherche actif, en particulier en ce qui concerne la reconnaissance des activités humaines et la détection des anomalies dans ces activités.

La surveillance automatique de la santé est un moyen peu coûteux d'obtenir les informations nécessaires pour fournir des soins efficaces et précis aux personnes âgées. Les Activités de la Vie Quotidienne (ADL) sont une méthode courante pour évaluer le bien-être cognitif et physique des personnes âgées. Ces activités, telles que la marche, la course et le sommeil, sont effectuées quotidiennement.

Cette thèse de doctorat examine la reconnaissance des activités humaines (HAR) et la détection des anomalies chez les personnes âgées vivant dans des maisons intelligentes, en utilisant des capteurs portables et des techniques avancées de deep learning. Les données sont collectées à l'aide d'accéléromètres, de gyroscopes et de magnétomètres, puis traitées à l'aide de réseaux de neurones convolutionnels (CNN), de réseaux de mémoire à long terme (LSTM) et de modèles hybrides CNN-LSTM. Ces modèles classifient les activités quotidiennes et détectent les anomalies en temps réel, améliorant ainsi la sécurité et l'autonomie des personnes âgées. Les résultats expérimentaux montrent une grande précision dans la reconnaissance des activités et une détection efficace des anomalies, offrant des avancées significatives pour la gestion de la santé et la réponse aux urgences dans les environnements de maisons intelligentes.

Ma thèse comporte deux contributions. Premièrement, j'ai développé un système intelligent de reconnaissance des activités humaines basé sur des capteurs portables et le deep learning. Les modèles proposés ont été mis en œuvre sur un ensemble de données réelles, et le modèle hybride CNN Multi-heads-LSTM a obtenu les meilleures performances, avec une précision de 98,69%. Deuxièmement, j'ai proposé un nouvel algorithme hybride inspiré de la nature PGTO pour améliorer la performance de DFFNN, et je l'ai testé sur le problème de la détection des anomalies dans les activités humaines

chez les personnes âgées. Les résultats montrent que l'approche proposée a amélioré les performances de détection.

Mots-Clés : Assistance à l'autonomie à domicile (AAL) ; Reconnaissance des activités humaines (HAR); Activités de la vie quotidienne (ADL); apprentissage profond; PGTO; détection d'anomalie.

المخلص

مع تقدم الأشخاص في السن، تزداد حاجتهم للرعاية. ومع نمو السكان المسنين في العالم، يصبح من المهم تقديم هذه الرعاية بشكل فعال. أدى ارتفاع تكاليف الرعاية التمريضية ونقص مقدمي الرعاية إلى زيادة الحاجة إلى العيش المساعد في المنزل. أصبحت الرعاية الصحية المنزلية للمسنين مجالاً نشطاً للبحث، خاصة فيما يتعلق بالتعرف على الأنشطة البشرية واكتشاف الشذوذ في هاته الأنشطة. تعد المراقبة الصحية الآلية وسيلة غير مكلفة للحصول على المعلومات اللازمة لتقديم رعاية فعالة ودقيقة للمسنين. تعد أنشطة الحياة اليومية (ADL) طريقة شائعة لتقييم الرفاهية المعرفية والجسدية لكبار السن. تُمارس هذه الأنشطة، مثل المشي والجري والنوم بشكل يوميًا.

تدرس هذه الأطروحة التعرف على الأنشطة البشرية (HAR) واكتشاف الشذوذ لدى كبار السن الذين يعيشون في المنازل الذكية، باستخدام أجهزة استشعار قابلة للارتداء وتقنيات التعلم العميق. يتم جمع البيانات باستخدام مقاييس التسارع والجيروسكوبات والمغناطيسات، ثم تتم معالجتها باستخدام الشبكات العصبية التلافيفية (CNNs) وشبكات الذاكرة طويلة-قصيرة الأجل (LSTMs) والنماذج الهجينة CNN-LSTM. تقوم هذه النماذج بتصنيف الأنشطة اليومية واكتشاف الشذوذ في الوقت الحقيقي، مما يحسن من أمان واستقلالية كبار السن. أظهرت النتائج التجريبية دقة عالية في التعرف على الأنشطة واكتشاف الشذوذ بشكل فعال، مما يوفر تقدمًا كبيرًا في إدارة الصحة والاستجابة للطوارئ في بيئات المنازل الذكية.

تتضمن أطروحتي مساهمتين. أولاً، طورت نظامًا ذكيًا للتعرف على الأنشطة البشرية يعتمد على أجهزة استشعار قابلة للارتداء والتعلم العميق. تم تنفيذ النماذج المقترحة على مجموعة بيانات حقيقية، وحقق النموذج الهجين CNN Multi-heads-LSTM أفضل أداء بدقة بلغت 98.69%. ثانيًا، اقترحت خوارزمية هجينة جديدة مستوحاة من الطبيعة (PGTO) لتحسين أداء DFFNN، واختبرتها على مشكلة اكتشاف الشذوذ في الأنشطة البشرية لدى كبار السن. أظهرت النتائج أن النهج المقترح قد حسن أداء الاكتشاف.

الكلمات المفتاحية: المعيشة بمساعدة البيئة (AAL)، التعرف على النشاط البشري (HAR)، أنشطة الحياة اليومية (ADL)، التعلم العميق، PGTO، إكتشاف الشذوذ.

Remerciements

Tout d'abord, Je remercie Dieu tout puissant de m'avoir donné le courage, la force et la patience d'achever ce modeste travail.

J'aimerais exprimer ma gratitude à mon superviseur, le Pr. Mohammed Charaf Eddine Meftah pour leur soutien et leurs conseils continus tout au long de ce doctorat.

Je voudrais également remercier les membres du jury pour avoir accepté d'évaluer ce travail.

J'aimerais remercier le Pr Laouid A/Kader, Dr. Abbas Messaoud, Dr. Saci Medileh, Dr. Kertio Ismail, Dr. Khelaifa Abd Enacer et tous les enseignants de faculté Sciences exactes à l'université de L'oued pour leurs soutiens inestimables.

Enfin, je voudrais remercier ma femme qui m'a soutenu inconditionnellement et m'a donné l'opportunité d'étudier, mes fils, mes frères, mes sœurs et mes amis.

Dédicace

A la mémoire de mon défunt père

Et ma défunte mère.

À ma femme qui a toujours été à mes côtés.

*A mes chères filles : Djana, Djouri, et mes chers fils : yahia et
Amin*

*À tous mes frères et sœurs, ainsi que leurs enfants À toute ma
famille*

Table des matières

Abstract	I
Résumé.....	III
الملخص.....	V
Table des matières	VIII
Liste des figures	XI
Liste des tableaux	XIII
Liste des abréviations	XIV
Chapitre 1 Introduction.....	1
1.1 Background	2
1.2 Problématique	3
1.3 Objectifs scientifiques.....	4
1.4 Organisation du travail de la thèse.....	4
Chapitre 2 Etat de l'art.....	6
2.1. Introduction.....	7
2.2. Internet des objets (IoT)	7
2.3. Assistance à l'autonomie à domicile	8
2.4. AAL Technologies.....	8
2.4.1. La maison intelligente (Smart Home)	8
2.4.2. Capteurs utilisés.....	13
a) Capteurs visuels.....	13
b) Les capteurs ambiants.....	14
c) Les capteurs portables	14
d) Les capteurs multimodaux	15
2.4.3. Reconnaissance des activités humaines	15
a) Les activités de la vie quotidienne	16
b) Taxinomie des activités humaines	16
c) Classification des approches de reconnaissance de l'activité humaine	17
2.4.4. La détection des anomalies	19
2.4.5. Les défis de la HAR	20
2.5. Conclusion	21
Chapitre 3 Revue de littérature	22
3.1 Introduction.....	23

3.2	Les modèles ML et DL.....	23
3.2.1	Apprentissage automatique (Machine Learning).....	23
a)	Apprentissage supervisé	24
b)	Apprentissage non-supervisé.....	24
c)	Apprentissage semi-supervisé	25
3.2.2	Modèles d'apprentissage profond (Deep Learning).....	25
3.2.2.1	Réseau de neurones artificiels (ANN)	26
3.2.2.2	Réseau de neurones convolutifs (CNN)	33
3.2.2.3	Réseau de neurones récurrents.....	35
3.2.2.4	Le réseau hybride.....	38
3.3	Travaux connexes	39
3.3.1	Reconnaissance des activités humaines (HAR)	39
3.3.2	La détection des anomalies dans les activités humaines	41
3.3.3	Discussion	44
A)	Travaux dans la reconnaissance des activités humaines (HAR).....	44
B)	Travaux dans la détection des anomalies dans les activités humaines.....	44
3.4	Conclusion	45
Chapitre 4	Contribution 1 : une approche DL pour la reconnaissance des activités des personnes âgées à l'aide de capteurs portables	46
4.1	Introduction	47
4.2	Méthodologie.....	48
4.3	Résultats Expérimentaux et discussion.....	58
4.4	Conclusion.....	66
Chapitre 5	Contribution 2 : Algorithme hybride PGTO d'optimisation du réseau DFFNN et application à la détection d'anomalies dans l'activité humaine	67
5.1	Introduction	68
5.2	Background	69
5.2.1	Algorithme d'optimisation Pelican (POA)	69
5.2.2	Algorithme d'optimisation de troupes de Gorilles (GTO)	72
5.3	Méthodologie.....	77
5.3.1	Framework proposé	77
5.4	Résultats et discussion	84
5.4.1	Métriques d'évaluation	84
5.4.2	Analyse comparative	85
5.5	Conclusion.....	94

Chapitre 6	Conclusion et travaux futurs	95
6.1	Contribution de la thèse	96
6.2	Travaux futurs	97
Bibliographie	98

Liste des figures

Figure 2-1: Quelques applications de maison intelligente (Abhishek Ghosh, 2019)	9
Figure 2-2 : Classification des approches de reconnaissance de l'activité humaine.....	17
Figure 3-1: Exemple d'un réseau ANN avec 3 couches cachées	27
Figure 3-2 : Principe de fonctionnement de la descente de gradient.....	32
Figure 3-3 Architecture du réseau CNN	33
Figure 3-4 : Trois types de l'opération de regroupement (Pooling)	35
<i>Figure 3-5 : Architecture RNN simple.....</i>	<i>36</i>
Figure 3-6 : Cellule unique de réseau LSTM	36
Figure 3-7 architecture du modèle LSTM	37
<i>Figure 4-1: A framework de système de reconnaissance des activités humaines basé sur l'apprentissage profond et utilisant des capteurs portables.</i>	<i>49</i>
<i>Figure 4-2 : Emplacement des capteurs sur le sujet</i>	<i>50</i>
Figure 4-3 : Segmentation des données de capteurs	52
<i>Figure 4-4: Architecture des modèles : (a) CNN, (b) LSTM</i>	<i>54</i>
<i>Figure 4-5: Architecture du modèle CNN-LSTM</i>	<i>55</i>
<i>Figure 4-6: Architecture du modèle CNN Multi_têtes-LSTM</i>	<i>57</i>
<i>Figure 4-7 : Courbes Accuracy et Loss du modèle : CNN</i>	<i>60</i>
<i>Figure 4-8 : Courbes Accuracy et Loss du modèle ;LSTM</i>	<i>60</i>
<i>Figure 4-9 : Courbes Accuracy et Loss du modèle : CNN-LSTM.....</i>	<i>60</i>
<i>Figure 4-10: Courbes Accuracy/Loss d'apprentissage/test du modèle CNN Multi_têtes- LSTM.....</i>	<i>61</i>
<i>Figure 4-11 : La matrice de confusion de jeu de données Mhealth pour le modèle : CNN</i>	<i>62</i>
<i>Figure 4-12 : La matrice de confusion de jeu de données Mhealth pour le modèle : LSTM.....</i>	<i>62</i>
<i>Figure 4-13 : La matrice de confusion de jeu de données Mhealth pour le modèle : CNN- LSTM.....</i>	<i>63</i>
<i>Figure 4-14 : La matrice de confusion de jeu de données Mhealth pour le modèle : CNN Multi_têtes-LSTM</i>	<i>63</i>
<i>Figure 5-1 : schéma fonctionnel du Framework proposé</i>	<i>78</i>
Figure 5-2 : DFFNN avec 3 couches cachées.....	80
<i>Figure 5-3 : Comparaison basé sur différents taux d'apprentissage/test en termes de (a) accuracy, (b) precision, (c) recall pour différents DL algorithmes et PGTO-DFFNN.....</i>	<i>87</i>
Figure 5-4 : Comparaison des performances pour différentes tailles de batch pour PGTO- DFFNN et divers DL modèles. La taille du batch est fixée à 5, 7 et 8 en termes de (a) Accuracy (b) Precision, (c) Recall	88
Figure 5-5 : Comparaison des performances des divers algorithmes d'optimisation basé sur différentes tailles de population (swarm size) en termes de (a) accuracy, (b) precision, (c) recall	90

Figure 5-6: Analyse Comparatives entre PGTO et les autres algorithmes d'optimisation de DFFNN : GA, ROA, POA et GTO basée sur différentes nombres d'itérations : 10,30 et 40 en termes de : (a) accuracy, (b) precision, (c) recall.....92

Figure 5-7 : Matrice de confusion normalisée utilisant l'approche PGTO-DFFNN (classe 1 : Relaxing, classe 2 : Sleeping, classe 3 : Working, et classe 4 : Emergency)93

Liste des tableaux

<i>Tableau 4-1 : Liste des activités</i>	51
<i>Tableau 4-2 : Résultats de classification des modèles proposés</i>	61
<i>Tableau 4-3 : Rapport de classification des modèles CNN et LSTM avec le dataset Mhealth</i>	64
<i>Tableau 4-4 : Résumé du rapport de classification des modèles CNN-LSTM et CNN Multi_têtes-LSTM avec le jeu de données Mhealth</i>	65
<i>Tableau 4-5 : Comparaison des performances des modèles proposée et des méthodes précédentes sur le jeu de données Mhealth.</i>	66
Tableau 5-1: Analyse comparative	93
Tableau 5-2: Analyse algorithmique.....	93
Tableau 5-3: Comparaison entre PGTO-DFNN et les autres méthodes.....	94

Liste des abréviations

AAL	Ambient Assisting Living
ADL	Activities of Daily Living
AI	Artificiel Intelligence
ANN	Artificial Neural Networks
BP	Back-Propagation
CNN	Convolutional Neural Network
LSTM	Long-Short Term Memory
HAR	Human Activity Recognition
RFID	Radio-Frequency IDentification
SM	Smart Home
SVM	Support Vector Machines
DFFNN	Deep Feed Forward neural network
ReLU	Rectified linear unit.
SMOTE	Synthetic Minority Oversampling TEchnique
MSE	Mean Square Error
IoT	Internet of Object
MSE	Mean Squared Error
PCA	Principal Component Analysis
POA	Pelican Optimization Algorithm
GTO	Gorilla Troop Optimization
IADL	Activities of Daily Living
USEFIL	Unobtrusive Smart Environments for Independent Living
PGTO	Pelican Gorilla Troupes Optimization

Liste des Publications

1. Pelican Gorilla Troop Optimization Based on Deep Feed Forward Neural Network for Human Activity Abnormality Detection in Smart Spaces
Chafik Berdjouh; Mohammed Charaf Eddine Meftah; Abdelkader Laouid; Mohammad Hammoudeh; Akshi Kumar
IEEE Internet of Things Journal
Year: 2023 | Volume: 10, Issue: 21 | Journal Article | Publisher: IEEE
2. Hybrid Architecture Based on Blockchain Technology and Smart Contracts for Healthcare Applications
Mohammed Charaf Eddine Meftah, Cherif Ahmed Cherif, Belgacem Meraghni, Chafik Berdjouh
Source Title: International Journal of Organizational and Collective Intelligence (IJOCI) 12(1)
Copyright: © 2022 | Pages: 20
DOI: 10.4018/IJOCI.313600
3. Human Activity Recognition with Anomaly Prediction for E-Health Systems using Lightweight AI
R. Priyanka Pramila; Anita Gehlot; Berdjouh Chafik; Ajith Sundaram; Lekha H; Finney Daniel Shadrach; Naveen I G ,
2022 International Conference on Knowledge Engineering and Communication Systems (ICKES)
Year: 2022 | Conference Paper | Publisher: IEEE

Chapitre 1 Introduction

1.1 Background

Au cours des dernières décennies, l'espérance de vie mondiale a augmenté considérablement en raison des progrès de la science et de la technologie médicale, de la médecine et de la santé publique, ainsi qu'une prise de conscience accrue de l'hygiène environnementale et personnelle. Cependant, l'allongement de l'espérance de vie a entraîné un vieillissement croissant de la population, mettant en péril la structure socio-économique de nombreux pays en termes de coûts liés aux soins de santé et au bien-être des personnes âgées. Des solutions de santé abordables, discrètes et faciles à utiliser sont essentielles pour faire face au besoin croissant de services de santé pour les personnes âgées. Les maisons intelligentes (Smart Houses, en anglais), qui intègrent des capteurs médicaux environnementaux et portables, des actionneurs et des technologies modernes de communication et d'information, peuvent permettre une surveillance continue et à distance de la santé et du bien-être des personnes âgées.

La technologie de la maison intelligente vise à améliorer la qualité de vie et à permettre aux personnes âgées de vivre de manière confortable et autonome [1]. Cette technologie est considérée comme un moyen de réduire les coûts de la vie et des soins tout en améliorant la qualité de vie des personnes dépendantes. Elle a été utilisée à diverses fins [2], notamment l'économie d'énergie, la sécurité et la sûreté, la détection des chutes, la gestion de la lumière, la détection des fumées et des incendies, etc. en utilisant diverses solutions telles que la surveillance vidéo, les alarmes, les planificateurs et calendriers intelligents, les rappels, etc. Les maisons intelligentes peuvent permettre aux systèmes automatiques ou aux soignants de contrôler l'environnement pour le compte des résidents, de prédire leurs activités et de suivre leur état de santé en utilisant des capteurs, des actionneurs et éventuellement des caméras pour collecter divers types de données sur le domicile et les résidents.

Les personnes âgées sont souvent confrontées à des défis liés à la mobilité, à la santé et aux activités de la vie quotidienne. La reconnaissance des activités humaines et la détection des anomalies sont apparues comme des technologies innovantes visant à améliorer les soins et l'accompagnement apportés à la population âgée.

La reconnaissance des activités humaines (HAR) implique l'utilisation d'algorithmes avancés d'intelligence artificielle (IA) et d'apprentissage automatique (ML)

ou profond (DL) pour identifier et comprendre automatiquement les différentes activités réalisées par les individus. Dans le contexte des personnes âgées, cette technologie peut jouer un rôle déterminant dans la surveillance de leurs routines quotidiennes, telles que marcher, s'asseoir, se tenir debout et s'engager dans des tâches spécifiques comme cuisiner ou prendre des médicaments. En reconnaissant avec précision ces activités, les soignants, les membres de la famille et les professionnels de la santé peuvent obtenir des informations précieuses sur le bien-être des personnes âgées, en identifiant des changements de comportement pouvant indiquer des problèmes de santé potentiels ou des changements dans leur état. Par exemple, l'enregistrement des activités quotidiennes d'une personne peut être utilisé pour calculer les calories qu'elle a consommées au cours d'une journée, ce qui peut en outre suggérer un régime alimentaire approprié pour qu'elle reste en bonne santé et en forme.

La détection des anomalies, d'autre part, complète la reconnaissance des activités humaines en se concentrant sur l'identification des comportements et des événements anormaux. L'intégration d'algorithmes de détection d'anomalies dans les systèmes de soins aux personnes âgées peut améliorer considérablement la sécurité et les temps de réponse, car elle permet des alertes rapides et des interventions appropriées lorsqu'un comportement irrégulier ou des incidents sont détectés. Par exemple, la détection de l'activité de chute des personnes âgées peut être utilisée pour déclencher une assistance d'urgence afin d'éviter de provoquer des accidents graves.

Cette approche combinée de reconnaissance des activités humaines et de détection des anomalies améliore non seulement les soins aux personnes âgées, mais permet également un accompagnement plus personnalisé et efficace. En surveillant en permanence les activités et en identifiant les écarts par rapport à la norme, les soignants et les prestataires de soins de santé peuvent adapter les interventions et les services aux besoins uniques de chaque individu, favorisant ainsi une meilleure qualité de vie pour la population âgée.

1.2 Problématique

La problématique de la reconnaissance des activités réside dans la complexité de surveiller et de comprendre les activités quotidiennes (ADLs), qui peuvent être variées et changeantes. Cela devient encore plus crucial pour les personnes âgées vivant seules ou

nécessitant des soins particuliers. La technologie de reconnaissance d'activités vise à automatiser cette surveillance en utilisant des méthodes comme le Deep learning. L'objectif est d'identifier et de suivre les activités normales, détecter les comportements anormaux ou les situations dangereuses, pour fournir des alertes et pour garantir le bien-être et la sécurité des personnes âgées, tout en améliorant leur qualité de vie et leur indépendance.

1.3 Objectifs scientifiques

L'objectif principal de la thèse était de proposer et de développer des méthodes pour reconnaître les activités humaines et détecter les anomalies dans ces activités pour faciliter un meilleur vieillissement et un maintien à domicile des personnes âgées, basées sur l'apprentissage profond avec des données issues de capteurs portables. Plus précisément, la combinaison de plusieurs objectifs sous-jacents permet de réaliser l'objectif global de la thèse :

- Développer un système intelligent pour la reconnaissance des activités physiques des personnes âgées basé sur l'apprentissage profond.
- Développer un nouvel algorithme hybride d'optimisation inspiré de la nature PGTO afin d'améliorer les performances du DFFNN [3], puis le tester au problème de la détection des anomalies dans les activités humaines.

1.4 Organisation du travail de la thèse

Cette thèse est organisée comme suit :

Le chapitre 1 : présente le contexte, la problématique, les objectifs de la thèse.

Le chapitre 2 : donne un aperçu général autour l'IoT et la maison intelligente, les capteurs, et la reconnaissance des activités humaines (HAR) et ses défis.

Le chapitre 3 : passe en revue de littérature des travaux connexes dans le domaine de la reconnaissance d'activités humaines (HAR) avec des capteurs portables à l'aide des modèles d'apprentissage automatique et profond.

Le chapitre 4 : présente une proposition d'un système intelligent de classification des activités humaines à l'aide des modèles d'apprentissage profond : CNN, LSTM, CNN-LSTM et CNN_Multi-têtes-LSTM basé sur les capteurs portables. En outre, le framework proposé, le dataset et les configurations expérimentales sont discutés en détail. Les performances du modèle sont

analysées à travers les courbes en termes de validation et de loss. Les résultats de la matrice de confusion du test sont également illustrés. Des métriques telles que l'accuracy et la precision sont calculés.

Le chapitre 5 : propose un nouveau modèle hybride PGTO-DFNN qui utilise l'algorithme PGTO pour optimiser les paramètres DFNN afin d'augmenter la performance de classification. Les résultats de l'application de l'algorithme hybride sur un dataset public pour détecter les anomalies sur les activités physiques sont présentés dans ce chapitre.

Le chapitre 6 : résume les contributions de la thèse à la recherche et propose des recommandations pour les travaux futurs.

Chapitre 2 Etat de l'art

2.1. Introduction

L'Internet des objets (IoT) promet de créer un monde dans lequel tout ce qui nous entoure est connecté à Internet et communique entre eux avec une intervention humaine minimale. L'objectif principal est de créer un monde meilleur pour les humains, dans lequel les objets qui nous entourent perçoivent le contexte, leur permettant de répondre à des questions telles que : ce que nous voulons, ce dont nous avons besoin et l'endroit où nous nous trouvons ? la maison intelligente, l'un des principaux domaines d'application de l'IoT, a fait l'objet d'une attention particulière de la part des chercheurs [4].

Les maisons intelligentes offrent un environnement sûr et sécurisé aux personnes âgées ou dépendantes. Ils offrent la possibilité de (1) surveiller les activités des résidents sans interférer avec leur vie quotidienne ; et (2) surveiller le comportement des résidents et surveiller leur santé à l'aide de capteurs installés dans leurs espaces de vie [5]. Les données collectées à partir des maisons intelligentes doivent être analysées et étudiées en profondeur pour en tirer des informations utiles sur les activités quotidiennes des résidents. Dans ce chapitre, nous présenterons les définitions de l'internet des objets (IoT) et l'assistance à l'autonomie à domicile (AAL). La section suivante passe en revue les technologies AAL qui peut apporter aux personnes âgées certains avantages qui facilitent l'autonomie, et les aident dans leurs activités quotidiennes et améliorent leur santé. Ensuite, un aperçu sur les capteurs utilisés, la reconnaissance des activités humaines et la détection des anomalies.

2.2. Internet des objets (IoT)

L'internet des objets (IoT) est défini comme un système numérique interconnecté composé d'une variété de dispositifs informatiques indépendants. Il possède une identité unique et est capable d'échanger des données via l'internet sans intervention humaine [6]. Il s'agit d'une plate-forme où les capteurs et les appareils peuvent communiquer de manière transparente dans un environnement intelligent. Chaque année, le nombre d'appareils compatibles IoT augmente ; des secteurs tels que le transport, la santé, la sécurité, la ville intelligente, l'éducation, l'agriculture et bien d'autres ont bénéficié de son développement. Le résultat sera une génération d'applications capables d'effectuer des tâches complexes de détection et de reconnaissance pour prendre en charge un nouveau monde d'interaction homme-objet. Par exemple, les applications IoT dans des environnements tels que les

maisons intelligentes [7] ont le potentiel de soutenir et d'aider les personnes âgées et de les aider à vivre de manière indépendante dans leur propre maison [8]. Ces systèmes peuvent également aider à indiquer la capacité d'une personne âgée à effectuer des activités quotidiennes de base telles que cuisiner et prendre un bain [9].

2.3. Assistance à l'autonomie à domicile

L'assistance à l'autonomie à domicile (en anglais Ambient Assisting Living ou AAL) peut être définie comme : "L'utilisation des technologies de l'information et de la communication (TIC) dans la vie quotidienne d'une personne pour lui permettre de rester active plus longtemps, de conserver des liens sociaux et de vivre de manière indépendante jusqu'à un âge avancé" [10].

L'AAL comprend la mise en œuvre et l'utilisation de technologies intelligentes pour permettre aux personnes âgées de résoudre les problèmes liés à l'assistance à l'autonomie et de vivre plus longtemps dans l'environnement qu'elles préfèrent [11].

Le domaine de recherche des (AAL) se concentre sur la création de technologies, de produits et de services innovants pour l'aide, les soins médicaux et la réadaptation des personnes âgées, dans le but d'augmenter le temps pendant lequel ces personnes peuvent vivre de manière indépendante, qu'elles souffrent de maladies neurodégénératives ou de handicaps. Ce domaine de recherche clé est responsable du développement de systèmes de reconnaissance d'activités, qui constituent un outil précieux pour identifier les types d'activités effectuées par les personnes âgées et leur fournir une assistance efficace leur permettant d'effectuer normalement leurs activités quotidiennes.

2.4. AAL Technologies

De nombreuses technologies et des projets intelligents peuvent aider les personnes âgées à vivre de manière autonome. Les maisons intelligentes, les capteurs portables et mobiles et la surveillance des activités humaines font partie des avancées récentes en matière d'AAL.

2.4.1. La maison intelligente (Smart Home)

1. Définition

[12] a défini une maison intelligente (SH) comme "une résidence équipée d'ordinateurs et de technologies de l'information, qui anticipe et répond aux besoins des

occupants, s'efforçant de promouvoir leur confort, leur commodité, leur sécurité et leur divertissement grâce à la gestion de la technologie à l'intérieur de la maison et aux connexions avec le monde extérieur". Leur définition comprend la partie technologique du phénomène, les services et les fonctions qu'il offre et les types de besoins des utilisateurs auxquels les maisons intelligentes visent à répondre (Figure 2-1).

La définition de [13] indique que "La maison intelligente est une résidence équipée d'un réseau de haute technologie, reliant des capteurs et des dispositifs domestiques, des appareils et des accessoires qui peuvent être surveillés, consultés ou contrôlés à distance, et fournir des services qui répondent aux besoins de ses habitants".

Grâce à l'utilisation de technologies intelligentes, telles que les capteurs, les actionneurs et l'intelligence artificielle (IA), ces environnements peuvent avoir un impact sur la qualité de vie, le bien-être, la sécurité et les économies d'énergie [14]. À cette fin, l'intelligence artificielle et Internet des objets (IoT) collectent des données entre les objets, les appareils et les occupants et analysent les modes de vie et les habitudes des résidents.

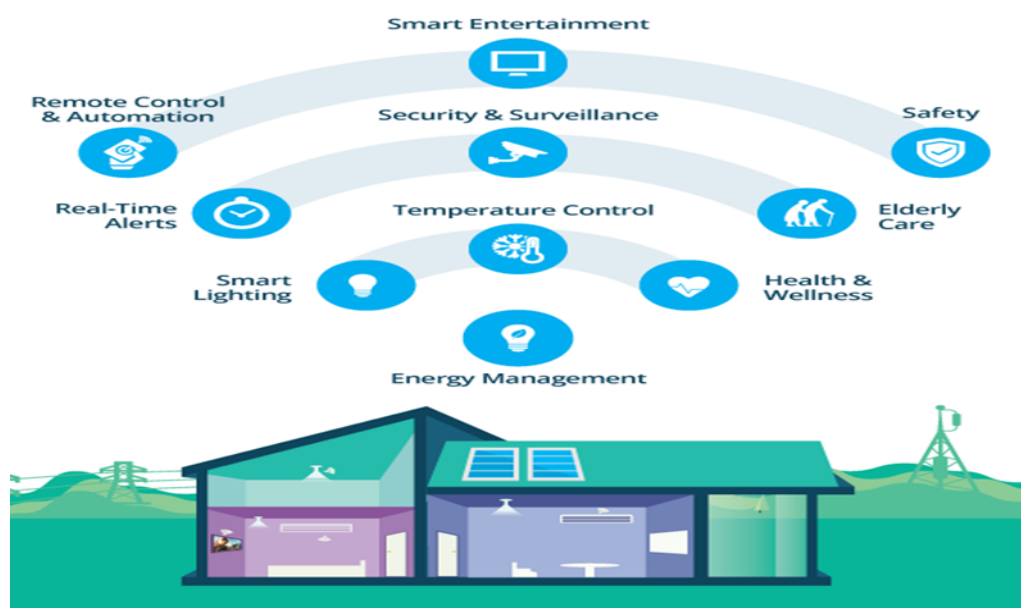


Figure 2-1: Quelques applications de maison intelligente (Abhishek Ghosh, 2019)

2. Principales fonctions des maisons intelligentes pour les soins aux personnes âgées

Les maisons intelligentes peuvent fournir de nombreuses fonctions pour améliorer la santé et le bien-être des personnes âgées [15]. Ces caractéristiques peuvent aider les personnes âgées à conserver leur indépendance tout en assurant leur sécurité et leur confort. Voici quelques caractéristiques clés de la maison intelligente pour la santé des personnes âgées :

➤ Prévenir les maladies

Depuis le début de la pandémie de COVID-19 en 2020, les habitants sont plus nombreux que jamais à se préoccuper de la prévention des maladies et de la propreté de leurs logements. La technologie de surveillance de la santé dans les maisons intelligentes s'est développée, y compris les lampes désinfectantes à UV pour éliminer les germes et les capteurs infrarouges qui détectent les températures élevées. Ces dispositifs font un effort supplémentaire pour tenir les bactéries et les virus à distance, ce qui permet aux résidents de se reposer en toute tranquillité.

➤ Détecter et prévenir les chutes

Les maisons intelligentes peuvent intégrer des capteurs et des caméras pour détecter les chutes ou les mouvements inhabituels. Ils peuvent ensuite envoyer des notifications aux soignants ou aux services d'urgence, réduisant ainsi les temps de réponse en cas d'accident.

➤ Gestion des médicaments

Le distributeur de médicaments intelligent peut rappeler aux personnes âgées de prendre leurs médicaments à temps. Ces systèmes peuvent également informer les soignants si une dose est oubliée.

➤ Surveillance de la santé

Des dispositifs comme les montres intelligentes et les trackers de santé portables peuvent suivre les signes vitaux comme la fréquence cardiaque, la pression artérielle et les habitudes de sommeil. Ces données peuvent être partagées avec des prestataires de soins de santé ou des membres de la famille pour une détection précoce des problèmes de santé.

➤ Systèmes d'intervention d'urgence

Les appareils à commande vocale comme Amazon Echo ou Google Home peuvent être utilisés pour appeler à l'aide en cas d'urgence, comme une chute ou une maladie soudaine.

➤ Soutien cognitif

Les appareils domestiques intelligents peuvent fournir des aides à la mémoire, des rappels et des jeux d'apprentissage cérébral pour aider les personnes âgées à maintenir leurs fonctions cognitives.

➤ Surveillance à distance par les soignants

Les soignants ou les membres de la famille peuvent surveiller à distance le bien-être des personnes âgées à l'aide d'appareils connectés, afin de s'assurer qu'elles sont en sécurité et à l'aise.

➤ Aide à la nutrition

Les réfrigérateurs intelligents peuvent aider les personnes âgées à suivre les stocks de nourriture, à suggérer des plans de repas et même à commander des plats en ligne.

➤ Sécurité et sûreté

Les systèmes de sécurité domestique intelligents peuvent apporter la tranquillité d'esprit en surveillant les intrusions, les incendies ou d'autres risques de sécurité.

3. Projets de maisons intelligentes

Dans cette section, nous menons une étude sur les projets de maisons intelligentes les plus importants qui visent à améliorer l'assistance à l'autonomie des personnes âgées dans le passé récent. Ces projets de maison intelligente simulent entièrement l'environnement d'une maison intelligente, y compris le déploiement d'une large gamme de capteurs.

- GatorTech [12], [16] est un projet précoce de maison intelligente réalisé à l'université de Floride, qui a intégré un ensemble de capteurs et de dispositifs pour fournir des services tels que la reconnaissance vocale et le suivi de l'activité des habitants.
- Le projet de maison intelligente CASAS [17], créé en 2007 à l'université de l'État de Washington, est un projet de recherche multidisciplinaire axé sur la création d'un environnement domestique intelligent à l'aide de capteurs et d'actionneurs discrets. Les domaines de recherche inclus dans CASAS sont la technologie d'assistance, l'intelligence artificielle, l'apprentissage automatique et la reconnaissance d'activité. Dans ses recherches récentes, cette même équipe a créé la "maison intelligente dans une boîte (smarthome in a box)" [18], une conception de maison intelligente légère et facile à installer qui fournit des capacités maison intelligente dès la sortie de la

boîte, sans aucun ajustement ou formation supplémentaires. Ces capacités incluent la reconnaissance d'activité, qui fournit un label d'activité en temps réel lorsque les événements du capteur arrivent dans un flux, et la découverte d'activité pour les données non marquées en utilisant un algorithme d'apprentissage non supervisé.

- SWEET-HOME [19], un projet de recherche financé par l'État en France, vise à créer un nouveau système de maison intelligent basé sur la technologie audio. Ce projet a trois objectifs principaux : fournir une technologie d'interaction basée sur l'audio qui permet aux utilisateurs de contrôler entièrement leur environnement domestique, détecter les situations de détresse et faciliter l'inclusion sociale des personnes âgées et fragiles. Un axe de recherche intéressant de leur système de maison intelligente est le processus de décision tenant compte du contexte, qui utilise une approche de réseau logique de Markov dédiée pour améliorer la capacité à faire face à des événements incertains déduits de données de capteurs réelles [20].
- Un projet récent de maison intelligente est "Unobtrusive Smart Environments for Independent Living" (USEFIL), qui a débuté en 2011. Il vise à fournir une assistance sanitaire avancée et abordable dans un environnement domestique intelligent, en adoptant une architecture à trois couches et un réseau de capteurs discret pour soutenir un système d'aide à la décision (SAD) afin de fournir des données aux applications de surveillance par le biais d'une interface conviviale [21]. Dans un environnement MI, un ensemble limité de capteurs et de périphériques, y compris une unité portable au poignet, une caméra, un microphone et un capteur Kinect, sont utilisés pour identifier les activités physiques de base des personnes âgées, telles que se coucher, s'asseoir, marcher, se tenir debout, cyclisme, courir, monter et descendre les escaliers [22]. Des plateformes open source et un système autonome à faible coût sont développés pour faciliter la création d'applications visant à combler le manque entre les progrès technologiques et le vieillissement de la population.
- Qing et Mohan [23] ont proposé une solution de maison plus intelligente et plus sûre au laboratoire Franco-Australien CSIRO pour améliorer la qualité de vie des personnes âgées. Pour réaliser cette solution, plusieurs capteurs environnementaux sont installés dans différents endroits de la maison intelligente. Ces capteurs agissent comme des dispositifs de surveillance non intrusifs qui peuvent identifier

les comportements humains. Sur la base de ce projet, Zhang et al [24] ont proposé une plateforme Smart Assistive Living (SAL) pour permettre aux personnes âgées de rester chez elles de manière autonome le plus longtemps possible. Les capteurs placés dans les maisons intelligentes sont censés fournir un flux de données continu à un serveur. L'extraction et l'analyse de ces données à l'aide de mécanismes d'apprentissage automatique permettent aux experts cliniques et aux soignants de procéder à des diagnostics et de prendre des décisions.

2.4.2. Capteurs utilisés

Aujourd'hui, de nombreux types de capteurs sont utilisés pour collecter des informations sur les activités humaines. Parmi ceux-ci, on peut citer principalement : des capteurs visuels, des capteurs ambiants, capteurs portables et capteurs multimodaux.

a) Capteurs visuels

Les capteurs visuels sont basés sur l'utilisation de dispositifs de détection visuelle, en particulier des caméras, pour extraire des données visuelles et les utiliser pour observer le comportement des utilisateurs et les fluctuations de l'environnement [25]. Les outils d'analyse de la vision par ordinateur sont utilisés dans cette catégorie pour reconnaître les activités quotidiennes d'un ou de plusieurs utilisateurs [26]. L'utilisation de la détection visuelle vidéo pour la reconnaissance d'activités est très répandue dans le domaine de la sécurité domestique. [27] propose de capturer sur vidéo les résidents d'une maison intelligente et de traiter la vidéo pour reconnaître les activités. Des travaux récents ont amélioré le domaine des données visuelles dans les maisons intelligentes, comme l'analyse de divers comportements anormaux en temps réel [28], le système d'assistance [29] et la détection d'activités humaines [30].

Cependant, les capteurs visuels ont encore de nombreux problèmes liés à la vie privée du propriétaire, aux coûts d'installation élevés, aux variations de lumière et aux inefficacités de la caméra face aux changements de luminosité et de couleur. De plus, la grande taille des données des caméras nécessite une grande capacité de stockage et de puissants systèmes d'analyse.

b) Les capteurs ambiants

La reconnaissance des activités humaines basée sur les signaux des capteurs ambiants prend de l'ampleur, car les capteurs peuvent collecter de nombreuses données précieuses de bas niveau sur les utilisateurs et leur environnement de vie. Ces données sont analysées à l'aide de techniques d'exploration de données et d'apprentissage automatique pour former des modèles d'activité. Par exemple, l'activation d'un capteur booléen au lit peut clairement indiquer l'activité du sommeil. Les capteurs de mouvement déduisent les activités en suivant les interactions homme-objet [31], et [32] utilisent des capteurs qui peuvent être installés rapidement et couramment dans les maisons pour détecter des activités connues telles qu'aller aux toilettes, se laver et nettoyer la maison. (Cottone et al., 2015) a utilisé des capteurs de propriétés différentes, capteur de proximité, capteur d'objet et capteur d'environnement, dont l'état peut être représenté par une variable binaire discrète ou continue, pour reconnaître les activités de la vie quotidienne de l'utilisateur. Certaines études sur la reconnaissance d'activité [33], [34] se concentrent davantage sur des capteurs à faible coût et à faible consommation d'énergie pour reconnaître l'activité humaine à l'intérieure de la maison.

Cependant, l'utilisation de capteurs ambiants est difficile en raison du bruit des données des capteurs et de la faible efficacité de la reconnaissance de modèles complexes d'activité humaine, tels que la cuisine et le nettoyage de la maison. Les capteurs ambiants génèrent des frais importants d'installation et de maintenance. De plus, ils ne sont pas en mesure de fournir des informations adéquates pour effectuer une reconnaissance précise des activités.

c) Les capteurs portables

Les capteurs portables utilisés dans la reconnaissance de l'activité humaine sont de petits dispositifs sans fil qui peuvent être portés sur le corps, les chaussures, des bracelets, des chemises, des patches, des montres et même des accessoires tels que des bagues. Ces capteurs sont équipés de diverses technologies telles que des accéléromètres, des gyroscopes et parfois même des moniteurs de fréquence cardiaque. Ils capturent et mesurent les mouvements physiques, l'orientation et parfois les signaux physiologiques du porteur. Ces données sont ensuite analysées pour identifier et classer différentes activités humaines, telles que la marche, la course, le cyclisme ou même des actions plus complexes. Ces capteurs ont des applications dans les domaines de la santé, du sport, de l'interaction

homme-robot et dans d'autres domaines. Ils peuvent fournir des informations sur l'état de santé, les modèles de comportement et les mesures de performance, contribuant ainsi à améliorer le bien-être et à fournir une assistance sur mesure dans divers aspects de la vie [35]. Cependant, les capteurs portables sont un défi en raison de la nécessité d'apporter des capteurs partout et à tout moment à l'intérieur de la maison, les systèmes et les dispositifs portables doivent avoir des logiciels et du matériel efficaces, discrets et résistants à l'eau et à la température. En outre, les données des capteurs portables ne permettent pas à elles seules de différencier certaines activités instrumentales, par exemple la préparation du thé et du café.

d) Les capteurs multimodaux

La détection multimodale implique l'intégration et l'utilisation de données provenant de plusieurs modalités sensorielles afin d'améliorer la collecte et la compréhension des informations. La détection multimodale dans la reconnaissance de l'activité humaine implique l'intégration de données provenant de plusieurs modalités de capteurs afin de détecter et de classer avec précision les différentes activités effectuées. En combinant les informations provenant de capteurs qui surveillent les mouvements du corps, la fréquence cardiaque et d'autres signaux physiologiques, il est possible d'obtenir une compréhension globale des activités de la personne. Cette approche améliore la précision et la fiabilité de la reconnaissance des activités dans la vie quotidienne. Le projet SWEET-HOME [36], lancé en 2009, est basé sur des capteurs multimodaux qui détectent les situations de détresse d'une personne et fournissent une assistance par le biais de commandes vocales et tactiles afin d'assurer la sécurité à tout moment et en tout lieu de la maison.

2.4.3. Reconnaissance des activités humaines

La technologie de reconnaissance des activités humaines (Human Activity Recognition ou HAR) est l'une des technologies les plus importantes pour l'assistance à l'autonomie, la sécurité basée sur la surveillance, les activités sportives et de remise en forme, et les soins de santé des personnes âgées. Les activités exercées par la personne âgée sont pertinentes pour notre travail. À cet effet, nous examinons les activités de la vie quotidienne proposée dans la littérature et leurs classifications.

a) Les activités de la vie quotidienne

Les activités de la vie quotidienne (Activities of daily living en anglais ou ADL) sont un terme utilisé pour décrire collectivement les compétences fondamentales requises pour prendre soin de soi de manière autonome, comme manger, prendre un bain et se déplacer. Le terme activités de la vie quotidienne a été inventé pour la première fois par Sidney Katz en 1950 [37].

Les ADL sont utilisés comme indicateur de l'état fonctionnel d'une personne. L'incapacité d'accomplir les ADL entraîne une dépendance à l'égard d'autres personnes et/ou d'appareils mécaniques. Elle peut conduire à des conditions dangereuses et à une mauvaise qualité de vie [38].

b) Taxinomie des activités humaines

Les activités de la vie quotidienne sont classées en ADL de base et en activités instrumentales de la vie quotidienne (IADL). Les AVQ de base (BADL) ou AVQ physiques sont les compétences requises pour gérer les besoins physiques de base d'une personne, notamment l'hygiène personnelle ou la toilette, l'habillage, le transfert ou la marche et l'alimentation. Les Activités Instrumentales de la Vie Quotidienne (Instrumental Activities of Daily Living (IADL)) comprennent des activités plus complexes liées à la capacité de vivre de manière autonome au sein de la communauté. Cela comprendrait des activités telles que, par exemple, la gestion des finances et des médicaments, la préparation des aliments, l'entretien ménager et la lessive.

1) ADL de base

Les ADL de base telles que décrites par l'indice d'indépendance de Katz dans les activités de la vie quotidienne, comprennent six compétences essentielles [39] :

- Alimentation : La capacité humaine à se nourrir.
- S'habiller : La capacité de choisir des vêtements adaptés et de les porter.
- Hygiène personnelle : La capacité à se laver soi-même, à se toiletter, d'assurer l'hygiène dentaire, le soin des ongles et des cheveux.
- Déplacement : Le degré de capacité d'un individu à se déplacer d'un endroit à un autre et à marcher de manière autonome.
- Continence : La capacité à contrôler la fonction vésicale et intestinale.

- Toilette : La capacité à marcher jusqu'aux toilettes, à les utiliser correctement et à se nettoyer.

2) ADL instrumentales

Les AVQ instrumentales sont celles qui nécessitent des capacités de réflexion plus complexes, notamment des compétences organisationnelles. Elle, telles que définies par l'échelle de Lawton-Brody, comprennent les éléments suivants [40]:

- Nettoyage et entretien ménager
- Gestion de l'argent
- Gérer les médicaments et prendre les médicaments comme indiqué
- Préparation des repas
- Faire les achats d'épicerie et autres nécessité
- Transport, y compris changement de résidence et déménagement
- Utilisation d'appareils de communication, y compris le téléphone ou l'ordinateur

c) Classification des approches de reconnaissance de l'activité humaine

Diverses méthodes ont été utilisées pour reconnaître des activités. Ces approches peuvent être classées en deux grandes catégories : les approches basées sur la vision et les approches basées sur les capteurs [41], comme le montre la Figure 2-2.

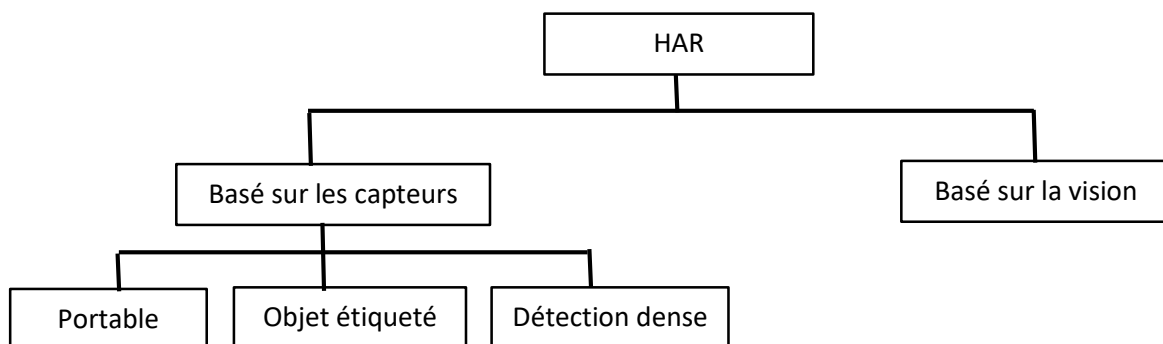


Figure 2-2 : Classification des approches de reconnaissance de l'activité humaine

L'une des premières approches dans ce domaine est l'approche basée sur la vision, dans laquelle des caméras sont utilisées pour collecter des informations sur les activités humaines. En appliquant des techniques de vision par ordinateur à ces données collectées, différentes activités peuvent être reconnues. Bien que les techniques basées sur la vision par ordinateur soient faciles à utiliser et puissent donner de bons résultats, cette approche

pose de nombreux problèmes. La confidentialité est la principale préoccupation. Un autre problème de cette méthode est sa dépendance à la lumière. Les caméras traditionnelles ne fonctionnent pas dans des conditions où il n'y a pas assez de lumière (par exemple la nuit). De nombreuses recherches ont été consacrées aux approches basées sur la vision, car il s'agissait de l'une des premières méthodes utilisées pour la reconnaissance d'activité [42]. Par conséquent, les techniques basées sur la vision n'ont pas été incluses dans notre thèse. En raison des faibles coûts et des progrès de la technologie des capteurs, la plupart des recherches dans le domaine du HAR se sont orientées vers une approche basée sur les capteurs. Dans cette approche, différents capteurs sont utilisés pour capturer le comportement humain lors qu'il effectue des activités de la vie quotidienne. Les solutions basées sur des capteurs peuvent être divisées en trois catégories principales en fonction du déploiement des capteurs : i) portables, ii) marquage d'objets (fixés à l'appareil ou objet-tagged) et iii) détection dense (marqué de manière environnementale/libres appareils ou device free) [43].

Dans la méthode portable, l'utilisateur doit porter le capteur tout en effectuant une activité. De nombreux travaux ont été réalisés sur la reconnaissance d'activité à l'aide de capteurs portables, mais le principal problème de cette approche est qu'il n'est pas toujours possible de porter le capteur. Par exemple, les personnes peuvent oublier de porter des capteurs ou refuser de les porter.

Pour les solutions utilisant l'étiquetage des objets, les capteurs sont attachés aux objets d'usage quotidien. En fonction de l'interaction de l'utilisateur avec ces objets, différentes activités seront reconnues. Il s'agit d'une approche liée au dispositif, c'est-à-dire que les utilisateurs ne doivent utiliser que des objets spécifiques (objets étiquetés). Comme l'approche précédente, elle peut également ne pas être réalisable en permanence, car elle oblige les utilisateurs à utiliser des objets étiquetés.

Ces dernières années, les chercheurs se sont concentrés sur les approches sans dispositif (capteurs denses), dans lesquelles les utilisateurs ne sont pas obligés de transporter une étiquette ou un dispositif avec eux. L'idée est de déployer des capteurs dans l'environnement (l'installation où l'activité est exercée) et lorsqu'une personne effectue une activité, des données sont collectées par ces capteurs, qui peuvent ensuite être utilisées pour la reconnaissance de l'activité. Une approche sans dispositif est plus pratique car elle

n'oblige pas l'utilisateur à porter un dispositif pendant qu'il effectue d'une activité. Mais cette approche présente également des difficultés, telles que l'interférence environnementale. Les données collectées peuvent être perturbées par l'environnement, ce qui peut provoquer du bruit dans les données.

2.4.4. La détection des anomalies

La détection d'anomalies dans les activités quotidiennes des patients âgés fait référence au processus d'identification de comportements ou de modèles inhabituels ou anormaux dans les activités quotidiennes des personnes âgées.

Elle implique une surveillance continue des actions, des mouvements, des signes vitaux et des habitudes des patients âgés. Lorsque des écarts ou des anomalies significatives sont détectés, ce système génère des alertes ou des notifications aux soignants, aux professionnels de la santé ou aux membres de la famille, ce qui permet d'intervenir à temps et d'améliorer les soins prodigués aux personnes âgées. Pour ce faire, on utilise souvent la technologie et des techniques d'analyse de données pour surveiller et évaluer leurs activités, à la recherche d'écarts par rapport aux normes ou aux routines établies.

Les raisons d'appliquer la détection d'anomalies aux activités quotidiennes des patients âgés peuvent varier :

- *Surveillance de la santé* : La détection d'anomalies peut contribuer à l'identification précoce de problèmes de santé. Par exemple, une diminution soudaine de la mobilité peut être le signe d'une chute ou d'une détérioration de l'état de santé (Paudel, 2018).
- *Sécurité* : Les anomalies peuvent indiquer des situations dangereuses, comme une cuisinière laissée allumée ou une porte laissée ouverte, qui peuvent présenter des risques pour la personne âgée.
- *Soins aux personnes atteintes de démence* : en cas de démence ou de déclin cognitif, le suivi des anomalies peut fournir des indications sur la progression de la maladie et aider les soignants à prendre les mesures qui s'imposent.
- *Aide à l'autonomie* : La détection des anomalies peut permettre aux personnes âgées de conserver leur indépendance en n'alertant les soignants qu'en cas de nécessité, ce qui minimise l'intrusion dans leur vie.

- *Efficacité* : Les établissements de soins peuvent utiliser la détection des anomalies pour optimiser l'affectation des ressources. Par exemple, le personnel peut concentrer son attention sur les cas où des anomalies sont détectées.

2.4.5. Les défis de la HAR

La reconnaissance de l'activité humaine est un domaine de recherche en plein essor qui présente de nombreux défis à relever. Bien que la reconnaissance de l'activité humaine soit un domaine bien étudié, ces défis doivent être étudiés de manière plus approfondie pour la réalisation efficace des systèmes de reconnaissance de l'activité humaine. Ces défis sont les suivants :

- a) *Le nombre de résidents* : Plus il y a de résidents, plus les traces d'activation des capteurs correspondant à des activités différentes sont entrelacées. Par exemple, si un résident "regarde la télévision" tandis qu'un autre "prend sa douche" au même moment, les changements d'état des capteurs correspondant à chaque activité de l'un et de l'autre ont tendance à se mélanger. En outre, même si un résident vit seul, d'autres résidents en visite temporaire peuvent effectuer des activités pendant qu'ils sont présents.
- b) *Compatibilité avec les données du monde réel* : Les données du monde réel sont souvent différentes des données de laboratoire en raison de l'environnement limité. La plupart des données du monde réel proviennent de flux et ne sont pas étiquetées ; les systèmes HAR doivent donc être suffisamment robustes pour les scénarios du monde réel.
- c) *Placement du capteur* : La position du capteur joue un rôle important dans la précision de la reconnaissance. Différentes positions sont possibles : bras droit/gauche, cheville, pied, hanche, poitrine, etc. Les lectures des signaux des capteurs varient en fonction de la position pour une même activité.
- d) *Le chevauchement des activités* : La plupart des systèmes HAR reconnaissent une seule activité à la fois, comme marcher, se tenir debout, s'asseoir ou se brosser les dents, mais il peut y avoir des chevauchements d'activités, comme prendre un café tout en regardant la télévision ou marcher tout en buvant de l'eau. Peu de recherches ont été effectuées dans ce domaine, mais il existe encore de bonnes possibilités de recherche dans cette direction.

- e) *Réglage des hyper-paramètres* : la précision des modèles profonds dépend fortement de l'ajustement des paramètres du réseau tels que le taux d'apprentissage, choix de fonction d'activation, choix d'algorithme d'optimisation, la taille du filtre, le nombre d'unités et de couches profondes, etc. Dans la plupart des recherches, ces paramètres sont définis à l'aide de méthodes heuristiques [44], il est donc nécessaire d'utiliser des algorithmes d'optimisation pour ajuster ces hyper-paramètres.
- f) *Sécurité* : le système HAR est affecté par des risques en matière de sécurité et de confidentialité, compte tenu du nombre croissant de capteurs, d'appareils et de différents types de participants tels que la personne âgée, les médecins et la famille. En outre, la transmission de données augmente le risque de détournement et d'attaques d'écoute clandestine dans les canaux de communication. Récemment, la technologie blockchain a été utilisée pour fournir un système fiable et efficace dans l'Internet des objets (IoT) tel que SHCS (Smart Home Control System) [45] ou proposition d'une architecture garantissant la confidentialité des données liées à la santé, telle que nous proposons dans [46] .

2.5. Conclusion

Dans ce chapitre, nous avons présentée une brève introduction aux concepts de HAR et les éléments à prendre en compte pour reconnaître les activités humaines à l'intérieur du domicile. À la fin, divers défis de recherche sont discutés qui peuvent être abordés pour rendre les systèmes HAR plus robustes et réalisables dans des scénarios du monde réel.

Dans le chapitre suivant, nous allons examiner en détail les approches connexes basées sur l'apprentissage automatique et profond pour la reconnaissance des activités et utilisées dans la littérature récente.

Chapitre 3 Revue de littérature

3.1 Introduction

La reconnaissance de l'activité humaine (HAR) à l'aide des capteurs portables est devenue un domaine de recherche et de développement essentiel, qui s'est d'abord appuyé sur des techniques d'apprentissage automatique avant d'intégrer des méthodes d'apprentissage profond. La reconnaissance des activités humaines implique l'application de modèles informatiques pour détecter, classer et analyser automatiquement diverses activités physiques effectuées par les personnes âgées dans leur vie quotidienne.

L'apprentissage automatique, avec ses méthodes établies de reconnaissance des modèles et de modélisation prédictive, a posé les bases du HAR dans ce contexte. Il a permis le développement de systèmes capables d'identifier les activités et les besoins des personnes âgées et d'y répondre. Toutefois, à mesure que la complexité des tâches de reconnaissance et la richesse des sources de données ont augmenté, l'apprentissage profond s'est imposé par sa capacité à extraire automatiquement des modèles complexes à partir de données de capteurs, améliorant ainsi de manière significative la précision des systèmes de HAR.

Dans ce chapitre, nous présenterons un aperçu général sur une variété de modèles d'apprentissage profond et même d'hybridation de modèles pour la reconnaissance automatique de l'activité humaine, ensuite nous allons effectuer une étude des principaux travaux en matière de HAR des capteurs portables et des modèles d'apprentissage automatique (ML) et d'apprentissage profond (DL). Ensuite, nous explorerons quelques travaux dans le domaine de la détection d'anomalies dans les activités des personnes âgées.

3.2 Les modèles ML et DL

3.2.1 Apprentissage automatique (Machine Learning)

L'apprentissage automatique est un domaine de l'intelligence artificielle (IA) qui se concentre sur le développement d'algorithmes et de modèles informatiques capables d'apprendre à partir de données et de faire des prédictions ou de prendre des décisions sans être explicitement programmés. Contrairement aux programmes informatiques traditionnels, dans lesquels les programmeurs humains écrivent des instructions explicites sur la manière dont une tâche doit être exécutée, les algorithmes d'apprentissage automatique permettent aux ordinateurs d'apprendre à partir de données passées afin d'améliorer leurs performances dans des tâches spécifiques. Ce processus d'apprentissage implique l'identification de modèles, de relations et de tendances dans les données, ce qui

permet aux machines de faire des prédictions ou de prendre des décisions éclairées lorsqu'elles sont confrontées à de nouvelles données non vues [47].

L'apprentissage automatique peut être divisé en trois catégories principales, chacune d'entre elles ayant un objectif différent et utilisant des techniques différentes : apprentissage supervisé, apprentissage non-supervisé et apprentissage semi-supervisé.

a) Apprentissage supervisé

L'apprentissage supervisé est une catégorie fondamentale de l'apprentissage automatique dans laquelle un algorithme apprend à partir d'un ensemble de données étiquetées pour faire des prédictions. Dans l'apprentissage supervisé, le jeu de données (dataset) utilisé pour l'apprentissage comprend des paires entrée-sortie, où les données d'entrée sont accompagnées d'étiquettes de sortie correspondantes ou de valeurs cibles [47]. L'apprentissage supervisé peut être appliqué à diverses tâches, notamment la classification et la régression :

- *Classification* : Dans les tâches de classification, l'objectif est d'affecter les données d'entrée à des catégories ou classes prédéfinies. Par exemple, classer des courriels comme étant du spam ou non, ou identifier des objets dans des images (par exemple, des chats ou des chiens).
- *Régression* : Dans les tâches de régression, l'objectif est de prédire une valeur numérique continue en sortie. Par exemple, prédire les prix des logements en fonction de caractéristiques telles que la superficie et l'emplacement.

b) Apprentissage non-supervisé

L'apprentissage non supervisé est une catégorie d'apprentissage automatique dans laquelle l'algorithme est entraîné sur un jeu de données sans supervision explicite ni résultats étiquetés. En d'autres termes, l'algorithme ne reçoit pas d'étiquettes cibles prédéfinies ou de résultats à partir desquels il peut apprendre. Au lieu de cela, son objectif principal est de découvrir par lui-même des modèles, des structures ou des relations au sein des données [47].

L'objectif principal de l'apprentissage non supervisé est souvent de découvrir des structures cachées ou latentes dans les données. Il peut être utilisé pour diverses tâches, notamment :

- *Clustering (regroupement)* : Regroupement de points de données similaires sur la base d'une certaine métrique de similarité. Les algorithmes de regroupement les plus courants sont K-Means, le regroupement hiérarchique et DBSCAN.
- *Réduction de la dimensionnalité* : Réduction du nombre de caractéristiques ou de variables dans les données tout en préservant les informations essentielles. Parmi les techniques populaires de réduction de la dimensionnalité, on peut citer l'analyse en composantes principales (PCA).

c) Apprentissage semi-supervisé

L'apprentissage semi-supervisé est un modèle d'apprentissage automatique qui combine des éléments d'apprentissage supervisé et non supervisé. Dans l'apprentissage semi-supervisé, l'ensemble de données d'apprentissage contient un mélange de données étiquetées et non étiquetées [47]. Cette approche vise à exploiter les avantages des données étiquetées (où le modèle apprend à partir d'exemples évidents) et de grandes quantités de données non étiquetées (où le modèle peut découvrir des modèles et des relations). L'apprentissage semi-supervisé est appliqué dans une variété de domaines, y compris le traitement du langage naturel, la vision par ordinateur et la reconnaissance vocale. Par exemple, dans la classification d'images, certaines images peuvent être étiquetées manuellement (par exemple, les chats et les chiens), tandis qu'un ensemble beaucoup plus important d'images non étiquetées est utilisé pour aider le modèle à se généraliser à une gamme plus large d'objets.

3.2.2 Modèles d'apprentissage profond (Deep Learning)

L'apprentissage en profondeur est une branche de l'apprentissage automatique dans laquelle les informations provenant des données sont traitées par chaque couche contenant un algorithme unifié avec une sorte de fonction d'activation. Dans chaque couche, des caractéristiques significatives des données sont construites pour l'apprentissage. L'histoire de l'apprentissage profond remonte à 1943, lorsque Walter Pitts et Warren McCulloch [48] ont créé des réseaux neuronaux basés sur un modèle informatique du cerveau humain. Le premier réseau neuronal convolutif a été développé par Kunihiko Fukushima en 1979 [49], qui utilise un réseau neuronal artificiel de conception hiérarchique et multicouche appelé Neocognitron.

À la fin des années 90, des progrès significatifs ont été réalisés, tels que le développement de machines à vecteurs de support (SVM) [50] et de mémoire long à court terme (LSTM) [51]. Mais au début des années 2000, avec le développement de la puissance de calcul des ordinateurs et des unités de traitement graphique (GPU), ont permis de développer de grands modèles avec une efficacité et une précision accrues qui pouvaient être entraînés avec des données volumineuses.

Dans la section suivante, nous décrirons certaines des architectures des réseaux d'apprentissage profond.

3.2.2.1 Réseau de neurones artificiels (ANN)

Les réseaux neuronaux artificiels sont des modèles informatiques inspirés des neurones biologiques du cerveau humain qui tentent de simuler un traitement de l'information et une exécution des tâches similaires. Les ANN sont composés de plusieurs couches interconnectées et successives. Chaque couche est représentée par un ensemble de neurones artificiels appelés nœuds. Chaque nœud est relié aux nœuds de la couche suivante par des liens, de la même manière que les neurones naturels et leurs connexions. Chaque lien est représenté par un poids numérique correspondant (weight) à l'impact de chaque nœud sur le nœud de la couche suivante. Les nœuds sont généralement organisés en plusieurs couches successives. La première couche est appelée couche d'entrée (Input layer) car elle reçoit les données d'entrée qui seront introduites dans le réseau. Les couches intermédiaires sont appelées couches cachées (hidden layers) et la dernière est appelée couche de sortie (Output layer). La typologie de la couche de sortie dépend essentiellement de la tâche du réseau, qu'il s'agisse de régression ou de classification, etc. La forme de base des ANN est également connue sous le nom de perceptron multicouches (MLP) ou de réseau Feed-forward qui contient seulement trois couches : entrée, cachée et sortie. Tandis que, un réseau neuronal profond avec deux couches cachées ou plus s'appelle Deep Feed-Forward neural network (DFNN) (Figure 3-1).

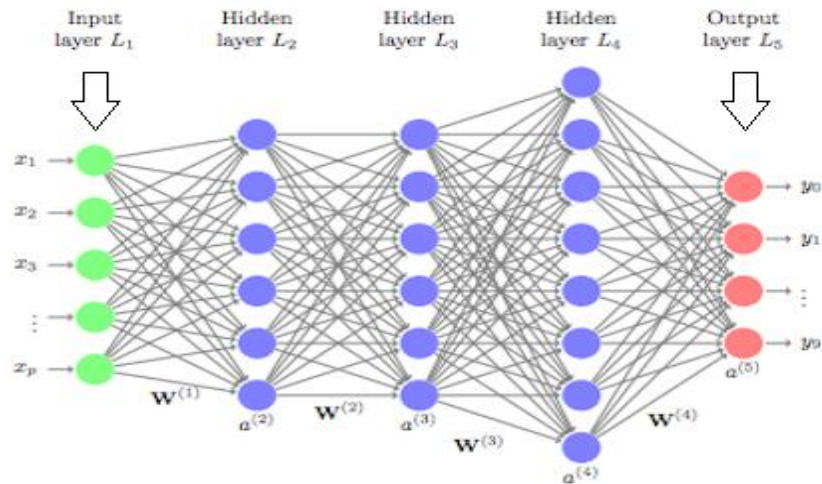


Figure 3-1: Exemple d'un réseau ANN avec 3 couches cachées

Ces réseaux sont conçus pour résoudre des problèmes non linéaires. C'est pourquoi ils sont souvent utilisés dans la classification des formes, la traduction des langues et la reconnaissance d'images. À travers chaque couche cachée, les valeurs de la couche précédente sont transformées avec une sommation pondérée $w_1x_1 + w_2x_2 + \dots + w_mx_m$. Plus formellement, elle s'écrit comme suit :

$$a = f\left(\sum_j w_j x_j + b\right) \quad (3-1)$$

Où : w_j est le vecteur de poids, x_j sont les valeurs d'entrée, b est un vecteur de biais, f est une fonction d'activation non linéaire et a est l'activation du neurone.

Ces réseaux nécessitent un réglage d'un certain nombre d'hyperparamètres. Cela se fait généralement en évaluant les résultats du modèle dans l'ensemble de test/validation. Une vue générale des hyperparamètres d'un modèle de réseau neuronal est donnée ci-après.

A) Hyperparameters

- **La fonction d'activation**

Dans le deep learning, les fonctions d'activation sont l'un des paramètres essentiels dans l'apprentissage et la construction d'un modèle de deep learning qui fait des prédictions précises.

Une fonction d'activation décide si un neurone doit être activé (déclenché) ou non en fonction de sa pertinence pour la prédiction finale du modèle. De plus, une fonction d'activation non linéaire introduit une non-linéarité dans le modèle permettant de s'adapter à des tâches plus complexes. Pour améliorer l'ensemble du modèle, lorsque cette fonction

de perte est optimisée via un algorithme pour trouver le nombre minimum d'erreurs possible dans le modèle, on parle de descente de gradient.

Plusieurs fonctions d'activation différentes ont été utilisées dans la littérature. Cependant, Linear, Sigmoid, Tanh et ReLU sont les plus couramment utilisés.

1) Sigmoid

La fonction sigmoïde est l'une des fonctions d'activation les plus utilisées en apprentissage automatique et en apprentissage profond. Il peut être utilisé dans les couches cachées, qui prennent la sortie de la couche précédente et amènent les valeurs d'entrée entre 0 et 1. La formule de la fonction d'activation sigmoïde est :

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3-2)$$

2) ReLU

ReLU (Rectified Linear Unit), elle est devenue l'une des fonctions d'activation les plus utilisées en apprentissage profond. Elle produit une sortie maximale entre 0 et x. Sa formule est indiquée dans l'équation (3-3).

$$f(x) = \max(0, x) \quad (3-3)$$

Où, x est la valeur d'entrée.

3) Tanh

Tanh réduit la sortie entre [-1, 1], elle est centrée sur zéro et est souvent utilisée pour les tâches de classification binaire.

$$f(x) = \tanh(x) \quad (3-4)$$

4) Softmax

La fonction softmax est principalement utilisée après la dernière couche d'un réseau neuronal qui joue le rôle de classificateur. La fonction softmax réduit les scores de sortie d'un noeud j dans un intervalle de [0, 1] et la somme totale sur l'ensemble des K unités est égale à 1 (Equation (3-5)).

$$f(x) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad (3-5)$$

Ici, le z représente les valeurs des neurones de la couche de sortie.

- **La fonction Loss (perte)**

Tout modèle d'apprentissage profond a pour objectif principal de réduire l'erreur ou la perte entre les sorties réelles et prédites. Les fonctions de perte sont utilisées pour représenter mathématiquement cette erreur. Pendant l'entraînement, le modèle apprend à améliorer ses prédictions et à prendre des décisions précises en optimisant la fonction loss. Voici une liste étendue de différents types de fonctions de perte couramment utilisées en apprentissage profond :

1) *Mean Absolute Error (MAE)*

MAE (Erreur absolue moyenne), également connu sous le nom de perte L1, surmonte l'inconvénient de l'erreur de moyenne (ME) mentionné précédemment. Il calcule la distance absolue entre la sortie réelle et la sortie prévue.

L'erreur moyenne absolue (MAE) est également la fonction de perte la plus simple. Pour calculer le MAE, vous prenez la différence entre la valeur réelle et la prédiction du modèle et vous en faites la moyenne sur l'ensemble de données (Equation (3-6)).

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (3-6)$$

Où : y_i : valeurs réelles

\hat{y}_i : Prédiction du réseau neuronal

2) *Mean Squared Error (MSE)*

MSE (Erreur quadratique moyenne), également connu sous le nom de perte L2, est utilisé le plus fréquemment et améliore avec succès les inconvénients du ME et du MAE. Il calcule le «carré» de la distance entre la sortie réelle et la sortie prévue, évitant ainsi les possibilités d'erreur négative. Ses extensions les plus améliorées sont l'erreur quadratique moyenne (RMSE) et l'erreur logarithmique quadratique moyenne (RMSLE).

La formule de calcul du MSE est donnée par l'équation (3-7) :

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \quad (3-7)$$

3) *Binary Cross Entropy*

Binary Cross Entropy (Entropie croisée binaire) compare chacune des probabilités prédites à la sortie réelle de la classe qui peut être 0 ou 1. Ensuite, elle calcule le score qui pénalise les probabilités en fonction de la distance par rapport à la valeur attendue. Cela indique à quel point la valeur est proche ou éloignée de sa valeur réelle (Equation (3-8)).

$$BCE = -\frac{1}{n} \sum_i^n (y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)) \quad (3-8)$$

4) Categorical Cross Entropy ou Softmax Loss

Categorical Cross Entropy (la perte d'entropie croisée catégorielle) est utilisée dans les tâches de classification multi-classes. Ce type de fonction comme la fonction binaire, quantifie la différence entre les probabilités prédites et les étiquettes catégorielles réelles. La formule de perte d'entropie croisée de la manière suivante :

$$Cross - Entropy = -\sum_i^c y_i \log(\hat{y}_i) \quad (3-9)$$

Où, c représente le nombre total de classes (output), \hat{y}_i spécifie la sortie réelle et y_i est la sortie prédite.

- **Epoch**

Dans l'apprentissage profond, un epoch correspond à un passage unique sur l'ensemble d'apprentissage [52]. Cela inclut la propagation vers l'avant et vers l'arrière, de sorte que le modèle a mis à jour les paramètres une fois. Au fur et à mesure que le nombre d'epoch augmente, l'ajustement du modèle s'améliore, bien qu'un sur-apprentissage (overfitting) puisse se produire avec un grand nombre d'époques.

- **Batch Size (lot)**

Un batch est un sous-ensemble de données passant par le réseau. Chaque batch fait l'objet d'un passage complet en avant et en arrière avant qu'une epoch ne soit atteinte. La taille du batch est le nombre d'exemples d'apprentissage utilisés dans chaque passage (batch). Lorsque tous les lots sont passés par le réseau, un epoch est terminée. L'utilisation d'une petite taille de batch permet d'économiser de la mémoire, ce qui devient important lorsque l'on travaille avec de grands ensembles de données. En règle générale, les réseaux neuronaux s'entraînent généralement plus rapidement avec des lots plus petits.

- **Learning Rate (taux d'apprentissage)**

Le taux d'apprentissage est un hyperparamètre utilisé dans l'algorithme de gradient descent. L'algorithme estime le gradient de l'erreur et met à jour les poids du modèle en conséquence.

L'échelle à laquelle les poids sont ajustés est connue sous le nom de taux d'apprentissage. Il s'agit d'une petite valeur positive généralement comprise entre 0 et 1. Le choix de la valeur du taux d'apprentissage a un impact à la fois sur la vitesse et l'accuracy de la convergence du modèle. S'il est trop bas, le modèle aura besoin de plus d'époques pour converger. Il peut également ne pas converger ou rester "coincé" dans un minima local. S'il est trop grand, le modèle nécessite moins d'époques, mais les performances du modèle varient de manière significative et il peut se terminer par un ensemble de poids sous-optimal [53].

B) Backpropagation

La backpropagation est un algorithme d'apprentissage supervisé utilisé pour entraîner des réseaux de neurones artificiels. Il s'agit d'une méthode de calcul du gradient d'une fonction de perte par rapport aux poids du réseau, qui permet de mettre à jour les poids lors de l'entraînement. L'idée de base de la backpropagation est de propager l'erreur de la couche de sortie du réseau vers la couche d'entrée, tout en ajustant les poids en cours du chemin. Ce processus se déroule en deux étapes : propagation vers l'avant et propagation vers l'arrière.

Lors de la propagation vers l'avant, les entrées sont transmises au réseau et la sortie est calculée. La sortie est ensuite comparée à la sortie souhaitée et l'erreur est calculée.

Lors de la propagation vers l'arrière, l'erreur se propage à travers le réseau, en commençant par la couche de sortie et en remontant jusqu'à la couche d'entrée. L'erreur au niveau de chaque couche est utilisée pour calculer le gradient de la fonction de perte par rapport aux poids de cette couche.

Le gradient est ensuite utilisé pour mettre à jour les poids du réseau, à l'aide d'un algorithme d'optimisation tel que la descente de gradient stochastique. Ce processus est répété plusieurs itérations, jusqu'à ce que le réseau ait appris à classer correctement les entrées.

Algorithme de Gradient Descent

Gradient Descent est un algorithme d'optimisation permettant de trouver un minimum local d'une fonction. Dans le contexte des réseaux de neurones, l'objectif de l'algorithme de gradient descent est d'ajuster les paramètres du modèle (comme les poids et les biais) pour minimiser la fonction de perte, rendant le modèle aussi précis que possible.

Le gradient descent est un processus itératif dans lequel le modèle converge progressivement vers une valeur minimale, et si le modèle itère au-delà de ce point, il ne produit que peu ou pas de changements de perte (voir Figure 3-2)).

En terme général, l'algorithme se déroule comme suit :

- Pour chaque exemple d'apprentissage, calculer le gradient de la fonction de perte en fonction de chaque paramètre de poids et de biais.
- Calculer le gradient moyen pour tous les poids et biais.
- Mettre à jour les poids et les biais à l'aide de la règle de mise à jour : (Equations (3-10), (3-11)).

$$w = w - \epsilon \frac{\partial F}{\partial w} \quad (3-10)$$

$$b = b - \epsilon \frac{\partial F}{\partial b} \quad (3-11)$$

Où, ϵ : learning rate, F : Fonction de perte, W : matrice de poids, b : matrice de biais, la dérivée de F dans w ou b peut être calculée en utilisant les dérivées partielles de F dans les poids ou biais individuels.

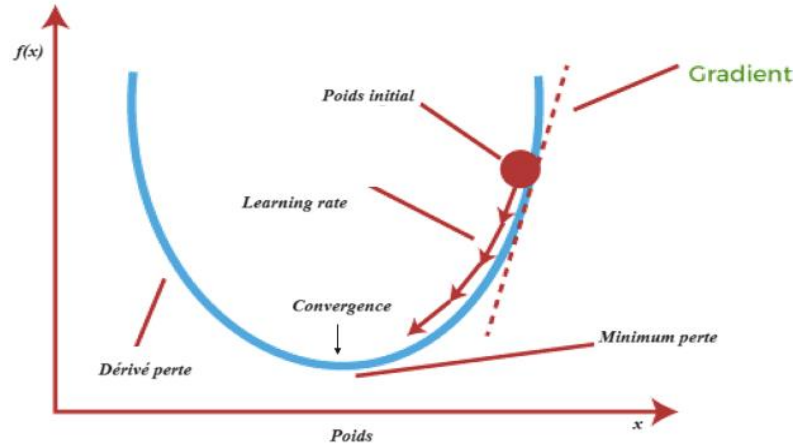


Figure 3-2 : Principe de fonctionnement de la descente de gradient

Il existe de nombreuses techniques d'optimisation qui sont généralement basées sur des méthodes de descente de gradient, mais certaines des plus populaires sont :

- Stochastic gradient descent (SGD)
- Root Mean Squared Propagation (RMSProp)
- Adaptive Moment Estimation (Adam)

3.2.2.2 Réseau de neurones convolutifs (CNN)

Un réseau de neurones convolutifs (CNN) fait partie de l'apprentissage profond qui reconnaît et classe les caractéristiques des images d'entrée en leur attribuant des poids et des biais pouvant être appris, et en les séparant les unes des autres. Ces dernières années, il a été grandement amélioré dans de nombreux domaines tels que la classification d'images et de vidéos, la classification d'images de vision par ordinateur, la reconnaissance faciale, l'analyse d'images et le traitement du langage naturel. L'un des principaux avantages d'un réseau CNN est sa capacité à capturer les dépendances spatiales et temporelles de l'image et à réduire l'image sans perdre de caractéristiques, permettant de concevoir un modèle plus évolutif lors de la prédiction.

L'architecture de base d'un CNN se compose de deux parties. Les parties d'extraction de caractéristiques et les parties de classification comme le montre la Figure 3-3.

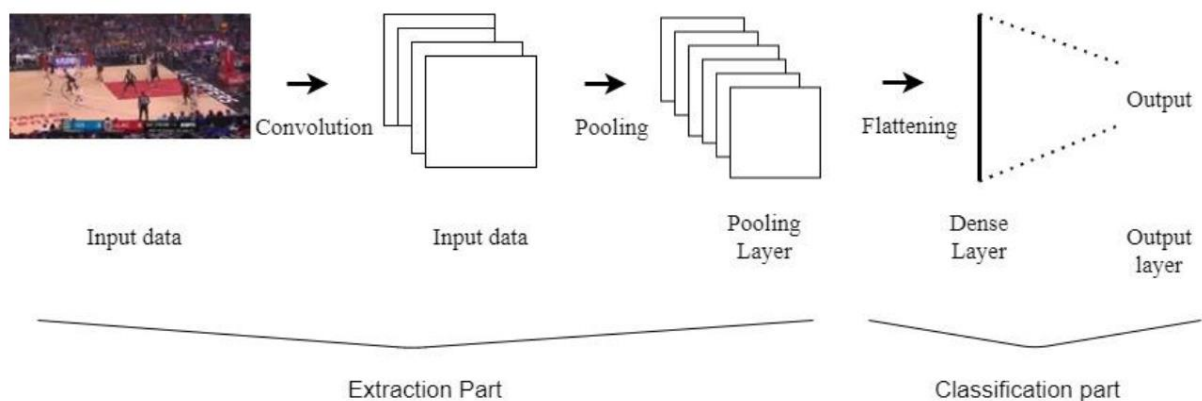


Figure 3-3 Architecture du réseau CNN

L'extraction de caractéristiques se compose généralement de deux couches. La couche de convolution et la couche de regroupement. La couche de convolution est la première couche de l'architecture CNN. Elle permet d'extraire différentes caractéristiques de l'image d'entrée ou des images vidéo à différentes étiquettes. Dans la couche de convolution, l'image d'entrée est convoluée avec un filtre en faisant glisser le filtre sur l'image d'entrée. La sortie de la couche de convolution est appelée carte de caractéristiques (feature map) ou carte d'activation. L'étape de convolution peut être 2D ou 3D en fonction du type d'image d'entrée. La couche initiale extrait diverses caractéristiques de bas niveau de l'image, telles que les bords, la couleur des coins, l'orientation des gradients, etc. Ces données seront transmises à d'autres couches d'étiquettes supérieures afin d'extraire d'autres caractéristiques de l'image d'entrée. Les couches de convolution d'étiquettes supérieures

extraient des caractéristiques de haut niveau qui fournissent des informations complètes sur l'image d'entrée.

La deuxième couche est la couche de regroupement (Pooling). Cette couche est une opération d'échantillonnage descendant qui diminue la taille spatiale de la carte des caractéristiques de la couche de convolution en réduisant la dimensionnalité de la connexion entre les couches. Cela permet de réduire la puissance de calcul nécessaire. Cette couche permet également d'extraire les caractéristiques dominantes de la carte des caractéristiques.

Le regroupement maximal (Max pooling), le regroupement moyen (average pooling) et le regroupement de la somme (sum pooling) et Global Average Pooling ou GAP pooling sont quelques-unes des méthodes de regroupement qui ont été utilisées avec succès pour l'architecture CNN (voir Figure 3-4)).

L'architecture CNN se compose d'un certain nombre de couches convolutives et de couches de regroupement, en fonction de la complexité de l'image, afin de capturer davantage de caractéristiques.

Les couches entièrement connectées sont les couches des parties de classification où les caractéristiques extraites de la partie d'extraction des caractéristiques sont aplaties et ensuite utilisées pour prédire les étiquettes de sortie ou la classe de l'image d'entrée. Cette couche utilise les poids et les biais qui relient les neurones pour apprendre des fonctions non linéaires dans cet espace. Elle est utilisée pour optimiser des objectifs tels que les scores de classe. Cette couche est suivie par la couche de sortie qui classifie l'image à l'aide des techniques de classification SoftMax. Outre ces couches, une autre couche très utilisée dans l'architecture du CNN est la couche d'exclusion (dropout). Cette couche permet de réduire le problème de sur-apprentissage (overfitting) en désactivant certains neurones du réseau au cours de l'apprentissage.

De nombreuses architectures CNN ont été développées au cours des deux dernières décennies et ont permis d'obtenir des résultats remarquables dans le domaine de l'intelligence artificielle. La première est LeNet-5, développée en 1998. Parmi les autres modèles principaux et leur architecture, citons Alex Net (2000) [54], GoogleLeNet/Inception, Inception ResNets (2017), ResNext-50 (2019).

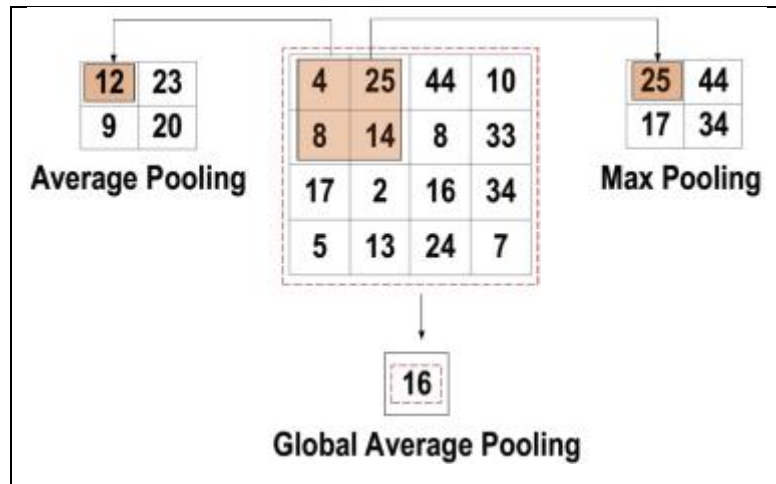


Figure 3-4 : Trois types de l'opération de regroupement (Pooling)

3.2.2.3 Réseau de neurones récurrents

Les réseaux neuronaux récurrents (RNN) sont une classe de modèles d'apprentissage profond qui possèdent une mémoire interne, ce qui leur permet de capturer les dépendances séquentielles. Contrairement aux réseaux neuronaux traditionnels qui traitent les entrées comme des entités indépendantes, les RNN prennent en compte l'ordre temporel des entrées, ce qui les rend adaptés aux tâches impliquant des données séquentielles [55]. En utilisant une boucle, les RNN appliquent la même opération à chaque élément d'une série, le calcul actuel dépendant à la fois de l'entrée actuelle et des calculs précédents [56].

La capacité des RNN à utiliser des informations contextuelles est particulièrement précieuse dans des tâches telles que le traitement du langage naturel, la classification vidéo et la reconnaissance vocale. Par exemple, dans la modélisation du langage, la compréhension des mots précédents dans une phrase est cruciale pour prédire le mot suivant. Les RNN excellent dans la capture de ces dépendances en raison de leur nature récurrente [57].

Toutefois, les RNN simples sont limités par leur mémoire à court terme, qui restreint leur capacité à retenir des informations sur de longues séquences [58]. Pour y remédier, des variantes de RNN plus avancées ont été développées, notamment la mémoire à long terme (Long short-term Memory ou LSTM) [51], [59].

Figure 3-5 illustre un réseau neuronal récurrent simple, où l'état caché (h_t) est calculés à l'aide de l'équations (3-12) [60] :

$$h_t = \sigma(Wh_{t-1} + Ux_t + b_t) \quad (3-12)$$

Où, σ est une fonction sigmoïde, x_t est un vecteur d'entrée à l'instant t , h_t est un vecteur d'état caché à l'instant t , W est une matrice de poids entrée vers caché, U est une matrice de poids caché vers caché, et b_t est un terme de biais.

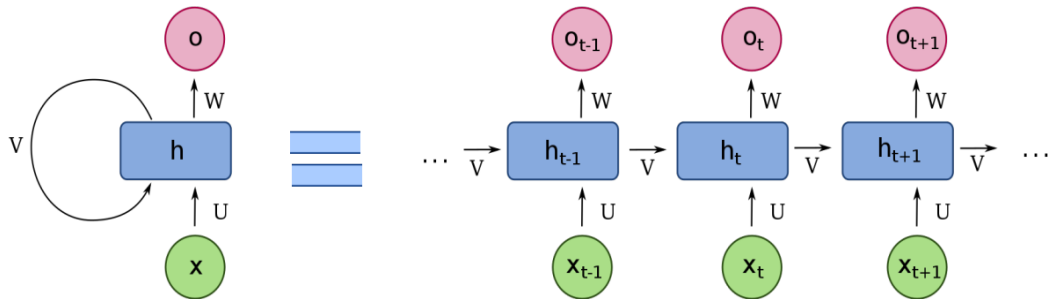


Figure 3-5 : Architecture RNN simple

➤ **Réseau mémoire à court et long terme (LSTM)**

La mémoire à court et long terme (LSTM) est une variante avancée des réseaux neuronaux récurrents (RNN) qui permet de capturer les dépendances à long terme. LSTM a été initialement introduite par [51] en 1997 et améliorée par [61] en 2013, gagnant une popularité significative dans la communauté de l'apprentissage profond. Comparés aux RNN standard, les modèles LSTM se sont avérés plus efficaces pour retenir et utiliser les informations sur des séquences plus longues [62].

Dans un réseau LSTM, l'entrée actuelle à un pas de temps (timestep) spécifique et la sortie du pas de temps précédent sont introduites dans l'unité LSTM, qui génère alors une sortie qui est transmise au pas de temps suivant. La dernière couche cachée du dernier pas de temps, est généralement utilisée à des fins de classification. L'architecture globale d'un réseau LSTM est illustrée à la Figure 3-6.

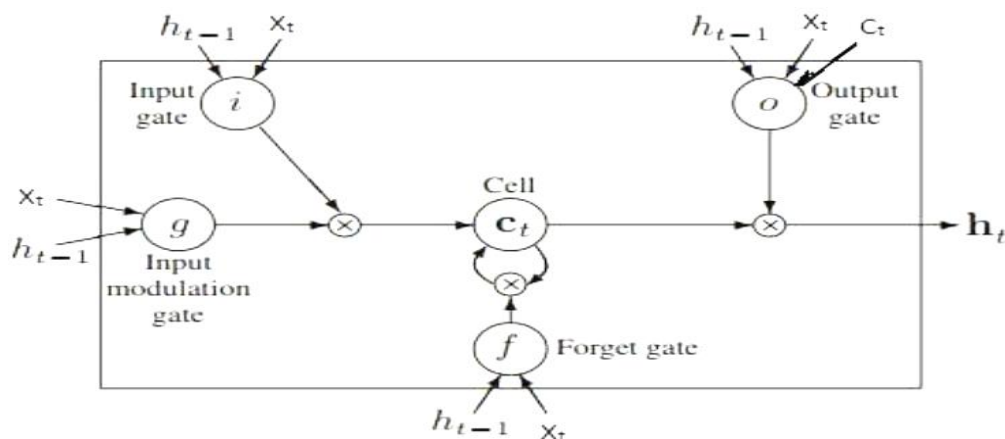


Figure 3-6 : Cellule unique de réseau LSTM

Dans le réseau LSTM, les cellules des couches cachées possèdent trois portes : une porte d'entrée (input gate), une porte de sortie (output gate) et une porte d'oubli (forget gate), ainsi qu'un neurone qui se connecte lui-même et peut stocker et supprimer des données dans l'état de la cellule. La Figure 3-7 représente l'architecture d'une seule cellule d'un réseau LSTM.

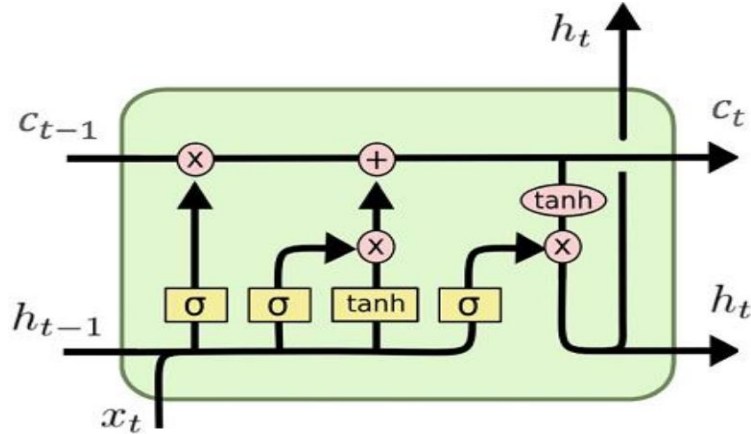


Figure 3-7 architecture du modèle LSTM

Toutes les portes sont généralement contrôlées par une fonction d'activation de la fonction sigmoïde ou tanh.

Par exemple, la porte d'entrée i_t est responsable de la mise à jour des informations. La porte d'oubli f_t traite les informations de la porte d'entrée i_t et de l'état de cellule précédent C_{t-1} . Si nécessaire, les informations sont également supprimées de l'état actuel C_t . La porte de sortie o_t transmet la sortie finale à l'unité LSTM suivante et conserve la valeur de sortie pour prédire la séquence suivante. D'autre part, l'unité récurrente \hat{C}_t estime l'état de la cellule précédente C_{t-1} et la valeur d'entrée actuelle x_t à l'aide de la fonction d'activation tanh. La valeur de h_t peut être calculée par le produit scalaire de o_t et tanh de C_t . Enfin, nous pouvons obtenir la sortie finale en passant h_t au classificateur Softmax. Mathématiquement, les opérations des portes susmentionnées peuvent être exprimées comme par les équations ci-après :

$$i_t = \sigma(W_{it} \cdot x_t + W_{it} \cdot h_{t-1} + b_i) \quad (3-13)$$

$$f_t = \sigma(W_{ft} \cdot x_t + W_{ft} \cdot h_{t-1} + b_f) \quad (3-14)$$

$$o_t = \sigma(W_{ot} \cdot x_t + W_{ot} \cdot h_{t-1} + b_o) \quad (3-15)$$

$$g_t = \tanh(W_{gt} \cdot x_t + W_{gt} \cdot h_{t-1} + b_g) \quad (3-16)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot g_t \quad (3-17)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (3-18)$$

Où x_t représente l'entrée au pas de temps t , σ désigne la fonction d'activation sigmoïde, W et b sont les matrices de poids et les vecteurs de biais, h_t : l'unité cachée dans le temps t , i_t : porte d'entrée, f_t : porte d'oubli, o_t : porte de sortie et $(.)$ indique la multiplication de la matrice.

3.2.2.4 Le réseau hybride

Un réseau hybride est conçu en fusionnant deux classificateurs ou plus. Il intègre deux ou plusieurs modèles, les entraîne et combine leurs prédictions pour obtenir de meilleurs résultats. Cette méthode est également connue sous le nom d'apprentissage d'ensemble en termes d'apprentissage automatique. Il utilise plusieurs modèles pour prédire le même problème, où les sorties de différents modèles sont moyennées pour prédire la sortie. De plus, des modèles tels que CNN et LSTM sont combinés pour apprendre les caractéristiques spatiales et temporelles des données d'entrée [63], [64]. Par exemple, pour la classification d'images, CNN génère des informations spatiales de haut niveau de l'activité des images, tandis que le modèle RNN est utilisé pour extraire la corrélation temporelle entre des séquences d'images consécutives en conservant la mémoire des images précédentes [65].

3.3 Travaux connexes

3.3.1 Reconnaissance des activités humaines (HAR)

- **Apprentissage automatique (ML) dans HAR**

Pendant la pandémie précédente de Covid-19, de nombreuses personnes sont mis en quarantaine dans des endroits désignés ou isolés à la maison ; Ainsi, il y a un besoin accru de surveiller leurs activités physiques.

Plusieurs études sur la reconnaissance de l'activité humaine ont été réalisées, en particulier des approches conventionnelles utilisant des algorithmes d'apprentissage automatique tels que Random Forest (RF), Naive Bayes, k-nearest Neighbours (kNN) ou Support Vector Machines (SVM)etc. Dans ce contexte, on peut citer quelques études antérieures et plus récentes en HAR :

Une étude présentée par Tran et Phan [66] pour la reconnaissance de six activités principales qui sont : la marcher, la position debout, la position assise, la position allongée, la montée et la descente. Les activités ont été capturées par les capteurs du mobile. Un ensemble de caractéristiques a été sélectionné, comme la différence entre les données des capteurs, la magnitude, la moyenne, l'écart type, le maximum, le minimum, l'énergie et l'entropie. Le classificateur SVM a été utilisé et la qualité des caractéristiques et la qualité de l'apprentissage du modèle contribuent à améliorer les performances. Le nombre de caractéristiques obtenues à partir de ce modèle est de 248. L'accuracy obtenu avec le système proposé est de 89,59 %.

Un travail étonnant sur la classification des activités pour les personnes âgées est rapporté dans [67] , où l'idée de chaussures instrumentées capables d'enregistrer les mouvements est introduite dans le but de discriminer les transitions posturales, la locomotion et les activités de marche à l'aide d'arbres de décision. Les résultats indiquent un taux global de détection erronée de 2,8 %. L'accuracy atteint par le système proposé était de 97,2%.

De même, Vakili et al [68] ont proposé un modèle HAR en temps réel pour la prédiction en ligne des mouvements physiques humains basé sur des capteurs inertiels de smartphone. Au total, 20 activités différentes ont été sélectionnées et six algorithmes d'apprentissage incrémental ont été utilisés pour vérifier les performances du système, qui ont ensuite tous été comparés aux algorithmes HAR de l'état de l'art tels que les arbres de décision (DT),

AdaBoost, etc. Les algorithmes d'apprentissage incrémental k-NN et naive Bayes donnent le plus grand accuracy de 95%.

Dans l'article de [69], ils ont présenté un nouveau système de surveillance des personnes âgées à l'aide d'un algorithme d'apprentissage automatique. Dans cette étude, ils ont utilisé un modèle k-NN avec une valeur k de 5, qui a atteint l'accuracy le plus élevé de 96,40% pour détecter neuf activités en temps réel des personnes âgées. Le système développé peut surveiller les personnes âgées en temps réel et fournir des alertes via un dispositif de messagerie.

- **Apprentissage profond (DL) dans HAR**

Bien que les méthodes conventionnelles aient obtenu des résultats prometteurs avec une grande précision de classification dans divers environnements contrôlés, ces méthodes sont basées sur une extraction des caractéristiques manuelles, ce qui nécessite des efforts considérables dans l'étape de prétraitement des données. En outre, la discrimination d'activités très similaires pour ces méthodes reste une tâche difficile. Les algorithmes d'apprentissage profond (DL) semblent être une bonne solution pour surmonter ces problèmes, car ils effectuent une modélisation structurelle couche par couche pour extraire des caractéristiques spécifiques et permettre le processus de classification.

Les architectures d'apprentissage profond, telles que les réseaux neuronaux convolutifs (CNN), LSTM, ainsi que des modèles hybrides CNN-LSTM ont également été exploitées pour la reconnaissance des activités humaines.

Chernbumroong et al. [70] ont proposé un modèle de reconnaissance et de classification des activités pour les personnes âgées. Ils ont recueilli des données à l'aide de capteurs portés au poignet. Cependant, leur classification n'était précise qu'à 90 % au maximum. Pour remédier à ce problème, les auteurs ont proposé de combiner les données de température et d'accélération.

Dans l'étude [71], les auteurs ont utilisé une approche d'apprentissage profond pour la reconnaissance de l'activité humaine à l'aide de capteurs portables. Ils ont proposé des modèles profonds, convolutifs et récurrents, et ont évalué leurs modèles à l'aide de trois jeux de données. Leurs résultats de classification ont atteint un score F1-score de 90,4 % pour le DNN (Réseau neuronal profond), 93,7% pour le CNN, 88,2 % pour le LSTM et 86,8% pour le bi-directional LSTM.

Cheng et al. [72] ont développé un modèle neuronal profond basé sur la combinaison entre le réseau neuronal ANN et le réseau neuronal récurrent. Trois jeux de données différents ont été utilisés : Opportunity, PAMAP2 et smartphones. Les résultats obtenus (accuracy) avec le réseau neuronal d'origine uniquement sont respectivement de 85,1%, 88% et 91,5% sur les trois jeux de données, tandis que le modèle RNN a montré de précision de 91,5%, 90,5% et 91% respectivement.

Yashi Nan et al [73] ont utilisé une combinaison de CNN et de LSTM pour détecter l'activité chez les personnes âgées. Ils ont développé un algorithme de reconnaissance d'activité pour les données d'accéléromètre des smartphones des 53 personnes âgées. Les activités ont été classées en position couchée, assise, debout, transition, marche, montée et descente. Ce modèle a donné l'accuracy la plus élevée de 81,1 %.

Md Zia Uddin et al. [35] ont proposé un système de reconnaissance d'activités basé sur des capteurs corporels et utilisant un apprentissage neuronal profond basé sur la mémoire à long-moyen terme (LSTM). Le modèle proposé atteint un recall de 99 %, qui est ensuite comparé aux modèles d'apprentissage profond existants tels que RNN, CNN.

Baoding Zhou et al [74] ont proposé un CNN pour la reconnaissance de l'activité humaine en intérieur. Au total, neuf activités différentes sont reconnues sur la base d'un accéléromètre, d'un magnétomètre, d'un gyroscope et d'un baromètre collectés par smartphone. La méthode proposée a atteint un excellent accuracy de 98%.

Hayat et al. [75] ont développé un système de surveillance automatique de l'activité qui détectera automatiquement l'activité des personnes âgées à l'aide de différentes méthodes d'apprentissage automatique et profond telles que RF, k-NN, SVM, ANN et LSTM. Parmi toutes les méthodes, la méthode LSTM a donné le meilleur accuracy de 95,05 %.

3.3.2 La détection des anomalies dans les activités humaines

La détection d'anomalies fait référence au problème de trouver des modèles dans les données qui ne sont pas conformes au comportement attendu [76]. Les anomalies sont généralement rares et inattendues, ce qui entraîne un manque de données d'apprentissage. Cependant, quelques études ont examiné la possibilité d'identifier des comportements anormaux en trouvant des modèles de comportement différents des modèles normaux appris. De nombreuses études ont démontré la faisabilité de l'apprentissage d'un classificateur pour détecter un événement spécifique, en particulier les chutes. Cette section passe en revue les approches récentes de la détection des anomalies.

Arifoglu et al [77] ont proposé une méthode de détection d'anomalies utilisant un algorithme de réseau neuronal récurrent (RNN) pour observer et détecter des anomalies dans la séquence d'activité humaine chez les personnes âgées présentant des symptômes de démence. Pour évaluer les performances de la détection des comportements anormaux, ils ont utilisés les mesures d'évaluation suivantes : Taux de vrais positifs (TPR) et Taux de faux positifs (FPR). Le TPR est le pourcentage d'activités anormales correctement détectées par rapport au nombre total d'activités anormales, le FPR est le pourcentage d'activités normales faussement détectées comme des activités anormales par l'algorithme par rapport au nombre total d'activités normales. Le modèle obtenu un taux de 91,43% de vrais cas positifs (TPR) et un taux de 40,96% de cas faux positifs (FPR).

Yahaya et al [78] ont proposé un algorithme de détection d'anomalie, appelé machine à vecteur de support monocouche (SVM), qu'ils ont appliqué pour détecter les anomalies dans les activités de la vie quotidienne. Plus précisément, ils ont étudié les anomalies des habitudes de sommeil, qui pourraient être le signe d'un léger déclin cognitif chez les personnes âgées ou d'autres problèmes de santé.

Aran et al [79] ont proposé une approche pour observer et modéliser automatiquement le comportement quotidien des personnes âgées et détecter d'éventuelles anomalies dans les données des capteurs. Dans la méthode qu'ils suggèrent, les anomalies pourraient être le signe d'un problème de santé. Ils ont donc créé un modèle probabiliste spatio-temporel pour résumer le comportement quotidien.

Zerkouk, M. et Chikhaoui, B. [80] ont proposé une méthode d'extraction de caractéristiques multiples utilisant LSTM pour identifier et prédire les comportements anormaux des personnes âgées dans les maisons intelligentes. Leur travail a ensuite été étendu dans [81] CNN, LSTM, CNN-LSTM et l'auto-encodeur CNN-LSTM sont utilisés pour détecter des activités anormales chez les personnes âgées. Deux jeux de données sont utilisés : SIMADL et MobiAct. MobiAct est enregistré à l'aide d'un smartphone dans la poche d'un sujet pendant qu'il effectue des activités physiques telles que sauter, s'asseoir et marcher. Puisque le nombre d'activités anormales est inférieur au nombre d'activités normales, ils ont utilisés la technique statistique SMOTE pour suréchantillonner les activités anormales et résoudre le problème de déséquilibre des classes. Les résultats montrent que la méthode CNN-LSTM prend en compte à la fois les informations temporelles et spatiales, atteignant ainsi un accuracy de 93% sur MobiAct et de 98 % sur SIMADL.

3.3.3 Discussion

A) Travaux dans la reconnaissance des activités humaines (HAR)

La plupart des études axées sur la reconnaissance des activités humaines globales ne sont pas spécifiques aux personnes âgées et ne sont pas évaluées par rapport à des critères communs. En plus, Elles ont été limitées par l'utilisation de caractéristiques de jeux de données publiques, de données collectées dans un environnement contrôlé, et d'un nombre limité de participants âgés et d'activités.

De manière générale, la plupart des études rapportées indiquent que les modèles DL surpassent les modèles ML traditionnels en ce qui concerne le problème HAR. Cependant, chaque modèle DL a ses propres forces et faiblesses, et aucun modèle DL ne peut relever tous les défis du HAR. Par exemple, les CNN sont très puissants pour extraire les caractéristiques locales appropriées à partir des données des capteurs. Cependant, ils sont sans mémoire et ignorent les dépendances temporelles entre les enregistrements de données. D'un autre côté, les LSTM conviennent bien aux problèmes dans lesquels les dépendances temporelles jouent un rôle important.

Dans notre thèse, nous proposons un système de reconnaissances des activités physiques chez les personnes âgées basé sur l'apprentissage profond à travers quelques expériences sur divers modèles CNN et LSTM et leur hybridation en ce qui concerne les mesures de performance les plus importantes, par exemple l'accuracy, la precision et le recall.

B) Travaux dans la détection des anomalies dans les activités humaines

Les modèles de détection d'anomalies dans les activités humaines proposés dans les travaux précédents sont complexes et nécessitent un temps d'apprentissage long ainsi qu'une puissance de calcul élevée. De plus, divers paramètres doivent être réglés. Ils peuvent toujours être déployés dans des applications réelles, mais cela prendra plus de temps qu'avec des algorithmes d'apprentissage automatique plus simples.

Dans notre travail, nous proposons un modèle DFFNN basé sur un algorithme hybride d'optimisation PGTO pour détecter les anomalies dans les activités des personnes âgées. Le réseau neuronal DFFNN est exploité pour détecter les anomalies, l'entraînement (training) du modèle étant effectuée à l'aide d'un algorithme hybride d'optimisation PGTO.

Le processus d'optimisation permet d'obtenir les poids et les biais optimaux pour avoir un meilleur accuracy.

3.4 Conclusion

La recherche dans le domaine de la reconnaissance de l'activité humaine (HAR) à l'aide de capteurs portables et de techniques d'apprentissage automatique ou d'apprentissage profond pour les personnes âgées a fait des progrès considérables. Son intégration pourrait révolutionner les soins aux personnes âgées, en permettant une vie indépendante tout en garantissant leur sécurité et leur bien-être.

Au cours de ce chapitre nous avons donné une revue de la littérature qui représente un aperçu des principales études et tendances récentes en matière de HAR pour les personnes âgées à l'aide de capteurs portables en utilisant des approches d'apprentissage automatique (ML) et d'apprentissage profond (DP).

Dans le chapitre suivant, nous allons présenter notre méthodologie de reconnaissance des activités humaines chez les personnes âgées en utilisant des capteurs portables basée sur l'apprentissage profond.

Chapitre 4 Contribution 1 : une approche DL
pour la reconnaissance des activités des
personnes âgées à l'aide de capteurs portables

4.1 Introduction

Le domaine de recherche sur les environnements de vie assistés – AAL (Ambient Assisted Living) se concentre sur la génération de technologies, de produits et de services innovants pour aider, soigner et réadapter les personnes âgées, afin d'augmenter la durée de vie de ces personnes. Indépendamment, qu'ils souffrent de maladies neurodégénératives ou d'un handicap. Ce domaine important est à l'origine du développement des systèmes de reconnaissance d'activités humaines – HARS (Human Activity Recognition Systems), qui constituent un outil précieux lorsqu'il s'agit d'identifier le type d'activité réalisée par les personnes âgées, afin de leur apporter de l'aide qui vous permet de réaliser leurs activités quotidiennes en toute normalité.

La reconnaissance des activités humaines (HAR) est un processus d'identification des activités humaines à partir de données brutes collectées au cours d'une période donnée. Les données sont séquentielles temporelles (time-series) et peuvent être présentées sous la forme de mesures discrètes effectuées par des appareils numériques portables, tels que des accéléromètres, des gyroscopes, des magnétomètres, etc. Les capteurs peuvent être environnementaux, géolocalisés ou portables.

Selon une étude récente [82], plus de 50 % des activités de reconnaissance des activités humaines se concentrent sur les données de capteurs plutôt que sur les données visuelles, RFID et WiFi. Le domaine de recherche moderne de HAR se concentre principalement sur la reconnaissance basée sur les capteurs pour les raisons suivantes : i) le développement de la technologie des capteurs ; ii) le coût et la consommation d'énergie réduits ; iii) le respect de la vie privée amélioré ; iv) les appareils portables nécessitant moins d'infrastructure ; et les données provenant des capteurs plus fiables et plus précises.

Le HAR basé sur des capteurs portables est devenu populaire dans divers domaines, notamment les maisons intelligentes et les soins de santé pour les personnes âgées.

L'apprentissage profond (D.L.) a obtenu des résultats remarquables ces dernières années dans le domaine HAR. L'extraction automatique des caractéristiques et la classification avec une grande précision sont l'un des aspects importants de l'apprentissage profond, ce qui a suscité un intérêt dans le domaine HAR [83].

De nombreux chercheurs ont proposé divers systèmes de surveillance des activités humaines utilisant des dispositifs de détection portables. Cependant, l'accuracy de la classification des activités obtenue jusqu'à présent doit être améliorée pour relever le défi d'une surveillance plus précise des activités.

Les données HAR sont des données de séries chronologiques comprenant des informations spatiales et temporelles qui nécessitent un modèle robuste avec le potentiel d'extraire les deux informations à la fois.

Dans ce chapitre, nous allons étudier un ensemble de modèles d'apprentissage profond : des modèles de base (CNN, LSTM) et des modèles hybrides (CNN-LSTM et CNN Multi_têtes-LSTM) pour reconnaître et prédire avec précision les activités physiques. Ces modèles sont testés et entraînés sur le jeu de données de capteurs corporels portables MHealth pour classement des mouvements corporels sur la base des enregistrements des signes vitaux.

4.2 Méthodologie

Dans cette section, nous proposons un framework de système HAR pour surveillance des activités humaines basé sur l'apprentissage profond piloté par des données des capteurs portables. Dans un premier temps, les données d'activité sont collectées à l'aide de capteurs portables. Dans la deuxième étape, les données brutes sont prétraitées avec des opérations de normalisation et de segmentation des données, etc. Dans la troisième étape, les données traitées sont introduites dans le modèle d'apprentissage profond proposé pour l'extraction de caractéristiques et classées à l'aide d'une couche de classification avec activation SoftMax. Enfin, les performances du modèle sont évaluées par des métriques d'évaluation standards telles que l'accuracy, la precision, le recall et le F1-score. L'architecture globale du Framework est illustrée par la Figure 4-1.

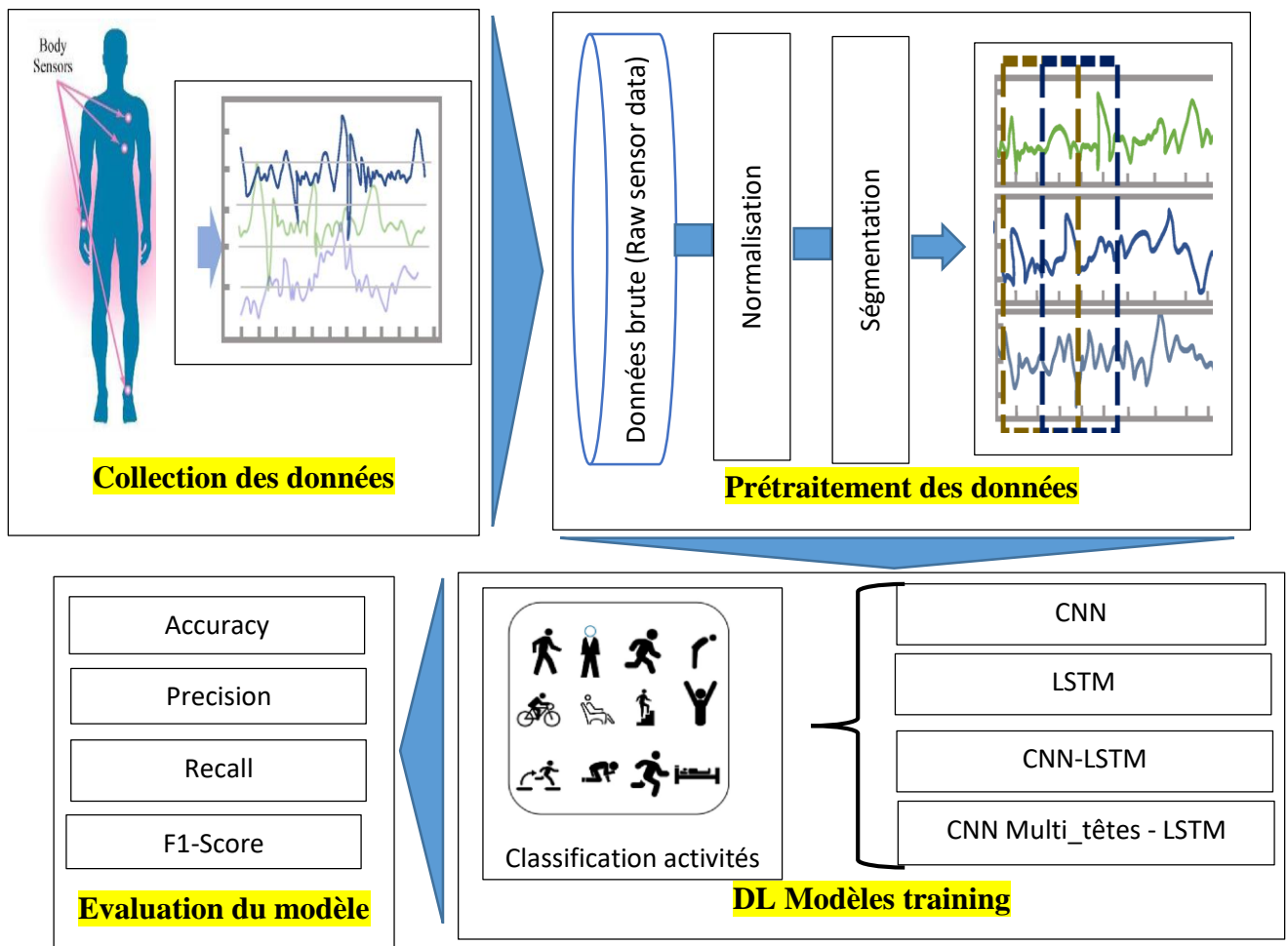


Figure 4-1: A framework de système de reconnaissance des activités humaines basé sur l'apprentissage profond et utilisant des capteurs portables.

4.2.1 Description du jeu de données

Le jeu de données (dataset) utilisé dans notre étude est une base de données de santé mobile MHEALTH (Mobile Health) [84]. Ce dataset comprenant des enregistrements de mouvements corporels et de signes vitaux pour dix volontaires de profils divers lors de l'exécution de 12 activités physiques, en utilisant des capteurs portables Shimmer2 [85]. Les chercheurs ont placé des capteurs sur la poitrine, le poignet droit et la cheville gauche du sujet et fixés avec des sangles élastiques comme le montre la Figure 4-2. L'objectif est de mesurer les mouvements du corps, notamment pour calculer l'accélération, le taux de rotation et l'orientation du champ magnétique dans les différentes parties du corps. En outre, les capteurs placés sur la poitrine fournissent des mesures ECG qui peuvent être

utilisées pour la surveillance du cœur et ne sont pas utilisées pour le développement du modèle de reconnaissance mais plutôt collectées à des fins de travaux futurs.

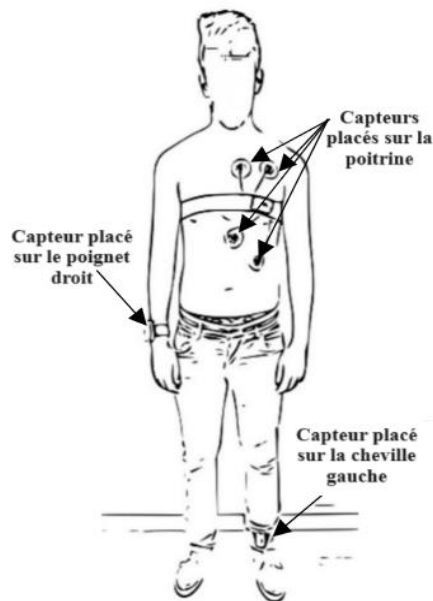


Figure 4-2 : Emplacement des capteurs sur le sujet

Toutes les modalités de détection sont enregistrées à une fréquence d'échantillonnage de 50 Hz, considérée comme suffisante pour capturer l'activité humaine. Ce jeu de données prend en considération la variété des activités courantes de la vie quotidienne, compte tenu de la diversité des parties du corps impliquées dans chacune d'elles (par exemple, élévation frontale des bras contre flexion des genoux), de l'intensité des actions (par exemple, faire du vélo contre s'asseoir et se détendre) et de leur vitesse d'exécution (par exemple, courir contre rester immobile).

Les activités ont été collectées dans un environnement hors laboratoire, sans aucune contrainte sur la manière dont elles doivent être exécutées, à l'exception du fait que le sujet doit faire de son mieux pour les exécuter.

La liste d'activités et leurs instances sont indiquées dans le Tableau 4-1 : *Liste des activités*

Tableau 4-1 : Liste des activités

Code activité	Nom d'activité	Nombre d'instances
L1	Rester immobile (1min)	3072
L2	S'asseoir et se détendre (1 min)	3072
L3	S'allonger (1min)	3072
L4	Marcher (1min)	3072
L5	Monter les escaliers (1 min)	3072
L6	La taille se penche en avant (20x)	3072
L7	Élévation frontale des bras (20x)	3072
L8	Flexion des genoux (accroupissement) (20x)	3379
L9	Cycling (1 min)	3072
L10	Jogging (1 min)	3072
L11	Course (1 min)	3072
L12	Saut avant et arrière (20x)	1075
Total		35174

Remarque : Entre parenthèses indique le nombre de répétitions (Nx) ou la durée des exercices (min).

4.2.2 Prétraitement des données

Les données du monde réel sont généralement incomplètes et bruyantes et sont susceptibles de contenir des informations non pertinentes et redondantes ou des erreurs. Le prétraitement des données permet de transformer les données brutes (raw data) en un format compréhensible pour améliorer la qualité des données, et contribuant ainsi à améliorer la précision et l'efficacité du processus d'exploration. Les méthodes de prétraitement des données qui seront utilisées ici sont :

A) Normalisation (mise à l'échelle du centre) avec Z-score

La distribution des valeurs est remise à l'échelle à l'aide de la méthode de Z-score ou mise à l'échelle centrale. Dans cette méthode, la moyenne des données observées sera "0" et l'écart-type sera "1". Comme les données des capteurs ont des valeurs dans différentes échelles, un échantillon est normalisé selon l'équation (4-1) :

$$y = \frac{(x - Mean)}{var} \tag{4-1}$$

Où la moyenne est estimée comme suit :

$$Mean = \frac{\sum_{i=1}^n X_i}{n} \tag{4-2}$$

Et l'écart-type est estimé comme suit :

$$Var = \sqrt{\frac{\sum_{i=1}^n (X_i - Mean)^2}{n}} \tag{4-3}$$

B) *Segmentation* : Les données des capteurs sont d'abord divisées en fenêtres de taille fixe w à l'aide de l'approche de la fenêtre coulissante produisant N fenêtres afin de simplifier le processus de classification. On utilise une fenêtre glissante à chevauchement de longueur fixe de 2,56 s et un chevauchement de 50 % (128 lectures par fenêtre) (voir Figure 4-3).

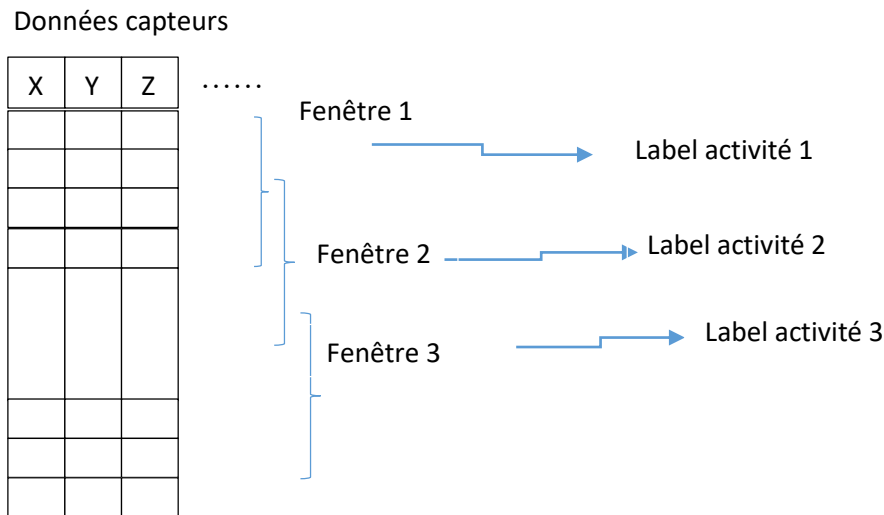


Figure 4-3 : Segmentation des données de capteurs

4.2.3 Modèles d'apprentissage profond proposés

a) Modèles de base

1. Modèle CNN

Les réseaux CNN sont intensivement utilisés dans de nombreux domaines de recherche, y compris la reconnaissance d'activité, et ont atteint d'excellentes performances [86]. Les CNN consistent en une combinaison de couches de convolution, pooling, de normalisation et entièrement connectées (full-connected) qui extraient progressivement des caractéristiques riches des données d'entrée. Les CNN sont conçus pour les problèmes

de classification d'images dans lesquels le modèle apprend une représentation interne de l'image d'entrée et produit un ensemble de caractéristiques 2D. Le même processus peut être utilisé pour les séquences de données unidimensionnelles comme pour les capteurs accéléromètres et gyroscopes en adoptant des réseaux 1D-CNN. Le modèle apprend à extraire des caractéristiques d'un ensemble d'observations et à établir des corrélations entre les caractéristiques internes et différents types d'activités.

Dans ce travail, nous adoptons une architecture qui contient de quatre couches convolutives (Conv1D), et une couche One MaxPooling1D et une couche de GlobalAveragePooling pour prendre les moyennes des cartes de caractéristiques (features map) complètes resultants des couches convolutives et qui rend les réseaux plus robustes aux translations spatiales et leurs décisions plus faciles à interpréter et réduit chaque carte de caractéristiques à une seule valeur, créant ainsi une représentation aplatie (Flatten). Cette dernière couche est suivie par une couche Dropout qui permet de réduire l'overfitting lors de l'entraînement via la désactivation temporaire de certaines neurones. Et enfin une couche de classification SoftMax. La fonction Softmax normalise les scores de sortie en probabilités, représentant la probabilité que l'entrée appartienne à chaque classe. La classe ayant la probabilité la plus élevée est considérée comme la classe prédite pour l'entrée. Après chaque opération de convolution, la sortie passe généralement par une fonction d'activation non linéaire pour former la sortie de la couche de convolution. La fonction d'activation 'ReLU' est utilisée dans les couches convolutives. Cette fonction permet de convertir les valeurs négatives des données extraites par CNN en 0, et les valeurs positives des données restent inchangées. L'architecture CNN proposée est représentée dans la Figure 4-4 (a).

2. Le modèle LSTM

LSTM est un cas particulier des réseaux de neurones récurrents (RNN). Ils ont la capacité d'extraire des dépendances dans des séries de données temporelles et d'apprendre des informations progressivement au fil des intervalles de temps. Il convient donc à l'analyse de données sensorielles dans le cadre de la reconnaissance de l'activité humaine. L'architecture interne du LSTM contient plusieurs portes, chaque porte traitant l'entrée de la porte précédente et la transmettant à la porte suivante, contrôlant ainsi le flux d'informations vers la sortie finale.

Les LSTM ont été appliquées aux séquences vidéo, à la reconnaissance vocale, et à l'analyse de texte et à HAR et des résultats très encourageants ont été obtenus [87].

L'architecture LSTM proposé, se compose de deux couches LSTM, suivis par une couche dense avec une fonction d'activation Softmax, comme indiquée dans la Figure 4-4 (b).

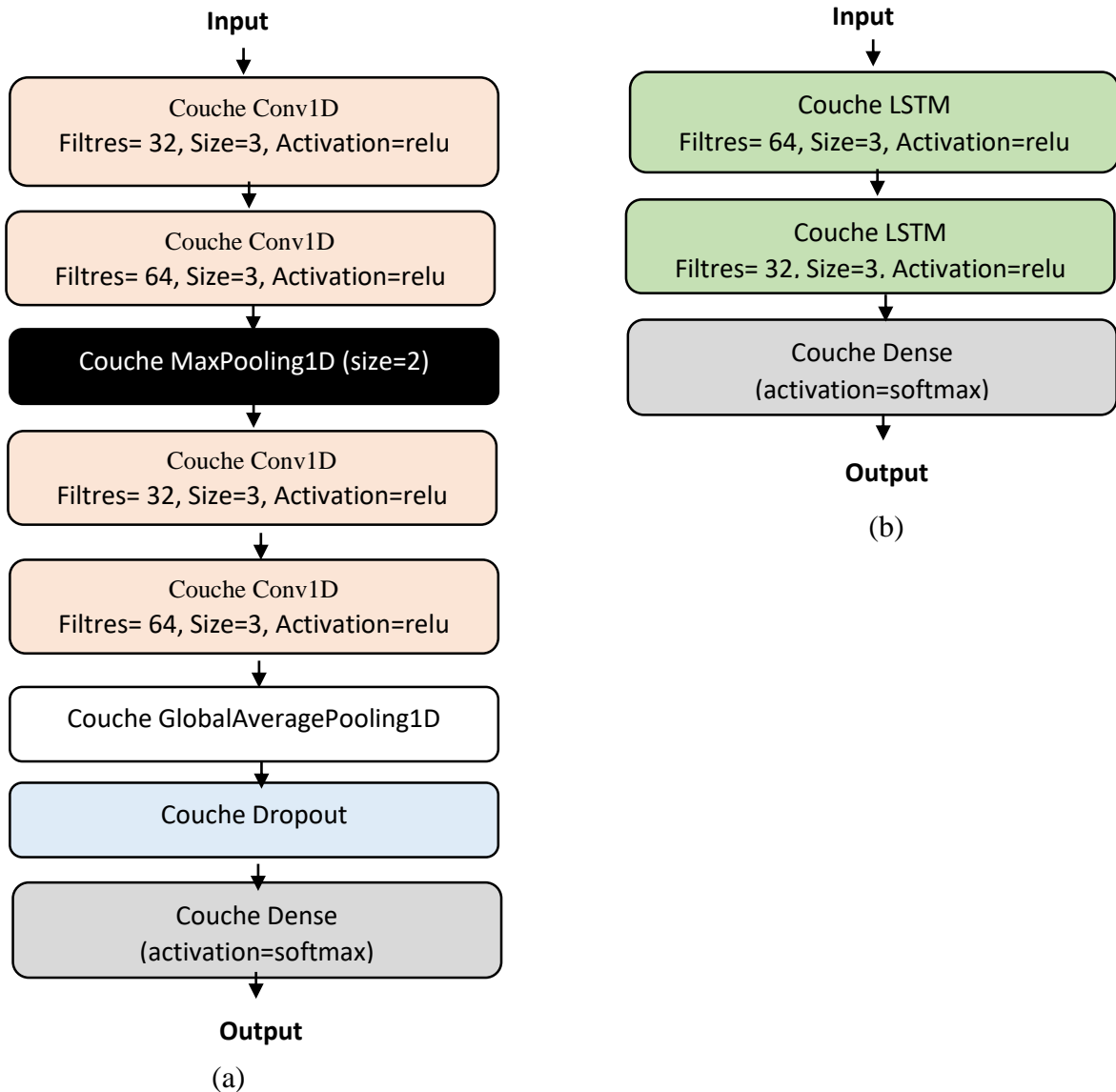


Figure 4-4: Architecture des modèles : (a) CNN, (b) LSTM

b) Modèles Deep learning hybrides

1. Modèle CNN-LSTM

Il s'agit d'un modèle hybride qui vise à combiner les atouts du CNN et du LSTM. Essentiellement, CNN est impliqué dans l'extraction de caractéristiques dans les données d'entrée, couplé avec LSTM pour exploiter les dépendances spatio-temporelles. Ici, la couche max-pooling est utilisée pour introduire les informations dans le LSTM.

L'architecture du modèle hybride CNN-LSTM est constituée de quatre couches convolutives (Conv1D) et deux couches Maxpooling, suivis par deux couches LSTM, puis une couche Dropout, et à la fin, une couche entièrement connectée (Dense) utilise une fonction d'activation Softmax. Comme le montre la Figure 4-5.

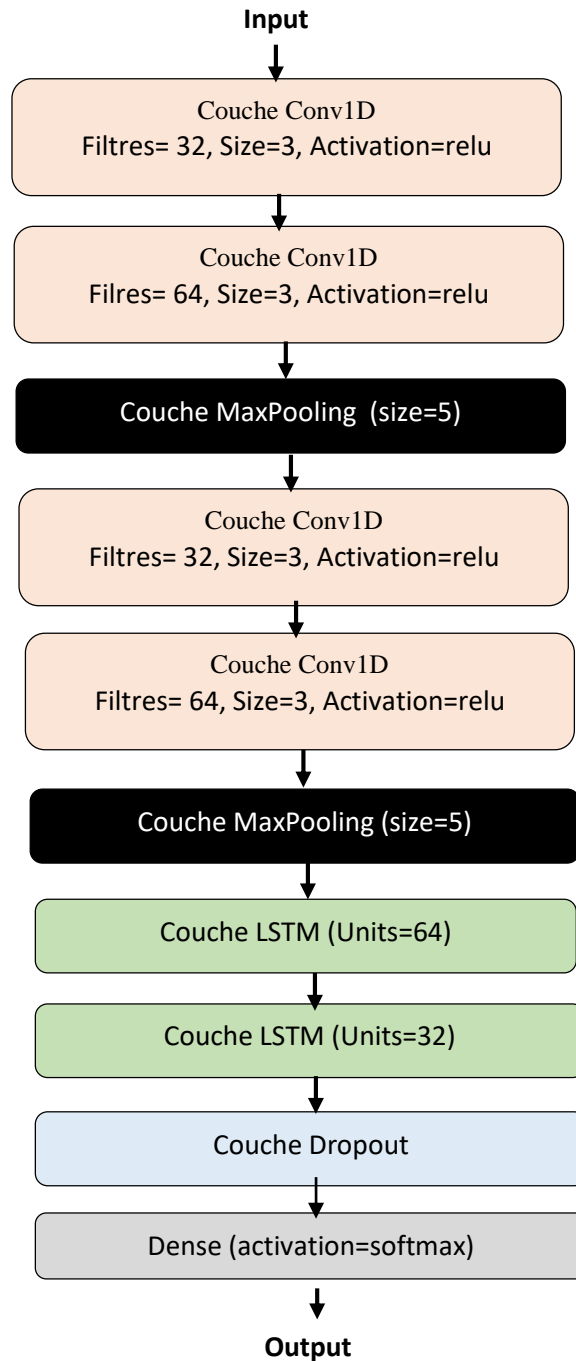


Figure 4-5: Architecture du modèle CNN-LSTM

2. Modèle CNN Multi_têtes -LSTM

Un CNN multi_têtes est utilisé pour extraire des cartes de caractéristiques (features map) des données des capteurs portables. Il se compose de plusieurs têtes convolutives indépendantes, chaque tête étant chargée de capturer différents niveaux d'information à partir des données. Par conséquent, chaque tête convolutive du modèle HAR proposé est conçue pour utiliser différentes tailles de filtre (3, 5, 7) afin d'extraire des cartes de caractéristiques à différents niveaux. Chaque tête contient trois couches convolutives (Conv1D) et deux couches MaxPooling, ensuite ces trois têtes sont concaténées en parallèle, et fournies en tant qu'entrée à la couche LSTM. Cette dernière couche est suivie par une autre couche LSTM, suivie par une couche Dropout, puis une couche Flatten, une couche entièrement connectée Dense, et Dropout. Et à la fin, une couche de entièrement connectée (Dense) utilise une fonction d'activation Softmax pour la classification des activités. Comme le montre-la Figure 4-6.

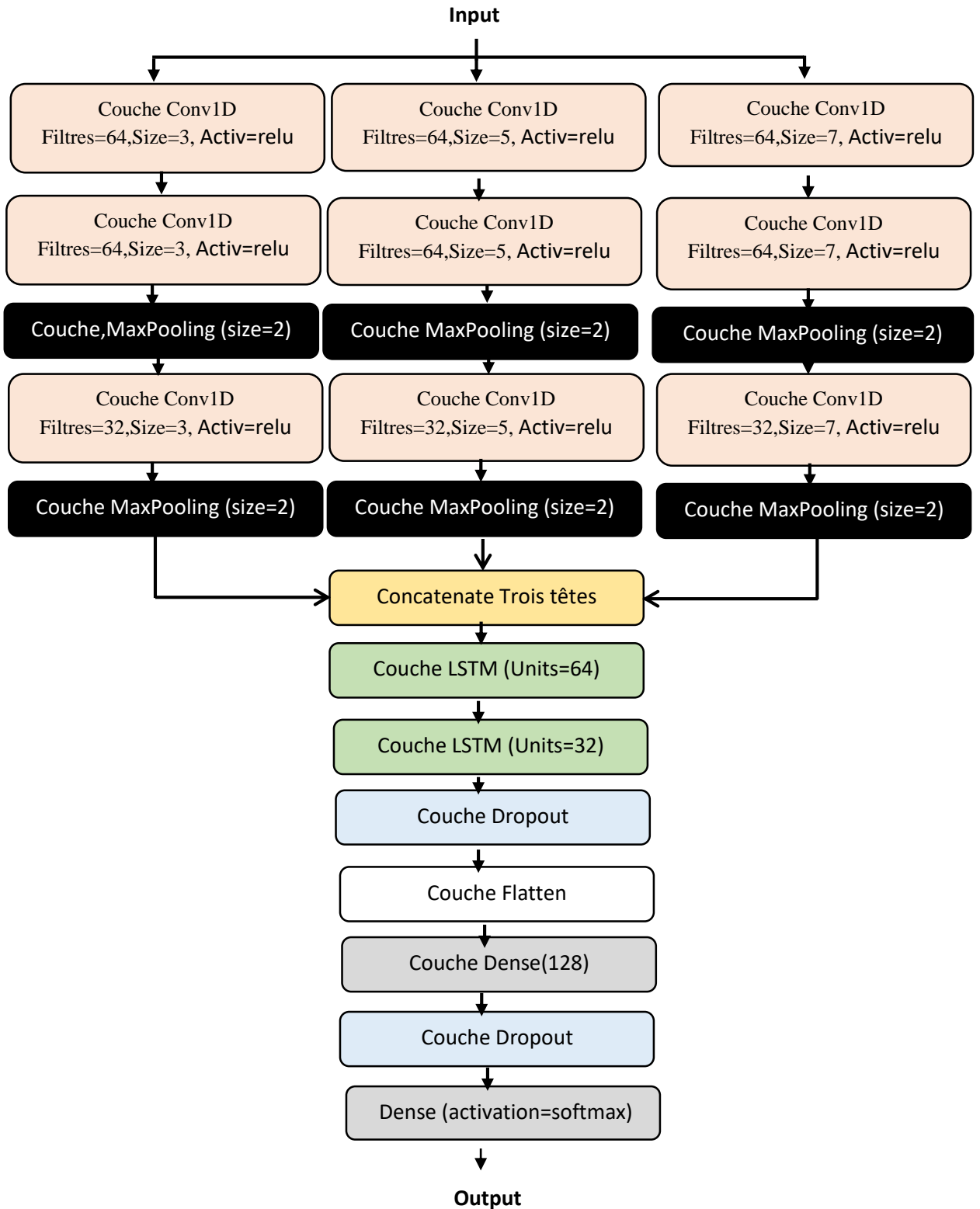


Figure 4-6: Architecture du modèle CNN Multi_têtes-LSTM

4.3 Résultats Expérimentaux et discussion

Dans cette section, on décrit la configuration expérimentale et nous présentons et discutons les résultats expérimentaux de toutes les modèles d'apprentissage profond proposées pour la reconnaissance d'activités humaines à l'aide de métriques d'évaluation courantes.

4.3.1 Détails d'implémentation

L'expérience est réalisée sur un ordinateur doté d'un processeur Intel(R) Core(TM) i7-8850H CPU @ 2.60GHz et de 16 Go de RAM et une carte graphique NVIDIA Quadro P600. Au niveau logiciel, le serveur Google COLAB est utilisé pour compiler et faire des analyses expérimentales. Nous avons utilisé Tensorflow comme bibliothèque Python pour effectuer un apprentissage approfondie sur les réseaux neuronaux. Pour d'autres opérations sur les tableaux, Numpy est utilisé, tandis que matplotlib et seaborn sont utilisés pour la visualisation des données.

Toutes les expériences sont réalisées en utilisant les mêmes hyper paramètres tels que la taille du batch size est 32, le taux d'apprentissage est 0,001, l'optimiseur choisi est Adam, la fonction loss choisie est cross-entropy et nombre d'itérations est 100. Ces paramètres optimaux sont sélectionnés après avoir effectué un nombre d'expériences sur différents paramètres.

Nous avons fusionné toutes les données de chacun des participants de datase Mhealth et l'avons divisée de manière aléatoire en apprentissage (80%) et en test (20%).

4.3.2 Métriques d'évaluation

L'efficacité du modèle sera évaluée à l'aide de mesures d'évaluation ou métriques telles que l'accuracy de la classification, recall, precision, et F1-score.

Ces métriques sont basées sur le nombre de faux positifs (F_{Pos}), de vrais positifs (T_{Pos}), de faux négatifs (F_{Neg}) et de vrais négatifs (T_{Neg}) dans les prédictions des classificateurs.

- L'accuracy est définie comme le pourcentage de prédictions correctes (vraies).
- Le recall est défini comme le rapport entre les observations correctement prédites positives et toutes les observations de la classe réelle.

- La precision est définie comme le rapport entre les observations positives correctement prédites et le total des observations positivement prédites.
- Le F1-score est défini comme la moyenne pondérée du rappel et de la précision.

Les équations de chaque métrique sont définies comme suit :

$$Accuracy (Acc) = \frac{T_{Pos} + T_{Neg}}{T_{Pos} + T_{Neg} + F_{Pos} + F_{Neg}} \quad (4-4)$$

$$Recall(Rec) = \frac{T_{Pos}}{T_{Pos} + F_{Neg}} \quad (4-5)$$

$$Precision(Pre) = \frac{T_{Pos}}{T_{Pos} + F_{Pos}} \quad (4-6)$$

$$F1 - score(F1 - S) = \frac{2 * Pre * Rec}{Pre + Rec} \quad (4-7)$$

4.3.3 Analyse des modèles d'apprentissage profond

L'objectif était d'effectuer une analyse comparative de différents modèles d'apprentissage profond. Plus précisément, nous avons examiné quatre modèles différents: CNN, LSTM, CNN-LSTM et CNN Multi_têtes-LSTM. En évaluant ces modèles et leurs performances à l'aide de mesures d'évaluation.

Pour comparer les performances de ces modèles, nous avons utilisé les diagrammes de l'accuracy et loss d'apprentissage/test. Ces diagrammes fournissent un aperçu de la manière dont les modèles apprennent à partir des données et aident à évaluer leur efficacité pour les tâches de classification des activités. Les Figure 4-7, Figure 4-8, Figure 4-9 et Figure 4-10 décrivent respectivement la courbe de l'accuracy et la courbe de loss à chaque epoch des modèles proposés CNN, LSTM, CNN-LSTM et CNN Multi_têtes-LSTM respectivement. À partir de ces courbes, on remarque les courbes de l'accuracy d'apprentissage et de test des modèles sont parallèles et proche l'une de l'autre, et ne montrant aucun signe de sur-apprentissage (overfitting) et une valeur de loss moyenne inférieure de 0,4.

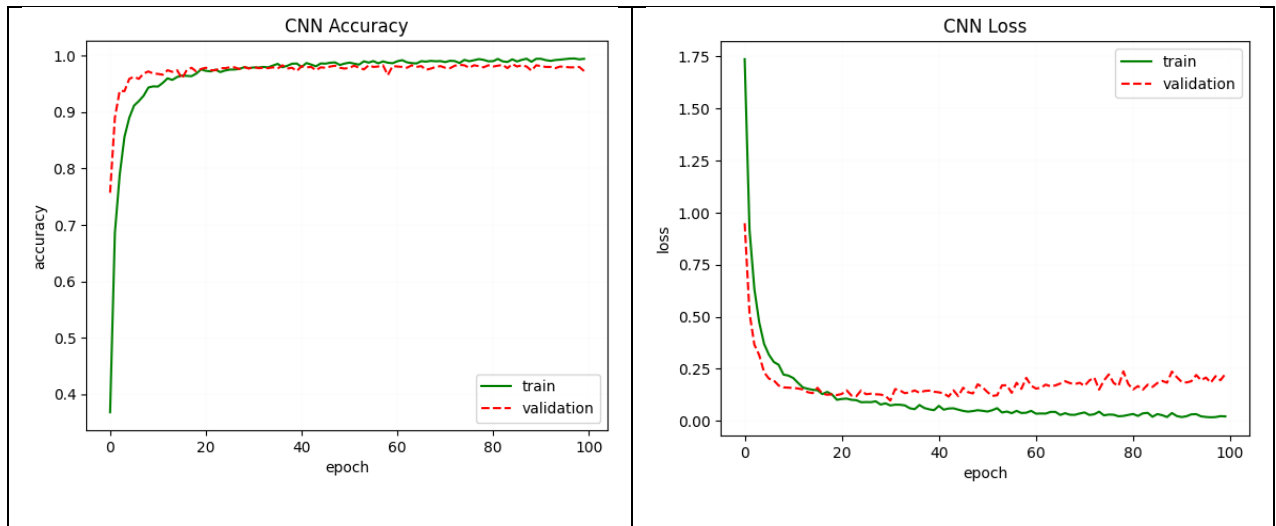


Figure 4-7 : Courbes Accuracy et Loss du modèle : CNN

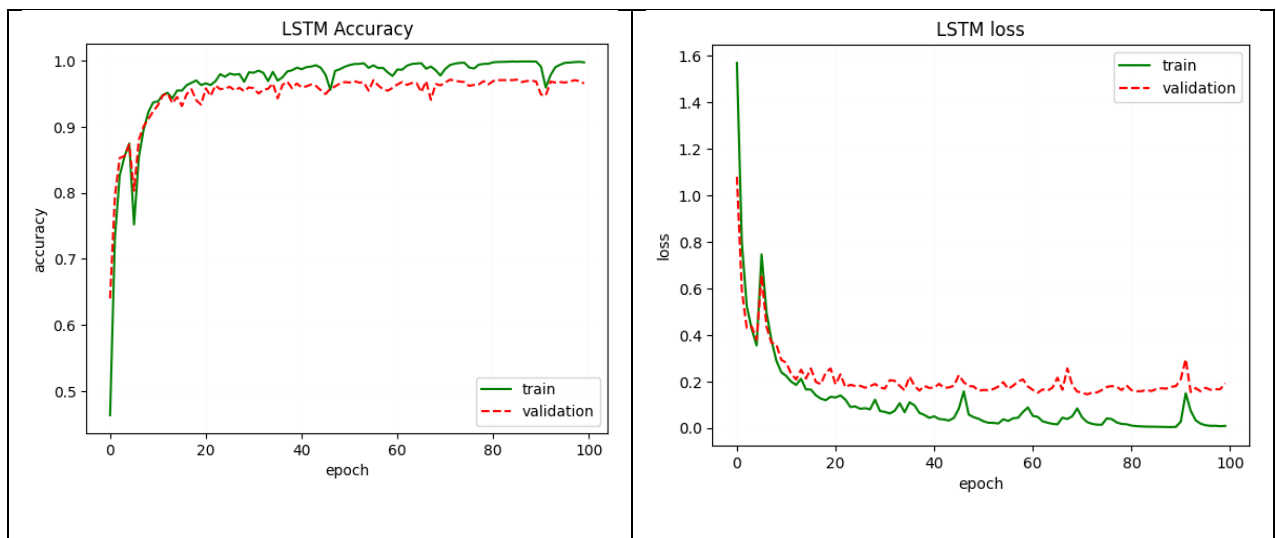


Figure 4-8 : Courbes Accuracy et Loss du modèle :,LSTM

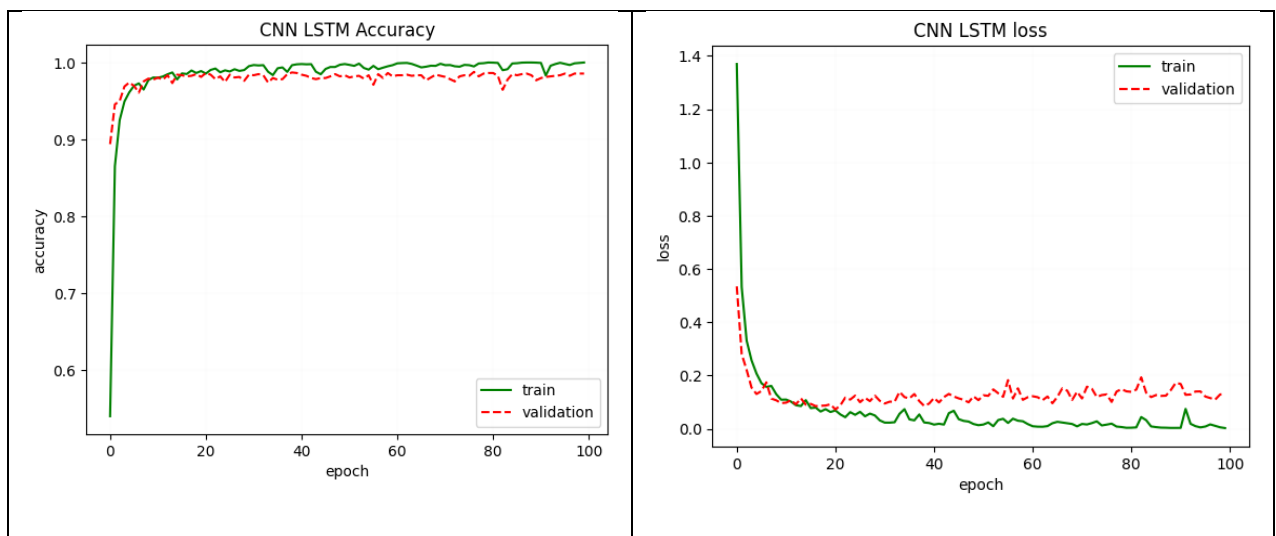


Figure 4-9 : Courbes Accuracy et Loss du modèle : CNN-LSTM

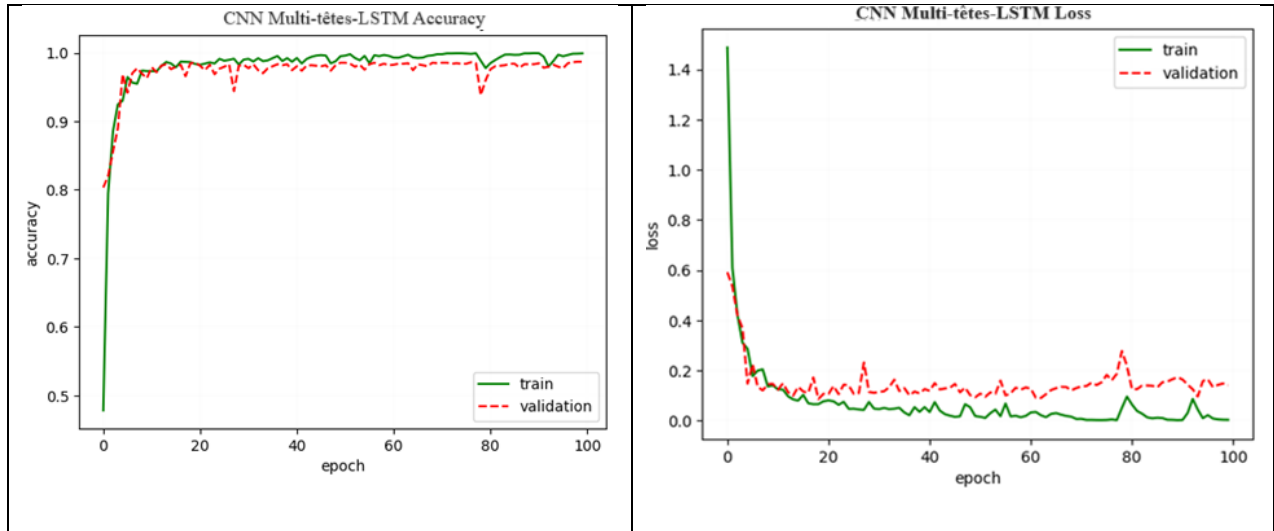


Figure 4-10: Courbes Accuracy/Loss d'apprentissage/test du modèle CNN Multi_têtes-LSTM

Les résultats obtenus sur le dataset Mhealth sont présentés dans le Tableau 4-2. La meilleure accuracy obtenu est de **98,69%** pour le modèle CNN Multi_têtes-LSTM, où les autres mesures de performances telles que le taux de precision, le taux de recall, et F1-score sont de 98,41 %, 98,41 % et 98,50 % respectivement. La deuxième meilleure accuracy est du modèle CNN-LSTM de 98,54 %, avec un taux de precision de 98,25%, un taux de recall de 98,25% et un taux de F1-score de 98,41 %. Les autres modèles ont un taux d'accuracy de 98,25% (CNN) et 96,58 % (LSTM).

Tableau 4-2 : Résultats de classification des modèles proposés

Modèle	Accuracy	Precision	Recall	F1-Score
CNN	98,25	97,91%	97,83%	98%
LSTM	96,58	96,25%	96,25%	96,08%
CNN-LSTM	98,54	98,25%	98,25%	98,41%
CNN Multi_têtes-LSTM	98,69	98,41%	98,41%	98,50%

Figure 4-11, Figure 4-12, Figure 4-13 et Figure 4-14 montrent la matrice de confusion pour chaque modèle qui indique les performances de l'algorithme de classification. Chaque ligne de cette matrice est utilisée pour représenter une instance de classe prédite et les colonnes sont utilisées pour représenter une instance de classe réelle. En regardant l'ensemble des matrices de confusion, nous pouvons dire que CNN Multi_têtes-LSTM a donné les meilleurs résultats par rapport aux valeurs réelles. Malgré,

certaines activités sont confondues avec d'autres activités qui présentent des mouvements similaires et qui ne sont pas pertinentes pour prédire l'activité en question. Par exemple, le jogging est souvent confondu avec Running.

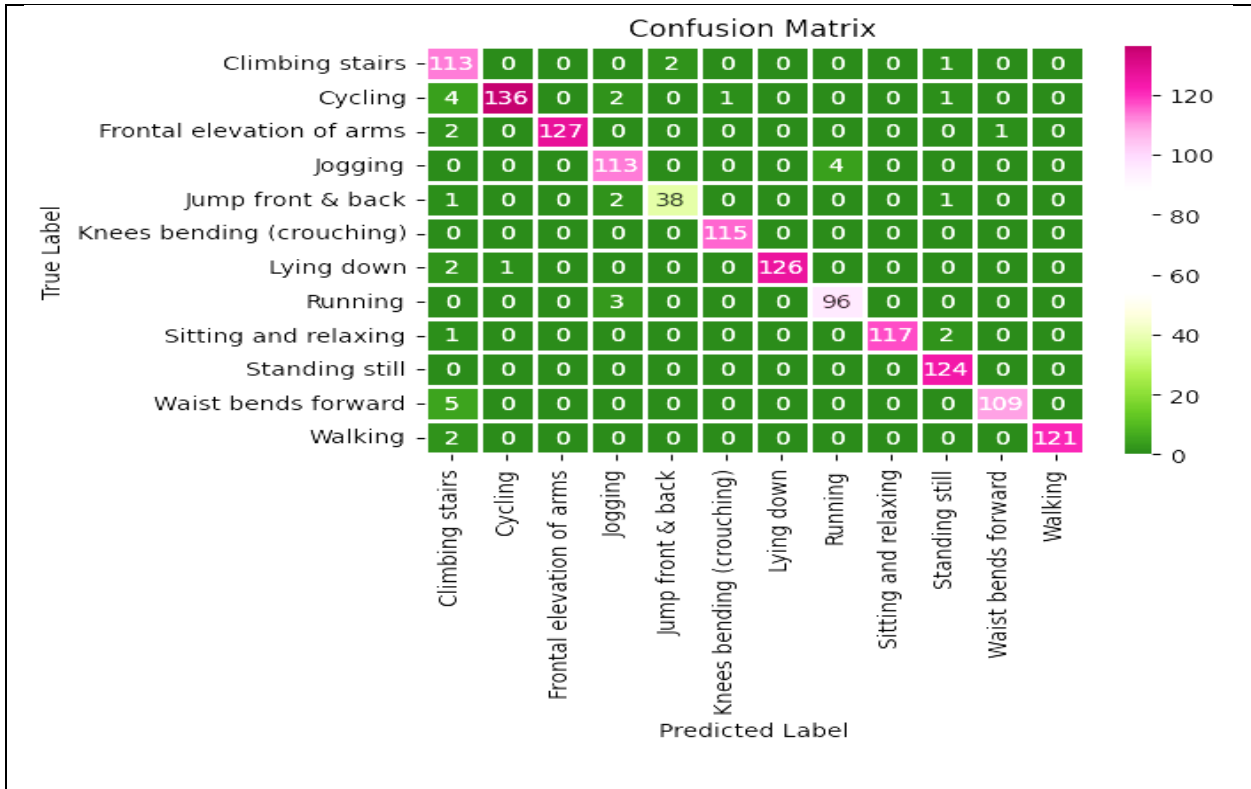


Figure 4-11 : La matrice de confusion de jeu de données Mhealth pour le modèle : CNN

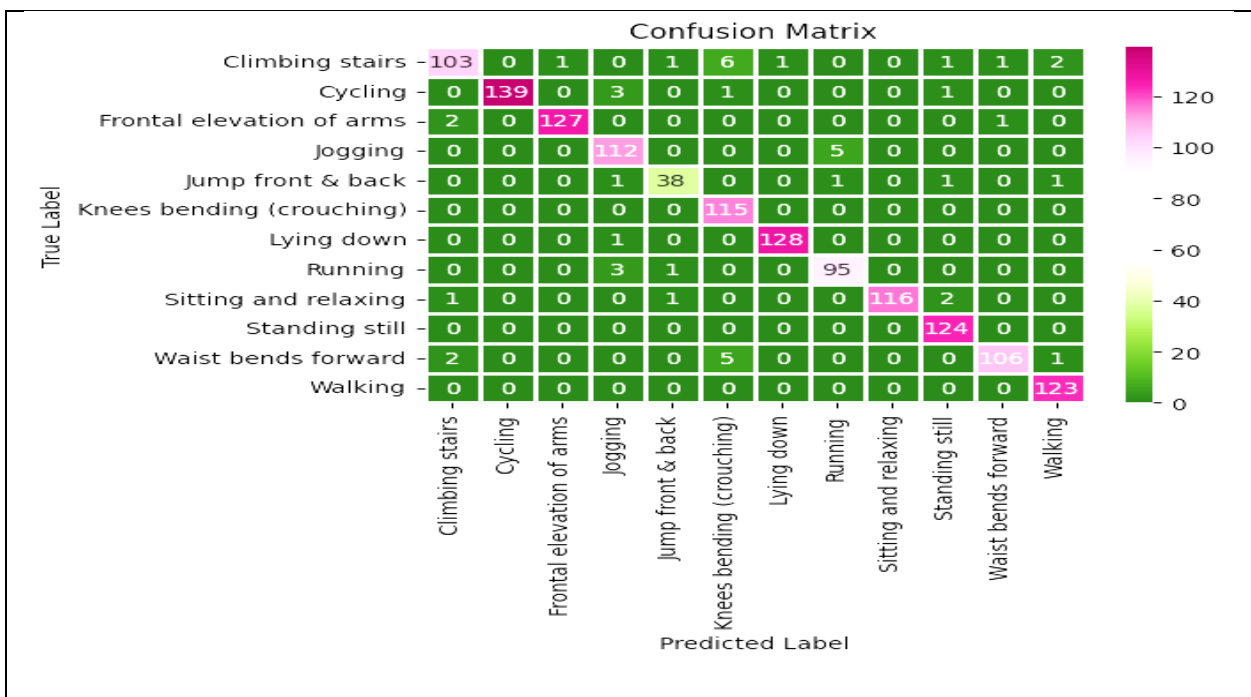


Figure 4-12 : La matrice de confusion de jeu de données Mhealth pour le modèle : LSTM

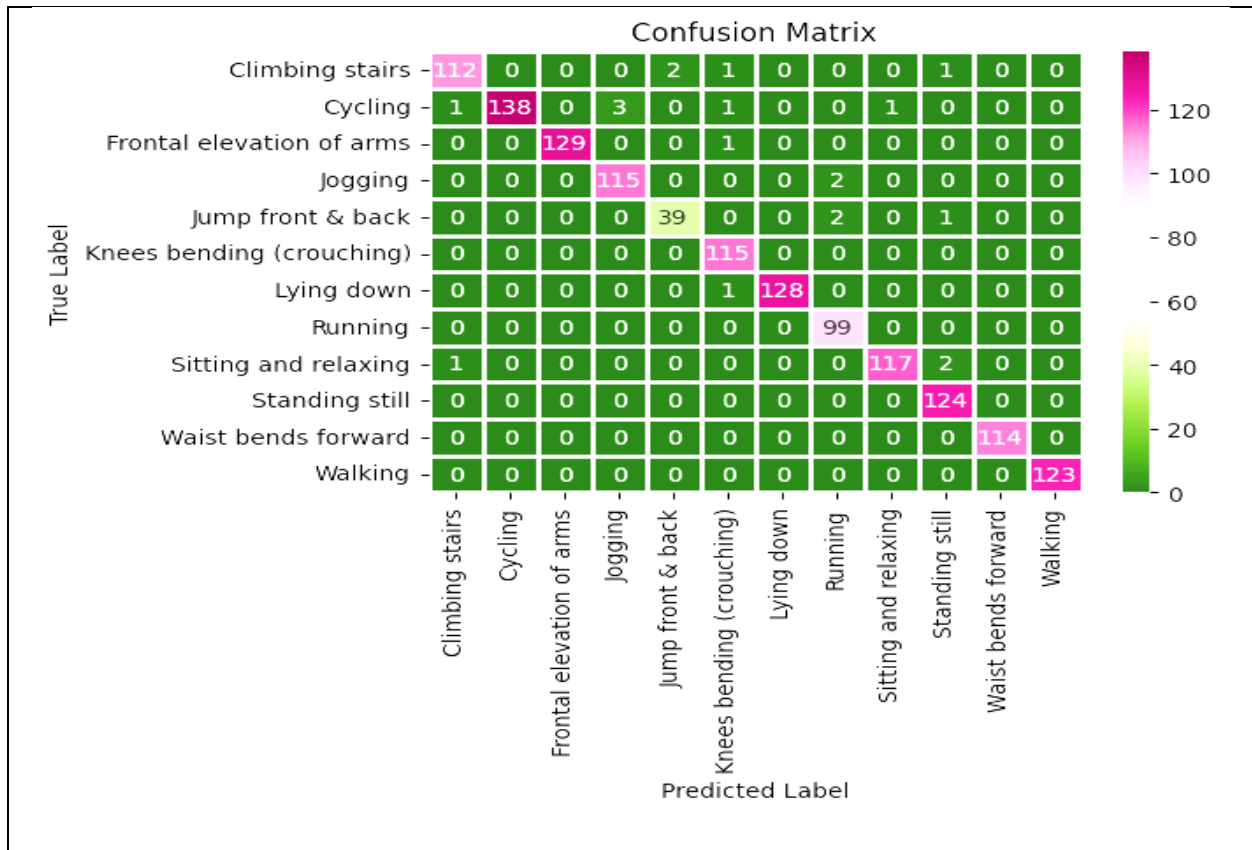


Figure 4-13 : La matrice de confusion de jeu de données Mhealth pour le modèle : CNN-LSTM

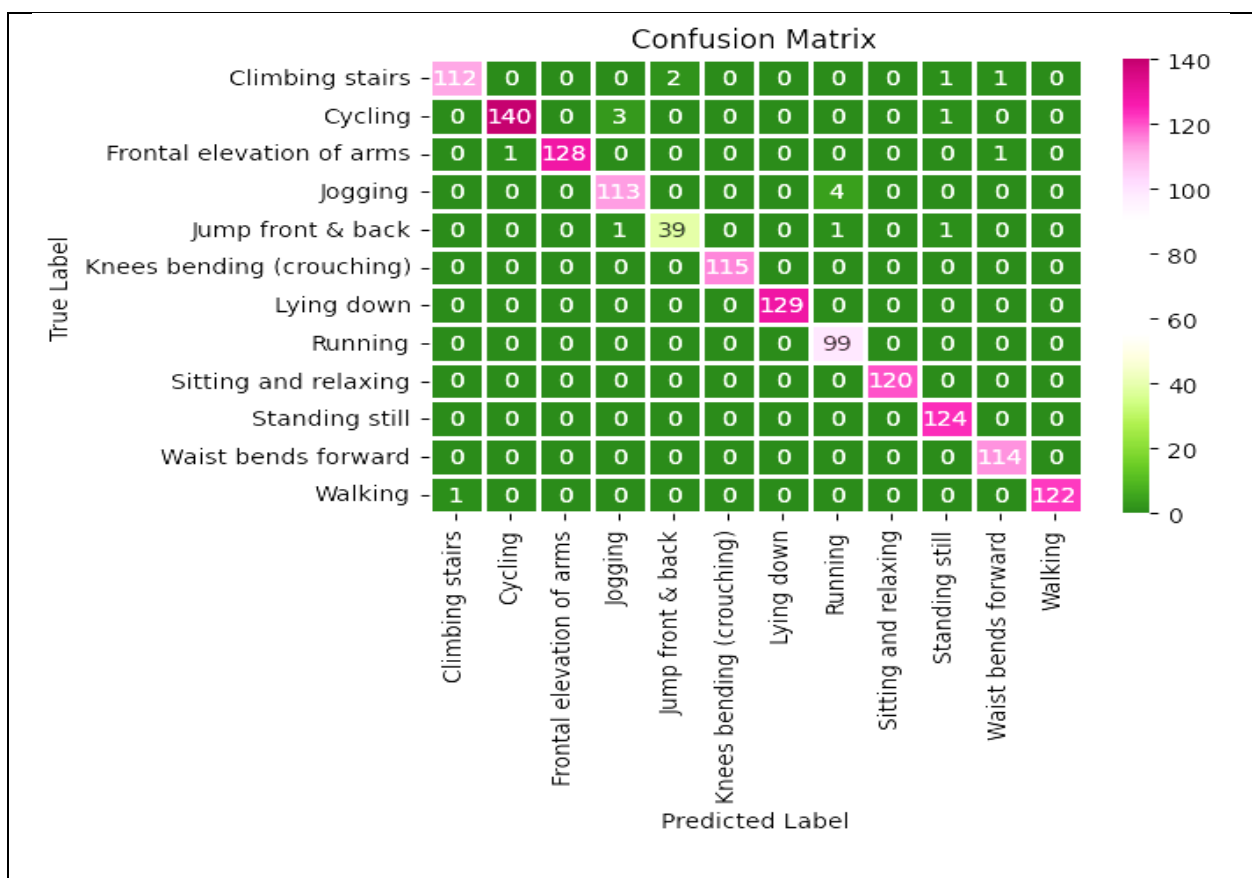


Figure 4-14 : La matrice de confusion de jeu de données Mhealth pour le modèle : CNN Multi_têtes-LSTM

Tableau 4-3 et Tableau 4-4 présentent des résumés du rapport de classification des modèles proposés en termes de précision, de recall, et de F1-score pour chaque activité du dataset mHealth.

Tableau 4-3 : Rapport de classification des modèles CNN et LSTM avec le dataset Mhealth

Modèle \ Activité	CNN			LSTM		
	Precision	Recall	F1-score	Precision	Recall	F1-score
Climbing stairs	0.97	0.97	0.97	0.95	0.89	0.92
Cycling	1.00	0.97	0.98	1.00	0.97	0.98
Front. elevation of arms	1.00	0.99	1.00	0.99	0.98	0.98
Jogging	0.93	0.97	0.95	0.93	0.96	0.95
Jump front & back	0.93	0.93	0.93	0.93	0.90	0.92
Knees bending	0.99	1.00	1.00	0.91	1.00	0.95
Lying down	1.00	1.00	1.00	0.99	0.99	0.99
Running	0.97	0.95	0.96	0.94	0.96	0.95
Sitting and relaxing	0.99	0.98	0.99	1.00	0.97	0.98
Standing still	0.98	1.00	0.99	0.96	1.00	0.98
Waist bends forward	0.99	0.98	0.99	0.98	0.93	0.95
Walking	1.00	1.00	1.00	0.97	1.00	0.98

Tableau 4-4 : Résumé du rapport de classification des modèles CNN-LSTM et CNN Multi_têtes-LSTM avec le jeu de données Mhealth

Modèle Activité	CNN-LSTM			CNN Multi_têtes-LSTM		
	Precision	Recall	F1-score	Precision	Recall	F1-score
Climbing stairs	0.98	0.97	0.97	0.99	0.97	0.98
Cycling	1.00	0.96	0.98	0.99	0.97	0.98
Front. elevation of arms	1.00	0.99	1.00	1.00	0.98	0.99
Jogging	0.97	0.98	0.98	0.97	0.97	0.97
Jump front & back	0.95	0.93	0.94	0.95	0.93	0.94
Knees bending	0.97	1.00	0.98	1.00	1.00	1.00
Lying down	1.00	0.99	1.00	1.00	1.00	1.00
Running	0.96	1.00	0.98	0.95	1.00	0.98
Sitting and relaxing	0.99	0.97	0.98	1.00	1.00	1.00
Standing still	0.97	1.00	0.98	0.98	1.00	0.99
Waist bends forward	1.00	1.00	1.00	0.98	1.00	0.99
Walking	1.00	1.00	1.00	1.00	0.99	1.00

Pour évaluer l'efficacité des modèles d'apprentissage profond proposés, nous les avons comparés à un certain nombre de modèles utilisant le même dataset dans des travaux précédents. Le Tableau 4-5 montre que l'accuracy des modèles proposés est amélioré par rapport aux techniques existantes sur dataset sélectionné.

Tableau 4-5 : Comparaison des performances des modèles proposée et des méthodes précédentes sur le jeu de données Mhealth.

	Modèle	Accuracy (%)
Nguyen et al. [88]	Machine Learning	94.72
Qin et al. [89]	CNN	94.29
	LSTM	93.22
	CNN-LSTM	95.87
Sena et al. [90]	Deep convolutional neural networks (DCNN)	96.27
Proposée	CNN	98.25
	LSTM	96.58
	CNN-LSTM	98.54
	CNN Multi_têtes-LSTM	98.69

4.4 Conclusion

Dans ce chapitre, nous avons étudié des modèles de base : CNN, LSTM, et deux modèles hybrides combinant CNN et LSTM pour reconnaître efficacement l'activité humaine avec un plus grand accuracy. CNN apprend les caractéristiques locales à partir des données originales des capteurs portables, et LSTM extrait les relations temporelles entre les caractéristiques. Nous avons choisi un dataset mHealth qui est collecté auprès de dix participants où chaque participant effectue douze activités physiques : statiques telles que : Lying down, sitting and relaxing et dynamiques telles que : Climbing stairs, Cycling, Front elevation of arms, Joggingetc. La reconnaissance des différentes activités physiques du dataset peuvent aider à améliorer la santé de personne âgée. D'après des expériences sur les modèles d'apprentissage profond proposés. Le modèle CNN Multi_Têtes-LSTM a atteint un accuracy élevée de 98,69 % par rapport à d'autres modèles d'apprentissage profond.

Chapitre 5 Contribution 2 : Algorithme hybride
PGTO d'optimisation du réseau DFFNN et
application à la détection d'anomalies dans
l'activité humaine

5.1 Introduction

Les maisons intelligentes offrent un environnement non invasif et sûr aux personnes âgées et aux personnes handicapées. Plusieurs techniques de soins de santé intelligents ont été conçues pour surveiller à distance les indicateurs de santé et les modèles d'activité des personnes à l'aide de capteurs et de petits appareils portables. Dans l'architecture d'un système de maison intelligente, les appareils de l'internet des objets (IoT) sont gérés par des plateformes de maison intelligente pour automatiser le contrôle en fonction du comportement et des besoins du résident. Grâce aux progrès récents de la technologie de détection, il devient possible de collecter facilement des données telles que la glycémie, la saturation en oxygène, la posture et le mouvement. La variété et la précision des capteurs permettent de reconnaître, de surveiller à distance les activités humaines et d'offrir des services intelligents en fonction des scénarios sans porter atteinte à la vie privée de la personne.

Aujourd'hui, le caractère abordable des appareils intelligents a entraîné une augmentation de l'utilisation des technologies IoT portables [91]. Ces dispositifs intelligents peuvent être intégrés dans des techniques de reconnaissance de l'activité humaine (HAR) afin de fournir des systèmes de santé intelligents. La reconnaissance des activités humaines est une stratégie de l'informatique omniprésente qui vise à faciliter la classification des comportements humains, ce qui permet diverses applications contextuelles dans les maisons intelligentes.

Ces interactions entre les appareils et les humains offrent des canaux de reconnaissance des activités humaines. La caractérisation des activités anormales des patients est une étape essentielle pour des soins de santé intelligents efficaces. La détection d'un comportement anormal dans la vie quotidienne d'une personne se fait généralement en identifiant les non-conformités par rapport aux activités normales de la vie quotidienne [81].

Équiper l'IoT d'un système d'apprentissage profond améliore la reconnaissance des activités et permet d'extraire des caractéristiques complexes à partir du jeu de données sensorielles complexes. L'apprentissage profond est capable d'extraire des caractéristiques temporelles et spatiales significatives et pertinentes à partir des jeux de données [92]. L'application de l'apprentissage profond à ces caractéristiques extraites permet de détecter efficacement les comportements anormaux dans le trafic réseau à haut volume, ce qui n'est pas possible avec les techniques d'analyse de données conventionnelles [93].

Pour entraîner un réseau DFFNN, le poids de chaque neurone est modifié plusieurs fois jusqu'à ce qu'il atteigne une valeur optimale. La technique courante d'apprentissage par backpropagation utilise un algorithme de gradient descent. Les algorithmes basés sur le gradient descent peuvent bloquer dans le minimum local (local minima), alors que le résultat idéal exige de trouver le minimum global [94]. Dans le DFFNN proposé, nous remplaçons l'approche d'entraînement (training) par backpropagation par une stratégie d'optimisation : algorithme PGTO.

Dans ce chapitre nous proposons une approche qui permet de détecter les anomalies dans les activités physiques chez la personne âgée en utilisant un réseau de Deep Feed Forward Neural Network (DFFNN) entraîné par un algorithme hybride des deux algorithmes inspirés de la nature, l'algorithme d'optimisation Pélican (POA) [95] et l'algorithme Gorilla Troops Optimizer (GTO) [96], qui sont utilisés pour établir les poids et biases optimaux du DFFNN.

5.2 Background

5.2.1 Algorithme d'optimisation Pelican (POA)

Le monde naturel a servi de source de motivation pour le développement de plusieurs algorithmes méta-heuristiques. Dans le domaine de la recherche scientifique, l'intelligence en essaim est véritablement un outil important pour résoudre de nombreux problèmes difficiles [97]. Un nouvel algorithme inspiré par la nature est développé par Dehghani & Trojovský en 2022 [95]. POA se concentre sur le comportement social et les tactiques de chasse des pélicans.

POA est un algorithme basé sur la population, et les pélicans sont considérés comme membres de cette population. Dans cet algorithme, chaque membre représente une solution candidate.

Chaque membre a une position dans l'espace de recherche, qui détermine les valeurs qu'il donne aux variables qui composent le problème d'optimisation.

Les pélicans coopèrent fréquemment lorsqu'ils poursuivent des proies. Lorsque les pélicans localisent une cible, ils descendent d'une hauteur de 10 à 20 m. Ils déploient ensuite leurs ailes pour forcer les poissons vers les eaux peu profondes, ce qui leur permet d'attraper plus facilement leurs proies.

Le bec du pélican absorbe une quantité importante d'eau lorsqu'il attrape du poisson, ce qui amène le pélican à pousser la tête vers l'avant pour expulser l'excès de liquide avant d'avaler le poisson.

Cet algorithme se divise en deux étapes principales : la phase 1, qui consiste à se déplacer vers la proie (phase exploratoire) et la phase 2, qui consiste à se déplacer au-dessus de l'eau (phase d'exploitation) [95].

Au début l'algorithme commence par l'initialisation :

- Initialisation

L'initialisation aléatoire de chaque membre de la population marque le début de la procédure d'optimisation, en utilisant l'équation (5-1):

$$x_{i,j} = l_j + rand. (u_j - l_j) \quad i = 1,2 \dots n, \quad j = 1,2, \dots m \quad (5-1)$$

Où $x_{i,j}$ est la valeur de la solution candidate, n représente le nombre de membres de la population, m représente le nombre de variables du problème, $rand$ signifie un nombre aléatoire compris entre $[0, 1]$. l_j signifie la $j^{\text{ème}}$ borne inférieure et u_j spécifie la $j^{\text{ème}}$ borne supérieure.

- Phase 1 : Déplacement vers la source de nourriture (*phase d'exploration*)

Au cours de cette phase, le pélican localise sa proie et descend rapidement sur elle depuis une grande hauteur. La disposition aléatoire des proies dans la zone de recherche améliore la capacité d'exploration du pélican. L'équation (5-2) simule mathématiquement la mise à jour de la position du pélican à chaque itération, ainsi que la représentation de sa position.

$$x_{i,j}^{P1} = \begin{cases} x_{i,j} + rand. (p_j - I. x_{i,j}) & \text{si } F_p < F_i \\ x_{i,j} + rand. (x_{i,j} - p_j) & \text{else} \end{cases} \quad (5-2)$$

Où $x_{i,j}^{P1}$ représente la position mise à jour du $i^{\text{ème}}$ pélican dans la $j^{\text{ème}}$ dimension en fonction de la phase 1, p_j représente la position de la proie dans la $j^{\text{ème}}$ dimension et F représente la valeur de fonction objective (fitness). La valeur du paramètre I est fixée aléatoirement à 1 ou 2.

Si la valeur de la fonction de fitness est meilleure dans une nouvelle position pour un pélican, cette nouvelle position sera considérée comme acceptable. Ce type de mise à jour,

connu sous le nom de mise à jour efficace, empêche l'algorithme de se rendre à des endroits qui ne sont pas optimaux. L'équation utilisée est présentée dans l'équation (5-3).

$$x_i = \begin{cases} x_i^{P1} & \text{si } F_i^{P1} < F_i \\ x_i & \text{sinon} \end{cases} \quad (5-3)$$

Où le nouveau état du $j^{\text{ème}}$ pélican est désigné par x_i^{P1} et F_i^{P1} représente la valeur de la fonction fitness (objective) basée sur la phase d'exploration.

- *Phase 2 : S'envoler à la surface de l'eau (phase d'exploitation)*

Lorsqu'ils atteignent la surface de l'eau, les pélicans étendent leurs ailes pour pousser les poissons vers des eaux moins profondes, ce qui leur permet d'attraper plus facilement leurs proies. Ce comportement peut être modélisé mathématiquement par l'équation (5-4):

$$x_{i,j}^{P2} = x_{i,j} + R \cdot \left(1 - \frac{t}{T}\right) \cdot (2 \cdot \text{rand} - 1) \cdot x_{i,j} \quad (5-4)$$

Où, $x_{i,j}^{P2}$ représente la position mise à jour du $i^{\text{ème}}$ pélican dans la $j^{\text{ème}}$ dimension, en tenant compte de la phase 2, R est une constante de valeur 0.2 et le rayon $x_{i,j}$ est représenté par $R \cdot (1 - t/T)$, où t représente le compteur d'itérations et T le nombre maximum d'itérations. $R \cdot (1 - t/T)$ représente le $x_{i,j}$ rayon du voisinage

À ce stade, une mise à jour efficace a également été utilisée pour accepter ou rejeter la nouvelle position du pélican proposée, qui est décrite dans l'équation (5-5).

$$x_i = \begin{cases} x_i^{P2} & \text{si } F_i^{P2} < F_i \\ x_i & \text{sinon} \end{cases} \quad (5-5)$$

Où la nouvelle position du $j^{\text{ème}}$ pélican est désigné par x_i^{P2} et F_i^{P2} représente la valeur de la fonction de fitness basée sur la phase d'exploitation.

Algorithme POA –Pseudo code

Saisir les informations relatives au problème d'optimisation.

Déterminer la taille de la population de POA (N) et le nombre d'itérations (T).

Initialisation de la position des pélicans à l'aide de l'équation (5-1) et calcul de la fonction Fitness (objective).

Pour t = 1:T

 Générer la position de la proie au hasard.

 Pour I = 1:N

Phase 1 : déplacement vers la proie (phase d'exploration).

Pour $j = 1:m$

Calculer la nouvelle position de la $j^{\text{ème}}$ dimension à l'aide de l'équation (5-2).

Mettre à jour le $i^{\text{ème}}$ membre de la population à l'aide de l'équation (5-3).

Phase 2 : survoler à la surface de l'eau (phase d'exploitation).

Pour $j = 1:m$

Calculer la nouvelle position de la $j^{\text{ème}}$ dimension à l'aide de l'équation (5-4).

Mettre à jour le $i^{\text{ème}}$ membre de la population à l'aide de l'équation (5-5).

Mettre à jour la meilleure solution candidate.

Produire la meilleure solution candidate obtenue par la POA.

5.2.2 Algorithme d'optimisation de troupes de Gorilles (GTO)

GTO est une nouvelle technique d'optimisation proposée par Abdollahzadeh et al. en 2021 [98]; il simule le comportement social et les mouvements des gorilles dans la nature. Les gorilles sont des animaux sociables qui vivent en groupes, appelés troupes. Chaque troupe est dirigée par un gorille au dos argenté (silverback) qui prend des décisions importantes et protège la troupe, et tous les autres membres de la troupe le suivent. Les jeunes gorilles mâles, appelés dos noirs (black backs), occupent la deuxième place dans la hiérarchie du groupe. Les blackbacks suivent également le silverback et protègent le groupe.

La technique GTO est similaire aux autres techniques d'optimisation basées sur les phases d'exploration et d'exploitation. L'exploration dans GTO comprend trois stratégies : La première est basée sur le déplacement du gorille vers des endroits inconnus, tandis que les deuxième et troisième sont basées sur le déplacement d'un gorille vers un autre gorille ou vers un endroit connu. La phase d'exploitation comprend deux stratégies : la première est basée sur les déplacements de silverback tandis que la seconde décrit le déplacement des femelles adultes.

- *Étape d'exploration*

Dans la communauté des gorilles, nous savons qu'il existe un gorille silverback qui gère toutes les décisions. Les gorilles se rendent parfois dans des endroits pour chercher de la nourriture. Le gorille silverback est considéré comme la solution candidate optimal à chaque étape d'exploration d'optimisation.

Les trois mécanismes de la phase d'exploration sont exprimés dans l'équation (5-6) où p est un paramètre compris entre 0 et 1 qui contrôle la stratégie de migration des positions inconnues. Lorsque $\text{rand} < p$, la position actuelle du gorille se déplacera vers une position inconnue. Cela permet à l'algorithme GTO de mieux surveiller l'ensemble de l'espace du problème, ce qui rend la distribution des solutions plus dispersée et plus complète. A l'inverse, si $\text{rand} \geq p$, deux autres mécanismes seront choisis. Ensuite, si $\text{rand} \geq 0,5$, le gorille se déplacera en direction des autres gorilles. Ce mécanisme rapproche la solution actuelle des autres solutions et améliore les performances d'exploration de l'algorithme GTO. Lorsque le $\text{rand} < 0,5$, les gorilles migrent vers la position connue. Cela a amélioré la capacité de l'algorithme GTO à échapper aux solutions optimales locales. L'équation (5-6) est la suivante :

$$Gx(t + 1) = \begin{cases} (UB - LB) \times r_1 + LB & \text{si } \text{rand} < p \\ (r_2 - C) \times X_A(t) + L \times H & \text{si } \text{rand} \geq 0,5 \\ X(t) - L \times (L \times (X(t) - X_B(t)) + r_3 \times (X(t) - X_B(t))) & \text{si } \text{rand} < 0,5 \end{cases} \quad (5-6)$$

Où t indique le temps de l'itération en cours, $X(t)$ désigne la position courante du gorille, et $Gx(t + 1)$ représente le vecteur de position du gorille dans l'itération suivante et sa dimension indique la dimension du problème. En outre, r_1 , r_2 et r_3 sont des valeurs aléatoires comprises entre 0 et 1. $X_A(t)$ et $X_B(t)$ sont deux positions de gorilles choisies au hasard dans la population actuelle. p est une constante. Les bornes supérieures et inférieures des variables sont exprimées par UB et LB . Le paramètre C est calculé selon l'équation (5-7).

$$C = F \times \left(1 - \frac{t}{\text{MaxIt}}\right) \quad (5-7)$$

Où t représente le nombre actuel d'itérations et MaxIt représente le nombre maximal d'itérations. Et F est calculé comme suit :

$$F = \text{Cos}(2 \times r_4) + 1 \quad (5-8)$$

Où r_4 est un nombre aléatoire compris entre 0 et 1 dans une distribution uniforme afin de garantir le caractère aléatoire de la recherche, ce qui est utile pour trouver la solution optimale globale.

L est un paramètre utilisé pour simuler le leadership des gorilles silverback, qui est calculé à l'aide de l'équation (5-9).

$$L = C \times l \quad (5-9)$$

Où, l est un nombre aléatoire entre -1 et 1.

Le paramètre H est calculé par l'équation (5-10).

$$H = Z \times X(t) \quad (5-10)$$

Où, Z est une valeur aléatoire dans les dimensions du problème et dans l'intervalle $[-C, C]$.

A la fin de la phase d'exploration, les valeurs de fitness de G_X et X sont calculées. Si la valeur de fitness de $G_X(t)$ est inférieure à la valeur de fitness de $X(t)$, alors la position de $G_X(t)$ remplacera la position de $X(t)$. En outre, la solution optimale à cette période est sélectionnée en tant que silverback : $X_{\text{silverback}}$.

- *Étape d'exploitation*

Dans la phase d'exploitation, deux comportements sont pris, continuer à suivre le silverback et entrer en compétition pour les femelles matures gorilles. L'un des deux comportements peut être sélectionné en fonction des résultats de la comparaison (équation (5-6)) avec sa valeur initialisée au départ.

Dans la phase d'exploitation de l'algorithme GTO, deux comportements sont appliqués : le suivi du gorille silverback et la compétition pour les femelles adultes. Le gorille silverback dirige un groupe, prend toutes les décisions, détermine les mouvements du groupe et dirige les gorilles vers les sources de nourriture. D'un autre côté, le gorille silverback peut s'affaiblir, vieillir et finalement mourir, et le blackback du groupe peut devenir le leader du groupe, ou d'autres gorilles mâles peuvent engager le gorille silverback et dominer le groupe. Comme décrit avec les deux mécanismes utilisés dans la phase d'exploitation, il est possible de choisir entre suivre le silverback ou la compétition pour les femelles adultes en utilisant la valeur C dans l'équation (5-7). Si $C \geq W$, le mécanisme de suivi du silverback est sélectionné, mais si $C < W$, c'est le mécanisme de compétition des femelles adultes qui est retenu. W est un paramètre à définir avant l'opération d'optimisation.

1. *Suivre le silverback*

Le silverback est considéré comme le leadership d'un groupe de gorilles qui font des choix et conduisent le groupe vers des sources de nourriture. Si, cette approche est choisie. L'équation (5-11) peut être utilisée pour décrire numériquement ce comportement.

$$GX(t + 1) = L \times M \times (X(t) - X_{Silverback}) + X(t) \quad (5-11)$$

Où, X(t) est le vecteur de la position du gorille, tandis que la position du silverback est représentée par le vecteur X_{silverback} (solution optimale). Et M est représenté par l'équation (5-12).

$$M = \left(\left| \frac{1}{N} \sum_{i=1}^N GX_i(t) \right|^g \right)^{1/g} \quad (5-12)$$

Où GX_i(t) représente le vecteur de position de chaque gorille candidat à l'itération courante. N représente le nombre total de gorilles, et g peut être calculé avec l'équation (5-13) comme suit :

$$g = 2^L \quad (5-13)$$

En outre, L est calculé à l'aide de l'équation (5-9).

2. *Compétition pour les femelles adultes*

Si C < W, le second mécanisme est sélectionné pour la phase d'exploitation. Au bout d'un certain temps, lorsque les jeunes gorilles atteignent la maturité, ils se battent avec les autres gorilles mâles de leur groupe pour choisir les femelles adultes, et cette compétition est souvent violente. Ces combats peuvent durer plusieurs jours et impliquer les membres du groupe. L'équation (5-14):

$$GX(t) = X_{silverback} - (X_{silverback} \times Q - X(t) \times Q) \times A \quad (5-14)$$

$$Q = 2 \times r_5 - 1 \quad (5-15)$$

$$A = \beta \times E \quad (5-16)$$

$$E = \begin{cases} N_1 & \text{si } rand \geq 0,5 \\ N_2 & \text{si } rand < 0,5 \end{cases} \quad (5-17)$$

Où, X_{silverback} est le vecteur de position du silverback (meilleure solution) et X(t) est le vecteur de position courante du gorille. Q est utilisé pour simuler la force d'impact, calculée à l'aide de l'équation (5-15). r₅ est un nombre aléatoire compris entre 0 et 1, A est le vecteur de coefficient utilisé pour déterminer le degré de violence dans les conflits, β est une valeur

donnée avant l'opération d'optimisation. E est évalué à l'aide de l'équation (5-17) tout en étant utilisé pour simuler l'effet de la violence sur les dimensions des solutions. Si $\text{rand} \geq 0,5$, la valeur de E sera égale à des valeurs aléatoires dans la distribution normale et les dimensions du problème, mais si $\text{rand} < 0,5$, E sera égale à une valeur aléatoire dans la distribution normale. Lorsque $\text{rand} \geq 0,5$, E sera égal à la valeur aléatoire dans la distribution normale et la dimension du problème ; sinon, E sera égal au nombre aléatoire dans la distribution normale. rand est également une valeur aléatoire comprise entre 0 et 1

A la fin de la phase d'exploitation, une opération de formation de groupe est effectuée, au cours de laquelle le coût de toutes les solutions GX est estimé et si le coût de $GX(t) < X(t)$ la solution GX(t) est utilisée comme solution X(t) et la meilleure solution obtenue parmi l'ensemble de la population est considérée comme un silverback.

Algorithm GTO- Pseudo code

Entrées: La taille de la population N et le nombre maximum d'itérations T et les paramètres β et p

Sorties : La position de Gorille et sa valeur de fitness

%initialisation

Initialiser la population aléatoire X_i ($i= 1, 2, \dots, N$)

Calculer les valeurs de fitness de Gorilles

Tant que (la condition d'arrêt n'est pas satisfaite) do

Mettre à jour le C à l'aide de l'équation (5-7)

Mettre à jour le L à l'aide de l'équation (5-9)

%Phase d'exploration

Pour (chaque Gorille (X_i)) faire

Mettre à jour la position du gorille en utilisant l'équation (5-6)

fin pour

%Création d'un groupe

Calculer les valeurs de fitness des gorilles

Si GX est meilleur que X, les remplacer

Définir $X_{\text{silverback}}$ comme la position de silverback (meilleur position)

%Phase d'exploitation

Pour (chaque Gorille (X_i)) faire

si ($C \geq W$) alors

Mettre à jour la position du gorille en utilisant l'équation (5-11)

```

Sinon
    Mettre à jour la position du gorille à l'aide de l'équation (5-14)
Fin si
Fin pour
% créer un groupe
Calculer les valeurs de fitness du Gorille
Si les nouvelles solutions sont meilleures que les solutions précédentes, les remplacer
Définir  $X_{silverback}$  comme l'emplacement du silverback (meilleure position)
Fin Tq
Return  $X_{BestGorilla}, BestFitness$ 

```

5.3 Méthodologie

5.3.1 Framework proposé

L'objectif est de modéliser le DFFNN basé sur l'algorithme d'optimisation hybride PGTO pour la détection d'anomalies dans les activités de la personne âgée. Le framework proposée se compose de différentes phases. Dans un premier temps, les données d'entrée sont récupérées à partir du dataset, puis passées au module de prétraitement, qui est une étape fondamentale qui permet de convertir des données brutes en informations utiles. Ensuite, le modèle DFFNN est utilisé pour détecter les anomalies de telle sorte que l'entraînement (training) du classificateur est effectué à l'aide de l'algorithme PGTO. La

Figure 5-1 illustre le schéma fonctionnel du Framework proposé.

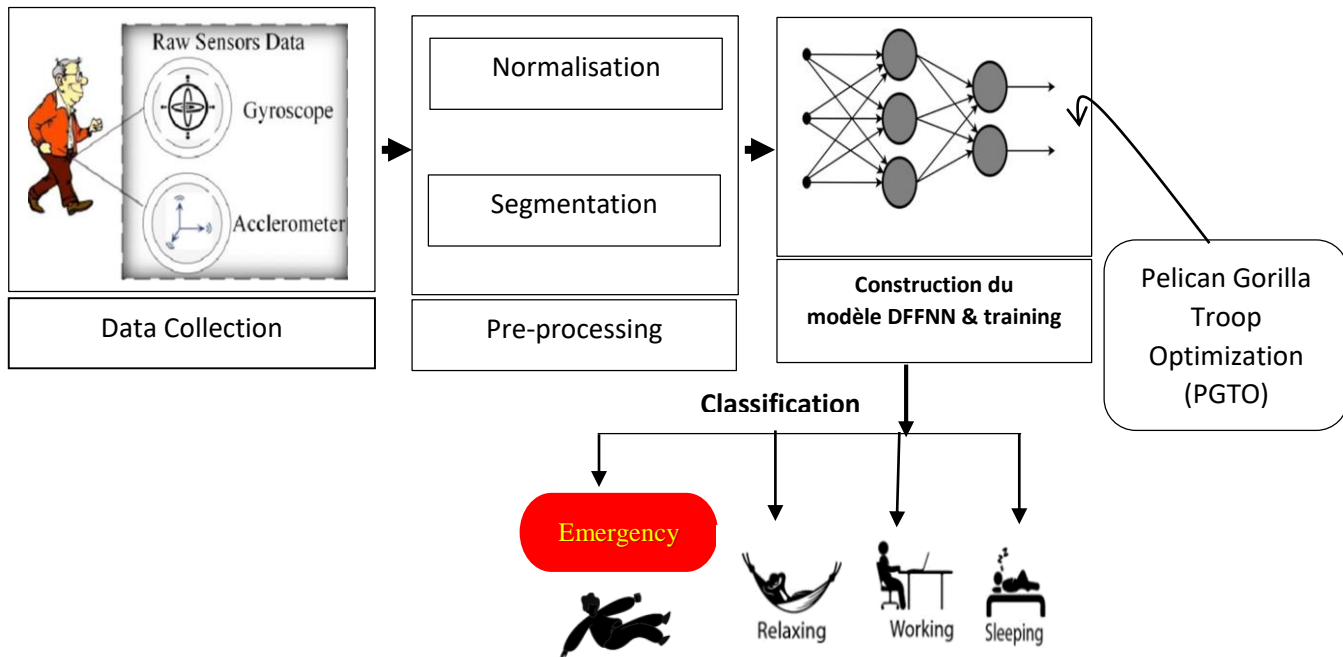


Figure 5-1 : schéma fonctionnel du Framework proposé

a) Description de jeu de données

La détection des anomalies dans les activités des personnes âgées se fait à l'aide d'un dataset "Action recognition dataset" de personne âgée [99]. Le jeu de données utilisé contient les données acquises à partir des systèmes d'activité physique et de localisation intérieure. Les données du système de suivi de l'activité physique sont l'accéléromètre (axe X, axe Y, axe Z) et le gyroscope (axe X, axe Y, axe Z) et le nom de l'activité réalisée par le volontaire obtenu à partir du capteur MPU6050 à une fréquence d'échantillonnage de 20 Hz. Ce capteur a été placé à l'intérieur du dispositif portable porté par le sujet.

Le dataset sur les emplacements intérieurs est enregistré dans le Digital Smart Home Lab (DSHL) situé à l'Université multimédia. DSHL se compose de 5 pièces, à savoir le salon (10 m × 5 m), la cuisine (4,5 m × 4,5 m), la chambre (5 m × 4,5 m), le bureau (6m × 4 m), le hall (4 m × 3,8 m) et les toilettes (1,6 m × 0,95m).

Les données d'actions obtenues en intégrant les résultats de la localisation intérieure et des systèmes de suivi des activités pour étudier les actions. Les actions désignent le résultat de la réalisation d'une activité à un endroit donné, car la même activité à un endroit différent peut illustrer quelque chose de différent. Par exemple, s'allonger dans la chambre signifie dormir (Sleeping), mais s'allonger dans le salon signifie se détendre tandis qu'être allongé dans les toilettes signifie une urgence. Les actions étudiées sont : Relaxing,

Sleeping, Working, Emergency (chute). La taille du jeu de données est de 600 000 points de données (samples).

b) Prétraitement des données (Pre-processing)

Une fois l'acquisition des données terminée, le prétraitement des données est effectué en utilisant la normalisation du Z-score [100] pour normaliser les données d'entrée. Cette technique met à l'échelle les valeurs d'une caractéristique pour avoir une moyenne de 0 et un écart type de 1.

Pour extraire les aspects temporels des activités, la segmentation nécessite de séparer les informations du capteur en segments ou fenêtres plus compacts. La taille et le chevauchement des fenêtres sont déterminés par la durée et l'intensité de l'activité observée. On utilise une fenêtre glissante de longueur fixe de 4 s et un chevauchement de 50 % (80 lectures par fenêtre).

c) Sélection du modèle

Le modèle DFFNN (Deep Feed Forward Neural Network) est un modèle mathématique conçu pour imiter la manière dont le cerveau humain traite les informations. Les DFFNNs sont une classe de réseaux neuronaux artificiels dans lesquels les connexions entre les différentes couches de neurones suivent une direction unique des nœuds d'entrée aux nœuds de sortie. Ces modèles [101] sont généralement composés de couches principales: une couche d'entrée, plusieurs couches cachées et une couche de sortie. Le nombre de couches cachées définit la profondeur de l'architecture. Avec une grande quantité de couches cachées et quantité de neurones dans chaque couche cachée, le DFNN peut fournir une représentation multiniveau de jeu de données. L'architecture d'un modèle DFNN général est montrée dans la Figure 5-2.

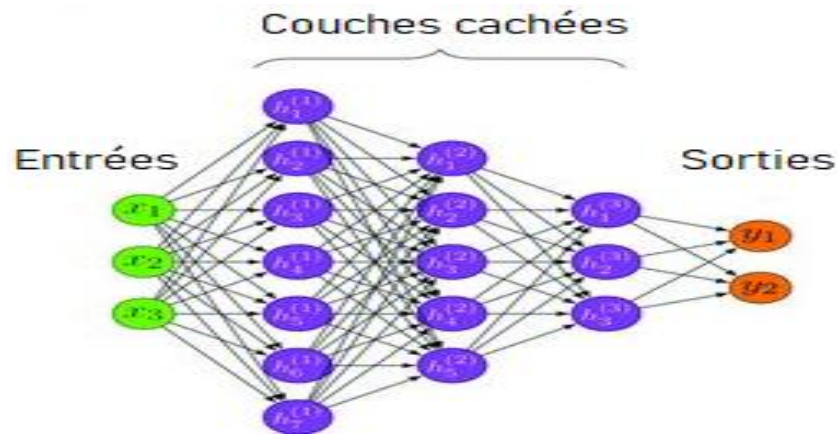


Figure 5-2 : DFFNN avec 3 couches cachées

Le processus d'apprentissage d'un tel réseau est généralement basé sur l'algorithme de rétropropagation (backpropagation) [102], qui met à jour les poids et biais correspondant aux différentes synapses entre les neurones, et par conséquent leur activation éventuelle, afin de minimiser l'erreur entre la sortie fournie par le réseau et l'étiquette correcte.

Pour notre travail, nous considérons un DFFNN avec une couche d'entrée, deux couches cachées avec une fonction d'activation 'Relu' et une couche de sortie avec une fonction d'activation 'softmax'.

La fonction d'activation Rectified Linear Unit (ReLu) est utilisée dans le DFFNN pour améliorer le processus de détection. Cette fonction utilisée est définie comme suit :

$$f(r) = \begin{cases} 0 & \text{Si } r < 0 \\ r & \text{Sinon} \end{cases} \quad (5-18)$$

Où r est la valeur d'entrée.

d) Détails de l'algorithme PGTO

Cette section décrit l'algorithme hybride PGTO, qui est formé avec de deux algorithmes POA et GTO. L'algorithme PGTO est conçu en incluant la phase d'exploration des pélicans dans la localisation actualisée des gorilles. Le comportement d'exploration de l'algorithme POA est inefficace puisqu'il ne prend en compte que l'itération actuelle pour mettre à jour la position. La phase d'exploration de l'algorithme GTO est donc adaptée et appliquée avec l'algorithme POA pour améliorer la mise à jour des positions, car l'algorithme GTO prend en compte l'itération suivante dans la phase d'exploration. L'emplacement actualisé de GTO est adapté pour améliorer l'efficacité de la détection et la

vitesse de traitement de l'algorithme conçu. Nous présentons ci-après les étapes de l'algorithme PGTO proposé.

Étape 1 : Initialisation

La POA est un algorithme d'optimisation dépendant de la population dans lequel les pélicans sont considérés comme les membres de la population. Les membres de la population sont initialisés sur la base des valeurs des bornes inférieure et supérieure et sont définis dans l'équation (5-1).

Étape 2 : Calcul de fonction de fitness

La fonction de fitness de l'algorithme PGTO est calculée sur la base de fonction de de l'erreur quadratique moyenne (MSE) pour atteindre la solution optimale en minimisant l'erreur. Elle est définie par l'équation (5-19).

$$F_p = \frac{1}{n} \sum_i^n (\hat{y}_i - y_i)^2 \quad (5-19)$$

Où, F_p spécifie la fonction de fitness, \hat{y}_i spécifie la sortie réelle et y_i est la sortie prédite du DFFNN, n est le nombre de samples (échantillons).

Après le calcul de la fonction de fitness, la stratégie de chasse est déroulée en deux étapes : les phases d'exploration et d'exploitation.

Étape 3 : Phase d'exploration (Se diriger vers la source de nourriture)

Dans la phase d'exploration, les pélicans trouvent la localisation exacte des proies en balayant les différents emplacements de l'espace de recherche et en se déplaçant ensuite vers les emplacements identifiés, ce qui améliore le pouvoir d'exploration de l'algorithme. Ainsi, l'expression du déplacement des pélicans vers la proie est définie dans l'équation (5-2).

Lorsque $F_p < F_i$, l'expression est donnée par :

$$x_{ij}(t+1) = x_{i,j}(t) + rand. \cdot (p_j - I \cdot x_{i,j}(t)) \quad (5-20)$$

$$x_{ij}(t+1) = rand. \cdot p_j + x_{ij}(t) - rand. \cdot I \cdot x_{ij}(t) \quad (5-21)$$

$$x_{ij}(t+1) = rand. \cdot p_j + x_{ij}(t)[1 - rand. \cdot I] \quad (5-22)$$

Pour améliorer la phase d'exploration de la POA, l'expression d'exploration de GTO est appliquée à l'emplacement actualisé du POA. D'après Algorithme GTO l'équation (5-6), lorsque $\text{rand} \geq 0,5$:

$$x_{ij}(t + 1) = (r_2 - C) * X_A(t) + L \times Z \times X_{ij}(t) \quad (5-23)$$

$$L \times Z \times X_{ij}(t) = x_{ij}(t + 1) - (r_2 - C) * X_A(t) \quad (5-24)$$

$$x_{ij}(t) = \frac{x_{ij}(t + 1) - (r_2 - C) * X_A(t)}{L \times Z} \quad (5-25)$$

Substituer l'équation (5-25) à l'équation (5-22) et $X_A(t)$ par p_j , on obtient :

$$x_{ij}(t + 1) = \text{rand} \cdot p_j + \frac{x_{ij}(t + 1) - (r_2 - C) * p_j}{L \times Z} [1 - \text{rand} \cdot I] \quad (5-26)$$

Signifie que :

$$x_{ij}(t + 1) = \frac{p_j}{L \times Z - 1 + \text{rand} \cdot I} \times (\text{rand} \times L \times Z - (r_2 - C)(1 - \text{rand} \cdot I)) \quad (5-27)$$

La même manière, lorsque $F_p \geq F_i$, on obtient :

$$x_{ij}(t + 1) = \frac{p_j}{L \times Z - 1 - \text{rand}} \times (\text{rand} \times L \times Z - (C - r_2)(1 + \text{rand})) \quad (5-28)$$

Où, $x_{ij}(t + 1)$ est l'emplacement de la solution $i^{\text{ème}}$ dans la $j^{\text{ème}}$ dimension à l'itération $t+1$, p_j signifie l'emplacement de la proie dans la $j^{\text{ème}}$ dimension, I représente la contrainte arbitraire dans $[1, 2]$, rand , r_2 sont deux nombres aléatoires dans $[0, 1]$.

Dans cette phase, la position mise à jour est considérée comme permettant de rejeter ou d'accepter la nouvelle localisation du pélican en utilisant l'équation (5-3).

Étape 4 : Phase d'exploitation

Dans la nature, les pélicans déploient leurs ailes au-dessus de l'eau après avoir repéré la zone de nourriture. Ce comportement conduit les poissons à se déplacer aléatoirement dans l'eau, de sorte que les pélicans rassemblent les poissons dans leur poche. Cette stratégie est utilisée pour améliorer le taux de convergence dans la région de chasse. En outre, la phase d'exploitation augmente la puissance de recherche ainsi que le taux d'exploitation des pélicans.

En outre, la modélisation analytique du processus de chasse s'exprime avec l'équation (5-4).

Dans cette phase, la localisation mise à jour est utilisée pour rejeter ou accepter la nouvelle localisation du pélican et qui s'exprime avec l'équation (5-5).

L'algorithme ci-après présente le pseudo-code de PGTO conçu.

Algorithm PGTO : Pseudo code

Saisir les informations relatives au problème d'optimisation.

Déterminer la taille de la population (N) et le nombre d'itérations (T).

Initialisation de la position des pélicans à l'aide de l'équation (5-1) et calcul de la fonction de fitness selon l'équation (5-19)

Pour $t = 1$ à T faire

Générer la position de la proie au hasard.

Pour $i = 1$ à N faire

Phase 1 : déplacement vers la proie (phase d'exploration).

Pour $j = 1$ à m faire

Calculer la nouvelle position de la $j^{\text{ème}}$ dimension à l'aide de l'équation (5-27) ou l'équation (5-28)

Mettre à jour le $i^{\text{ème}}$ membre de la population à l'aide de l'équation (5-3).

Phase 2 : Survoler à la surface de l'eau (phase d'exploitation).

Pour $j = 1$ à m faire

Calculer le nouvel état de la $j^{\text{ème}}$ dimension à l'aide de l'équation (5-4)

Mettre à jour le $i^{\text{ème}}$ membre de la population à l'aide de l'équation (5-5).

Mettre à jour la meilleure solution candidate.

Produire la meilleure solution candidate obtenue par la POA.

e) Entraînement de DFFNN à l'aide de PGTO

L'utilisation de PGTO avec DFFNN exploite la capacité de recherche globale de PGTO pour faire trouver les poids et les biais optimaux du DFFNN.

PGTO étant un algorithme basé sur une population, il nécessite une initialisation de positions de pélicans.

Dans notre travail, les positions de pélicans dans PGTO sont initialisées de manière aléatoire entre -1 et 1.

Chaque composante du vecteur de position d'un pélican représente des poids et des biais de réseau DFFNN.

Au cours de chaque itération, les pélicans sont soumis à deux phases, à savoir l'exploration et l'exploitation.

Les étapes du processus d'entraînement sont :

- 1) Diviser le dataset en ensembles d'apprentissage et de test
- 2) Initier les paramètres d'optimisation de l'algorithme PGTO en déterminant le nombre d'itérations T et la taille de la population (nombre de pélicans) N et les positions des pélicans qui sont générés de manière aléatoire à l'aide de l'équation (5-1) de sorte que la dimension de chaque pelican (solution) est le total des poids et des biais du DFFN, puis calcule de la fonction de fitness selon l'équation (5-19).
- 3) Générer la position de la proie au hasard et calcule sa fonction de fitness
- 4) Calculer la nouvelle position de du $i^{\text{ème}}$ pélican à l'aide de l'équation (5-27) ou l'équation (5-28) et mettre à jour la position $i^{\text{ème}}$ pelican de la population à l'aide de l'équation (5-3) (phase d'exploration).
- 5) Calculer la nouvelle position de du $i^{\text{ème}}$ pélican à l'aide de l'équation (5-4) et mettre à jour la position $i^{\text{ème}}$ pelican de la population à l'aide de l'équation (5-5) (phase d'exploitation).
- 6) Répéter l'étape 4-5 pour chaque pélican
- 7) Mettre à jour la meilleure solution candidate.
- 8) Répéter les étapes 3-7 jusqu'à ce que le nombre d'itérations soit atteint.
- 9) Les poids optimaux sont attribués au modèle DFFNN.
- 10) Evaluer le modèle DFFNN avec l'ensemble de test.

5.4 Résultats et discussion

L'expérimentation, les résultats obtenus et l'analyse de PGTO-DFFNN conçu seront présentés dans cette section.

Tous les traitements sont effectués sur un PC/Intel Core(TM) i7-8850H CPU @ 2.60GHz, 16 Go de mémoire (RAM), système d'exploitation Windows 10 (64 bits) et environnement python 3.5.

5.4.1 Métriques d'évaluation

1. **Accuracy** : l'accuracy est une mesure proportionnelle du nombre de prédictions correctes sur l'ensemble des prédictions. Il est indiqué par l'équation (5-29).

$$Accuracy (Acc) = \frac{T_{pos} + T_{nég}}{T_{pos} + T_{nég} + F_{pos} + F_{nég}} \quad (5-29)$$

Où : T_{pos} est le vrai positif et $T_{nég}$ le vrai négatif,

F_{pos} est le faux positif et $F_{nég}$ est le faux négatif.

2. **Precision** : La precision est évaluée par le rapport entre le nombre d'échantillons positifs correctement identifiés et le nombre total d'échantillons positivement détectés (équation (5-30)).

$$Precision (Prec) = \frac{T_{pos}}{T_{pos} + F_{pos}} \quad (5-30)$$

3. **Recall (Rappel)** : La valeur de recall est calculée comme le rapport entre le nombre d'échantillons positifs correctement identifiés et le nombre total d'échantillons positifs (voir l'équation (5-31)).

$$Recall (Rec) = \frac{T_{pos}}{T_{pos} + F_{nég}} \quad (5-31)$$

4. **F1-score** : Il prend la moyenne harmonique de la précision et du rappel d'un classificateur pour fournir une mesure unique et il est défini dans l'équation (5-32).

$$F1-Score = 2 * \frac{Prec * Rec}{Prec + Rec} \quad (5-32)$$

5.4.2 Analyse comparative

1) Comparaison des résultats de l'apprentissage DFFNN à l'aide de différents algorithmes Deep learning

Dans notre analyse expérimentale, nous avons utilisé le jeu de données choisi "Action recognition dataset". L'objectif était de réaliser une analyse comparative de différents modèles d'apprentissage profond basés sur l'algorithme d'optimisation backpropagation et le modèle proposé DFFNN basé sur l'algorithme d'optimisation hybride PGTO. Plus précisément, nous allons examiner quatre modèles DL différents : DFFNN, LSTM, CNN et CNN-LSTM. En évaluant ces modèles sur les données de test, nous avons cherché à évaluer leurs performances à l'aide de plusieurs mesures telles que l'accuracy, la precision et le Recall.

Dans un premier temps, on évalue l'influence du taux d'apprentissage et de test sur la performance du PGTO-DFFNN et les divers modèles DL, suivie de l'étude des effets de choix de la taille de batch sur la performance.

a) Analyse basée sur taux d'apprentissage et de test

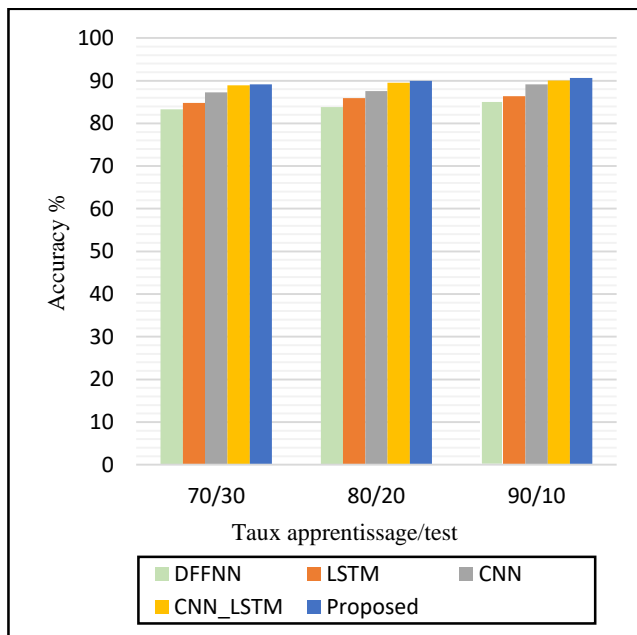
La Figure 5-3 illustre le résultat graphique à l'aide des métriques d'évaluations précédentes (Accuracy, precision et recall).

La Figure 5-3(a) montre l'accuracy des différentes modèles DL et DFFNN-PGTO proposé sur la base de taux d'apprentissage et de test en tant que mesure variable (90:10%, 80:20% , 70:30%).

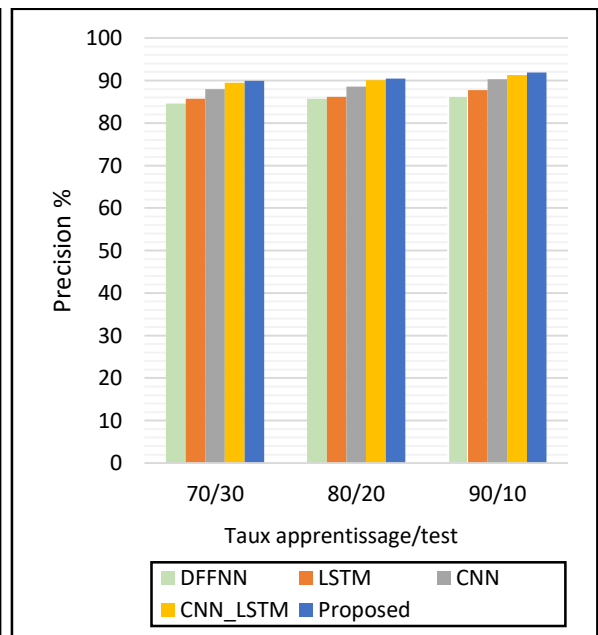
Lorsque nous définissons la valeur de taux 90:10%, les approches conventionnelles DFFNN, LSTM, CNN et CNN-LSTM basé sur l'algorithme d'optimisation backpropagation et notre approche DFFNN-PGTO atteignent un accuracy de 85.24%, 86.33%, 89.11%, 90.02% et 90.65%, respectivement (voir Figure 5-3(a)).

Lorsque le taux de l'apprentissage et de test est fixé à 80:20 % , la precision des différents algorithmes DL et PGTO-DFFNN atteignent respectivement 85.72%, 86.16%, 88.55%, 90.08% et 90.42% (voir Figure 5-3 (b)).

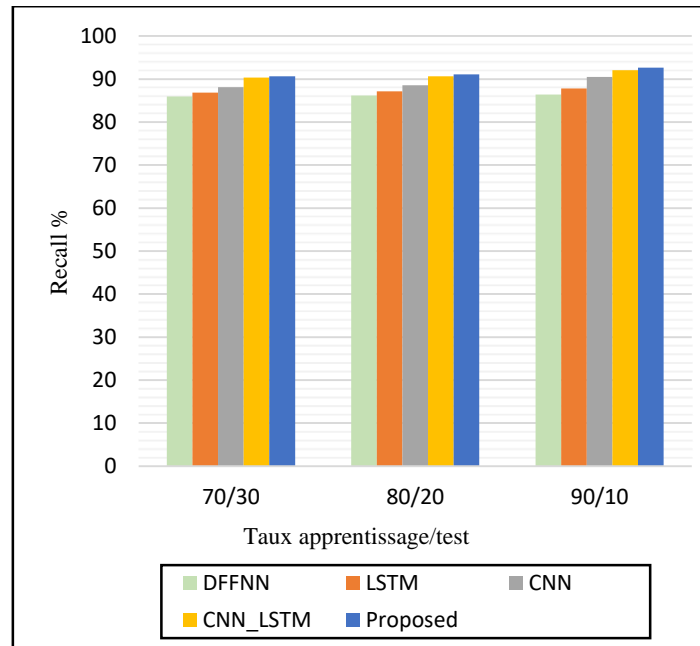
En fixant le taux d'apprentissage et de test à 70:30 % , la valeur de recall pour DFFNN atteint 85.92%, LSTM atteint 86.82%, CNN atteint 88.11%, CNN-LSTM atteint 90.32% et celle du PGTO-DFFNN proposé atteint 90.63% (voir Figure 5-3(c)).



(a)



(b)



(c)

Figure 5-3 : Comparaison basé sur différents taux d'apprentissage/test en termes de (a) accuracy, (b) precision, (c) recall pour différents DL algorithmes et PGTO-DFNN

Dans l'ensemble, les résultats ont montré que les performances des modèles DL considérés et notre approche PGTO-DFNN étaient sensiblement modifiées en fonction des différents rapports apprentissage/test et que notre approche a obtenu la meilleure performance.

b) Analyse basée sur le batch size

Nous explorons l'influence de différentes tailles de batch sur l'approche proposée et les différents modèles de DL, comme le montre la Figure 5-4. Les valeurs expérimentées pour la taille du batch sont 5, 7 et 8. On constate que l'accuracy, précisions, et recall s'améliorent avec l'augmentation de la taille des batch. Par exemple, lorsque la taille de batch égale à 8, l'accuracy pour PGTO-DFNN est de 90.51%, pour DFFNN est de 81.45%, de LSTM model est de 83.48%, de CNN model de 88.65% et de CNN-LSTM de 89.84%, comme montre la Figure 5-4 (a). En revanche, selon la Figure 5-4(b), la precision du modèle avec une taille de batch égale à 7 est de 90.23% pour DFFNN-PGTO, et 82.01%, 84.51%, 88.21% et 89.36% pour les autres modèles DL : DFFNN, LSTM, CNN, et CNN-LSTM respectivement.

Cependant, lorsque nous fixons la valeur de batch size à 5, les valeurs de recall des modèles DL sont de 81.09%, 84.56%, 88.53%, 89.77% respectivement, et 90.16% est obtenue par PGTO-DFFNN (Figure 5-4(c)).

Dans l'ensemble, les résultats expérimentaux montrent que les diverses métriques s'améliorent avec l'augmentation de la taille de batch et notre approche ayant les meilleures métriques.

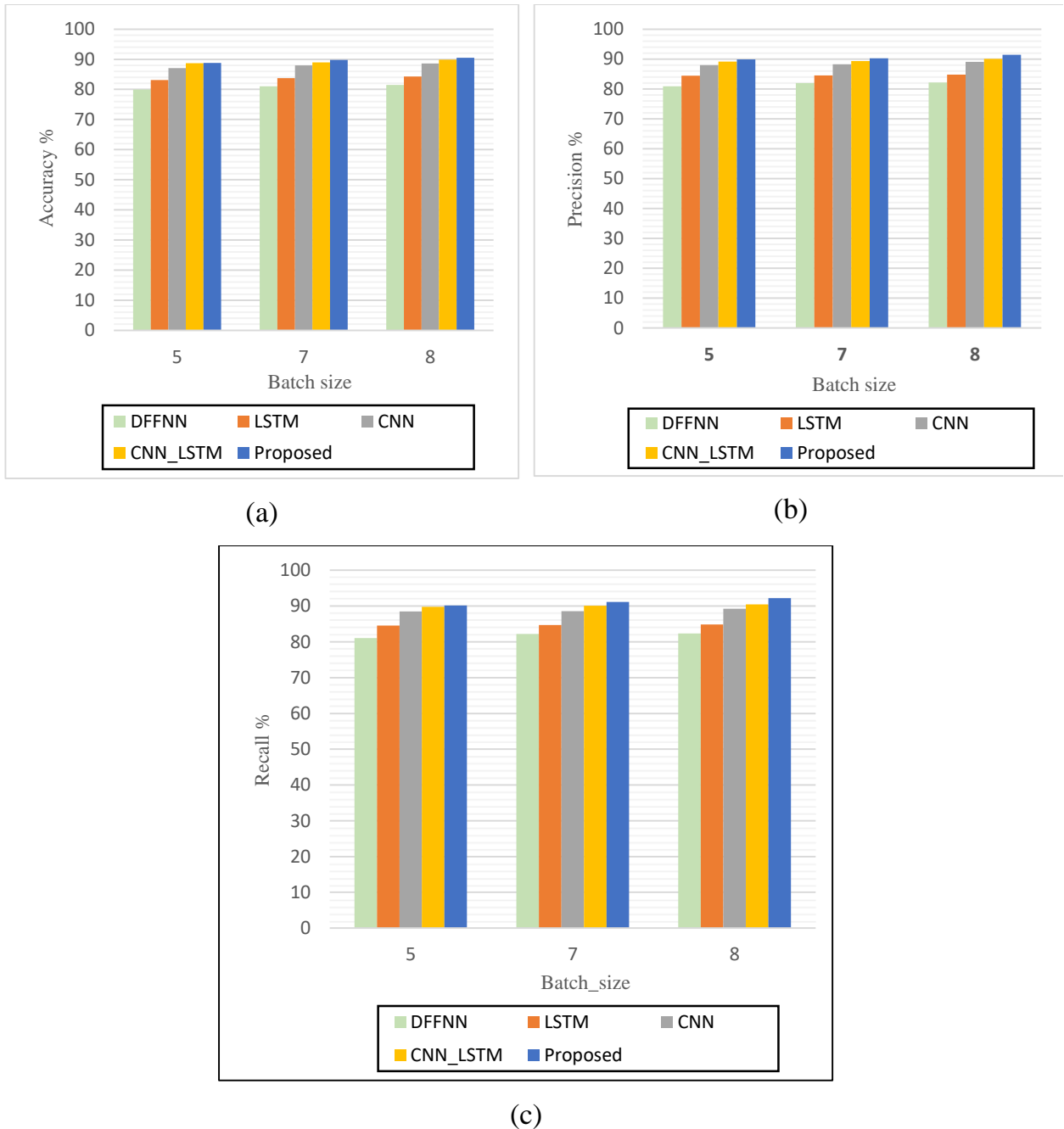


Figure 5-4 : Comparaison des performances pour différentes tailles de batch pour PGTO-DFFNN et divers DL modèles. La taille du batch est fixée à 5, 7 et 8 en termes de (a) Accuracy (b) Precision, (c) Recall

2) Comparaison des résultats de l'entraînement de DFFNN à l'aide de différents algorithmes d'optimisation

La performance de l'approche d'optimisation proposée est analysée en la comparant avec quatre algorithmes d'optimisation : Algorithme génétique (GA)[101], Algorithme Rider (ROA)[103], Algorithme Pelican (POA) et Algorithme Gorilla troop Optimization (GTO) pour l'entraînement du réseau DFFNN. Dans cette comparaison, l'analyse algorithmique est basée sur le choix de l'ajustement de taille de population (swarm size) et de l'ajustement du nombre d'itérations.

a) Analyse par ajustement de la taille de population

La taille de la population (swarm size) peut être considérée comme le paramètre de contrôle le plus « basique » des algorithmes d'optimisation, car elle définit le nombre d'individus dans la population et son réglage peut donc difficilement être évité. Elle a un impact sur les performances des algorithmes d'optimisation.

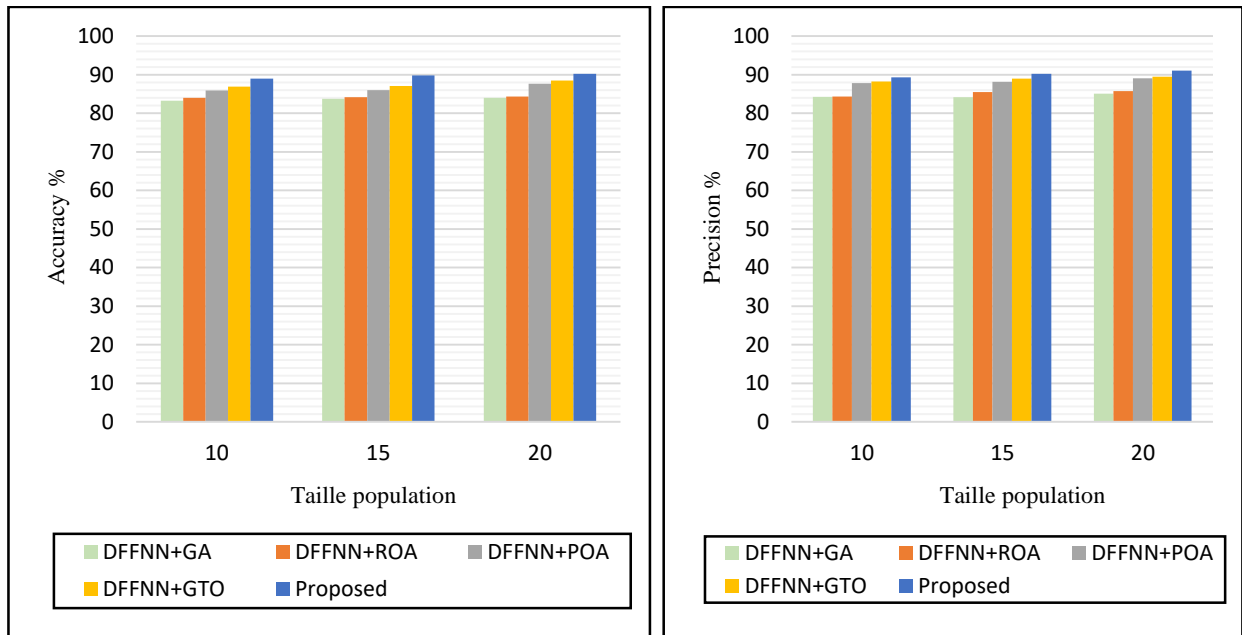
Dans le processus de l'entraînement, nous choisissons pour l'expérimentation trois tailles de population différentes : 10, 15 et 20, des résultats ont été obtenus pour chaque algorithme d'optimisation en utilisant le réseau DFFNN. On constate que l'accuracy, la precision, et le recall s'améliorent avec l'augmentation de la taille de population comme le montre la Figure 5-5. Par exemple, lorsque la taille de population égale à 20, l'accuracy pour DFFNN+PGTO est égal à 90.25%, pour DFFNN+GA est 84.02%, pour DFFNN+ROA est de 84.35%, pour DFFNN+POA de 87.62% et pour DFFNN+GTO est de 88.51%, comme montre la Figure 5-5(a).

En revanche, selon la Figure 5-5(b), la precision du modèle avec une taille de population égale à 15 est de 90.22% pour DFFNN-PGTO proposé, et 84.2%, 85.52%, 88.13% et 88.95% pour les autres algorithmes d'optimisation : DFFNN-GA, DFFNN-ROA, DFFNN-POA et DFFNN-GTO respectivement.

Cependant, lorsque nous fixons la valeur de taille de population à 10, nous obtenons les valeurs de recall suivantes : 86.23% (DFFNN-GA), 86.31% (DFFNN-ROA), 89.94% (DFFNN-POA), 90.11% (DFFNN-GTO), et 90.35% est obtenue par PGTO (Figure 5-5(c)).

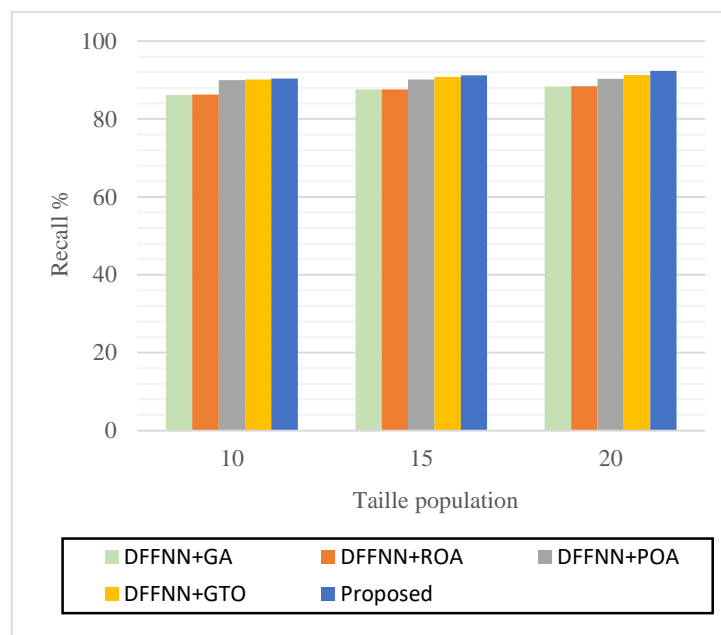
Les résultats obtenus montrent que l'algorithme hybride PGTO est plus performant que les autres algorithmes concurrents en fournissant des solutions optimales au problème

d'optimisation des poids et biais du modèle DFFNN, grâce à un équilibre proportionnel entre l'exploration et l'exploitation.



(a)

(b)



(c)

Figure 5-5 : Comparaison des performances des divers algorithmes d'optimisation basé sur différentes tailles de population (swarm size) en termes de (a) accuracy, (b) precision, (c) recall

b) Analyse par ajustement de nombre d'itérations

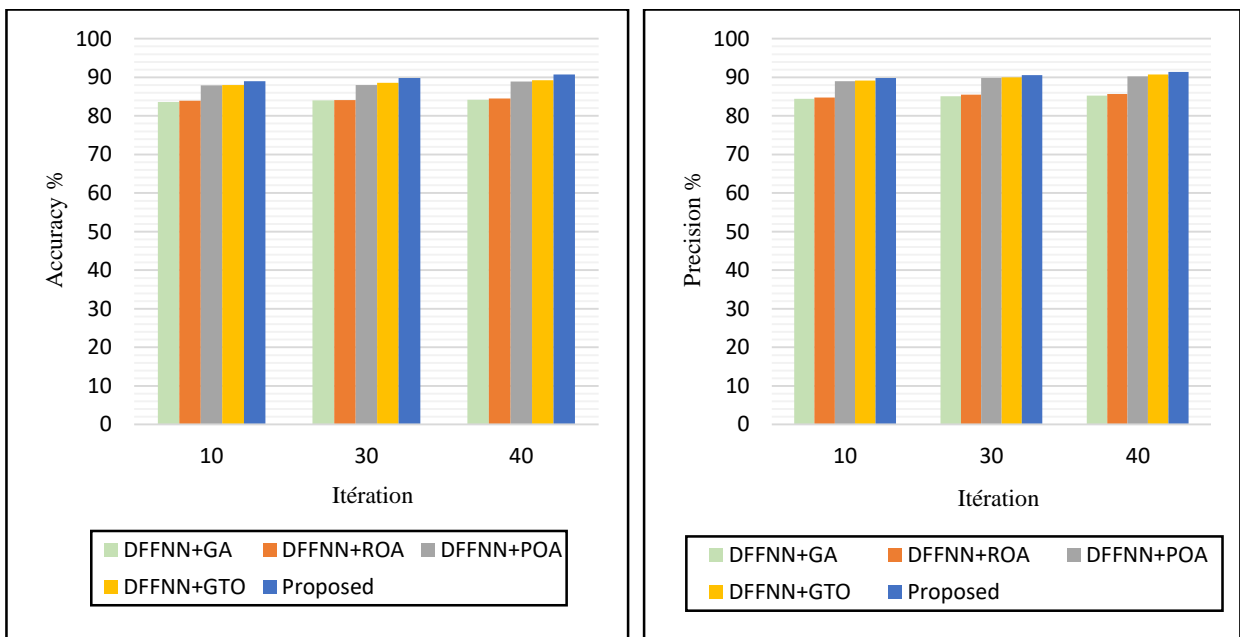
Pour optimiser le processus d'entraînement, nous avons testé et comparé notre algorithme hybride PGTO avec les autres algorithmes d'optimisation : GA, ROA, POA et

GTO avec le réseau DFFNN sur des valeurs d'itération différentes : 10, 30 et 40 comme indiqué dans la Figure 5-6.

Par exemple, dans la Figure 5-6(a) lorsque le nombre d'itérations égal à 40, nous avons obtenu l'accuracy de DFFNN avec GA est de 84.15%, avec ROA est de 84.51%, avec POA est de 88.91%, avec GTO est de 89.25% et avec l'algorithme hybride PGTO est de 90.75%. En revanche, selon la Figure 5-6(b), lorsque l'itération égale à 30 est la precision est de 90.56% pour DFFNN avec PGTO proposé, et 85.11%, 85.52%, 89.84%, et 90.02% pour DFFNN avec les autres algorithmes d'optimisation : GA, ROA, POA et GTO respectivement.

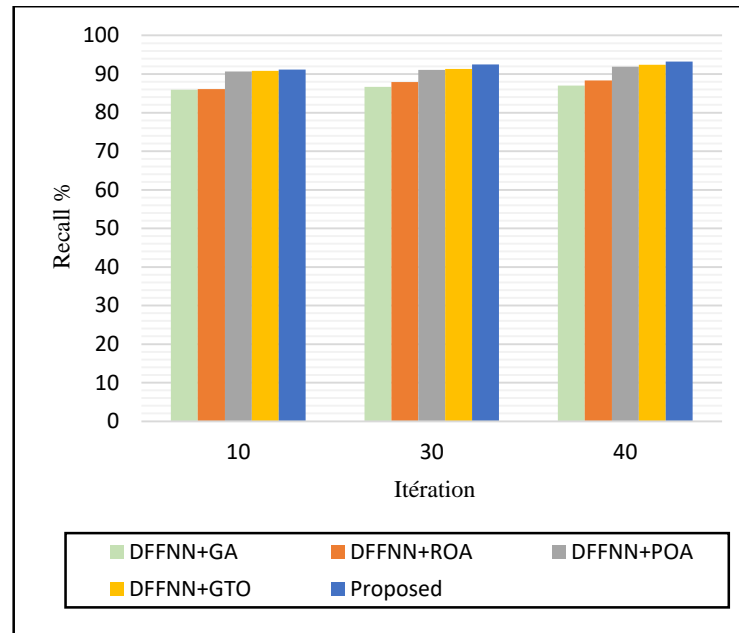
Cependant, lorsque nous fixons la valeur du nombre d'itérations à 10, nous obtenons les valeurs de recall pour DFFNN avec les algorithmes d'optimisation GA, ROA, POA et GTO: 85.91%, 86.07%, 90.64%, 90.85% respectivement, et la valeur 91.17% est obtenue pour DFFNN avec PGTO (Figure 5-6(c)).

Une analyse des résultats présentés nous permet de conclure que le réseau DFFNN entraîné à l'aide de PGTO fournit des meilleurs scores.



(a)

(b)



(c)

Figure 5-6: Analyse Comparatives entre PGTO et les autres algorithmes d'optimisation de DFFNN : GA, ROA, POA et GTO basée sur différents nombres d'itérations : 10,30 et 40 en termes de : (a) accuracy, (b) precision, (c) recall

3) Discussion

Les résultats obtenus montrent que l'approche PGTO-DFFNN proposé présente des valeurs de métriques acceptables : accuracy, Precision, Recall et F1-score.

La Figure 5-7 montre la matrice de confusion normalisée pour le dataset choisi en utilisant l'approche PGTO-DFFNN. Les valeurs affichées dans cette figure prouvent que le classificateur fonctionne correctement. D'après cette matrice, les valeurs les plus élevées pour chaque type de classes se situent sur la diagonale : 0.92 est l'accuracy de classification de deux classes Sleeping et Relaxing, tandis que 0.97 pour la classe Working et 0.91 pour la classe Emergency. D'autre part, la plupart des valeurs non diagonales sont des faibles erreurs. Certaines actions sont confondues avec d'autres qui présentent des actions similaires. Par exemple, l'action sleeping est souvent confondue avec l'action Emergency (chute).

True label	Class-1	0.92	0.03	0.01	0.03
	Class-2	0.03	0.92	0.01	0.04
	Class-3	0.01	0.01	0.97	0.01
	Class-4	0.02	0.04	0.02	0.91
		Class-1	Class-2	Class-3	Class-4
		Predicted label			

Figure 5-7 : Matrice de confusion normalisée utilisant l'approche PGTO-DFFNN (classe 1 : Relaxing, classe 2 : Sleeping, classe 3 : Working, et classe 4 : Emergency)

Le Tableau 5-1 présente les résultats en tenant compte des paramètres optimaux : le taux d'apprentissage/test et la valeur de batch size pour les différents modèles : DFFNN, LSTM, CNN et CNN-LSTM et notre approche. Les résultats obtenus montre que PGTO-DFFNN obtient un meilleur score sur la base des métriques : accuracy, precision, recall et F1-score qui sont : 90.65%, 91.92%, 92.66% et 92.29% respectivement.

Tableau 5-1: Analyse comparative

Paramètres	Métriques	DFFNN	LSTM	CNN	CNN-LSTM	Notre approche
Taux apprentissage /test = (90 :10) batch size =8	Accuracy	85.24	86.33	89.11	90.02	90.65
	Precision	86.36	87.75	90.33	91.26	91.92
	Recall	86.41	87.80	90.48	92.03	92.66
	F1-Score	86.38	87.77	90.40	91.64	92.29

Le Tableau 5-2 présente les résultats en tenant compte les paramètres optimaux : la taille de la population et le nombre d'itérations pour notre approche PGTO-DFFNN et les autres approches DFFNN+GA, DFFNN+ROA, DFFNN+POA et DFFNN+GTO. L'approche PGTO-DFFNN proposée a atteint un meilleur score d'accuracy, de precision, de recall et de F1-score qui sont 90.75%, 91.39%, 93.21% et 92.29% respectivement, ce qui confirme ses gains de performance par rapport aux autres approches.

Tableau 5-2: Analyse algorithmique

Paramètres	Métriques	DFFNN + GA	DFFNN + ROA	DFFNN + Pelican	DFFNN + GTO	Notre approche
Taille population = 20 Nb. iterations =40	Accuracy	84.15	84.51	88.91	89.25	90.75
	Precision	85.26	85.67	90.22	90.75	91.39
	Recall	87.05	88.32	91.87	92.44	93.21
	F1-Score	86.15	86.97	91.04	91.59	92.29

Pour évaluer l'efficacité de l'approche proposée, nous les avons comparés à un certain nombre de modèles dans d'autres travaux précédents. Le Tableau 5-3 compare l'approche PGTO-DFFNN aux modèles conventionnelles (SVM, CNN).

Tableau 5-3: Comparaison entre PGTO-DFFNN et les autres méthodes

	Méthode	Accuracy	Precision	Recall	F1-Score
[104]	SVM	86.5	87.5	85.9	86.6
[105]	CNN	87.1	88.6	90.2	89.3
Proposé	DFFNN-PGTO	90.75	91.39	93.21	92.29

Généralement, la bonne performance de l'approche proposée est obtenue en concaténant les avantages des deux algorithmes d'optimisation du pélican et des gorilles, ce qui permet d'améliorer la détection.

5.5 Conclusion

L'entraînement d'un réseau DFFNN est un problème d'optimisation et, par conséquent, l'algorithme hybride d'optimisation utilisé est d'une importance primordiale. La backpropagation est l'algorithme le plus largement utilisé pour entraîner les réseaux DFFNN. Cependant, la backpropagation n'est pas idéale et est souvent incapable de trouver le minimum global. Les algorithmes d'optimisation inspirés de la nature (méta-heuristiques) peuvent être utilisés pour entraîner efficacement les réseaux DFFNN. Dans ce chapitre, nous avons proposé une approche d'entraînement des réseaux DFFNN qui utilise un algorithme hybride PGTO pour déterminer les valeurs optimales des poids et de biais pour améliorer les performances lors de la résolution de problèmes du monde réel.

Nous avons appliqué cette approche dans le problème de détection des anomalies dans l'activité chez la personne âgée. Les résultats expérimentaux montrent que l'approche proposée a permis d'améliorer les performances de détection sur la base des métriques d'accuracy, de precision et de recall.

Chapitre 6 Conclusion et travaux futurs

Dans cette section, nous présentons les principales contributions de la présente étude de recherche. Ensuite, des travaux futurs pouvant être envisagées ont été introduites.

6.1 Contribution de la thèse

Ces dernières années, il a été observé que le secteur de la population ciblée comme étant âgée augmente plus rapidement qu'au cours des décennies précédentes. En outre, les problèmes de santé liés à l'âge et les différentes maladies chroniques commencent à se manifester à ce stade de la vie. Par ailleurs, l'augmentation du coût des soins infirmiers et la pénurie de soignants ont entraîné une augmentation du besoin d'aide à la vie à domicile. Par conséquent, il est essentiel d'aider les personnes âgées à mener une vie indépendante plus sûre.

La surveillance et la reconnaissance des activités humaines (HAR) dans des maisons intelligentes sont importantes pour évaluer les changements dans les profils physiques et comportementaux de la population au fil du temps. En particulier, les systèmes HAR basées sur des capteurs portables qui présentent de nombreux avantages par rapport aux méthodes HAR basées sur la vision, qui soulèvent des problèmes de confidentialité. Les algorithmes de reconnaissance d'activité basés sur DL sont au cœur de HAR.

A cet égard, nos principales contributions dans cette thèse sont :

- Un état de l'art relatif au sujet proposé : Dans le but de récupérer des informations servant de base aux approches introduites, des études connexes a été réalisée et une discussion autour des solutions existantes.
- la proposition d'un système de reconnaissance des activités physiques basé sur l'apprentissage profond à travers des modèles de bases (CNN et LSTM) et des modèles hybrides (CNN-LSTM) et CNN Multi-têtes-LSTM. Les résultats de cette étude ont démontré que les modèles CNN et LSTM peuvent être utilisées ensemble pour générer une solution hybride avec une capacité élevée de reconnaissance et de classification d'activités physiques. Les résultats expérimentaux sont faits sur le dataset Mhealth et montrent que le modèle CNN multi-têtes-LSTM proposé est

meilleur en termes d'accuracy, de precision et de recall par rapport à d'autres modèles, suivi par le modèle CNN-LSTM.

Le modèle CNN multi-têtes-LSTM a atteint un accuracy de 98,69% pour la prédiction des douze activités physiques de dix personnes, tandis que la precision, le recall et F1-Score sont de 98,41 % et 98,41%, 98,50% respectivement.

- le développement d'un nouvel algorithme hybride PGTO qui surmonte la possibilité de stagnation de l'algorithme de backpropagation en local-minimum pendant l'entraînement de réseau DFFNN, et il est utilisé pour trouver les poids et les biais optimaux de DFFNN. Nous avons appliqué cet algorithme hybride dans un problème de détection des anomalies dans les activités d'une personne âgée en utilisant DFFNN. Les résultats expérimentaux montrent que cette approche proposée PGTO-DFFNN à améliorer les performances de détection sur la base des métriques d'accuracy, de precision et de recall.

6.2 Travaux futurs

Les travaux futurs recommandés pour cette recherche sont les suivants :

- La reconnaissance des activités humaines vues pourrait également être étendue à d'autres types d'activités, notamment aux activités contextuelles telles que cuisiner et regarder la télévision, faire sa toilette, ...etc. pour aider les soignants à surveiller la santé des personnes âgées vivant de manière autonome à domicile et à identifier toute anomalie liée à leur vie quotidienne.
- Nous prévoyons d'étendre les algorithmes en introduisant les signes vitaux tels que la pression artérielle, fréquence cardiaque ...etc. dans le système de suivi et surveillance afin de détecter les situations d'urgence.
- Nous explorerons des techniques avancées basées sur l'apprentissage profond telles que la combinaison de CNN avec un mécanisme d'attention ou transfer learning pour la reconnaissance des activités. De plus, nous prévoyons de développer un vaste dataset HAR qui inclura plusieurs activités physiques.
- l'approche DFFNN-PGTO proposée sera appliquée à des datasets plus complexes et plus importants, avec d'avantage d'activités.

Bibliographie

- [1] G. Demiris *et al.*, « Older adults' attitudes towards and perceptions of 'smart home' technologies: a pilot study », *Med. Inform. Internet Med.*, vol. 29, n° 2, p. 87-94, juin 2004, doi: 10.1080/14639230410001684387.
- [2] F. G. Miskelly, « Assistive technology in elderly care », *Age Ageing*, vol. 30, n° 6, p. 455-458, nov. 2001, doi: 10.1093/ageing/30.6.455.
- [3] C. Berdjouh, M. C. E. Meftah, A. Laouid, M. Hammoudeh, et A. Kumar, « Pelican Gorilla Troop Optimization Based on Deep Feed Forward Neural Network for Human Activity Abnormality Detection in Smart Spaces », *IEEE Internet Things J.*, vol. 10, n° 21, p. 18495-18504, nov. 2023, doi: 10.1109/JIOT.2023.3271831.
- [4] C. Perera, A. Zaslavsky, P. Christen, et D. Georgakopoulos, « Context Aware Computing for The Internet of Things: A Survey », *IEEE Commun. Surv. Tutor.*, vol. 16, n° 1, p. 414-454, 2014, doi: 10.1109/SURV.2013.042313.00197.
- [5] Q. Lê, H. B. Nguyen, et T. Barnett, « Smart Homes for Older People: Positive Aging in a Digital World », *Future Internet*, vol. 4, n° 2, Art. n° 2, juin 2012, doi: 10.3390/fi4020607.
- [6] M. Bhagat, D. Kumar, et S. M. Balgi, « Application of Internet of Things in Digital Pedagogy », in *Computational Intelligence in Digital Pedagogy*, A. Deyasi, S. Mukherjee, A. Mukherjee, A. K. Bhattacharjee, et A. Mondal, Éd., Singapore: Springer Singapore, 2021, p. 219-234. doi: 10.1007/978-981-15-8744-3_11.
- [7] H. T. Malazi et M. Davari, « Combining emerging patterns with random forest for complex activity recognition in smart homes », *Appl. Intell.*, vol. 48, n° 2, p. 315, 2017.
- [8] G. Singla, D. J. Cook, et M. Schmitter-Edgecombe, « Recognizing independent and joint activities among multiple residents in smart environments », *J. Ambient Intell. Humaniz. Comput.*, vol. 1, n° 1, p. 57-63, mars 2010, doi: 10.1007/s12652-009-0007-1.
- [9] P. Rashidi et A. Mihailidis, « A Survey on Ambient-Assisted Living Tools for Older Adults », *IEEE J. Biomed. Health Inform.*, vol. 17, n° 3, p. 579-590, mai 2013, doi: 10.1109/JBHI.2012.2234129.
- [10] « AAL Home 2020 », AAL Programme. Consulté le: 3 septembre 2023. [En ligne]. Disponible sur: <http://www.aal-europe.eu/>
- [11] T. Kleinberger, M. Becker, E. Ras, A. Holzinger, et P. Müller, « Ambient Intelligence in Assisted Living: Enable Elderly People to Handle Future Interfaces », in *Universal Access in Human-Computer Interaction. Ambient Interaction*, C. Stephanidis, Éd., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2007, p. 103-112. doi: 10.1007/978-3-540-73281-5_11.
- [12] F. K. Aldrich, « Smart Homes: Past, Present and Future », in *Inside the Smart Home*, R. Harper, Éd., London: Springer, 2003, p. 17-39. doi: 10.1007/1-85233-854-7_2.
- [13] N. Balta-Ozkan, R. Davidson, M. Bicket, et L. Whitmarsh, « Social barriers to the adoption of smart homes », *Energy Policy*, vol. 63, p. 363-374, déc. 2013, doi: 10.1016/j.enpol.2013.08.043.

- [14] S. Sepasgozar *et al.*, « A Systematic Content Review of Artificial Intelligence and the Internet of Things Applications in Smart Home », *Appl. Sci.*, vol. 10, n° 9, Art. n° 9, janv. 2020, doi: 10.3390/app10093074.
- [15] C. Carra et K. Tabia, « Smart Home for Seniors: Opportunities and Challenges for AI », févr. 2020, Consulté le: 5 septembre 2023. [En ligne]. Disponible sur: <https://lilloa.univ-lille.fr/handle/20.500.12210/19027>
- [16] S. Helal, W. Mann, H. El-Zabadani, J. King, Y. Kaddoura, et E. Jansen, « The Gator Tech Smart House: a programmable pervasive space », *Computer*, vol. 38, n° 3, p. 50-60, mars 2005, doi: 10.1109/MC.2005.107.
- [17] M. Chan, D. Estève, C. Escriba, et E. Campo, « A review of smart homes—Present state and future challenges », *Comput. Methods Programs Biomed.*, vol. 91, n° 1, p. 55-81, juill. 2008, doi: 10.1016/j.cmpb.2008.02.001.
- [18] D. J. Cook, A. S. Crandall, B. L. Thomas, et N. C. Krishnan, « CASAS: A Smart Home in a Box », *Computer*, vol. 46, n° 7, p. 10.1109/MC.2012.328, juill. 2013, doi: 10.1109/MC.2012.328.
- [19] M. Vacher *et al.*, « The sweet-home project: audio technology in smart homes to improve well-being and reliance », *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2011, p. 5291-5294, 2011, doi: 10.1109/IEMBS.2011.6091309.
- [20] P. Chahuara, F. Portet, et M. Vacher, « Making Context Aware Decision from Uncertain Information in a Smart Home: A Markov Logic Network Approach », in *Ambient Intelligence*, J. C. Augusto, R. Wichert, R. Collier, D. Keyson, A. A. Salah, et A.-H. Tan, Éd., Cham: Springer International Publishing, 2013, p. 78-93.
- [21] P. E. Antoniou, E. I. Konstantinidis, A. S. Billis, et P. D. Bamidis, « Integrating the USEFIL Assisted Living Platform; Observation from the Field », in *6th European Conference of the International Federation for Medical and Biological Engineering*, I. Lacković et D. Vasic, Éd., in IFMBE Proceedings. Cham: Springer International Publishing, 2015, p. 657-660. doi: 10.1007/978-3-319-11128-5_164.
- [22] A. S. Billis, E. I. Papageorgiou, C. Frantzidis, E. I. Konstantinidis, et P. D. Bamidis, « Towards a hierarchically-structured decision support tool for improving seniors' independent living: the USEFIL decision support system », in *Proceedings of the 6th International Conference on Pervasive Technologies Related to Assistive Environments*, in PETRA '13. New York, NY, USA: Association for Computing Machinery, mai 2013, p. 1-4. doi: 10.1145/2504335.2504362.
- [23] Q. Zhang, M. Karunanithi, R. Rana, et J. Liu, « Determination of Activities of Daily Living of independent living older people using environmentally placed sensors », *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2013, p. 7044-7047, 2013, doi: 10.1109/EMBC.2013.6611180.
- [24] Q. Zhang, Y. Su, et P. Yu, « Assisting an Elderly with Early Dementia Using Wireless Sensors Data in Smarter Safer Home », in *Service Science and Knowledge Innovation*, K. Liu, S. R. Gulliver, W. Li, et C. Yu, Éd., in IFIP Advances in Information and Communication Technology. Berlin, Heidelberg: Springer, 2014, p. 398-404. doi: 10.1007/978-3-642-55355-4_41.

- [25] T. B. Moeslund, A. Hilton, et V. Krüger, « A survey of advances in vision-based human motion capture and analysis », *Comput. Vis. Image Underst.*, vol. 104, n° 2-3, p. 90-126, nov. 2006, doi: 10.1016/j.cviu.2006.08.002.
- [26] P. Turaga, R. Chellappa, V. S. Subrahmanian, et O. Udrea, « Machine Recognition of Human Activities: A Survey », *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, n° 11, p. 1473-1488, nov. 2008, doi: 10.1109/TCSVT.2008.2005594.
- [27] O. Brdiczka, J. L. Crowley, et P. Reignier, « Learning Situation Models in a Smart Home », *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 39, n° 1, p. 56-63, févr. 2009, doi: 10.1109/TSMCB.2008.923526.
- [28] J. Zhang, Y. Shan, et K. Huang, « ISEE Smart Home (ISH): Smart video analysis for home security », *Neurocomputing*, vol. 149, p. 752-766, févr. 2015, doi: 10.1016/j.neucom.2014.08.002.
- [29] C. Peters, T. Hermann, S. Wachsmuth, et J. Hoey, « Automatic Task Assistance for People with Cognitive Disabilities in Brushing Teeth - A User Study with the TEBRA System », *ACM Trans. Access. Comput.*, vol. 5, n° 4, p. 10:1-10:34, mars 2014, doi: 10.1145/2579700.
- [30] S. J. Ismail, M. A. A. Rahman, S. A. Mazlan, et H. Zamzuri, « Human gesture recognition using a low cost stereo vision in rehab activities », in *2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*, oct. 2015, p. 220-225. doi: 10.1109/IRIS.2015.7451615.
- [31] M. Philipose *et al.*, « Inferring activities from interactions with objects », *IEEE Pervasive Comput.*, vol. 3, n° 4, p. 50-57, oct. 2004, doi: 10.1109/MPRV.2004.7.
- [32] E. M. Tapia, S. S. Intille, et K. Larson, « Activity Recognition in the Home Using Simple and Ubiquitous Sensors », in *Pervasive Computing*, A. Ferscha et F. Mattern, Éd., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2004, p. 158-175. doi: 10.1007/978-3-540-24646-6_10.
- [33] K. D. Feuz, D. J. Cook, C. Rosasco, K. Robertson, et M. Schmitter-Edgecombe, « Automated Detection of Activity Transitions for Prompting », *IEEE Trans. Hum.-Mach. Syst.*, vol. 45, n° 5, p. 575-585, oct. 2015, doi: 10.1109/THMS.2014.2362529.
- [34] L. Gillani Fahad, A. Khan, et M. Rajarajan, « Activity recognition in smart homes with self verification of assignments », *Neurocomputing*, vol. 149, p. 1286-1298, févr. 2015, doi: 10.1016/j.neucom.2014.08.069.
- [35] M. Z. Uddin et A. Soylu, « Human activity recognition using wearable sensors, discriminant analysis, and long short-term memory-based neural structured learning », *Sci. Rep.*, vol. 11, n° 1, Art. n° 1, août 2021, doi: 10.1038/s41598-021-95947-y.
- [36] « Projet SWEET-HOME - ANR », SWEET-HOME « Système Domotique d'Assistance au Domicile ». Consulté le: 3 septembre 2023. [En ligne]. Disponible sur: <http://sweet-home.imag.fr/index.php?choix=projet>
- [37] S. Katz, « Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living », *J. Am. Geriatr. Soc.*, vol. 31, n° 12, p. 721-727, déc. 1983, doi: 10.1111/j.1532-5415.1983.tb03391.x.

- [38] B. Guidet *et al.*, « The contribution of frailty, cognition, activity of daily life and comorbidities on outcome in acutely admitted patients over 80 years in European ICUs: the VIP2 study », *Intensive Care Med.*, vol. 46, n° 1, p. 57-69, janv. 2020, doi: 10.1007/s00134-019-05853-1.
- [39] S. Katz, A. B. Ford, R. W. Moskowitz, B. A. Jackson, et M. W. Jaffe, « STUDIES OF ILLNESS IN THE AGED. THE INDEX OF ADL: A STANDARDIZED MEASURE OF BIOLOGICAL AND PSYCHOSOCIAL FUNCTION », *JAMA*, vol. 185, p. 914-919, sept. 1963, doi: 10.1001/jama.1963.03060120024016.
- [40] M. P. Lawton et E. M. Brody, « Assessment of older people: Self-maintaining and instrumental activities of daily living », *The Gerontologist*, vol. 9, n° 3, Pt 1, p. 179-186, 1969, doi: 10.1093/geront/9.3_Part_1.179.
- [41] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, et Z. Yu, « Sensor-Based Activity Recognition », *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 42, n° 6, p. 790-808, nov. 2012, doi: 10.1109/TSMCC.2012.2198883.
- [42] J. K. Aggarwal et L. Xia, « Human activity recognition from 3D data: A review », *Pattern Recognit. Lett.*, vol. 48, p. 70-80, oct. 2014, doi: 10.1016/j.patrec.2014.04.011.
- [43] S. Wang et G. Zhou, « A review on radio based activity recognition », *Digit. Commun. Netw.*, vol. 1, n° 1, p. 20-29, févr. 2015, doi: 10.1016/j.dcan.2015.02.006.
- [44] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, et F. E. Alsaadi, « A survey of deep neural network architectures and their applications », *Neurocomputing*, vol. 234, p. 11-26, avr. 2017, doi: 10.1016/j.neucom.2016.12.038.
- [45] W. Wang, H. Xu, M. Alazab, T. R. Gadekallu, Z. Han, et C. Su, « Blockchain-Based Reliable and Efficient Certificateless Signature for IIoT Devices », *IEEE Trans. Ind. Inform.*, vol. 18, n° 10, p. 7059-7067, oct. 2022, doi: 10.1109/TII.2021.3084753.
- [46] M. C. E. Meftah, C. A. Cherif, B. Meraghni, et C. Berdjouh, « Hybrid Architecture Based on Blockchain Technology and Smart Contracts for Healthcare Applications », *Int. J. Organ. Collect. Intell.*, vol. 12, n° 1, p. 1-20, nov. 2022, doi: 10.4018/IJOCI.313600.
- [47] A. C. Müller et S. Guido, *Introduction to machine learning with Python: a guide for data scientists*, First edition. Sebastopol, CA: O'Reilly Media, Inc, 2016.
- [48] W. Pitts, « Some observations on the simple neuron circuit », *Bull. Math. Biophys.*, vol. 4, n° 3, p. 121-129, sept. 1942, doi: 10.1007/BF02477942.
- [49] K. Fukushima, « Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position », *Biol. Cybern.*, vol. 36, n° 4, p. 193-202, avr. 1980, doi: 10.1007/BF00344251.
- [50] C. Cortes et V. Vapnik, « Support-vector networks », *Mach. Learn.*, vol. 20, n° 3, p. 273-297, sept. 1995, doi: 10.1007/BF00994018.
- [51] S. Hochreiter et J. Schmidhuber, « Long Short-Term Memory », *Neural Comput.*, vol. 9, n° 8, p. 1735-1780, nov. 1997, doi: 10.1162/neco.1997.9.8.1735.
- [52] M. A. Nielsen, *Neural Networks and Deep Learning*. Determination Press, 2015.

- [53] Jeremy Jordan et Jeremy, « Setting the learning rate of your neural network. », Jeremy Jordan. Consulté le: 15 février 2024. [En ligne]. Disponible sur: <https://www.jeremyjordan.me/nn-learning-rate/>
- [54] A. Krizhevsky, I. Sutskever, et G. E. Hinton, « ImageNet classification with deep convolutional neural networks », *Commun. ACM*, vol. 60, n° 6, p. 84-90, mai 2017, doi: 10.1145/3065386.
- [55] S. Abbaspour, F. Fotouhi, A. Sedaghatbaf, H. Fotouhi, M. Vahabi, et M. Linden, « A Comparative Analysis of Hybrid Deep Learning Models for Human Activity Recognition », *Sensors*, vol. 20, n° 19, p. 5707, oct. 2020, doi: 10.3390/s20195707.
- [56] W. Fang, Y. Chen, et Q. Xue, « Survey on Research of RNN-Based Spatio-Temporal Sequence Prediction Algorithms », *J. Big Data*, vol. 3, n° 3, p. 97-110, 2021, doi: 10.32604/jbd.2021.016993.
- [57] J. Xiao et Z. Zhou, « Research Progress of RNN Language Model », in *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, Dalian, China: IEEE, juin 2020, p. 1285-1288. doi: 10.1109/ICAICA50127.2020.9182390.
- [58] H. Apaydin, H. Feizi, M. T. Sattari, M. S. Colak, S. Shamshirband, et K.-W. Chau, « Comparative Analysis of Recurrent Neural Network Architectures for Reservoir Inflow Forecasting », *Water*, vol. 12, n° 5, Art. n° 5, mai 2020, doi: 10.3390/w12051500.
- [59] A. Shewalkar, D. Nyavanandi, et S. A. Ludwig, « Performance Evaluation of Deep Neural Networks Applied to Speech Recognition: RNN, LSTM and GRU », *J. Artif. Intell. Soft Comput. Res.*, vol. 9, n° 4, p. 235-245, oct. 2019, doi: 10.2478/jaiscr-2019-0006.
- [60] F. Kratzert, D. Klotz, C. Brenner, K. Schulz, et M. Herrnegger, « Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks », *Hydrol. Earth Syst. Sci.*, vol. 22, n° 11, p. 6005-6022, nov. 2018, doi: 10.5194/hess-22-6005-2018.
- [61] A. Graves, « Generating Sequences With Recurrent Neural Networks ». arXiv, 5 juin 2014. Consulté le: 4 septembre 2023. [En ligne]. Disponible sur: <http://arxiv.org/abs/1308.0850>
- [62] F. M. Shiri, T. Perumal, N. Mustapha, R. Mohamed, M. A. B. Ahmadon, et S. Yamaguchi, « A Survey on Multi-Resident Activity Recognition in Smart Environments ». arXiv, 24 avril 2023. doi: 10.48550/arXiv.2304.12304.
- [63] C. Etienne, G. Fidanza, A. Petrovskii, L. Devillers, et B. Schmauch, « CNN+LSTM Architecture for Speech Emotion Recognition with Data Augmentation », in *Workshop on Speech, Music and Mind (SMM 2018)*, sept. 2018, p. 21-25. doi: 10.21437/SMM.2018-5.
- [64] J. Zegers et H. Van hamme, « CNN-LSTM models for Multi-Speaker Source Separation using Bayesian Hyper Parameter Optimization ». arXiv, 19 décembre 2019. doi: 10.48550/arXiv.1912.09254.
- [65] T. Akilan, Q. J. Wu, A. Safaei, J. Huo, et Y. Yang, « A 3D CNN-LSTM-Based Image-to-Image Foreground Segmentation », *IEEE Trans. Intell. Transp. Syst.*, vol. 21, n° 3, p. 959-971, mars 2020, doi: 10.1109/TITS.2019.2900426.
- [66] D. N. Tran et D. D. Phan, « Human Activities Recognition in Android Smartphone Using Support Vector Machine », in *2016 7th International Conference on Intelligent Systems*,

- Modelling and Simulation (ISMS)*, Bangkok, Thailand: IEEE, janv. 2016, p. 64-68. doi: 10.1109/ISMS.2016.51.
- [67] J. Liu, J. Sohn, et S. Kim, « Classification of Daily Activities for the Elderly Using Wearable Sensors », *J. Healthc. Eng.*, vol. 2017, p. 1-7, 2017, doi: 10.1155/2017/8934816.
- [68] M. Vakili et M. Rezaei, « Incremental Learning Techniques for Online Human Activity Recognition ». arXiv, 20 septembre 2021. Consulté le: 18 octobre 2023. [En ligne]. Disponible sur: <http://arxiv.org/abs/2109.09435>
- [69] J. Muangprathub, A. Sriwichian, A. Wanichsombat, S. Kajornkasirat, P. Nillaor, et V. Boonjing, « A Novel Elderly Tracking System Using Machine Learning to Classify Signals from Mobile and Wearable Sensors », *Int. J. Environ. Res. Public Health*, vol. 18, n° 23, p. 12652, nov. 2021, doi: 10.3390/ijerph182312652.
- [70] S. Chernbumroong, S. Cang, A. Atkins, et H. Yu, « Elderly activities recognition and classification for applications in assisted living », *Expert Syst. Appl.*, vol. 40, n° 5, p. 1662-1674, avr. 2013, doi: 10.1016/j.eswa.2012.09.004.
- [71] N. Y. Hammerla, S. Halloran, et T. Ploetz, « Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables ». arXiv, 29 avril 2016. Consulté le: 17 octobre 2023. [En ligne]. Disponible sur: <http://arxiv.org/abs/1604.08880>
- [72] C. Xu, D. Chai, J. He, X. Zhang, et S. Duan, « InnoHAR: A Deep Neural Network for Complex Human Activity Recognition », *IEEE Access*, vol. 7, p. 9893-9902, 2019, doi: 10.1109/ACCESS.2018.2890675.
- [73] Y. Nan, N. H. Lovell, S. J. Redmond, K. Wang, K. Delbaere, et K. S. Van Schooten, « Deep Learning for Activity Recognition in Older People Using a Pocket-Worn Smartphone », *Sensors*, vol. 20, n° 24, p. 7195, déc. 2020, doi: 10.3390/s20247195.
- [74] A. Murad et J.-Y. Pyun, « Deep Recurrent Neural Networks for Human Activity Recognition », *Sensors*, vol. 17, n° 11, p. 2556, nov. 2017, doi: 10.3390/s17112556.
- [75] A. Hayat, F. Morgado-Dias, B. P. Bhuyan, et R. Tomar, « Human Activity Recognition for Elderly People Using Machine and Deep Learning Approaches », *Information*, vol. 13, n° 6, Art. n° 6, juin 2022, doi: 10.3390/info13060275.
- [76] V. Chandola, A. Banerjee, et V. Kumar, « Anomaly detection: A survey », *ACM Comput. Surv.*, vol. 41, n° 3, p. 1-58, juill. 2009, doi: 10.1145/1541880.1541882.
- [77] D. Arifoglu et H. Bouchachia, « Activity Recognition and Abnormal Behaviour Detection with Recurrent Neural Networks », *Procedia Comput. Sci.*, vol. 110, p. 86-93, déc. 2017, doi: 10.1016/j.procs.2017.06.121.
- [78] A. Lotfi, H. Bouchachia, A. Gegov, C. Langensiepen, et M. McGinnity, *Advances in Computational Intelligence Systems: Contributions Presented at the 18th UK Workshop on Computational Intelligence, September 5-7, 2018, Nottingham, UK*. Springer International Publishing, 2018.
- [79] J. Zhang, C. Wu, et Y. Wang, « Human Fall Detection Based on Body Posture Spatio-Temporal Evolution », *Sensors*, vol. 20, n° 3, p. 946, févr. 2020, doi: 10.3390/s20030946.
- [80] M. Zerkouk et B. Chikhaoui, « Long Short Term Memory Based Model for Abnormal Behavior Prediction in Elderly Persons », in *How AI Impacts Urban Living and Public*

- Health*, J. Pagán, M. Mokhtari, H. Aloulou, B. Abdulrazak, et M. F. Cabrera, Éd., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019, p. 36-45. doi: 10.1007/978-3-030-32785-9_4.
- [81] M. Zerkouk et B. Chikhaoui, « Spatio-Temporal Abnormal Behavior Prediction in Elderly Persons Using Deep Learning Models », *Sensors*, vol. 20, n° 8, p. 2359, avr. 2020, doi: 10.3390/s20082359.
- [82] N. Gupta, S. K. Gupta, R. K. Pathak, V. Jain, P. Rashidi, et J. S. Suri, « Human activity recognition in artificial intelligence framework: a narrative review », *Artif. Intell. Rev.*, vol. 55, n° 6, p. 4755-4808, août 2022, doi: 10.1007/s10462-021-10116-x.
- [83] W. Li, X. Liu, J. Liu, P. Chen, S. Wan, et X. Cui, « On Improving the accuracy with Auto-Encoder on Conjunctivitis », *Appl. Soft Comput.*, vol. 81, p. 105489, août 2019, doi: 10.1016/j.asoc.2019.105489.
- [84] R. G. Oresti Banos, « MHEALTH Dataset ». UCI Machine Learning Repository, 2014. doi: 10.24432/C5TW22.
- [85] A. Burns *et al.*, « SHIMMERTM – A Wireless Sensor Platform for Noninvasive Biomedical Research », *IEEE Sens. J.*, vol. 10, n° 9, p. 1527-1534, sept. 2010, doi: 10.1109/JSEN.2010.2045498.
- [86] S. Münzner, P. Schmidt, A. Reiss, M. Hanselmann, R. Stiefelhagen, et R. Dürichen, « CNN-based sensor fusion techniques for multimodal human activity recognition », *Proc. 2017 ACM Int. Symp. Wearable Comput.*, p. 158-165, sept. 2017, doi: 10.1145/3123021.3123046.
- [87] S. Benavidez et D. McCreight, « A Deep Learning Approach for Human Activity Recognition Project Category: Other (Time-Series Classification) », 2019. Consulté le: 6 février 2024. [En ligne]. Disponible sur: <https://www.semanticscholar.org/paper/A-Deep-Learning-Approach-for-Human-Activity-Project-Benavidez-McCreight/f18879998cf2a97317bea1275bc8c440128776ed>
- [88] H. D. Nguyen, K. P. Tran, X. Zeng, L. Koehl, et G. Tartare, « Wearable Sensor Data Based Human Activity Recognition using Machine Learning: A new approach ». arXiv, 9 mai 2019. Consulté le: 6 février 2024. [En ligne]. Disponible sur: <http://arxiv.org/abs/1905.03809>
- [89] Z. Qin, Y. Zhang, S. Meng, Z. Qin, et K.-K. R. Choo, « Imaging and fusing time series for wearable sensor-based human activity recognition », *Inf. Fusion*, vol. 53, p. 80-87, janv. 2020, doi: 10.1016/j.inffus.2019.06.014.
- [90] J. Sena, J. Barreto, C. Caetano, G. Cramer, et W. R. Schwartz, « Human activity recognition based on smartphone and wearable sensors using multiscale DCNN ensemble », *Neurocomputing*, vol. 444, p. 226-243, juill. 2021, doi: 10.1016/j.neucom.2020.04.151.
- [91] M. Hammoudeh et R. Newman, « Information extraction from sensor networks using the Watershed transform algorithm », *Inf. Fusion*, vol. 22, p. 39-49, mars 2015, doi: 10.1016/j.inffus.2013.07.001.
- [92] G. Pang, C. Shen, L. Cao, et A. van den Hengel, « Deep Learning for Anomaly Detection: A Review », *ACM Comput. Surv.*, vol. 54, n° 2, p. 1-38, mars 2022, doi: 10.1145/3439950.

- [93] N. Marir, H. Wang, G. Feng, B. Li, et M. Jia, « Distributed Abnormal Behavior Detection Approach Based on Deep Belief Network and Ensemble SVM Using Spark », *IEEE Access*, vol. 6, p. 59657-59671, 2018, doi: 10.1109/ACCESS.2018.2875045.
- [94] H. Han et U. Chong, « Neural Network Based Detection of Drowsiness with Eyes Open using AR Modelling », *IETE Tech. Rev.*, vol. 33, n° 5, p. 518-524, sept. 2016, doi: 10.1080/02564602.2015.1118362.
- [95] P. Trojovský et M. Dehghani, « Pelican Optimization Algorithm: A Novel Nature-Inspired Algorithm for Engineering Applications », *Sensors*, vol. 22, n° 3, p. 855, janv. 2022, doi: 10.3390/s22030855.
- [96] Y.-N. Xiao, X. Sun, Y. Guo, S. Li, Y. Zhang, et Y. Wang, « An Improved Gorilla Troops Optimizer Based on Lens Opposition-Based Learning and Adaptive β -Hill Climbing for Global Optimization », *Comput. Model. Eng. Sci.*, vol. 130, p. 1-36, janv. 2022, doi: 10.32604/cmescs.2022.019198.
- [97] A. M. Fathollahi-Fard, M. Hajiaghahi-Keshteli, et R. Tavakkoli-Moghaddam, « Red deer algorithm (RDA): a new nature-inspired meta-heuristic », *Soft Comput.*, vol. 24, n° 19, p. 14637-14665, oct. 2020, doi: 10.1007/s00500-020-04812-z.
- [98] B. Abdollahzadeh, F. Soleimanian Gharehchopogh, et S. Mirjalili, « Artificial gorilla troops optimizer: A new nature-inspired metaheuristic algorithm for global optimization problems », *Int. J. Intell. Syst.*, vol. 36, n° 10, p. 5887-5958, 2021, doi: 10.1002/int.22535.
- [99] N. E. Tabbakha, C. P. Ooi, et W. H. Tan, « A dataset for elderly action recognition using indoor location and activity tracking data », vol. 3, avr. 2020, doi: 10.17632/sy3kcttdtx.3.
- [100] S. Mirjalili, « Genetic Algorithm », in *Evolutionary Algorithms and Neural Networks: Theory and Applications*, S. Mirjalili, Éd., in Studies in Computational Intelligence. , Cham: Springer International Publishing, 2019, p. 43-55. doi: 10.1007/978-3-319-93025-1_4.
- [101] X. J. Luo *et al.*, « Genetic algorithm-determined deep feedforward neural network architecture for predicting electricity consumption in real buildings », *Energy AI*, vol. 2, p. 100015, nov. 2020, doi: 10.1016/j.egyai.2020.100015.
- [102] R. Hecht-nielsen, « III.3 - Theory of the Backpropagation Neural Network**Based on “nonindent” by Robert Hecht-Nielsen, which appeared in Proceedings of the International Joint Conference on Neural Networks 1, 593–611, June 1989. © 1989 IEEE. », in *Neural Networks for Perception*, H. Wechsler, Éd., Academic Press, 1992, p. 65-93. doi: 10.1016/B978-0-12-741252-8.50010-8.
- [103] D. Binu et B. S. Kariyappa, « RideNN: A New Rider Optimization Algorithm-Based Neural Network for Fault Diagnosis in Analog Circuits », *IEEE Trans. Instrum. Meas.*, vol. 68, n° 1, p. 2-26, janv. 2019, doi: 10.1109/TIM.2018.2836058.
- [104] N. Tabbakha, C. Ooi, W. H. Tan, et Y.-F. Tan, « A wearable device for machine learning based elderly’s activity tracking and indoor location system », *Bull. Electr. Eng. Inform.*, vol. 10, p. 927-939, avr. 2021, doi: 10.11591/eei.v10i2.2737.
- [105] A. Akula, A. K. Shah, et R. Ghosh, « Deep learning approach for human action recognition », *Cogn. Syst. Res.*, vol. 50, p. 146-154, août 2018, doi: 10.1016/j.cogsys.2018.04.002.