



**AMÉLIORATION DU VIEILLISSEMENT À DOMICILE PAR LA RECONNAISSANCE
D'ACTIVITÉS À L'AIDE D'UNE BROSSE À DENTS INTELLIGENTE**

par Sawadogo Wendenda Aniisah

**Mémoire présenté à l'Université du Québec à Chicoutimi en vue de l'obtention du grade de
maîtrise ès sciences (M. Sc.) en Informatique**

Québec, Canada

© Wendenda Aniisah Sawadogo, 2025

RÉSUMÉ

Le vieillissement de la population soulève des défis importants en matière de suivi de la santé et de maintien de l'autonomie à domicile. Ce mémoire étudie l'utilisation d'une brosse à dents intelligente équipée de capteurs inertiels pour la reconnaissance d'activités humaines (HAR) dans le cadre du suivi des routines de brossage dentaire. En s'appuyant sur une base de données issue d'une étude antérieure, plusieurs étapes ont été mises en œuvre : prétraitement des signaux, fusion des données capteurs, segmentation temporelle (CluMing, détection de pics) et extraction de caractéristiques. Trois familles de modèles ont été testées : les méthodes classiques (random forest), les réseaux de neurones récurrents (LSTM), et les modèles transformers appliqués à la HAR.

Les performances obtenues révèlent une précision de 98,7 % pour random forest, 99,79 % pour LSTM, et 96,17 % pour Transformer. LSTM s'est ainsi démarqué par son efficacité à capturer la dynamique des gestes de brossage, tandis que Transformer a montré une stabilité prometteuse malgré une légère baisse de performance. Une validation Leave-One-Subject-Out (LOSO) a été mise en œuvre pour évaluer la capacité des modèles à généraliser à de nouveaux utilisateurs, tandis que l'approche de One-Session-Out (OSO) s'est concentrée sur la stabilité du modèle à travers différentes sessions du même utilisateur. Les résultats mettent en lumière l'intérêt d'approches adaptatives, en intégrant potentiellement des connaissances expertes ou des mécanismes de personnalisation. Aussi, ils suggèrent que des systèmes embarqués de suivi du brossage pourraient être envisagés à moyen terme, notamment pour le soutien à domicile des populations âgées.

ABSTRACT

The aging of the population presents major challenges in terms of healthcare monitoring and maintaining autonomy at home. This research explores the use of a smart toothbrush equipped with inertial sensors as a tool for human activity recognition (HAR) to evaluate the quality and regularity of toothbrushing routines. Based on a dataset collected in a previous study, we implemented a multi-step pipeline including signal preprocessing, sensor fusion, temporal segmentation (using algorithms such as CluMing and peak detection), and discriminative feature extraction. Three model families were evaluated for the classification of brushing regions: traditional machine learning (Random Forest), recurrent neural networks (LSTM), and Transformers.

Experimental results showed that LSTM achieved the highest accuracy (99.79%), followed by Random Forest (98.7%) and Transformer (96.17%). LSTM proved particularly effective in capturing the sequential dynamics of brushing gestures, while Transformer models offered stable performance across different conditions. A Leave-One-Subject-Out (LOSO) validation was conducted to assess generalization across participants. In addition, a One-Session-Out (OSO) strategy was used to assess temporal robustness, testing the models' capacity to maintain performance over time for the same individual. The observed performance variations highlight the need for personalized or hybrid approaches, integrating expert knowledge or adaptive learning. This study contributes to the development of intelligent assistive technologies tailored to the needs of older adults, with potential applications in daily health monitoring and autonomy support.

TABLE DES MATIÈRES

RÉSUMÉ	ii
ABSTRACT	ii
TABLE DES MATIÈRES	iii
LISTE DES TABLEAUX	vii
LISTE DES FIGURES	viii
LISTE DES ABRÉVIATIONS	x
DÉDICACE	xi
REMERCIEMENTS	xii
CHAPITRE 1 INTRODUCTION	1
1.1 CONTEXTE DE LA RECHERCHE	2
1.1.2 RECONNAISSANCE D'ACTIVITES HUMAINES	3
1.1.3 TECHNOLOGIES INTELLIGENTES POUR LE VIEILLISSEMENT A DOMICILE	4
1.2 PROBLÉMATIQUE DE RECHERCHE	6
1.3 OBJECTIFS DE RECHERCHE	7
1.4 ORGANISATION DU DOCUMENT	8
CHAPITRE 2	10
REVUE DE LA LITTÉRATURE	10
2.1 RÉGIONS DE BROSSAGE	10
2.2 TECHNIQUE DE BROSSAGE BASS	11
2.3 CAPTEURS DE MOUVEMENT ET ESTIMATION D'ORIENTATION	13
2.3.1 CAPTEURS DE MOUVEMENTS	13
2.3.2 MODELISATION DE L'ORIENTATION	15
2.3.3 ESTIMATION DE L'ORIENTATION	17

2.4 TRAVAUX CONNEXES ET EXPLORATION DES METHODES D'ANALYSE DES SERIES TEMPORELLES.....	19
2.5 TRAVAUX CONNEXES SUR LA SUPERVISION DE L'ACTIVITE DE BROSSAGE DENTAIRE	22
2.5.1 SURVEILLANCE DU BROSSAGE DES DENTS A L'AIDE DE CAPTEURS FIXES SUR LES BROSSES A DENTS	22
2.5.2 SURVEILLANCE DU BROSSAGE DES DENTS A L'AIDE DE CAPTEURS PORTES AU POIGNET	25
2.5.3 SYNTHESE.....	27
CHAPITRE 3	29
ETUDE ET EXPLORATION DES DONNEES	29
3.1 PRESENTATION DES DONNEES.....	30
3.1.1 COLLECTE DES DONNEES.....	30
3.1.2 DESCRIPTION DES DONNEES	32
3.2 PRETRAITEMENT DES DONNEES	33
3.3 FILTRAGE ET FUSION DES CAPTEURS	35
3.4 DETECTION DES POINTS DE CHANGEMENT.....	37
3.4.1 ALGORITHME CLUMING.....	38
3.4.2 SEGMENTATION BASEE SUR LES PICS	41
3.4.3 LA DETECTION LOCALE DE POINTS DE CHANGEMENT	42
3.4.4 LA DETECTION DES MOMENTS D'IMMOBILITE.....	43
3.4.5 RÉSULTATS DES ALGORITHMES DE SEGMENTATION	45
3.5 EXTRACTION DES CARACTERISTIQUES.....	46
3.6 CONCLUSION	48

CHAPITRE 4	49
RÉSULTATS	49
4.1 MODELES DE CLASSIFICATION POUR LA RECONNAISSANCE D'ACTIVITES	49
4.1.1 ALGORITHMES CLASSIQUES	50
INTRODUCTION AUX ALGORITHMES CLASSIQUES	50
IMPLÉMENTATION DES ALGORITHMES CLASSIQUES	51
4.1.2 RÉSEAUX DE NEURONES	53
4.1.3 TRANSFORMER	59
4.2 MESURES DE PERFORMANCES	64
4.3 RÉSULTATS DE CLASSIFICATION	66
4.3.1 VALIDATION CROISÉE	66
4.3.2 MÉTHODE LEAVE-ONE-SUBJECT-OUT (LOSO)	71
4.3.3 MÉTHODE ONE-SESSION-OUT (OSO)	75
4.4 OPTIMISATION DES ALGORITHMES	80
4.4.1 ANALYSE PAR ABLATION DU MODÈLE TRANSFORMER	80
4.4.2 OPTIMISATION PAR GRIDSEARCHCV ET RANDOMIZEDSEARCHCV	81
4.5 EXPÉRIMENTATIONS SUR LES PROPRIÉTÉS DU DATASET	82
4.5.1 EFFET DE L'EMPLACEMENT DU CAPTEUR	83
4.5.2 IMPACT DU TYPE DE BROSSE	84
4.5.3 IMPACT DU FILTRAGE DES DONNÉES	84
4.5.4 IMPACT DE LA TAILLE DE LA FENÊTRE	85
4.5.5 IMPACT DE LA TAILLE DE LA FRÉQUENCE D'ÉCHANTILLONNAGE	86
4.5.6 REBALANCEMENT DES CLASSES	87

4.6 CONCLUSION	88
CHAPITRE 5	89
CONCLUSION ET PERSPECTIVES	89
5.1 RESUME DES CONTRIBUTIONS	89
5.2 LIMITES DE L'ETUDE	91
5.3 PERSPECTIVES	92

LISTE DES TABLEAUX

TABLEAU 1: LISTE ET DESCRIPTION DES COLONNES DE NOS DONNEES	32
TABLEAU 2: ÉVALUATION DES ALGORITHMES DE SEGMENTATION SELON DIFFERENTS CONTEXTES.....	46
TABLEAU 3: PERFORMANCES DES MODELES CLASSIQUES EN VALIDATION CROISEE	68
TABLEAU 4:PERFORMANCES DES MODELES RECURRENTS EN VALIDATION CROISEE	70
TABLEAU 5: STATISTIQUES DESCRIPTIVES DES SCORES F1 PAR MODELE EN LOSO	72
TABLEAU 6: STATISTIQUES DESCRIPTIVES DES SCORES F1 PAR MODELE EN OSO	76
TABLEAU 7: F1-SCORE DES MODELES SELON L'EMPLACEMENT DU CAPTEUR (POIGNET VS BROSSE).....	83
TABLEAU 8: F1-SCORE DES MODELES EN FONCTION DU TYPE DE BROSSE (MANUELLE VS ELECTRIQUE)	84
TABLEAU 9: EFFET DU FILTRAGE DES DONNEES ET DE L'ESTIMATION D'ORIENTATION SUR LE F1-SCORE DES MODELES	84
TABLEAU 10: F1-SCORE DES MODELES EN FONCTION DE LA TAILLE DES FENETRES TEMPORELLES	85
TABLEAU 11: F1-SCORE DES MODELES SELON LA FREQUENCE D'ECHANTILLONNAGE ...	86
TABLEAU 12: F1-SCORE DES ALGORITHMES AVEC SMOTE ET UNDERSAMPLING	87

LISTE DES FIGURES

FIGURE 1: PROCESSUS DE LA RECONNAISSANCE D'ACTIVITE HUMAINE	4
FIGURE 2: REPRESENTATION DES 16 REGIONS DENTAIRES.....	11
FIGURE 3: REPRESENTATION DE L'ORIENTATION SUR LES AXES.....	17
FIGURE 4: ÉTAPES DE LA PREPARATION DES DONNEES INERTIELLES.....	30
FIGURE 5: ÉCHANTILLON DE DONNEES DE BROSSAGE ISSUES DES CONFIGURATIONS 1 ET 2. LA CONFIGURATION 2, AVEC PAUSES, CORRESPOND A UN HOMME DROITIER UTILISANT UNE BROSSE MANUELLE ; LA CONFIGURATION 1, SANS PAUSES, A UN HOMME DROITIER UTILISANT UNE BROSSE ELECTRIQUE.	31
FIGURE 6: ÉCHANTILLON DE NOS DONNEES ISSUES D'UNE SESSION DE BROSSAGE.....	33
FIGURE 7: EXTRAIT DU TABLEAU SYNCHRONISE ISSU DE L'INTERPOLATION DES DONNEES CAPTEURS	35
FIGURE 8: SIGNAUX D'ACCELERATION BRUTS ET FILTRES PROVENANT D'UNE BROSSE MANUELLE	37
FIGURE 9: SIGNAUX D'ACCELERATION BRUTS ET FILTRES PROVENANT D'UNE BROSSE ELECTRIQUE.....	37
FIGURE 10: ÉTAPES DU TRAITEMENT PAR L'ALGORITHME CLUMING APPLIQUE AUX DONNEES DE BROSSAGE. LA PREMIERE IMAGE (A) ÉTAPE 1 DE CLUMING REPRESENTE LE REGROUPEMENT DES DONNEES EN CLUSTERS VIA K-MEANS++. LA SECONDE (B) ÉTAPE 1 DE CLUMING MONTRE LES SEGMENTS FINAUX REPRESENTANT LES REGIONS DE BROSSAGE APRES FUSION.....	40
FIGURE 11: ÉVOLUTION DE L'ACCELERATION SUR LES TROIS AXES (X, Y, Z) AVEC LES SEGMENTS DE BROSSAGE DETECTES PAR L'ALGORITHME CLUMING.....	40
FIGURE 12: DETECTION DES CHANGEMENTS D'ANGLE SUR LES AXES X, Y ET Z. LES CROIX ROUGES REPRESENTENT LES PICS DETECTES, CONSIDERES COMME DES TRANSITIONS DANS LES GESTES DE BROSSAGE.	42

FIGURE 13: ILLUSTRATION DU PROCESSUS DE SEGMENTATION DES DONNEES D'UNE SESSION DE BROSSAGE	45
FIGURE 14: ARCHITECTURE COMPLETE DU TRANSFORMER INCLUANT L'ENCODEUR ET LE DECODEUR. DANS LE CADRE DE CETTE ETUDE, SEULE LA PARTIE ENCODEUR A ETE UTILISEE.....	62
FIGURE 15: EXEMPLE ILLUSTRATIF D'UNE MATRICE DE CONFUSION POUR 4 CLASSES. LES CELLULES DE LA DIAGONALE PRINCIPALE, ANNOTEES VP (VRAIS POSITIFS), INDIQUENT LES PREDICTIONS CORRECTES POUR CHAQUE CLASSE.	66
FIGURE 16: MATRICE DE CONFUSION (%) POUR L'EVALUATION DE NOTRE MODELE DE RANDOM FOREST	69
FIGURE 17: F1-SCORE PAR CLASSE (%) POUR L'EVALUATION DE NOTRE MODELE SVC ..	70
FIGURE 18: F1-SCORE DE L'EVALUATION DU LEAVE-ONE-SUBJECT-OUT (LOSO) AVEC LE RANDOM FOREST	74
FIGURE 19: F1-SCORE DE L'EVALUATION DU LEAVE-ONE-SUBJECT-OUT (LOSO) AVEC LSTM	74
FIGURE 20: F1-SCORE DE L'EVALUATION DU LEAVE-ONE-SUBJECT-OUT (LOSO) AVEC TRANSFORMER.....	75
FIGURE 21: MATRICES DE CONFUSION POUR TROIS SESSIONS A FAIBLE PERFORMANCE	78
FIGURE 22: F1-SCORE PAR SESSION EN VALIDATION OSO, SELON LE TYPE DE BROSSE UTILISEE. CHAQUE POINT REPRESENTE UNE SESSION TEST DANS LE PROTOCOLE ONE-SESSION-OUT. LES COULEURS DISTINGUENT LES SESSIONS REALISEES AVEC UNE BROSSE ELECTRIQUE (BLEU) OU MANUELLE (ORANGE).	79
FIGURE 23: DISTRIBUTION DES F1-SCORES SELON LE SEXE DES PARTICIPANTS.....	79
FIGURE 24: DISTRIBUTION DES F1-SCORES EN FONCTION DU TYPE DE BROSSE	79
FIGURE 25: COMPARAISON DES PERFORMANCES DU TRANSFORMER SELON DIFFERENTES VARIANTES ARCHITECTURALES. LES MESURES AFFICHEES SONT LES SCORES MOYEN OBTENU SUR 5 PLIS DE VALIDATION CROISEE.....	81

LISTE DES ABRÉVIATIONS

IMU : Inertial Measurement Unit (Unité de mesure inertielle)

MARG : Magnetic, Angular Rate and Gravity (Capteurs magnéto-inertiels)

RF : Random Forest (Forêt aléatoire)

GBDT : Gradient Boosted Decision Trees (Arbres de décision à gradient boosté)

SVM : Support Vector Machines (Machines à vecteurs de support)

GNB : Gaussian Naïve Bayes (Naïve Bayes Gaussien)

KNN : K-Nearest Neighbors (K plus proches voisins)

OvR : One-vs-Rest (Un contre tous)

XGBoost : eXtreme Gradient Boosting (Boosting de gradient extrême)

LightGBM : Light Gradient Boosting Machine (Machine de boosting de gradient légère)

CPD : Change Point Detection (Détection de points de changement)

CluMing : Clustering-Based Segmentation Algorithm (Algorithme de segmentation basé sur le clustering)

AMVD : Acceleration Moving Variance Detector (DéTECTeur de variance mobile de l'accélération)

AMD : Acceleration Magnitude Detector (DéTECTeur de magnitude de l'accélération)

SLR : Segment Length Ratio (Ratio de longueur de segment)

RNN : Recurrent Neural Network (Réseau de neurones récurrent)

LSTM : Long Short-Term Memory (Mémoire à long terme)

GRU : Gated Recurrent Unit (Unité récurrente à portes)

HAR : Human Activity Recognition (Reconnaissance d'activités humaines)

LOSO : Leave-One-Subject-Out (Validation croisée avec exclusion d'un sujet)

FFT : Fast Fourier Transform (Transformée de Fourier rapide)

DTW : Dynamic Time Warping (Alignement temporel dynamique)

EKF : Extended Kalman Filter (Filtre de Kalman étendu)

DÉDICACE

À mon père Hamidou Sawadogo

Toi qui, souvent sans un mot, savais m'orienter sans imposer. Tes regards où se devinait une fierté et tes conseils glissés comme des vérités simples ont marqué mon chemin. Ce travail, c'est une empreinte de tout ce que tu m'as transmis.

À ma mère Hawa Kaboré

Toi dont l'amour a toujours été un repère, dont la force tranquille a apaisé bien des doutes et m'a appris, tant bien que mal, à avancer malgré les tempêtes. Ce mémoire, c'est un peu le tien aussi, un morceau de tout ce que tu as laissé en moi.

À mes sœurs Sawadogo Wendlassida Rabiia'h et Sawadogo Lina Janaan Wendtouin,

Les premières à croire en mes idées parfois un peu folles, celles qui ont partagé mes fous rires et soutenu mes silences fatigués. Vous êtes ce rappel, discret mais toujours là, que peu importe où la vie nous entraîne, on emporte avec nous les rires d'enfance et la certitude de ne jamais être seule. À vous, je dédie chaque page, chaque ligne, chaque mot.

Car au fond, ce chemin a toujours été le nôtre.

REMERCIEMENTS

La réalisation de ce mémoire marque l'aboutissement d'un parcours aussi exigeant qu'enrichissant. Il n'aurait pas été possible sans le soutien, l'encadrement et l'accompagnement de plusieurs personnes, auxquelles j'aimerais exprimer ma profonde gratitude.

Tout d'abord, je tiens à adresser mes remerciements les plus sincères à mon directeur de mémoire, M. Kévin Bouchard, pour son expertise précieuse, son écoute et ses conseils avisés. Son accompagnement rigoureux et son engagement tout au long de ma maîtrise ont été des éléments déterminants dans la construction et l'avancement de ce projet.

Je souhaite également exprimer ma reconnaissance envers le directeur du programme et l'équipe du Département d'informatique et de mathématique (DIM) de l'UQAC, dont l'attention et le soutien ont été d'une aide précieuse. Leur bienveillance et leur engagement ont grandement contribué à mon apprentissage et à ma progression tout au long de cette aventure académique.

À mes parents et mes sœurs, je vous adresse toute ma gratitude. Votre amour inconditionnel, votre soutien sans faille et vos encouragements constants ont été mon moteur et ma force tout au long de ce parcours. Ce mémoire vous est dédié, en témoignage de ma profonde reconnaissance.

Je remercie également mes amis et collègues, qui ont rendu cette aventure plus riche et plus stimulante. Vos échanges, votre soutien et vos encouragements ont été essentiels dans les moments de fatigue et d'incertitude.

Enfin, à toutes celles et ceux qui, de près ou de loin, ont contribué à cette aventure par un conseil, un mot d'encouragement ou une présence réconfortante : merci infiniment. Ce mémoire est le fruit de toutes ces précieuses contributions.



CHAPITRE 1

INTRODUCTION

Ce chapitre pose le cadre général de la recherche en s'intéressant aux défis liés au vieillissement à domicile et aux solutions technologiques qui permettent d'accompagner les personnes âgées dans leur quotidien. Avec l'augmentation de l'espérance de vie, il devient essentiel de développer des outils capables de surveiller discrètement les habitudes de vie et de détecter des signes précurseurs d'une perte d'autonomie. La reconnaissance d'activités humaines (HAR) fait partie des approches utilisées pour analyser les comportements et identifier des changements pouvant refléter des difficultés motrices ou cognitives.

Nous commençons par explorer les enjeux du vieillissement de la population et les moyens mis en place pour favoriser le maintien à domicile. Ensuite, nous nous penchons sur l'apport des technologies intelligentes dans ce domaine, notamment celles qui permettent de suivre les gestes du quotidien sans être intrusives. Parmi ces solutions, la brosse à dents intelligente présente un intérêt particulier, car l'hygiène bucco-dentaire est souvent un bon indicateur de l'état de santé général. L'objectif est donc d'exploiter les données inertielles issues de cette brosse pour détecter et classifier automatiquement les régions de brossage.

Enfin, nous précisons la problématique de recherche et les objectifs visés par cette étude, avant de conclure par une présentation de l'organisation du mémoire et des étapes qui seront suivies dans les chapitres à venir.

1.1 CONTEXTE DE LA RECHERCHE

Le vieillissement de la population représente un défi majeur pour de nombreuses sociétés industrialisées, y compris le Canada et le Québec. Selon Statistique Canada, la population des personnes âgées de 85 ans et plus a doublé depuis 2001 et devrait tripler d'ici 2046, en raison de l'augmentation de l'espérance de vie et de l'arrivée des premières cohortes du baby-boom dans cette tranche (Statistique Canada, 2022). Au Québec, environ 21 % de la population est maintenant âgée de 65 ans et plus, et cette proportion continue de croître, surtout dans les régions plus éloignées, comme la Gaspésie–Îles-de-la-Madeleine et le Bas-Saint-Laurent, où les personnes âgées représentent plus de 26 % de la population (Institut De La Statistique Du Québec, 2024).

Cette transformation démographique entraîne un besoin accru de solutions permettant aux personnes âgées de vieillir dans leur domicile le plus longtemps possible, ce que l'on appelle le vieillissement à domicile. Vieillir chez soi permet aux aînés de conserver leur autonomie plus longtemps et de préserver leurs liens sociaux et leur qualité de vie, tout en réduisant la pression sur les ressources du système de santé et des établissements de (Beaudoin-Benjamin et al., 2022; Savard, 2023). Cependant, ce choix de vie comporte des défis, notamment en matière de mobilité, d'accès aux soins et de sécurité. En effet, avec l'âge, les personnes âgées sont davantage sujettes à des chutes, qui représentent la première cause d'hospitalisation et de mortalité accidentelle dans cette population (Beaudoin-Benjamin et al., 2022).

Pour soutenir le vieillissement à domicile, divers aménagements et technologies peuvent être mis en place. Le Centre de référence sur l'environnement bâti et la santé (CREBS) souligne l'importance d'un environnement bâti favorable, comprenant des logements adaptés, des infrastructures de transport sécuritaires et des espaces publics accessibles. Ces aménagements sont essentiels pour maintenir la santé, la mobilité et la participation sociale des personnes âgées, contribuant ainsi à leur bien-être et à leur inclusion dans la société (Beaudoin-Benjamin et al., 2022).

Les technologies intelligentes jouent également un rôle croissant dans le soutien au vieillissement à domicile. Ces dispositifs permettent une surveillance discrète des activités

quotidiennes et aident à prévenir les risques. Par exemple, les capteurs de mouvement, les systèmes de surveillance de la santé et les rappels pour la prise de médicaments offrent aux personnes âgées un accompagnement sécuritaire tout en leur permettant de conserver une certaine autonomie. Ces technologies offrent également une tranquillité d'esprit aux familles et aux aidants, en leur permettant de rester informés de l'état de leurs proches (Institut De La Statistique Du Québec, 2024; Savard, 2023).

1.1.2 RECONNAISSANCE D'ACTIVITES HUMAINES

La reconnaissance d'activités humaines (HAR) est un domaine de recherche visant à décrire, analyser et identifier les activités ou mouvements des individus à partir de données recueillies par divers dispositifs de capteurs ([Voir page 27](#)) (Nia et al., 2023). Elle utilise des techniques de traitement des données et des algorithmes d'apprentissage automatique pour classifier les actions et identifier les comportements dans des contextes variés (Gupta et al., 2022).

Les applications de la HAR sont vastes et touchent divers secteurs. Dans le secteur de la santé, elle est utilisée pour surveiller les patients, détecter d'éventuelles chutes et suivre les progrès des activités de rééducation. Dans le domaine de la sécurité et de la surveillance, elle permet d'identifier des comportements inhabituels et contribue ainsi à renforcer les systèmes de protection dans les espaces publics et privés. Les environnements intelligents et la domotique bénéficient également de la HAR, en adaptant automatiquement les paramètres de l'habitat en fonction des activités détectées, ce qui améliore le confort et l'efficacité énergétique des logements connectés (Hussain, 2022; Vanleerberghe et al., 2017).

Dans le contexte du vieillissement à domicile, la HAR revêt une importance particulière. Avec l'augmentation de la population âgée, il devient essentiel de pouvoir suivre les routines quotidiennes des personnes âgées sans empiéter sur leur vie privée. La HAR permet ainsi de surveiller l'autonomie en fournissant des informations précieuses sur l'état fonctionnel des personnes âgées en détectant des changements dans les habitudes, comme une baisse de l'activité ou des anomalies dans les

routines. Les variations dans les comportements peuvent également être des indicateurs précoce d'un déclin cognitif ou physique, facilitant une intervention rapide des aidants ou des professionnels de santé. En fournissant une surveillance discrète et continue des activités, la HAR contribue à préserver un équilibre entre autonomie et sécurité, un facteur clé pour le bien-être des personnes âgées vivant (Hussain, 2022; Vanleerberghe et al., 2017).

La reconnaissance d'activités humaines repose sur un processus en plusieurs étapes comme représenté sur la Figure 1.

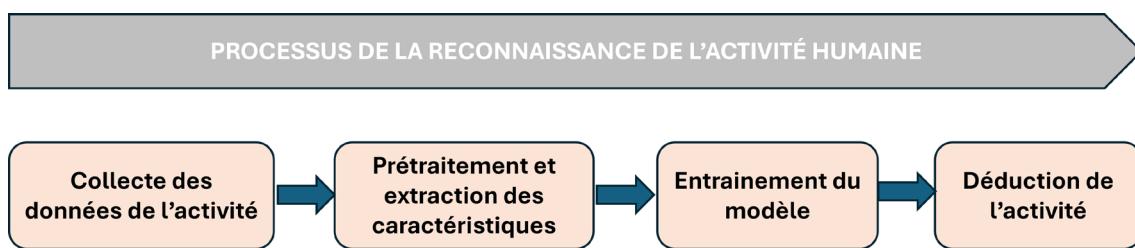


Figure 1: Processus de la reconnaissance d'activité humaine

©Sawadogo Wendenda Aniisah

1.1.3 TECHNOLOGIES INTELLIGENTES POUR LE VIEILLISSEMENT A DOMICILE

Avec l'augmentation de la population âgée, le vieillissement à domicile devient une priorité pour les systèmes de santé. Les personnes âgées préfèrent rester dans leur propre environnement aussi longtemps que possible (Vanleerberghe et al., 2017; W. H. O. Centre for Health Development, 2004), mais cela nécessite des solutions de suivi non intrusives pour assurer leur sécurité et leur bien-être. Les technologies intelligentes, en permettant de surveiller discrètement les activités quotidiennes et en détectant des changements comportementaux, jouent un rôle crucial dans l'autonomie des personnes âgées tout en soutenant les aidants et les professionnels de santé.

Parmi les dispositifs les plus couramment utilisés, on retrouve les montres connectées et autres objets portables, qui permettent de suivre des indicateurs de santé tels que le rythme cardiaque, la mobilité ou la qualité du sommeil. Les capteurs de mouvement, souvent installés dans le domicile, sont quant à eux capables de détecter les déplacements, les chutes ou une absence prolongée

d'activité, offrant ainsi une couche supplémentaire de sécurité. Enfin, les appareils électroménagers intelligents peuvent contribuer à la gestion de la prise de médicaments ou à la surveillance de l'utilisation de certains équipements essentiels, réduisant ainsi les risques d'accidents (Aggar et al., 2023; Kim et al., 2022).

Parmi les technologies intelligentes conçues pour le vieillissement à domicile, le suivi du brossage des dents reste encore peu étudié. Toutefois, l'hygiène bucco-dentaire est essentielle, surtout chez les personnes âgées qui, avec l'âge, peuvent négliger cette routine à cause de troubles cognitifs ou physiques (Bélanger et al., 2016). Dans ce contexte, la brosse à dents intelligente se distingue comme une application spécifique et innovante des technologies HAR. Cet appareil est équipé de capteurs qui capturent les données liées aux mouvements, à l'orientation et à la pression appliquée lors du brossage. Grâce à ces capteurs, la brosse à dents peut identifier les 16 régions de la bouche et suivre la durée et la fréquence du brossage dans chaque région.

En analysant les habitudes de brossage, il devient possible d'observer la régularité et la qualité des gestes. Des variations dans ces routines, comme une baisse de coordination ou de constance, peuvent indiquer des difficultés motrices ou cognitives. Ces technologies offrent ainsi un aperçu utile pour surveiller de manière proactive l'état de santé général des (Hussain, 2022).

L'utilisation de la brosse à dents intelligente pour le suivi du vieillissement à domicile offre des avantages qui vont bien au-delà de la simple supervision et classification des zones de brossage. Bien que cela dépasse le cadre de notre étude, il est pertinent de souligner son potentiel en tant qu'outil d'accompagnement des personnes âgées.

En effet, la brosse à dents intelligente permet un suivi précis des soins personnels en surveillant de manière autonome les habitudes de soins bucco-dentaires, et fournit ainsi des indices indirects sur l'état de santé général de l'utilisateur. Elle joue également un rôle dans la détection des changements comportementaux : une diminution de la fréquence ou de la qualité du brossage peut révéler le début de troubles cognitifs ou une perte de capacités physiques, alertant ainsi les proches ou les professionnels de santé sur des signes précoce de (Vanleerberghe et al., 2017). De plus, les

données collectées facilitent la prise de décision pour les aidants et les professionnels de santé, en leur permettant de réagir rapidement à des changements significatifs dans les habitudes d'hygiène, et de planifier des interventions préventives (American Dental Association, 2022) et ciblées si nécessaire (Hussain, 2022).

1.2 PROBLÉMATIQUE DE RECHERCHE

Face au vieillissement de la population, les personnes âgées expriment une préférence croissante pour vieillir chez elles afin de conserver leur autonomie et de rester dans leur environnement familial. Les technologies de maison intelligente et de reconnaissance d'activités humaines (HAR) se développent pour répondre à ce besoin en assurant un suivi discret des activités quotidiennes, réduisant ainsi la pression sur le système de santé et offrant une tranquillité d'esprit aux familles et aux aidants. Cependant, bien que la surveillance des activités de vie quotidienne (AVQ) soit cruciale pour détecter les signes de déclin, certaines activités, comme le brossage des dents, sont souvent négligées.

Dans ce contexte, la brosse à dents intelligente constitue une solution innovante et non intrusive pour évaluer l'hygiène personnelle, en capturant des données sur les mouvements de brossage via une unité centrale inertielle. Ces données peuvent révéler des changements dans la fréquence, la durée, et la qualité du brossage, permettant ainsi de déceler des indices précoce de troubles physiques ou cognitifs. Toutefois, le défi consiste à développer un modèle de HAR capable de détecter et de classifier automatiquement les zones de brossage à partir des données inertielles, afin de transformer ces observations en indicateurs cliniquement exploitables.

La problématique de cette recherche peut alors être formulée comme suit. Comment développer un système de reconnaissance d'activités humaines (HAR) basé sur les données inertielles d'une brosse à dents intelligente, capable de détecter et de classer automatiquement les zones de brossage, afin de fournir des indicateurs fiables et non intrusifs pour le suivi de l'autonomie et de la santé des personnes âgées vivant à domicile ?

Cette problématique met en avant les défis techniques de la HAR appliquée aux soins personnels dans un contexte de vieillissement à domicile, tout en soulignant la pertinence clinique et l'utilité préventive de la brosse à dents intelligente pour le suivi de la santé et de l'autonomie des utilisateurs.

1.3 OBJECTIFS DE RECHERCHE

L'objectif principal de ce mémoire est de concevoir un système de reconnaissance d'activités humaines (HAR) permettant de détecter et de classifier automatiquement les zones de brossage dentaire en utilisant les données inertielles capturées par une brosse à dents intelligente. Cette recherche s'inscrit dans une démarche de soutien au vieillissement à domicile, avec pour ambition d'améliorer la qualité de vie et l'autonomie des personnes âgées, grâce à des indicateurs non invasifs de leur hygiène personnelle.

Le processus de reconnaissance d'activités humaines repose sur quatre phases essentielles. Chacune de ces phases présente des défis spécifiques à relever pour assurer une détection précise et fiable des activités. Partant de là, nous pouvons diviser l'objectif principal en des objectifs plus spécifiques qui sont les suivants :

Objectif 1 : Mettre en place un cadre de prétraitement : Établir des procédures de prétraitement pour nettoyer les données des capteurs, réduire le bruit, garantissant des signaux cohérents pour l'analyse.

Objectif 2 : Tester et proposer des méthodes de détection de points de changement : Examiner plusieurs méthodes de détection de points de changement et analyser l'impact de chaque méthode sur la performance de la classification des régions de brossage.

Objectif 3 : Explorer et comparer différentes approches de classification : Tester des modèles classiques d'apprentissage automatique et des réseaux neuronaux, pour identifier la meilleure approche de classification des régions de brossage dentaire.

Objectif 4 : Évaluer et comparer les performances des modèles : Mesurer les performances des modèles de reconnaissance d'activité avec des indicateurs tels que la justesse, la précision, le rappel, et le score F1 ; comparer les modèles classiques avec les réseaux de neurones, et évaluer la généralisation à l'aide du test Leave-One-Subject-Out et One-Session-Out pour garantir la robustesse du modèle.

1.4 ORGANISATION DU DOCUMENT

Afin de donner vie aux objectifs précédemment définis, ce mémoire déploie une structure en cinq chapitres, guidant le lecteur depuis le contexte général jusqu'aux perspectives ouvertes par cette recherche.

Le premier chapitre introduit le contexte de la recherche en présentant les enjeux du vieillissement à domicile, les technologies de reconnaissance d'activités humaines (HAR) et l'importance d'une surveillance non intrusive de l'hygiène personnelle à travers des dispositifs intelligents comme la brosse à dents. Ce chapitre expose également la problématique et les objectifs poursuivis par cette étude.

Le second chapitre est une revue de la littérature qui examine les travaux existants sur les techniques de brossage, les capteurs de mouvement, l'estimation de l'orientation, ainsi que les approches de surveillance de l'activité de brossage. L'objectif est de contextualiser l'usage des capteurs inertielles pour la reconnaissance d'activités dans le cadre du vieillissement à domicile et de dégager les solutions les plus prometteuses pour le suivi des activités de la vie quotidienne.

Le troisième chapitre est consacré à l'exploration et l'étude de notre jeu de données. Il présente les étapes de prétraitement des données, les techniques de segmentation des données de capteurs pour la détection des points de changement, et les méthodes d'extraction de caractéristiques.

Le quatrième chapitre présente les résultats expérimentaux et les analyses en fonction des mesures d'évaluation. Il introduit ensuite les différents modèles d'apprentissage automatique et

d'apprentissage profond employés pour la classification des régions de brossage, en justifiant le choix de chaque approche et en comparant les configurations de déploiement des capteurs. Une évaluation approfondie utilisant les tests Leave-One-Subject-Out et One-Session-Out pour valider la solidité des résultats a également été réalisée. Ce chapitre discute des performances obtenues et de la pertinence de chaque méthode pour la reconnaissance des zones de brossage dans le cadre du vieillissement à domicile.

Enfin, le cinquième chapitre propose une conclusion synthétisant les contributions de cette recherche. Il identifie les limites rencontrées et formule des recommandations pour les travaux futurs, en explorant les perspectives de développement de technologies HAR adaptées aux habitats intelligents pour favoriser l'autonomie des personnes âgées.

CHAPITRE 2

REVUE DE LA LITTÉRATURE

Ce chapitre présente une revue de la littérature sur les différentes approches et technologies utilisées pour la reconnaissance des activités humaines (HAR) appliquées au suivi du brossage dentaire. L'objectif est d'examiner les méthodologies existantes, de comprendre comment les capteurs inertiels et les techniques de segmentation contribuent à l'analyse des mouvements, et d'identifier les limites actuelles dans les solutions proposées. Nous abordons d'abord les notions fondamentales, telles que la division des régions de brossage et la technique Bass, qui constituent une base pour l'étude. Ensuite, les capteurs de mouvement et les méthodes d'estimation de l'orientation sont explorés. Enfin, les travaux connexes sont analysés afin de comparer les approches utilisant des capteurs fixés sur les brosses à dents avec celles exploitant des dispositifs portés au poignet. Cette revue offre un cadre de référence pour les expérimentations et aide à identifier des opportunités d'amélioration pour une surveillance plus efficace et précise du brossage dentaire.

2.1 RÉGIONS DE BROSSAGE

Les régions de brossage dentaire sont divisées en fonction de la structure de la bouche et des surfaces dentaires spécifiques à nettoyer. Ces divisions permettent une approche méthodique et complète pour couvrir toutes les parties de la cavité buccale. Plusieurs travaux s'appuient sur une segmentation standardisée qui divise la bouche en 16 régions distinctes, comme on peut le voir sur la Figure 2 et chacune correspondant à des surfaces spécifiques des dents.

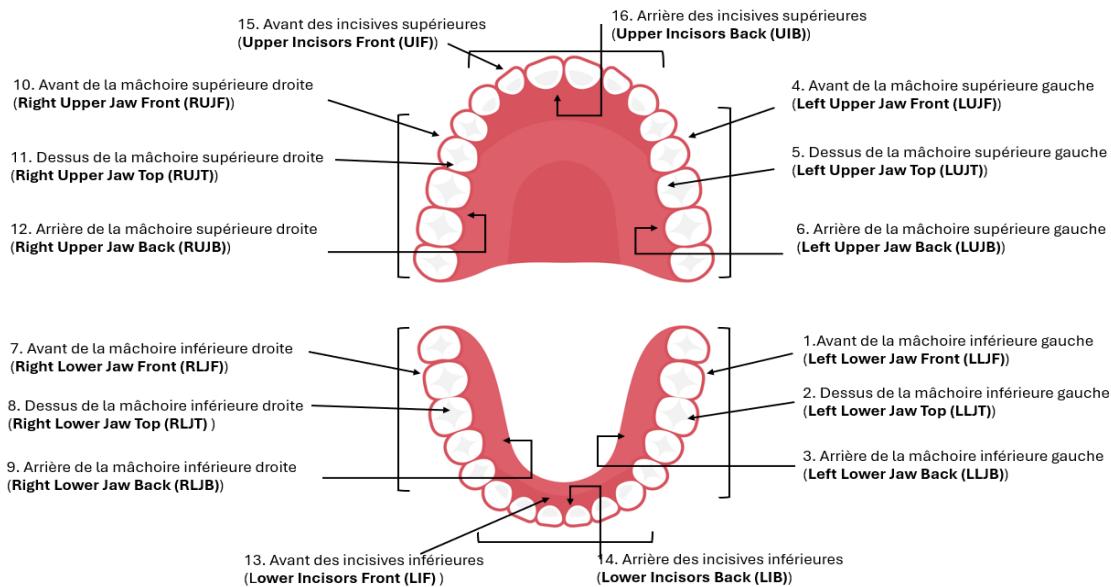


Figure 2: Représentation des 16 régions dentaires

©Sawadogo Wendenda Aniisah

2.2 TECHNIQUE DE BROSSAGE BASS

Le brossage dentaire est le principal moyen mécanique pour retirer des plaques des dents, contribuant ainsi à la prévention des maladies bucco-dentaires, telles que la carie dentaire et la gingivite (Canadian Dental Hygienists Association, 2022). Malheureusement, nombreuses sont les personnes qui ne sont pas informées sur les techniques adéquates pour un brossage efficace. Chez les personnes du troisième âge, cette situation est exacerbée par les limitations physiques qui peuvent compliquer leur capacité à effectuer un brossage optimal (Menon et al., 2024).

Les défis de l'amélioration de la santé bucco-dentaire, et plus spécifiquement du brossage dentaire, sont nombreux et complexes. Ils incluent plusieurs recommandations, dont certaines sont difficiles à respecter régulièrement. En effet, un brossage efficace des dents doit tenir compte de plusieurs paramètres, notamment la méthode de brossage, le temps consacré, la force appliquée et la fréquence de brossage. Cependant, ces consignes ne sont pas toujours suivies dans la pratique quotidienne. Chez les personnes âgées, des facteurs supplémentaires, tels que la dégradation des capacités physiques, peuvent également influencer la qualité du brossage. La surveillance de

l'activité de brossage dentaire peut donc avoir un double objectif : d'une part, améliorer la qualité du brossage des sujets et, d'autre part, détecter une dégradation des conditions de santé à travers une baisse de qualité du brossage.

Dans ce contexte, la technique Bass, largement reconnue comme une méthode de référence, offre des avantages cliniques et pratiques importants pour assurer un brossage efficace et prévenir les maladies bucco-dentaires. Elle est recommandée par l'Association Dentaire Américaine (ADA) et constitue la méthode de référence dans de nombreuses recherches liées à la supervision du brossage (American Dental Association, 2022; Canadian Dental Hygienists Association, 2022).

La technique Bass repose sur un positionnement spécifique de la brosse à dents. Les poils doivent être placés à un angle de 45 degrés par rapport à la ligne des gencives, permettant un accès optimal aux poches gingivales où s'accumulent fréquemment les bactéries. L'utilisateur doit effectuer de petits mouvements vibratoires et circulaires, avec un léger va-et-vient, en appliquant une pression modérée pour éviter d'endommager les gencives. Chaque série de dents doit être brossée pendant environ 20 secondes avant de passer à la suivante, garantissant une couverture complète de toutes les surfaces dentaires. La méthode inclut également un brossage doux du palais et de la langue afin d'éliminer les bactéries responsables de la mauvaise haleine et d'améliorer l'hygiène bucco-dentaire (Menon et al., 2024; Pramesti et al., 2023).

Plusieurs études ont utilisé la technique Bass comme cadre pour évaluer et surveiller l'efficacité du (Duffau & Lostal, 2021; Janakiram et al., 2020). Une segmentation standardisée des dents en 16 régions est souvent employée pour analyser si toutes les zones ont été correctement nettoyées, conformément aux recommandations de la méthode Bass. Certaines recherches ont cependant proposé une segmentation en 15 régions, ajustant légèrement les critères pour s'adapter à des besoins spécifiques d'analyse.

La technique Bass est particulièrement efficace pour réduire la plaque dentaire, en permettant un nettoyage en profondeur de la ligne gingivale, et pour limiter l'accumulation bactérienne, un facteur clé dans le développement des caries et des maladies parodontales. En outre, elle aide à prévenir

les maladies gingivales grâce à des mouvements doux et ciblés qui réduisent le risque de gingivite et protègent contre la parodontite. Enfin, la méthode est polyvalente et peut être utilisée avec une brosse manuelle ou électrique, rendant son application universelle et (American Dental Association, 2022; Canadian Dental Hygienists Association, 2022).

Malgré ses nombreux avantages, la technique Bass présente quelques défis. La précision des mouvements et l'angle spécifique à 45 degrés demandent une formation adéquate pour être bien maîtrisés, ce qui peut poser problème aux enfants ou aux personnes âgées. Par ailleurs, une pression excessive ou une mauvaise application des mouvements vibratoires peut entraîner des irritations gingivales ou des abrasions dentaires. Cela souligne l'importance d'un apprentissage rigoureux et, si possible, d'un suivi supervisé pour maximiser les bénéfices de cette méthode (American Dental Association, 2022; Canadian Dental Hygienists Association, 2022).

2.3 CAPTEURS DE MOUVEMENT ET ESTIMATION D'ORIENTATION

Dans les systèmes de reconnaissance d'activités humaines (HAR), la compréhension fine des mouvements passe par l'analyse des données fournies par les capteurs inertiels. Ceux-ci permettent non seulement de mesurer les accélérations et les rotations, mais aussi d'estimer l'orientation des objets ou parties du corps dans l'espace. Cette section présente les principaux types de capteurs utilisés, les méthodes de modélisation de l'orientation, ainsi que les techniques de fusion permettant une estimation fiable et stable des gestes, notamment dans le contexte de l'utilisation d'une brosse à dents intelligente.

2.3.1 CAPTEURS DE MOUVEMENTS

Les capteurs de mouvement jouent un rôle central dans la reconnaissance d'activités humaines (HAR) et les environnements intelligents tels que les maisons connectées. Ils permettent de relier le monde physique au digital en fournissant des données en temps réel sur les mouvements et comportements des utilisateurs. Grâce à leur disponibilité sous diverses formes, tailles et

capacités, ces capteurs sont intégrés dans des systèmes variés pour détecter, analyser et répondre aux activités (Bian et al., 2022).

Les systèmes de détection de mouvement reposent généralement sur l'utilisation de capteurs inertIELS, comprenant un ou plusieurs des composants (Faisal et al., 2019; Lopez-Nava & Munoz-Melendez, 2016; Mukhopadhyay, 2014) : Les accéléromètres, gyroscopes et magnétomètres. Les accéléromètres triaxiaux mesurent l'accélération sur les trois axes (X, Y, Z). Ils capturent les forces linéaires, qu'elles soient dues au mouvement ou à la gravité, permettant ainsi de mesurer des activités telles que marcher, courir ou sauter. Ces capteurs permettent également de calculer la vitesse et le déplacement en intégrant les données d'accélération. En outre, ils détectent les changements de posture ou les vibrations, ce qui en fait un outil polyvalent pour les appareils électroniques et les systèmes structurels.

Les gyroscopes triaxiaux, quant à eux, mesurent la vitesse angulaire sur les trois axes, détectant ainsi les rotations et les changements d'orientation. Ils permettent de suivre des mouvements complexes tels que ceux du poignet ou de la tête. Ils jouent également un rôle clé dans les systèmes de navigation inertielle, améliorant la précision de la localisation et du suivi dans les environnements où les signaux GPS sont limités.

Les magnétomètres triaxiaux complètent les données des accéléromètres et des gyroscopes en mesurant le champ magnétique terrestre. Ces capteurs fournissent une orientation absolue, indispensable pour estimer la position d'un objet dans l'espace. Ils sont couramment utilisés pour stabiliser les systèmes d'orientation dans des environnements dynamiques ou lorsque des références externes sont indisponibles.

Les unités de mesure inertielle (IMU) combinent ces trois types de capteurs pour offrir des mesures complètes en six ou neuf axes. Grâce à leur petite taille, leur coût abordable et leurs capacités en temps réel, les IMU sont devenues populaires dans de nombreuses applications. Elles peuvent être fixées à des objets ou des parties du corps pour surveiller des mouvements spécifiques. Par exemple, elles sont utilisées dans les montres intelligentes pour suivre l'activité physique, ou

encore dans les aspirateurs robots pour optimiser leur navigation. Ces capteurs offrent une solution polyvalente et efficace pour une variété de scénarios technologiques et (Hussain, 2022).

2.3.2 MODELISATION DE L'ORIENTATION

La modélisation de l'orientation consiste à représenter mathématiquement la position angulaire d'un objet dans un espace tridimensionnel (3D). Cette représentation est essentielle pour les systèmes utilisant des capteurs inertIELS, tels que les brosses à dents intelligentes ou les robots, car elle permet de convertir les données brutes des capteurs en une compréhension utile et exploitable de (Kok et al., 2017). (Lee et al., 2011), l'orientation d'un objet est définie par trois degrés de liberté (roulis, tangage, lacet), qui correspondent respectivement aux rotations autour des axes X, Y et Z.

La modélisation de l'orientation repose sur trois principales méthodes mathématiques (Kok et al., 2017; Kubelka & Reinstein, 2012). La première méthode utilise les angles d'Euler, qui décrivent l'orientation comme une séquence de trois rotations successives autour des axes principaux d'un référentiel. Ces rotations incluent le lacet(ψ), correspondant à une rotation horizontale autour de l'axe Z (comme le nez d'un avion déviant à gauche ou à droite), le tangage (θ), correspondant à une rotation verticale avant-arrière autour de l'axe Y (comme une inclinaison vers le haut ou le bas), et le roulis (ϕ), correspondant à une bascule latérale autour de l'axe X (comme le balancement d'un avion de gauche à droite). Cette méthode est largement utilisée pour représenter l'orientation dans l'espace tridimensionnel en raison de sa simplicité, bien que le problème de *gimbal lock* (une perte de degré de liberté due à l'alignement de deux axes de rotation, empêchant certains mouvements) puisse limiter son utilisation dans certains cas.

La deuxième méthode repose sur les matrices de rotation, qui sont des matrices 3x3 utilisées pour transformer un vecteur d'un référentiel à un autre. Ces matrices offrent une représentation fiable et sans ambiguïté de l'orientation, ce qui les rend précieuses pour de nombreuses applications nécessitant une précision élevée. Toutefois, elles sont coûteuses en termes de stockage et de

calculs, ce qui peut limiter leur utilisation dans des systèmes embarqués ou des applications nécessitant des ressources limitées.

Enfin, les quaternions constituent une extension des nombres complexes utilisée pour représenter des rotations dans un espace tridimensionnel (Essalat, 2023). Un quaternion est composé de quatre éléments :

- Un composant réel (q_1) et
- Trois composants imaginaires (q_2, q_3, q_4) .

Il peut décrire une rotation arbitraire autour d'un axe quelconque comme on peut le voir sur la Figure 3, ce qui en fait une méthode compacte et efficace pour modéliser les orientations. Les quaternions sont particulièrement adaptés aux systèmes nécessitant une grande précision et une gestion facile des interpolations, comme en robotique, dans l'animation 3D, ou encore dans les applications de suivi des mouvements.

Ces trois méthodes offrent des solutions variées pour la modélisation de l'orientation, chacune ayant ses avantages et ses limitations en fonction des exigences spécifiques des systèmes ou des environnements dans lesquels elles sont déployées.

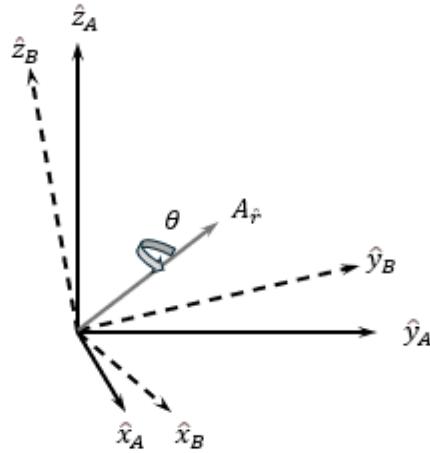


Figure 3: Représentation de l'orientation sur les axes

©Sawadogo Wendenda Aniisah

2.3.3 ESTIMATION DE L'ORIENTATION

L'estimation de l'orientation à l'aide des capteurs inertiels repose sur une combinaison des données fournies par le gyroscope, l'accéléromètre et le magnétomètre, un processus appelé fusion de capteurs. Chaque capteur contribue à l'estimation de manière complémentaire.

Le gyroscope mesure les vitesses angulaires (ω_x , ω_y , ω_z) qui, par intégration dans le temps, permettent de calculer les angles d'orientation relatifs, comme illustré dans l'équation (2.1).

$$\theta(t) = \int_{\{t_0\}}^t \omega(t), dt \quad (2.1)$$

Cette méthode offre une estimation rapide et précise des changements d'orientation sur de courtes périodes. Cependant, l'intégration des vitesses angulaires accumule des erreurs mineures au fil du temps, ce qui provoque une dérive gyroscopique, rendant nécessaire l'utilisation d'autres capteurs pour corriger ces erreurs (Essalat, 2023; Kok et al., 2017).

L'accéléromètre, quant à lui, mesure l'accélération totale (a_x, a_y, a_z) incluant la gravité. En conditions statiques, il fournit une référence absolue pour l'axe (z) du référentiel mondial, permettant de calculer les angles d'inclinaison, à savoir le roulis (ϕ) et le tangage (θ), à l'aide des équations conformément aux équations (2.2) et (2.3) :

$$\phi = \arctan\left(\frac{a_z}{a_y}\right) \quad (2.2)$$

et

$$\theta = \arctan\left(\frac{-a_x}{\sqrt{a_y^2 + a_z^2}}\right) \quad (2.3)$$

Toutefois, l'accéléromètre ne permet pas de déterminer l'angle de lacet (ϕ) ou de cap et ses mesures sont perturbées par les accélérations dynamiques, comme les mouvements rapides ou les vibrations.

Le magnétomètre complète ces deux capteurs en mesurant le champ magnétique terrestre (B_x, B_y, B_z), ce qui fournit une référence absolue pour l'axe (x) du référentiel mondial. Il permet de calculer l'angle de lacet (ϕ) ou de cap à l'aide de la relation décrite dans l'équation (2.4).

$$\psi = \arctan\left(\frac{B_y}{B_x}\right) \quad (2.4)$$

Cependant, les mesures du magnétomètre sont sensibles aux interférences magnétiques provenant d'objets métalliques ou de champs électromagnétiques, nécessitant une calibration régulière pour garantir des mesures précises.

Pour combiner les données de ces capteurs et surmonter leurs limitations respectives, des filtres sont utilisés. Le filtre complémentaire, par exemple, associe les données dynamiques du gyroscope et les références absolues de l'accéléromètre et du magnétomètre en pondérant leurs contributions respectives. Cette approche simple et efficace est adaptée aux systèmes embarqués avec des ressources limitées. Pour des applications nécessitant une précision accrue, le filtre de

Kalman étendu (EKF)(S. A. Ludwig & K. D. Burnham, 2018) propose une approche probabiliste, modélisant l'état dynamique et corrigéant les prédictions à l'aide des observations des capteurs. Cependant, cette méthode exige une calibration minutieuse des paramètres et une puissance de calcul plus élevée. Les filtres Madgwick et Mahony(Jansen, 2024; S. A. Ludwig & K. D. Burnham, 2018), conçus spécifiquement pour les capteurs inertIELS et magnéto-inertIELS à faible coût, offrent une alternative rapide et fiable, bien que dépendant fortement de la qualité de la calibration des (Essalat, 2023; Kok et al., 2017).

L'estimation de l'orientation dans des applications comme une brosse à dents intelligente implique de combiner les contributions de ces capteurs. Le gyroscope permet de suivre les changements rapides d'orientation, l'accéléromètre fournit une référence pour détecter l'inclinaison, et le magnétomètre aide à distinguer les orientations absolues. En fusionnant ces données, il est possible d'obtenir une estimation précise et stable, essentielle pour détecter les régions brossées et analyser les mouvements des (Essalat, 2023; Kok et al., 2017).

2.4 TRAVAUX CONNEXES ET EXPLORATION DES METHODES D'ANALYSE DES SERIES TEMPORELLES

L'étude des séries temporelles est essentielle dans de nombreux domaines tels que la santé, la finance, et l'industrie. Avec l'essor des algorithmes modernes d'apprentissage automatique et d'apprentissage profond, de nouvelles méthodes ont été développées pour améliorer la classification et le clustering des séries temporelles, notamment multivariées.

En 2022, (Weiss et al., 2022) explore l'efficacité de différentes techniques de prétraitement des séries temporelles pour améliorer les performances des modèles d'apprentissage automatique tout en minimisant le coût computationnel. Les auteurs se concentrent sur des méthodes simples et efficaces, regroupées en trois catégories : la réduction de la dimensionnalité temporelle, l'extraction de résumés statistiques et les transformations fréquentielles. Dans la première approche, des techniques comme l'échantillonnage (sous-échantillonnage) et l'agrégation sont appliquées pour

réduire la taille des données tout en conservant les tendances principales. La deuxième catégorie extrait des caractéristiques globales telles que la moyenne, l'écart-type, les valeurs maximales et minimales afin de résumer les séries temporelles sous une forme plus compacte. Enfin, la troisième méthode utilise des transformations fréquentielles, notamment la transformée de Fourier rapide (FFT), pour analyser les données dans le domaine fréquentiel et capturer les motifs périodiques. Les expériences menées sur divers jeux de données multivariées démontrent que les résumés statistiques et l'échantillonnage offrent des résultats particulièrement satisfaisants, en améliorant les performances des modèles tout en restant peu coûteux en calcul. Les modèles d'apprentissage classique, tels que random forest et SVM, atteignent des précisions supérieures à 95 % sur certaines tâches de classification grâce à ces techniques de prétraitement. Bien que les transformations fréquentielles soient plus performantes pour des séries présentant des motifs cycliques, elles présentent un coût computationnel plus élevé. Ainsi, les auteurs concluent que des méthodes simples comme l'agrégation et l'échantillonnage constituent des solutions stables et optimales pour le prétraitement des séries temporelles, en particulier dans des contextes où les ressources computationnelles sont limitées. Cette étude met en lumière l'importance d'un prétraitement adéquat pour exploiter efficacement les séries temporelles dans des applications d'apprentissage automatique en aval.

En 2021, (D. Guijo-Rubio et al., 2021) proposent une approche innovante pour le clustering des séries temporelles multivariées en se basant sur l'analyse des segments typologiques. Le but de cette étude est d'améliorer l'identification des motifs locaux dans les séries temporelles en segmentant celles-ci et en utilisant les caractéristiques extraites pour regrouper les séries selon leurs comportements segmentaires. Pour atteindre cet objectif, les auteurs appliquent une segmentation polynomiale qui divise les séries temporelles en segments de longueurs variables. Chaque segment est représenté par un polynôme de faible degré, permettant de capturer sa forme générale. Ensuite, des caractéristiques descriptives telles que l'inclinaison, la longueur et l'amplitude sont extraites pour décrire précisément les segments. Une fois les segments caractérisés, un profil segmentaire est créé pour chaque série temporelle, constitué des caractéristiques de ses segments. Enfin, ces profils sont utilisés comme entrée pour des algorithmes de clustering classiques comme k-means, afin de

regrouper les séries en fonction des similarités observées dans leurs comportements locaux. Les résultats expérimentaux obtenus sur des jeux de données publics montrent que cette méthode est capable d'identifier des groupes cohérents de séries temporelles en tenant compte des variations locales. Comparée aux méthodes traditionnelles telles que le DTW (Dynamic Time Warping), l'approche proposée présente une meilleure interprétabilité des résultats tout en étant capable de capturer efficacement les changements structurels locaux. Cette méthode offre ainsi une solution fiable pour le clustering des séries temporelles tout en facilitant leur analyse et leur compréhension.

En 2023, (Dhariyal et al., 2023) évalue la performance des algorithmes modernes de classification des séries temporelles en les comparant à des méthodes plus simples et classiques. Les auteurs observent que, malgré l'essor des approches complexes basées sur l'apprentissage profond (comme les réseaux de neurones convolutifs CNN, ResNet, ou InceptionTime), peu d'études testent rigoureusement si ces modèles surpassent réellement des méthodes de référence. Pour répondre à cette question, ils utilisent comme méthode de base K-Nearest Neighbors (k-NN) combiné à la Dynamic Time Warping (DTW), une mesure de distance classique utilisée pour comparer des séries temporelles. Leur expérimentation s'appuie sur l'ensemble de données UCR/UEA Time Series Archive, qui est une référence standard pour la classification des séries temporelles, comprenant des données univariées et multivariées. Les résultats montrent que le k-NN avec DTW rivalise, voire surpassé, les modèles modernes dans plusieurs cas, tout en étant beaucoup plus simples et transparents à implémenter. Les modèles tabulaires classiques (Random Forest, Logistic Regression, Ridge Regression, Linear Discriminant Analysis) surpassent par ailleurs les classificateurs modernes comme ROCKET (Random Convolutional Kernel Transform) sur environ 19 % des ensembles de données univariés et 28 % des ensembles multivariés. De plus, dans 50 % des ensembles de données, les modèles tabulaires atteignent une précision à moins de 10 points de pourcentage des méthodes modernes de pointe, tout en étant beaucoup plus rapides et plus faciles à déployer. Les auteurs soulignent que les méthodes modernes, bien que souvent plus complexes et gourmandes en ressources, n'apportent pas toujours des gains de performance significatifs qui justifieraient leur utilisation.

2.5 TRAVAUX CONNEXES SUR LA SUPERVISION DE L'ACTIVITE DE BROSSAGE DENTAIRE

L'analyse du brossage dentaire à l'aide de capteurs est une approche prometteuse pour mieux comprendre les habitudes d'hygiène et détecter d'éventuelles lacunes dans la couverture des différentes zones de la bouche. Plusieurs études ont exploré cette problématique en utilisant soit des capteurs intégrés directement sur les brosses à dents, soit des dispositifs portés au poignet pour analyser les mouvements associés au brossage. Cette section passe en revue ces travaux, en mettant en évidence les choix méthodologiques, les performances obtenues et les défis rencontrés. Ces recherches apportent un éclairage précieux sur les solutions existantes et permettent d'identifier les pistes d'améliorations pour une reconnaissance plus précise et fiable des régions de brossage.

2.5.1 SURVEILLANCE DU BROSSAGE DES DENTS A L'AIDE DE CAPTEURS FIXES SUR LES BROSSES A DENTS

Dans le domaine de la supervision du brossage dentaire basé sur des brosses à dents intelligentes ou modifiées, (Hussain et al., 2021) proposent une solution basée sur une unité de mesure inertielle (IMU) pour surveiller l'activité de brossage. Des capteurs détachables sont fixés au manche d'une brosse à dents afin de reconnaître les mouvements spécifiques aux différentes régions de la bouche. Les chercheurs ont appliqué une division de la bouche en 16 régions en appliquant les recommandations de brossage de la technique Bass. Leur méthode a permis d'atteindre une justesse globale de 97,15 % grâce à un modèle de random forest. Cette précision dépasse celle des approches utilisant des capteurs portés au poignet, grâce à une meilleure capture des mouvements directement au niveau de la brosse. L'étude a également exploré plusieurs aspects techniques. Par exemple, l'utilisation de filtres complémentaires pour nettoyer les données des capteurs et éliminer les variations dues aux différents types de brosses, qu'elles soient manuelles ou électriques. Les chercheurs ont aussi intégré des mesures comme le tangage et le roulis pour améliorer la reconnaissance des régions brossées. Ce niveau de détails montre leur souci d'optimiser la robustesse du système tout en restant pratique : le capteur détachable rend leur solution économique et adaptable, un avantage indéniable pour une utilisation quotidienne. Les expériences ont impliqué

17 participants, avec des sessions enregistrées dans cinq environnements différents, ce qui donne une bonne diversité contextuelle. Ils ont également comparé leur approche avec d'autres modèles, comme les SVM ou les GBDT, et ont démontré la supériorité du RF dans ce cadre précis. Cependant, cette étude présente certaines limites. Les restrictions imposées (ordre des régions et immobilité de la tête) réduisent l'applicabilité des résultats à des environnements réels non supervisés. De plus, bien que le modèle soit performant face au type de brosse, il ne tient pas compte de la variabilité interindividuelle dans les styles de brossage.

L'étude menée en 2023 par (Essalat et al., 2023) propose une méthode innovante pour surveiller le brossage dentaire et détecter les surfaces dentaires brossées en utilisant des capteurs inertiels et magnéto-inertiels (MARG). Ces capteurs sont fixés sur la manche des brosses à dents (manuelles ou électriques) et intégrés à un bracelet porté par chaque participant. Bien que les données des bracelets aient été collectées, elles n'ont pas été exploitées dans cette étude. Un filtre complémentaire a été utilisé pour estimer les angles de roulis et de tangage afin d'extraire les caractéristiques nécessaires à la classification des régions brossées. L'étude a impliqué 12 participants (10 droitiers et 2 gauchers) et a permis d'enregistrer un total de 187 sessions de brossage. Ce qui rend cette étude intéressante, c'est son double contexte : un brossage scripté, bien contrôlé, et un brossage libre, plus réaliste. Cela a permis de capturer des données riches et variées auprès de 12 participants, même si l'échantillon reste limité. Une des décisions clés de cette étude a été de réduire le nombre de régions dentaires de 16 à 9. Ce compromis visait à simplifier l'analyse et à améliorer la précision des algorithmes, comme les réseaux LSTM et les Transformers, utilisés pour classer les régions. Bien sûr, cette simplification a ses limites : elle sacrifie des détails sur des zones spécifiques, mais elle rend les résultats plus robustes, surtout face à la variabilité des mouvements de brossage. Les résultats montrent un déséquilibre dans les habitudes de brossage. Par exemple, les surfaces buccales sont systématiquement mieux brossées que les linguales, notamment celles des molaires supérieures, souvent oubliées. Cela reflète non seulement des habitudes humaines, mais aussi les défis techniques des capteurs pour capter avec précision les mouvements dans des zones moins accessibles. Les chercheurs ont noté des difficultés liées à la généralisation des résultats pour de nouveaux utilisateurs, en raison des variations individuelles dans les styles de

brossage et les mouvements de la tête. De plus, la taille limitée de l'échantillon et le manque de diversité des participants (en termes d'âge et de techniques de brossage) réduisent la portée des conclusions. Enfin, bien que le brossage libre reflète mieux les comportements réels, il est également plus complexe à analyser en raison de sa variabilité. Malgré cela, l'approche retenue ouvre des perspectives passionnantes pour améliorer la détection et le suivi des habitudes de brossage dans des conditions réelles.

Récemment, (Li & Jing, 2023) ont introduit un système de surveillance multimodal du brossage dentaire basé sur une brosse à dents intelligente intégrant à la fois des capteurs inertiels et des capteurs de force. Ce système se distingue par sa capacité à détecter les régions dentaires brossées et à mesurer la force appliquée lors du brossage, offrant ainsi un retour détaillé à l'utilisateur pour améliorer la qualité du brossage. L'étude a impliqué 20 participants adultes qui ont suivi la technique de Bass avec des instructions strictes, comme l'interdiction de bouger la tête pour minimiser les interférences. Les données collectées ont été analysées hors ligne à l'aide d'un modèle de random forest. Pour la classification des régions, la bouche a été divisée en six zones principales : quadrants supérieur gauche, supérieur droit, inférieur gauche, inférieur droit, et incisives. Cette simplification réduit la complexité du modèle tout en maintenant une justesse impressionnante hors ligne (analyse post-enregistrement, sans traitement en temps réel) de 97,6 %. Cette division contraste avec d'autres études qui subdivisent la bouche en 16 régions. Évidemment, la détection en ligne reste plus difficile, avec une justesse réduite à 74 %, ce qui montre bien à quel point les mouvements individuels et les positions variables compliquent la tâche. Cependant, les défis liés à la variabilité des comportements et à la fiabilité en temps réel montrent qu'il reste encore du travail pour perfectionner le système. Aussi, la population étudiée manque de diversité en termes d'âge et de styles de brossage, ce qui limite la généralisation des résultats. De plus, comme dans d'autres études similaires, des contraintes strictes imposées aux participants (notamment l'immobilité de la tête) réduisent l'applicabilité des résultats dans des contextes réels, où les utilisateurs se brossent les dents de manière moins contrôlée.

L'étude réalisée par (Chen et al., 2021) propose un modèle de reconnaissance des postures de brossage dentaire en utilisant une brosse à dents intelligente équipée de capteurs inertIELS à 9 axes (accéléromètre, gyroscope, magnétomètre) et d'un module Bluetooth. L'objectif principal est de détecter les régions dentaires brossées et de fournir un retour détaillé à l'utilisateur via un système de visualisation mobile. Les chercheurs ont divisé les dents en 15 régions distinctes en suivant la technique de Bass, avec un accent sur la posture de brossage comme caractéristique clé pour l'identification des zones. Pour analyser les données, les chercheurs ont testé trois modèles d'apprentissage profond : les CNN, LSTM ainsi que les réseaux de neurones probabilistes réCURRENTS (RPNN). Le modèle RPNN donne des résultats impressionnantes avec une justesse de 99,08 %, bien supérieure à celle des CNN (82,88 %) et des LSTM (77,87 %). Le fait qu'ils aient utilisé l'algorithme PSO (Particle Swarm Optimization) pour optimiser les paramètres est un choix astucieux. Le modèle a également été adapté aux habitudes individuelles, ce qui le rend parfait pour une utilisation sur des appareils à faible puissance comme les smartphones. Cependant, l'étude présente plusieurs limites. Elle ne prend pas en compte la diversité des participants, comme les différences d'âge ou de style de brossage, ce qui limite la généralisation des résultats. De plus, elle n'aborde pas l'intégration de différents types de brosses à dents (manuelles ou électriques), ce qui restreint son applicabilité dans des contextes variés. Malgré ces défis, cette étude illustre le potentiel des modèles RPNN pour améliorer la personnalisation et la précision du suivi de brossage, tout en ouvrant la voie à des solutions pratiques pour une meilleure hygiène dentaire.

2.5.2 SURVEILLANCE DU BROSSAGE DES DENTS A L'AIDE DE CAPTEURS PORTES AU POIGNET

Dans le domaine des solutions utilisant des dispositifs portables, (Huang & Lin, 2016) proposent un système basé sur une montre intelligente équipée de capteurs inertIELS (accéléromètre, gyroscope, magnétomètre). Pour améliorer la détection des rotations, des aimants ont été fixés sur les brosses manuelles utilisées, ce qui permet de surmonter les limitations des capteurs inertIELS seuls. En complément, le microphone intégré à la montre a été utilisé pour enregistrer le son des

coups de brossage, fournissant une source supplémentaire de données pour l'analyse. Pour classifier les régions brossées, les chercheurs ont utilisé un classificateur Naive Bayes, qui a permis d'atteindre une justesse de 85,6 %. La bouche a été divisée en 15 régions distinctes, alignées sur la technique de brossage de Bass, afin de garantir une couverture complète des zones dentaires. En complément, un modèle de Markov caché a été mis en œuvre pour exploiter les habitudes spécifiques des utilisateurs, améliorant ainsi la précision dans la reconnaissance des régions. Le modèle prend également en compte les mouvements du poignet, de l'avant-bras et de l'épaule pour différencier plus précisément les zones brossées. L'étude s'est déroulée sur trois semaines et a impliqué 12 participants formés à la technique Bass. Ces derniers ont été invités à brosser leurs dents dans des conditions strictement supervisées, avec des restrictions imposées sur les mouvements de la tête et du corps pour limiter les interférences. Bien que cette méthodologie garantisse des données de qualité, elle limite la généralisation des résultats à des scénarios non supervisés. De plus, le système s'est concentré exclusivement sur des brosses manuelles, laissant de côté les brosses électriques, ce qui pourrait restreindre son applicabilité pour certains utilisateurs.

En 2019, (Luo et al., 2019) introduisent le système Hygiea, exploitant uniquement les données d'accéléromètre d'un bracelet intelligent. Les données capturées par les capteurs sont traitées après l'application d'un filtre passe-bas afin de réduire les vibrations parasites et de différencier les mouvements associés aux brosses à dents manuelles et électriques. Avec un modèle LSTM basé sur l'attention, ils atteignent une justesse de 97,3 % pour la détection de 15 activités de brossage, incluant des mouvements supplémentaires comme « lever la main ». Cependant, plusieurs limites ont été identifiées dans cette étude. Les données ont été collectées dans des conditions fortement encadrées, ce qui peut réduire leur applicabilité dans des contextes réels et non supervisés. De plus, l'annotation des données a été effectuée manuellement par les chercheurs, une méthode susceptible d'introduire des biais ou des erreurs. Enfin, bien que Hygiea puisse traiter des brosses à dents manuelles et électriques, l'analyse détaillée s'est concentrée principalement sur les brosses manuelles, limitant l'évaluation des performances pour d'autres types de brosses. Ces études démontrent que les capteurs intégrés sur les brosses offrent une précision élevée pour la détection des régions spécifiques, tandis que les dispositifs portés au poignet fournissent une

alternative plus pratique et moins invasive pour des applications grand public. Cependant, l'applicabilité réelle de ces solutions est souvent limitée par les conditions supervisées dans lesquelles les données ont été collectées. L'intégration de méthodes hybrides ou la diversification des jeux de données pourrait permettre de surmonter ces défis.

2.5.3 SYNTHESE

Les travaux (Chen et al., 2021; Essalat et al., 2023; Huang & Lin, 2016; Hussain et al., 2021; Li & Jing, 2023) offrent une vue globale des approches innovantes pour la surveillance du brossage dentaire en utilisant des capteurs inertIELS. Ces études se divisent principalement en deux catégories : les solutions utilisant des capteurs intégrés aux brosses à dents et celles exploitant des dispositifs portés au poignet.

Dans les approches basées sur des capteurs intégrés aux brosses à dents, les études de (Essalat et al., 2023), (Li & Jing, 2023), et (Chen et al., 2021) explorent les avantages des capteurs placés directement sur les brosses à dents, permettant une capture précise des mouvements et une analyse détaillée des régions brossées. Ces études utilisent des unités de mesure inertIELLE (IMU) et des capteurs MARG (intégrant accéléromètre, gyroscope et magnétomètre) pour collecter des données sur les mouvements et l'orientation de la brosse. Les données sont prétraitées avec des filtres, comme le filtre complémentaire ou Madgwick, pour réduire le bruit et estimer les angles de roulis et de tangage. (Li & Jing, 2023) se distinguent par l'intégration de capteurs de force, une innovation permettant d'évaluer l'intensité du brossage.

Les études impliquent entre 12 et 20 participants, utilisant majoritairement la technique de Bass. Bien que certaines études, comme (Essalat et al., 2023), aient inclus des scénarios libres, la plupart imposent des contraintes strictes, telles que l'ordre de brossage et l'immobilité de la tête, limitant ainsi leur applicabilité dans des contextes réels. Les modèles d'apprentissage utilisés (RF, RPNN) atteignent des précisions élevées, variant de 97 % à 99 %. Cependant, la généralisation reste un défi en raison de la variabilité interindividuelle et du manque de diversité des participants (âge,

style de brossage). De plus, les contraintes imposées réduisent leur pertinence pour des environnements non supervisés.

En ce qui concerne les approches basées sur des dispositifs portés au poignet, les travaux de (Huang & Lin, 2016) et (Luo et al., 2019) explorent des solutions utilisant des bracelets ou montres intelligents pour analyser les gestes associés au brossage. Ces approches se concentrent sur la portabilité et la commodité, bien que leur précision soit légèrement inférieure à celle des capteurs intégrés aux brosses. Les dispositifs portés exploitent principalement des accéléromètres et des gyroscopes pour capturer les mouvements. (Huang & Lin, 2016) intègrent des aimants sur les brosses manuelles pour améliorer la détection des rotations, tandis que (Luo et al., 2019) utilisent un modèle LSTM basé sur l'attention, atteignant une précision remarquable de 97,3 %. Ces études incluent un échantillon similaire (12 à 15 participants) mais se concentrent uniquement sur les brosses manuelles. Les protocoles imposent des conditions strictes, comme l'utilisation exclusive de la technique de Bass et des restrictions sur les mouvements de la tête. Bien que prometteuses, ces approches sont sensibles aux mouvements parasites (gestes non liés au brossage) et peinent à s'adapter aux spécificités des brosses électriques. De plus, les données collectées en conditions supervisées limitent leur applicabilité dans des environnements naturels. Nous remarquons également que les capteurs intégrés aux brosses à dents offrent une précision supérieure et permettent une segmentation détaillée des régions brossées, ce qui est bien adapté aux applications cliniques et éducatives. Cependant, leur coût et leurs contraintes d'utilisation peuvent limiter leur adoption à grande échelle.

En revanche, les dispositifs portés au poignet sont plus pratiques et moins coûteux, mais leur précision est affectée par les mouvements parasites, et leur capacité à analyser les brosses électriques reste limitée. De plus, les deux approches partagent des limites communes, notamment l'absence de diversité des participants et la collecte de données dans des environnements supervisés, ce qui réduit leur applicabilité à des scénarios réels.

CHAPITRE 3

ETUDE ET EXPLORATION DES DONNEES

Ce chapitre est consacré à l'étude approfondie du jeu de données utilisé dans cette recherche, ainsi qu'aux différentes étapes de son exploration et de son traitement. Il couvre l'ensemble du processus, depuis la collecte des données jusqu'à l'extraction des caractéristiques pour les tâches de classification. Pour notre étude, nous avons exploité un jeu de données complet et riche, collecté par (Hussain, 2022), qui se concentre sur l'analyse des activités de brossage de dents. Ce jeu de données, accessible publiquement, a été conçu pour répondre à un manque de ressources dans ce domaine. Nous avons spécifiquement utilisé le jeu de données (Zawar, 2021), qui contient les données brutes issues d'expériences réalisées en environnement contrôlé, avec des participants recrutés pour enregistrer des sessions selon deux configurations distinctes. Il a été collecté à partir d'expériences réalisées dans des environnements contrôlés, avec des participants recrutés pour enregistrer des sessions dans deux configurations distinctes. Ce jeu de données modélise le brossage de dents comme un problème de reconnaissance d'activités humaines (HAR). Les chercheurs ont enregistré les mouvements et orientations de la brosse pour 16 sous-régions des dents, représentant autant de sous-activités distinctes.

Le processus d'exploration représenté sur la Figure 4 commence par une phase de prétraitement, comprenant le nettoyage des données brutes et leur synchronisation. Ensuite, un filtrage et une fusion des données des capteurs sont appliqués afin de réduire le bruit et d'unifier les informations issues de plusieurs sources. Des techniques de détection de points de changement sont ensuite appliquées afin de segmenter et délimiter les séquences de gestes correspondant aux différentes régions de brossage. À partir de ces segments, des caractéristiques discriminantes sont extraites pour permettre une classification précise des régions de brossage.

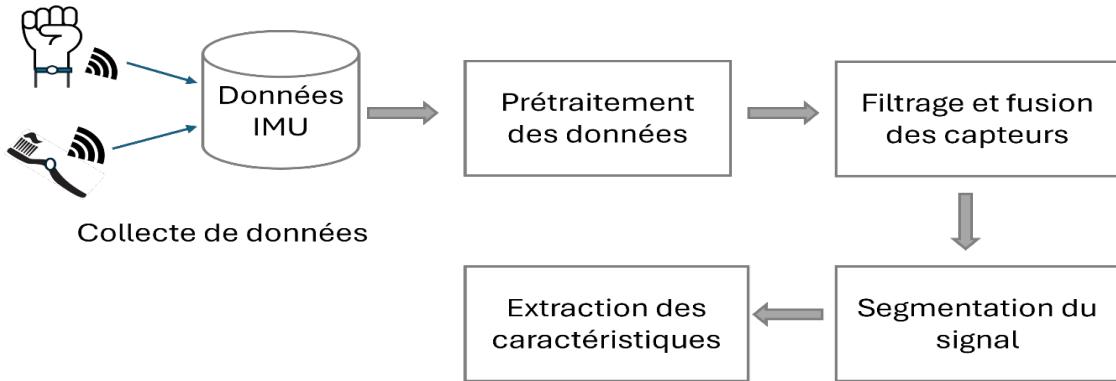


Figure 4: Étapes de la préparation des données inertielles

©Sawadogo Wendenda Aniisah 3.1 PRESENTATION DES DONNEES

Cette section présente les données utilisées dans notre étude, en détaillant d'abord les conditions de collecte puis la structure et les caractéristiques des données enregistrées. Deux configurations expérimentales ont été mises en place pour capturer les gestes de brossage de manière fidèle et représentative.

3.1.1 COLLECTE DES DONNEES

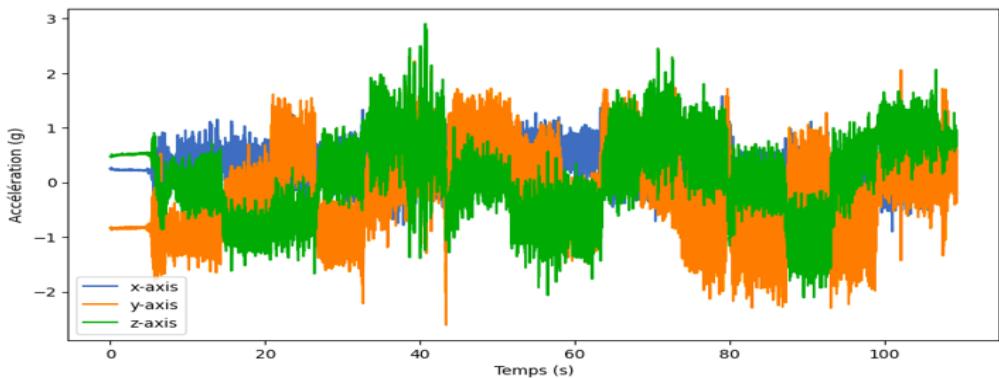
Les données ont été collectées dans un environnement contrôlé sous la supervision d'un chercheur. Des capteurs IMU (MMR) légers et compacts ont été fixés à la poignée de la brosse à dents et portés au poignet des participants. Deux configurations expérimentales ont été utilisées.

Dans la configuration 1, 17 participants (9 hommes et 8 femmes) ont été recrutés pour réaliser les sessions de brossage. Les sessions de brossage ont été réalisées de manière continue, sans pause entre les sous-régions. Un total de 62 sessions d'une durée moyenne de deux minutes a été enregistré, correspondant aux recommandations standard des dentistes.

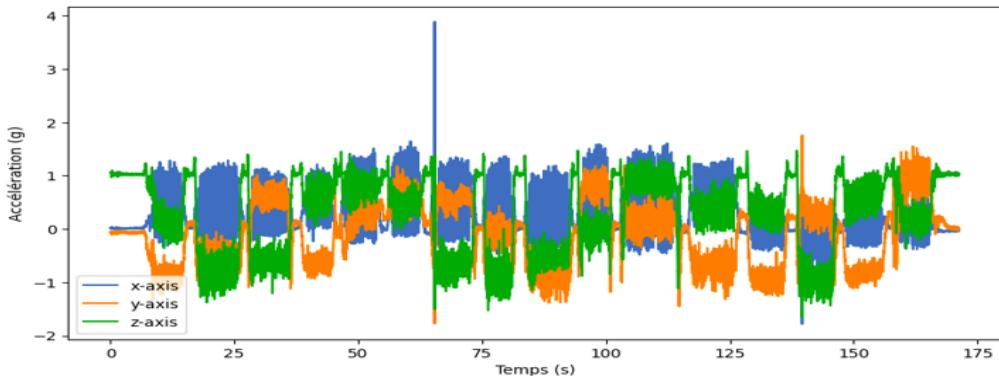
Dans la configuration 2, 17 participants (8 hommes et 9 femmes) ont été recrutés pour réaliser les sessions de brossage (dont 12 participants de la configuration 1 et 5 nouveaux). Les participants

devaient marquer une pause de quelques secondes entre le brossage des sous-régions ramenant la brosse à un point de référence. Un total de 58 sessions d'environ trois minutes a été collecté.

Dans les deux cas, les participants ont reçu des instructions détaillées pour suivre une séquence de brossage couvrant les 16 sous-régions dentaires. Les données ont été transférées via Bluetooth, puis sauvegardées sur le cloud. Un exemple des signaux collectés est présenté dans la Figure 5.



(b) Configuration 1: brossage sans pauses entre les sous-activités



(a) Configuration 2: brossage avec pauses entre les sous-activités

Figure 5: Échantillon de données de brossage issues des configurations 1 et 2. La configuration 2, avec pauses, correspond à un homme droitier utilisant une brosse manuelle ; la configuration 1, sans pauses, à un homme droitier utilisant une brosse électrique.

©Sawadogo Wendenda Aniisah

3.1.2 DESCRIPTION DES DONNEES

Le jeu de données final regroupe 122 sessions impliquant 22 participants (11 hommes et 11 femmes), âgés de 22 à 40 ans, répartis sur sept lieux différents. Les sessions sont enregistrées sous forme de fichiers CSV bruts, avec une structure distincte selon la configuration expérimentale (voir Tableau 1).

En configuration 1, chaque session génère quatre fichiers correspondant aux mesures d'accéléromètre et de gyroscope, pour la brosse à dents et le poignet. Cette configuration représente environ 1,5 million de points par capteur et par appareil.

En configuration 2, les sessions incluent également un magnétomètre, pour un total de six fichiers. Le volume des données atteint environ 2 millions de points pour les accéléromètres et gyroscopes, et 200 000 pour le magnétomètre.

Les données ont été enregistrées à un taux d'échantillonnage de 200 Hz pour l'accéléromètre et le gyroscope, et de 20 Hz pour le magnétomètre. Elles permettent de suivre les mouvements et orientations de la brosse au fil des 16 sous-régions, chacune représentant une sous-activité distincte.

La Figure 6 illustre des extraits typiques des fichiers CSV générés par les différents capteurs.

Tableau 1: Liste et description des colonnes de nos données

Nom de la colonne	Description
epoch (ms)	Temps Unix Epoch
timestamp (+1000)	Date et heure dans le fuseau horaire local
elapsed (s)	Temps écoulé depuis le début de l'enregistrement
x-axis (g)	Axe x des données brutes de l'accéléromètre
y-axis (g)	Axe y des données brutes de l'accéléromètre
z-axis (g)	Axe z des données brutes de l'accéléromètre
x-axis (deg/s)	Axe x des données brutes du gyroscope
y-axis (deg/s)	Axe x des données brutes du gyroscope
z-axis (deg/s)	Axe x des données brutes du gyroscope
x-axis (T)	Axe x des données brutes du magnétomètre
y-axis (T)	Axe x des données brutes du magnétomètre
z-axis (T)	Axe x des données brutes du magnétomètre

	epoch (ms)	timestamp (+1000)	elapsed (s)	x-axis (g)	y-axis (g)	z-axis (g)
0	1607826399794	2020-12-13T13.26.39.794	0.000	0.051	0.175	0.991
1	1607826399798	2020-12-13T13.26.39.798	0.004	0.049	0.179	1.005
2	1607826399804	2020-12-13T13.26.39.804	0.010	0.048	0.179	1.011
3	1607826399809	2020-12-13T13.26.39.809	0.015	0.044	0.181	1.011
4	1607826399815	2020-12-13T13.26.39.815	0.021	0.043	0.183	1.015

Échantillon de données d'accéléromètre

	epoch (ms)	timestamp (+1000)	elapsed (s)	x-axis (deg/s)	y-axis (deg/s)	z-axis (deg/s)
0	1607826399815	2020-12-13T13.26.39.815	0.000	-5.244	3.476	0.549
1	1607826399819	2020-12-13T13.26.39.819	0.004	-5.732	2.683	0.488
2	1607826399825	2020-12-13T13.26.39.825	0.010	-5.793	2.744	0.366
3	1607826399829	2020-12-13T13.26.39.829	0.014	-5.122	2.134	0.732
4	1607826399834	2020-12-13T13.26.39.834	0.019	-3.293	1.098	1.280

Échantillon de données de gyroscope

	epoch (ms)	timestamp (+1000)	elapsed (s)	x-axis (T)	y-axis (T)	z-axis (T)
0	1607826399772	2020-12-13T13.26.39.772	0.000	0.000025	-0.000070	0.000077
1	1607826399822	2020-12-13T13.26.39.822	0.050	0.000025	-0.000070	0.000077
2	1607826399870	2020-12-13T13.26.39.870	0.098	0.000026	-0.000069	0.000077
3	1607826399920	2020-12-13T13.26.39.920	0.148	0.000026	-0.000069	0.000076
4	1607826399970	2020-12-13T13.26.39.970	0.198	0.000025	-0.000069	0.000078

Échantillon de données de magnétomètre

Figure 6: Échantillon de nos données issues d'une session de brossage

©Sawadogo Wendenda Aniisah

3.2 PRETRAITEMENT DES DONNEES

Le prétraitement des séries temporelles constitue une étape structurante de notre pipeline, visant à garantir la fiabilité des données analysées. Cette phase vise à structurer et à nettoyer les données afin de garantir leur qualité et leur adéquation aux algorithmes d'apprentissage automatique utilisés. L'objectif est de préparer les données brutes pour qu'elles puissent être exploitées efficacement, tout en réduisant les biais et en maximisant la pertinence des analyses. Une attention particulière a été portée à la personnalisation des techniques utilisées, pour répondre aux spécificités uniques des séries temporelles issues des capteurs.

La première étape du prétraitement consiste à synchroniser les données qui viennent de différentes sources. Ici, on utilise deux capteurs : un sur la brosse à dents et un autre sur un bracelet. Bien que ces dispositifs enregistrent simultanément les mouvements, leurs horodatages peuvent différer en raison de latences ou de pertes liées à la transmission Bluetooth. Pour corriger ces décalages, une synchronisation temporelle est effectuée. Les données brutes de chaque capteur sont chargées séparément et les fichiers contenant les mesures d'accéléromètre, de gyroscope sont lus, et leurs horodatages sont convertis dans un format de date (datetime) uniforme. Cela facilite la gestion des informations temporelles. Ensuite, les fichiers sont indexés à partir de leur colonne temporelle (*timestamp*), ce qui permet d'opérer des manipulations basées sur le temps. Une plage commune est ensuite définie à partir des limites minimales et maximales des enregistrements de chaque capteur. Sur cet intervalle, un nouvel index temporel est généré à une fréquence standardisée (200 Hz), et les données sont réalignées par interpolation (*forward fill*), qui propage la dernière valeur connue pour combler les éventuelles lacunes. Ce procédé permet à la fois de synchroniser les mesures et de corriger les pertes mineures liées à la transmission. Enfin, les données interpolées des différents capteurs sont combinées en un tableau unique par concaténation des colonnes. Un extrait de ce tableau final, comprenant les mesures issues de l'accéléromètre et du gyroscope, est présenté dans la Figure 7.

La seconde étape du prétraitement concerne le traitement des données manquantes et des segments non exploitables. Deux cas ont été identifiés : le premier type concerne les pertes de données dues à la transmission Bluetooth entre les capteurs et le dispositif d'enregistrement. Ces pertes se manifestent par des trous dans les enregistrements, créant des discontinuités temporelles dans les séries temporelles. Pour combler ces lacunes, une interpolation (*forward fill*) est utilisée. Le second type ne correspond pas à des valeurs absentes au sens strict, mais à des segments de données enregistrés mais jugés non exploitables. Il s'agit notamment d'interruptions imprévues ou de gestes incohérents survenus pendant le brossage, tels que des arrêts prolongés ou des mouvements ne correspondant pas aux consignes. Ces interruptions, bien qu'enregistrées, ne sont pas pertinentes pour l'analyse des gestes de brossage. Elles sont détectées à l'aide de seuils prédéfinis, basés par exemple sur une faible variance sur plusieurs échantillons consécutifs. Une fois

identifiés, les lignes correspondant à ces segments non conformes sont supprimées pour garantir une meilleure qualité et cohérence des données utilisées dans les analyses.

	epoch (ms)	timestamp (+1000)	elapsed (s)	x-axis (g)	y-axis (g)	z-axis (g)	x-axis (deg/s)	y-axis (deg/s)	z-axis (deg/s)
0	1.598149e+12	2020-08-23 12:22:37.332		5.366	0.179	-1.170	0.476	-11.768	-4.451
1	1.598149e+12	2020-08-23 12:22:37.336		5.370	0.145	-0.754	0.565	-8.537	0.305
2	1.598149e+12	2020-08-23 12:22:37.342		5.376	0.214	-0.755	0.435	-7.134	-4.146
3	1.598149e+12	2020-08-23 12:22:37.346		5.380	0.118	-0.983	0.526	-2.012	-4.451
4	1.598149e+12	2020-08-23 12:22:37.351		5.385	0.231	-1.042	0.517	9.390	-4.268

Figure 7: Extrait du tableau synchronisé issu de l'interpolation des données capteurs

©Sawadogo Wendenda Aniisah

3.3 FILTRAGE ET FUSION DES CAPTEURS

Les mesures issues des capteurs inertIELS accéléromètres, gyroscopes et magnétomètres sont particulièrement sensibles à diverses perturbations, qui introduisent du bruit dans les données collectées. Ce bruit peut avoir une origine interne, liée aux limitations technologiques des capteurs eux-mêmes : vibrations mécaniques produites par les moteurs des brosses électriques, bruit électrique de fond, ou encore dérive progressive des gyroscopes. D'autres sources de perturbation proviennent de l'environnement d'utilisation : gestes brusques ou irréguliers du bras lors du brossage manuel, champs magnétiques perturbés par la présence d'objets métalliques, variations de température, ou encore vibrations ambiantes.

Dans le contexte du brossage dentaire, la nature de ces perturbations varie selon le type de brosse utilisée. Une brosse électrique génère des vibrations continues qui introduisent des composantes de haute fréquence qui ne correspondent pas aux mouvements réels du bras. En revanche, l'utilisation d'une brosse manuelle implique des gestes souvent plus saccadés et moins réguliers, produisant des mouvements brusques et imprévisibles.

Pour atténuer ces perturbations, des filtres Butterworth (Oppenheim et al., 1997) sont appliqués selon le type de brosse. Un filtre passe-bas à 2 Hz (Shouran & Elgaml, 2020) est utilisé pour la brosse électrique, afin d'éliminer les hautes fréquences dues au moteur. Pour la brosse

manuelle, un filtre passe-bande entre 2 et 6 Hz cible les fréquences typiques du brossage, en excluant les composantes liées à la gravité ou au bruit. L'effet de ce filtrage peut être observé à la Figure 8 et Figure 9, qui présentent respectivement pour une brosse manuelle et électrique, les signaux bruts et filtrés issus des capteurs au poignet et sur la brosse

À la suite de cette étape, un filtre complémentaire (Mahony et al., 2008) est appliqué pour combiner les données des accéléromètres et des gyroscopes. Cela permet d'estimer précisément les angles d'orientation roulis et tangage. Le gyroscope mesure les variations rapides de mouvement, mais il accumule des erreurs à long terme, appelées dérives. L'accéléromètre, quant à lui, est stable à long terme grâce à la gravité, mais il est souvent bruité par des mouvements parasites. Le filtre complémentaire compense ces défauts. Un passe-haut corrige la dérive du gyroscope, tandis qu'un passe-bas réduit le bruit des données de l'accéléromètre.

La combinaison des données des deux capteurs se fait selon cette formule :

$$\phi = \beta \cdot \phi_g + (1 - \beta) \cdot \phi_a \quad (3.1)$$

où :

- ϕ représente l'angle final estimé (roulis ou tangage),
- ϕ_g est l'angle estimé par le gyroscope,
- ϕ_a est l'angle estimé par l'accéléromètre,
- β est un coefficient (souvent entre 0,95 et 0,98) qui ajuste l'équilibre entre les deux sources.

Ce système tire parti des forces des deux capteurs tout en limitant leurs défauts. Cela permet d'obtenir des données fiables pour analyser les gestes complexes et rapides du brossage, où les vibrations et mouvements imprévus pourraient perturber les mesures. Grâce à ces filtres, les signaux sont plus clairs et adaptés aux analyses et traitements ultérieurs.

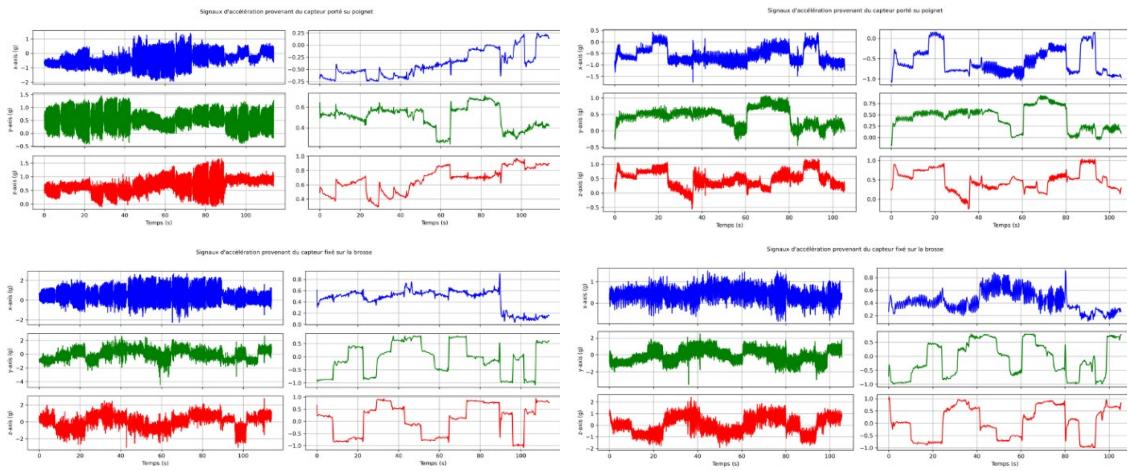


Figure 8: Signaux d'accélération bruts et filtrés provenant d'une brosse manuelle

Figure 9: Signaux d'accélération bruts et filtrés provenant d'une brosse électrique

©Sawadogo Wendenda Aniisah

3.4 DETECTION DES POINTS DE CHANGEMENT

Les points de transition, ou moments où la brosse change de région, se traduisent par des variations dans les signaux, notamment en termes d'orientation et de mouvement. Leur détection joue un rôle clé pour analyser les gestes complexes associés au brossage de dents. Cependant, les méthodes classiques de détection de points de changement (CPD), bien qu'efficaces dans d'autres contextes, ne sont souvent pas adaptées à notre problématique (Truong et al., 2018). Ces techniques peuvent être trop coûteuses en termes de calcul ou trop spécifiques à d'autres applications, ce qui limite leur performance face à des signaux inertIELS bruyants et enregistrés à haute fréquence.

Dans ce contexte, plusieurs algorithmes de détection ont été explorés pour segmenter les séries temporelles collectées lors des sessions de brossage. En premier lieu, nous avons utilisé CluMing, une méthode basée sur le clustering K-means. Cette approche se démarque par sa capacité à regrouper les signaux en clusters et à gérer les transitions en fusionnant les segments courts. Par la suite, d'autres méthodes complémentaires ont été testées : la segmentation basée sur les pics, qui identifie les événements marquants dans les signaux, et la détection locale de points de

changement, qui utilise une analyse statistique en fenêtres temporelles pour repérer les variations significatives. Enfin, une méthode de détection des moments d'immobilité a été intégrée pour isoler les périodes de faible activité ou de pauses, permettant ainsi de séparer clairement les mouvements continus des transitions. Ces différentes approches ont été choisies pour leurs avantages respectifs et leur capacité à s'adapter aux particularités des signaux inertiels. Cette section explore leur fonctionnement et leur application dans le cadre de notre étude.

3.4.1 ALGORITHME CLUMING

L'algorithme CluMing sert à analyser les données des capteurs pour repérer les moments où la brosse passe d'une zone à une autre pendant une session de brossage (Essalat, 2023). Il se divise en deux étapes principales : le regroupement des données (clustering) et la fusion des segments. Ces étapes sont essentielles pour obtenir des résultats fiables et représentatifs.

La première étape consiste à regrouper les données en fonction de leur similitude. Les informations collectées par les capteurs, comme celles de l'accéléromètre et gyroscope, sont regroupées en plusieurs catégories appelées clusters grâce à l'algorithme de clustering K-means++ (Deshpande et al., 2020; Vassilvitskii & Arthur, 2006). Cet algorithme minimise la distance entre les données et leur centre de cluster. Pour choisir le bon nombre de clusters (k), on utilise une méthode appelée *méthode du coude* (Humaira & Rasyidah, 2020), qui aide à trouver un équilibre entre la précision et la simplicité. Pour garantir un résultat pertinent, on applique une réduction minimale de 20 % de l'inertie, qui est une mesure de la variation non expliquée. Cette étape permet d'attribuer chaque échantillon à un cluster, ce qui aide à détecter les changements de zones (Figure 10 (a)).

La seconde étape, la fusion des segments, s'attaque aux transitions identifiées lors du clustering. Parfois, certains segments sont très courts et ne représentent pas réellement les gestes de brossage. Pour régler ce problème, les segments trop petits (moins de 1 seconde, soit $\tau=200\backslash\tauau = 200$ échantillons à 200 Hz) sont fusionnés avec leurs segments voisins. L'idée est de regrouper ces petits segments avec celui de gauche ou de droite en fonction de leur proximité et de leur

longueur. Pour décider avec quel segment fusionner, un critère appelé SLR est utilisé. Ce critère compare la longueur des segments voisins et leur similarité, c'est-à-dire à quel point leurs données (accéléromètre et gyroscope) se ressemblent. Plus un segment est long et proche en termes de caractéristiques, plus il est probable qu'il soit fusionné. La formule utilisée pour calculer ce score est indiquée dans l'équation (3.2) :

$$SLR = \frac{l_R}{l_L} - \lambda \frac{\|F - F_R\|^2}{\|F - F_L\|^2} \quad (3.2)$$

où F est la moyenne des données du segment étudié, l_R et l_L sont celles des segments voisins gauche et droit, et λ est un facteur qui équilibre la longueur et la distance. Le processus continue jusqu'à ce qu'il n'y ait plus de segments trop courts. À la fin, les segments obtenus sont suffisamment longs et cohérents pour représenter les gestes réels de brossage, en éliminant les transitions trop rapides ou erronées comme illustré en Figure 10(b).

Cet algorithme est utile pour étudier des gestes complexes et dynamiques, comme ceux impliqués dans le brossage. Il permet de délimiter les zones de brossage de manière plus structurée tout en réduisant les erreurs dues au bruit ou aux mouvements rapides. Grâce à lui, on peut analyser les comportements de brossage de façon efficace et fiable. Les segments finaux sont visibles à travers les délimitations verticales (en rouge) représentant le passage d'une région de brossage à une autre. La Figure 11 montre l'évolution des signaux bruts de l'accéléromètre sur les axes x, y et z, avec les indices des segments associés au brossage de différentes zones de la bouche.

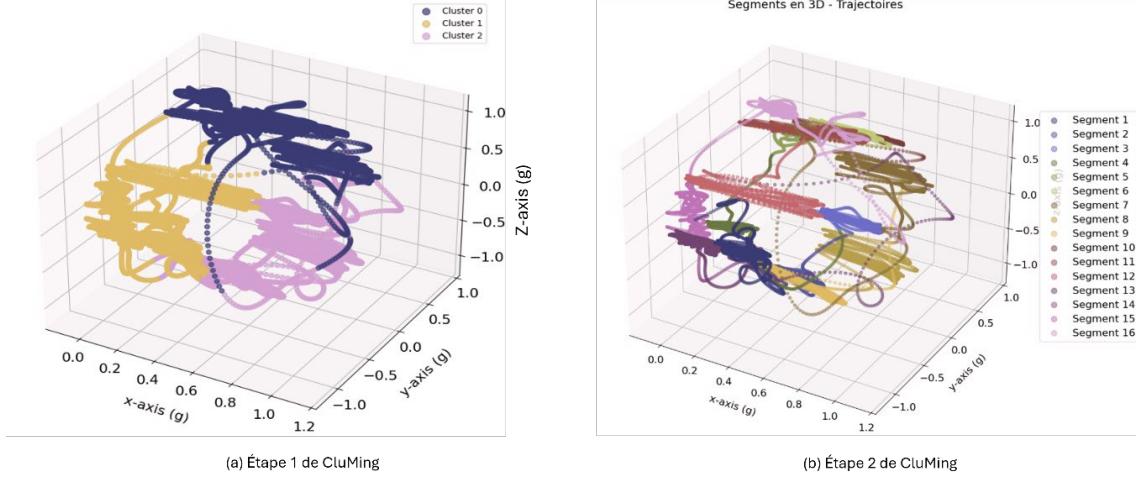


Figure 10: Étapes du traitement par l'algorithme CluMing appliquée aux données de brossage. La première image (a) Étape 1 de CluMing représente le regroupement des données en clusters via K-means++. La seconde (b) Étape 1 de CluMing montre les segments finaux représentant les régions de brossage après fusion.

©Sawadogo Wendenda Aniisah

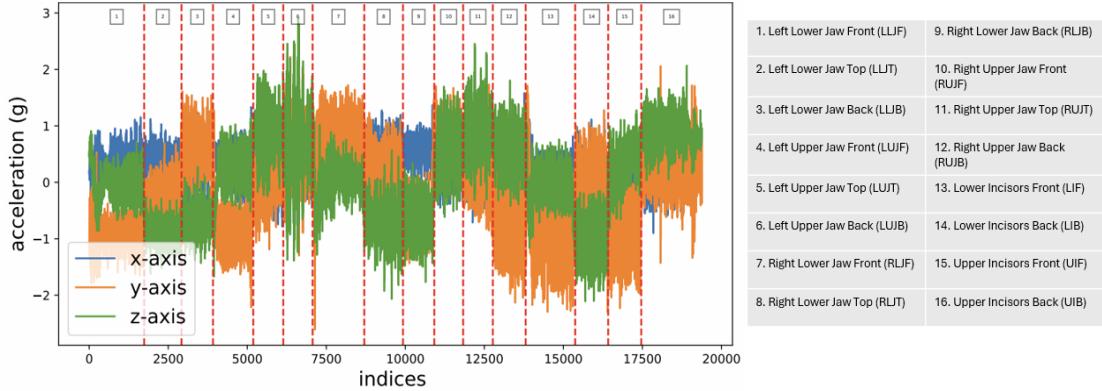


Figure 11: Évolution de l'accélération sur les trois axes (x, y, z) avec les segments de brossage détectés par l'algorithme CluMing

©Sawadogo Wendenda Aniisah

3.4.2 SEGMENTATION BASEE SUR LES PICS

La détection des changements dans les données des capteurs, comme l'accéléromètre et le gyroscope, est une méthode qui nous a montré son efficacité dans la détection des points de changements. L'idée est de repérer les pics dans les signaux, qui correspondent à des gestes importants. Pour cela, on utilise la fonction *find_peaks* de SciPy (SciPy, 2025), qui localise ces pics selon plusieurs critères. Par exemple, on ne garde que les pics dépassant une certaine hauteur, ce qui permet d'éliminer les petites variations dues au bruit. On impose aussi une distance minimale entre les pics pour éviter de compter plusieurs fois le même mouvement mais également de privilégier les segments longs. Enfin, on vérifie la prominence des pics, c'est-à-dire leur capacité à se démarquer des fluctuations environnantes, pour ne garder que les événements les plus marquants.

Un point $p \in W(t)$ est retenu comme un pic valide si :

$$S(p) > \text{seuil_hauteur} \wedge \text{Prominence}(p) > \text{seuil_prominence} \wedge \forall p_i \in P, |p - p_i| \geq \text{distance_min}$$

- $S(p)$: valeur du signal au point p ;
- $\text{Prominence}(p)$: proéminence du pic p ;
- P : ensemble des pics déjà détectés ;
- $|p - p_j|$: distance (temporelle) entre le pic p et un autre pic p_j ;
- $W(t)$: fenêtre temporelle en cours d'analyse.

Les résultats peuvent être affinés en cherchant des variations similaires entre les deux capteurs. Cette vérification, faite sur une petite période autour de chaque pic, aide à exclure les faux positifs qui pourraient venir de bruits propres à un seul capteur.

En combinant ces critères et cette corrélation entre les capteurs, on obtient une méthode fiable pour identifier les transitions de mouvements. Cela permet de se concentrer sur les gestes pertinents tout en réduisant l'effet des bruits et des anomalies, ce qui est crucial pour segmenter correctement les données. La Figure 12 illustre cette détection des changements à travers les angles d'orientation

extraits des données gyroscopiques. Chaque pic, représenté par une croix rouge, correspond à un changement significatif dans la trajectoire du geste de brossage.

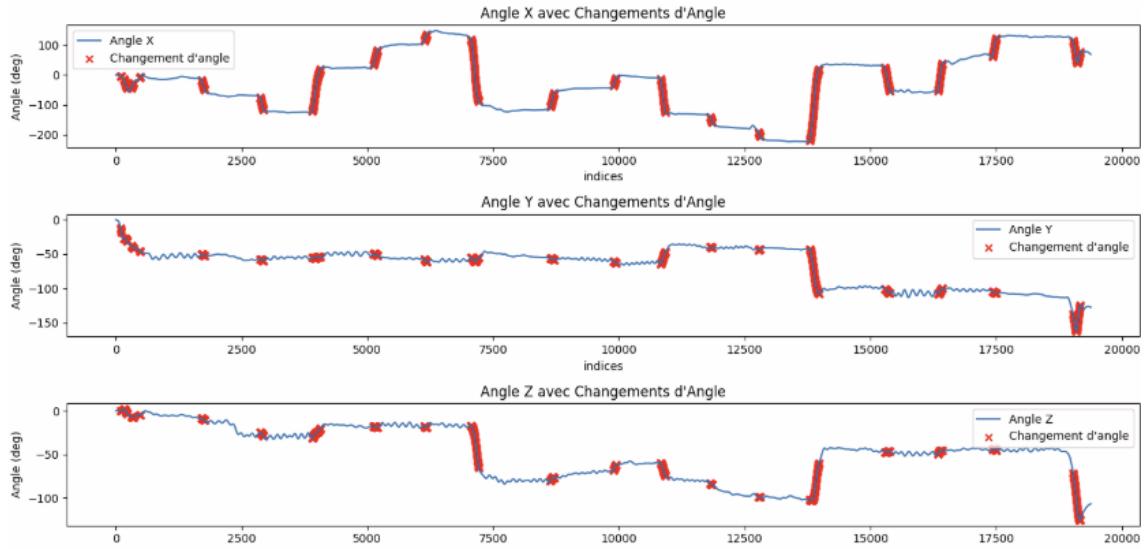


Figure 12: Détection des changements d'angle sur les axes X, Y et Z. Les croix rouges représentent les pics détectés, considérés comme des transitions dans les gestes de brossage.

©Sawadogo Wendenda Aniisah

3.4.3 LA DETECTION LOCALE DE POINTS DE CHANGEMENT

L'algorithme de détection locale de points de changement est une méthode simple pour repérer les transitions importantes dans les données issues de capteurs comme l'accéléromètre et le gyroscope. Il fonctionne en divisant les signaux en petites fenêtres de longueur fixe L , qui se chevauchent partiellement d'un pourcentage P . Ces fenêtres permettent d'analyser les signaux par petites sections, avec des paramètres typiques comme $L=50$ et $P=20$, choisis pour offrir un bon compromis entre précision et rapidité.

Dans chaque fenêtre, on calcule deux mesures importantes : la moyenne des valeurs (\bar{F}_i) et l'écart-type (σ_i). La moyenne reflète l'intensité générale des signaux, tandis que l'écart-type mesure leur variabilité. Ensuite, on compare la différence entre les moyennes des fenêtres consécutives à un seuil. La formule est définie comme suit dans l'équation (3.3) :

$$\|\bar{F}_i - \bar{F}_{i-1}\| > \|\sigma_{i-1}\| \cdot \alpha + \beta \quad (3.3)$$

Le seuil s'ajuste automatiquement en fonction des variations des signaux récents. Plus précisément, le paramètre α rend le seuil proportionnel au niveau de bruit ou d'activité récente, tandis que β ajoute une marge fixe de sécurité permettant de stabiliser la détection. Cela rend l'algorithme capable de s'adapter aux fluctuations normales tout en évitant de détecter des changements liés au bruit ou à de petites anomalies.

Un changement est identifié lorsqu'une variation significative dépasse ce seuil. Pour éviter de détecter trop de transitions rapprochées, on impose un délai minimum entre deux points de changement. Ces points permettent ensuite de découper les signaux en segments distincts, en s'assurant d'inclure le début et la fin des données pour une segmentation complète.

Cet algorithme est particulièrement pratique pour analyser des signaux complexes. Grâce à son approche flexible et adaptable, il identifie précisément les transitions importantes tout en éliminant les faux positifs causés par le bruit. Cela en fait un outil idéal pour des applications comme l'étude des gestes, par exemple pour comprendre les mouvements lors du brossage de dents.

3.4.4 LA DETECTION DES MOMENTS D'IMMOBILITE

L'algorithme de détection des moments d'immobilité sert à repérer les moments où les mouvements captés par un accéléromètre et un gyroscope sont très faibles ou presque absents, ce qui correspond à des pauses. Pour ça, il commence par calculer une mesure simplifiée de l'activité, qu'on appelle *magnitude*, et qui combine les trois axes (x, y, z) des capteurs en une seule valeur. Cette *magnitude* est calculée avec la formule (3.4)

$$\text{Magnitude} = \sqrt{x^2 + y^2 + z^2} \quad (3.4)$$

Ensuite, l'algorithme fixe des seuils ajustables, basés sur les données : par exemple, le seuil pour l'accéléromètre est mis au 85^e percentile, et celui du gyroscope au 50^e percentile. Ces seuils permettent d'ignorer les petites variations ou les anomalies.

Pour identifier les pauses, l'algorithme analyse les données par fenêtre. Ces fenêtres sont de taille fixe ($L=50$). Dans chaque fenêtre, il compare la magnitude des signaux aux seuils définis. Une pause est détectée si les magnitudes restent en dessous de ces seuils pour une durée minimale de (T_m) échantillons. Cette méthode par fenêtres permet d'être précis tout en étant insensible aux petites fluctuations dans les données.

Même si cet algorithme n'a pas été conçu pour détecter les transitions et changement de régions, il peut très bien être utilisé avec d'autres algorithmes de segmentation. Une fois les segments définis, l'algorithme de détection des moments d'immobilité peut ensuite déterminer lesquels de ces segments correspondent à des pauses. Cela aide à différencier clairement les segments actifs (mouvements de brossage) des segments inactifs (pauses) comme on peut le voir sur la Figure 13.

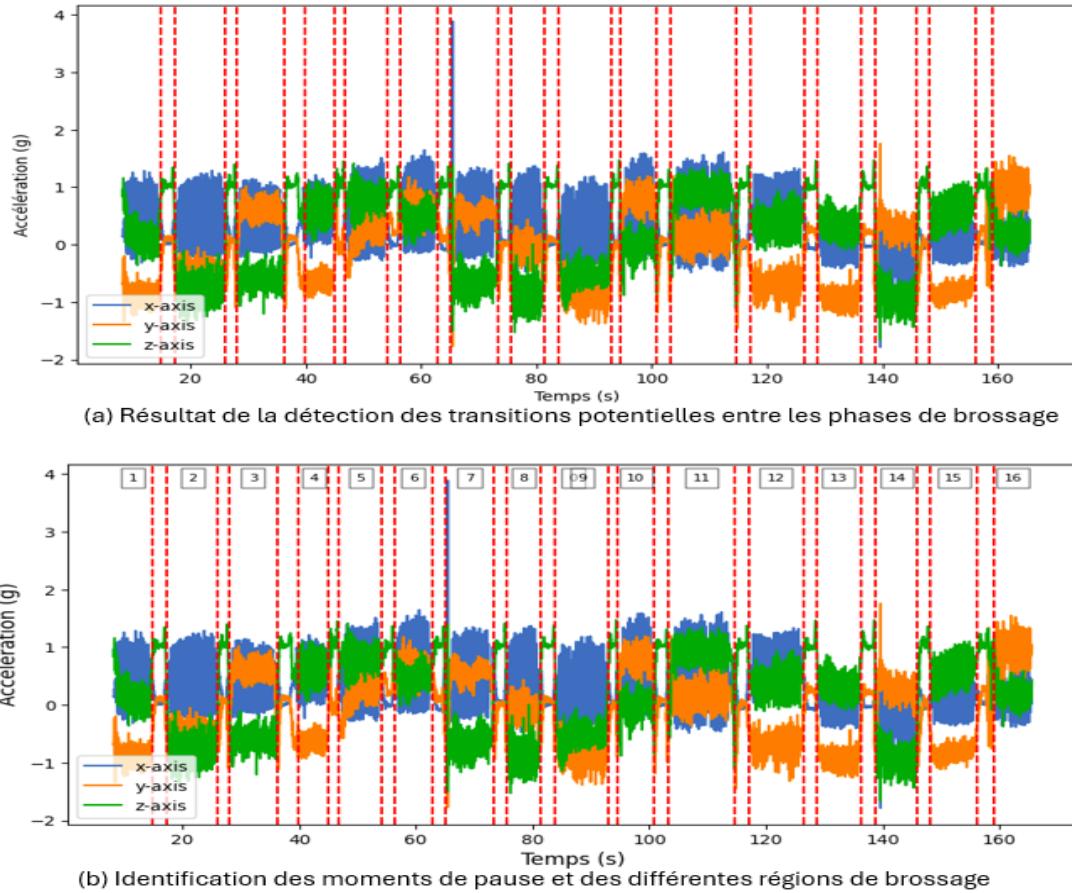


Figure 13: Illustration du processus de segmentation des données d'une session de brossage

©Sawadogo Wendenda Aniisah

3.4.5 RÉSULTATS DES ALGORITHMES DE SEGMENTATION

Pour juger de la pertinence de nos différentes méthodes de segmentation, une comparaison a été réalisée entre les segments produits automatiquement et ceux définis manuellement dans les données de référence. Comme on peut le remarquer sur le Tableau 2, les performances de tous les algorithmes se dégradent nettement sans filtrage, en particulier pour la méthode basée sur les pics. La présence de pauses marque également une différence décisive : les scores de CluMing et la détection locale de points de changement (LCP) augmentent fortement, ce qui montre que ces méthodes dépendent de ruptures nettes pour bien segmenter. Le positionnement du capteur influence aussi les résultats. Toutes les méthodes sont plus efficaces lorsqu'elles sont placées sur la brosse, car les gestes y sont plus distincts, mais on remarque que l'approche par pics tolère mieux

un capteur au poignet. Le LCP reste le plus stable, quels que soient les contextes, ce qui confirme l'intérêt de son mécanisme adaptatif basé sur la dynamique locale des signaux. En raison de cette stabilité, la méthode LCP a été retenue pour la suite des expérimentations.

Tableau 2: Évaluation des algorithmes de segmentation selon différents contextes

	Algorithme CluMing	Peak-based	LCP
Filtrées	94,58 %	92,72 %	92,36 %
Non Filtrées	88,75 %	85,34 %	89,20 %
Brosse Electrique	91,23 %	90,35 %	84,15 %
Brosse Manuelle	94 %	93,69 %	95,64 %
Capteur - brosse	96,08 %	94 %	94,24 %
Capteur au poignet	88,12 %	91,10 %	90,33 %
Configuration 1 (sans pause)	72,92 %	85 %	77,64 %
Configuration 2 (avec pauses)	93,28 %	95,98 %	98,54 %

3.5 EXTRACTION DES CARACTÉRISTIQUES

L'extraction des caractéristiques transforme les fenêtres temporelles de données brutes en vecteurs d'attributs de taille fixe, plus simples d'utilisation pour les algorithmes d'apprentissage automatique. Cette étape augmente la dimensionnalité en générant plusieurs variables à partir des signaux d'origine, tout en mettant en évidence des éléments essentiels comme les tendances générales, les variations ou les motifs spécifiques. Dans notre étude, cette étape prend toute son importance, car les données inertielles capturent des mouvements rapides et souvent irréguliers du brossage dentaire, parfois perturbés par des gestes moins précis, etc. En extrayant des caractéristiques pertinentes pour chaque région de brossage, nous facilitons l'apprentissage des modèles et améliorons leur capacité à classifier correctement les zones ciblées.

En s'inspirant des indications de (Hussain, 2022), l'extraction des caractéristiques dans le code commence par la structuration des données brutes issues des capteurs inertIELS. Après la segmentation, chaque session de brossage dentaire est divisée en 16 régions correspondant aux zones de brossage. Les pauses ou transitions identifiées, marquées par le numéro de segment 0, sont exclues de l'analyse pour se concentrer uniquement sur les périodes actives. Les colonnes utilisées pour l'extraction des caractéristiques sont spécifiques : "pitch", "roll", "x-acc_0Hz", "y-

"acc_0Hz", "z-acc_0Hz", "x-acc_4Hz", "y-acc_4Hz", et "z-acc_4Hz". Ces données permettent d'obtenir des informations sur les orientations et accélérations qui sont utiles pour différencier les régions dentaires.

Pour la segmentation, les signaux inertiels (accéléromètre et gyroscope) extraits pour chaque région de brossage sont découpés en fenêtres temporelles. Ces fenêtres sont définies par deux paramètres : la taille de la fenêtre (L) et le chevauchement (O), avec le pas calculé comme $L \times O$. Chaque fenêtre reçoit un identifiant unique (window_id) pour faciliter leur traitement. Une fois les fenêtres créées, les données des capteurs sont normalisées pour assurer une échelle comparable entre les caractéristiques, ce qui est important pour améliorer les performances des modèles de classification. Les données normalisées sont ensuite structurées pour être compatibles avec la bibliothèque *tsfresh* (Christ et al., 2018), qui reformule les informations dans un format où chaque variable est liée à un identifiant de fenêtre et un marqueur temporel.

Pour extraire les caractéristiques, des fonctions statistiques spécifiques sont appliquées aux données de chaque fenêtre. Parmi ces fonctions, on trouve la moyenne, l'écart-type, la kurtosis, l'asymétrie, ainsi que les valeurs minimale et maximale. Ces fonctions sont définies dans le dictionnaire *statistical_fc_parameters* et utilisées dans la fonction *extract_features* de *tsfresh* pour transformer les données segmentées en vecteurs consolidés. Ce processus réduit la matrice brute *X_raw* en un vecteur *x* en appliquant les fonctions aux axes pertinents : tangage, roulis, et les accélérations filtrées ou non filtrées. Une autre étape du code consiste à gérer les valeurs manquantes (NaN) dans les caractéristiques extraites. La fonction *impute* est utilisée pour combler les lacunes, suivie d'une vérification pour s'assurer qu'il n'en reste pas. Les colonnes contenant uniquement des NaN sont supprimées pour maintenir la qualité des données. Le vecteur final contient 48 caractéristiques, résumant les informations issues des fonctions appliquées. Une sélection supplémentaire des caractéristiques avec la fonction *feature_selection* est réalisée pour les données d'entraînement. Les fenêtres sont regroupées par leurs identifiants (window_id), et une majorité est appliquée sur les étiquettes des fenêtres pour générer une colonne cible (y).

3.6 CONCLUSION

Ce chapitre a permis d'explorer en détail le processus de préparation des données utilisées pour l'analyse des gestes de brossage. À travers les étapes de collecte, de synchronisation, de filtrage et d'extraction de caractéristiques, un pipeline rigoureux a été mis en place afin d'assurer la qualité des signaux traités. La richesse du jeu de données, provenant de deux configurations expérimentales distinctes, l'une avec un brossage continu et l'autre avec des pauses marquées entre les sous-régions, constitue un atout majeur pour évaluer la fiabilité des algorithmes dans des conditions proches du réel.

Par ailleurs, plusieurs méthodes de segmentation ont été testées pour détecter les transitions entre les régions de brossage. Parmi elles, la détection locale de points de changement (LCP) a été retenue pour les expérimentations, en raison de sa stabilité dans divers contextes, de sa rapidité d'exécution, et de ses résultats plus constants par rapport à la détection par pics. Elle s'est ainsi imposée comme le meilleur compromis entre performance et efficacité pour la suite de l'étude. Les techniques de prétraitement et de segmentation présentées permettent d'atteindre l'objectif spécifique 1, en transformant les signaux bruts en données exploitables, et l'objectif spécifique 2, en identifiant efficacement les transitions entre les gestes de brossage.

CHAPITRE 4

RÉSULTATS

Ce chapitre présente les résultats obtenus à travers l'ensemble des expérimentations menées pour évaluer la capacité des modèles à reconnaître les régions de brossage à partir des données issues de capteurs inertiels. Il s'agit d'analyser leur performance en tenant compte des variations entre utilisateurs, des conditions expérimentales et des caractéristiques du signal. Nous avons d'abord évalué plusieurs modèles de classification, allant d'algorithmes classiques comme la FA, le k-NN ou les SVM à des approches plus avancées, telles que les réseaux de neurones récurrents (RNN, LSTM, GRU) et le Transformer, mieux adaptés à la modélisation des séquences temporelles.

Plusieurs stratégies de validation ont été utilisées : la validation croisée pour estimer la performance globale des modèles, la validation Leave-One-Subject-Out (LOSO) pour évaluer leur capacité de généralisation à de nouveaux utilisateurs, et la validation One-Session-Out (OSO) pour tester leur stabilité temporelle intra-sujet. L'analyse repose sur diverses mesures (justesse, précision, rappel, F1-score), ainsi que sur des matrices de confusion pour mieux cerner les erreurs. La seconde partie du chapitre est consacrée aux expérimentations paramétriques, examinant l'influence de facteurs comme la fréquence d'échantillonnage, la taille des fenêtres, la position du capteur ou le type de brosse, afin d'identifier les conditions optimales d'utilisation.

4.1 MODELES DE CLASSIFICATION POUR LA RECONNAISSANCE D'ACTIVITES

Dans un premier temps, nous avons choisi de reproduire la méthodologie décrite par (Hussain, 2022), en utilisant des algorithmes classiques d'apprentissage supervisé tels que Random Forest (RF), Gradient Boosted Decision Trees (GBDT), Support Vector Machines (SVM), Gaussian Naïve Bayes (GNB) et K-Nearest Neighbors (KNN).

Afin d'approfondir notre analyse et d'optimiser les performances de classification, nous avons également intégré des modèles avancés, tels que les réseaux de neurones récurrents (RNN),

notamment Long Short-Term Memory (LSTM) et Gated Recurrent Units (GRU). Ces architectures, connues pour leur capacité à capturer les dépendances temporelles dans des séries longues, sont particulièrement adaptées pour modéliser les transitions entre les régions de brossage et tirer parti des caractéristiques temporelles des signaux.

4.1.1 ALGORITHMES CLASSIQUES

La classification supervisée repose d'abord sur des algorithmes classiques, éprouvés pour leur simplicité d'usage et leur efficacité sur des jeux de données bien structurés. Ces modèles fournissent une première base de référence dans l'analyse de nos signaux issus des capteurs. Cette section présente ces approches traditionnelles et détaille leur mise en œuvre dans le cadre de notre étude.

INTRODUCTION AUX ALGORITHMES CLASSIQUES

Logistic Regression, appréciée pour sa simplicité et son interprétabilité, est utilisée avec une approche One-Vs-Rest pour la classification multi-classes. Elle offre une première base de référence, avec une régularisation L1 ou L2 pour limiter le surapprentissage. L'optimisation des hyperparamètres est assurée par GridSearchCV, et un nombre élevé d'itérations est permis pour garantir la convergence sur des signaux bruités.

Les Decision Trees permettent de modéliser les relations non linéaires entre les caractéristiques des capteurs IMU et les classes, tout en restant facilement interprétables. Leur capacité à hiérarchiser les critères les plus discriminants est particulièrement utile face à la diversité des styles de brossage. Une optimisation fine de leur profondeur et de leurs critères d'impureté est réalisée par GridSearchCV.

Random forest renforce la stabilité des prédictions en combinant plusieurs arbres entraînés sur des sous-échantillons. Elle est bien adaptée aux données bruitées et aux variations inter-

utilisateurs. Une exploration large des hyperparamètres via RandomizedSearchCV permet d'ajuster la profondeur, le nombre d'arbres et les critères de sélection.

Contrairement aux approches à apprentissage explicite, le k-NN classe les exemples par simple comparaison avec les plus proches voisins, sans phase d'entraînement, ce qui le rend robuste à la variabilité gestuelle. Différentes valeurs de k, de pondération et de distances (euclidienne, manhattan) sont testées via GridSearchCV avec validation croisée.

Les SVM cherchent à maximiser la séparation entre classes par un hyperplan optimal, avec une bonne capacité de généralisation, même en présence de signaux complexes. L'approche multi-classes combine les stratégies OvR et OvO, tandis que les paramètres (C, noyau, gamma) sont ajustés par RandomizedSearchCV.

Naive Bayes, malgré son hypothèse d'indépendance des caractéristiques, offre une solution rapide et légère, particulièrement adaptée aux petits volumes de données. La version gaussienne est retenue, avec régularisation du paramètre alpha pour limiter les biais liés aux classes minoritaires.

Le Gradient Boosting quant à lui combine plusieurs arbres construits séquentiellement, chacun corrigeant les erreurs du précédent. Cette dynamique permet d'améliorer la précision sur des tâches complexes comme la classification fine des régions de brossage. Trois variantes sont comparées : XGBoost (efficace et régularisé), LightGBM (rapide et précis) et CatBoost (adapté aux variables catégorielles). Chacune bénéficie d'une optimisation dédiée des paramètres clés via RandomizedSearchCV.

IMPLÉMENTATION DES ALGORITHMES CLASSIQUES

L'ensemble des modèles classiques a été intégré dans un pipeline complet, allant du prétraitement des signaux à l'évaluation finale des performances. Le découpage des données a été réalisé à l'aide de la fonction `train_test_split` de scikit-learn, avec une répartition 80/20 entre l'entraînement et le test, en appliquant une stratification sur les classes afin de préserver leur distribution initiale.

Les plages des hyperparamètres testées pour les modèles les plus performants :

Pour Random Forest, la profondeur des arbres a été explorée entre 10 et 25, ou laissée illimitée, avec un nombre d'arbres variant entre 50 et 200. Les critères d'impureté (*gini*, *entropy*) ont été comparés, de même que les méthodes de sélection des attributs (*sqrt*, *log2*).

Le modèle XGBoost a été testé avec des profondeurs d'arbre comprises entre 3 et 7, un taux d'apprentissage allant de 0.01 à 0.2, et des ratios de sous-échantillonnage (*subsample* et *colsample_bytree*) compris entre 0.7 et 1.0.

Pour LightGBM, les tests ont couvert des profondeurs jusqu'à 10, un nombre de feuilles variant de 20 à 50, et un taux d'apprentissage entre 0.01 et 0.1. Les paramètres de régularisation *min_gain_to_split* et *min_data_in_leaf* ont été respectivement évalués entre 0.01 et 0.1, et entre 20 et 50.

Le modèle CatBoost a été entraîné avec des profondeurs allant de 4 à 10, un taux d'apprentissage entre 0.01 et 0.2, et un paramètre de régularisation *l2_leaf_reg* compris entre 1 et 5.

Logistic regression a été évaluée à l'aide du solveur *liblinear*, avec des pénalités L1 et L2, et une variation du paramètre de régularisation C entre 5 et 100.

Pour les Decision Tree, la profondeur a été ajustée entre 3 et 20, en testant les deux critères classiques d'impureté (*gini*, *entropy*), et une régularisation par *ccp_alpha* a été appliquée entre 0.001 et 0.1.

Le modèle k-NN a été évalué avec des valeurs de k comprises entre 3 et 9, en testant une pondération uniforme ou par la distance, et en comparant les métriques de distance euclidienne et Manhattan.

Enfin, le Naive Bayes gaussien a simplement été régularisé par le paramètre *alpha*, sans ajustement structurel supplémentaire.

4.1.2 RÉSEAUX DE NEURONES

L'analyse des données organisées en séquences temporelles, nécessite des modèles capables de capturer la dynamique temporelle des gestes. Les réseaux de neurones récurrents (RNN) et leurs variantes améliorées, comme les LSTM et GRU, se sont imposés comme des solutions efficaces pour modéliser ces dépendances. Cette partie explore progressivement ces architectures, en commençant par les RNN simples, avant d'aborder les évolutions majeures que représentent les LSTM et GRU, largement utilisés dans la reconnaissance d'activités humaines (HAR).

INTRODUCTION AUX RÉSEAUX DE NEURONES RÉCURRENTS (RNN)

Les réseaux de neurones récurrents (RNN) sont une classe de modèles d'apprentissage profond conçus pour traiter des données séquentielles, où l'ordre des observations est important. Contrairement aux réseaux de neurones classiques, qui traitent les données de manière indépendante, les RNN possèdent une mémoire interne, leur permettant de conserver des informations sur les entrées précédentes.

RESEAUX DE NEURONES RECURRENTS SIMPLES

Un RNN simple repose sur trois éléments principaux : une couche d'entrée, une couche cachée, et une couche de sortie. La couche d'entrée reçoit les données séquentielles, par exemple, une série de mots dans une phrase ou des valeurs dans une série temporelle. La couche cachée traite ces données tout en gardant une *mémoire* des informations des étapes précédentes. Enfin, la couche de sortie génère des prédictions pour chaque étape, comme une étiquette associée à un mot ou une valeur future dans une série (Pascanu et al., 2013).

Les RNN simples fonctionnent en traitant l'information au fil du temps, étape par étape. À chaque moment t , un état caché est calculé à partir de l'entrée actuelle (x_t) et de l'état caché précédent ($h_{\{t-1\}}$). Cette dynamique des états cachés est décrite par l'équation (2.5) (Pascanu et al., 2013) :

$$h_t = f(W_h \cdot h_{\{t-1\}} + W_x \cdot x_t + b) \quad (2.5)$$

où f est une fonction d'activation comme tanh ou ReLU, W_h et W_x sont des poids appris par le modèle, et b est un biais. La sortie, quant à elle, est générée comme indiqué dans l'équation (2.6) (Pascanu et al., 2013):

$$y_t = g(W_y \cdot h_t + b_y) \quad (2.6)$$

Où g est souvent une fonction comme softmax ou sigmoid.

Les RNN peuvent traiter des séquences de longueur variable, ce qui les rend adaptés à des tâches comme la reconnaissance vocale ou la prédiction de séries temporelles. Leur entraînement repose sur la rétropropagation à travers le temps (BPTT), mais cette méthode souffre de la disparition ou de l'explosion des gradients, limitant leur efficacité sur les longues séquences séquence (Pascanu et al., 2013). Ces limites rendent les RNN simples peu adaptés à des activités comme le brossage de dents, où la mémoire des régions précédemment traitées est essentielle. Pour y remédier, des variantes comme les LSTM et GRU ont été introduites. Grâce à leurs mécanismes de régulation de la mémoire, ces modèles gèrent mieux les dépendances à long terme et offrent une analyse plus robuste des signaux inertiels liés au brossage.

LONG SHORT-TERM MEMORY

Les Long Short-Term Memory (LSTM) sont une version améliorée des réseaux de neurones récurrents (RNN) conçue pour résoudre certains problèmes majeurs des RNN, comme la disparition ou l'explosion des gradients. Ces problèmes rendent difficile l'apprentissage sur des séquences longues, car les RNN standards ont du mal à retenir des informations importantes au fil des étapes. Introduits en 1997 par Hochreiter et Schmidhuber, les LSTM se sont imposés comme des outils

essentiels pour des tâches variées, comme le traitement du langage naturel, la reconnaissance vocale ou la prévision des séries temporelles.

Les LSTM se distinguent par leur architecture particulière, qui leur permet de conserver des informations importantes sur une longue période. Cette architecture repose sur des cellules mémoires et des portes qui régulent le passage des informations. Les portes principales incluent la porte d'oubli, qui élimine les données inutiles de l'état précédent, la porte d'entrée, qui détermine quelles nouvelles informations doivent être ajoutées, et la porte de sortie, qui décide des informations à transmettre comme résultat de l'étape en cours (Kawakami, 2008). Ces mécanismes permettent aux LSTM de choisir de manière flexible quelles informations doivent être retenues ou ignorées.

Pour mieux comprendre leur fonctionnement, voici un aperçu des étapes clés dans un LSTM:

Par exemple, la porte d'oubli, qui décide quelles informations de l'état précédent C_{t-1} doivent être oubliées, est calculée selon l'équation (2.7) en utilisant une fonction sigmoid :

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2.7)$$

De manière similaire, la porte d'entrée, qui détermine quelles nouvelles informations doivent être ajoutées, est définie par les équations (2.8) et (2.9) :

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2.8)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (2.9)$$

L'état de la cellule est ensuite mis à jour selon l'équation (2.10), qui combine les informations conservées et les nouvelles données :

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (2.10)$$

Enfin, l'état caché actuel (h_t) utilisé comme sortie et comme entrée pour l'étape suivante est généré comme indiqué dans l'équation (2.11) :

$$h_t = o_t \cdot \tanh(C_t) \quad (2.11)$$

Où $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$

Cette chaîne de calculs permet aux LSTM de mémoriser efficacement les informations utiles tout au long des séquences

Contrairement aux RNN standards, les LSTM parviennent à mieux gérer les dépendances à long terme grâce à leurs cellules mémoire et mécanismes de contrôle. Cette capacité à conserver l'information dans le temps en fait une solution fiable et adaptée aux tâches séquentielles complexes.

Les LSTM sont utilisés dans de nombreux domaines, notamment en traitement du langage, reconnaissance vocale et prévision de séries temporelles. Ils sont aussi employés dans des contextes créatifs, comme la composition musicale ou la génération de contenus visuels à partir de données séquentielles. Malgré leur efficacité, les LSTM restent coûteux en calcul du fait de leur architecture complexe. Ils sont également de plus en plus concurrencés par des modèles comme les transformers, qui exploitent l'attention pour mieux gérer les longues dépendances (Pascanu et al., 2013).

Même si les LSTM ne sont pas parfaits, ils répondent à de nombreux problèmes que les RNN standards ne peuvent résoudre. Dans le cadre de notre étude, l'utilisation des LSTM est particulièrement adaptée car notre problématique implique l'analyse de séquences longues, couvrant potentiellement tout un cycle de brossage d'une région à l'autre. Les LSTM permettent ainsi de conserver la mémoire des zones précédemment brossées et d'améliorer la classification des mouvements dans le temps.

GATED RECURRENT UNIT

Les Gated Recurrent Units (GRU) sont une version simplifiée des réseaux de neurones récurrents (RNN), introduites par (Cho et al., 2014). Elles ont été conçues pour surmonter certains

problèmes majeurs des RNN simples, tels que la disparition des gradients. La disparition des gradients, un problème courant dans les RNN standards, se produit lorsque les mises à jour du modèle deviennent si petites qu'elles n'ont plus d'effet, rendant difficile l'apprentissage sur des séquences longues. Contrairement aux LSTM, les GRU adoptent une architecture plus légère tout en offrant des performances comparables pour capturer des dépendances temporelles (Lin et al., 2022). Cela les d'autant plus pertinents dans des domaines variés tels que le traitement du langage naturel, les séries temporelles et la reconnaissance vocale.

Les GRU simplifient le fonctionnement des RNN en utilisant uniquement deux portes principales : la porte de réinitialisation et la porte de mise à jour. Ces portes contrôlent le flux d'informations entre les étapes temporelles, permettant de déterminer quelles parties de l'état précédent doivent être conservées, oubliées ou modifiées. Contrairement aux LSTM, les GRU n'utilisent pas de cellule mémoire distincte, ce qui réduit leur complexité tout en maintenant leur efficacité (Lin et al., 2022).

La porte de réinitialisation, décrite par l'équation (2.12) contrôle combien d'informations de l'état caché précédent doivent être oubliées. Elle est calculée à l'aide de la fonction sigmoid, qui produit une valeur entre 0 et 1 (Lin et al., 2022):

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (2.12)$$

Ce mécanisme permet d'effacer sélectivement des informations inutiles pour chaque étape temporelle.

La porte de mise à jour module la proportion d'informations anciennes à conserver par rapport aux nouvelles informations. Elle est également calculée avec une fonction sigmoid comme définie par l'équation (2.13) (Lin et al., 2022):

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (2.13)$$

Cela permet au modèle d'ajuster dynamiquement l'équilibre entre les données passées et présentes, selon les besoins de la tâche.

Cette dépendance temporelle est capturée par l'équation (2.14), qui calcule l'état candidat en générant en combinant l'entrée actuelle et l'état précédent pondéré par la porte de réinitialisation (Lin et al., 2022). Le symbole \odot représente ici le produit de Hadamard, c'est-à-dire une multiplication élément par élément entre les vecteurs r_t et h_{t-1} . Cette opération permet à la porte de contrôler quelles composantes de la mémoire passée doivent être prises en compte dans le calcul du nouvel état :

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h) \quad (2.14)$$

Enfin, l'équation (2.15) décrit comment l'état caché actuel est obtenu en combinant linéairement l'état précédent et l'état candidat, pondérés par la porte de mise à jour (Lin et al., 2022):

$$h_t = z_t \cdot h_{t-1} + (1 - z_t) \cdot \tilde{h}_t \quad (2.16)$$

Les GRU se démarquent par leur capacité à atténuer la disparition des gradients grâce à un contrôle efficace du flux d'information. Plus légères que les LSTM, elles sont aussi plus rapides et moins coûteuses en calcul (Lin et al., 2022).

Les GRU sont utilisées dans de nombreux domaines, notamment en traitement du langage, reconnaissance vocale, prévision de séries temporelles et modélisation du comportement utilisateur dans des applications interactives. Bien que plus rapides, les GRU restent moins efficaces que les LSTM pour modéliser des dépendances longues, en raison de l'absence de cellule mémoire. Leur simplicité en fait néanmoins un bon compromis dans de nombreuses applications. Dans notre étude, elles se révèlent surtout utile pour traiter les séquences de brossage en temps réel ou dans des environnements embarqués, offrant des performances proches des LSTM pour un coût de calcul réduit.

IMPLÉMENTATION DES RNN

Avant l'entraînement, les données brutes extraites des capteurs subissent une série de prétraitements, comme décrit dans la section méthodologie. Pour exploiter efficacement la capacité des RNN à modéliser les dépendances temporelles, notre pipeline n'inclut pas l'extraction des caractéristiques dans la section méthodologie. Une fois normalisées, les données sont découpées en fenêtres de 50 pas de temps, puis réorganisées sous forme de tenseurs tridimensionnels de type (nombre d'échantillons, fenêtre, nombre de caractéristiques). Cette représentation permet de préserver l'ordre temporel des séquences tout en assurant leur compatibilité avec les réseaux récurrents.

Chaque fenêtre étant susceptible de contenir plusieurs étiquettes successives, une agrégation est réalisée : l'étiquette dominante est attribuée à l'ensemble de la séquence. Cette méthode vise à éviter les changements abrupts de classe au sein d'une même fenêtre et favorise un apprentissage plus stable. Une fois les données ainsi préparées, elles peuvent être directement utilisées pour entraîner les modèles récurrents. L'étape suivante consiste donc à initier le processus d'apprentissage.

Chacun de ces modèles comprend une couche récurrente de 128 unités (avec activation tanh), suivie d'un dropout de 30 % pour prévenir le surajustement, puis d'une couche dense avec activation softmax pour la classification multi-classes. Tous les modèles sont optimisés avec l'algorithme Adam et entraînés avec la fonction de perte *sparse categorical crossentropy*, sur 20 époques avec une taille de lot de 32.

4.1.3 TRANSFORMER

Avant de décrire l'usage que nous faisons du Transformer dans notre étude, il est utile de revenir sur ses origines, son fonctionnement interne et les raisons pour lesquelles il se prête particulièrement bien à l'analyse de données séquentielles issues de capteurs. Une attention

particulière est ensuite portée à son adaptation à la reconnaissance d'activités humaines, en lien avec notre problématique.

CONTEXTE ET ORIGINE DU TRANSFORMER

Le Transformer est une architecture de réseau de neurones introduite en 2017 par Vaswani et ses collègues (Vaswani et al., 2017) lors de la conférence NIPS (Neural Information Processing Systems). Cet article, intitulé *Attention is All You Need*, a marqué un tournant dans le domaine du traitement automatique du langage (NLP) en proposant une approche totalement nouvelle pour le traitement des séquences. Contrairement aux modèles classiques comme les RNN, le Transformer repose exclusivement sur un mécanisme appelé self-attention. Cette innovation permet d'atteindre des performances de pointe tout en réduisant considérablement le coût de calcul et le temps d'entraînement.

Avant le Transformer, les modèles séquentiels comme les RNN traitaient les données dans un ordre temporel strict, ce qui limitait la parallélisation et la gestion des dépendances longues. Le Transformer, en revanche, traite l'ensemble de la séquence en parallèle tout en capturant les relations contextuelles entre les éléments, quelle que soit leur position.

Bien que développé initialement pour le traitement du langage, le Transformer s'applique particulièrement bien aux séquences temporelles multivariées, comme celles issues de capteurs. Dans notre cas, l'attention auto-référencée permet au modèle d'associer des motifs de mouvement à différentes positions temporelles dans une fenêtre d'observation. Cette capacité à capturer des relations complexes, indépendamment de leur position dans le temps, est cruciale pour la reconnaissance des gestes de brossage. De plus, le traitement parallèle s'intègre naturellement à notre stratégie d'analyse par fenêtres glissantes, optimisant ainsi l'efficacité du modèle.

ARCHITECTURE

L'architecture du Transformer suit une structure classique encodeur–décodeur, mais dans notre étude, nous nous concentrons uniquement sur la partie encodeur. Celui-ci est constitué d'un empilement de six couches identiques, chacune comportant deux sous-couches principales :

- une couche de *self-attention* multi-tête,
- une couche de réseau *feed-forward* entièrement connecté.

La Figure 14 illustre l'architecture complète du Transformer telle que proposée par Vaswani et al. (2017). Dans notre cas, seule la partie encodeur (encadrée en rouge) est exploitée. Chaque sous-couche est précédée d'une connexion résiduelle et d'une normalisation de couche (*layer normalization*), ce qui facilite la propagation du gradient et améliore la stabilité de l'apprentissage.

Le mécanisme d'attention repose sur trois vecteurs linéaires : requêtes (Q), clés (K) et valeurs (V). Le produit scalaire entre Q et K permet de mesurer la pertinence de chaque élément de la séquence, et les valeurs V sont pondérées en conséquence. La version multi-head attention applique ce mécanisme en parallèle sur plusieurs sous-espaces pour capturer différents types de dépendances contextuelles.

Le réseau *feed-forward*, appliqué individuellement à chaque position, se compose généralement de deux couches linéaires séparées par une activation non linéaire (telle que ReLU ou GELU), ce qui apporte une capacité d'abstraction supplémentaire.

Enfin, pour intégrer une notion de position (puisque le Transformer ne repose ni sur la récurrence ni sur la convolution), un encodage positionnel est ajouté aux vecteurs d'entrée. Celui-ci est calculé de manière déterministe à l'aide de fonctions sinusoïdales, permettant au modèle de distinguer l'ordre relatif des éléments dans une séquence.

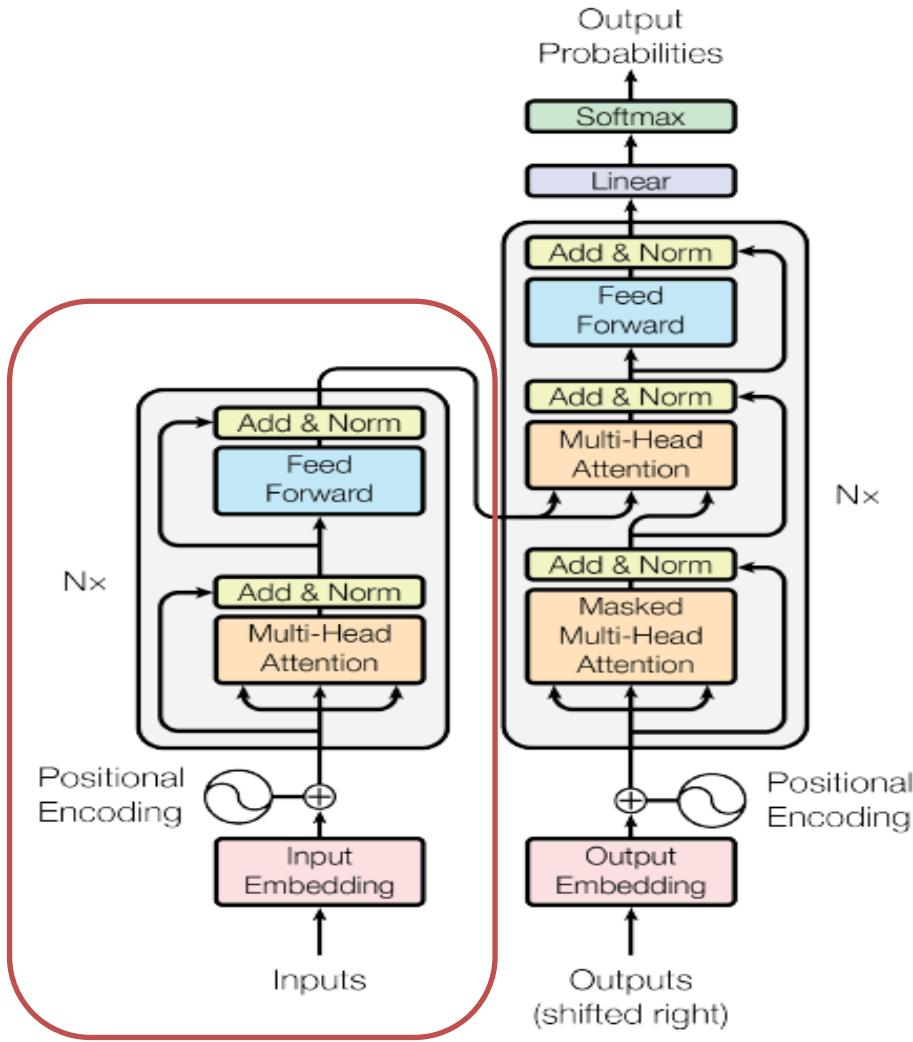


Figure 14: Architecture complète du Transformer incluant l'encodeur et le décodeur. Dans le cadre de cette étude, seule la partie encodeur a été utilisée.

@ (Vaswani et al., 2017)

ADAPTATION À LA CLASSIFICATION EN RECONNAISSANCE D'ACTIVITÉS HUMAINES (HAR)

Dans le cadre de la reconnaissance d'activités humaines (HAR), le transformer présente plusieurs avantages significatifs. Il est capable de modéliser des dépendances temporelles longues et de traiter les séquences dans leur globalité, sans contrainte d'ordre, tout en permettant une parallélisation complète du traitement.

Cette capacité est particulièrement pertinente pour l'analyse de gestes fins et complexes, comme ceux liés au brossage dentaire, où des régions proches anatomiquement (par exemple, les molaires inférieures et supérieures d'un même côté) peuvent produire des signaux similaires. Grâce au mécanisme d'attention multi-tête, le Transformer contextualise chaque instant de la séquence en tenant compte de l'ensemble de la fenêtre temporelle, ce qui améliore la différenciation entre les régions brossées.

Par ailleurs, son fonctionnement s'adapte naturellement au format de données segmentées utilisé dans notre étude, basé sur des fenêtres temporelles glissantes, ce qui en fait un modèle efficace pour l'analyse de séquences multivariées de capteurs inertIELS.

MODÈLE TRANSFORMER ADAPTÉ

Pour répondre à la tâche de classification, nous avons construit une architecture Transformer personnalisée inspirée directement de l'encodeur du modèle Transformer (Vaswani et al., 2017). Comme dans l'architecture originale, notre modèle repose sur un empilement de blocs d'attention multi-tête suivis de réseaux *feed-forward*, chacun encadré par une connexion résiduelle et une normalisation de couche. Toutefois, plusieurs adaptations ont été introduites pour mieux répondre aux contraintes liées à la nature des données (séries temporelles multivariées, contexte de capteurs portés). L'entrée du modèle est une séquence temporelle de vecteurs de caractéristiques, correspondant à une fenêtre fixe de 50 échantillons. Nous initialisons ensuite un vecteur nul, dont les deux premières dimensions sont utilisées pour injecter deux encodages auxiliaires : la dextérité du participant (gauche ou droitier) et le type de brosse utilisée (électrique ou manuelle). Cette étape permet au modèle d'intégrer une information dès le début du traitement de la séquence. Ces encodages contextuels peuvent un rôle important dans la performance du modèle, les signaux inertIELS étant en effet influencés par des facteurs tels que la main utilisée ou la technologie de la brosse. En injectant ces informations, le modèle est guidé dès les premières couches à interpréter les signaux dans leur contexte réel. Cela peut réduire l'ambiguïté, améliorer la généralisation à différents profils utilisateurs et accélérer l'apprentissage. Ensuite, la séquence étendue passe par

une projection linéaire via une couche Dense, qui transforme chaque vecteur en une dimension cachée commune (hidden_size = 128). Afin de préserver l'ordre temporel des données, un encodage de position sinusoïdal est ajouté à chaque vecteur de la séquence. Cet encodage injecte une notion de position dans l'espace d'attention, ce que le Transformer ne capte pas naturellement.

L'architecture comprend ensuite un empilement de quatre blocs Transformer Encoder. Chaque bloc se compose d'une couche *de Multi-Head Self-Attention* permettant à chaque position d'interagir avec toutes les autres, suivie d'une normalisation de couche avec connexion résiduelle. Puis, une sous-couche *Feed-Forward* composée de deux couches *Denses* (la première avec activation GELU) traite chaque position indépendamment, suivie là aussi d'une normalisation et d'une connexion résiduelle. Enfin, la couche de sortie est une couche *Dense* avec une activation softmax, qui retourne une probabilité pour chaque région de brossage possible.

4.2 MESURES DE PERFORMANCES

L'évaluation des modèles dans notre étude repose sur plusieurs mesures qui permettent d'analyser leur capacité à bien identifier les régions de brossage en fonction des données capturées par les capteurs. Pour mieux comprendre les performances et les erreurs du modèle, nous utilisons des indicateurs globaux comme la justesse. Mais les variations de mouvements et les différences entre utilisateurs compliquent l'évaluation du modèle, rendant une simple justesse insuffisante pour garantir de bonnes performances sur toutes les régions. Pour une analyse plus précise, nous combinons plusieurs mesures telles que :

La justesse (Accuracy), qui mesure la proportion d'observations correctement classées et donne un aperçu global de la performance.

Le Rappel (Recall) indique la capacité du modèle à identifier correctement chaque région de brossage. Nous utilisons cette mesure afin de s'assurer que toutes les zones sont bien détectées, même celles qui sont moins représentées dans les données.

La Précision (Precision) quant à elle, mesure la proportion des prédictions correctes parmi celles effectuées pour une région donnée.

Le Score F1, enfin qui est une moyenne harmonique entre la précision et le rappel. Il est particulièrement utile dans notre étude, car notre ensemble de données comporte des classes avec des distributions inégales. Nous avons optimisé nos modèles en nous basant principalement sur cette mesure pour assurer un bon équilibre entre les bonnes détections et la réduction des erreurs.

Matrice de confusion : Analyse des erreurs

La matrice de confusion permet de visualiser les performances du modèle en comparant les classes réelles (lignes) aux classes prédictes (colonnes). Les valeurs sur la diagonale indiquent les classifications correctes, tandis que les valeurs hors diagonale révèlent les erreurs les plus fréquentes. Un exemple illustratif d'une matrice de confusion pour quatre classes est présenté à la Figure 15, afin d'en faciliter la compréhension.

Dans notre étude, elle met en évidence les régions de brossage bien détectées, ainsi que les confusions fréquentes, notamment entre des zones proches comme les molaires supérieures et inférieures. Elle permet également de repérer d'éventuels déséquilibres ou difficultés de classification pour certaines régions.

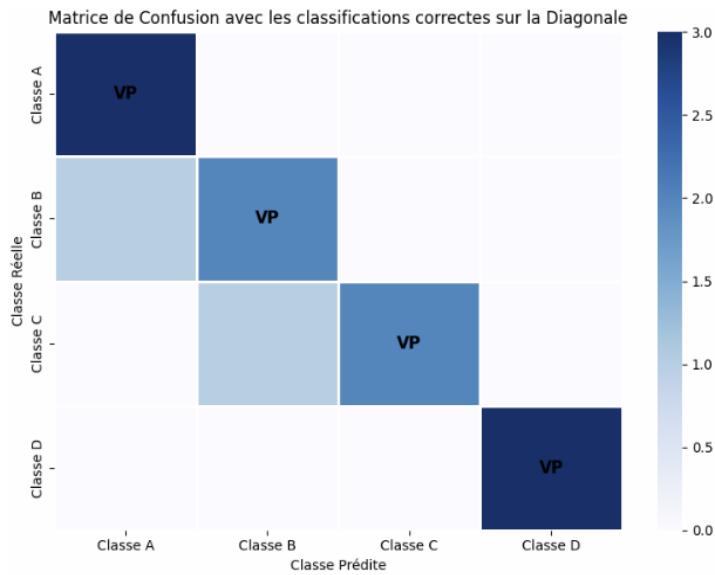


Figure 15: Exemple illustratif d'une matrice de confusion pour 4 classes. Les cellules de la diagonale principale, annotées VP (vrais positifs), indiquent les prédictions correctes pour chaque classe.

©Sawadogo Wendenda Aniisah

4.3 RÉSULTATS DE CLASSIFICATION

Cette section vise à confronter les algorithmes testés à des situations réalistes, en tenant compte à la fois de la distribution déséquilibrée des classes, des styles propres à chaque utilisateur, et des changements entre sessions. Pour cela, plusieurs stratégies d'évaluation sont mobilisées, allant d'une validation croisée à des validations plus exigeantes comme LOSO et OSO. Ces tests permettent de révéler non seulement les meilleures performances moyennes, mais aussi la stabilité, la sensibilité aux biais et la capacité de généralisation des modèles.

4.3.1 VALIDATION CROISÉE

Dans notre étude, nous avons utilisé la validation croisée pour évaluer la robustesse de nos différents modèles et nous assurer qu'ils fonctionnent bien sur de nouvelles données. Plutôt que de simplement diviser les données en un ensemble d'entraînement et un ensemble de test, cette méthode permet d'exploiter toutes les données disponibles tout en évitant que le modèle s'adapte

trop aux spécificités d'un seul jeu de données. Nous avons opté pour une validation croisée stratifiée k-plis avec k=5. Concrètement, les données sont divisées en cinq parties à peu près égales, tout en s'assurant que chaque groupe garde la même répartition des classes que l'ensemble initial. Le modèle est entraîné sur quatre de ces plis et évalué sur le cinquième, en répétant le processus jusqu'à ce que chaque pli ait été utilisé comme ensemble de test. Cette approche a l'avantage d'être plus équilibrée et de ne pas trop dépendre d'un seul découpage des données. Dans notre cas, cela présente un avantage, car certaines régions de brossage sont moins représentées que d'autres. En s'assurant que chaque pli conserve une répartition correcte des classes, on évite que le modèle ne privilégie trop certaines régions au détriment des autres. Enfin, pour optimiser les hyperparamètres, nous avons combiné RandomizedSearchCV et GridSearchCV, en utilisant *cross_val_score* avec cette même validation croisée à 5 plis, afin d'évaluer les performances des modèles sélectionnés

Pour nos modèles de réseaux de neurones et transformer, nous avons implémenté la validation croisée avec la fonction TimeSeriesSplit(). Cette méthode divise les données en plusieurs parties tout en gardant leur ordre. À chaque itération, le modèle est entraîné sur les données passées et évalué sur les données futures, ce qui permet de simuler un apprentissage en conditions réelles. Les scores moyens obtenus sur l'ensemble des plis offrent une évaluation fiable et représentative des performances du modèle.

RÉSULTATS

Tableau 3: Performances des modèles classiques en validation croisée

	Justesse	Précision	F1-score	Rappel
Logistic Regression	39,56 %	40,29 %	39,53 %	39,56 %
Decision Tree	52,91 %	53,02 %	52,31 %	52,19 %
Random Forest	98,48 %	98,48 %	98,46 %	98,45 %
K-Nearest Neighbors (KNN)	97,58 %	97,58 %	97,58 %	96,58 %
Support Vector Classifier (SVC)	44,38 %	44,53 %	44,51 %	44,40 %
Naïve Bayes	35,34 %	35,09 %	34,77 %	35,34 %
Gradient Boosting (XGBoost)	97,73 %	97,86 %	97,58 %	97,96 %
Gradient Boosting (CatBoost)	87,09 %	87,07 %	87,14 %	87,09 %
Gradient Boosting (LightGBM)	78,99 %	79,00 %	79,00 %	79,00 %

Les scores présentés dans le Tableau 3 permettent de comparer les performances des principaux modèles classiques testés en validation croisée. Les meilleurs résultats ont été obtenus avec random forest, XGBoost et KNN, qui affichent tous des scores F1 supérieurs à 97 %. Ils maintiennent une homogénéité remarquable entre les différentes classes traduisant leur aptitude à capturer la complexité et la diversité des données, comme l'illustre la matrice de confusion du modèle de random forest présentée dans la Figure 16. À l'inverse, plusieurs modèles affichent des performances nettement plus faibles. Naïve Bayes, avec un F1-score moyen de 35 %, présente une forte inégalité entre les classes. Decision Tree, avec un F1-score de 53 %, reste stable en validation croisée mais limité par sa structure simple, incapable de modéliser toute la variabilité des gestes. Le Support Vector Classifier (SVC) atteint un F1-score de 44 %, mais présente une forte disparité entre classes. En effet, comme on peut le voir sur la Figure 17 certaines, comme la classe 1, sont correctement reconnues (plus de 80 % sur le F1-score), tandis que d'autres, telles que la classe 10, sont inférieurs à 30 % indiquant des difficultés à séparer correctement certaines catégories. Quant à logistic regression, sa précision autour de 40 % reflète son incapacité à modéliser des relations

complexes au-delà de frontières linéaires. Ces résultats confirment la nécessité d'utiliser des modèles capables de traiter efficacement la non-linéarité et la structure complexe des données pour obtenir des performances satisfaisantes.

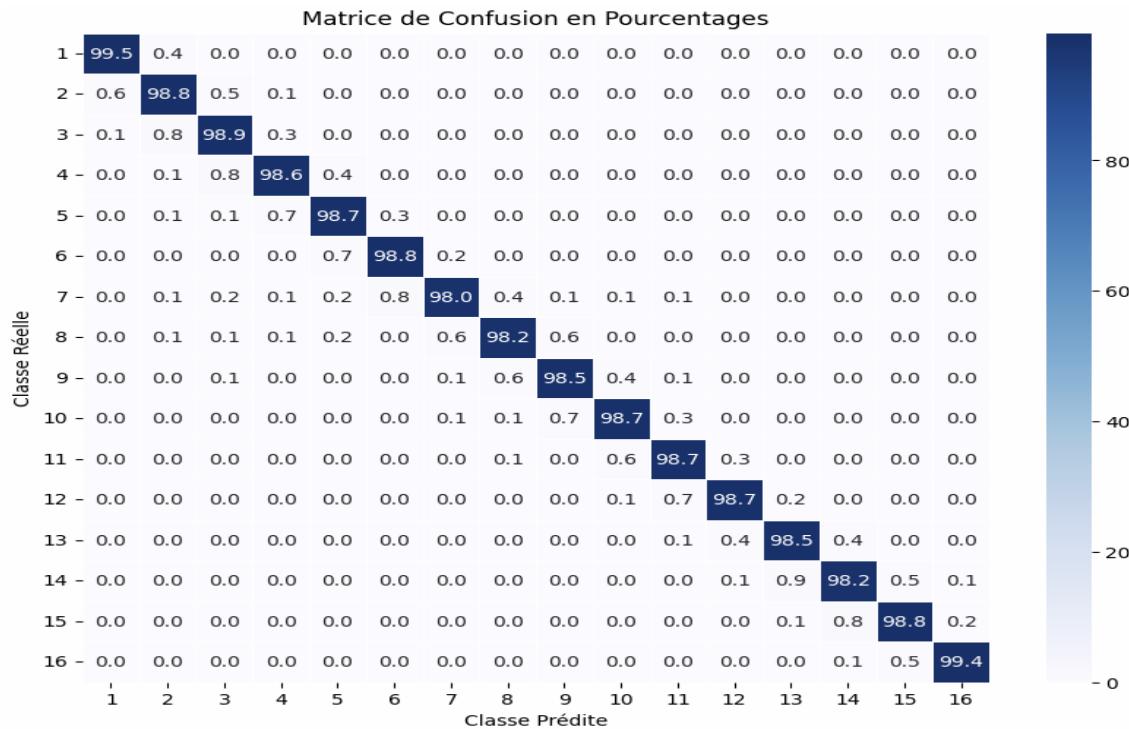


Figure 16: Matrice de confusion (%) pour l'évaluation de notre modèle de random forest

©Sawadogo Wendenda Aniisah

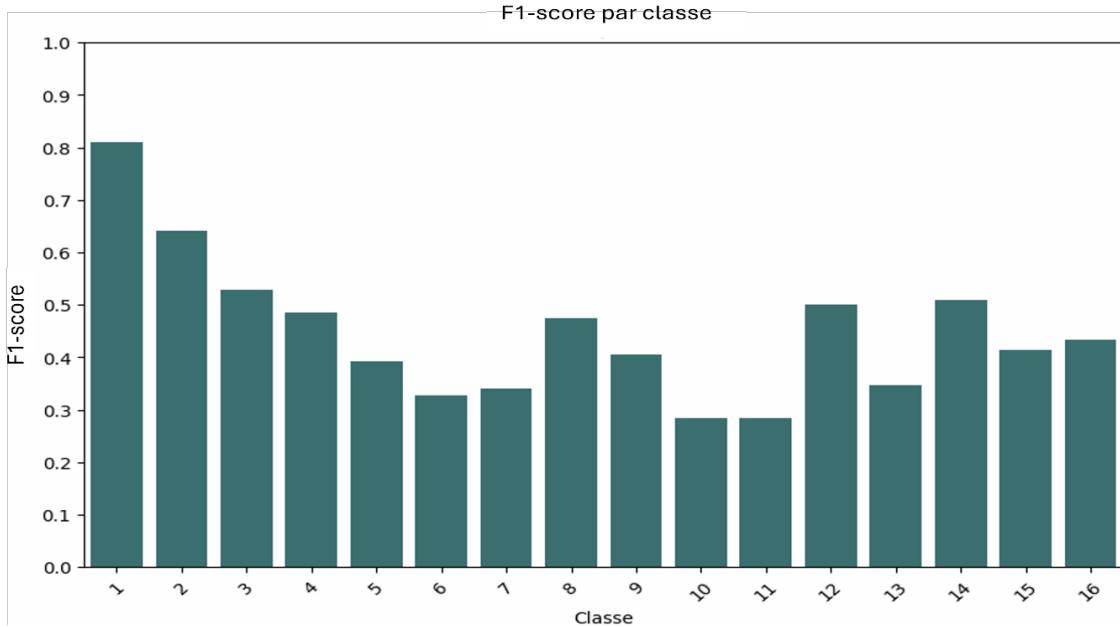


Figure 17: F1-score par classe (%) pour l'évaluation de notre modèle SVC

©Sawadogo Wendenda Aniisah

Tableau 4:Performances des modèles récurrents en validation croisée

	Justesse	Précision	Rappel	F1-score
RNN Simple	97,95 %	97,99 %	97,95 %	97,94 %
LSTM	99,79 %	99,79 %	99,79 %	99,79 %
GRU	99,65 %	99,65 %	99,65 %	99,65 %
TRANSFORMER	96,17 %	96,33 %	96,17 %	96,14 %

D'après les résultats résumés dans le Tableau 4, les modèles de réseaux de neurones récurrents, à savoir le RNN simple, le LSTM et le GRU, ont tous obtenu d'excellentes performances. Le LSTM se distingue avec un score final de 99,79 %, pour chaque mesure de performance, ce qui reflète une bonne stabilité. Le GRU suit de près avec des résultats très similaires (99,65 % sur l'ensemble des mesures) et une bonne capacité de généralisation, malgré quelques variations initiales. Le RNN simple, pour sa part, atteint un F1-score de 97,95 %, ce qui est inférieure à celle

des deux autres architectures, mais reste globalement performant. Une analyse approfondie pour chacun des algorithmes montre que les erreurs de classification sont devenues marginales, la majorité des prédictions se concentrant sur la diagonale principale.

Le modèle Transformer complet, incluant l'encodage positionnel, les informations contextuelles et un empilement de quatre couches d'encodeur à deux têtes d'attention, a montré des performances très satisfaisantes. Avec une précision de 96,33 %, un F1-score de 96,14 % et des chiffres sensiblement égaux pour les autres mesures. Le modèle s'est révélé capable de capturer avec précision les dynamiques gestuelles du brossage. Ces résultats témoignent de l'efficacité combinée des mécanismes attentionnels et des encodages temporels et contextuels dans la représentation des séquences. L'architecture Transformer se distingue ainsi par sa capacité à extraire des patterns complexes à partir des données inertielles, tout en maintenant une bonne stabilité au cours des différentes partitions.

4.3.2 MÉTHODE LEAVE-ONE-SUBJECT-OUT (LOSO)

Un modèle performant en validation croisée classique ne garantit pas nécessairement une bonne généralisation à de nouveaux utilisateurs. En effet, dans ce type de validation, les données d'un même individu peuvent être utilisées à la fois pour l'entraînement et le test, ce qui peut fausser l'évaluation réelle des capacités du modèle. Pour mieux évaluer la robustesse des approches retenues, nous avons donc mis en place une validation Leave-One-Subject-Out (LOSO), appliquée spécifiquement aux trois modèles les plus performants : random forest, LSTM et transformer. La validation Leave-One-Subject-Out (LOSO) fonctionne différemment des méthodes classiques. Au lieu de diviser aléatoirement les données en plusieurs groupes, on entraîne le modèle en excluant complètement un participant à chaque itération, puis on le teste sur les données de ce participant. On recommence jusqu'à ce que chaque personne ait servi une fois comme test. Après avoir testé sur tous les participants, les scores globaux sont calculés en moyennant les mesures obtenues.

Le processus suit globalement les étapes définies dans la section méthodologie, avec des ajustements spécifiques à chaque algorithme. Pour le modèle LSTM, les données sont organisées sous forme de fenêtres temporelles afin de capturer les séquences d'actions et de mieux exploiter la structure des mouvements. Pour nos trois modèles RF, LSTM et Transformer, la stratégie d'évaluation Leave-One-Subject-Out (LOSO) a été implémentée à l'aide de la classe LeaveOneGroupOut de Scikit-learn. Cela permet d'assurer que chaque sujet est tour à tour utilisé comme ensemble de test, tandis que les autres servent à l'entraînement. Pour le modèle Transformer, l'encodage positionnel est toujours ajouté selon la méthodologie principale pour capturer la structure temporelle.

L'évaluation des performances repose sur le test du modèle sur le sujet exclu, avec le calcul de mesures comme la justesse, le F1-score, la précision et le rappel. Cette approche assure une évaluation rigoureuse et réaliste de la capacité du modèle à généraliser sur de nouveaux utilisateurs. Cette méthode s'avère utile ici, car l'objectif est de voir si le modèle fonctionne bien sur de nouveaux utilisateurs. Comme les habitudes de brossage varient d'une personne à l'autre, il est important de tester si le modèle peut s'adapter à des individus qu'il n'a jamais vus. L'inconvénient, c'est que cette approche demande plus de temps de calcul, car on doit entraîner le modèle autant de fois qu'il y a de participants.

RÉSULTATS

Tableau 5: Statistiques descriptives des scores F1 par modèle en LOSO

	Minimum	Maximum	Écart-type	Moyenne
RF	13 %	56 %	11,52 %	34,05 %
LSTM	73,97 %	99,69 %	7 %	94,17 %
Transformer	92,30 %	99,38 %	2,2 %	96,94 %

La validation LOSO met en lumière les faiblesses du modèle random forest. Alors qu'il affiche d'excellentes performances en validation croisée classique, ses scores chutent fortement en validation LOSO, avec un F1-score moyen de 34 %. De plus, une analyse plus détaillée montre une grande hétérogénéité des performances selon les sujets, comme l'indique un écart-type élevé de 11,52 % dans le Tableau 5. Certains sujets atteignent plus de 50 % de justesse, tandis que d'autres

comme le sujet 7 (14 %) et le sujet 15 (19 %), ont des performances catastrophiques (Figure 18). Le modèle semble avoir de la peine à extraire des motifs de brossage suffisamment généraux pour être applicables à de nouveaux utilisateurs. En d'autres termes, il semble apprendre des spécificités propres aux sujets présents dans l'entraînement, mais ne parvient pas à bien transférer cet apprentissage à des personnes jamais vues auparavant.

À l'opposé, les réseaux de neurones récurrents montrent une bien meilleure stabilité en validation LOSO. En effet, comme on peut l'apercevoir sur le Tableau 5, LSTM affiche un F1-score moyen élevé et un écart-type de 7%, avec des performances stables dépassant souvent 97 %, atteignant même 99,69 % pour certains sujets. Toutefois, une variabilité subsiste : pour certains individus, comme le sujet 3 (74,10 %) et le sujet 20 (82,85 %), les performances sont nettement inférieures comme on peut le constater sur la Figure 19, suggérant une difficulté du modèle à généraliser pour certains styles de brossage.

Par ailleurs, nous avons également évaluer le modèle basé sur l'architecture Transformer en validation LOSO. Les résultats obtenus sont globalement très bons, avec un F1-score de 96,94 % et un écart type de seulement 2,2%. Ces performances montrent que le Transformer arrive à bien généraliser à des sujets qui n'ont pas été vus pendant l'entraînement. Même si certaines personnes, comme les sujets 2 ou 21, obtiennent des scores un peu plus faibles (autour de 93 %), la plupart des sujets dépassent largement les 96 %, et certains atteignent même des performances proches de la perfection, comme les sujets 7 ou 9 avec plus de 99 % de F1-score comme on peut le constater sur la Figure 20. Cela indique que le Transformer réussit à extraire des représentations plus générales du geste de brossage, ce qui lui permet de mieux s'adapter à la diversité des utilisateurs.

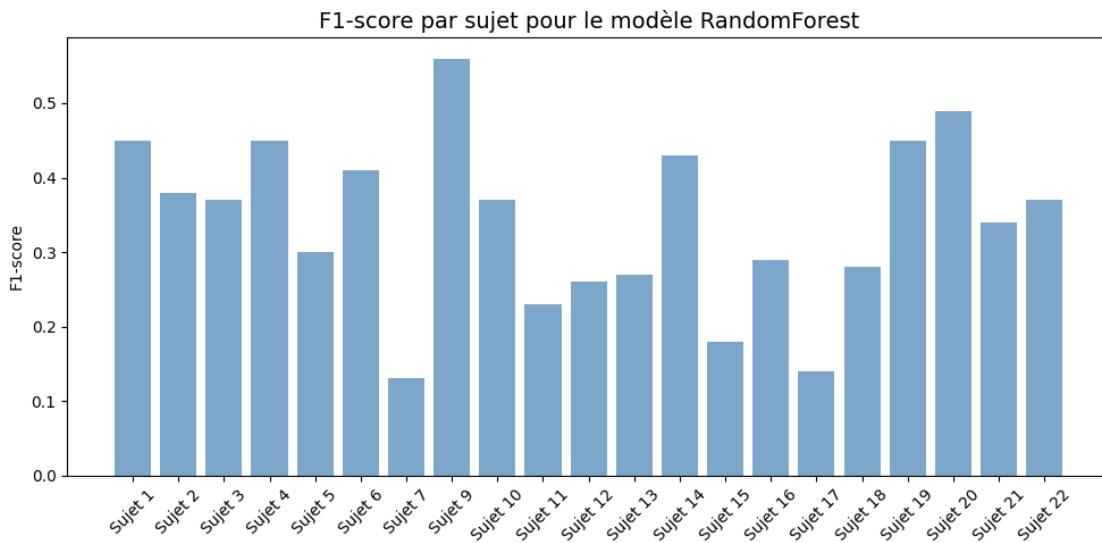


Figure 18: F1-score de l'évaluation du Leave-One-Subject-Out (LOSO) avec le random forest

©Sawadogo Wendenda Aniisah

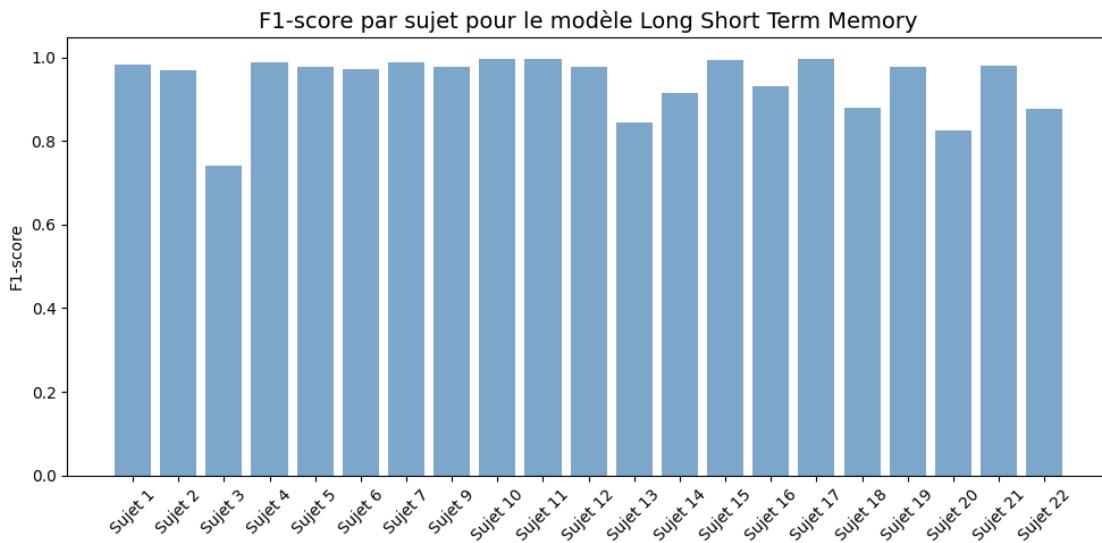


Figure 19: F1-score de l'évaluation du Leave-One-Subject-Out (LOSO) avec LSTM

©Sawadogo Wendenda Aniisah

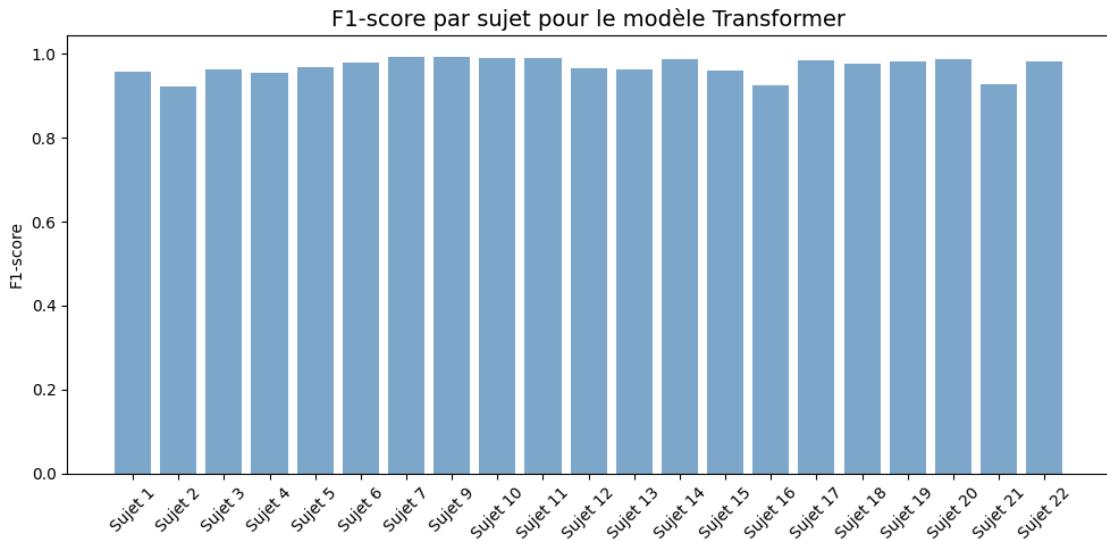


Figure 20: F1-score de l'évaluation du Leave-One-Subject-Out (LOSO) avec Transformer

©Sawadogo Wendenda Aniisah

4.3.3 MÉTHODE ONE-SESSION-OUT (OSO)

En complément de la validation inter-sujets, nous avons mis en place une stratégie One-Session-Out (OSO) afin d'évaluer la capacité des modèles à généraliser entre différentes sessions d'un même utilisateur. Concrètement, les données de chaque participant sont divisées en sessions distinctes. Chaque session est identifiée par un code structuré qui encode la configuration expérimentale, l'identifiant du participant et le numéro de la session. Lors de chaque itération, une session est entièrement exclue et utilisée comme ensemble de test, tandis que les autres servent à l'entraînement. Ce processus est appliqué à l'ensemble des sessions du sujet. Dans notre cas, il s'agit d'un OSO intra-sujet, où seules les données d'un même individu sont impliquées dans chaque itération. Cela se distingue d'un OSO inter-sujets, qui impliquerait de tester sur une session d'un participant tout en s'entraînant sur celles d'un autre. Le choix d'un OSO intra-sujet est motivé par notre objectif principal : évaluer la stabilité temporelle du modèle pour un utilisateur donné, c'est-à-dire sa capacité à reconnaître correctement les régions de brossage au fil du temps, malgré des variations naturelles entre les séances (par exemple, changement de posture ou de rythme).

Dans notre implémentation, une session test contient les données issues des deux capteurs la brosse à dents intelligente et la montre inertielle qui sont fusionnées pour chaque session afin de capturer toute la richesse des données disponibles. Lors de l'entraînement, la session utilisée pour le test est strictement exclue de tout processus d'optimisation. En particulier, aucune mesure issue de cette session n'est utilisée pour ajuster les poids du modèle ou pour déclencher l'arrêt précoce (*EarlyStopping*). Les modèles sont entraînés indépendamment pour chaque session testée, et les performances (justesse, F1-score, précision, rappel) sont extraites et ensuite moyennées sur l'ensemble des sessions pour obtenir une évaluation globale.

RÉSULTATS

Tableau 6: Statistiques descriptives des scores F1 par modèle en OSO

	Minimum	Maximum	Écart-type	Moyenne
RF	19,27 %	96,12 %	19,98 %	62,79 %
LSTM	14 %	98,34 %	31,67 %	64,28 %
Transformer	79,31 %	99,99 %	4,23 %	97,45 %

Du côté du modèle random forest, les résultats en OSO sont très inégaux. Si certaines sessions parviennent à des scores acceptables, la majorité reste en dessous des attentes. Plusieurs sessions, comme S1_S10_S2 ou S1_S11_S5, présentent des F1-scores inférieurs à 40 %, ce qui reflète une réelle difficulté à s'adapter aux variations entre les enregistrements. À l'inverse, quelques cas isolés, notamment S2_S2_S3 ou S2_S14_S2, atteignent des scores bien plus élevés, parfois au-delà des 90 %. Ces bonnes performances restent rares et semblent dépendre de conditions favorables un style de brossage stable ou peu de différences entre les sessions, par exemple. Globalement, le modèle peine à reconnaître les schémas propres à chaque individu lorsque ceux-ci ne sont pas présents dans les données d'entraînement. Cette variabilité se reflète dans un écart-type élevé de 19,98 %, comme résumé dans le Tableau 6, traduisant des performances très dispersées d'une session à l'autre.

Les performances du LSTM en validation One-Session-Out (OSO) sont nettement plus faibles que celles observées en LOSO. Certaines, comme S1_S1_S4 ou S2_S2_S3, affichent d'excellentes performances (au-delà de 94 %), tandis que d'autres, notamment S2_S16_S2, S1_S13_S4 ou

S1_S14_S1, atteignent à peine 20 %, voire moins dans certains cas extrêmes. Cette tendance à une performance plus faible en OSO qu'en LOSO va à l'encontre des observations habituellement rapportées dans la littérature. La dégradation des performances peut toutefois s'expliquer par la nature même de la validation OSO. Ici, le modèle doit faire face aux différences entre deux enregistrements réalisés à des moments différents, avec des variations possibles dans les gestes, la posture ou le rythme du brossage. L'écart-type très élevé de 31,67 %, illustré dans le Tableau 6, met en évidence l'instabilité du modèle selon les sessions testées.

En revanche, les résultats obtenus avec le modèle Transformer indiquent une bonne cohérence intra-sujet. Dans la majorité des cas, les F1-scores dépassent les 97 %, avec une moyenne globale de 97,89 %. Ce niveau de performance suggère que le modèle est capable de reconnaître des motifs de brossage caractéristiques d'un individu, même lorsque ceux-ci proviennent de sessions différentes. Bien que quelques dégradations ponctuelles aient été observées, l'écart-type très faible de 4,23 % comme le montre Tableau 6 le confirme une stabilité remarquable du modèle à travers les enregistrements.

Après avoir mis en évidence la supériorité globale du Transformer (voir Tableau 6), il est pertinent d'examiner de plus près les sessions où ses performances s'écartent de la tendance générale. En effet, certaines sessions, telles que S1_S15_S1, S2_S6_S1 et S2_S21_S3, présentent des F1-scores inférieurs à 85 %, ce qui contraste fortement avec la tendance générale. Une observation plus détaillée de ces sessions révèle des erreurs localisées, souvent concentrées sur des régions spécifiques. Pour S2_S6_S1 sur la Figure 21(a), ce sont les premières régions du geste (classes 1 et 2) qui déstabilisent le modèle. Les F1-scores y sont très faibles, proches de zéro, alors que les autres régions sont correctement identifiées. Dans le cas de S1_S15_S1 représentée sur la Figure 21(b), les classes 15 et 16 posent particulièrement un problème. Le modèle parvient difficilement à détecter la classe 15 (F1-score : 0,05), tandis que la classe 16 est prédite de manière excessive (F1-score : 0,53, rappel maximal mais précision faible). Cette configuration évoque une confusion persistante entre ces deux régions, qui semblent mal différencier dans cette session. La session S2_S21_S3 sur la Figure 21(c) montre quant à elle un autre type d'erreur : la classe 11 est

entièrement absente des prédictions, tandis que la classe 10 est fortement sur-représentée malgré une précision inférieure à 50 %. Le modèle semble donc confondre ces deux zones, sans parvenir à les distinguer de manière fiable.

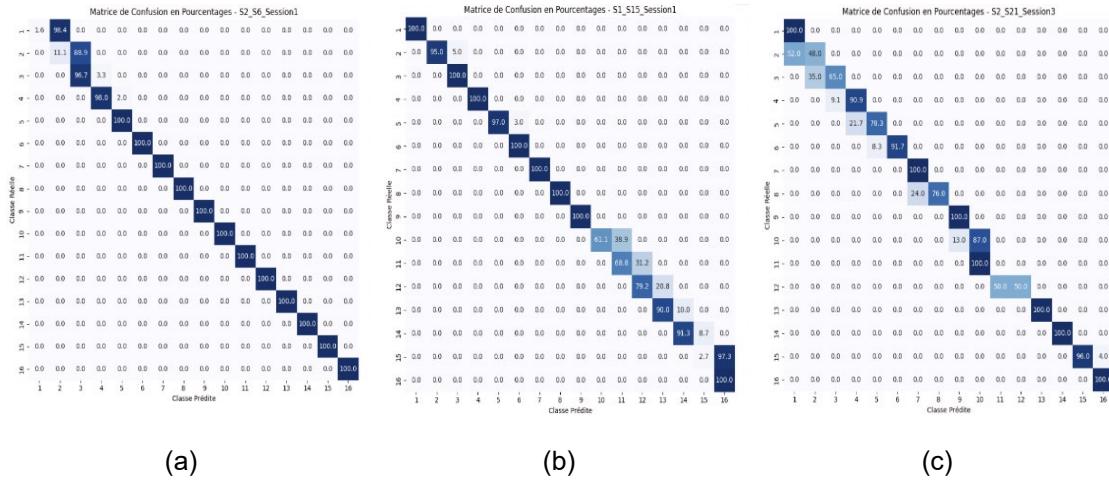


Figure 21: Matrices de confusion pour trois sessions à faible performance

©Sawadogo Wendenda Aniisah

IMPACTS DES MÉTADONNÉES

Afin d'approfondir l'interprétation des écarts observés dans certaines sessions, nous avons examiné les métadonnées associées aux enregistrements. Nous cherchons à déterminer si certaines caractéristiques, telles que le type de brosse utilisée ou le sexe du participant, peuvent expliquer les variations de performance en OSO. Cette analyse révèle notamment une surreprésentation des brosses manuelles parmi les sessions les moins performantes. Comme le montre la Figure 22 et la Figure 24, la dispersion des scores est plus marquée pour ce type de brosse, avec une concentration d'observations sous le seuil de 90 %. Cette variabilité peut être due à l'absence de moteur guidant les gestes, ce qui rend le mouvement plus difficile à reproduire de manière stable d'une session à l'autre.

En revanche, la Figure 23 ne met en évidence aucune différence notable entre les sexes. La distribution des F1-scores est similaire pour les participantes et les participants, et les médianes sont

quasi identiques. Les quelques valeurs aberrantes identifiées dans chaque groupe ne permettent pas d'identifier une tendance claire.

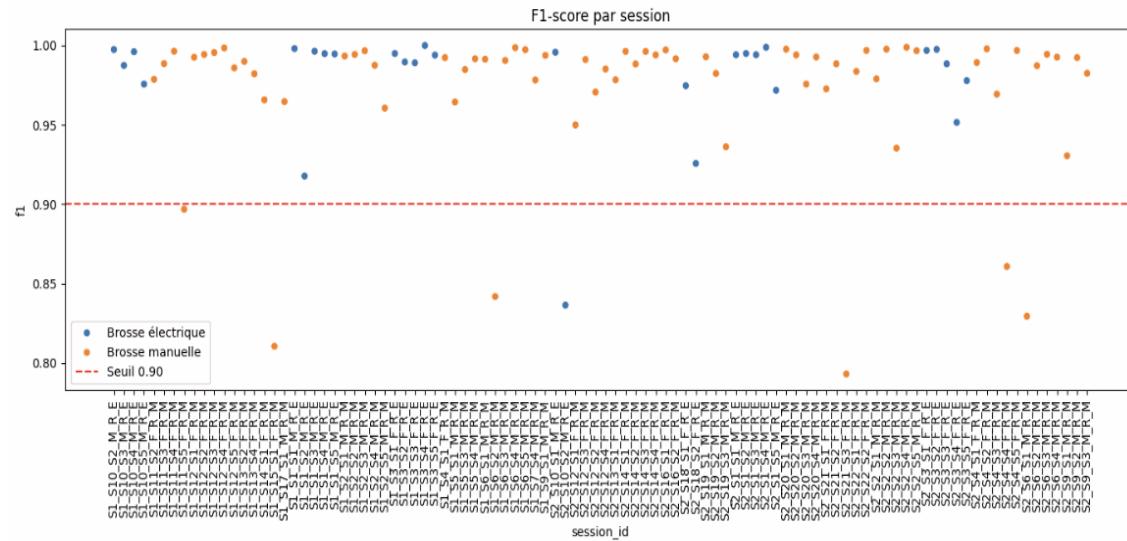


Figure 22: F1-score par session en validation OSO, selon le type de brosse utilisée. Chaque point représente une session test dans le protocole One-Session-Out. Les couleurs distinguent les sessions réalisées avec une brosse électrique (bleu) ou manuelle (orange).

©Sawadogo Wendenda Aniisah

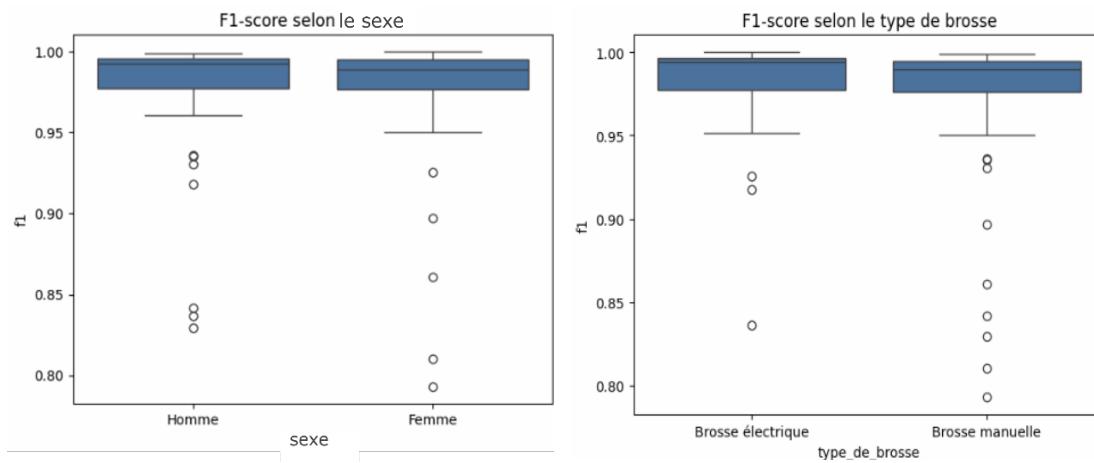


Figure 23: Distribution des F1-scores selon le sexe des participants

Figure 24: Distribution des F1-scores en fonction du type de brosse

©Sawadogo Wendenda Aniisah

4.4 OPTIMISATION DES ALGORITHMES

L'amélioration des performances des modèles passe non seulement par le choix d'une architecture adaptée, mais aussi par l'ajustement fin de ses composantes et de ses paramètres. Cette section présente deux démarches complémentaires mises en œuvre dans ce travail : d'une part, une analyse par ablation visant à mieux comprendre l'apport de chaque module du Transformer, et d'autre part, une optimisation systématique des hyperparamètres pour les modèles classiques à l'aide de RandomizedSearchCV et GridSearchCV.

4.4.1 ANALYSE PAR ABLATION DU MODÈLE TRANSFORMER

Afin de mieux comprendre la contribution de chaque composant de notre modèle de transformer, nous avons mené plusieurs expérimentations en désactivant sélectivement certains modules.

En désactivant l'encodage contextuel, les performances sont légèrement améliorées, atteignant 96,52 % de F1-score. Cette amélioration, bien que modeste, suggère que le modèle est déjà capable d'inférer implicitement les différences liées à la latéralité ou au type de brosse directement à partir des séquences de données. L'ajout explicite de ces informations n'apporte donc pas un avantage systématique, mais peut néanmoins favoriser la stabilité et la généralisation sur des cas moins fréquents ou plus variés.

En revanche, retirer l'encodage positionnel sinusoïdal entraîne une chute marquée des performances. Avec moyenne temporelle (80,58 %), le modèle peine à extraire une représentation temporelle fiable des séquences. Cela confirme l'importance de fournir au modèle un repère explicite de position dans le temps, sans lequel il ne peut capturer les variations de mouvement propres à chaque région de brossage. Ces différentes configurations sont résumées dans la Figure 1, qui compare les performances obtenues selon l'activation ou non des différents composants du modèle Transformer.

Enfin, des ajustements structurels ont été testés. Augmenter le nombre de têtes d'attention de 2 à 4 améliore les performances, avec un F1-score maximale atteignant 96,88 %, probablement grâce à une meilleure capacité à modéliser différentes perspectives spatiales et dynamiques dans la séquence. En revanche, augmenter le nombre de couches de 4 à 6 diminue significativement les performances (93,25 %), indiquant que la profondeur excessive peut nuire au modèle, par dégradation du signal.

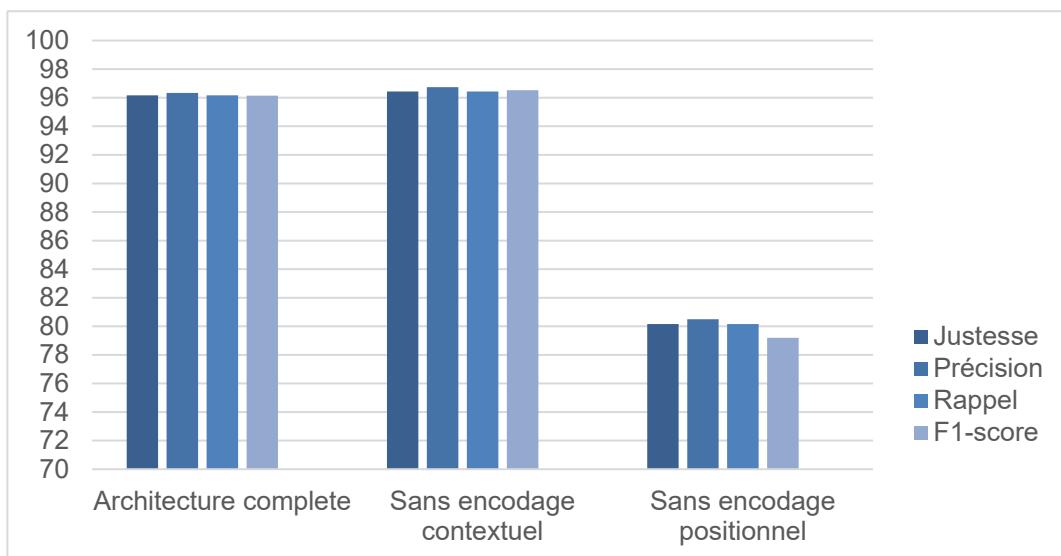


Figure 25: Comparaison des performances du Transformer selon différentes variantes architecturales. Les mesures affichées sont les scores moyen obtenu sur 5 plis de validation croisée.

©Sawadogo Wendenda Aniisah

4.4.2 OPTIMISATION PAR GRIDSEARCHCV ET RANDOMIZEDSEARCHCV

L'optimisation des modèles a reposé sur deux approches complémentaires : GridSearchCV pour les modèles à espace de recherche modéré, et RandomizedSearchCV pour les modèles plus complexes. Ces méthodes ont permis d'examiner comment les hyperparamètres influencent la capacité des algorithmes à généraliser.

Avec RandomizedSearchCV, les améliorations les plus nettes ont été observées sur random forest. Dans les premiers tests, avec 120 arbres, une profondeur fixée et une sélection des

caractéristiques par racine carrée ont permis d'atteindre 97,40 % de F1-score. En poursuivant l'exploration des paramètres, une profondeur illimitée combinée à une sélection logarithmique des attributs a permis d'améliorer ces performances jusqu'à 98,46 %, tout en maintenant une bonne généralisation sur les classes. Du côté des modèles de boosting, XGBoost s'est montré performant dès les premières itérations : changer la profondeur ou le nombre d'arbres n'a pratiquement rien changé. À l'inverse, CatBoost et LightGBM sont restés en retrait. On a essayé différents réglages, mais les résultats sont restés bloqués autour de 87 % pour l'un, et à peine 79 % pour l'autre, ce qui montre que certains modèles sont peu sensibles à l'optimisation ou simplement moins bien adaptés aux données.

Dans un second temps, nous avons exploré une stratégie plus systématique avec GridSearchCV, en particulier pour voir si un réglage plus fin pouvait rehausser les performances des modèles initialement moins efficaces. Cette approche a été particulièrement bénéfique pour KNN. En affinant précisément le nombre de voisins, la méthode de pondération et la métrique de distance, le modèle est parvenu à atteindre jusqu'à 97,36 % de F1-score, un score supérieur à celui obtenu avec RandomizedSearchCV (90,18 %). Ce contraste montre que KNN, bien qu'un modèle simple, bénéficie largement d'un réglage fin. En revanche, pour des modèles comme SVC, Logistic Regression ou Naïve Bayes, même une grille d'hyperparamètres bien pensée n'a pas permis de réel bond en performance. Les scores sont restés modestes, en dessous de 45 % pour la plupart. L'optimisation, bien que nécessaire, ne suffit pas toujours : quand le modèle de base n'est pas en mesure de représenter correctement les données, ajuster ses paramètres ne change pas grand-chose.

4.5 EXPÉRIMENTATIONS SUR LES PROPRIÉTÉS DU DATASET

Cette section examine l'impact de différentes propriétés du jeu de données sur les performances des modèles de classification, en s'appuyant sur une validation croisée pour garantir la fiabilité des résultats. Les expérimentations sont menées à l'aide des trois algorithmes : random forest, le réseau de neurones récurrent (LSTM) et transformer.

Chacun de ces paramètres est étudié séparément, à l'aide de la mesure comparative basée sur le F1-score. Cette analyse permet non seulement d'évaluer la résilience des modèles face à la variabilité des conditions expérimentales, mais aussi d'identifier les combinaisons les plus favorables entre structure du signal et choix algorithmique.

4.5.1 EFFET DE L'EMPLACEMENT DU CAPTEUR

Tableau 7: F1-score des modèles selon l'emplacement du capteur (poignet vs brosse)

	Random Forest	LSTM	Transformer
Capteur au poignet	97,87 %	97,07 %	96,62 %
Capteur sur la brosse	98,48 %	95,83 %	99,31 %

Les résultats obtenus selon l'emplacement du capteur révèlent des comportements contrastés entre les modèles. D'après les données présentées dans le Tableau 7, lorsqu'il est positionné sur la brosse, le transformer atteint sa meilleure performance, avec un F1-score de 99,31 %, contre seulement 96,62 % au poignet. Cette différence de plus de 2 points indique que ce modèle tire mieux parti des signaux directement issus des mouvements fins de la main. À l'inverse, le LSTM semble favoriser les données enregistrées au poignet, où il obtient 97,07 %, soit un gain de plus d'un point par rapport à la brosse (95,83 %). Quant à random forest, il se démarque par sa stabilité : les écarts entre les deux configurations restent limités (98,48 % sur la brosse contre 97,87 % au poignet).

Ce manque de cohérence entre les tendances suggère que chaque architecture ne répond pas de la même manière aux caractéristiques du signal. Certains modèles semblent plus sensibles aux micro-variations captées au plus près de l'action (comme avec la brosse), tandis que d'autres valorisent davantage la régularité globale des gestes mesurée au poignet.

4.5.2 IMPACT DU TYPE DE BROSSE

Tableau 8: F1-score des modèles en fonction du type de brosse (manuelle vs électrique)

	Random Forest	LSTM	Transformer
Brosse Manuelle	98,11 %	96,91 %	99,44 %
Brosse électrique	98,82 %	97,01 %	98,25 %

Les performances varient selon le type de brosse, mais pas pour tous les modèles. On constate en effet dans le Tableau 8 que random forest obtient un F1-score légèrement plus élevé avec la brosse électrique (98,82 %) qu'avec la manuelle (98,11 %), ce qui laisse penser qu'il profite d'un signal plus régulier. Le LSTM progresse légèrement avec la brosse électrique, passant de 96,91 % à 97,01 %, ce modèle montre une bonne capacité d'adaptation, indépendamment du type de mouvement. Le transformer, quant à lui, présente une tendance inverse : ses performances sont meilleures avec la brosse manuelle, atteignant 99,44 %, contre 98,25 % avec la brosse électrique. Bien que cette baisse reste modérée en valeur absolue, elle est notable au regard des autres modèles.

Comme dans l'impact de l'emplacement du capteur, les différences de tendances observées selon le type de brosse montrent que les modèles n'exploitent pas tous les mêmes aspects du signal.

4.5.3 IMPACT DU FILTRAGE DES DONNÉES

Tableau 9: Effet du filtrage des données et de l'estimation d'orientation sur le F1-score des modèles

	Random Forest	LSTM	Transformer
Non filtrées	88,28 %	98,20 %	99,64 %
Non filtrées + orientation	96,49 %	98,03 %	99,22 %
Brosse Électrique	93,28 %	98,23 %	91 %
Brosse Manuelle	88,69 %	98,56 %	99,43 %

Les résultats reportés dans le Tableau 9 révèlent que la qualité du signal a un effet très différent selon les modèles. Random forest est clairement le plus sensible au bruit : son F1-score chute à 88,28 % sur données brutes, ce qui révèle une forte dépendance au prétraitement. À l'inverse, le

LSTM reste très stable, avec 98,20 % sans filtrage, ce qui confirme sa capacité à extraire des motifs discriminants même dans un environnement bruité. Le Transformer atteint même son meilleur score toutes expérimentations confondues (99,64 %) sur données non filtrées. Ce modèle exploite efficacement la richesse du signal brut, sans nécessiter de nettoyage préalable.

Une expérimentation complémentaire a également été réalisée en utilisant les données non filtrées, mais avec une estimation de l'orientation basée sur les angles de tangage et de roulis, calculés à l'aide d'un filtre complémentaire fusionnant les données de l'accéléromètre et du gyroscope. L'ajout de cette information améliore nettement les résultats de random forest, qui passe de 88,28 % à 96,49 %. En revanche, le LSTM et le Transformer affichent des performances légèrement inférieures à celles obtenues sans orientation, suggérant que cette information n'apporte pas un gain systématique pour les modèles profonds. Cela indique que les bénéfices de l'estimation d'orientation dépendent du type de modèle utilisé.

4.5.4 IMPACT DE LA TAILLE DE LA FENÊTRE

Tableau 10: F1-score des modèles en fonction de la taille des fenêtres temporelles

	Random Forest	LSTM	Transformer
50	98,48 %	99,31 %	96,88 %
100	93,49 %	96,47 %	93,50 %
150	92,23 %	87,17 %	93,23 %
200	90,88 %	85,44 %	87,78 %

Les performances des modèles diminuent globalement à mesure que la taille des fenêtres augmente, comme l'illustrent les chiffres du Tableau 10. Sur des fenêtres courtes de 50 échantillons, les résultats sont élevés pour tous les modèles : random forest atteint 98,48 %, le LSTM affiche 99,31 %, et le transformer 96,88 %. À cette taille, les gestes sont bien capturés, et la distinction entre les classes reste nette.

Dès que la fenêtre est étendue à 100 échantillons, une baisse est observable. L'allongement du segment semble diluer l'information utile et introduire des transitions plus floues au sein des

séquences. Cette tendance se confirme avec des fenêtres de 150 et 200 échantillons, où les trois modèles atteignent leurs scores les plus bas. Le LSTM, pourtant spécialisé dans le traitement de séquences, chute à 85,44 % et devient le modèle le plus impacté, tandis que les autres tombent également sous les 91 %. LSTM bien que résistant au bruit reste donc ici sensible à la structuration temporelle des séquences.

4.5.5 IMPACT DE LA TAILLE DE LA FRÉQUENCE D'ÉCHANTILLONNAGE

Tableau 11: F1-Score des modèles selon la fréquence d'échantillonnage

	Random Forest	LSTM	Transformer
50	91,3 %	98,02 %	89,56 %
100	94,84 %	98,93 %	97,57 %
150	95,77 %	99,08 %	98,17 %
200	98,48 %	99,31 %	96,88 %

L'évolution des performances en fonction de la fréquence d'échantillonnage, résumée dans le Tableau 11 montre que tous les modèles bénéficient d'un signal plus dense, mais de manière variable selon leur architecture. Pour random forest une fréquence plus élevée améliore progressivement les résultats : le score-F1 passe de 91,3 % à 50 Hz à 98,48 % à 200 Hz. Le modèle bénéficie ainsi d'un signal plus riche, bien qu'il ne tienne pas compte des relations temporelles entre les observations.

Le LSTM affiche des performances élevées dès les plus basses fréquences. À 50 Hz, il atteint déjà 98,02 %, et franchit les 99 % à partir de 100 Hz, avec une progression modérée jusqu'à 99,31 %. Il reste performant, même lorsque le signal est moins détaillé. Le transformer commence plus bas (89,56 % à 50 Hz), puis s'améliore nettement jusqu'à 98,17 % à 150 Hz, avant de redescendre légèrement à 96,88 % à 200 Hz. Il semble réagir favorablement à une fréquence intermédiaire, mais une densité trop importante tend, au contraire, à nuire à la stabilité de l'apprentissage. Ces résultats soulignent des comportements différents : le LSTM reste constant, le transformer plus fluctuant, et random forest dépendant de la richesse du signal pour atteindre ses meilleures performances.

4.5.6 REBALANCEMENT DES CLASSES

Les classes correspondant aux différentes régions de brossage ne sont pas représentées de manière équilibrée, certaines apparaissant bien plus fréquemment que d'autres dans les données. Cette distribution asymétrique constitue une source de biais pour l'apprentissage supervisé, en particulier pour les algorithmes sensibles à la fréquence des classes telles que les forêts aléatoires.

Pour atténuer l'effet du déséquilibre entre classes, deux méthodes de rééchantillonnage ont été appliquées : SMOTE pour le suréchantillonnage et un sous-échantillonnage aléatoire des classes majoritaires avec UnderSampling. Aucune de ces approches n'apporte une réelle amélioration, et que leur impact varie selon les algorithmes comme le montre le Tableau 12. Logistic regression et le SVC, par exemple, voient leurs performances stagner, voire se dégrader légèrement. En revanche, des modèles plus robustes comme random forest ou XGBoost conservent un bon niveau de performance, suggérant une meilleure capacité à gérer le déséquilibre initial. L'undersampling semble parfois plus efficace que SMOTE, notamment pour LightGBM ou random forest, mais les différences restent modestes. Notons également que ces techniques ne règlent pas le déséquilibre des performances entre classes : certaines restent systématiquement moins bien prédites. Le rééchantillonnage peut améliorer ponctuellement la stabilité de l'apprentissage, sans pour autant garantir une meilleure équité entre les classes.

Tableau 12: F1-score des algorithmes avec SMOTE et undersampling

Algorithme	SMOTE	UNDERSAMPLE
Logistic Regression	37,27 %	37,20 %
Decision Tree	52,19 %	52,11 %
Random Forest	98,18 %	98,52 %
K-Nearest Neighbors (KNN)	97,60 %	96,60 %
Support Vector Classifier (SVC)	45 %	42,05 %
Naives Bayes	34,12 %	34,05 %
Gradient Boosting (XGBoost)	97,47 %	97,58 %
Gradient Boosting (CatBoost)	87,14 %	86,65 %
Gradient Boosting (LightGBM)	78,46 %	78,87 %

4.6 CONCLUSION

L'ensemble des expérimentations confirme que les modèles avancés surpassent globalement les approches classiques. Le LSTM se démarque comme le modèle le plus performant dans des environnements riches, avec une stabilité remarquable lorsque l'ensemble des données est disponible. Cependant, ses performances diminuent dès que la quantité ou la diversité des données est réduite, comme en validation OSO ou dans certaines configurations spécifiques (type de brosse, position du capteur), ce qui suggère une dépendance à une exposition suffisante aux séquences d'entraînement. À l'inverse, le Transformer, bien que plus sensible aux paramètres expérimentaux en validation croisée (fenêtre, fréquence, capteur), démontre une bonne capacité de généralisation en se classant premier en validation LOSO et en OSO. Il se distingue également par des comportements atypiques, ne suivant pas toujours les tendances des autres modèles face aux variations du signal, ce qui témoigne d'un mode d'apprentissage différent.

Les modèles classiques, comme random forest, donnent de bons résultats en conditions contrôlées, mais ont plus de mal à faire face à des données réelles, bruitées ou nouvelles. L'analyse menée, fondée sur des métriques variées et des validations réalistes, permet de conclure à l'atteinte de l'objectif fixé pour cette phase comparative.

CHAPITRE 5

CONCLUSION ET PERSPECTIVES

Cette étude a exploré la reconnaissance des régions de brossage à l'aide des données inertielles issues d'une brosse à dents intelligente. L'objectif principal était de développer un système automatisé permettant d'identifier et de classifier avec précision les différentes zones de brossage, dans une perspective d'amélioration du suivi de l'hygiène bucco-dentaire, notamment chez les personnes âgées. Ce chapitre dresse un bilan des contributions de cette recherche, en mettant en évidence les avancées réalisées en matière de segmentation des signaux et de classification des régions brossées. Il souligne également les limites identifiées, qu'elles concernent la diversité des données, la généralisation des modèles ou encore les contraintes expérimentales. Enfin, plusieurs perspectives sont proposées pour améliorer la stabilité et l'applicabilité du modèle, en tenant compte des variations interindividuelles, de l'optimisation des performances en temps réel et des enjeux liés à la protection des données personnelles.

5.1 RESUME DES CONTRIBUTIONS

Cette étude s'est concentrée sur la reconnaissance des régions de brossage à l'aide des données inertielles capturées par une brosse à dents intelligente, avec pour objectif principal le développement d'un système automatisé de détection et de classification de ces régions. L'enjeu était de fournir des indicateurs exploitables pour le suivi de la santé et de l'autonomie des personnes âgées vivant à domicile. Dans cette optique, un cadre méthodologique rigoureux a été mis en place afin de garantir un prétraitement efficace des données issues des capteurs inertIELS. Les différentes étapes de nettoyage, de filtrage et de segmentation ont permis d'améliorer la qualité des signaux enregistrés et de faciliter l'identification précise des transitions entre les zones de brossage. L'évaluation de plusieurs méthodes a permis d'optimiser la segmentation, garantissant ainsi une meilleure détection des points de changement.

Le travail réalisé a également porté sur l'expérimentation et la comparaison de plusieurs stratégies de classification des régions de brossage. Les performances de modèles classiques, tels que logistic regression, decision tree et les méthodes de boosting comme XGBoost, CatBoost et LightGBM, ont été étudiées en parallèle de celles des réseaux de neurones récurrents, notamment les architectures RNN, LSTM et GRU ainsi que du modèle transformer. L'analyse comparative a permis de mettre en évidence les forces et les limites de ces différentes approches, en se basant sur des mesures d'évaluation comme la justesse, la précision, le rappel et le score F1.

Afin d'assurer une évaluation rigoureuse des performances des modèles, plusieurs protocoles de validation ont été mis en place. L'approche adoptée a combiné la validation croisée stratifiée, utilisée pour mesurer la fiabilité globale des modèles, la validation Leave-One-Subject-Out, qui a permis de tester leur capacité à s'adapter à de nouveaux utilisateurs et la validation One-Session-Out destinée à évaluer leur stabilité temporelle au sein d'un même individu, en confrontant le modèle à des sessions indépendantes. Cette triple approche a fourni une analyse approfondie des erreurs et a mis en lumière les paramètres influençant la qualité des prédictions.

L'étude s'est également intéressée à l'impact de divers facteurs expérimentaux sur la classification des régions de brossage. L'emplacement du capteur, qu'il soit fixé sur la brosse ou porté au poignet, s'est révélé déterminant dans la précision de la détection. De même, le type de brosse utilisée, électrique ou manuelle, a influencé les motifs des signaux enregistrés. L'efficacité du modèle a également été évaluée en fonction du filtrage des données, de la taille des fenêtres d'analyse et de la fréquence d'échantillonnage, permettant d'identifier les configurations les plus adaptées pour améliorer la solidité et la fiabilité du système.

En comparant ces différentes conditions expérimentales, il a été possible d'identifier les modèles les plus performants ainsi que les paramètres optimaux pour garantir une reconnaissance efficace des zones de brossage. Les réseaux de neurones récurrents, en particulier LSTM, ont affiché les meilleures performances en validation croisée, avec des taux de F1-score dépassant 99 %. Toutefois, c'est le Transformer qui s'est montré le plus fiable en généralisation, notamment dans des contextes à forte variabilité interindividuelle et intra-individuelle.

Enfin, cette recherche souligne le potentiel des brosses à dents intelligentes comme outil non intrusif de suivi de l'hygiène personnelle et de l'autonomie des personnes âgées. En exploitant la reconnaissance des schémas de brossage, il serait possible de détecter précocement des troubles moteurs ou cognitifs, ouvrant ainsi des perspectives d'application prometteuses dans les domaines de la télésanté et de la médecine préventive. En proposant un cadre méthodologique rigoureux et en explorant l'efficacité de différentes stratégies de classification, ce travail apporte une contribution utile à l'étude de l'utilisation des capteurs inertiels pour la reconnaissance d'activités quotidiennes et la prévention des risques associés au vieillissement.

5.2 LIMITES DE L'ETUDE

Cette étude contribue à une meilleure compréhension de la reconnaissance des régions de brossage. Ces limitations concernent à la fois la variabilité des données collectées ainsi que les contraintes techniques et expérimentales rencontrées au cours de la recherche.

Une première limite réside dans la diversité restreinte des participants et des conditions expérimentales. Comme dans plusieurs études antérieures, notamment celles de (Huang & Lin, 2016; Luo et al., 2019), les données utilisées ont été collectées dans un environnement contrôlé avec un protocole défini. Or, cette approche, bien qu'indispensable pour garantir la qualité des annotations, limite la capacité du modèle à généraliser à des situations plus naturelles et spontanées. De plus, notre jeu de données présente une limitation majeure en lien avec la population ciblée par notre problématique : bien que l'objectif de cette recherche soit d'améliorer le suivi du brossage chez les personnes âgées, aucun participant de notre échantillon ne correspond à cette tranche d'âge. En effet, l'âge le plus élevé parmi les participants ne dépasse pas 40 ans, ce qui signifie que notre modèle a été entraîné et évalué sur une population plus jeune. Or, les habitudes de brossage peuvent être influencées par des facteurs tels que la dextérité, la force musculaire ou encore d'éventuels troubles moteurs, qui sont plus fréquents chez les seniors. Cette absence de données représentatives des utilisateurs finaux limite donc la portée clinique de nos résultats et souligne la nécessité de tests complémentaires sur une population plus âgée.

Par ailleurs, la généralisation des modèles à de nouveaux utilisateurs demeure une problématique centrale. Si la validation Leave-One-Subject-Out (LOSO) a permis d'évaluer la capacité du modèle à s'adapter à des utilisateurs non vus lors de l'entraînement, certaines différences interindividuelles persistent et limitent la fiabilité du système pour tous les profils. Un modèle entraîné sur un groupe spécifique d'utilisateurs pourrait ne pas fonctionner aussi efficacement sur d'autres populations aux habitudes de brossage différentes. Cette difficulté est également présente dans les études antérieures, où la faible diversité des participants constitue un frein à la généralisation des résultats. Dans notre cas, l'absence de données issues de personnes âgées rend impossible une véritable évaluation des performances du modèle sur la population cible, ce qui constitue un biais important.

5.3 PERSPECTIVES

Cette étude a montré qu'il est possible d'utiliser des capteurs inertIELS intégrés à une brosse à dents intelligente pour reconnaître les différentes régions de brossage. Malgré ces résultats encourageants, plusieurs pistes restent à explorer pour améliorer la robustesse du modèle et son utilisation dans des conditions plus réalistes. Il serait intéressant d'examiner comment adapter le modèle aux différences entre les utilisateurs, détecter les changements dans les habitudes de brossage, rendre le système plus léger pour fonctionner sur une brosse connectée ou un smartphone, explorer la prédiction des mouvements futurs, et réfléchir à la gestion et la protection des données.

PRENDRE EN COMPTE LES DIFFÉRENCES ENTRE LES UTILISATEURS

Un des défis majeurs est que tout le monde ne se brosse pas les dents de la même manière. La vitesse, la pression exercée, l'ordre des régions brossées, tout cela varie d'une personne à l'autre. Dans les données utilisées pour cette étude, les participants suivaient un protocole précis, mais en conditions réelles, les habitudes peuvent être très différentes.

Pour que le modèle fonctionne bien pour tout le monde, il faudrait trouver un moyen de le rendre plus flexible et capable d'apprendre des habitudes de chaque utilisateur. Une piste serait d'intégrer des méthodes d'adaptation progressive, qui ajusteraient les prédictions au fil des utilisations. Un autre point important serait d'élargir la collecte de données à des utilisateurs plus variés, en particulier des personnes âgées, pour mieux correspondre à la population cible.

REPÉRER LES CHANGEMENTS INHABITUELS DANS LE BROSSAGE

Un brossage qui devient plus court ou plus irrégulier peut être un signe d'un problème de santé ou d'une baisse de motivation. Un modèle capable de détecter ces changements pourrait être utile pour prévenir des troubles comme des difficultés motrices ou une perte d'autonomie. Par exemple, si la durée du brossage diminue progressivement, si certaines zones sont systématiquement oubliées ou si des gestes anormalement lents ou irréguliers apparaissent, cela pourrait être le signe d'une perte de dextérité ou d'un début de trouble moteur. En intégrant une analyse des tendances à long terme, il serait possible d'identifier ces changements et d'envoyer des recommandations ou des alertes aux soignants ou aux proches (Hussain, 2022). En effet, les approches basées sur l'analyse temporelle et les autoencodeurs (Fahad & Tahir, 2021) peuvent permettre d'identifier les comportements inhabituels en comparant chaque session avec les schémas de brossage antérieurs. Plusieurs autres travaux ont exploré la détection des anomalies dans les routines quotidiennes. Les réseaux de neurones récurrents (RNN, LSTM, GRU) ont notamment été utilisés pour analyser les comportements et repérer des écarts inhabituels chez les personnes atteintes de démence (Arifoglu & Bouchachia, 2017). Une approche similaire appliquée au brossage pourrait aider à identifier des modifications progressives du geste, signalant une perte de dextérité.

Toutefois, ces variations ne sont pas toujours dues à l'utilisateur lui-même. Des études sur les maisons intelligentes montrent que l'environnement influence aussi les comportements (Sfar & Bouzeghoub, 2018). Un changement de routine ou une gêne physique temporaire peuvent modifier le brossage, soulignant l'importance d'une approche qui intègre à la fois les tendances individuelles et le contexte global.

ALLÉGER LE MODÈLE POUR QU'IL FONCTIONNE EN TEMPS RÉEL

Pour qu'un tel système soit vraiment utile, il faudrait qu'il puisse fonctionner directement sur une brosse connectée ou une application mobile, sans nécessiter un ordinateur puissant. Or, les modèles de réseaux de neurones utilisés ici, bien qu'efficaces, demandent pas mal de ressources. Il faudrait donc tester des modèles plus légers et optimisés, qui consomment moins d'énergie tout en gardant de bonnes performances. Une autre approche serait de voir si réduire la fréquence d'échantillonnage des capteurs permettrait d'économiser de la puissance sans trop perdre en précision.

ANTICIPER LES MOUVEMENTS POUR GUIDER LE BROSSAGE

Actuellement, le modèle reconnaît les régions après coup, mais il pourrait aussi être intéressant d'essayer de prédire les prochaines actions de l'utilisateur. Par exemple, si quelqu'un passe trop vite sur une région ou oublie une zone, le système pourrait lui suggérer en temps réel de corriger son brossage. Ce type de retour en direct pourrait être utile sous forme d'une application mobile interactive ou même d'un signal sur la brosse (vibration, LED). Cela permettrait d'aider les utilisateurs à mieux répartir leur temps et à s'assurer qu'aucune zone n'est oubliée.

GÉRER ET PROTÉGER LES DONNÉES DES UTILISATEURS

Le suivi des habitudes de brossage implique de collecter des données personnelles, ce qui pose la question de leur sécurité et de leur confidentialité. Il serait donc important de réfléchir à des solutions qui protègent les informations des utilisateurs tout en permettant une analyse efficace. Une solution pourrait être de traiter directement les données sur l'appareil, plutôt que de les envoyer sur un serveur distant. Il faudrait aussi envisager des méthodes de chiffrement pour garantir que seules les personnes autorisées puissent y accéder.

RÉFÉRENCES

Aggar, C., Sorwar, G., Seton, C., Penman, O., & Ward, A. (2023). Smart home technology to support older people's quality of life : A longitudinal pilot study. *International journal of older people nursing*, 18(1), e12489.

American Dental Association. (2022). *Toothbrushes*. <https://www.ada.org/resources/ada-library/oral-health-topics/toothbrushes>

Arifoglu, D., & Bouchachia, A. (2017). Activity recognition and abnormal behaviour detection with recurrent neural networks. *Procedia Computer Science*, 110, 86-93.

Beaudoin-Benjamin, A. S.-L.-M., Labesse, D. D.-B.-M. E., & Tournigny, É. R.-A. (2022). *CREBS OPUSS-Des lieux pour vieillir en bonne santé*.

Bélanger, R., Blanchet, C., & Hamel, D. (2016). *La santé buccodentaire des aînés québécois* [Collection Vieillissement en santé]. Institut national de santé publique du Québec. https://www.inspq.qc.ca/sites/default/files/publications/2097_sante_buccodentaire_aines_quebecois.pdf

Bian, S., Liu, M., Zhou, B., & Lukowicz, P. (2022). The state-of-the-art sensing techniques in human activity recognition : A survey. *Sensors*, 22(12), 4596.

Canadian Dental Hygienists Association. (2022). *Toothbrushing Toolkit for Total Health*. Dental Hygiene Canada. https://www.dentalhygienecanada.ca/dhcanada/DHCanada/Preventive_Oral_Care/Toothbrushing_Toolkit.aspx

Chen, C.-H., Wang, C.-C., & Chen, Y.-Z. (2021). Intelligent brushing monitoring using a smart toothbrush with recurrent probabilistic neural network. *Sensors*, 21(4), 1238.

Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.

Christ, M., Braun, N., Neuffer, J., & Kempa-Liehr, A. W. (2018). Time series feature extraction on basis of scalable hypothesis tests (tsfresh—a python package). *Neurocomputing*, 307, 72-77.

D. Guijo-Rubio, A. M. Durán-Rosal, P. A. Gutiérrez, A. Troncoso, & C. Hervás-Martínez. (2021).

Time-Series Clustering Based on the Characterization of Segment Typologies. *IEEE Transactions on Cybernetics*, 51(11), 5409-5422.

<https://doi.org/10.1109/TCYB.2019.2962584>

Deshpande, A., Kacham, P., & Pratap, R. (2020). *Robust \$ k \\$-means++*. 799-808.

Dhariyal, B., Le Nguyen, T., & Ifrim, G. (2023). *Back to basics : A sanity check on modern time series classification algorithms*. 205-229.

Duffau, F., & Lostal, L. (2021). Temps par temps pour motiver et enseigner le contrôle de plaque. *CLINIC*, 42, 726-733.

Essalat, M. (2023). *Time-series Classification for Monitoring Toothbrushing Behavior Using Motion Sensors*. University of California, Los Angeles.

Essalat, M., Padilla, O. H. M., Shetty, V., & Pottie, G. (2023). Monitoring brushing behaviors using toothbrush embedded motion-sensors. *Authorea Preprints*.

Fahad, L. G., & Tahir, S. F. (2021). Activity recognition and anomaly detection in smart homes. *Neurocomputing*, 423, 362-372.

Faisal, I. A., Purboyo, T. W., & Ansori, A. S. R. (2019). A review of accelerometer sensor and gyroscope sensor in IMU sensors on motion capture. *J. Eng. Appl. Sci*, 15(3), 826-829.

Gupta, N., Gupta, S. K., Pathak, R. K., Jain, V., Rashidi, P., & Suri, J. S. (2022). Human activity recognition in artificial intelligence framework : A narrative review. *Artificial intelligence review*, 55(6), 4755-4808.

Huang, H., & Lin, S. (2016). *Toothbrushing monitoring using wrist watch*. 202-215.

Humaira, H., & Rasyidah, R. (2020). *Determining the appropriate cluster number using elbow method for k-means algorithm*. 1-8.

Hussain, Z. (2022). *Ubiquitous human activity recognition for healthcare applications : A study*. Macquarie University.

Hussain, Z., Waterworth, D., Aldeer, M., Zhang, W. E., Sheng, Q. Z., & Ortiz, J. (2021). *Do you brush your teeth properly ? An off-body sensor-based approach for toothbrushing monitoring*. 59-69.

Institut De La Statistique Du Québec. (2024). *Le bilan démographique du Québec. Édition 2024*.

Janakiram, C., Varghese, N., Venkitachalam, R., Joseph, J., & Vineetha, K. (2020). Comparison of modified Bass, Fones and normal tooth brushing technique for the efficacy of plaque control in young adults-A randomized clinical trial. *Journal of clinical and experimental dentistry*, 12(2), e123.

Jansen, W. (2024). *Acquiring Head Orientation through using an IMU with Madgwick Orientation Filtering*. University of Twente.

Kim, D., Bian, H., Chang, C. K., Dong, L., & Margrett, J. (2022). In-home monitoring technology for aging in place : Scoping review. *Interactive journal of medical research*, 11(2), e39005.

Kok, M., Hol, J. D., & Schön, T. B. (2017). Using inertial sensors for position and orientation estimation. *arXiv preprint arXiv:1704.06053*.

Kubelka, V., & Reinstein, M. (2012). *Complementary filtering approach to orientation estimation using inertial sensors only*. 599-605.

Lee, B., Bang, W.-C., Kim, J. D., & Kim, C. Y. (2011). Orientation estimation in mobile virtual environments with inertial sensors. *IEEE Transactions on Consumer Electronics*, 57(2), 802-810.

Li, H., & Jing, L. (2023). 3d monitoring of toothbrushing regions and force using multimodal sensors and unity. *IEEE Access*, 11, 94753-94771.

Lopez-Nava, I. H., & Munoz-Melendez, A. (2016). Wearable inertial sensors for human motion analysis : A review. *IEEE Sensors Journal*, 16(22), 7821-7834.

Luo, C., Feng, X., Chen, J., Li, J., Xu, W., Li, W., Zhang, L., Tari, Z., & Zomaya, A. Y. (2019). *Brush like a dentist : Accurate monitoring of toothbrushing via wrist-worn gesture sensing*. 1234-1242.

Mahony, R., Hamel, T., & Pflimlin, J.-M. (2008). Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on automatic control*, 53(5), 1203-1218.

Menon, A., Schroth, R. J., Hai-Santiago, K., Yerex, K., & Bertone, M. (2024). The Canadian dental care plan and the senior population. *Frontiers in Oral Health*, 5, 1385482.

Mukhopadhyay, S. C. (2014). Wearable sensors for human activity monitoring : A review. *IEEE Sensors Journal*, 15(3), 1321-1330.

Nia, N. G., Kaplanoglu, E., Nasab, A., & Qin, H. (2023). *Human activity recognition using machine learning algorithms based on IMU data*. 1-8.

Oppenheim, A. V., Willsky, A. S., & Nawab, S. H. (1997). *Signals & systems*. Pearson Educación.

Pramesti, T. A., Hadi, S., & Ulfah, S. F. (2023). *THE EFFECTIVENESS OF TOOTH BRUSHING WITH BASS AND ROLL METHOD ON DECREASIN PLAQUE INDEX IN DEAF STUDENTS*. 2.

S. A. Ludwig, & K. D. Burnham. (2018). *Comparison of Euler Estimate using Extended Kalman Filter, Madgwick and Mahony on Quadcopter Flight Data*. 1236-1241.
<https://doi.org/10.1109/ICUAS.2018.8453465>

Savard, L. (2023). *POUR UNE SOCIÉTÉ OÙ IL FAIT BON VIEILLIR-DOCUMENT DE CONSULTATION*.

SciPy, C. (2025). *Scipy.signal.find_peaks – Find peaks in a signal*.

Sfar, H., & Bouzeghoub, A. (2018). Activity recognition for anomalous situations detection. *IRBM*, 39(6), 400-406.

Statistique Canada. (2022). *Portrait de la population croissante des personnes âgées de 85 ans et plus au Canada selon le Recensement de 2021* (Nos. 98-200-X2021004). Statistique Canada. <https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-x/2021004/98-200-x2021004-fra.cfm>

Truong, C., Oudre, L., & Vayatis, N. (2018). ruptures : Change point detection in Python. *arXiv preprint arXiv:1801.00826*.

Vanleerberghe, P., De Witte, N., Claes, C., Schalock, R. L., & Verté, D. (2017). The quality of life of older people aging in place : A literature review. *Quality of Life Research*, 26, 2899-2907.

Vassilvitskii, S., & Arthur, D. (2006). *k-means++ : The advantages of careful seeding*. 1027-1035.

W. H. O. Centre for Health Development. (2004). *A glossary of terms for community health care and services for older persons*. Kobe, Japan : WHO Centre for Health Development.

Weiss, M. L., McDonald, J., Bestor, D., Yee, C., Edelman, D., Jones, M., Prout, A., Bowne, A., McEvoy, L., & Gadepally, V. (2022). *An Evaluation of Low Overhead Time Series Preprocessing Techniques for Downstream Machine Learning*. 1-6.

Zawar, H. (2021). *Dataset for Toothbrushing Activity Using Brush-Attached and Wearable Sensors* (Version 1) [Dataset]. <https://doi.org/10.17632/hx5kkbr3j.1>