



Université Rennes 2 – EUR Digisport
Master 2 Sciences du Numérique et Sport

Développement d'un algorithme personnalisé pour la détection du
Freezing of Gait chez les personnes atteintes de la Maladie de Parkinson

Année universitaire 2023-2024
Présenté par : Anthony PLACÉ

Encadrants : Karim JAMAL et Sébastien CORDILLET

Remerciements

Je tiens à exprimer ma profonde gratitude envers Karim Jamal et Sébastien Cordillet pour leur incroyable soutien tout au long de cette expérience. Leur disponibilité et leur engagement ont été des piliers sur lesquels j'ai pu m'appuyer sans cesse. Leur bonne humeur constante et leur légèreté ont su créer un environnement de travail agréable et stimulant, transformant ce parcours en un véritable plaisir. Leur approche équilibrée entre professionnalisme et moments de rigolade a rendu chaque jour plus enrichissant et plaisant.

Mais je souhaite exprimer encore une fois ma reconnaissance envers Sébastien. En effet, son expertise exceptionnelle m'a permis d'acquérir des compétences précieuses et d'approfondir mes connaissances dans des domaines essentiels.

Je remercie également mes collègues de l'UCAM, qui m'ont chaleureusement accueilli et intégré à l'équipe. Leur soutien, leurs échanges constructifs et leur camaraderie ont grandement contribué à la richesse de cette expérience.

Un grand merci à Florian pour son efficacité dans l'inclusion des patients. Sans son dévouement, je n'aurais pas eu la matière nécessaire pour mener à bien mon étude.

Enfin, je n'oublierai jamais Crousti-pied, qui m'a fait découvrir des aspects insoupçonnés de mon travail et m'a rappelé la réalité parfois difficile de ce métier. Son impact a été profond et inoubliable.

Liste des figures

Figure 1: Les étapes de pré-traitement des données.....	9
Figure 2: Performance de détection du FOG et latence de détection en fonction de la longueur de la fenêtre, pour un RF.....	10
Figure 3: Diagramme sur le fenêtrage des données avec la composition des classes cibles : (a) structure des fenêtres, (b) Composition des classes pour un algorithme de prédiction et de détection, (c) Composition des classes pour un algorithme de détection, (d) Composition des classes pour un algorithme de prédiction.....	12
Figure 4: Les étapes de traitement des données.....	13
Figure 5: Parcours de la tâche de marche. La case de gauche représente la chaise au début et à la fin de l'essai. Les triangles sont des cônes et la zone grise délimite un couloir étroit. Les octogones rouges indiquent des arrêts volontaires momentanés (2 s).....	18
Figure 6: Organisation de l'étude.....	18
Figure 7: Placement des IMUs.....	19
Figure 8: Labellisation des fenêtres.....	20
Figure 9: Organigramme du processus de création de l'algorithme.....	23
Figure 10: Répartition du nombre de fenêtre en fonction des classes et des méthodes de rééquilibrage.....	24
Figure 11: Répartition des 20 caractéristiques les plus importantes en fonction de la localisation des IMUs et des patients affichant également la nature des caractéristiques pour chaque méthode de rééquilibrage.....	25
Figure 12: Comparaison des résultats du Random Forest issus des différentes méthodes appliquées pour n nombre de caractéristiques.....	26

Index des tableaux

Tableau 1: Détails concernant l'extraction des caractéristiques.....	21
Tableau 2: Nombre de fenêtres pour chaque label extrait de chaque patient.....	36
Tableau 3: Nombre de fenêtres regroupées en classes et servant à la création de l'algorithme.....	36
Tableau 4: Performances du Random Forest.....	36

Liste des abréviations

ACC : Accéléromètre

AUC : Area Under the Curve

DP : Deep Learning

FOG : Freezing Of Gaits

GYRO : Gyroscope

IMUs : Unités de Mesure Inertielle

k-NN : k-nearest neighbors

LOSO : Leave-One-Subject-Out

ML : Machine Learning

MP : Maladie de Parkinson

RF : Random Forest

SMOTE : Synthetic Minority Over-sampling Technique

SVM : Support Vector Machines

Sommaire

Introduction.....	1
I. Approches technologiques pour la détection du FOG.....	2
1. Placement des IMUs.....	2
2. Utilisation des algorithmes de Machine Learning (ML).....	3
II. Pré-traitement des données.....	4
1. Filtrage et Normalisation.....	4
2. Fenêtrage des données.....	5
a. La longueur.....	5
b. Le chevauchement.....	5
c. La labellisation.....	6
III. Traitement des données.....	8
1. Extraction des caractéristiques.....	8
2. Sélection des caractéristiques.....	9
3. Équilibrage des données.....	9
4. Expérimentation et évaluation de l'algorithme.....	10
IV. Objectifs.....	12
V. Matériel et Méthodes.....	13
1. Protocole.....	13
2. Collecte des données et labellisation du FOG.....	14
3. Pré-traitement des données.....	15
4. Traitement des données.....	16
a. Extraction des caractéristiques.....	16
b. Sélection des caractéristiques.....	17
5. Expérimentation et validation.....	18
VI. Résultats.....	19
1. Rééquilibrage des données.....	19
2. Sélection des caractéristiques.....	20
3. Expérimentation et validation.....	21
VII. Discussions.....	22
VIII. Conclusion.....	24
Résumé.....	25
Abstract.....	26
Bibliographie.....	27
Annexes.....	31

Introduction

La maladie de Parkinson (MP) est un trouble neurodégénératif chronique qui affecte environ 8,5 millions de personnes à travers le monde (Organisation Mondiale de la santé, 2023). Elle affecte les neurones dopaminergiques, conduisant à une diminution de la dopamine, un neurotransmetteur essentiel pour le contrôle des mouvements volontaires, ce qui peut se manifester par des symptômes moteurs tels que le tremblement, la rigidité, la bradykinésie et l'instabilité posturale (Beitz, 2014), induisant des risques importants de chutes. En plus des troubles moteurs, la MP est associée à des symptômes non moteurs chez presque tous les patients, comme la perte de mémoire, la dépression et les troubles du sommeil (Tolosa et al., 2021).

Parmi les symptômes les plus invalidants, le Freezing of Gait (FOG) se caractérise par une "absence brève et épisodique ou une réduction marquée de la progression des pieds vers l'avant malgré l'intention de marcher" (Nutt et al., 2011). Les patients décrivent souvent une sensation de pieds collés au sol, rendant difficiles les démarrages ou les transitions (Schaafsma et al., 2003). Ces épisodes surviennent principalement lors de l'initiation de la marche, dans des couloirs étroits, lors de virages ou dans des situations stressantes et durent généralement quelques secondes (Cockx et al., 2023).

Sur le plan physiopathologique, le FOG semble lié à des dysfonctionnements dans les régions cérébrales responsables de la planification et du contrôle moteur, exacerbés par la variabilité de la réponse à la dopamine, particulièrement imprévisible chez les patients avancés de la MP. Le FOG pourrait également être associé à une déconnexion entre les régions cérébrales qui coordonnent les mouvements automatiques et celles nécessitant une attention consciente, influençant l'apparition du FOG en fonction du type de tâche effectuée (Gao et al., 2020; Snijders et al., 2016).

La prise en charge actuelle du FOG repose principalement sur des traitements pharmacologiques comme la lévodopa, un précurseur de la dopamine. Cependant, son efficacité peut diminuer avec le temps et des effets secondaires significatifs peuvent apparaître. Concernant son évaluation, elle est le plus souvent basée sur des observations subjectives et des rapports de patients, ce qui manque de précision et complique l'ajustement optimal des traitements (Bikias et al., 2021).

Pour pallier ces limitations, une détection fiable des épisodes de FOG pourrait améliorer la précision des diagnostics et permettre des interventions mieux ciblées. Les Unités de Mesure Inertielle (IMUs), capturant des données cinématiques continues et objectives, offrent une solution pour détecter les patterns de marche associés au FOG. Ces dispositifs permettent de surmonter les limitations des méthodes actuelles, en fournissant des données détaillées et exploitables pour une surveillance précise, même en dehors du cadre clinique (O'Day et al., 2022). Cependant, l'usage des IMUs associé aux algorithmes varie tels que le placement des capteurs, les caractéristiques d'inclusion, les paramètres de l'algorithme. Ce mémoire vise donc non seulement à développer un algorithme de détection du FOG, mais également à proposer des pratiques standardisées pour offrir des diagnostics plus précis.

I. Approches technologiques pour la détection du FOG

Pour comprendre les avancées récentes dans la détection du FOG chez les patients atteints de la maladie de Parkinson, il est essentiel d'examiner les travaux existants utilisant les IMUs. Dans cette partie nous allons discuter de l'importance du placement des IMUs et des différents algorithmes de Machine Learning (ML) utilisés.

1. Placement des IMUs

L'utilisation des IMUs pour la détection du FOG a démontré son utilité dans de nombreuses recherches. Toutefois, l'identification de l'emplacement optimal de ces dispositifs reste un sujet de débat, étant donné que leur positionnement et leur nombre peuvent influencer de manière significative la qualité des données recueillies. En fonction de l'emplacement des IMUs, les types de mouvements et d'événements détectés peuvent varier considérablement. Les capteurs placés aux pieds, par exemple, sont plus susceptibles de détecter les arrêts brusques et les mouvements fins des orteils, tandis que les capteurs situés au niveau du bas du dos capturent les changements de posture et d'équilibre. Certaines études suggèrent que les placements au niveau du bas du dos et/ou de la cheville peuvent être optimaux pour capturer les données les plus pertinentes (Moore et al., 2008; O'Day et al., 2022; Tripoliti et al., 2013).

Cette diversité dans la détection des données en fonction de l'emplacement des capteurs complique la standardisation des protocoles de détection du FOG. De plus, les critères d'inclusion des participants et les méthodes de traitement des données varient d'une étude à l'autre, ce qui peut conduire à des résultats divergents même pour des capteurs placés aux mêmes endroits. Par exemple, des études utilisant des algorithmes différents pour analyser les données peuvent aboutir à des conclusions différentes, soulignant l'importance de la cohérence dans les approches méthodologiques. Il est donc essentiel de standardiser non seulement les emplacements des capteurs, mais aussi les algorithmes et les protocoles de traitement des données pour améliorer la comparabilité des résultats entre les études. Cela permettra d'assurer une détection plus fiable et cohérente des épisodes de FOG, contribuant ainsi à une meilleure gestion clinique des patients atteints de la MP.

2. Utilisation des algorithmes de Machine Learning (ML)

L'utilisation des algorithmes de détection du FOG a évolué de manière significative au fil des années. Les approches initiales reposaient principalement sur des méthodes de seuils, utilisant des caractéristiques simples, c'est à dire des valeurs numériques calculées à partir des signaux des capteurs capturant des paramètres spécifiques des données. Un exemple typique est le Freeze Index (FI) introduit par Moore et al. (2008), qui mesure le rapport de la puissance dans la bande de fréquences du FOG (3-8 Hz) par rapport à celle de la bande de locomotion (0-3 Hz). Un seuil était défini en utilisant la moyenne et un écart type du FI de la marche normale, permettant d'identifier le FOG lorsque ce seuil était dépassé. Bien que cette méthode soit efficace en termes de temps de traitement des données, elle présente des performances de détection faibles.

Pour surmonter les limitations des méthodes basées sur des seuils, un panel de caractéristiques plus diversifié a été développé. Les caractéristiques utilisées pour l'identification du FOG intègrent des caractéristiques temporelles ainsi que des mesures basées sur la transformée de Fourier pour extraire des caractéristiques fréquentielles.

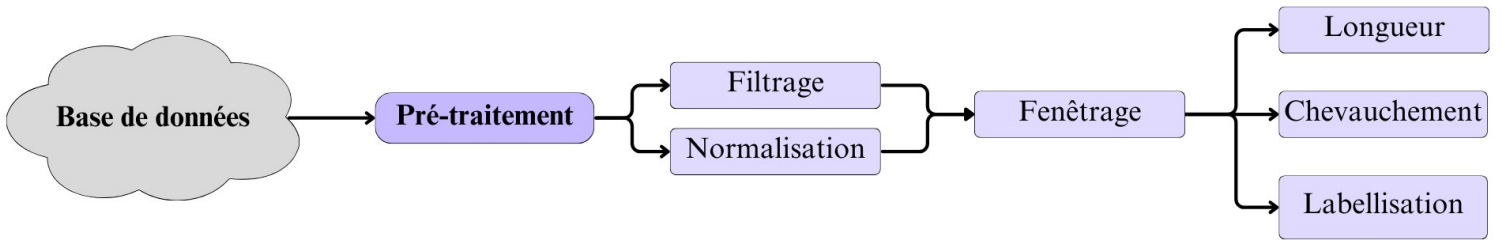
L'amélioration des méthodes de détection du FOG a conduit au développement d'algorithmes d'apprentissage supervisé. Dans ce type de classification, les données sont étiquetées en classes spécifiques (par exemple, FOG ou Non-FOG) et utilisées pour entraîner un modèle de classification. Une fois le modèle entraîné, des données de test non étiquetées lui sont fournies. Le modèle utilise les connaissances acquises pendant l'entraînement pour classer ces nouvelles données dans les classes appropriées. Ainsi, un modèle de détection de FOG classifierait les données comme FOG ou Non-FOG en se basant sur les exemples d'entraînement précédents (Pardoel, 2021).

Parmi ces algorithmes, on retrouve le ML avec le Random Forest (RF), les Support Vector Machines (SVM), les k-nearest neighbors (k-NN) et les Naïve Bayes qui montrent des performances particulièrement robustes (Mazilu et al., 2012; Reches et al., 2020; Rodríguez-Martín et al., 2017; Samà et al., 2018; Sigcha et al., 2020; Tripoliti et al., 2013). Mais il y a également des techniques de Deep Learning (DL) utilisant des réseaux de neurones artificiels pour modéliser des données complexes, tels que les réseaux de neurones convolutifs (CNN) (Ashfaq Mostafa et al., 2021; Bikias et al., 2021; Camps et al., 2018; Sigcha et al., 2020). Cependant, bien que les techniques de DP aient montré des résultats prometteurs, elles ne sont pas toujours nécessaires pour atteindre une précision élevée dans la détection du FOG. En effet, les modèles de ML offrent une complexité moindre et une interprétabilité plus facile, ce qui est crucial pour une application clinique.

Pour garantir la pertinence des données utilisées par les modèles prédictifs, une gestion minutieuse est essentielle. Cela comprend le pré-traitement des données pour assurer des entrées cohérentes et représentatives (Figure 1), suivies de techniques avancées pour optimiser les performances des algorithmes.

II. Pré-traitement des données

Figure 1: Les étapes de pré-traitement des données



1. Filtrage et Normalisation

La première étape du pré-traitement consiste à filtrer les données pour éliminer le bruit et les artefacts qui pourraient fausser les résultats. Les filtres passe-bas, atténuant les hautes fréquences, sont largement utilisés car ils maintiennent la forme des signaux de basses fréquences tout en réduisant efficacement le bruit. Étant donné, que les fréquences de marche sont entre 0.5-3 Hz et les FOG entre 3-8 Hz (Moore et al., 2008), il est inutile de conserver des fréquences largement supérieures à 8 Hz. C'est pourquoi de nombreuses études appliquent un filtre passe-bas avec une fréquence de coupure de 15 Hz, permettant de préserver les données significatives relatives au FOG tout en éliminant les fluctuations rapides dues au bruit (Rodríguez-Martín et al., 2017; Sigcha et al., 2020).

Par ailleurs, la normalisation est effectuée pour mettre à échelle les données de manière à ce que différentes sessions d'enregistrement ou différents sujets soient comparables. Cela inclut souvent la normalisation centrée sur zéro sur l'intégralité du signal (Borzi et al., 2021), ce qui permet une cohérence globale entre les sessions et réduit les variations dues aux conditions d'enregistrement et la variabilité inter-individuelle liée à la vitesse de marche. Cependant, cette méthode peut être influencée par des valeurs extrêmes, qui peuvent fausser la moyenne et l'écart type, masquant ainsi les dynamiques locales.

Pour pallier à cela, une autre méthode consiste à normaliser les données à l'intérieur de différentes fenêtres temporelles, c'est-à-dire des segments de durée fixe du signal (O'Day et al., 2022). Cette approche permet de mieux capturer les variations locales spécifiques à chaque fenêtre, offrant une adaptation plus fine aux dynamiques du signal. Cependant, elle peut introduire des variations artificielles entre les fenêtres successives et compliquer l'analyse globale du signal.

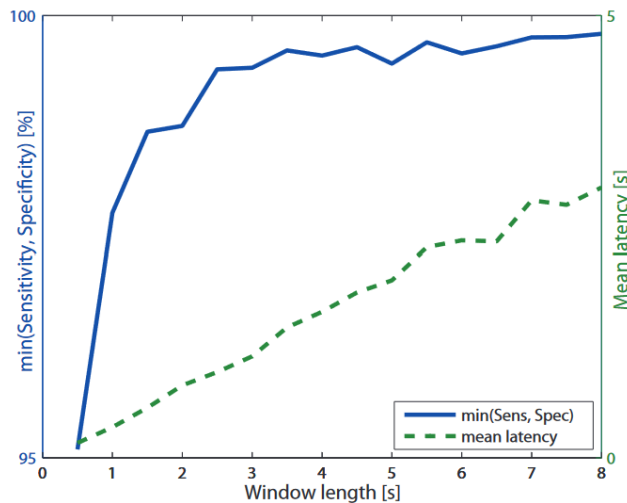
2. Fenêtrage des données

a. La longueur

Le fenêtrage des données est une technique préparatoire qui permet d'analyser des segments de données sur des intervalles courts et réguliers. Cette approche est particulièrement utile pour identifier les changements dans les motifs de marche associés au FOG. En effet, une fenêtre trop grande peut diluer des événements transitoires importants, comme le début ou la fin d'un épisode de FOG, rendant difficile la détection précise de ces événements. A l'inverse, une fenêtre trop petite risque de ne pas fournir suffisamment de contexte, compromettant ainsi l'analyse globale. En outre, la sélection de la longueur de la fenêtre doit également s'accompagner par la considération du temps de latence lié à l'exécution du modèle. En effet, une fenêtre plus longue peut améliorer les performances de détection, mais cela se fait au prix d'une latence accrue (Mazilu et al., 2012).

Des recherches, telles que celles de Mazilu et al. (2012), ont montré que des fenêtres fixes de 1 à 3 secondes sont optimales pour capturer les dynamiques du FOG. Cette durée permet de conserver un bon équilibre entre la précision de détection et un temps de latence acceptable. Comme illustré dans la figure ci-dessous (Figure 2), il est crucial de trouver ce juste milieu pour maximiser les performances de l'algorithme tout en minimisant les délais.

Figure 2: Performance de détection du FOG et latence de détection en fonction de la longueur de la fenêtre, pour un RF



b. Le chevauchement

La difficulté liée à l'optimisation des paramètres du fenêtrage dépend également des conditions de chevauchement entre les fenêtres. Cette méthode consiste à faire en sorte que chaque nouvelle fenêtre de données commence avant que la précédente ne se termine, créant ainsi une zone de superposition entre deux segments consécutifs. Ainsi, les segments de données aux extrémités sont ré-analysés dans le contexte de la fenêtre suivante, permettant une meilleure intégrité du signal traité. L'objectif de cette approche est de minimiser les pertes d'information importantes qui pourraient survenir aux frontières des fenêtres, là où les événements de FOG peuvent être tronqués ou mal représentés si les fenêtres sont totalement disjointes.

Diverses méthodes de chevauchement des fenêtres ont été explorées dans la littérature. Borzi et al. (2023), Reches et al. (2020), suggèrent l'utilisation de chevauchements dynamiques adaptés en fonction de la nature du signal. Dans cette approche, le chevauchement est amplifié pour les fenêtres contenant des segments de FOG et réduit pour les fenêtres sans FOG, ce qui permet de mieux gérer le déséquilibre des classes. D'autres études utilisent des chevauchements fixes (Pardoel, 2021; Rodríguez-Martín et al., 2017; San-Segundo et al., 2019; Tripoliti et al., 2013), avec des paramètres variant de 50 % à 80 %. Certaines recherches ont même exploré l'impact de l'absence de chevauchement (Ashfaque Mostafa et al., 2021; Camps et al., 2018). Bien qu'il n'existe pas de consensus clair sur la meilleure méthode de chevauchement, la tendance générale indique que des chevauchements compris entre 50 % et 80 % sont les plus efficaces pour maximiser la capture des événements FOG tout en maintenant une bonne performance de détection (Borzi et al., 2021; Pardoel, 2021; Rodríguez-Martín et al., 2017; San-Segundo et al., 2019; Tripoliti et al., 2013).

c. La labellisation

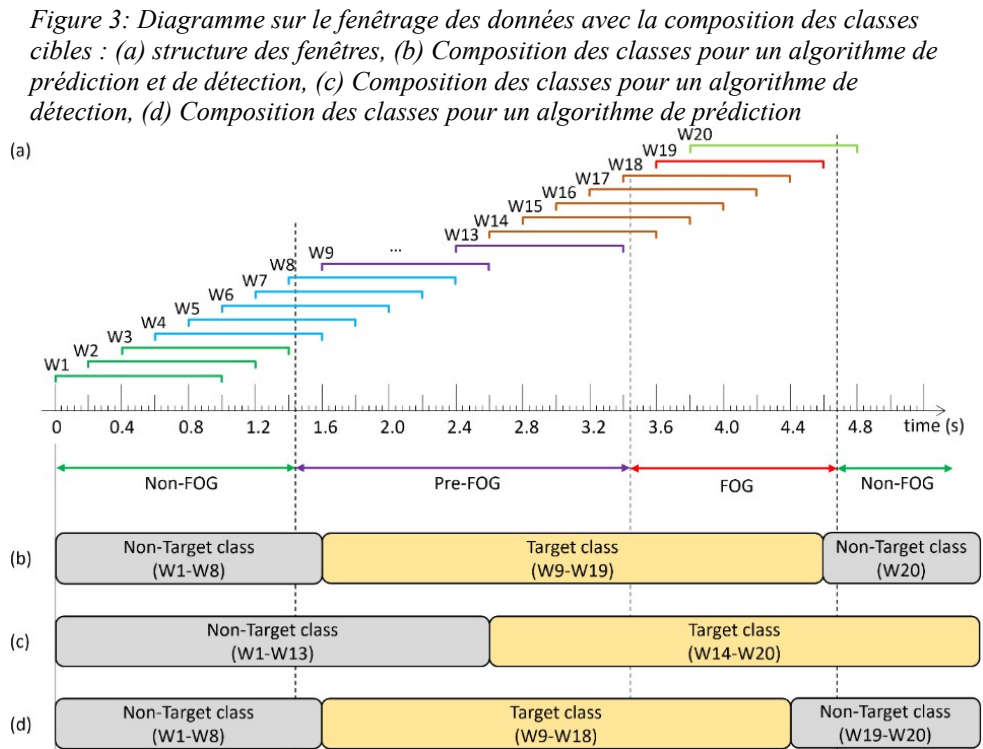
La labellisation des fenêtres est un enjeu majeur dans la détection et la prédiction des FOG, en particulier pour la gestion des épisodes de FOG de courte durée. Ces événements fugaces sont souvent difficiles à capturer avec précision en raison de leur brièveté, ce qui les rend facilement négligés par les systèmes automatiques (Delval et al., 2010).

Pour traiter ces épisodes, différentes stratégies ont été adoptées. Certains ont choisi de fusionner les épisodes courts de FOG pour créer des fenêtres homogènes, augmentant ainsi la probabilité de capturer des données significatives dans chaque fenêtre (Camps et al., 2018; San-Segundo et al., 2019; Sigcha et al., 2020). D'autres, en revanche, ont opté pour l'exclusion des épisodes de moins de 1 à 3 secondes, simplifiant ainsi l'analyse mais risquant d'éliminer des informations cruciales sur la dynamique de la marche et la fréquence des interruptions (Reches et al., 2020; Tripoliti et al., 2013). Cependant, Pardoel (2021) a démontré que la prise en compte ou l'exclusion des épisodes courts de FOG, définis comme ceux apparaissant dans un intervalle de moins de 3 secondes, n'affectait pas significativement les performances des algorithmes de détection. Cette observation suggère que des approches plus inclusives peuvent être adoptées sans compromettre l'efficacité des systèmes de détection, permettant une représentation plus fidèle et complète des schémas de marche des patients.

Un autre aspect critique de la labellisation est la composition des fenêtres et les règles de décision pour déterminer si une fenêtre est étiquetée comme FOG. Une fenêtre peut être classée comme FOG si elle contient un certain pourcentage de données FOG, généralement fixé à 50 % (Camps et al., 2018; San-Segundo et al., 2019; Sigcha et al., 2020). Cette règle assure que les fenêtres mixtes, contenant à la fois des périodes de marche normale et de FOG, ne brouillent pas la distinction entre les deux états. Pour s'assurer d'éviter toute confusion entre ces états, certaines approches suppriment les fenêtres qui ne sont pas composées à 100 % de segments Non-FOG (Camps et al., 2018; San-Segundo et al., 2019; Sigcha et al., 2020). Cependant, cette méthode peut aussi conduire à une perte d'informations significatives sur les transitions entre ces états et empêcher la compréhension des motifs initiaux du FOG.

Concernant la prédiction du FOG, elle repose sur la capacité à identifier des patterns moteurs ou des signaux précurseurs qui précèdent l'épisode de FOG. Ces derniers sont délicats à identifier, car ils peuvent être confondus avec des variations normales de la marche. Pour optimiser la labellisation, il est crucial de trouver un équilibre dans la durée de pré-FOG à considérer. Plus on s'éloigne temporellement de l'épisode de FOG, moins les prédictions sont précises, car les données capturées deviennent davantage représentatives de la marche normale plutôt que de la phase de transition critique. À l'inverse, des fenêtres trop courtes peuvent retarder la détection et la réponse préventive, réduisant ainsi l'efficacité des interventions (Borzi et al., 2021; Pardoel, 2021).

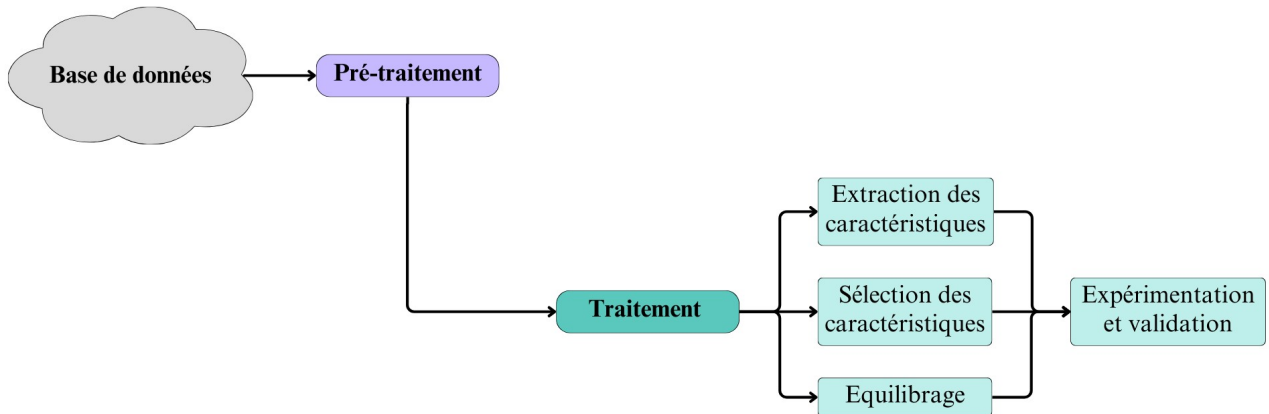
Pardoel (2021) explore la complexité liée à la construction d'un algorithme de détection et/ou de prédiction. En proposant une labellisation exhaustive, intégrant les transitions entre la marche normale, le pré-FOG, et le FOG lui-même, il montre que, selon les objectifs, les classes cibles à considérer varient. La figure ci-dessous (Figure 3) illustre comment les fenêtres sont étiquetées et quelles sont celles à considérer selon les objectifs de détection ou de prédiction du FOG, soulignant l'importance de cette distinction pour le développement de modèles efficaces.



En définitive, la clé pour une prédiction et/ou une détection efficace du FOG réside dans l'utilisation judicieuse de fenêtres de temps adaptées, où la longueur de la fenêtre et la granularité des labels sont soigneusement calibrées pour capturer les nuances dans la marche du patient. Une granularité plus fine permet de détecter des transitions rapides et des événements fugaces, mais elle peut aussi introduire des variations artificielles et compliquer l'analyse globale. À l'inverse, une granularité plus grossière pourrait simplifier le traitement des données et réduire les faux positifs, mais elle risquerait de manquer des détails critiques dans les dynamiques du FOG (Pardoel, 2021). Il est donc légitime de se demander s'il n'existe pas une granularité optimale qui permettrait de concilier détection précoce et efficacité, ouvrant ainsi la voie à des approches de prédiction et de détection plus robustes.

III. Traitement des données

Figure 4: Les étapes de traitement des données



L'étape suivante est l'extraction et la sélection minutieuse des caractéristiques, qui sont essentielles pour alimenter efficacement les modèles de ML. Cette démarche permet non seulement de réduire le temps d'apprentissage de l'algorithme, mais aussi d'identifier les éléments discriminants du signal qui différencient un état de marche normal d'un épisode de FOG.

1. Extraction des caractéristiques

La gestion optimale des caractéristiques pour la détection du FOG nécessite une sélection méticuleuse pour capturer les distinctions entre une démarche normale, d'un état de FOG. L'utilisation de caractéristiques issues des domaines temporel et fréquentiel permet d'aborder la complexité des manifestations du FOG. Les caractéristiques temporelles appliquées à chaque fenêtre, telles que la moyenne, l'écart type, etc. offrent des détails précis sur la mécanique de la marche. Simultanément, les caractéristiques fréquentielles sont également cruciales pour identifier les subtiles oscillations corporelles spécifiques au FOG, comme les tremblements capturés dans certaines bandes de fréquences (Ashfaq Mostafa et al., 2021; Mazilu et al., 2012; Pardoel, 2021; Rodríguez-Martín et al., 2017; Samà et al., 2018).

Cependant, l'intégration d'un large éventail de caractéristiques peut poser des défis, notamment en termes de charge computationnelle. Pour optimiser le temps de traitement et maximiser les performances, il est important de parvenir à un équilibre entre le nombre et la complexité des caractéristiques. Une approche minimaliste, privilégiant des caractéristiques faciles à calculer, peut être avantageuse pour la rapidité du système, mais elle doit être soigneusement gérée pour éviter de compromettre la précision et la capacité de généralisation du modèle.

2. Sélection des caractéristiques

Pour résoudre la difficulté liée à la grande dimensionnalité des données, il est nécessaire d'utiliser des méthodes de sélection des caractéristiques pour réduire cette dimensionnalité afin d'améliorer la performance des modèles, diminuer le surapprentissage et réduire le temps d'exécution des modèles. Ces méthodes peuvent être classées en trois grandes catégories : les méthodes de filtrage, les méthodes d'emballage (wrapper), et les méthodes embarquées (embedded). Les méthodes de filtrage, comme celles basées sur ReliefF, évaluent l'importance des caractéristiques avant la modélisation tout en considérant les interactions entre elles sans avoir besoin d'examiner toutes les combinaisons possibles. Ces méthodes sont généralement rapides et ne dépendent pas des modèles, ce qui les rend polyvalentes pour diverses applications (Urbanowicz et al., 2018). Les méthodes d'emballage utilisent un modèle prédictif pour évaluer et sélectionner les ensembles de caractéristiques, offrant souvent une meilleure performance au prix d'une plus grande charge computationnelle. Enfin, les méthodes embarquées intègrent la sélection de caractéristiques directement dans l'algorithme d'apprentissage, améliorant ainsi l'efficacité mais avec une dépendance au modèle utilisé (Urbanowicz et al., 2018). Chaque approche a ses avantages : les méthodes de filtrage sont efficaces et flexibles, les méthodes d'emballage sont précises mais coûteuses en ressources, et les méthodes embarquées offrent un équilibre entre performance et efficacité, bien que leur applicabilité puisse être limitée par le choix du modèle spécifique. Ainsi, la sélection appropriée des caractéristiques est cruciale pour adapter les algorithmes aux spécificités du problème traité, maximisant à la fois la performance du modèle et l'efficacité computationnelle.

Pardoel (2021) a démontré à l'aide de ReliefF que certaines caractéristiques, telles que celles extraites des semelles de pression plantaire, affichent une meilleure capacité de distinction entre les épisodes de FOG et la démarche normale que les données issues des IMUs. Toutefois, les caractéristiques utilisées dans les modèles ont été identifiées grâce à ReliefF appliqué globalement sur tous les patients. Or, la variabilité interindividuelle dans les manifestations du FOG suggère que les caractéristiques pertinentes peuvent varier d'un patient à l'autre (Ehgoetz Martens et al., 2018).

Ainsi, il serait pertinent d'explorer une application individualisée de ReliefF en adoptant une approche dynamique. Cette approche permettrait de déterminer les caractéristiques importantes spécifiquement pour chaque patient avant de les intégrer dans le modèle.

3. Équilibrage des données

Lorsqu'on traite des données, il est fréquent de rencontrer des ensembles déséquilibrés où certaines classes sont beaucoup plus représentées que d'autres. Cette disproportion peut entraîner des biais dans les modèles de prédiction, où les classes majoritaires sont mieux reconnues au détriment des classes minoritaires, réduisant ainsi la performance globale du modèle sur des données moins fréquentes mais souvent cruciales. On peut alors se poser la question de l'effet que ce déséquilibre peut avoir sur les résultats de la sélection des caractéristiques si les classes sont déséquilibrées. En effet, un déséquilibre marqué entre les classes peut fausser l'évaluation de l'importance des caractéristiques, car les algorithmes de sélection peuvent privilégier les caractéristiques discriminant bien les classes majoritaires, au détriment des classes minoritaires. Cela peut entraîner une surreprésentation des caractéristiques non pertinentes et une sous-

représentation des caractéristiques cruciales pour la détection des épisodes de FOG (Chawla et al., 2002).

Diverses méthodes existent pour équilibrer la distribution des classes. Le sur-échantillonnage et le sous-échantillonnage sont deux approches couramment utilisées. Le sur-échantillonnage, par exemple à travers la technique Synthetic Minority Over-sampling Technique (SMOTE) (Chawla et al., 2002), génère des instances synthétiques en interpolant entre des exemples existants de la classe minoritaire, augmentant ainsi leur représentation. À l'inverse, le sous-échantillonnage élimine aléatoirement des exemples de la classe majoritaire pour réduire les biais (Liu et al., 2009). Enfin, ajuster les poids des classes dans les algorithmes est également une méthode efficace pour améliorer la représentation des classes minoritaires (Krawczyk, 2016). Cette approche permet de renforcer l'importance des classes minoritaires lors de l'apprentissage, équilibrant ainsi l'impact des différentes classes sur le modèle final. En somme, ces approches permettent d'ajuster la distribution des classes, facilitant un apprentissage plus équilibré et améliorant la capacité des modèles à identifier correctement les épisodes de FOG. Cela renforce ainsi la fiabilité et l'efficacité de la détection dans les applications cliniques.

4. Expérimentation et évaluation de l'algorithme

L'ensemble des traitements précédemment évoqué, avait pour but ultime d'optimiser l'application des algorithmes de ML pour la détection et la prédiction du FOG. Cependant, l'utilisation de ces algorithmes est particulièrement délicate et nécessite une attention méticuleuse, car leurs performances peuvent varier considérablement en fonction des méthodes employées.

Concernant les algorithmes de détection, les RF et les SVM sont parmi les algorithmes les plus performants, affichant des sensibilités allant de 66% à 98% et des spécificités de 79% à 99% (Mazilu et al., 2012; Reches et al., 2020; Rodríguez-Martín et al., 2017; Samà et al., 2018; Tripoliti et al., 2013). En particulier, Mazilu et al. (2012) ont montré que les modèles spécifiques à chaque individu peuvent atteindre des sensibilités (98.35%) et spécificités (99.72%) supérieures, par rapport à des modèles indépendants avec une sensibilité (66.25%) et spécificité (95.38%). Enfin, les algorithmes de prédiction ont montré des sensibilités de 83% à 85% et des spécificités de 67% à 86% (Borzi et al., 2021; Palmerini et al., 2017). Des résultats prometteurs, qui renforcent l'idée qu'il existe des patterns initiateurs du FOG.

La performance de ces algorithmes est validée par des méthodes rigoureuses telles que la validation croisée K-fold et la validation Leave-One-Subject-Out (LOSO). La validation K-fold divise les données en K sous-ensembles, entraînant le modèle sur K-1 sous-ensembles et testant sur le sous-ensemble restant, répétant ce processus K fois pour une évaluation complète et minimisant le biais. La validation LOSO, quant à elle, utilise les données de tous les sujets sauf un pour l'entraînement et le sujet laissé de côté pour le test, répétant ce processus pour chaque sujet. Ces méthodes permettent de mesurer des métriques essentielles comme la précision, la sensibilité, la spécificité et l'AUC (Area Under the Curve), (Mazilu et al., 2012; Reches et al., 2020; Rodríguez-Martín et al., 2017; Samà et al., 2018; Sigcha et al., 2020; Tripoliti et al., 2013).

La précision mesure la proportion de prédictions correctes parmi l'ensemble des prédictions faites par l'algorithme. Ensuite, la sensibilité mesure la capacité d'un algorithme à identifier

correctement les épisodes de FOG, autrement dit, elle indique la proportion de vrais positifs détectés. Une haute sensibilité signifie que l'algorithme manque peu d'épisodes de FOG. La spécificité, en revanche, évalue la capacité de l'algorithme à éviter les faux positifs, c'est-à-dire à ne pas signaler de FOG lorsqu'il n'y en a pas. Une haute spécificité garantit que les patients ne sont pas faussement diagnostiqués, minimisant les alarmes inutiles. Enfin, l'AUC est une mesure de performance pour les modèles de classification. Elle représente la probabilité qu'un classificateur, tel que le RF, classe correctement un positif aléatoire plus haut qu'un négatif aléatoire. Une AUC de 100 % indique un classificateur parfait, tandis qu'une AUC de 50 % indique un classificateur sans compétence. Ensemble, ces métriques offrent une vue d'ensemble de la fiabilité et de la précision des algorithmes, essentielle pour leur application clinique.

Bien que les recherches scientifiques aient permis des avancées majeures dans la détection et la prédiction du FOG, plusieurs aspects nécessitent encore des recherches approfondies. Il apparaît clairement que l'individualisation des modèles, la gestion du déséquilibre des classes, l'exploration des combinaisons de capteurs, et l'élargissement des cohortes d'étude sont des étapes essentielles pour continuer à améliorer la précision et la fiabilité des diagnostics du FOG.

IV. Objectifs

L'objectif principal de cette étude est de développer et d'évaluer un algorithme de détection du FOG chez les patients atteints de la maladie de Parkinson en utilisant les IMUs. Cette recherche vise à améliorer la précision et la fiabilité des diagnostics, ainsi que la gestion clinique du FOG, en répondant à plusieurs questions critiques sur l'optimisation des techniques de détection.

En premier lieu, nous nous pencherons sur l'identification des caractéristiques cinématiques spécifiques au FOG à partir des données des IMUs. La variabilité des caractéristiques entre les patients sera évaluée afin de déterminer si une approche personnalisée est nécessaire pour chaque individu. Nous émettons l'hypothèse que les caractéristiques discriminantes extraites des IMUs varient entre les patients atteints de FOG, nécessitant ainsi une personnalisation des modèles de détection.

En deuxième lieu, nous analyserons l'impact des techniques de rééquilibrage des données, telles que le sur-échantillonnage et le sous-échantillonnage, sur les performances des algorithmes de ML. L'objectif est d'évaluer si ces techniques peuvent améliorer la sensibilité et la spécificité des modèles. Nous postulons que l'application de ces techniques de rééquilibrage des données améliore les performances des algorithmes de détection du FOG, en augmentant la sensibilité et la spécificité. Ce mémoire vise à apporter des recommandations pour développer un algorithme et pour évaluer la sévérité du FOG dans le cadre clinique.

V. Matériel et Méthodes

1. Protocole

Cette étude clinique non interventionnelle, monocentrique, implique des patients atteints de la MP déclarant des épisodes FOG recrutés au Centre Hospitalier Universitaire de Rennes. Les patients sont sélectionnés lors de leurs visites chez le neurologue. Au total, 20 patients seront recrutés, mais pour la présente étude, nous analyserons les données de 5 patients. Ces derniers participent à deux visites séparées d'environ deux semaines, pendant les phases ON et OFF du traitement à la lévodopa. La première visite (visite d'inclusion) a lieu pendant la phase ON, et la seconde visite (visite de suivi) pendant la phase OFF. L'évaluation de la marche en phase OFF est effectuée après 12 heures sans traitement, de préférence le matin avant la première dose de lévodopa. Lors de chaque visite, les patients marchent à une vitesse confortable sur un parcours conçu pour induire le FOG, comprenant plusieurs virages, des arrêts à deux reprises, et un couloir étroit (Figure 5 et Figure 6).

Figure 5: Parcours de la tâche de marche. La case de gauche représente la chaise au début et à la fin de l'essai. Les triangles sont des cônes et la zone grise délimite un couloir étroit. Les octogones rouges indiquent des arrêts volontaires momentanés (2 s).

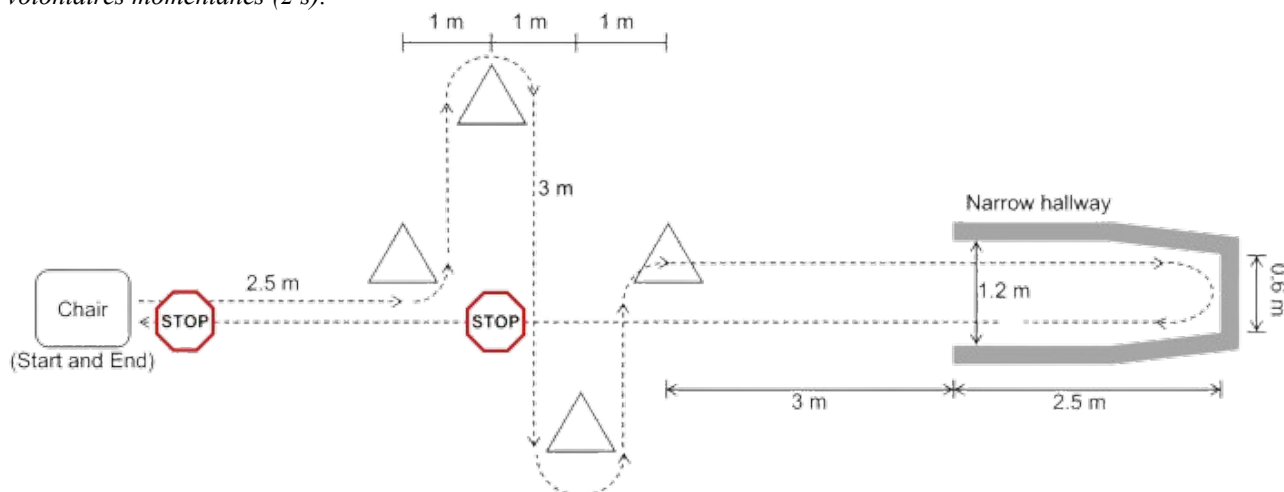
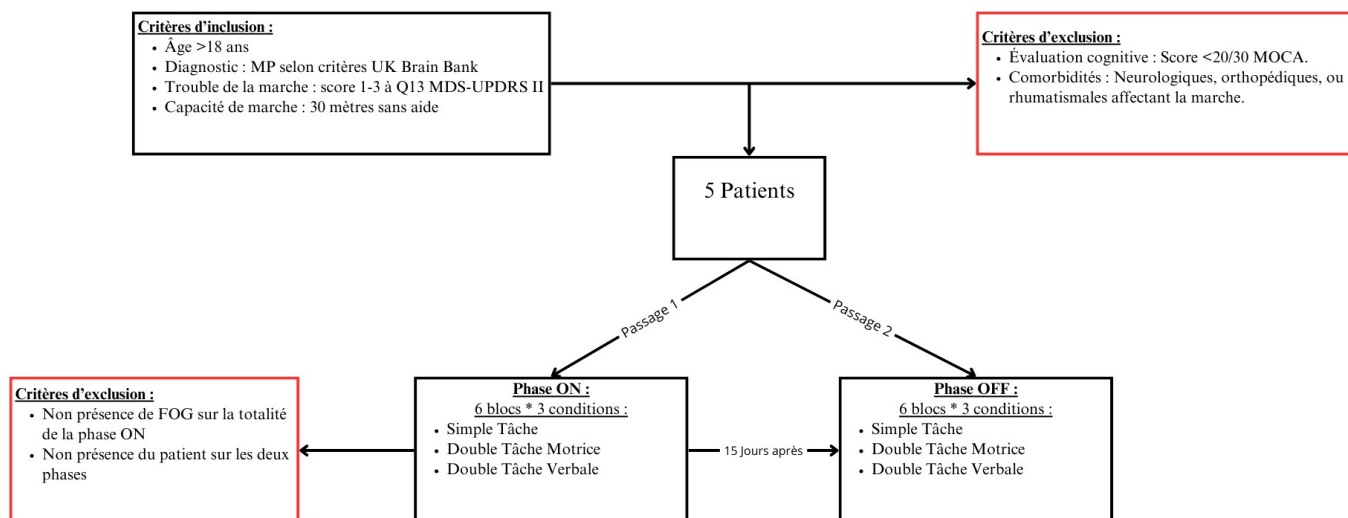


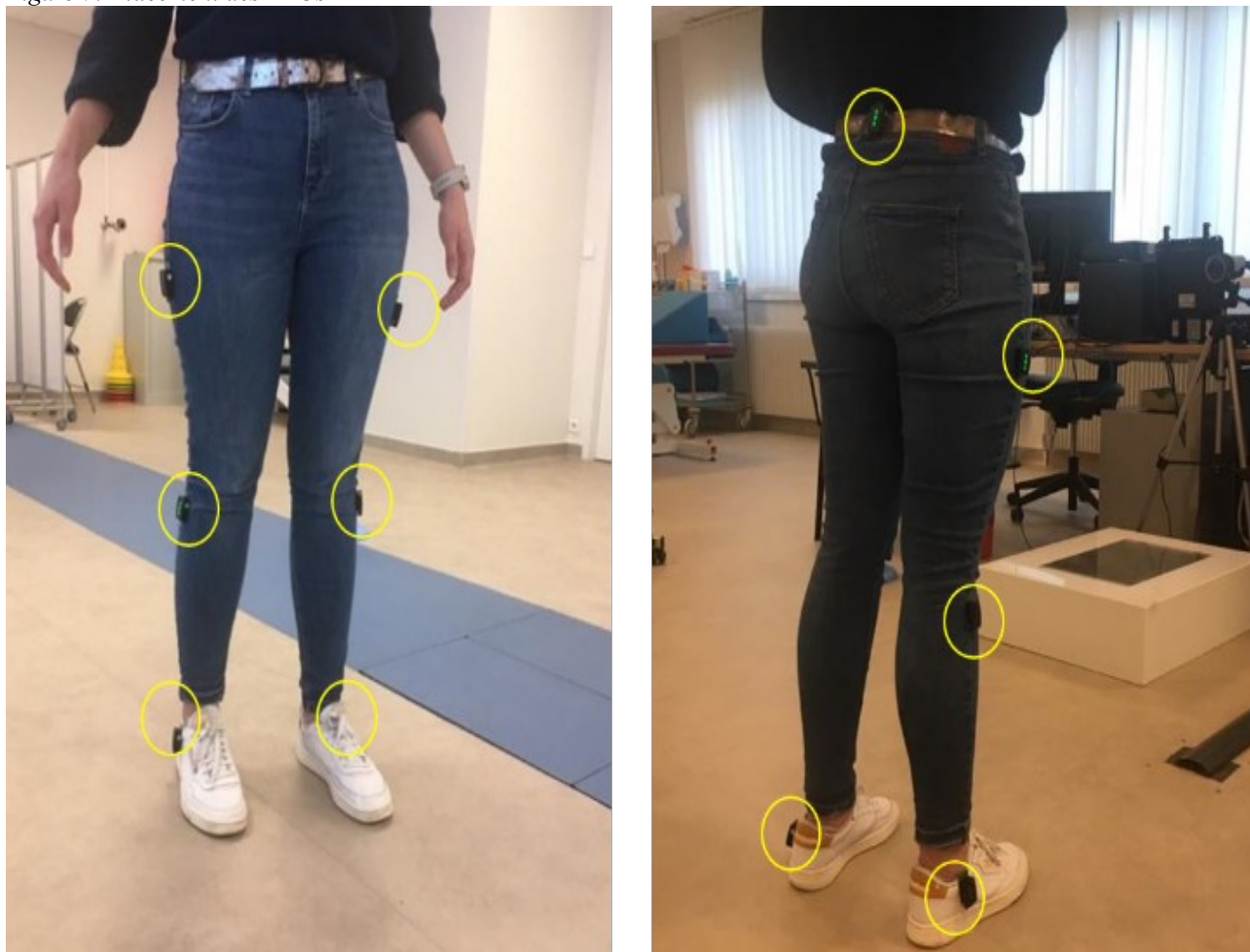
Figure 6: Organisation de l'étude



2. Collecte des données et labellisation du FOG

Les patients sont équipés de capteurs IMUs TRIGNO Avanci (Delsys), comprenant un accéléromètre triaxial, un gyroscope et un magnétomètre. Les données, transmises par Bluetooth, sont collectées à l'aide de l'API de Delsys et enregistrées à une fréquence d'échantillonnage de 148,15 Hz dans le logiciel QTM 2020 (Qualisys, Suède). Les capteurs sont placés sur chaque cuisse (côté latéral, tiers supérieur), chaque péroné (tiers supérieur), chaque pied (sous la malléole latérale) et au niveau lombaire (L5) (Figure 7).

Figure 7: Placement des IMUs



À chaque visite, les patients sont filmés pour chaque essai à l'aide de 14 caméras Miquis M3 synchronisées (Qualisys) enregistrant à une résolution de 1824x1088 pixels et à une fréquence de 25 Hz. Les épisodes de FOG sont déterminés ultérieurement à partir des vidéos par deux spécialistes des troubles du mouvement, qui examinent les vidéos indépendamment. Le début d'un épisode de FOG est défini comme le moment où le pied ne quitte pas le sol malgré une intention claire de marcher, et la fin de l'épisode comme le moment où le pied commence ou reprend un pas effectif. En cas de désaccord, une discussion est organisée pour établir un consensus.

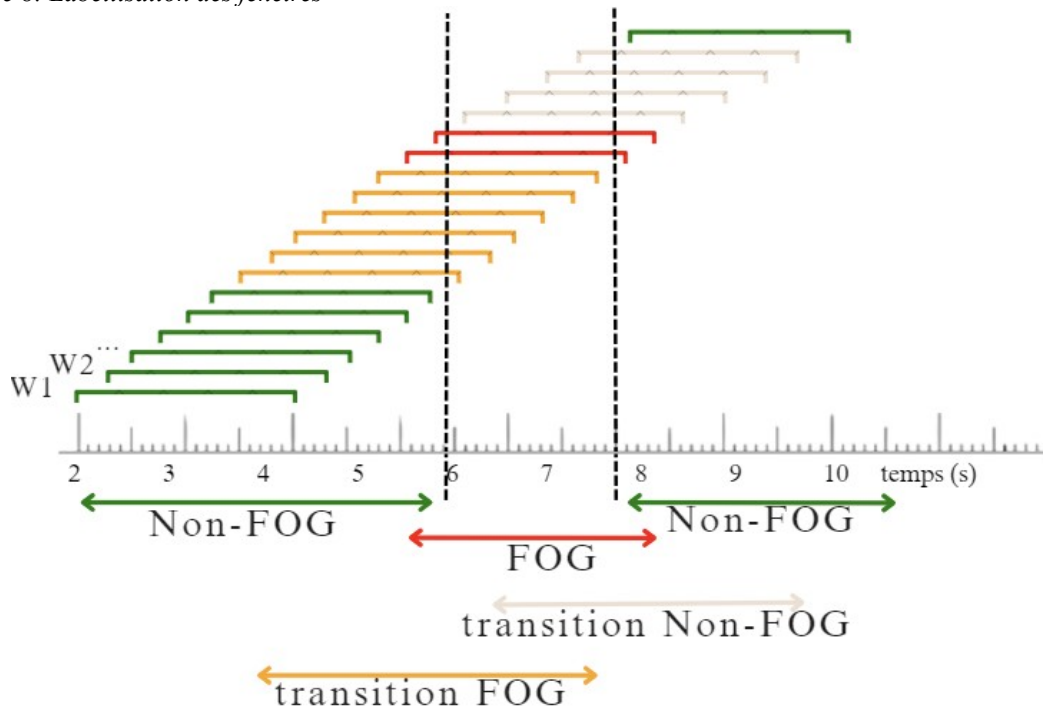
3. Pré-traitement des données

Conformément aux études précédentes, les données sont échantillonnées à 50 Hz, et un filtre passe-bas de deuxième ordre avec une fréquence de coupure de 20 Hz est appliqué pour éliminer les bruits. Les données sont ensuite normalisées en utilisant la moyenne nulle et la variance unitaire pour assurer une comparaison entre les enregistrements et les patients.

Les données sont divisées en fenêtres de 2 secondes avec un chevauchement de 80 % (décalage de 0,4 secondes). La labellisation des fenêtres est réalisée en regroupant différentes méthodologie (Pardoel, 2021; San-Segundo et al., 2019; Sigcha et al., 2020), pour répondre aux objectifs de prédiction et de détection (Figure 8). Les fenêtres sont labellisées FOG si les événements de début et de fin de FOG apparaissent à l'intérieur de la même fenêtre ou si plus de 50 % des échantillons sont étiquetés comme FOG. Les fenêtres contenant moins de 50 % d'échantillons FOG sont considérées comme des transitions Non-FOG et sont finalement exclues pour réduire toute confusion avec un potentiel pattern de FOG.

Les fenêtres de transition FOG sont appliquées à toutes les fenêtres contenant seulement le début d'un épisode de FOG, offrant des détails sur les patterns de marche jusqu'à 2 secondes avant le FOG. Cette approche permet d'identifier des caractéristiques spécifiques précédant les épisodes de FOG, comme suggéré par les recherches antérieures (Borzi et al., 2021; Palmerini et al., 2017).

Figure 8: Labellisation des fenêtres



4. Traitement des données

a. Extraction des caractéristiques

Pour chaque capteur et pour chaque axe, ainsi que la norme formée par les trois axes, nous créons les caractéristiques temporelles et nous appliquons la transformée de Fourier discrète pour extraire les caractéristiques fréquentielles. Cette approche nous permet de générer un total de 2352 caractéristiques (Tableau 1).

Tableau 1: Détails concernant l'extraction des caractéristiques

Caractéristique	Description de la caractéristique	Source	Total Caractéristiques (n capteur * ACC et GYRO * n axes (X, Y, Z, norme))
Domaine Temporel			
Moyenne		Ashafaque Mostafa et al. (2021), Rodriguez et al. (2017), Samà et al. (2018), Camps et al. (2018)	56
Variance	Dispersion des données par rapport à la moyenne	Ashafaque Mostafa et al. (2021), Samà et al. (2018), Camps et al. (2018)	56
Ecart-Type	Carré de l'écart type	Ashafaque Mostafa et al. (2021)	56
Energie	Somme des carrés des valeurs	Ashafaque Mostafa et al. (2021)	56
Range	Différence entre la valeur maximale et la valeur minimale	Ashafaque Mostafa et al. (2021)	56
Root Mean Square	Carré moyen des valeurs	Ashafaque Mostafa et al. (2021)	56
Moyenne Ajustée	Moyenne calculée après avoir retiré un pourcentage des valeurs extrêmes	Ashafaque Mostafa et al. (2021)	56
Médiane	Valeur centrale	Ashafaque Mostafa et al. (2021)	56
Valeur Absolue Moyenne	Moyenne des valeurs absolues	Ashafaque Mostafa et al. (2021)	56
Écart Absolu Médian	Moyenne des écarts absolues par rapport à la médiane	Ashafaque Mostafa et al. (2021)	56
1 ^{er} Quartile	Valeur au-dessous de laquelle se trouvent 25 % des données lorsqu'elles sont arrangées en ordre croissant	Ashafaque Mostafa et al. (2021)	56
3 ^e Quartile	Valeur au-dessous de laquelle se trouvent 75 % des données lorsqu'elles sont arrangées en ordre croissant	Ashafaque Mostafa et al. (2021)	56
Interquartile	Valeur entre le 1 ^{er} et 3 ^e quartile	Ashafaque Mostafa et al. (2021)	56
Skewness	Asymétrie de la distribution des données par rapport à la moyenne	Ashafaque Mostafa et al. (2021), Rodriguez et al. (2017), Camps et al. (2018)	56
Kurtosis	Coefficient d'aplatissement de la courbe formée par les données	Ashafaque Mostafa et al. (2021), Rodriguez et al. (2017)	56
Moyenne Incrémentale	Différence entre la valeurs moyenne de la fenêtre n et n-1	Rodriguez et al. (2017), Samà et al. (2018), Camps et al. (2018)	56
Coefficient de Variation	Rapport entre l'écart-type et la moyenne	Caramia et al. 2018	56
Total Caractéristiques			952
Domaine Fréquentiel			
Amplitude Première et Deuxième Harmonique	Amplitude maximale	Samà et al. (2018), Camps et al. (2018), Borzi et al (2021)	112
Fréquence Première et Deuxième Harmonique	Fréquence Maximale	Samà et al. (2018), Camps et al. (2018), Borzi et al (2021)	112
Distance entre les Fréquences et Amplitudes de la Première et Deuxième Harmonique	Écart entre les fréquences du premier et deuxième pics d'amplitude	Samà et al. (2018)	112
Largeur Première Harmonique	Largeur de la bande de fréquence à laquelle se trouve l'harmonique principale	Borzi et al (2021)	56
Écart type amplitude (sans fréquence spécifique et entre 0,68–3 Hz / 3–8 Hz / 8–20 Hz et 0,1–8 Hz)	Mesure de la dispersion des amplitudes des harmoniques dans différentes bandes de fréquences	Rodriguez et al. (2017), Samà et al. (2018)	280
Entropie Spectrale de Puissance	Mesure de la désorganisation ou de l'irrégularité	Borzi et al (2021)	56
Centre de Densité Spectrale	la moyenne pondérée des fréquences présentes dans le signal, où les pondérations sont les magnitudes du spectre de fréquence du signal	Rodriguez et al. (2017)	56
Centre de Densité Spectrale de Puissance	la moyenne pondérée des fréquences présentes dans le signal, où les pondérations sont les magnitudes au carré du spectre de fréquence du signal		56
Rapport Harmonique Fréquence et amplitude	Détail de la première harmonique divisée par la deuxième harmonique	Borzi et al (2021)	112
Crête Spectrale de Puissance Pondéré GPT			56
Crête Spectrale de Puissance Pondéré (Borzi)	Produit de l'amplitude et de la fréquence de l'harmonique principale	Borzi et al (2021)	56
Ratio de Faible Puissance entre 0 et 2Hz	Rapport entre la puissance dans la largeur de bande 0-2 Hz et la puissance totale du signal	Borzi et al (2021)	56
Freeze Band Power	Somme de la puissance à l'intérieur de la bande de freezing divisée par la fréquence d'échantillonnage	Ashafaque Mostafa et al. (2021)	56
Locomotion Band Power	Somme de la puissance de la bande de locomotion divisée par la fréquence d'échantillonnage	Ashafaque Mostafa et al. (2021)	56
Freeze Index	Puissance de la bande de freezing (3-8Hz) divisée par la puissance de la bande de locomotion (0.5-3 Hz)	Ashafaque Mostafa et al. (2021)	56
Band Power	Somme de la puissance de la bande de locomotion et de la bande freezing	Ashafaque Mostafa et al. (2021)	56
Energie		Ashafaque Mostafa et al. (2021)	56
Total Caractéristiques			1400
Total Caractéristiques Temporelles + Fréquentielles			2352

b. Sélection des caractéristiques

Étant donné la grande quantité de données brutes issues des capteurs IMUs, il est essentiel de réduire la dimensionnalité des données sans perdre d'informations pertinentes. Pour ce faire, nous avons utilisé la méthode ReliefF, qui évalue l'importance de chaque caractéristique en fonction de sa capacité à distinguer les échantillons proches appartenant à des classes différentes. ReliefF prend en compte les interactions entre les caractéristiques, ce qui est crucial pour capturer les dynamiques complexes du FOG (Urbanowicz et al., 2018). Nous avons utilisé les paramètres de base de la fonction ReliefF, avec 100 voisins. Dans cette méthode, le terme "voisins" fait référence aux échantillons de données les plus proches à prendre en compte pour évaluer chaque caractéristique.

Pour appliquer cette sélection, nous avons préalablement divisé les données en entraînement (70 %) et en test (30%), afin d'appliquer relief F uniquement sur les données d'entraînement. Quant aux données de test, elles serviront à évaluer la capacité du modèle final à s'adapter à de nouvelles données non vues lors de l'apprentissage.

Cependant, nos données affichent des déséquilibres entre les classes FOG et Non-FOG. Donc, nous avons examiné l'effet potentiel de ce biais sur la performance de la méthode de sélection des caractéristiques. Pour aborder cette problématique, nous avons appliqué par patient différentes techniques d'échantillonnage, avant l'exécution de ReliefF, en excluant le sous-échantillonnage afin de ne pas perdre d'informations pertinentes (Ashfaq Mostafa et al., 2021) :

- Sur-échantillonnage avec SMOTE (Ashfaq Mostafa et al., 2021; Chawla et al., 2002).
- Echantillonnage optimisé : Cette méthode combine le sur-échantillonnage de la classe minoritaire avec SMOTE et un sous-échantillonnage aléatoire de la classe majoritaire (Chawla et al., 2002). La proportion optimale a été choisie en fonction du score AUC, calculé grâce à une validation croisée K-fold (K=10). Cette approche vise à limiter la création de données synthétiques.

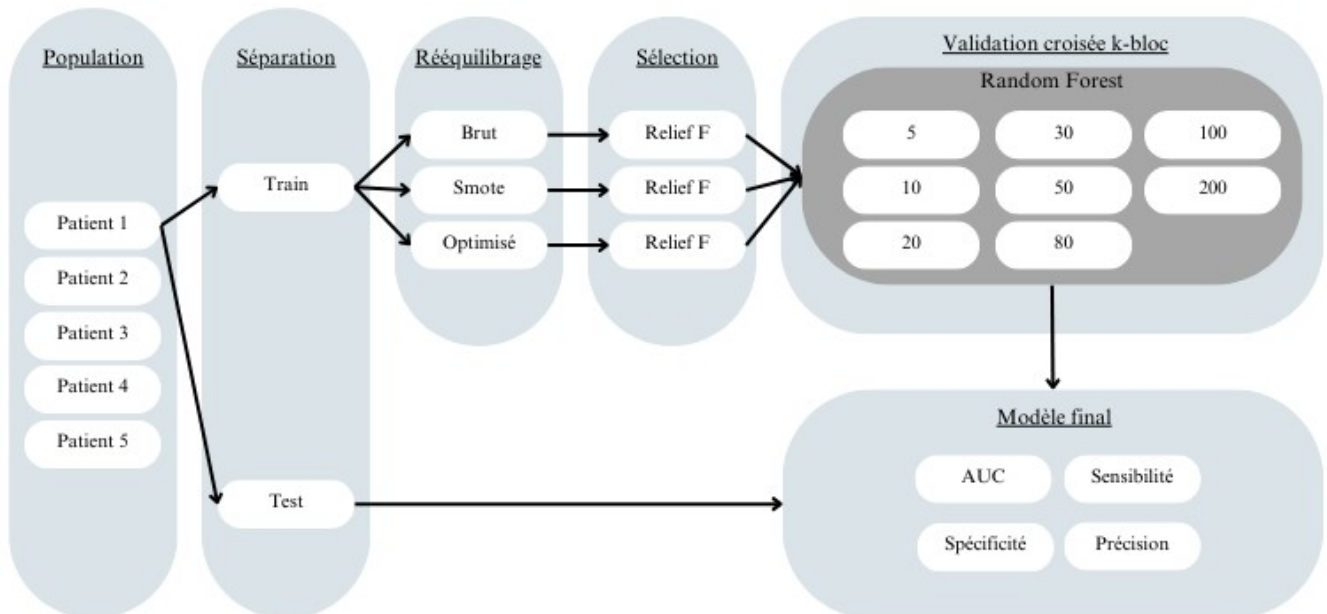
En utilisant ces techniques de rééquilibrage et en les comparant aux données sans rééquilibrage, nous avons pu aborder le déséquilibre des classes et permettre une sélection de caractéristiques plus équilibrée. Cela nous a permis d'identifier un ensemble de caractéristiques optimales, essentielles pour entraîner et tester nos modèles de détection du FOG.

5. Expérimentation et validation

À partir du classement établi par ReliefF, nous avons procédé à une validation croisée K-bloc ($k=10$) sur les données d'entraînement. Cette approche nous a permis d'essayer différentes quantités de caractéristiques (5, 10, 20, 30, 50, 80, 100 et 200 caractéristiques les plus importantes) pour chaque méthode de rééquilibrage (brute, optimisée, sur-échantillonnée). Ainsi, pour chaque ensemble de caractéristiques, un modèle de Random Forest a été entraîné avec les paramètres par défaut, et les métriques de performance telles que la précision, la sensibilité, la spécificité et le score AUC ont été calculées.

La validation croisée nous a fourni des performances moyennes d'entraînement pour chaque combinaison de caractéristiques et de méthode de rééquilibrage, nous permettant ainsi de comprendre l'impact de chaque méthode sur les performances des modèles. Ensuite, nous avons appliqué ces ensembles de caractéristiques aux données de test pour évaluer la performance finale des modèles (Figure 9).

Figure 9: Organigramme du processus de création de l'algorithme

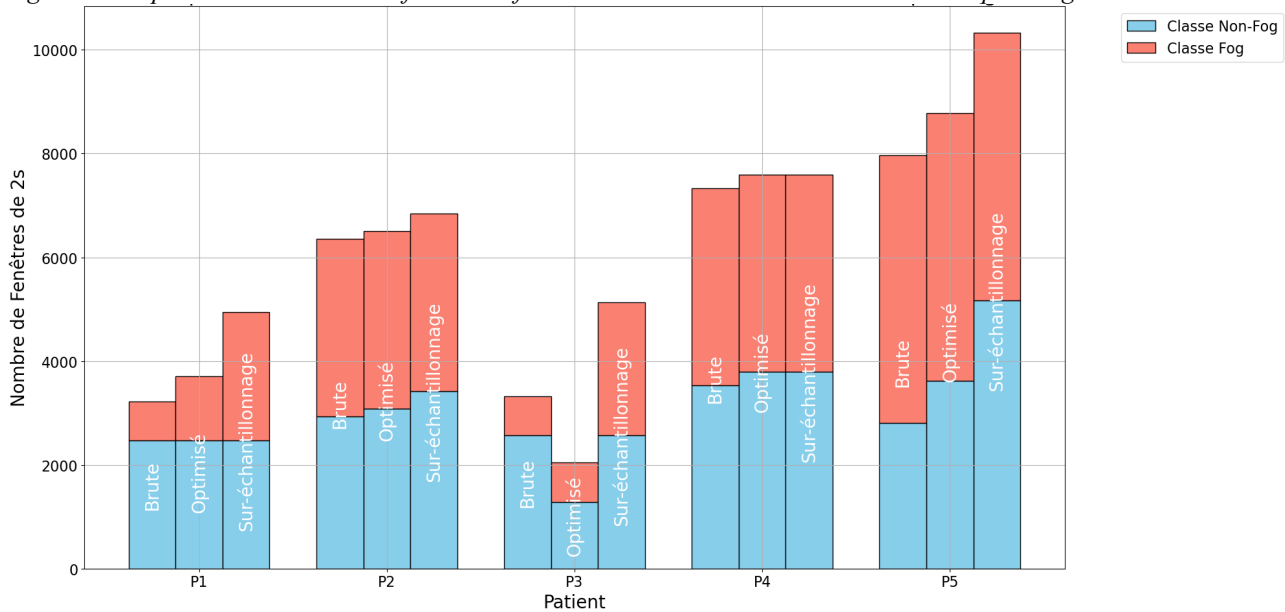


VI. Résultats

1. Rééquilibrage des données

Le rééquilibrage des données a été effectué en utilisant deux méthodes différentes : optimisée et suréchantillonnée, comparées avec les données brutes, pour lesquelles les détails de la segmentation de nos fenêtres sont présentés (Tableau 2 et Tableau 3). Les graphiques ci-dessous (Figure 10) montrent la distribution des classes Non-FOG et FOG pour chaque patient selon les différentes méthodes de rééquilibrage.

Figure 10: Répartition du nombre de fenêtre en fonction des classes et des méthodes de rééquilibrage



On peut voir que la distribution brute varie en fonction des patients. En effet, pour P1, P3 on constate une dominance nette de la classe Non-FOG, alors que pour P2 et P4 la distribution est plus équilibrée avec une légère dominance de la classe FOG. Pour le patient P5, on obtient une prédominance très nette de la classe FOG.

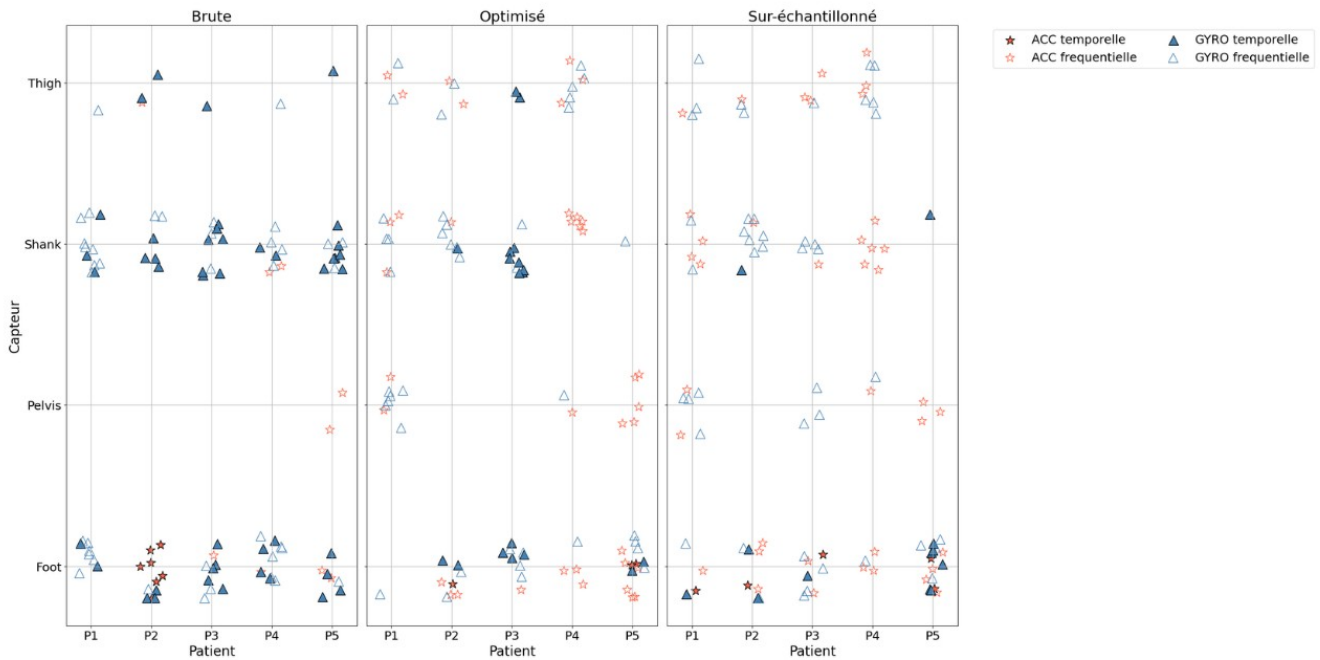
Ensuite, la méthode optimisée et sur-échantillonnée agissent différemment sur la proportion des classes, excepté pour P4 où la méthode optimisée affiche la même distribution que la méthode sur-échantillonnée. La méthode optimisée crée de manière globale des nouvelles données, sauf pour P3, pour lequel une réduction de classe majoritaire est appliquée. Quant à la méthode sur-échantillonnée, il y a une augmentation du nombre total de fenêtres avec une distribution parfaitement équilibrée entre la classe avec FOG et Non-FOG. Enfin, les patients P1 et P3 affichent de manière générale moins de fenêtres. Alors que, le patient P5 présente le plus de fenêtres.

Il semblerait que la méthode optimisée soit plutôt adaptée pour des patients ayant recensé un nombre important de données. Tandis que la méthode sur-échantillonnée serait plutôt favorable lorsqu'il y a très peu de données.

2. Sélection des caractéristiques

Pour chaque méthode le graphique suivant (Figure 11) tend à montrer la nature des caractéristiques les plus discriminantes pour chaque patient, par rapport à la localisation du capteur (cuisse, tibia, pelvis, pied) et à la nature des caractéristiques (accéléromètre temporel, accéléromètre fréquentiel, gyroscope temporel, gyroscope fréquentiel).

Figure 11: Répartition des 20 caractéristiques les plus importantes en fonction de la localisation des IMUs et des patients affichant également la nature des caractéristiques pour chaque méthode de rééquilibrage



Concernant la méthode brute, les caractéristiques les plus importantes se concentrent principalement sur les capteurs placés au niveau des tibias et des pieds, avec une prédominance des caractéristiques provenant du gyroscope (GYRO) et un équilibre entre caractéristiques temporelles et fréquentielles pour chaque patient. En comparaison, les résultats des méthodes optimisée et sur-échantillonnée sont similaires, mais présentent des différences notables par rapport à la méthode brute. En effet, on observe davantage de variabilité dans la nature des caractéristiques et l'emplacement des capteurs entre les patients.

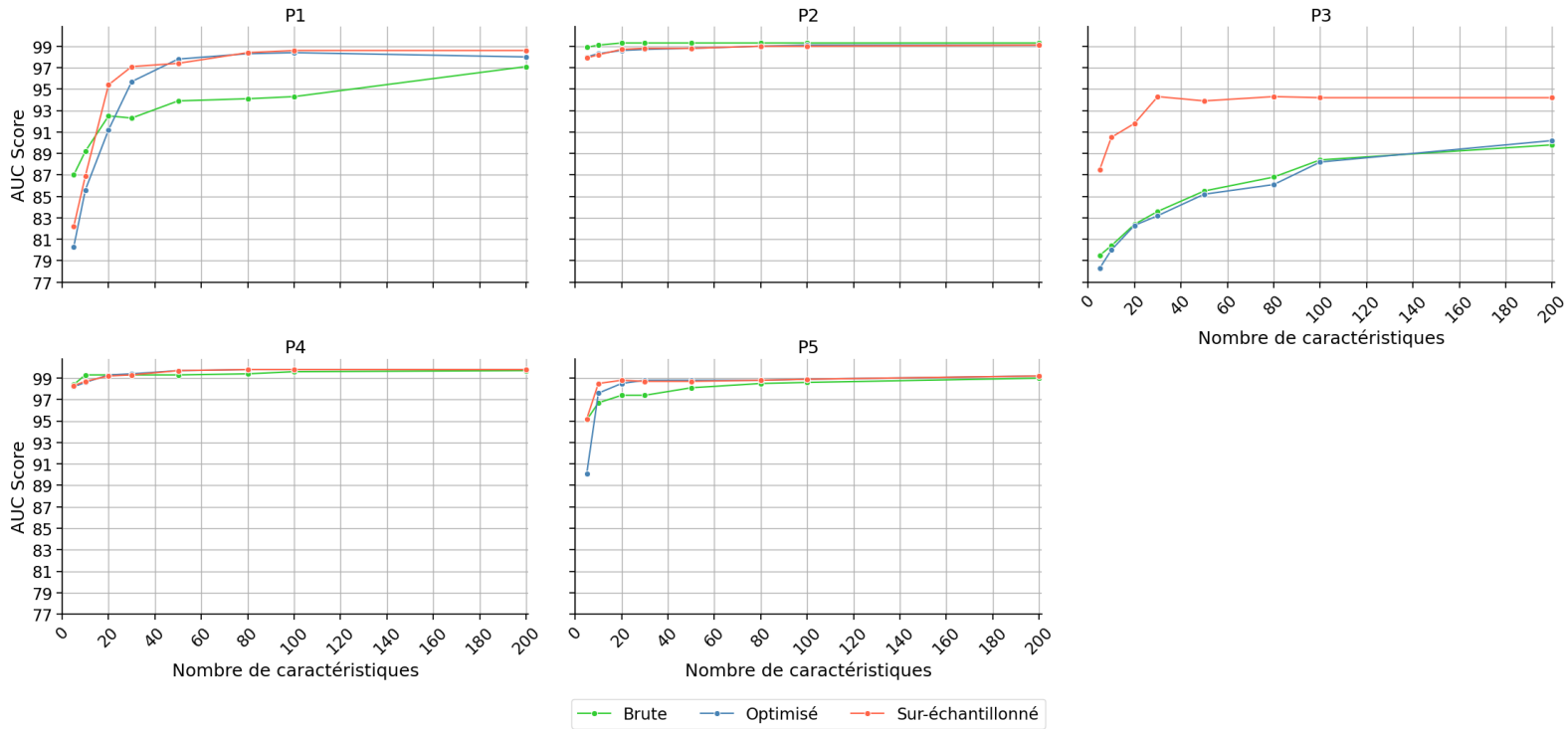
Pour les méthodes optimisée et sur-échantillonnée, il n'y a presque aucune caractéristique temporelle, excepté pour les patients P3 et P5, qui dépend fortement de la méthode utilisée. Enfin, les caractéristiques fréquentielles sont majoritairement réparties de manière équilibrée entre les GYRO et les accéléromètres (ACC).

La localisation des capteurs affiche une légère prédominance des capteurs situés au niveau des tibias et des pieds, même si la tendance semble moins évidente et nécessite une étude au cas par cas. Par exemple, pour le patient P5, une grande majorité des caractéristiques concernent le pied. Pour le patient P1, il y a une prédominance des caractéristiques provenant du bassin et du tibia. En revanche, pour les patients P3 et P2, les caractéristiques les plus importantes proviennent majoritairement du pied et du tibia. Enfin, pour le patient P4, les caractéristiques dominantes proviennent de la cuisse et du tibia.

3. Expérimentation et validation

Le graphique ci-dessous (Figure 12) présente les performances du RF appliquées sur plusieurs nombres de caractéristiques. Les courbes représentent les scores AUC pour les trois méthodes de rééquilibrage : brute, optimisée, et sur-échantillonnée.

Figure 12: Comparaison des résultats du Random Forest issus des différentes méthodes appliquées pour n nombre de caractéristiques



On remarque que les performances s'améliorent avec l'augmentation du nombre de caractéristiques sélectionnées, mais l'évolution est différente en fonction des méthodes et des patients. En effet, pour P1 et P3, on observe une grande augmentation au fur et à mesure de l'ajout de caractéristiques avec des performances plateaux atteintes à partir de 100 caractéristiques environ, pour la méthode sur-échantillonnée et optimisée concernant P1 et 30 caractéristiques pour P3 seulement pour la méthode sur-échantillonnée. Alors que pour les patients P2, P4 et P5, on constate très peu de variation avec des performances plateaux atteintes pour 5-10 caractéristiques (Tableau 4). Finalement, la méthode de sur-échantillonnage affiche les meilleurs scores AUC pour tous patients confondus.

VII. Discussions

L'objectif principal de cette étude est de développer un algorithme individualisé pour la détection du FOG chez les patients atteints de la MP afin d'améliorer leur prise en charge clinique. Un aspect clé de notre recherche est de déterminer si les caractéristiques discriminantes extraites des données des IMUs varient entre les patients. Nous explorons également l'impact des méthodes de rééquilibrage des données sur les performances des algorithmes et la pertinence d'individualiser ces algorithmes pour chaque patient. En outre, nous examinons le rôle du positionnement des capteurs et comment les différences interindividuelles dans les épisodes de FOG pourraient influencer la détection et la prédiction des algorithmes.

Les résultats de notre étude montrent que les méthodes de rééquilibrage, affichent de meilleurs résultats par rapport aux données sans rééquilibrage. En effet, nous avons obtenu d'excellentes performances (Tableau 4), avec des scores AUC élevés, ainsi qu'une sensibilité et une spécificité élevées pour la plupart des patients. Par exemple, pour une sélection des 100 meilleures caractéristiques, la méthode de sur-échantillonnage a atteint une sensibilité de 91 % et une spécificité de 94,6 %, tandis que la méthode optimisée a obtenu une sensibilité de 88,6 % et une spécificité de 94 %. Cependant, P3 a affiché des performances inférieures, avec une sensibilité de 73,2 % pour la méthode de sur-échantillonnage et de 67,4 % pour la méthode optimisée, qui peut s'expliquer par un déséquilibre important entre les classes pour la méthode optimisée et un nombre limité de données. Le modèle a eu très peu de données à entraîner et n'a pas encore pu saisir une distinction efficace entre la classe FOG et Non-FOG. Dans de tels cas, la méthode de sur-échantillonnage s'avère plus adaptée en augmentant le volume de données, assurant ainsi un nombre suffisant d'exemples pour un entraînement efficace des modèles.

Pour les autres patients, à performances égales, il est préférable d'utiliser la méthode optimisée, car elle minimise la création excessive de données synthétiques et maintient une distribution des données plus proche de la réalité. Cette approche permet de limiter l'introduction d'artefacts ou d'anomalies et assure que les modèles s'appuient sur des données fidèles aux conditions réelles des patients, ce qui est essentiel pour des analyses précises et fiables. Ainsi, les modèles entraînés sont plus robustes et moins susceptibles de sur-adapter aux données synthétiques, ce qui est particulièrement pertinent pour nos applications cliniques (Chawla et al., 2002).

L'analyse des caractéristiques extraites des données des IMUs révèle une dominance des caractéristiques provenant du gyroscope situées aux pieds et aux tibias pour la méthode brute, ce qui ne rejoint pas totalement les recherches précédentes (Moore et al., 2008; O'Day et al., 2022; Tripoliti et al., 2013), qui suggèrent une utilisation optimale au niveau du pied et du pelvis. En revanche, les méthodes de rééquilibrage affichent une plus grande variabilité, avec des représentations distinctes en fonction des patients, favorisant toutefois les caractéristiques fréquentielles. Par exemple, pour P3, les caractéristiques proviennent principalement du gyroscope au niveau du pied et du tibia, tandis que pour P5, elles proviennent majoritairement de l'accéléromètre et du gyroscope, principalement au niveau du pied.

Ces observations montrent qu'il existe des caractéristiques de FOG différentes entre les patients, soulignant la nécessité d'individualiser les modèles pour améliorer les performances de détection. Cela confirme les résultats de Mazilu et al. (2012) et Rodriguez et al. (2017), qui ont

également montré l'intérêt d'individualiser les données pour obtenir de meilleures performances. Nos résultats surpassent ceux des méthodes plus générales, comme celles appliquées par Pardoel (2021) avec le RUSBoosted Decision Tree (sensibilité : 81,4 %, spécificité : 91,6 % avec 25 caractéristiques). Notre modèle rivalise également avec les résultats de Tripoliti et al. (2013), qui ont obtenu 81,94 % de sensibilité et 98,74 % de spécificité avec un RF utilisant uniquement l'entropie des gyroscopes et des accéléromètres. Ces résultats mettent en évidence l'importance de personnaliser les modèles de détection du FOG pour chaque patient, afin d'améliorer la prise en charge clinique.

Dans notre étude, nous avons utilisé 2352 caractéristiques, principalement fréquentielles et temporelles. D'autres recherches utilisent des nombres différents de caractéristiques : par exemple, Pardoel (2021) utilise 861 caractéristiques, mélangeant des données de pression plantaire et des IMUs, tandis que Trabassi et al. (2022) utilisent 22 caractéristiques spatio-temporelles, et Tripoliti et al. (2013) se concentrent uniquement sur l'entropie sur chaque axe. Les conditions de passage varient également : nous avons réalisé 18 passages en phase ON et OFF, ce qui nous a offert une grande quantité de données, comparé à Pardoel (2021) qui n'a effectué des passages qu'en phase ON. L'emplacement des capteurs et le nombre de patients étudiés varient aussi, compliquant ainsi la comparaison directe des résultats.

Malgré les résultats prometteurs obtenus, certaines limites doivent être soulignées. Tout d'abord, le nombre de patients inclus est limité à cinq, ce qui restreint notre capacité à généraliser nos interprétations. Avec un échantillon aussi restreint, nous ne pouvons qu'identifier des tendances plutôt que des résultats définitifs et généralisables. De plus, les performances presque similaires entre les méthodes de rééquilibrage (optimisée et sur-échantillonnée) pourraient s'expliquer par des différences minimales dans la distribution finale des données rééquilibrées, suggérant que nous comparons des méthodes presque similaires. Enfin, l'absence d'un modèle de référence général pour comparaison directe limite la portée de nos conclusions sur l'efficacité relative des approches individualisées.

Pour aller plus loin, il serait bénéfique de réaliser une validation croisée des différents résultats obtenus entre ReliefF et nos données de validation. Par exemple, en appliquant Random Forest sur les données rééquilibrées avec la sélection des caractéristiques des données brutes et vice versa. Ensuite, d'autres modèles de ML, tels que les SVM, le RUSboosting Decision Tree, pourraient être comparés avec le Random Forest. Il pourrait également être intéressant d'envisager d'autres approches de sélection, telles que les méthodes embarquées, pour observer leur influence sur la sélection des caractéristiques.

Ainsi, bien que cette étude ait apporté des contributions intéressantes dans la détection du FOG chez les patients atteints de la MP, de nombreuses pistes de recherche restent à explorer pour perfectionner ces outils et optimiser la prise en charge clinique.

VIII. Conclusion

Cette étude avait pour objectif de développer un algorithme personnalisé pour la détection du FOG chez les personnes atteintes de la MP en utilisant les IMUs. Nos résultats montrent que les caractéristiques discriminantes extraites des IMUs varient entre les patients, soulignant la nécessité de personnaliser les modèles de détection pour chaque individu. L'application de techniques de rééquilibrage des données, telles que le suréchantillonnage et l'alternance entre sur-échantillonnage et sous-échantillonnage, a démontré une amélioration des performances des algorithmes, avec une utilisation adaptée au cas par cas.

Nous avons également constaté que les données sans rééquilibrage affichent une dominance des caractéristiques provenant des gyroscopes situés aux pieds et aux tibias. En revanche, les méthodes de rééquilibrage ont révélé une variabilité accrue dans la localisation des capteurs et les caractéristiques dominantes entre les patients, favorisant les caractéristiques fréquentielles et, dans une moindre mesure, les IMUs provenant des tibias.

Ces résultats mettent en évidence l'importance de personnaliser les modèles de détection du FOG pour chaque patient afin d'optimiser la prise en charge clinique. En développant des algorithmes performants et individualisés, nous pouvons espérer améliorer la qualité de vie des patients atteints de la MP en fournissant des diagnostics plus précis et des interventions cliniques mieux ciblées.

Résumé

Développement d'un algorithme personnalisé pour la détection du Freezing of Gait chez les personnes atteintes de la Maladie de Parkinson

Problématique

La maladie de Parkinson affecte gravement la mobilité des patients, notamment par le Freezing of Gait (FOG). Les méthodes actuelles de détection du FOG sont souvent appliquées de manière générale sans prendre en compte l'individualité des patients. Nous nous intéressons donc à la nécessité d'intégrer les variations individuelles des symptômes de FOG dans la création d'un algorithme de détection.

Méthodologie

Cette étude propose de développer un algorithme pour la détection du FOG en utilisant des Unités de Mesure Inertielle (IMUs). Les données sont collectées auprès de patients durant des sessions de marche, incluant des tâches simples et doubles. Des techniques de rééquilibrage des données, telles que le sur-échantillonnage avec SMOTE et une alternance optimisée entre SMOTE et sous-échantillonnage, sont appliquées pour améliorer la sélection des caractéristiques de ReliefF et les performance de Random Forest.

Discussion/Conclusion

Les résultats montrent que les méthodes de rééquilibrage améliorent la détection du FOG, avec une préférence pour une utilisation au cas par cas de ces méthodes. Les caractéristiques discriminantes varient significativement entre les patients, renforçant l'importance d'individualiser les modèles de détection pour optimiser la prise en charge clinique.

Perspectives

Les futures recherches devraient approfondir l'individualisation des algorithmes avec d'autres méthodes de sélection et explorer d'autres techniques de machine learning. De plus, il serait pertinent de réaliser une validation croisée entre les différents résultats de sélection et les différentes méthodes de rééquilibrage.

Abstract

Development of a Customized Algorithm for Detecting Freezing of Gait in People with Parkinson's Disease

Problem Statement

Parkinson's disease severely affects patients' mobility, particularly through Freezing of Gait (FOG). Current methods of detecting FOG are often applied generally without considering the individuality of patients. Therefore, we are interested in the need to integrate individual variations of FOG symptoms into the creation of a detection algorithm.

Methodology

This study proposes developing an algorithm for detecting FOG using Inertial Measurement Units (IMUs). Data are collected from patients during walking sessions, including both simple and dual tasks. Data rebalancing techniques, such as oversampling with SMOTE and an optimized alternation between SMOTE and undersampling, are applied to improve ReliefF feature selection and Random Forest performance.

Discussion/Conclusion

The results show that rebalancing methods improve FOG detection, with a preference for case-by-case application of these methods. Discriminative features vary significantly between patients, reinforcing the importance of individualizing detection models to optimize clinical management.

Perspectives

Future research should delve deeper into the individualization of algorithms with other selection methods and explore additional machine learning techniques. Moreover, it would be relevant to conduct cross-validation between different selection results and rebalancing methods.

Bibliographie

- Ashfaque Mostafa, T., Soltaninejad, S., McIsaac, T. L., & Cheng, I. (2021). A Comparative Study of Time Frequency Representation Techniques for Freeze of Gait Detection and Prediction. *Sensors (Basel, Switzerland)*, 21(19), 6446. <https://doi.org/10.3390/s21196446>
- Beitz, J. M. (2014). Parkinson's disease : A review. *Frontiers in Bioscience-Scholar*, 6(1), Article 1. <https://doi.org/10.2741/S415>
- Bella, S. D., Benoit, C.-E., Farrugia, N., Keller, P. E., Obrig, H., Mainka, S., & Kotz, S. A. (2017). Gait improvement via rhythmic stimulation in Parkinson's disease is linked to rhythmic skills. *Scientific Reports*, 7(1), 42005. <https://doi.org/10.1038/srep42005>
- Bikias, T., Iakovakis, D., Hadjidimitriou, S., Charisis, V., & Hadjileontiadis, L. J. (2021). DeepFoG : An IMU-Based Detection of Freezing of Gait Episodes in Parkinson's Disease Patients via Deep Learning. *Frontiers in Robotics and AI*, 8. <https://www.frontiersin.org/articles/10.3389/frobt.2021.537384>
- Borzi, L., Mazzetta, I., Zampogna, A., Suppa, A., Olmo, G., & Irrera, F. (2021). Prediction of Freezing of Gait in Parkinson's Disease Using Wearables and Machine Learning. *Sensors (Basel, Switzerland)*, 21(2), 614. <https://doi.org/10.3390/s21020614>
- Camps, J., Samà, A., Martín, M., Rodríguez-Martín, D., Pérez-López, C., Moreno Arostegui, J. M., Cabestany, J., Català, A., Alcaine, S., Mestre, B., Prats, A., Crespo-Maraver, M. C., Counihan, T. J., Browne, P., Quinlan, L. R., Laighin, G. Ó., Sweeney, D., Lewy, H., Vainstein, G., ... Rodríguez-Molinero, A. (2018). Deep learning for freezing of gait detection in Parkinson's disease patients in their homes using a waist-worn inertial measurement unit. *Knowledge-Based Systems*, 139, 119-131. <https://doi.org/10.1016/j.knosys.2017.10.017>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE : Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321-357. <https://doi.org/10.1613/jair.953>

- Cockx, H. M., Lemmen, E. M., van Wezel, R. J. A., & Cameron, I. G. M. (2023). The effect of doorway characteristics on freezing of gait in Parkinson's disease. *Frontiers in Neurology*, 14. <https://doi.org/10.3389/fneur.2023.1265409>
- Delval, A., Snijders, A. H., Weerdesteyn, V., Duysens, J. E., Defebvre, L., Giladi, N., & Bloem, B. R. (2010). Objective detection of subtle freezing of gait episodes in Parkinson's disease. *Movement Disorders*, 25(11), 1684-1693. <https://doi.org/10.1002/mds.23159>
- Ehgoetz Martens, K. A., Shine, J. M., Walton, C. C., Georgiades, M. J., Gilat, M., Hall, J. M., Muller, A. J., Szeto, J. Y. Y., & Lewis, S. J. G. (2018). Evidence for subtypes of freezing of gait in Parkinson's disease. *Movement Disorders*, 33(7), 1174-1178. <https://doi.org/10.1002/mds.27417>
- Gao, C., Liu, J., Tan, Y., & Chen, S. (2020). Freezing of gait in Parkinson's disease : Pathophysiology, risk factors and treatments. *Translational Neurodegeneration*, 9(1), 12. <https://doi.org/10.1186/s40035-020-00191-5>
- Krawczyk, B. (2016). Learning from imbalanced data : Open challenges and future directions. *Progress in Artificial Intelligence*, 5(4), 221-232. <https://doi.org/10.1007/s13748-016-0094-0>
- Liu, X.-Y., Wu, J., & Zhou, Z.-H. (2009). Exploratory Undersampling for Class-Imbalance Learning. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 39, 539-550. <https://doi.org/10.1109/TSMCB.2008.2007853>
- Mazilu, S., Hardegger, M., Zhu, Z., Roggen, D., Tröster, G., Plotnik, M., & Hausdorff, J. M. (2012). Online detection of freezing of gait with smartphones and machine learning techniques. *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*, 123-130. <https://doi.org/10.4108/icst.pervasivehealth.2012.248680>
- Moore, S. T., MacDougall, H. G., & Ondo, W. G. (2008). Ambulatory monitoring of freezing of gait in Parkinson's disease. *Journal of Neuroscience Methods*, 167(2), 340-348. <https://doi.org/10.1016/j.jneumeth.2007.08.023>

- Nutt, J. G., Bloem, B. R., Giladi, N., Hallett, M., Horak, F. B., & Nieuwboer, A. (2011). Freezing of gait : Moving forward on a mysterious clinical phenomenon. *The Lancet. Neurology*, 10(8), 734-744. [https://doi.org/10.1016/S1474-4422\(11\)70143-0](https://doi.org/10.1016/S1474-4422(11)70143-0)
- O'Day, J., Lee, M., Seagers, K., Hoffman, S., Jih-Schiff, A., Kidziński, Ł., Delp, S., & Bronte-Stewart, H. (2022). Assessing inertial measurement unit locations for freezing of gait detection and patient preference. *Journal of NeuroEngineering and Rehabilitation*, 19(1), 20. <https://doi.org/10.1186/s12984-022-00992-x>
- Palmerini, L., Rocchi, L., Mazilu, S., Gazit, E., Hausdorff, J. M., & Chiari, L. (2017). Identification of Characteristic Motor Patterns Preceding Freezing of Gait in Parkinson's Disease Using Wearable Sensors. *Frontiers in Neurology*, 8. <https://doi.org/10.3389/fneur.2017.00394>
- Pardoel, S. (2021). *Detection and Prediction of Freezing of Gait in Parkinson's Disease using Wearable Sensors and Machine Learning* [Doctoral Thesis, University of Waterloo]. <https://uwspace.uwaterloo.ca/handle/10012/17634>
- Reches, T., Dagan, M., Herman, T., Gazit, E., Gouskova, N. A., Giladi, N., Manor, B., & Hausdorff, J. M. (2020). Using Wearable Sensors and Machine Learning to Automatically Detect Freezing of Gait during a FOG-Provoking Test. *Sensors*, 20(16), Article 16. <https://doi.org/10.3390/s20164474>
- Rodríguez-Martín, D., Samà, A., Pérez-López, C., Català, A., Moreno Arostegui, J. M., Cabestany, J., Bayés, À., Alcaine, S., Mestre, B., Prats, A., Crespo, M. C., Counihan, T. J., Browne, P., Quinlan, L. R., ÓLaighin, G., Sweeney, D., Lewy, H., Azuri, J., Vainstein, G., ... Rodríguez-Molinero, A. (2017). Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer. *PLoS ONE*, 12(2), e0171764. <https://doi.org/10.1371/journal.pone.0171764>
- Samà, A., Rodríguez-Martín, D., Pérez-López, C., Català, A., Alcaine, S., Mestre, B., Prats, A., Crespo, M. C., & Bayés, À. (2018). Determining the optimal features in freezing of gait detection through a single waist accelerometer in home environments. *Pattern Recognition Letters*, 105, 135-143. <https://doi.org/10.1016/j.patrec.2017.05.009>

- San-Segundo, R., Navarro-Hellín, H., Torres-Sánchez, R., Hodgins, J., & De la Torre, F. (2019). Increasing Robustness in the Detection of Freezing of Gait in Parkinson's Disease. *Electronics*, 8(2), Article 2. <https://doi.org/10.3390/electronics8020119>
- Schaafsma, J. D., Balash, Y., Gurevich, T., Bartels, A. L., Hausdorff, J. M., & Giladi, N. (2003). Characterization of freezing of gait subtypes and the response of each to levodopa in Parkinson's disease. *European Journal of Neurology*, 10(4), 391-398. <https://doi.org/10.1046/j.1468-1331.2003.00611.x>
- Sigcha, L., Costa, N., Pavón, I., Costa, S., Arezes, P., López, J. M., & De Arcas, G. (2020). Deep Learning Approaches for Detecting Freezing of Gait in Parkinson's Disease Patients through On-Body Acceleration Sensors. *Sensors (Basel, Switzerland)*, 20(7), 1895. <https://doi.org/10.3390/s20071895>
- Snijders, A. H., Takakusaki, K., Debu, B., Lozano, A. M., Krishna, V., Fasano, A., Aziz, T. Z., Papa, S. M., Factor, S. A., & Hallett, M. (2016). Physiology of freezing of gait. *Annals of Neurology*, 80(5), 644-659. <https://doi.org/10.1002/ana.24778>
- Tolosa, E., Garrido, A., Scholz, S. W., & Poewe, W. (2021). Challenges in the diagnosis of Parkinson's disease. *The Lancet. Neurology*, 20(5), 385-397. [https://doi.org/10.1016/S1474-4422\(21\)00030-2](https://doi.org/10.1016/S1474-4422(21)00030-2)
- Tripoliti, E. E., Tzallas, A. T., Tsipouras, M. G., Rigas, G., Bougia, P., Leontiou, M., Konitsiotis, S., Chondrogiorgi, M., Tsouli, S., & Fotiadis, D. I. (2013). Automatic detection of freezing of gait events in patients with Parkinson's disease. *Computer Methods and Programs in Biomedicine*, 110(1), 12-26. <https://doi.org/10.1016/j.cmpb.2012.10.016>
- Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018). Relief-based feature selection : Introduction and review. *Journal of Biomedical Informatics*, 85, 189-203. <https://doi.org/10.1016/j.jbi.2018.07.014>

Annexes

Tableau 2: Nombre de fenêtres pour chaque label extrait de chaque patient

Patient	Label des fenêtres			
	nonFog	fog	transitionFog	transitionNonFog
P1	3535	853	220	97
P2	4197	4678	213	118
P3	3666	740	343	136
P4	5050	4530	896	468
P5	4006	6922	456	235
Total	20454	17723	2128	1054

Tableau 3: Nombre de fenêtres regroupées en classes et servant à la création de l'algorithme

Patient	Label des fenêtres			
	train		test	
	Classe 1 : fog + transitionFog	Classe 0 : NonFog	Classe 1 : fog + transitionFog	Classe 0 : NonFog
P1	751	2474	322	1061
P2	3423	2938	1468	1259
P3	758	2566	325	1100
P4	3798	3535	1628	1515
P5	5164	2804	2214	1202
Total	13894	14317	5957	6137

Tableau 4: Performances du Random Forest

Patient	100 Meilleurs caractéristiques								
	Brute			Sur-échantillonné			Optimisé		
	Précision (%)	Sensibilité (%)	Spécificité (%)	Précision (%)	Sensibilité (%)	Spécificité (%)	Précision (%)	Sensibilité (%)	Spécificité (%)
P1	89,6	64,3	97,3	95,1	88,8	97	94,4	82	98,1
P2	95,3	97,5	92,6	95,2	97,7	92,4	95,1	97,8	92
P3	86,6	53,8	96,3	89,3	73,2	94	86,2	67,4	91,8
P4	96,9	98,6	95,1	97,8	98,9	96,7	97,7	98,9	96,4
P5	94	96,2	90	95,1	96,2	93	95	96,8	91,7
Moyenne	92,5	82,1	94,3	94,5	91,0	94,6	93,7	88,6	94,0
Écart-type	4,3	21,4	3,0	3,1	10,7	2,1	4,4	13,7	3,0

Patient	80 Meilleurs caractéristiques								
	Brute			Sur-échantillonné			Optimisé		
	Précision (%)	Sensibilité (%)	Spécificité (%)	Précision (%)	Sensibilité (%)	Spécificité (%)	Précision (%)	Sensibilité (%)	Spécificité (%)
P1	89,9	65,2	97,5	94,5	87	96,8	94,9	83,2	98,4
P2	95,6	98	92,9	94,8	97,5	91,6	94,9	97,8	91,4
P3	85,8	52,3	95,7	89,8	72,9	94,8	84,1	64,9	89,7
P4	96,3	98,2	94,4	97,5	98,6	96,3	97,6	98,6	96,6
P5	93,9	96	90	94,9	95,9	93,1	94,8	96,2	92,2
Moyenne	92,3	81,9	94,1	94,3	90,4	94,5	93,3	88,1	93,7
Écart-type	4,4	21,7	2,8	2,8	10,8	2,2	5,3	14,4	3,7

Patient	50 Meilleurs caractéristiques								
	Brute			Sur-échantillonné			Optimisé		
	Précision (%)	Sensibilité (%)	Spécificité (%)	Précision (%)	Sensibilité (%)	Spécificité (%)	Précision (%)	Sensibilité (%)	Spécificité (%)
P1	89,8	65,2	97,3	93	82,3	96,2	93,4	78,3	98
P2	95,1	97,5	92,4	94,9	97,3	92,1	94,8	97,6	91,5
P3	85,1	50,5	95,4	89	73,8	93,5	83,4	61,8	89,7
P4	95,9	97,4	94,3	97,6	98,5	96,7	97,7	98,5	96,8
P5	93,1	95,8	88,1	94,4	95,6	92,3	94,7	96,2	91,8
Moyenne	91,8	81,3	93,5	93,8	89,5	94,2	92,8	86,5	93,6
Écart-type	4,4	22,0	3,5	3,1	10,9	2,2	5,5	16,1	3,6

Patient	30 Meilleurs caractéristiques								
	Brute			Sur-échantillonné			Optimisé		
	Précision (%)	Sensibilité (%)	Spécificité (%)	Précision (%)	Sensibilité (%)	Spécificité (%)	Précision (%)	Sensibilité (%)	Spécificité (%)
P1	89,1	62,4	97,2	92,6	81,1	96	90,4	67,7	97,3
P2	95,2	97,1	93	94,8	97,4	91,8	94,9	97,8	91,6
P3	84,5	50,2	94,6	89,5	76,3	93,4	81,1	57,5	88
P4	96,1	97,5	94,5	96,9	98,4	95,4	97,4	98,6	96
P5	91,7	95,7	84,3	94,3	95,1	92,8	94,7	96,5	91,3
Moyenne	91,3	80,6	92,7	93,6	89,7	93,9	91,7	83,6	92,8
Écart-type	4,7	22,6	4,9	2,8	10,2	1,8	6,4	19,5	3,8