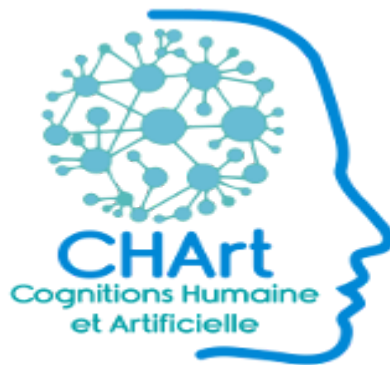


Rapport de stage de fin d'étude

Classification des phases de sommeil par ballistocardiographie



Trieu Xuong YU

Stage réalisé d'août à décembre 2020 sous la direction
de Guillaume CATHELAIN

Table des matières

Remerciements	3
Présentation du laboratoire CHArt et de l'entreprise Fealing	3
1. L'EPHE et le laboratoire CHArt	4
2. FEALING	4
I - Introduction	5
1. Les phases de sommeil	5
2. Les signaux de polysomnographie	5
3. Les signaux de ballistocardiographie	6
3.1. Définitions et mécanisme	6
II - Matériels et méthodes	9
1. Protocole d'enregistrement	9
1.1 Le capteur Emfit QS	9
1.2 Présentation des variables	9
2. Prétraitement des données	11
3. Différents modèles utilisés	11
3.1 LSTM - Bi-directional LSTM	12
3.2 U-net	13
4. Entraînement, validation, et test avec une fonction de coût personnalisée)	15
5. Évaluation de la performance	16
III - Résultats	18
1. Performance du réseau LSTM	18
2. U-net	18
3. Performance sur le jeu de données de test	19
IV - Discussion	22
1. Évaluation de la performance	22
2. Comparaison des performances avec d'autres méthodes	23
3. Limitations	23
V - Conclusion	24

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué au succès de mon stage et qui m'ont aidé lors de la rédaction de ce rapport.

Tout d'abord, je tiens à remercier vivement mon tuteur de stage, Monsieur Guillaume CATHELAIN, le fondateur de l'entreprise FEALING, pour son accueil, le temps passé ensemble et le partage de son expertise au quotidien. Grâce aussi à sa confiance j'ai pu m'accomplir totalement dans mes missions. Il fut d'une aide précieuse dans les moments les plus délicats.

Je remercie également à l'École Pratique des Hautes Études et l'hôpital Raymond Poincaré à Garches de m'avoir accueilli et m'avoir fourni toutes les conditions pour mener au mieux ce travail.

Présentation du laboratoire CHArt et de l'entreprise Fealing

1. L'EPHE et le laboratoire CHArt

L'École Pratique des Hautes Études(EPHE) - PSL, grand établissement d'enseignement supérieur, propose des formations à la recherche par la pratique de la recherche dans ses trois sections : Sciences de la vie et de la terre, Sciences historiques et philologiques et Sciences religieuses. Son enseignement est dispensé à partir du niveau Master. Créée en 1868 pour promouvoir des méthodes « pratiques » de formation à la recherche, l'EPHE - PSL apporte une contribution originale dans le domaine de la formation tout au long de la vie et de la formation continue en privilégiant un apprentissage de la recherche par la recherche.

Le laboratoire CHArt - Cognition Humaine et Artificielle - regroupe des équipes de recherche en sciences cognitives, en mathématiques et en informatique. En particulier, le co-directeur de ce laboratoire, François Jouen, étudie la douleur chez le bébé prématuré et ses conséquences épigénétiques sur le développement neuro-moteur du bébé. Le laboratoire CHArt développe des outils pour quantifier cette douleur et la limiter dans la mesure du possible lors des manipulations du nourrisson telles que les mesures physiologiques (fréquence cardiaque et respiratoire, saturation en dioxygène) qui sont parfois douloureuses. Des alternatives sans contact donc indolores à l'électrocardiogramme ou aux oxymètres de pouls sont en cours de développement, parmi lesquelles des méthodes optiques et mécaniques (photopléthysmographie, ballistocardiographie, séismocardiographie).

2. FEALING

FEALING est une startup issue de la recherche académique à l'EPHE et de son laboratoire CHArt. Elle compte comme actionnaire minoritaire la Fondation Paris Sciences Lettres (PSL) qui valorise la propriété intellectuelle issue des écoles et laboratoires de l'Université PSL. FEALING est également en collaboration de recherche avec l'APHP (Assistance Publique des Hôpitaux de recherche) et l'EPHE. FEALING a pour objet social le développement et la commercialisation de solutions algorithmiques et informatiques, d'objets connectés et de dispositifs médicaux pour la surveillance, le diagnostic et le soin de troubles physiologiques. En particulier, FEALING développe des solutions technologies de ballistocardiographie (enregistrement mécanique cardiaque sans contact) et de polygraphie pendant le sommeil pour la télésurveillance cardio-respiratoire, la classification des phases de sommeil, l'auscultation à distance, les matelas intelligents, le diagnostic de troubles du rythme cardiaque ou encore du syndrome d'apnée du sommeil.

I - Introduction

1. Les phases de sommeil

La qualité de sommeil est essentielle à la santé d'un être humain. Pour l'analyse, le sommeil est généralement classé en plusieurs phases dans lesquelles les signaux physiologiques suivent une dynamique différente, liés à diverses fonctions physiologiques.

Bien qu'il existe de nombreux critères pour classer des phases de sommeils, mais nous nous concentrons plutôt sur le critère des quatre phases de sommeils qui sont le réveil(Wake), le sommeil léger(Light sleep), le sommeil profond(Deep sleep), et le sommeil paradoxal ou REM(Rapid Eye Movement).

Par exemple, le sommeil léger est le début de notre cycle de sommeil et en même temps notre corps se calme. La respiration, la fréquence cardiaque et les changements musculaires préparent notre corps à un sommeil plus profond. Le sommeil profond est la partie de notre cycle de sommeil dans laquelle le corps récupère de la journée. Lorsque nous dormons suffisamment profondément, nous nous réveillons reposés. En cas de manque de sommeil profond, nous nous sentirons fatigués même si nous avons une nuit complète de repos. Le sommeil paradoxal est quant à lui la phase pendant laquelle le cerveau et les autres systèmes corporels deviennent plus actifs, et les muscles deviennent plus détendus. Le rêve se produit en raison d'une activité cérébrale accrue, mais les muscles volontaires s'immobilisent.

Dans cette étude, nous allons utiliser plusieurs modèles de machine learning pour effectuer une classification automatique des phases de sommeil en utilisant des paramètres extraits de signaux ballistocardiographiques.

2. Les signaux de polysomnographie

L'évaluation du sommeil consiste en la détection des différentes phases du sommeil grâce à la notation visuelle de l'enregistrement de polysomnographie (PSG). La PSG implique l'enregistrement nocturne de différents signaux électrophysiologiques tels que l'électroencéphalogramme (EEG), l'électromyogramme (EMG) et l'électro-oculogramme (EOG) [1].

Selon le manuel de "American academy of sleep medicine" (AASM) [1], les experts du sommeil notent visuellement les phases de sommeil toutes les 30 s en réveil, non-rapid eye movement (phases NREM 1 à 3) et rapid eye movement (REM). Cependant, bien que le PSG ait été utilisé pour évaluer le sommeil, il existe de nombreuses limitations :

1. De nombreux capteurs sont attachés au corps, ce qui est inconfortable ; ainsi, la surveillance à long terme est difficile. En outre, les inconvénients causés par les capteurs peuvent perturber le sommeil normal.
2. Un expert du sommeil qui peut mener le PSG basé sur le manuel AASM est nécessaire ; en outre, la réalisation du PSG et le processus de notation de la phase du sommeil sont des tâches exigeantes en main-d'œuvre et en temps.

Pour surmonter ces limitations, de nombreuses études ont tenté de classer automatiquement les phases de sommeil en utilisant un nombre minimum de signaux. Étant

donné que le manuel AASM pour la notation des phases de sommeil est basé sur les signaux EEG, EOG et EMG, des modèles de machine learning sont nécessaires pour la classification des phases de sommeil à l'aide des informations du système nerveux autonome (SNA).

Les paramètres de variabilité de la fréquence cardiaque (VFC), en anglais Heart Rate Variability(HRV), qui représentent l'activité SNA ont été considérés comme l'une des fonctionnalités les plus utiles pour la stadification du sommeil. Les paramètres de la VFC dérivés des signaux ECG ont été utilisés pour l'évaluation automatique du sommeil [18] [19]. La HRV est le degré de fluctuation de la durée des contractions du cœur, ou de l'intervalle entre deux contractions. Il s'agit d'une plus ou moins grande régularité, tandis que le rythme cardiaque s'intéresse à la fréquence des contractions.

3. Les signaux de ballistocardiographie

La technologie de surveillance ballistocardiographique a le potentiel d'augmenter la participation des patients aux programmes de télésanté et d'améliorer la sécurité et le bien-être des patients dans les hôpitaux. Cependant, cette technologie n'est pas encore mûre sur le plan clinique.[2]

3.1. Définitions et mécanisme

La ballistocardiographie est une méthode de surveillance sans contact de l'activité cardiaque. Elle repose sur la mesure des forces balistiques : pendant la systole ventriculaire, le sang est éjecté du ventricule gauche à travers la crosse aortique, générant une force balistique cardiaque pulsée.

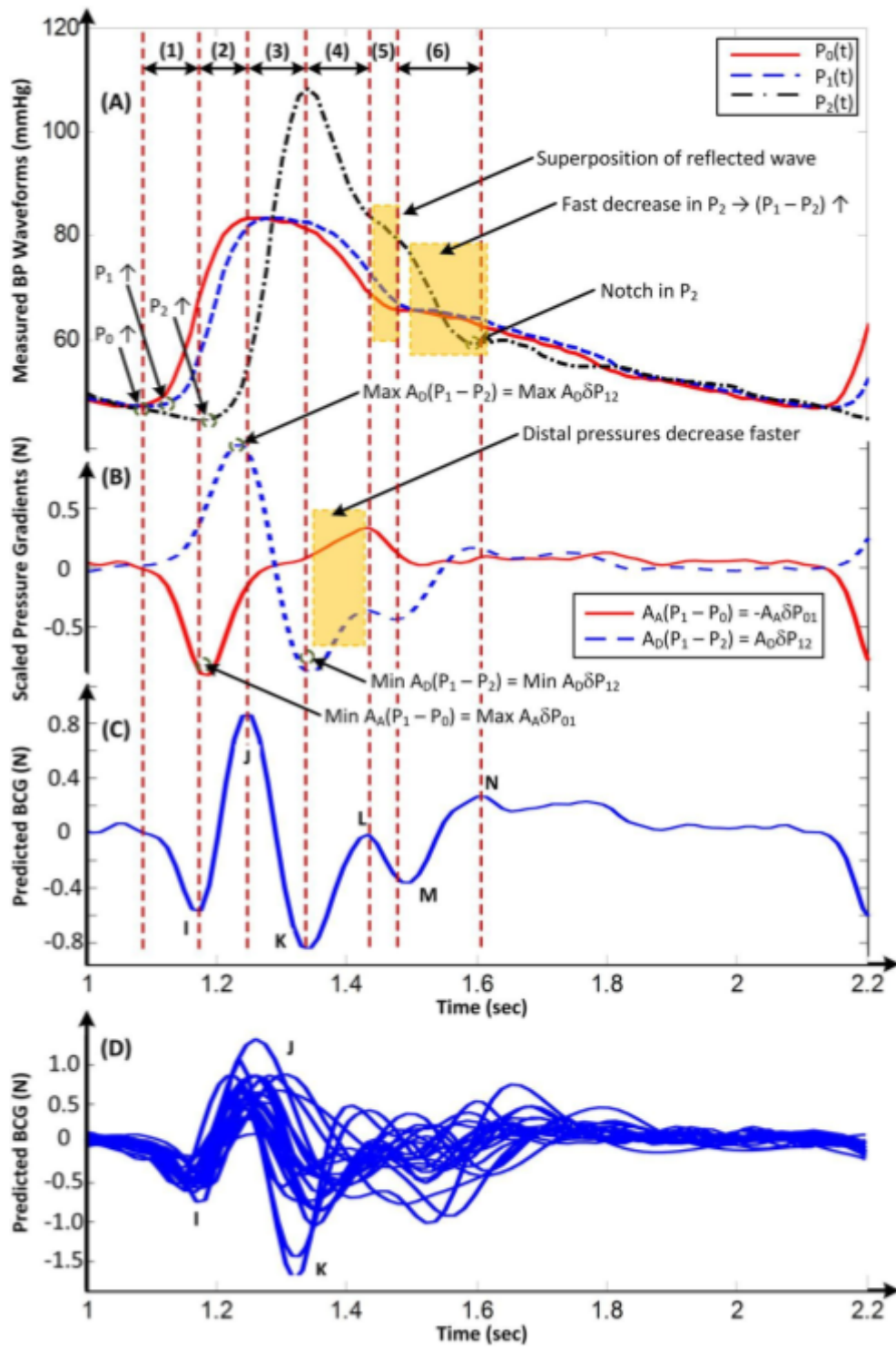
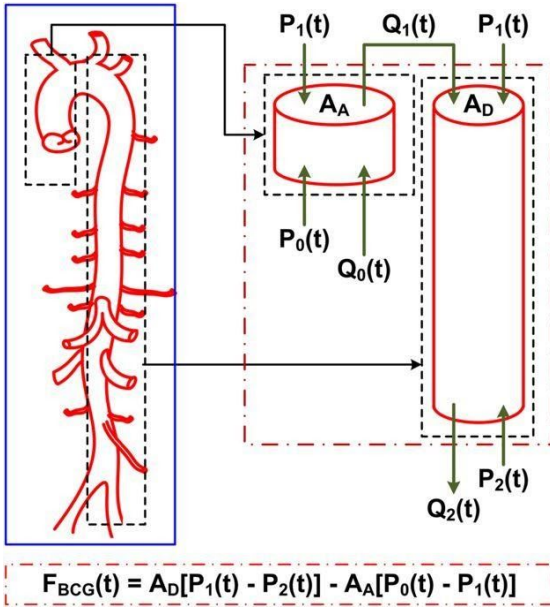


Figure SEQ Figure * ARABIC 3 : Ballistocardiogram waveform construction



La force du cardiogramme a été modélisée par des équations mathématiques [3], comme la perte hydraulique à travers l'arcade aortique. Dans ce modèle, la partie ascendante de l'aorte se situe entre la valve aortique et l'arc aortique d'une part, et après l'arcade aortique d'autre part, représentée par les indices A et D, respectivement.[2] La crosse aortique n'est pas représentée. Ces parties des artères aortiques sont des cylindres avec des sections de zones presque fixes A. P représente la pression artérielle tandis que Q représente le débit volumique, ces quantités étant dynamiques. L'entrée de l'aorte ascendante, l'apex de la crosse aortique et la sortie de l'aorte descendante sont ainsi respectivement définis par les $(P_0(t), Q_0(t))$, $(P_1(t), Q_1(t))$ et $(P_2(t), Q_2(t))$ avec A_A et A_D

pour les zones des parties ascendantes et descendantes. La deuxième loi de Newton implique que l'aorte ascendante et l'aorte descendante sont soumises aux forces suivantes, avec ρ la masse volumique sanguine :

$$F_A(t) = [P_0(t)A_A + \frac{\rho Q_0^2(t)}{A_A}] - [P_1(t)A_A + \frac{\rho Q_1^2(t)}{A_A}]$$

$$F_D(t) = [P_2(t)A_D + \frac{\rho Q_2^2(t)}{A_D}] - [P_1(t)A_D + \frac{\rho Q_1^2(t)}{A_D}]$$

Donc par la troisième loi de Newton, la crosse aortique est soumise à une force ballistocardiographique :

$$F_{BCG}(t) = -[F_A(t) + F_D(t)] = A_D \left[\left(P_1(t) + \rho \left[\frac{Q_1(t)}{A_D} \right]^2 \right) - \left(P_2(t) + \rho \left[\frac{Q_2(t)}{A_D} \right]^2 \right) \right]$$

$$= -A_A \left[\left(P_0(t) + \rho \left[\frac{Q_0(t)}{A_A} \right]^2 \right) - \left(P_1(t) + \rho \left[\frac{Q_1(t)}{A_A} \right]^2 \right) \right]$$

$$\approx A_D [P_1(t) - P_2(t)] - A_A [P_0(t) - P_1(t)]$$

Dans cette équation, la contribution de la vitesse est négligeable (de 1,6 à 1,9 mmHg) par rapport à la contribution de la tension artérielle (100 mmHg). La forme d'onde ballistocardiographique, appelée ballistocardiogramme, est illustrée à la Figure 3.

Le recul de cette force ballistocardiographique se répercute de la crosse aortique vers l'ensemble du corps à travers les tissus organiques, puis du corps vers le support du patient, par exemple un lit. Par conséquent, l'ensemble du lit est soumis à une contrainte mécanique courte et minime qui peut être mesurée comme un déplacement sur des surfaces à contrainte libre, par exemple sur la surface du matelas à l'aide d'un accéléromètre, ou en tant que contrainte sur des surfaces contraintes, par exemple sous le matelas à l'aide de jauges de contrainte.

II - Matériels et méthodes

1. Protocole d'enregistrement

1.1 Le capteur Emfit QS

Six patients dont 3 hommes et 3 femmes ont participé à cette étude. Chaque patient a participé pour une vingtième de nuits en moyenne, après une journée de travail ou le weekend. Les données et les signaux de ballistocardiographie sont obtenus par un capteur de lit Emfit QS.

Emfit est une société de technologie basée en Finlande qui conçoit et fabrique des biocapteurs depuis plusieurs décennies dans une gamme d'applications de santé différentes. La société a une solide histoire de recherche avec près de 100 articles de recherche scientifique publiés à son actif. Emfit QS (Quantified Sleep) est la première incursion de la société dans le monde du suivi du sommeil des consommateurs. Il s'agit d'une solution sophistiquée de surveillance du sommeil sans contact au lit, idéale pour l'analyse générale du sommeil. Mais le QS propose également un ensemble de fonctionnalités de mesures du sommeil très détaillées destinées aux athlètes, aux "Cross-Fitters" et aux sportifs professionnels. Emfit QS utilise des signaux de la ballistocardiographie, avec ses bio-capteurs logés dans une fine bande que nous plaçons sous nos matelas. Le suivi du sommeil est complètement autonome ; allongeons-nous simplement sur le lit et l'appareil commencera à enregistrer les mesures de nos corps.

1.2 Présentation des variables

L'ensemble de données se compose d'enregistrements de BCG et de données HRV(Heart Rate Variability) enregistrées par des participants pendant la nuit chez eux. Nous avons inclus des participants âgés de 25 à 56 ans et ne présentant aucun symptôme connu lié au sommeil. Les critères d'exclusion de cette étude étaient les suivants, les personnes ayant :

- des antécédents de maladies physiques ou psychologiques graves
- une instabilité de signes vitaux
- une arythmie cardiaque ou respiratoire
- troubles du sommeil (par exemple, le trouble des mouvements périodique de membre, le syndrome des jambes sans repos, le somnambulisme)

Tous nos enregistrements de BCG satisfaisant aux critères d'inclusion et d'exclusion ont été utilisés dans cette étude.

Le tableau I résume les variables liées au sommeil et les données démographiques(et aussi les labels) des participants :

Tableau I : RÉSUMÉ DES VARIABLES ET DÉMOGRAPHIQUES LIÉES AU SOMMEIL

Variables	Moyen ± écart type
Sexe(male/female)	3/3
Âge	36 ± 14.2
Durée des enregistrements(minutes)	602.5 ± 85.5
Label - Wake (en %)	8.5 ± 3.8
Label - REM (en %)	21.5 ± 4.7
Label - Light (en %)	51.4 ± 4.2
Label - Deep (en %)	18.6 ± 3.8

Deep : sommeil profond ; Light : sommeil léger ; REM : rapide eye movement i.e le sommeil paradoxal ; Wake : réveil

Le Emfit QS nous fournit des données vitales comme la fréquence respiratoire et la fréquence cardiaque, et des donnée VFC ou variabilité de la fréquence cardiaque(en anglais HRV, Heart Rate Variability) par exemple la moyenne quadratique des intervalles R-R successifs de HRV, les hautes fréquences normalisées, les basses fréquences normalisées. Le tableau II représente les variables utilisées pour cette étude.

Tableau II : Description des 8 variables associé avec le vital et le HRV

Association	Variables	Description	Unité
Vital	Hr	Fréquence cardiaque	battements par minute
	Rr	Fréquence respiratoire	respiration par minute
	Act	Activité	sans unité(0 - 32767)
HRV	Rmssd	Moyenne quadratique des intervalles R-R successifs de battement cardiaque	millisecondes
	Tp	Puissance spectrale totale de HRV entre 0.04 à 0.4 Hz	millisecondes ^ 2
	Lfn	Basses fréquences normalisées[1], l'activité fréquentielle dans l'intervalle de 0.04 à 0.15 Hz	%
	Hfn	Hautes fréquences normalisées[1], l'activité fréquentielle dans l'intervalle de 0.15 à 0.40 Hz	%
	R_hrv	Emfit Robust Heart Rate Variability Index[2]	millisecondes

[1] Lfn + Hfn = 100%

[2] La variable R_{hrv} est en phase expérimentale et Emfit n'a pas validé ses performances. Le R_{hrv} est conçu pour être calculable et stable pour une faible couverture des battements cardiaques.

2. Prétraitement des données

Selon des données de Emfit QS que nous avons récupérées, il existe 2 jeux de données pour 2 associations (Voir le tableau II) sur une nuit. D'une part, le premier jeu de données est un jeu de données de signes vitaux, qui contient des données de la fréquence cardiaque, la fréquence respiratoire et l'activité. Ce jeu de données contient des observations comme des époques de 4 secondes, c'est-à-dire qu'il y a 15 époques i.e 15 observations en une minute. D'autre part, le deuxième jeu de données contient des données de HRV. Ce jeu de données contient des observations comme des époques de 180 secondes, c'est-à-dire qu'il y a 20 observations en une heure. Il existe un autre jeu de données qui nous indique que les horaires auxquels le participant a quitté le lit et quand il est revenu dans le lit, ce jeu de données produit une classe de label qui s'appelle "bedexits".

Les deux jeux de données n'ont pas du tout la même taille à cause de différentes époques utilisées. Dans un premier temps, nous avons essayé plusieurs jeux de données pour entraîner différents modèles, par exemple les bases des données qui sont fixées à une durée déterminée pour une nuit (nous avons essayé de les fixer à 12 heures, 14 heures, 18 heures et 24 heures). Ensuite nous avons interpolé le jeu de données avec une époque de 4 secondes et de 180 secondes. Donc finalement nous avons décidé de faire une interpolation des époques de 180 secondes sur le jeu de données de HRV, mais en gardant des signes vitaux en colonne, c'est-à-dire que des signes vitaux deviennent des nouvelles variables. En 180 secondes, on aura 45 époques de 4 secondes, donc on aura $45 * 3 = 135$ nouvelles variables sur notre jeu de données final. Donc sur une nuit, on aura un jeu de données qui contient 288 lignes d'observations (des époques de 180 secondes) et 141 colonnes dont 140 variables explicatives et une variable cible.

3. Différents modèles utilisés

Récemment, l'apprentissage profond i.e Deep learning [4], une branche de l'apprentissage automatique qui utilise des réseaux de neurones constitués de plusieurs couches pour apprendre et extraire des caractéristiques de représentation des données, a été utilisé pour la détection d'objets, la reconnaissance vocale, la reconnaissance visuelle et de nombreux autres domaines. L'apprentissage profond a apporté des percées dans de nombreux domaines, et il a également été appliqué aux études de classification automatique des phases de sommeil. Phan et al. [5] ont introduit un framework de multi-classification et de prédiction en utilisant le réseau neuronal convolutif i.e Convolutional Neural Network (CNN) pour la classification automatique des phases de sommeil à l'aide de signaux EEG, EOG et EMG. Stephansen et coll. [6] ont utilisé un grand jeu de données comprenant 3000 enregistrements de sommeil normaux et anormaux et ont proposé un modèle CNN + LSTM avec des signaux EEG, EOG et EMG pour la notation automatique des phases de sommeil. Ces études ont atteint une très bonne performance pour la classification des phases de sommeil ; cependant, ils nécessitent un nombre important de signaux acquis par des capteurs intrusifs. Pour minimiser le nombre de signaux utilisés, plusieurs études ont

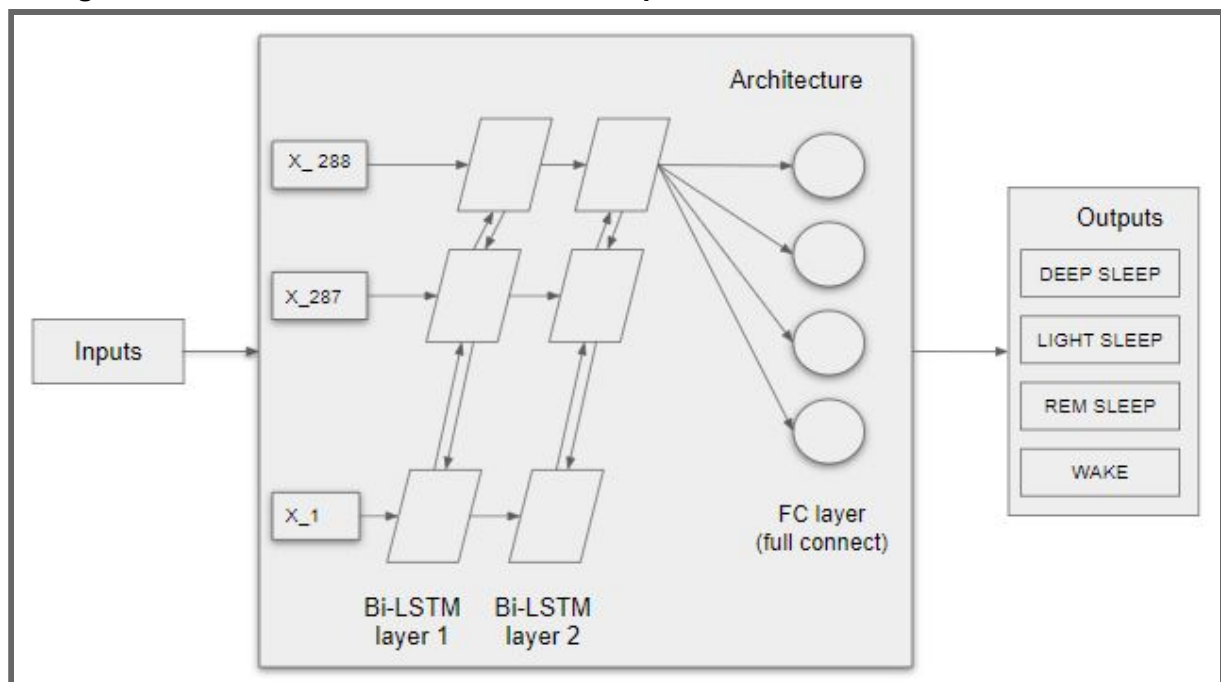
développé un modèle d'estimation des phases de sommeil basé sur le deep learning en utilisant un seul canal du signal EEG [7]. Bien que la phase de sommeil puisse être estimée en appliquant une méthode de deep learning à partir d'un signal EEG à un canal, le capteur doit être attaché au corps pour mesurer le signal EEG, ce qui est peu pratique pour la surveillance du sommeil à long terme.

Dans cette étude, nous avons évalué la capacité d'une architecture de réseau de neurones récurrents i.e Recurrent neural network - RNN, qui est la mémoire à court terme i.e Long Short-Term Memory LSTM ; et une certaine architecture de réseau neuronal convolutif CNN, ici un U-net à classifier les quatre phases de sommeil. Et nous pourrions comparer leurs performances avec un autre modèle de machine learning, l'Extreme Gradient Boosting - XGB.

3.1 LSTM - Bi-directional LSTM

Les réseaux de neurones récurrents (RNN) sont des réseaux à structure en boucle, qui prennent en compte les données en série temporelle. Ils ont démontré de meilleures performances que les autres méthodes de machine learning pour le traitement d'entrées séquentielles telles que la parole, le langage et les bio-sigaux de séries chronologiques [4]. En particulier, les RNN les plus performants sont des réseaux LSTM qui utilisent une unité cachée(hidden unit) spéciale, appelée cellule mémoire. Les réseaux LSTM sont conçus pour éviter le problème de gradient de fuite qui rend difficile l'apprentissage des RNN pour connecter les informations à mesure que la longueur de dépendance augmente. L'idée de base derrière un LSTM est que l'état de la cellule supprime ou ajoute des informations via une porte d'oubli, une porte d'entrée et une porte de sortie. Les équations détaillées des cellules LSTM sont décrites dans l'article original [8].

Figure 1 : Processus de classification des phases de sommeil basé sur le LSTM



LSTM, long short-term memory ; Bi-LSTM, bidirectional-LSTM ; FC layer, fully connected layer.

La figure 1 décrit le processus de classification des phases de sommeil et l'architecture LSTM sélectionnée comprenant deux couches LSTM bidirectionnelles et une couche entièrement connectée (FC). La couche LSTM comprend 256 unités par couche, 141 de dimensions d'entrée et 288 de longueurs de séquence qui impliquent 140 variables explicatives et 1 variable cible, et 288 époques de 180 secondes utilisées comme segment de données d'entrée. À partir de la couche LSTM, 256 unités de sortie sont connectées à une couche entièrement connectée (Fully connected layer - FC layer) qui classe les quatre classes (WAKE, REM SLEEP, LIGHT SLEEP, DEEP SLEEP). Pour éviter le surapprentissage (overfitting) du modèle, nous avons utilisé une couche de décrochage (dropout layer), qui est considéré comme une méthode de régularisation efficace [9].

Nous avons entraîné et validé plusieurs modèles de LSTM pour optimiser les hyperparamètres dans des intervalles prédéterminés. Une combinaison optimale d'hyperparamètres ayant produit les meilleures performances dans l'ensemble de données de validation a été recherchée, où les hyperparamètres étaient les suivants :

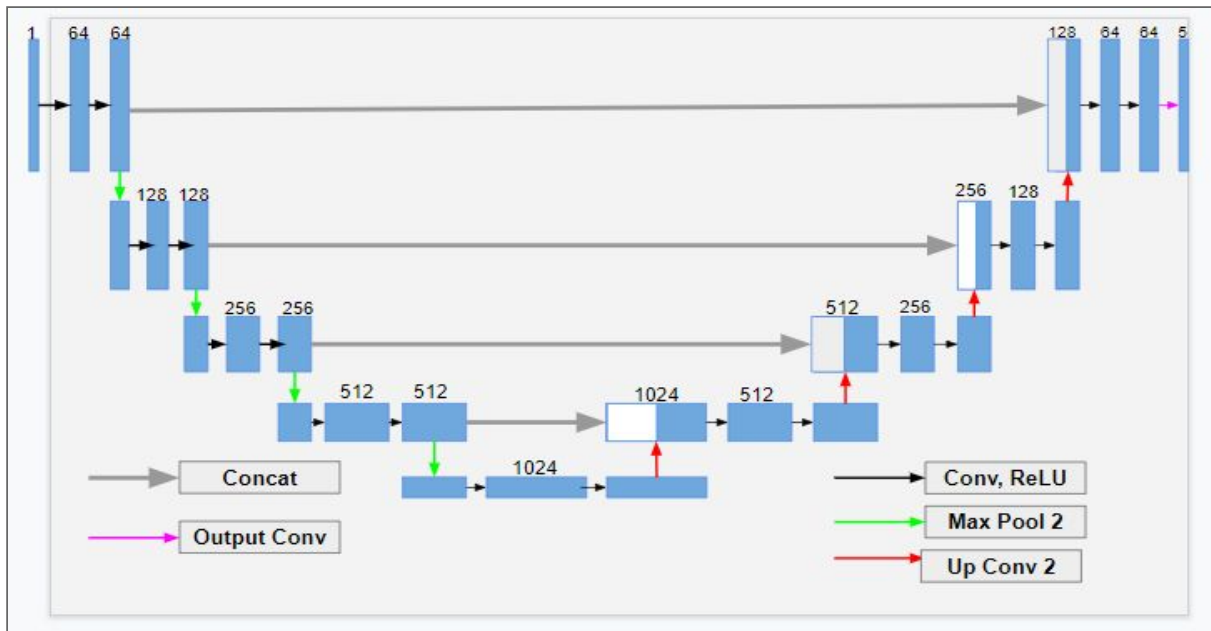
- nombre de couches de LSTM : {1, 2, 3}
- nombre d'unités caché de LSTM : {32, 64, 128, 256, 512}
- taux de décrochage : {0.1, 0.2, 0.3, 0.4, 0.5}
- taux de décrochage récurrente : {0.1, 0.2, 0.3, 0.4, 0.5}
- taux d'apprentissage : {0.1, 0.01, 0.001, 0.0001}
- type de structure de LSTM : {*unidirectionnelle*, *bidirectionnelle*}

Tanh et softmax ont été utilisés comme fonction d'activation pour LSTM et la couche dense, respectivement. La fonction d'activation récurrente du LSTM était une fonction sigmoïde.

3.2 U-net

Récemment, les réseaux convolutifs profonds (deep convolutional networks) ont surpassé l'état de l'art dans de nombreuses tâches de reconnaissance visuelle. Alors que les réseaux convolutifs existent déjà depuis longtemps [11], leur succès a été limité en raison de la taille des ensembles de formation disponibles et de la taille des réseaux considérés. La percée de Krizhevsky et al. [12] était dû à la formation supervisée d'un grand réseau avec 8 couches et des millions de paramètres sur l'ensemble de données ImageNet avec 1 million d'images d'entraînement. Depuis lors, des réseaux encore plus grands et plus profonds ont été formés [13].

Figure 2 : L'architecture de CNN



Dans cette étude, nous avons utilisé une architecture du réseau convolutif, qui ressemble à U-Net. L'architecture du réseau est illustrée à la figure 2. Il se compose d'un chemin de contraction (côté gauche) et d'un chemin expansif (côté droit). L'entrée du réseau neuronal est un vecteur de longueur l , où l est la longueur des époques de chaque nuit. Chaque fil est alimenté séparément à l'entrée du réseau neurone. La taille de sortie est $(5, l)$. Chaque colonne de la matrice de sortie contient 5 scores, qui caractérisent le degré de confiance du réseau neuronal que la valeur actuelle des phases de sommeil. Le réseau de neurones proposé comprend les couches suivantes :

- 1) 4 blocs, dont chacun comprend deux couches convolutives avec la fonction d'activation Relu ; ces blocs sont connectés séquentiellement avec des couches Max Pooling
- 2) La sortie de la couche précédente à travers la couche Max Pooling est envoyée à l'entrée d'un autre bloc contenant deux couches convolutives avec la fonction d'activation Relu
- 3) La sortie de la couche précédente à travers les couches de convolution transposé est concaténée avec la sortie de la couche 2) et est envoyée à l'entrée du bloc qui comprend deux couches convolutives chacune avec la fonction d'activation Relu
- 4) La sortie de la couche précédente à travers les couches de déconvolution et de remplissage nul est envoyée séquentiellement à l'entrée de 4 autres blocs contenant deux couches convolutives chacune avec la fonction d'activation Relu; chaque fois que la sortie est concaténée avec la sortie des couches correspondantes 1) dans l'ordre inverse
- 5) La sortie de la couche précédente est envoyée à l'entrée d'une autre couche convolutionnelle.

La différence entre le réseau proposé et U-Net est que nous utilisons des convolutions 1d au lieu de 2d convolutions.

Comme le réseau précédent LSTM, nous avons aussi entraîné et validé plusieurs modèles de U-net pour optimiser les hyperparamètres dans des intervalles prédéterminés. Une combinaison optimale d'hyper- paramètres ayant produit les meilleures performances dans l'ensemble de données de validation a été recherchée, où les hyperparamètres étaient les suivants :

- profondeur du réseau : {2, 3, 4}
- nombre de filtres de la couche de convolution au départ : {16, 32, 64, 128}
- longueur de la fenêtre de la couche de convolution : {9, 15, 24}
- taux de décrochage : {0.1, 0.2, 0.3, 0.4, 0.5}
- taux d'apprentissage : {0.1, 0.01, 0.001, 0.0001}

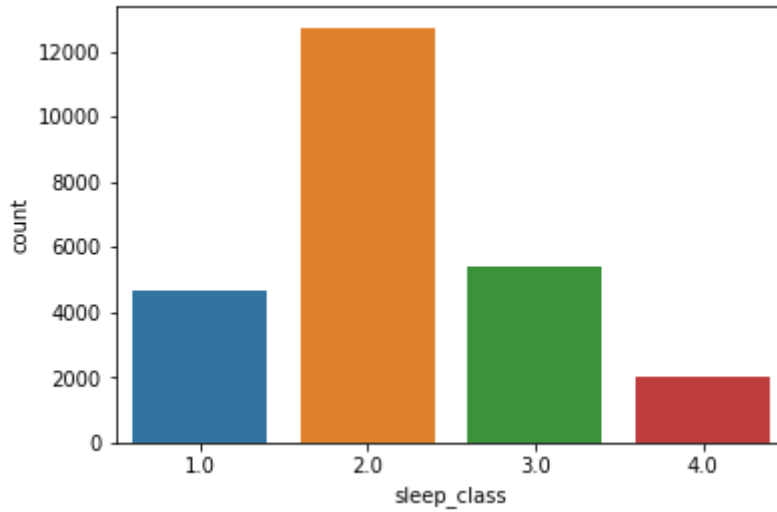
ReLU et softmax ont été utilisés comme fonction d'activation pour la couche de convolution et la couche dense, respectivement. La fonction d'activation récurrente du LSTM était une fonction sigmoïde.

4. Entraînement, validation, et test avec une fonction de coût personnalisée)

Nous avons entraîné le modèle en utilisant l'optimiseur Adam[10], et une fonction de perte(loss function) "categorical cross-entropy" personnalisé. La taille du lot(batch size), le taux d'apprentissage(learning rate) et les époques d'apprentissage(training epochs) ont été définis sur 32, 0.001, et 100, respectivement. Le nombre d'époques d'apprentissage définit le nombre de fois que l'algorithme d'apprentissage fonctionnera sur le jeu de données d'apprentissage.

Sur les 128 enregistrements de données de Emfit QS, 82 enregistrements ont été utilisés comme jeu de données de l'entraînement, 20 enregistrements ont été utilisés comme jeu de données de validation et les 26 enregistrements restants ont été utilisés pour un jeu de données de test. Lorsque nous avons entraîné le modèle LSTM et U-net, il y avait un problème de déséquilibre de classe(voir le tableau I) pour la répartition des phases de sommeil sur l'ensemble du jeu de données. Un équilibre entre le nombre d'observations par classe est nécessaire pour empêcher le modèle LSTM et U-net de se sur-apprendre à la classe majoritaire de sommeil léger(Light Sleep).

Figure 4 : Répartition de la variable cible



1 : Deep Sleep ; 2 : Light Sleep ; 3 : REM Sleep ; 4 : Wake

Il existe plusieurs méthodes pour équilibrer les classes, tels que le suréchantillonnage, le sous-échantillonnage et l'attribution de différents poids. Dans cette étude, nous avons adopté la méthode suggérée dans une étude précédente [14], qui attribue un poids différent à chaque classe et puis de les mettre dans une fonction de coût personnalisée. De cette manière, toutes les classes ont contribué de manière égale à la fonction de coût. Le poids de la classe X a été calculé à l'aide de l'équation (1) :

$$\text{Poids de Classe } X = \frac{\text{nombre total des époques}}{\text{nb de la classe} \times \text{nb de la classe } X} \quad (1)$$

En utilisant cette méthode, toutes les classes ont contribué de manière égale à la fonction de perte lors de la formation du modèle LSTM et U-net et ont eu pour effet d'éviter le sur-apprentissage pour la majorité des classes. Remarquons qu'il existe une 5^e classe de phase de sommeil qui est celle des sorties de lit, les "bedexits". Le poids de cette classe a été mis à 0 car cette classe ne nous intéresse pas. En effet, nous savons aisément prédire cette classe-là avec une simple analyse statistique de l'activité donc pas besoin de la prédire. De plus, une méthode de "Early Stopping" a été utilisée pour arrêter le processus de l'entraînement lorsque le coût de validation ne s'améliore pas. Au cours des 100 époques d'apprentissage, si le coût de validation n'a pas diminué pendant 10 itérations, l'entraînement est interrompu et les poids du modèle sont rétablis à la valeur qui présentait le score de AUC (Area Under Curve) le plus haut.

5. Évaluation de la performance

Pour évaluer les performances de classification des phases de sommeil, nous avons calculé le score de la précision (accuracy score) et l'AUC. L'AUC est calculée par le taux de vrai positif ou la sensibilité et le taux de faux positifs ou la spécificité. Et les formules sont :

1. $\text{sensibilité} = \frac{VP}{VP + FN}$ avec VP : vrai positif, FN : faux négatif
2. $\text{spécificité} = \frac{VN}{VN + FP}$ avec VN : vrai négatif, FP : faux positif

$$3. \text{ accuracy} = \frac{TP_D + TP_L + TP_{REM} + TP_W}{Total}$$

La performance de classification pour les quatre phases de sommeil a été calculée en utilisant une matrice de confusion (Fig. 3) et les équations (2) - (4). Lors du calcul des performances pour chaque phase de sommeil. Pendant le calcul des performances pour chaque phase de sommeil, la matrice de confusion et la classe de l'équation ont été réduites à 2 * 2 et à deux classes, respectivement.

Nous avons sélectionné le modèle LSTM et U-net qui présentait la valeur AUC la plus élevée dans le jeu de données de validation. Lorsque les performances d'une combinaison d'hyperparamètres sont les mêmes que celles de l'autre, un modèle de faible complexité a été sélectionné. Ensuite, nous avons évalué les performances du modèle sélectionné dans le jeu de données de test. Pour obtenir des résultats objectifs, nous avons comparé le modèle LSTM et le U-net avec une autre méthode de machine learning qui est XGB(Extrem Gradient Boosting). Nous avons effectué un processus d'optimisation hyperparamétrique pour le XGB. Dans le modèle de XGB, nous avons recherché le paramètre :

- Le profondeur("depth") maximale optimale, qui représente la profondeur maximale de l'arbre. Plus l'arborescence est profonde, plus il est divisé et plus il capture d'informations sur les données. Les candidats du paramètre de profondeur maximale étaient {2, 4, 6, 8, 10, 12, 14}
- Le nombre des estimateurs(n estimators), qui représente le nombre d'arbres. Les candidats du paramètre étaient {300, 500, 800, 1000, 1200}
- Le taux d'entraînement(learning rate), les candidats du paramètre étaient {0.001, 0.01, 0.1, 0.2, 0.3}

Après l'optimisation des hyperparamètres, les performances de chaque phase de sommeil et la classification totale des phases de sommeil selon les méthodes de classification ont été comparées dans un jeu de données de test.

Vrai classes	D	VP_D	$E_{D,L}$	$E_{D,R}$	$E_{D,W}$	$E_{D,B}$
	L	$E_{L,D}$	VP_L	$E_{L,R}$	$E_{L,W}$	$E_{L,B}$
	R	$E_{R,D}$	$E_{R,L}$	VP_R	$E_{R,W}$	$E_{R,B}$
	W	$E_{W,D}$	$E_{W,L}$	$E_{W,R}$	VP_W	$E_{W,B}$
	B	$E_{B,D}$	$E_{B,L}$	$E_{B,R}$	$E_{B,W}$	VP_B
		D	L	R	W	B
		classes prédites				

Figure 5 : Matrice de confusion pour la classification des phases de sommeils.

VP : Vrai positive ; E : erreur ; D : Deep Sleep ; L : Light Sleep ; R : REM Sleep ; W : Wake ; B : bedexits

III - Résultats

1. Performance du réseau LSTM

Pour sélectionner la combinaison optimale des hyperparamètres, nous avons comparé les performances du modèle LSTM dans le jeu de données de validation composé de 20 enregistrements. La tableau III présente les performances de classification des phases de sommeil dans le jeu de données de validation en modifiant des hyperparamètres du modèle LSTM.

Tableau III : Résultat de l'optimisation des hyperparamètres du modèle bidirectionnel-LSTM

Range	num_units	dropout_1	r_dropout_1	dropout_2	r_dropout_2	learning rate	AUC score
1	256	0.2	0.2	0.3	0.1	0.001	0.816
2	256	0.3	0.0	0.0	0.3	0.001	0.80
3	128	0.4	0.2	0.4	0.3	0.001	0.795
9	128	0.3	0.0	0.2	0.1	0.001	0.76
10	256	0.5	0	0.5	0	0.001	0.75

num_units : nombre d'unités caché de la couche ; dropout : le taux de décrochage ; learning rate : le taux d'apprentissage ; AUC score : Area Under the Receiver Operating Characteristic Curve score

Indépendamment de la structure LSTM (unidirectionnelle et bidirectionnelle) ou du nombre de couches, les performances de classification ont montré globalement une tendance croissante à mesure que le nombre d'unités caché de la couche augmente et les meilleurs modèles ont tous un taux d'apprentissage de 0.001 et de structure bidirectionnelle. Nous avons ensuite sélectionné chaque modèle LSTM qui présentait les performances les plus élevées en taux décrochage. Donc le meilleur modèle de la couche de LSTM sélectionnée est de structure bidirectionnelle, qui a deux couches LSTM de 256 unités cachées chacune, un taux d'apprentissage de 0.001, et {0.2, 0.2} de taux de décrochage et de taux de décrochage récurrente pour la première couche de LSTM et {0.3, 0.1} de taux de décrochage et de taux de décrochage récurrente pour la deuxième couche de LSTM.

2. U-net

Pour sélectionner la combinaison optimale des hyperparamètres, nous avons comparé les performances du modèle U-net dans le jeu de données de validation composé de 20 enregistrements. La tableau IV présente les performances de classification des phases de sommeil dans le jeu de données de validation en modifiant des hyperparamètres du modèle U-net.

Tableau VI: Résultat de l'optimisation des hyperparamètres du modèle U-net

Range	filter	kernel_size	dropouts_rate(1-8)	learning rate	AUC score
1	64	9	{0.2, 0.2, 0.3, 0, 0.2, 0.1, 0.4, 0.2}	0.001	0.62
2	16	15	{0.1, 0.5, 0.3, 0.5, 0, 0.5, 0, 0.3}	0.001	0.60
3	32	24	{0.4, 0.2, 0.1, 0.2, 0.2, 0.1, 0.3, 0.1}	0.001	0.59
9	64	9	{0.3, 0.3, 0, 0, 0.3, 0.3, 0.3, 0.2}	0.01	0.41
10	16	24	{0.4, 0.1, 0.1, 0.1, 0.5, 0.4, 0.2, 0}	0.01	0.40

filter : le nombre de filtres de sortie dans la convolution au départ ; kernel_size : la longueur de la fenêtre de la coche de convolution ; dropout : le taux de décrochage ; learning rate : le taux d'apprentissage ; AUC score : Area Under the Receiver Operating Characteristic Curve score

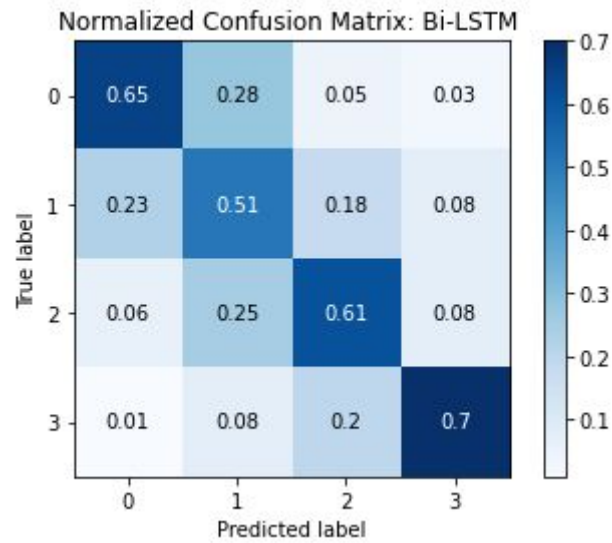
Les modèles de U-net qui présentaient les performances les plus élevées ont un taux d'apprentissage de 0.001, et de profondeur 4 c'est-à-dire qu'ils contiennent 4 blocs dans le chemin de contraction et le chemin expansif. Chaque bloc contient 2 couches de décrochage, et chaque couche de décrochage contient un taux de décrochage, nous les notons "dropout_rate_1, ..., dropout_rate_8", et ils sont {0.2, 0.2, 0.3, 0, 0.2, 0.1, 0.4, 0.2}. Toutes les couches convolutives ont les caractéristiques suivantes : kernel-size = 9, padding = 4. Toutes les couches de convolution transposée ont kernel-size = 8, stride = 2, padding = 3. Pour la dernière couche convolutionnelle kernel size = 1.

3. Performance sur le jeu de données de test

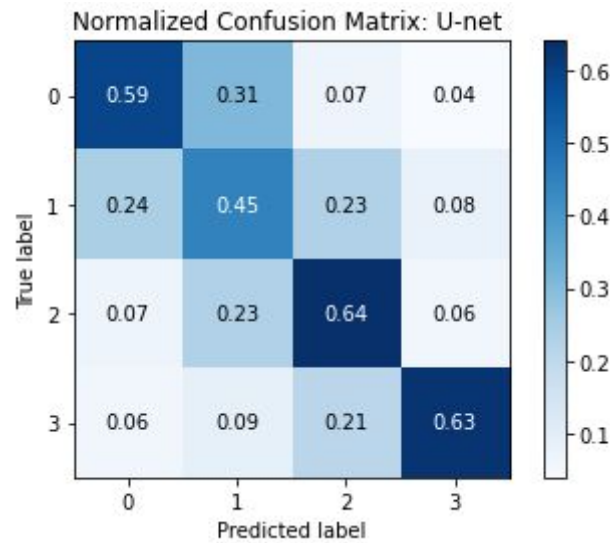
Nous avons évalué le modèle bi-LSTM de meilleure performance sélectionnée et le modèle U-net de meilleure performance sélectionnée sur le jeu de données de test, qui comprenait 26 enregistrements. Les performances de classification des quatre phases de sommeil ont été calculées et comparées entre le modèle bi-LSTM, U-net et XGB, la figure 6 nous a montré que les matrices de confusion sur le jeu de données de test en utilisant ces modèles. Sans compter la classe "bedexits" qui ne nous intéresse pas, le modèle bi-LSTM a montré les performances les plus élevées pour la classification des phases de sommeil, il obtient un score de la précision de 0.62, et le modèle de U-net a eu 0.59 et 0.49 pour le XGB.

En outre, à partir de la valeur AUC, le modèle bi-LSTM a considérablement amélioré les performances de classification de quatre phases de sommeil par rapport aux deux autres modèles utilisés. La figure 7 représente le courbe ROC de chaque modèle utilisé, nous pouvons voir que le LSTM a obtenu une meilleure score d'AUC qui est 0.87, alors que le U-net a eu seulement 0.68 et 0.61 pour le XGB. En plus, le LSTM a montré les performances les plus élevées sur les quatre phases de sommeil qui nous intéressent.

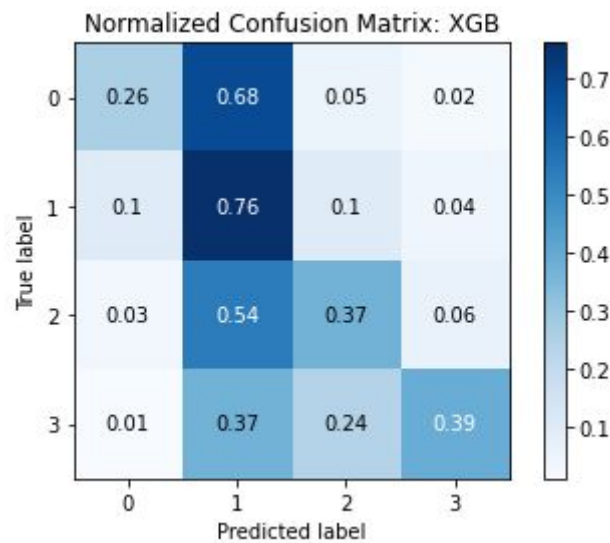
Figure 6: La matrice de confusion de modèles sélectionnés et de XGB
0:Deep sleep, 1: Light sleep, 2: REM sleep, 3: Wake, 4: Bedexits



Accuracy: 0.62

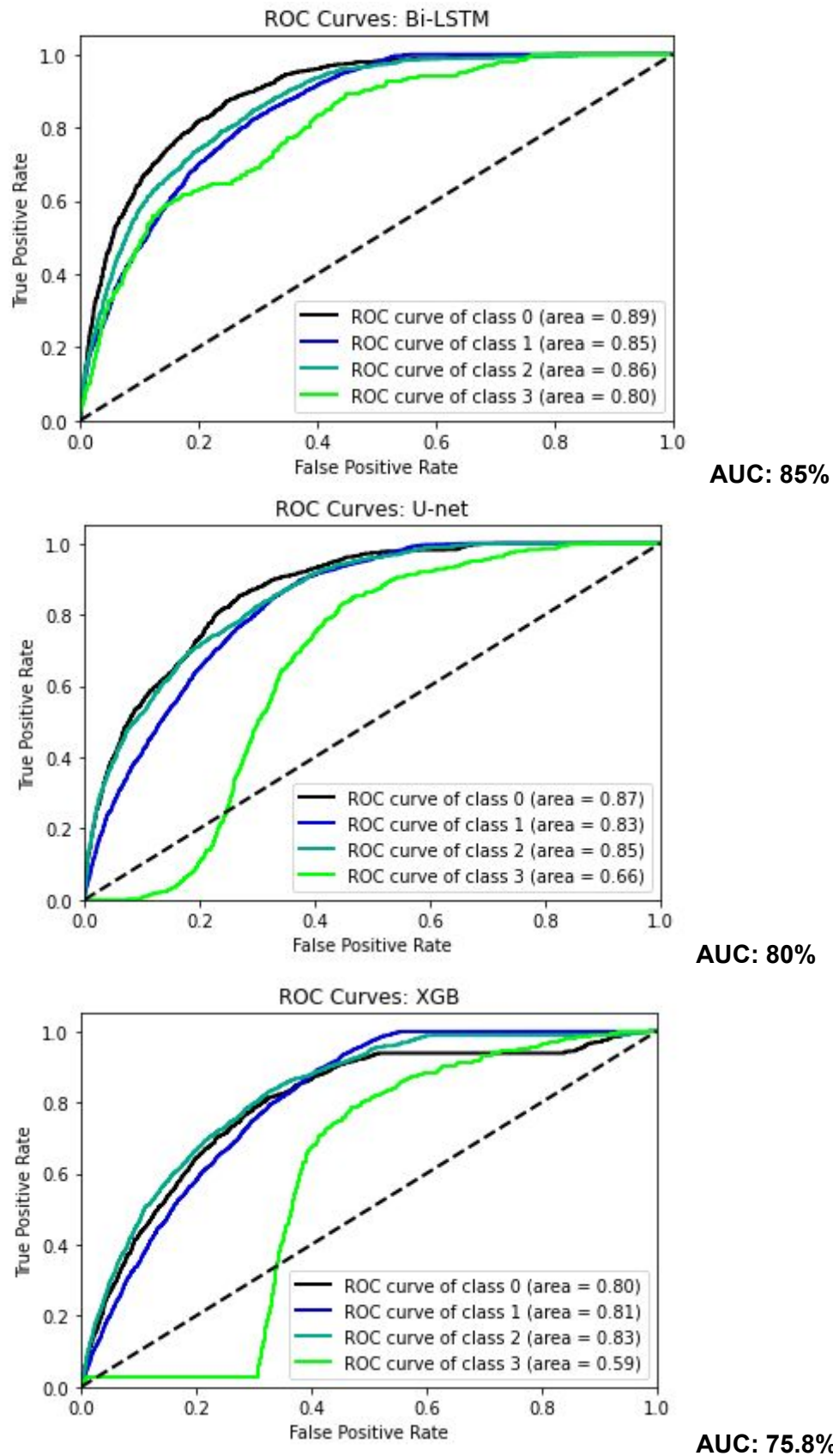


Accuracy: 0.58



Accuracy: 0.445

Figure 7 : La courbe ROC(receiver operating characteristic) de modèles sélectionnés et de XGB



IV - Discussion

Dans cet article, nous avons proposé un modèle LSTM et un modèle U-net optimal pour classer quatre phases de sommeil en utilisant des données d'un capteur de ballistocardiographie (huit paramètres liés au sommeil, y compris la fréquence cardiaque, la respiration et la HRV basée sur le BCG, ainsi que les phases de sommeil labellisés par l'algorithme EMFIT). Les performances de classification ont été évaluées sur la base de la combinaison d'hyperparamètres LSTM et U-net. De plus, nous avons comparé les performances de nos modèles proposés avec une des méthodes de machine learning conventionnelles, XGB.

1. Évaluation de la performance

En comparant les performances de la combinaison des hyperparamètres LSTM, les performances de classification avaient tendance à augmenter avec l'augmentation du nombre d'unités LSTM (voir tableau III). Le nombre d'unités affecte la capacité de la cellule LSTM à capturer les caractéristiques de la séquence d'entrée. Comme nous avons utilisé 8 variables explicatives, nous avons comparé les performances de classification en doublant le nombre d'unités de 32 à 512. Au fur et à mesure que le nombre d'unités augmentait, les performances s'amélioraient. L'augmentation du nombre d'unités de 16 à 256 a montré une amélioration des performances plus faible que lors de l'augmentation de 256 à 512. La principale raison de l'amélioration des performances différentielles peut être que les principales caractéristiques qui différencient les phases de sommeil sont extraites dans 256 unités, et les caractéristiques qui peuvent classer les phases de sommeil plus précisément sont extraites dans les unités supplémentaires.

Les caractéristiques de variation des paramètres de la HRV en fonction des phases de sommeil ont été bien étudiées et la tendance des variations de la HRV a été utilisée comme paramètre de classification des phases de sommeil dans les études précédentes [15] - [16]. Par conséquent, nous avons supposé qu'une séquence d'entrée plus longue contient des informations utiles liées à la phase de sommeil, et les réseaux LSTM et U-net ont été considérés comme des modèles efficaces pour extraire ces caractéristiques.

De plus, nous avons comparé les performances de classification entre les structures LSTM unidirectionnelles et bidirectionnelles. Bidirectional-LSTM [17] est une version étendue de unidirectional-LSTM, qui consiste en deux modèles LSTM traitant indépendamment les séquences d'entrée vers l'avant et vers l'arrière. Cette structure peut extraire des informations des états passés et futurs simultanément. Par conséquent, nous avons supposé que le modèle bidirectionnel-LSTM, qui extrait les informations d'état passé et futur, pourrait mieux classer les phases de sommeil que le modèle unidirectionnel-LSTM. En effet, la structure bidirectionnelle-LSTM a montré des performances de classification plus élevées que le LSTM unidirectionnel.

Lorsque nous avons entraîné les modèles de LSTM et de Unet, nous avons attribué des poids différents à chaque classe en utilisant une fonction de coût personnalisée. Chaque phase de sommeil n'était pas également distribuée dans le jeu de données d'entraînement, et si le modèle était entraîné dans l'état de déséquilibre de classe, le modèle apprendrait à ne classer que la majorité des phases de sommeil. Pour éviter que le modèle ne soit sur-appris, chaque classe doit être également à la fonction de coût dans l'ensemble de données d'apprentissage. Nous avons comparé les performances du modèle lors de l'utilisation de l'état d'équilibre de classe ou d'état de déséquilibre de classe.

2. Comparaison des performances avec d'autres méthodes

Nous avons comparé les performances des modèles LSTM et U-net proposés avec une des méthodes de machine learning conventionnelles tel que XGB. Nous avons utilisé les mêmes features et le même jeu de données pour entraîner le modèle. Le modèle LSTM, qui peut traiter des informations courtes et longues, s'est avéré être une meilleure méthode pour extraire des fonctionnalités utiles pour la mise en scène du sommeil par rapport aux autres méthodes. La figure 7 représente les performances de classification des phases de sommeil de notre modèle LSTM proposé, notre modèle U-net proposé et le XGB; la méthode LSTM a produit la valeur AUC la plus élevée par rapport aux deux autres méthodes. Les courbes ROC montrent que le modèle LSTM offre des performances améliorées pour la classification des phases de sommeil par rapport aux méthodes de CNN et de machine learning.

Pour chaque phase de sommeil, les LSTM ont les performances de classification les plus élevées pour les quatre phases de sommeils. (voir Fig.6-7) Par conséquent, le modèle LSTM pourrait extraire des caractéristiques de discrimination plus significatives pour chaque phase de sommeil que d'autres méthodes, et il est donc considéré comme la meilleure approche pour la classification des phases de sommeil.

3. Limitations

Bien que nos résultats soient corrects, notre étude présente quelques limites. Premièrement, nous avons formé et évalué le modèle en utilisant les signaux des participants sains. Pour obtenir un modèle généralisé, il est nécessaire d'acquérir des données auprès de personnes souffrant de troubles du sommeil tels que l'insomnie, l'apnée obstructive du sommeil et le PLMS et de tester les performances du modèle. Deuxièmement, bien que notre modèle ait amélioré les performances de classification des phases de sommeil, la complexité du modèle a également été augmentée par rapport aux autres méthodes d'apprentissage automatique. Dans cette étude, nous nous sommes concentrés sur l'amélioration des performances plutôt que sur la complexité du modèle. D'un point de vue informatique, le compromis entre les performances et la complexité du modèle doit être pris en compte, et un moyen de réduire la complexité du modèle d'apprentissage en profondeur doit être étudié. Et troisièmement, il existe certaines variables de Emfit qui sont propriétaires, nous n'avons des formules ou des explications de ces variables comme "R_hrv" et "activity", donc une régression de features doit être étudiée.

V - Conclusion

Dans cette étude, un modèle bidirectionnel-LSTM et un modèle de U-net ont été développés pour classer quatre phases de sommeil en utilisant des données de vital et HRV par un capteur de lit Emfit. Le modèle LSTM et U-net développés utilisent le mouvement, la respiration et le HRV pour évaluer les phases de sommeil, et ces modèles pourraient apprendre des modèles séquentiels de sommeil. En appliquant la méthode d'équilibrage des classes, le modèle LSTM a obtenu des performances de classification uniformément élevées à chaque étape du sommeil.

Notre étude contribue à une solution pour la surveillance du sommeil à long terme à domicile, ce qui pourrait avoir un impact positif sur les soins de santé. En effet, non seulement il fournit des informations utiles pour les soins de santé personnels, mais il pourrait également être utilisé pour suivre l'amélioration de la qualité du sommeil d'un patient après les traitements.

Références

- [1] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, C. L. Marcus, and B. V. Vaughn, "The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, Version 2.0," *Am. Acad. Sleep Med.*, 2012.
- [2] Guillaume Cathelain, "Ballistocardiography and applications", Thèse de doctorat, Université Paris Sciences Lettres, 2020.
- [3] C-S. Kim, S.L. Ober, M. S. McMurtry, B. A. Finegan, O. T. Inan, R. Mukkamala and J.-O. Hahn, "Ballistocardiogram: Mechanism and Potential for Unobtrusive Cardiovascular Health Monitoring," *Scientific Reports*, vol.6, 8 2016.
- [4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [5] H. Phan, F. Andreotti, N. Cooray, O. Y. Chen, and M. De Vos, "Joint Classification and Prediction CNN Framework for Automatic Sleep Stage Classification," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 5, pp. 1285–1296, 2018.
- [6] J. B. Stephansen et al., "Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy," *Nat. Commun.*, pp. 1–15, 2018.
- [7] H. Dong, A. Supratak, W. Pan, C. Wu, P. M. Matthews, and Y. Guo, "Mixed Neural Network Approach for Temporal Sleep Stage Classification," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 324–333, 2018.
- [8] S. Hochreiter, "Long Short-Term Memory," vol. 1780, pp. 1735–1780, 1997.
- [9] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.
- [11] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. *Neural Computation* 1(4), 541–551 (1989)
- [12] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: *NIPS*. pp. 1106–1114 (2012) 8. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R
- [13] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2014), [arXiv:1409.1556 \[cs.CV\]](https://arxiv.org/abs/1409.1556)
- [14] A. Malafeev et al., "Automatic Human Sleep Stage Scoring Using Deep Neural Networks," *Front. Neurosci.*, vol. 12, no. November, pp. 1–15, 2018.

- [15] H. Yoon et al., "Wakefulness evaluation during sleep for healthy subjects and OSA patients using a patch-type device," *Comput. Methods Programs Biomed.*, vol. 155, pp. 127–138, 2018.
- [16] S. M. Isa, I. Wasito, and A. M. Arymurthy, "Kernel Dimensionality Reduction on Sleep Stage Classification using ECG Signal," *Int. J. Comput. Sci. Issues*, vol. 8, no. 4, pp. 115–123, 2011.
- [17] M. Schuster and K. K. Paliwal, "Bidirectional Recurrent Neural Networks," *IEEE Trans. signal Process.*, vol. 45, no. 11, p. 6757, 1997.
- [18] B. Yilmaz, M. H. Asyali, E. Arikan, S. Yetkin, and F. Özgen, "Sleep stage and obstructive apneaic epoch classification using single-lead ECG," *Biomed. Eng. Online*, vol. 9, pp. 1–14, 2010.
- [19] H. Yoon, S. H. Hwang, J. W. Choi, Y. J. Lee, D. U. Jeong, and K. S. Park, "Slow-Wave Sleep Estimation for Healthy Subjects and OSA Patients Using R-R Intervals," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 119–128, 2018.
- [20] H. Lu, H. Zhang, Z. Lin, and N. S. Huat, "A Novel Deep Learning based Neural Network for Heartbeat Detection in Ballistocardiograph," 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018.
- [21] S. H. Choi, H. B. Kwon, H. W. Jin, H. N. Yoon, M. H. Lee, Y. J. Lee, and K. S. Park, "Long Short-Term Memory Networks for Unconstrained Sleep Stage Classification using Polyvinylidene Fluoride Film Sensor," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–1, 2020
- [22] R. Gargees, J. M. Keller, M. Popescu, and M. Skubic, "Non-invasive Classification of Sleep Stages with a Hydraulic Bed Sensor Using Deep Learning," *How AI Impacts Urban Living and Public Health Lecture Notes in Computer Science*, pp. 73–82, 2019.
- [23] W. Karlen, C. Mattiussi, and D. Floreano, "Improving actigraph sleep/wake classification with cardio-respiratory signals," 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008.
- [24] J. M. Kelly, R. E. Strecker, and M. T. Bianchi, "Recent Developments in Home Sleep-Monitoring Devices," *ISRN Neurology*, vol. 2012, pp. 1–10, 2012.
- [25] M. Migliorini, A. M. Bianchi, Nisticò D, J. Kortelainen, E. Arce-Santana, S. Cerutti, and M. O. Mendez, "Automatic sleep staging based on ballistocardiographic signals recorded through bed sensors," 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, 2010.
- [26] V. Moskalenko, N. Zolotykh, and G. Osipov, "Deep Learning for ECG Segmentation," *Studies in Computational Intelligence Advances in Neural Computation, Machine Learning, and Cognitive Research III*, pp. 246–254, 2019.

[27] J. Palotti, R. Mall, M. Aupetit, M. Rueschman, M. Singh, A. Sathyanarayana, S. Taheri, and L. Fernandez-Luque, "Benchmark on a large cohort for sleep-wake classification with machine learning techniques," *npj Digital Medicine*, vol. 2, no. 1, 2019.

[28] J. Palotti, R. Mall, M. Aupetit, M. Rueschman, M. Singh, A. Sathyanarayana, S. Taheri, and L. Fernandez-Luque, "Benchmark on a large cohort for sleep-wake classification with machine learning techniques," *npj Digital Medicine*, vol. 2, no. 1, 2019.

[29] Ronneberger, O.; P. Fischer & Brox, T. "U-Net: Convolutional Networks for Biomedical Image Segmentation *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*", Springer, 2015, 9351, 234-241

[30] X. Zhang, W. Kou, E. I.-C. Chang, H. Gao, Y. Fan, and Y. Xu, "Sleep stage classification based on multi-level feature learning and recurrent neural networks via wearable device," *Computers in Biology and Medicine*, vol. 103, pp. 71–81, 2018.