

République Algérienne Démocratique et Populaire  
الجمهورية الجزائرية الديمقراطية الشعبية  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
وزارة التعليم العالي و البحث العلمي  
École Nationale Polytechnique

---



المدرسة الوطنية المتعددة التقنيات  
Ecole Nationale Polytechnique

Département d'Électronique

### Mémoire de projet de fin d'études

Pour l'obtention du diplôme d'Ingénieur d'État en Électronique

---

# Estimation de la fréquence cardiaque à partir de signaux de photopléthysmographie affectés par des artefacts de mouvement

---

**OUMAZIZ Thiziri & MEZAIR Mayas**

Sous la direction de Dr. LALEG Taous-Meriem, HDR  
Et Dr. BOUSBIA-SALAH Hicham, Professeur

Présenté et soutenu publiquement le 26/06/2023 auprès des membres du jury :

<b>Président</b>	M. BELOUCHRANI Adel	Prof.	ENP, Alger
<b>Examinateuse</b>	Mme. LANI Fatiha	MAA	ENP, Alger
<b>Promotrice</b>	Mme. LALEG Taous-Meriem	HDR	INRIA-Saclay, Paris
<b>Co-promoteur</b>	M. BOUSBIA-SALAH Hicham	Prof.	ENP, Alger



République Algérienne Démocratique et Populaire  
الجمهورية الجزائرية الديمقراطية الشعبية  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
وزارة التعليم العالي و البحث العلمي  
École Nationale Polytechnique

---



المدرسة الوطنية المتعددة التقنيات  
Ecole Nationale Polytechnique

Département d'Électronique

Mémoire de projet de fin d'études  
Pour l'obtention du diplôme d'Ingénieur d'État en Électronique

---

# Estimation de la fréquence cardiaque à partir de signaux de photopléthysmographie affectés par des artefacts de mouvement

---

**OUMAZIZ Thiziri & MEZAIR Mayas**

Sous la direction de Dr. LALEG Taous-Meriem, HDR  
Et Dr. BOUSBIA-SALAH Hicham, Professeur

Présenté et soutenu publiquement le 26/06/2023 auprès des membres du jury :

<b>Président</b>	M. BELOUCHRANI Adel	Prof.	ENP, Alger
<b>Examinateuse</b>	Mme. LANI Fatiha	MAA	ENP, Alger
<b>Promotrice</b>	Mme. LALEG Taous-Meriem	HDR	INRIA-Saclay, Paris
<b>Co-promoteur</b>	M. BOUSBIA-SALAH Hicham	Prof.	ENP, Alger

## ملخص

متابعة معدل ضربات القلب باستخدام الأجهزة المرتدة على المعصم التي تدمج جهاز استشعار الفوتوبليزيموغرافيا (PPG) يحظى بشعبية متزايدة بسبب تكلفته المخفضة وسهولة استخدامه. لكن وجود تشوهات الحركة (MA) في إشارات PPG يعقد تقدير معدل ضربات القلب. في هذا العمل، قمنا بدراسة ثلاثة طرق مقترنة في المراجع التي تعامل مع هذه المشكلة: JOSS و WFPV و Q-PPG و قمنا بتجارب مختلفة عليها. أظهرت الطرق JOSS و WFPV أداءً جيداً في الأنشطة التي تنطوي على تشوهات حركة معتدلة، ولكنها أثبتت عدم فعالية في الأنشطة ذات تشوهات حركة أكثر تعقيداً. إضافة إلى ذلك، أظهرت طريقة JOSS قدرة تعميم منخفضة بسبب استعمالها لعدد كبير من المعلمات القابلة للضبط. أما طريقة Q-PPG فأثبتت قدرتها في تحسين نماذج التعلم العميق لتقدير معدل ضربات القلب تلقائياً، يجعلها أخف بهدف استخدامها في الأجهزة المرتدة ذات قدرة حوسية محدودة. ومع ذلك، فإن هذا الحل لا يزال محدوداً بتوفيرمجموعات بيانات كبيرة ومتنوعة بما فيه الكفاية.

**كلمات مفتاحية :** متابعة معدل ضربات القلب، PPG، تشوهات الحركة، JOSS، WFPV.

## Abstract

Heart rate (HR) monitoring using wrist-worn devices, integrating photoplethysmography (PPG) sensors, keeps gaining popularity due to its low cost and convenience. However, the presence of motion artifacts (MA) in PPG signals complicates HR estimation. In this work, we explored and conducted various experiments on three methods proposed in the literature that address this problem : JOSS, WFPV, and Q-PPG. The methods JOSS and WFPV, which performed well on activities involving moderate MA, proved to be ineffective on activities with more complex MA. Furthermore, JOSS demonstrated low generalization capability due to its numerous heuristic parameters and thresholds. The Q-PPG method proved flexibility in automatically optimizing deep learning models used for HR estimation, with the aim of deploying them on resource-limited hardware targets. However, this solution is still limited by the availability of sufficiently large and diverse datasets.

**Keywords :** Heart rate estimation, PPG, Motion artifacts, JOSS, WFPV, Q-PPG.

## Résumé

Le suivi de la fréquence cardiaque (FC) à l'aide de dispositifs portés au poignet, comprenant des capteurs de photopléthysmographie (PPG), ne cesse de gagner en popularité en raison de son faible coût et de sa grande commodité. Cependant, la présence d'artefacts de mouvement (MA) dans les signaux PPG compliquent l'estimation de la FC. Dans ce travail, nous avons exploré et mené diverses expérimentations sur trois méthodes de la littérature qui abordent cette problématique : JOSS, WFPV et Q-PPG. Les méthodes JOSS et WFPV qui avaient obtenu de très bons résultats d'estimation de la FC sur des activités impliquant des MA modérés se sont avérées inefficaces sur des activités ayant des MA plus complexes. De plus, JOSS a démontré une faible capacité de généralisation en raison de ses nombreux paramètres et seuils heuristiques. La méthode Q-PPG a montré une grande flexibilité dans l'optimisation automatique des modèles d'apprentissage profond utilisés pour l'estimation de la FC, en vue de leur déploiement sur des cibles matérielles à ressources limitées. Toutefois, cette solution reste limitée par la disponibilité d'ensembles de données assez larges et diversifiés.

**Mots clés :** Estimation de la fréquence cardiaque, PPG, Artefacts de mouvement, JOSS, WFPV, Q-PPG.

# Dédicaces

*À mes parents, à mes sœurs pour leur soutien inconditionnel, et à mes premiers éducateurs qui m'ont fait croire en moi.*

*- Mayas*

# Dédicaces

*Je dédie ce modeste travail en premier lieu à mes parents, qui ont tant fait pour moi. Les mots, à eux seuls, ne sauraient décrire tout l'amour et la gratitude que je ressens. Mon plus grand objectif est de vous rendre fiers.*

*À mon merveilleux frère, Arezki, merci pour les moments de rire et les précieuses leçons de vie.*

*À ma grand-mère maternelle, Yemma, et à la mémoire de ma grand-mère paternelle, Mamah, qui m'ont toujours inspirée par leur force et leur résilience.*

*À ma tante maternelle, Nora, à son mari Noredine, et à leurs enfants qui ont été d'un soutien immense tout au long de mon parcours universitaire.*

*À ma nounou, Fatiha, à son mari Cherif, et à leurs filles, Sarah et Rymah, qui sont comme ma seconde famille et qui ont toujours cru en moi.*

*À mes tantes et à mes oncles qui n'ont jamais cessé de m'encourager.*

*Aux merveilleuses rencontres que j'ai faites à l'École Nationale Polytechnique, qui m'ont tant fait évoluer humainement, et à tous ceux qui sont par la suite devenus mes amis.*

*À mes copines de chambre, Rym, Melissa et Kamelia, qui ont rendu les années en cité universitaire plus douces et moins pénibles.*

*À la meilleure promo, ELN Match, merci les amis d'avoir fait de ces années de spécialité de si belles années.*

- *Thiziri*

# Remerciements

Tout d'abord, nous tenons à exprimer nos sincères remerciements à notre promotrice, Mme. Laleg Taous Meriem, chercheuse à l'INRIA-Saclay et professeure à l'Université des sciences et technologies du roi Abdallah, pour sa présence, ses conseils avisés, ses encouragements constants et la bienveillance dont elle a fait preuve à notre égard. Nous la remercions pour l'opportunité qu'elle nous a offerte de travailler sous sa supervision, ainsi qu'avec son groupe de travail, dont les membres ont toujours été de très bons conseils.

Nous remercions également notre co-promoteur, M. Bousbia-Salah Hicham, professeur à l'École Nationale Polytechnique, pour ses remarques, ses conseils ainsi que sa gentillesse durant nos trois années de spécialité.

Nous tenons à remercier le président du jury, M. Belouchrani Adel, professeur à l'École Nationale Polytechnique, ainsi que l'examinatrice, Mme Lani Fatiha, enseignante à l'École Nationale Polytechnique, d'avoir accepté d'évaluer notre travail.

Nous remercions chaleureusement nos camarades de promotion pour leur soutien moral et leurs conseils concernant notre travail.

Enfin, nous souhaitons exprimer notre profonde gratitude envers les enseignants qui ont contribué à notre réussite. Leur soutien constant et leurs encouragements ont joué un rôle crucial dans notre développement académique et personnel.

# Table des matières

Liste des tableaux

Table des figures

Liste des abréviations

<b>Introduction générale</b>	<b>13</b>
<b>1 Généralités et contexte</b>	<b>15</b>
1.1 Généralités . . . . .	16
1.1.1 Suivi de la fréquence cardiaque . . . . .	16
1.1.2 Principe de la photopléthysmographie (PPG) . . . . .	17
1.1.3 Avantages de l'utilisation du PPG . . . . .	18
1.1.4 Artefacts de mouvement dans les signaux PPG . . . . .	19
1.2 État de l'art . . . . .	20
1.2.1 Approche Classique . . . . .	20
1.2.2 Approche par apprentissage profond . . . . .	21
1.3 Ensembles de données disponibles . . . . .	23
1.3.1 IEEE SPC 2015 – train . . . . .	23
1.3.2 IEEE SPC 2015 – test . . . . .	23
1.3.3 PPG-DaLiA . . . . .	24
1.3.4 Comparaison des ensembles de données . . . . .	27
1.4 Métriques d'évaluation . . . . .	28
1.5 Comparaison des méthodes . . . . .	29
<b>2 Approche Classique</b>	<b>32</b>
2.1 Méthode JOint Sparse Spectrum reconstruction (JOSS) . . . . .	33
2.1.1 Méthodologie proposée . . . . .	33
2.1.1.1 Reconstruction conjointe parcimonieuse des spectres en utilisant le modèle MMV . . . . .	34
2.1.1.2 Soustraction Spectrale . . . . .	36
2.1.1.3 Suivi du pic spectral (SPT) . . . . .	37
2.1.2 Paramètres de la méthode . . . . .	42
2.2 Méthode Wiener Filtering Phase Vocoder (WFPV) . . . . .	43
2.2.1 Méthodologie proposée . . . . .	43
2.2.1.1 Prétraitement . . . . .	43
2.2.1.2 Débruitage du signal PPG en utilisant le filtre de Wiener	43

2.2.1.3	Affinement de la fréquence cardiaque estimée en utilisant le vocodeur de phase . . . . .	46
2.2.1.4	Post-traitement . . . . .	48
2.2.2	Paramètres de la méthode . . . . .	48
2.3	Implémentation et évaluation de la méthode JOSS . . . . .	49
2.3.1	Implémentation sur MATLAB . . . . .	49
2.3.2	Évaluation de la robustesse de la méthode . . . . .	50
2.3.2.1	Impact du réglage des paramètres sur les performances .	50
2.3.2.2	Capacité de généralisation à de nouvelles activités . . . . .	52
2.3.3	Évaluation du temps d'exécution . . . . .	53
2.3.4	Conclusions de l'évaluation . . . . .	53
2.4	Implémentation et évaluation de la méthode WFPV . . . . .	54
2.4.1	Implémentation sur MATLAB . . . . .	55
2.4.2	Évaluation de la robustesse de la méthode . . . . .	55
2.4.2.1	Impact du réglage des paramètres sur les performances .	55
2.4.2.2	Capacité de généralisation à de nouvelles activités . . . . .	57
2.4.3	Évaluation du temps d'exécution . . . . .	59
2.4.4	Conclusion de l'évaluation . . . . .	59
<b>3</b>	<b>Approche par apprentissage profond</b>	<b>61</b>
3.1	Notions préliminaires . . . . .	62
3.1.1	Les Réseaux de Neurones Convolutifs (CNN) . . . . .	62
3.1.2	Normalisation par lots (Batch Normalization) . . . . .	65
3.1.3	Les Réseaux de Convolution Temporelle (TCN) . . . . .	66
3.1.4	Outils de recherche d'architectures neuronales (NAS) . . . . .	67
3.1.4.1	MorphNet . . . . .	68
3.1.4.2	Pruning In Time (PIT) : . . . . .	71
3.1.5	Architecture TEMPONet . . . . .	74
3.2	Méthode Q-PPG . . . . .	76
3.2.1	Réseau initial . . . . .	77
3.2.2	Optimisation de l'architecture . . . . .	77
3.2.3	Post-traitement . . . . .	79
3.3	Expérimentations et évaluation . . . . .	79
3.3.1	Expérimentations sur la méthodologie de recherche d'architectures proposée . . . . .	80
3.3.1.1	Effet du choix de la contrainte de ressources . . . . .	80
3.3.1.2	Effet du réglage des paramètres de régularisation . . . . .	81
3.3.1.3	Effet de l'étape de “warmup” . . . . .	82
3.3.2	Évaluation des performances . . . . .	83
3.3.2.1	Évaluation sur l'ensemble de données PPG-DaLiA . . . . .	83
3.3.2.2	Évaluation sur les ensembles de données IEEE SPC 2015 .	86
3.3.3	Expérimentations sur le post-traitement . . . . .	87
<b>Conclusion et perspectives</b>		<b>91</b>
<b>Bibliographie</b>		<b>94</b>

# Liste des tableaux

1.1	Description de l'ensemble de données IEEE SPC 2015 – train . . . . .	24
1.2	Description de l'ensemble de données IEEE SPC 2015 – test. T1 = exercices de rééducation du bras. T2 = mouvements intensifs du bras (boxe). * : un rythme cardiaque et une tension artérielle anormaux ont été observés. . . . .	24
1.3	Aperçu des sujets participant à la collecte de données dans PPG-DaLiA. Condition physique : fréquence de pratique sportive du sujet, sur une échelle de 1 à 6, où 1 correspond à moins d'une fois par mois et 6 correspond à 5 à 7 fois par semaine. . . . .	25
1.4	Durée approximative et fréquence cardiaque moyenne de chaque activité dans PPG-DaLiA. . . . .	26
1.5	Tableau de comparaison des méthodes de l'état de l'art. Les différents résultats de MAE (BPM) correspondent à différents ensembles de données	29
2.1	Paramètres réglables par l'utilisateur dans la méthode JOSS. . . . .	42
2.2	Comparaison des résultats en MAE (BPM) de notre implémentation de JOSS avec ceux de l'article de référence [4] sur l'ensemble de données IEEE SPC 2015 – train. . . . .	49
2.3	Résultats en MAE (BPM) de la méthode JOSS avec et sans optimisation des paramètres sur l'ensemble de données IEEE SPC 2015 – test. . . . .	50
2.4	Résultats en MAE (BPM) de la méthode WFPV sur l'ensemble de données IEEE SPC 2015 – train. . . . .	54
2.5	Résultats en MAE (BPM) de la méthode WFPV sur l'ensemble de données IEEE SPC 2015 – test. . . . .	54
2.6	Résultats en MAE (BPM) de la méthode WFPV sur l'ensemble de données PPG-DaLiA. . . . .	56
3.1	Comparaison du nombre de canaux par couche de deux architectures élaborées avec MorphNet, qui obtiennent des performances similaires, en ciblant l'optimisation soit de la taille du modèle ou du nombre de FLOPs. . . . .	81
3.2	Effet de l'étape de “warmup” sur le nombre de paramètres du modèle. . . . .	82
3.3	Résultats en MAE (BPM) de la méthode Q-PPG sur PPG-DaLiA. . . . .	83
3.4	Analyse des résultats obtenus avec Q-PPG et WFPV pour chaque activité incluse dans l'ensemble de données PPG-DaLiA, en précisant la fréquence cardiaque moyenne et la durée de chaque activité. . . . .	85
3.5	Résultats en MAE (BPM) de la méthode Q-PPG sur l'ensemble de données IEEE SPC 2015 - train. . . . .	86
3.6	Résultats en MAE (BPM) de la méthode Q-PPG sur l'ensemble de données IEEE SPC 2015 - test. . . . .	86

## Liste des tableaux

---

- 3.7 Résultats obtenus en MAE (BPM) par le modèle le plus précis sur l'ensemble de données PPG-DaLiA sans et avec différents post-traitements. . . 88

# Table des figures

1.1	Illustration du placement du capteur PPG en mode de transmission (en haut) et en mode de réflexion (en bas) [11]. . . . .	17
1.2	Illustration de la variation de l'atténuation de la lumière émise par la LED [13]. . . . .	18
1.3	Signaux PPG et ACC avec leurs spectres en l'absence de MA. Le pic vert, qui correspond à la vraie valeur de la FC, est aligné avec les pics dominants dans les spectres des signaux PPG1 et PPG2. PPG1 et PPG2 sont les canaux PPG, N1, N2 et N3 sont les canaux ACC [16]. . . . .	19
1.4	Signaux PPG et ACC avec leurs spectres en présence de MA. Le pic vert, qui correspond à la vraie valeur de la FC, n'est aligné avec aucun des pics dominants dans les spectres des signaux PPG1 et PPG2. PPG1 et PPG2 sont les canaux PPG, N1, N2 et N3 sont les canaux ACC [16]. . . . .	20
2.1	Organigramme de la méthode JOSS. . . . .	33
2.2	Illustration de l'étape de soustraction spectrale sur une fenêtre donnée pour la méthode JOSS. Le cercle rouge indique la valeur réelle de la FC en Hz. . . . .	36
2.3	Organigramme de la phase de sélection du pic de la méthode JOSS. . . . .	39
2.4	Organigramme de la phase de vérification du pic de la méthode JOSS. . . . .	40
2.5	Organigramme de la phase de découverte du pic de la méthode JOSS. . . . .	41
2.6	Organigramme de la méthode WFPV. . . . .	43
2.7	Débruitage du signal PPG et estimation initiale de la FC pour la méthode WFPV. Le point rouge correspond à l'estimation initiale de la FC. . . . .	45
2.8	Étapes de la transformée de Fourier à court terme (STFT). . . . .	46
2.9	Illustration du changement de phase d'un signal sinusoïdal. . . . .	47
2.10	Estimation faite sur le meilleur sujet (Sujet 8) de l'ensemble de données IEEE SPC 2015 – test avec réglage des paramètres de la méthode JOSS. . . . .	51
2.11	Estimation faite sur le pire sujet (Sujet 1) de l'ensemble de données IEEE SPC 2015 – test avec réglage des paramètres de la méthode JOSS. . . . .	51
2.12	Résultats en MAE (BPM) de JOSS par sujet et par activité sur PPG-DaLiA. . . . .	52
2.13	Résultats en MAE (BPM) de la méthode WFPV par sujet et par activité pour l'ensemble de données PPG-DaLiA. . . . .	57
2.14	Estimation faite sur le meilleur sujet (S7) de PPG-DaLiA avec WFPV. . . . .	58
2.15	Estimation faite sur le pire sujet (S5) de PPG-DaLiA avec WFPV. . . . .	59
3.1	Illustration de l'arrangement des couches dans un CNN. . . . .	62
3.2	Illustration de l'étape de convolution. . . . .	63
3.3	Illustration du Average-pooling et du Max-Pooling. . . . .	64
3.4	Illustration du principe de causalité. . . . .	66

3.5	Illustration du principe de dilatation pour différentes valeurs de $d$ . . . . .	67
3.6	Histogramme des valeurs de $\gamma$ pour l'une des couches de ResNet101 entraîné avec une régularisation de type FLOPs. Certains des $ \gamma $ ont été mis à zéro et sont séparés clairement des $ \gamma $ non nuls [26]. . . . .	70
3.7	Combinaison des éléments de $\gamma$ entre eux et avec les poids du filtre de convolution pour former différents schémas de dilatation. Exemple pour $rf_{max} = 9$ [27]. . . . .	72
3.8	Exemple de génération du vecteur de masque $M$ avec des opérations tensorielles différentiables[27]. . . . .	73
3.9	L'architecture TEMPONet [48]. . . . .	75
3.10	Structure du bloc 2 de convolution du réseau TEMPONet [48]. . . . .	75
3.11	Organigramme de la méthode Q-PPG. . . . .	76
3.12	Schéma expliquant la fonctionnalité des deux algorithmes NAS utilisés pour l'optimisation de l'architecture. Les couches de pooling et autres ne sont pas illustrées pour des raisons de simplicité [3]. . . . .	78
3.13	Résultats en MAE (BPM) obtenus avec des architectures possédant différents nombres de paramètres. L'étiquette MN correspond à l'utilisation de MorphNet uniquement, MN-PIT correspond à l'utilisation de MorphNet suivi de PIT. . . . .	82
3.14	Estimation faite sur le meilleur sujet (S13) de PPG-DaLiA avec Q-PPG. . . . .	84
3.15	Estimation faite sur le pire sujet (S5) de PPG-DaLiA avec Q-PPG. . . . .	84
3.16	Résultats en MAE (BPM) de la méthode Q-PPG par sujet et par activité pour l'ensemble de données PPG-DaLiA. . . . .	85
3.17	Estimations produites par le modèle le plus précis sur une partie de l'enregistrement du sujet S13 sans et avec différents post-traitements. . . . .	88
3.18	Estimations faites avec et sans post-traitement sur le sujet S1 de l'ensemble de données IEEE SPC 2015 - train avec la méthode Q-PPG. . . . .	89

# Liste des abréviations

**ACC** Accéléromètre

**BPM** Beat per minute (Battement par minute)

**CNN** Convolutional Neural Network (Réseau de Neurones Convolutif)

**DFT** Discrete Fourier Transform (Transformée de Fourier Discrète)

**ECG** Électrocardiogramme

**FC** Fréquence cardiaque

**FFT** Fast Fourier Transform (Transformée de Fourier Rapide)

**FLOPs** Floating Point Operations per inference (Opérations en virgule flottante par inférence)

**JOSS** JOint Sparse Spectrum reconstruction (Reconstruction Conjointe et Parcimonieuse des Spectres)

**M-FOCUSS** Multiple - FOCal Underdetermined System Solver (Résolveur Focalisé de Systèmes Sous-Déterminés)

**MA** Motion Artifacts (Artefacts de Mouvement)

**MAE** Mean Absolute Error (Erreur absolue moyenne)

**MMV** Multiple Measurement Vector (Vecteur de Mesures Multiples)

**NAS** Neural Architecture Search (Recherche d'architectures Neuronales)

**PIT** Pruning in Time (Elagage temporel)

**PPG** Photoplethysmographie

**SMV** Single Measurement Vector (Vecteur de Mesure Unique)

**SPT** Spectral Peak Tracking (Suivi du Pic Spectral)

**SSR** Sparse Signal Reconstruction (Reconstruction Parcimonieuse de Signaux)

**std** Standard Deviation (Ecart type)

**STFT** Short Term Fourier Transform (Transformée de Fourier à Court Terme)

**TCN** Temporal Convolutional Network (Réseau de Convolution Temporelle)

**WFPV** Wiener Filtering Phase Vocoder (Filtrage de Wiener Vocodeur de Phase)

# Introduction générale

Le suivi de la fréquence cardiaque présente plusieurs utilisations importantes dans le domaine de la santé et du bien-être. Il permet d'évaluer la condition physique, de détecter les problèmes cardiaques potentiels, d'ajuster l'intensité des entraînements physiques et de suivre la récupération après ces derniers.

Les dispositifs modernes portés au poignet, tels que les montres intelligentes, comprennent des ensembles de capteurs de plus en plus hétérogènes qui surveillent les mouvements et les paramètres vitaux, notamment la fréquence cardiaque (FC). Les premiers dispositifs de suivi de la FC portés au poignet étaient connectés à une bande thoracique séparée, équipée d'un simple capteur électrocardiographique (ECG). Bien que très précise, cette solution était coûteuse et inconfortable pour une utilisation quotidienne. Par conséquent, ces dernières années, les bandes thoraciques ECG ont été progressivement remplacées par des capteurs photopléthysmographiques (PPG) plus économiques et plus confortables, qui permettent d'estimer la FC directement à partir du poignet.

La photopléthysmographie (PPG) est une technologie non invasive qui utilise une source de lumière et un photodétecteur à la surface de la peau pour mesurer les variations volumétriques de la circulation sanguine [1]. Le signal PPG est principalement dominé par les ondes de pulsation générées à chaque battement de cœur ; par conséquent la fréquence cardiaque peut être estimée à partir de ce signal [2].

La principale source d'inexactitude lors de l'estimation de la fréquence cardiaque (FC) à partir des signaux PPG est la présence d'artefacts de mouvement (MA) causés par les mouvements du bras et de la main de l'utilisateur. Ces artefacts entraînent une variabilité de la pression exercée par le capteur sur la peau ou une fuite de lumière ambiante entre la photodiode et le poignet [3].

Afin de résoudre la problématique de l'estimation de la fréquence cardiaque à partir de signaux PPG affectés par des artefacts de mouvement, plusieurs méthodes, récemment proposées dans la littérature, exploitent des données inertielles provenant d'accéléromètres, dans le but de faire des estimations plus robustes de la FC. Ces méthodes suivent deux principales approches : l'approche classique basée sur les outils de traitement du signal et l'approche par apprentissage profond.

La plupart des méthodes de l'approche classique ont été développées grâce à la disponibilité de deux ensembles de données dédiés à cette problématique, à savoir IEEE SPC 2015 - train et IEEE SPC 2015 - test, qui ont été publiés lors de la compétition IEEE Signal Processing Cup en 2015. Cependant, la capacité de généralisation de ces méthodes a été remise en question, à plusieurs reprises, en raison de leur nombre élevé de paramètres et seuils heuristiques réglés manuellement. Il est toutefois important de noter que l'éva-

## **Introduction générale**

---

luation de ces dernières a été considérablement restreinte en raison de l'indisponibilité des implémentations de la plupart d'entre elles.

L'utilisation de l'approche par apprentissage profond a pendant longtemps été peu explorée pour la résolution de cette problématique pour deux principales raisons : la complexité du déploiement des modèles d'apprentissage profond sur des appareils portables à ressources limitées, et l'indisponibilité, jusqu'à récemment, d'un ensemble de données suffisamment large et diversifié pour l'entraînement des modèles. La publication du jeu de données PPG-DaLiA en 2019 a tout de même pu régler, à une certaine mesure, le second problème.

Dans ce travail, nous explorons trois méthodes proposées dans la littérature pour l'estimation de la fréquence cardiaque à partir de signaux PPG affectés par des artefacts de mouvement. Nous évaluons leurs performances et leur robustesse sur trois ensembles de données. Les méthodes investiguées sont les suivantes :

- La méthode JOint Sparse Spectrum reconstruction (JOSS) [4] qui est considérée comme l'un des travaux fondateurs de l'approche classique.
- La méthode Wiener Filtering Phase Vocoder (WFPV) [5] qui a été l'une des premières méthodes de l'approche classique à proposer une solution basée sur très peu de paramètres et seuils heuristiques à régler manuellement.
- La méthode Q-PPG [3] qui présente la première méthodologie pour l'optimisation automatique des modèles d'apprentissage profond afin de rendre leur implémentation possible.

La rédaction est organisée comme suit :

- Dans le premier chapitre, nous exposons l'utilité du suivi de la fréquence cardiaque, la technologie derrière la photopléthysmographie et les défis liés à la présence d'artefacts de mouvement dans les signaux PPG. Nous présentons également différents algorithmes et ensembles de données utilisés pour la résolution de notre problématique.
- Dans le deuxième chapitre, nous commençons par présenter les méthodologies proposées par JOSS et WFPV et les fondements théoriques qui les accompagnent. Par la suite, nous implementons les deux méthodes sur MATLAB afin de les évaluer en termes de robustesse et de temps d'exécution en utilisant des ensembles de données sur lesquels elles n'ont jamais été évaluées.
- Dans le troisième chapitre, nous menons différentes expérimentations sur la méthodologie proposée par Q-PPG et nous évaluons ses performances sur deux ensembles de données sur lesquels elle n'a jamais été évaluée.
- Nous terminons notre rapport par une conclusion qui résume les résultats de notre travail et ouvre un ensemble de perspectives de travaux futurs sur la problématique de l'estimation de la FC à partir de signaux PPG affectés par des artefacts de mouvement.

# Chapitre 1

## Généralités et contexte

Ce chapitre introductif expose l'utilité du suivi de la fréquence cardiaque. Il présente la technologie derrière la photopléthysmographie, ses avantages par rapport à l'ECG traditionnel et aborde également le défi posé par la corruption des signaux PPG par des artefacts de mouvement, ce qui complique l'estimation précise de la fréquence cardiaque. Pour surmonter ce défi, de nombreuses méthodes utilisant des outils de traitement du signal et des techniques d'apprentissage profond ont été proposées dans la littérature. Ces dernières sont passées en revue et sont comparées. De plus, ce chapitre présente et compare les trois ensembles de données disponibles publiquement, utilisés pour développer et évaluer les méthodes proposées.

### 1.1 Généralités

#### 1.1.1 Suivi de la fréquence cardiaque

La fréquence cardiaque (FC) est définie comme la fréquence des battements du cœur, mesurée par le nombre de contractions cardiaques par unité de temps, généralement exprimé en battements par minute (BPM) [6].

La FC peut varier en fonction des besoins du corps en oxygène, notamment pendant l'exercice physique. Lorsque le corps a besoin de plus d'oxygène, le rythme cardiaque augmente afin de faciliter la circulation sanguine et de garantir un apport d'oxygène adéquat aux organes. De nombreux facteurs influencent également la fréquence cardiaque, tels que la génétique, la condition physique, l'état psychologique ou le niveau de stress, l'alimentation, l'état hormonal et la présence de maladies ou de pathologies [7].

Le suivi régulier de la fréquence cardiaque possède plusieurs utilisations importantes dans le domaine de la santé et du bien-être. Par exemple, les personnes souffrant d'arythmie (rythme cardiaque irrégulier) ont souvent besoin de surveiller leur fréquence cardiaque quotidiennement pour détecter les variations et les anomalies. Cela leur permet de suivre l'évolution de leur condition [8]. D'autre part, la surveillance de la fréquence cardiaque est également couramment utilisée pour optimiser les routines d'entraînement. Les entraîneurs intègrent depuis longtemps la surveillance de la fréquence cardiaque dans les programmes d'entraînement des athlètes de haut niveau. Cela leur permet de mesurer l'intensité de l'exercice, de surveiller la récupération et d'ajuster l'effort en fonction des objectifs spécifiques de l'athlète [9].

L'électrocardiogramme (ECG) est la méthode privilégiée de surveillance de la fréquence cardiaque dans le domaine médical en raison de sa précision et de sa fiabilité dans la mesure de l'activité électrique du cœur. Cependant, cette méthode n'est pas adaptée pour une utilisation quotidienne en raison de son coût élevé et de l'inconfort qu'elle peut causer aux utilisateurs. Récemment, les moniteurs de fréquence cardiaque utilisant la photopléthysmographie (PPG) gagnent en popularité en raison de leur praticité et accessibilité.

### 1.1.2 Principe de la photopléthysmographie (PPG)

La technologie PPG permet de mesurer les variations du volume sanguin dans les tissus au cours d'un cycle cardiaque à l'aide d'une source lumineuse. Un capteur PPG est principalement composé de deux éléments électroniques : un émetteur de lumière et un détecteur d'intensité lumineuse. En général, une diode électroluminescente (LED) est utilisée comme émetteur de lumière et un photodétecteur pour détecter les variations d'intensité lumineuse [10].

Les capteurs PPG sont conçus sous deux formes distinctes : le mode de transmission et le mode de réflexion, comme illustré dans la figure 1.1. Dans le mode de transmission, la source de lumière et le photodétecteur sont positionnés de part et d'autre des tissus, tandis que dans le mode de réflexion, le photodétecteur est placé du même côté que la source de lumière, le long des tissus, pour mesurer la lumière réfléchie. Les deux types de capteurs permettent d'obtenir des mesures non invasives [1].

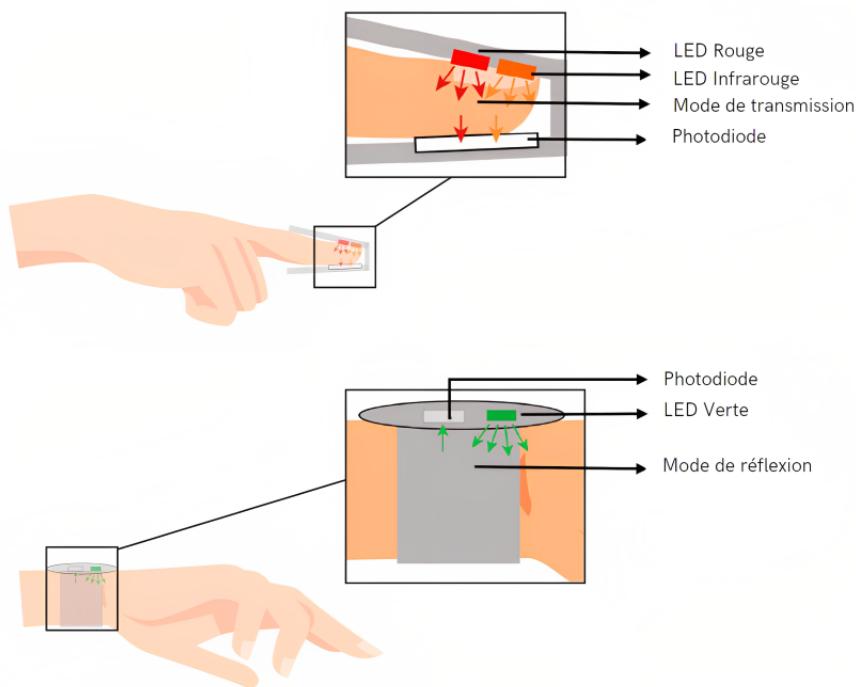


FIG. 1.1 : Illustration du placement du capteur PPG en mode de transmission (en haut) et en mode de réflexion (en bas) [11].

Une impulsion PPG correspondant à un battement cardiaque comprend les phases systolique et diastolique. Pendant la phase systolique, le volume de sang dans les artères est plus élevé car le cœur se contracte et propulse du sang riche en oxygène vers tous les tissus et organes. Cette phase entraîne une absorption accrue de lumière par les globules rouges. Par conséquent, la quantité de lumière détectée par le photodétecteur pendant la phase systolique est faible. Pendant la phase diastolique, le sang reflue vers le cœur. Par conséquent, pendant cette phase, la lumière détectée par le photodétecteur augmente en raison de la diminution du volume sanguin [12]. Les deux phases sont illustrées dans la figure 1.2.

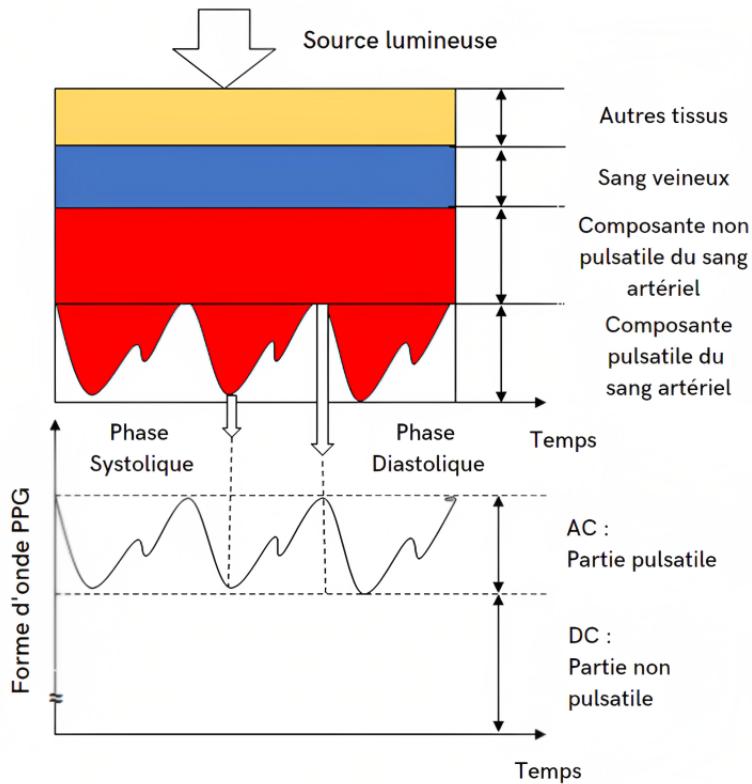


FIG. 1.2 : Illustration de la variation de l'atténuation de la lumière émise par la LED [13].

Un signal PPG est principalement composé de composantes AC et DC. La composante AC dans la forme d'onde de sortie du PPG indique le changement d'intensité lumineuse pendant la phase systolique et diastolique en raison du sang dans les artères [14].

La partie DC, constante de la forme d'onde du PPG, indique la lumière absorbée par les tissus, la peau et les os [13].

Le signal PPG est principalement dominé par les ondes de pulsation générées à chaque battement de cœur ; par conséquent, la fréquence cardiaque peut être estimée à partir de ce signal [2].

### 1.1.3 Avantages de l'utilisation du PPG

Les techniques d'estimation de la FC qui reposent sur des capteurs PPG présentent plusieurs avantages par rapport aux systèmes traditionnels basés sur l'ECG. Par exemple, les capteurs PPG utilisent une implémentation matérielle plus simple et sont moins coûteux. De plus, pour fonctionner, un seul capteur doit être placé sur le corps, ce qui contraste avec les enregistrements ECG traditionnels. Un système ECG traditionnel nécessite au moins trois bio-électrodes placées à différents endroits du corps (comme le bras droit, le bras gauche et la jambe droite) pour pouvoir fonctionner efficacement. Cette exigence limite considérablement la liberté de mouvement de l'utilisateur.

En comparaison avec les différents types d'appareils de surveillance de la fréquence cardiaque basés sur la technologie PPG qui existent, le bracelet au poignet, qui utilise le mode de réflexion, est considéré comme le plus populaire. La raison de sa popularité

réside dans ses propriétés telles que son faible coût, sa grande portabilité et sa grande commodité pour les utilisateurs [1].

### 1.1.4 Artefacts de mouvement dans les signaux PPG

Les signaux PPG acquis au niveau du poignet, pendant différents types de mouvements, sont sujets aux artefacts de mouvement (MA). Les MA sont généralement causés par le déplacement du capteur par rapport à la peau. Cela affecte la qualité du signal et l'extraction des paramètres physiologiques, en l'occurrence la fréquence cardiaque [15].

Au cours des dernières années, plusieurs initiatives de recherche ont obtenu des résultats prometteurs en matière de détection, de suppression ou d'atténuation des artefacts de mouvement [15]. La majorité de ces méthodes utilisent les signaux d'accéléromètre (ACC), plus spécifiquement d'accéléromètre triaxial, comme référence de bruit. Les accéléromètres, intégrés dans les dispositifs portés au poignet, mesurent l'accélération ou les changements de vitesse le long de différents axes. Ces informations peuvent être utilisées pour détecter les artefacts de mouvement (MA) présents dans les signaux PPG.

Les figures 1.3 et 1.4 montrent l'impact de la présence d'artefacts de mouvement sur l'estimation de la FC à partir de signaux PPG.

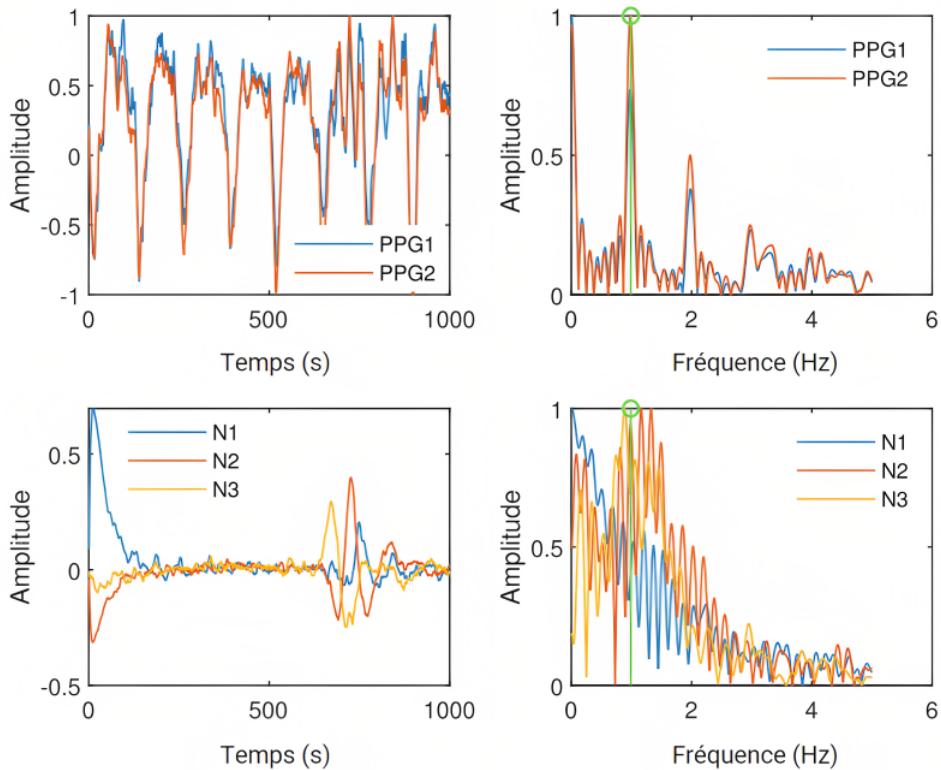


FIG. 1.3 : Signaux PPG et ACC avec leurs spectres en l'absence de MA. Le pic vert, qui correspond à la vraie valeur de la FC, est aligné avec les pics dominants dans les spectres des signaux PPG1 et PPG2. PPG1 et PPG2 sont les canaux PPG, N1, N2 et N3 sont les canaux ACC [16].

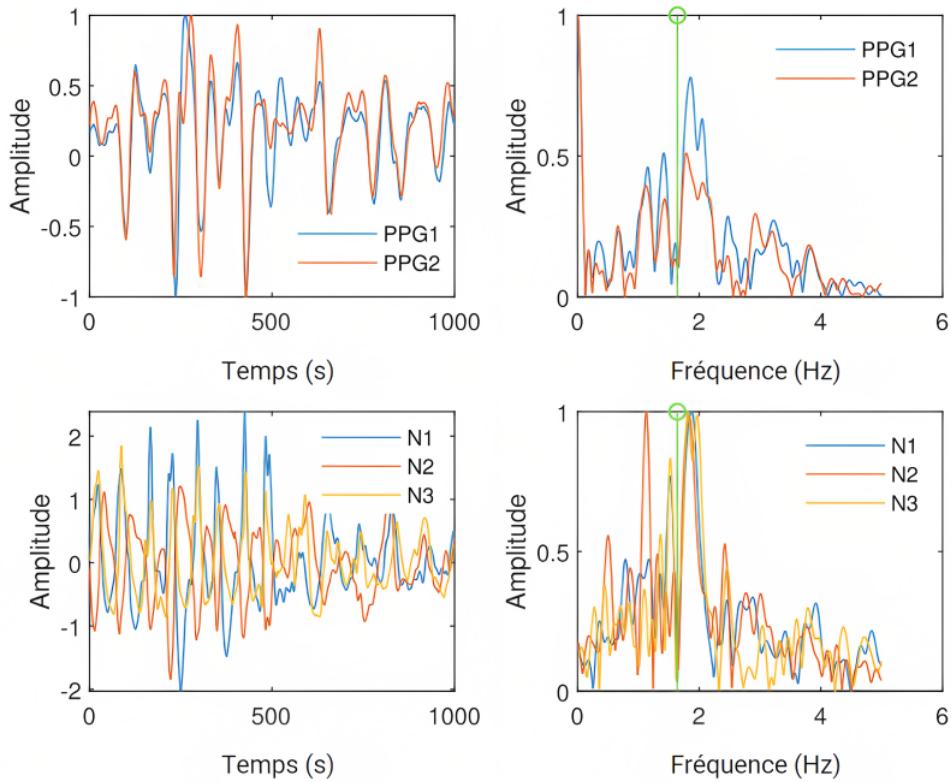


FIG. 1.4 : Signaux PPG et ACC avec leurs spectres en présence de MA. Le pic vert, qui correspond à la vraie valeur de la FC, n'est aligné avec aucun des pics dominants dans les spectres des signaux PPG1 et PPG2. PPG1 et PPG2 sont les canaux PPG, N1, N2 et N3 sont les canaux ACC [16].

## 1.2 État de l'art

Deux approches ont été développées pour aborder la problématique de l'estimation de la fréquence cardiaque à partir de signaux PPG et d'accéléromètre. La première concerne les méthodes classiques de traitement du signal. La seconde concerne l'apprentissage profond.

### 1.2.1 Approche Classique

La méthode TROIKA proposée par Z. Zhang en 2014 [17] est considérée comme la méthode qui a popularisé la problématique de l'estimation de la fréquence cardiaque à partir de signaux PPG fortement affectés par des artefacts de mouvement [18]. En effet, les travaux antérieurs ont été majoritairement proposés pour des scénarios cliniques ou bien dans le cadre d'activités physiques très modérées où les données PPG étaient enregistrées au niveau des doigts ou de l'oreille. Dans ces cas, les artefacts de mouvement n'étaient pas très forts et présentaient même des caractéristiques facilitant l'estimation de la fréquence cardiaque [17].

La méthode TROIKA se compose de trois parties clés, à savoir la décomposition du signal, la reconstruction du signal parcimonieux (SSR) [19] et le suivi du pic spectral (SPT). La décomposition du signal vise à éliminer partiellement les artefacts de mouvement dans le signal PPG brut et à rendre son spectre plus parcimonieux. La reconstruction du signal parcimonieux vise à estimer le spectre haute résolution du signal PPG. L'étape du suivi du pic spectral recherche le pic spectral correspondant à la fréquence cardiaque dans le spectre du signal PPG. Pour surmonter davantage les interférences des artefacts de mouvement et compléter l'approche de sélection du pic, des mécanismes de décision sont conçus pour vérifier le pic spectral sélectionné.

TROIKA a été améliorée par JOSS [4], proposée par le même auteur en 2015, en utilisant une reconstruction conjointe parcimonieuse des spectres des signaux PPG et d'accéléromètre obtenue grâce au modèle de Vecteur de Mesures Multiples (MMV). Une étape de soustraction spectrale est effectuée pour la réduction des pics correspondant aux artefacts de mouvement dans le spectre PPG après les avoir identifiés dans les spectres des signaux d'accéléromètre, ce qui élimine ainsi la nécessité de la décomposition du signal, des étapes de suivi du pic spectral et de vérification sont également proposées comme dans TROIKA.

En 2016, la méthode SpaMA [20] a été proposée par SM. Salehizadeh et al., l'approche SpaMA calcule d'abord la densité spectrale de puissance du signal PPG et des signaux d'accéléromètre par segment. Ensuite, les pics les plus élevés sont identifiés dans chaque spectre. Les pics présents dans les spectres d'accéléromètre sont supprimés dans le spectre PPG afin d'éliminer au maximum les artefacts de mouvement. Le pic le plus élevé dans le spectre PPG nettoyé est considéré comme la fréquence cardiaque recherchée.

En 2017, A. Temko a proposé la méthode WFPV [5]. Cette méthode repose sur l'utilisation du filtre de Wiener et du vocodeur de phase. Le filtre de Wiener est utilisé pour réduire les composantes de bruit présentes dans le signal PPG en utilisant une estimation de la signature du bruit obtenue à partir des signaux d'accéléromètre. Quant au vocodeur de phase, il est utilisé pour surmonter la résolution limitée de la transformée de Fourier discrète et améliorer l'estimation initiale de la fréquence dominante. Contrairement aux méthodes précédentes, la méthode WFPV se distingue par son nombre très réduit de paramètres et de seuils à ajuster.

En 2020, M. Zhou et N. Selvaraj ont proposé la méthode CurToSS [21] qui reprend le même principe que la méthode JOSS pour l'estimation des spectres des signaux PPG et d'accéléromètre et pour la soustraction spectrale mais qui propose une autre méthode pour l'étape du suivi du pic spectral (SPT).

### 1.2.2 Approche par apprentissage profond

L'utilisation de l'approche par apprentissage profond a pendant longtemps été peu explorée pour la résolution de la problématique de l'estimation de la FC à partir de signaux PPG affectés par des artefacts de mouvement pour deux principales raisons : la complexité du déploiement des modèles d'apprentissage profond sur des appareils portables à ressources limitées, et l'indisponibilité, jusqu'à récemment, d'un ensemble de données suffisamment large et diversifié pour l'entraînement des modèles. La publication du jeu

de données PPG-DaLiA [22] en 2019 a tout de même pu régler, à une certaine mesure, le second problème.

Le travail précurseur de cette approche est Deep PPG de A. Reiss et al. en 2019 [22]. La méthode proposée prend en entrée d'un réseau de neurones convolutif (CNN) les représentations temps-fréquence des signaux PPG et d'accéléromètre synchronisés, et fournit l'estimation de la FC en sortie. Différentes valeurs d'hyperparamètres du réseau CNN (nombre de couches de convolution, taille de la couche entièrement connectée, etc.) ont été investiguées par ajustement manuel afin d'optimiser les performances. Le travail a également exploré la possibilité de réduire la taille du réseau tout en maintenant une précision acceptable, à travers des ajustements manuels et des essais-erreurs. En effet, le modèle CNN proposé, qui offre les meilleurs résultats, s'avère trop complexe pour être déployé sur des dispositifs portables à ressources limitées.

Une autre solution d'apprentissage profond, appelée CorNET, a été proposée par D. Biswas et al. en 2019 [23]. Elle vise à estimer la fréquence cardiaque et à réaliser une identification biométrique en utilisant uniquement un signal PPG monocanal. Le modèle utilise un réseau à quatre couches comprenant des couches de convolution, de mémoire à court terme (LSTM), et une couche entièrement connectée en sortie. La couche entièrement connectée finale est personnalisée pour chaque application, comprenant une couche de régression pour prédire la fréquence cardiaque et une couche de classification pour identifier un sujet parmi un groupe.

NAS-PPG, proposée en 2021, par S. B. Song et al. [24] a adopté un modèle composé de plusieurs couches de convolution (CNN) et de couches de mémoire à court terme (LSTM). Une transformée de Fourier rapide (FFT) est appliquée aux signaux PPG et d'accéléromètre avant leur introduction dans le modèle. Un outil de recherche d'architecture neuronale (NAS) [25], qui automatise l'ajustement des hyperparamètres du réseau, a été appliqué pour la première fois dans le but de réduire la taille du modèle. Cette approche a permis d'obtenir un réseau CNN+LSTM de moindre complexité par rapport à Deep PPG.

En 2020, A. Burrello et al. ont introduit Q-PPG [3]. Cette approche propose d'utiliser comme modèle initial, un réseau de convolution temporelle (TCN) qui prend en entrée les segments des séries temporelles brutes du signal PPG et des signaux d'accéléromètre synchronisés. Le choix du TCN est motivé par sa complexité réduite par rapport aux CNN, ce qui le rend plus adapté à une implémentation matérielle. Par la suite, une famille de modèles dérivés, offrant différents compromis entre précision de l'estimation de la FC et complexité du réseau, est obtenue par élagage du réseau initial, en appliquant une cascade de deux outils de recherche automatique d'architecture neuronale (NAS) [26] [27]. Les modèles dérivés grâce à cette méthode ont surpassé les modèles précédemment proposés dans la littérature, que ce soit en termes de précision ou de légèreté.

## 1.3 Ensembles de données disponibles

Dans cette section, nous commençons par présenter les trois ensembles de données les plus utilisés pour la problématique de l'estimation de la fréquence cardiaque à partir de signaux PPG et d'accéléromètre. Ensuite, nous procédons à une comparaison de ces derniers.

### 1.3.1 IEEE SPC 2015 – train

Cet ensemble de données a été proposé à l'occasion de la compétition IEEE Signal Processing Cup en 2015, il contient **12 sessions d'enregistrements de 5 minutes** faits sur **12 sujets** masculins en bonne santé, âgés de 18 à 35 ans (les détails sont donnés dans le tableau 1.1). Pour chaque sujet, il inclut deux canaux de signaux PPG, trois canaux de signaux d'accéléromètre et un canal de signal ECG enregistrés simultanément. Les signaux PPG et d'accéléromètre ont été enregistrés à l'aide d'un dispositif porté au poignet, comprenant un capteur PPG à deux canaux (LED verte, longueur d'onde : 515 nm) et un accéléromètre triaxial (à la même position que le capteur PPG). Le signal ECG quant à lui a été enregistré depuis la poitrine à l'aide d'électrodes ECG humides. Tous les signaux ont été initialement échantillonnés à 125 Hz.

Pendant l'enregistrement des données, les sujets ont marché ou couru sur un tapis de course avec les vitesses suivantes (dans l'ordre) : une vitesse de 1 à 2 km/h pendant 0.5 minute, une vitesse de 6 à 8 km/h pendant 1 minute, une vitesse de 12 à 15 km/h pendant 1 minute, une vitesse de 6 à 8 km/h pendant 1 minute, une vitesse de 12 à 15 km/h pendant 1 minute et une vitesse de 1 à 2 km/h pendant 0.5 minute. Les sujets ont utilisé la main avec le bracelet au poignet pour appuyer sur les boutons du tapis de course, essuyer la sueur sur leurs fronts, etc.

L'ECG a été utilisé pour fournir la fréquence cardiaque de référence en BPM. Les fréquences cardiaques basées sur l'ECG ont été calculées pour chaque fenêtre de 8 secondes avec un décalage de 2 secondes, sans recourir à aucun algorithme de détection de pics.

### 1.3.2 IEEE SPC 2015 – test

Cet ensemble de données a également été proposé à l'occasion de la compétition IEEE Signal Processing Cup en 2015, il contient **10 sessions d'enregistrements de 5 minutes** faits sur **8 sujets** en bonne santé, 7 hommes et une femme, âgés de 19 à 58 ans (les détails sont donnés dans le tableau 1.2). La configuration matérielle ainsi que les signaux fournis sont identiques à ceux de l'ensemble de données **IEEE SPC 2015 – train**. Cependant, le type d'activités diffère.

Deux types d'activités ont été réalisés. L'activité de type 1 (**T1**) comprenait différents exercices de l'avant-bras et du bras supérieur couramment utilisés en rééducation du bras (par exemple, secouer les mains, s'étirer, faire des pompes, etc.). L'activité de type 2 (**T2**) consistait en des mouvements intensifs de l'avant-bras et du bras supérieur (par exemple, la boxe).

Session	Identifiant Sujet	Âge /Poids /taille	Sexe	En bonne santé ?
S1	1	18-35 ans (Pas plus d'informations)	H	Oui
S2	2		H	Oui
S3	3		H	Oui
S4	4		H	Oui
S5	5		H	Oui
S6	6		H	Oui
S7	7		H	Oui
S8	8		H	Oui
S9	9		H	Oui
S10	10		H	Oui
S11	11		H	Oui
S12	12		H	Oui

TAB. 1.1 : Description de l'ensemble de données IEEE SPC 2015 – train

Session	Identifiant Sujet	Type de l'activité	Âge /Poids /taille	Sexe	En bonne santé ?
S1	1	T2	29 ans /70 kg /169 cm	H	Oui
S2	2	T2	21 ans /77 kg /188 cm	H	Oui
S3	2	T3	21 ans /77 kg /188 cm	H	Oui
S4	3	T3	19 ans /54 kg /174 cm	H	Oui
S5	4	T3	20 ans /64 kg /162 cm	H	Oui
S6	5	T3	20 ans /57 kg /174 cm	H	Oui
S7	6	T2	19 ans /70 kg /180 cm	H	Oui
S8	6	T3	19 ans /70 kg /180 cm	H	Oui
S9	7	T3	21 ans /73 kg /180 cm	H	Oui
S10	8	T2	58 ans /70 kg /156 cm	F	Non*

TAB. 1.2 : Description de l'ensemble de données IEEE SPC 2015 – test. T1 = exercices de rééducation du bras. T2 = mouvements intensifs du bras (boxe). \* : un rythme cardiaque et une tension artérielle anormaux ont été observés.

→ Les deux ensembles de données peuvent être téléchargés depuis [28].

### 1.3.3 PPG-DaLiA

En 2019, un nouvel ensemble de données, appelé PPG-DaLiA (PPG dataset for motion compensation and heart rate estimation in Daily Life Activities), a été proposé dans [22]. PPG-DaLiA peut être téléchargé depuis [29]. Jusqu'à présent, ce jeu de données est le plus grand et le plus diversifié disponible publiquement pour la problématique d'estimation de la FC à partir de signaux PPG présentant des artefacts de mouvement.

L'ensemble de données comprend au total plus de **36 heures d'enregistrements** sur **15 sujets**. Il considère **8 types différents d'activités** typiquement réalisés dans la vie quotidienne et des périodes de transition entre celles-ci.

Le jeu de données contient 15 sessions d'enregistrements faits sur 15 sujets différents, 7 hommes et 8 femmes, âgés de  $30.60 \pm 9.59$  ans. Les détails spécifiques à chaque sujet sont présentés dans le tableau 1.3.

Session	Sujet	Âge /Poids /taille	Sexe	Condition physique
S1	1	34 ans /78 kg /182 cm	H	6
S2	2	28 ans /80 kg /189 cm	H	5
S3	3	25 ans /60 kg /170 cm	H	5
S4	4	28 ans /57 kg /168 cm	H	5
S5	5	21 ans /70 kg /180 cm	F	4
S6	6	37 ans /70 kg /176 cm	F	1
S7	7	21 ans /58 kg /168 cm	F	2
S8	8	43 ans /70 kg /179 cm	H	5
S9	9	28 ans /60 kg /167 cm	F	5
S10	10	58 ans /70 kg /156 cm	F	5
S11	11	24 ans /68 kg /162 cm	F	5
S12	12	43 ans /105 kg /195 cm	H	5
S13	13	21 ans /63 kg /170 cm	F	6
S14	14	26 ans /67 kg /170 cm	F	4
S15	15	28 ans /79 kg /183 cm	H	5

TAB. 1.3 : Aperçu des sujets participant à la collecte de données dans PPG-DaLiA. Condition physique : fréquence de pratique sportive du sujet, sur une échelle de 1 à 6, où 1 correspond à moins d'une fois par mois et 6 correspond à 5 à 7 fois par semaine.

Chaque sujet a effectué 8 activités différentes, dont les durées approximatives sont données dans le tableau 1.4. Cependant, afin de collecter des données dans des conditions naturelles proches de la vie quotidienne, les sujets étaient autorisés à s'écartier de la durée prévue si une tâche était terminée plus tôt ou prenait plus de temps.

Les activités incluses, dans l'ordre dans lequel elles ont été réalisées, sont les suivantes :

- **Assis** : Rester assis en lisant.
- **Escaliers** : Monter six étages et redescendre, en répétant cette opération deux fois.  
Note : pour les sujets S1 et S2, la descente a été effectuée une seule fois.
- **Babyfoot** : Jouer au baby-foot.
- **Vélo** : Effectué en extérieur, en suivant un itinéraire défini, d'environ 2 km, avec des conditions de route variables (graviers, pavés).

- **Conduite** : Réalisée sur un itinéraire défini, d'environ 15 minutes, comprenant la conduite sur différentes routes dans une petite ville ainsi que la conduite sur des routes de campagne.
- **Déjeuner** : Réalisée dans une cantine, comprenant la file d'attente, la récupération de la nourriture, la prise du repas et la conversation à table.
- **Marche** : Marche du retour de la cantine au bureau, avec un détour.
- **Travail** : Les sujets sont retournés à leurs bureaux et ont travaillé comme s'ils ne participaient pas à l'étude. Le travail consistait principalement à utiliser un ordinateur.

Des périodes de transition ont été incluses entre la plupart des activités pour se rendre sur le lieu de l'activité suivante. Dans l'ensemble, le protocole de collecte de données a duré environ 2.5 heures pour chaque sujet.

**Remarque** : En raison d'un problème matériel, les enregistrements du sujet S6 ne sont pas exploitables pour les 3 dernières activités (Déjeuner, Marche et Travail).

Activité	Durée (min)	FC moyenne (BPM)
Assis	10	61.07
Escaliers	5	118.53
Babyfoot	5	90.4
Vélo	8	123.05
Conduite	15	84.7
Déjeuner	30	83.58
Marche	10	99.18
Travail	20	76.4

TAB. 1.4 : Durée approximative et fréquence cardiaque moyenne de chaque activité dans PPG-DaLiA.

Le tableau 1.4 fournit également la fréquence cardiaque moyenne par activité. Cette information permettra d'expliquer les différences de performances par activités des solutions investiguées dans les chapitres ultérieurs.

Les trois capteurs utilisés dans le cadre de notre problématique sont : un capteur PPG porté au poignet, un accéléromètre triaxial intégré dans le même dispositif que le capteur PPG et un capteur ECG.

Le dispositif commercial Empatica E4 [30], porté au poignet le moins sollicité par le sujet, a été utilisé pour collecter les signaux PPG et d'accéléromètre triaxial. Le capteur PPG est composé de quatre LED (deux vertes et deux rouges) et de deux unités photodiode, sa sortie est échantillonnée à une fréquence de 64 Hz, tandis que les signaux d'accéléromètre sont échantillonnés à 32 Hz.

D'autre part, le dispositif commercial RespiBAN Professional [31], porté au niveau de la poitrine, a été utilisé pour l'enregistrement du signal ECG via un capteur ECG standard

à trois points avec des électrodes Ag/AgCl. Le signal ECG de sortie a été échantillonné à 700 Hz et filtré en bande passante de 0.5 à 100 Hz.

Tout comme les deux autres ensembles de données, le signal ECG a été utilisé pour déterminer la fréquence cardiaque de référence en BPM. Les fréquences cardiaques basées sur l'ECG ont été calculées pour chaque fenêtre de 8 secondes avec un décalage de 2 secondes selon la méthode suivante : un détecteur de pics R [32] a été utilisé, suivi d'une correction manuelle. À partir des pics R identifiés et corrigés, la fréquence cardiaque instantanée a été calculée et la fréquence cardiaque de référence a été définie comme la moyenne des fréquences cardiaques instantanées sur chaque fenêtre de 8 secondes.

### 1.3.4 Comparaison des ensembles de données

Les deux jeux de données introduits dans le cadre de la compétition IEEE Signal Processing Cup en 2015 (IEEE SPC 2015 - train et IEEE SPC 2015 - test) présentent plusieurs limitations :

- Leur taille : bien que le nombre de sujets puisse être considéré comme suffisant (respectivement 12 et 8 sujets), la durée d'enregistrement de chaque session est très courte (environ 5 minutes en moyenne) et les segments de l'enregistrement les plus affectés par les artefacts de mouvement, qui sont ceux qui nous intéressent du point de vue de l'évaluation, ne représentent qu'une petite partie de la totalité.
- Le nombre d'activités incluses est limité : il s'agit uniquement de la course sur un tapis de course à une vitesse variable dans IEEE SPC 2015 - train et d'exercices impliquant l'utilisation des bras dans IEEE SPC 2015 - test. Cependant, il est important de noter que différentes activités physiques produisent différents artefacts de mouvement. Ainsi, une plus grande variété d'activités est nécessaire pour évaluer la capacité d'une méthode à traiter différents scénarios.
- L'enregistrement a été réalisé dans un environnement contrôlé : la course sur tapis a été effectuée dans un laboratoire sur une machine d'exercice. Cependant, il est important de noter que les artefacts de mouvement générés par les activités de la vie quotidienne peuvent différer de ceux présents dans ces ensembles de données.

L'ensemble de données PPG-DaLiA, introduit plus récemment, est considérablement plus volumineux. Il comprend 36h d'enregistrement au total (environ 2.5 heures par session), ce qui correspond à environ 65 000 échantillons, en comparaison des 1 800 et 1 300 échantillons des deux jeux de données IEEE SPC 2015. Chaque échantillon se réfère à un segment de données de 8 secondes, pour lequel une étiquette (valeur réelle de la fréquence cardiaque) est définie.

De plus, PPG-DaLiA propose une gamme d'activités plus diversifiée. Elle comprend des activités à faible intensité (Conduite), des activités à intensité modérée (Marche), et des mouvements intenses des bras (Babyfoot), afin de générer des artefacts de mouvement à différentes amplitudes. Les activités prises en compte sont pour certaines périodiques (Marche et Escalier), et pour d'autres apériodiques (Manger et Babyfoot). De plus, afin

de générer des fréquences cardiaques très variables, des activités nécessitant différents niveaux d'efforts physiques ont été selectionnées, par exemple, Conduite par rapport à Escalier.

Enfin, bien que les données dans PPG-DaLiA n'aient pas été tirées directement à partir de la vie quotidienne des sujets, ces derniers ont réalisé les activités de manière naturelle, dans des conditions proches de la réalité. Les instructions étaient vagues, la durée approximative et les sujets autorisés à s'en écarter si une tâche était terminée plus tôt ou prenait plus de temps.

**Remarque :** Pour PPG-DaLiA et IEEE SPC 2015 –train, les termes **session** et **sujet** sont interchangeables car chaque session a été effectuée par un sujet différent. Cependant, pour IEEE SPC 2015 – test, les 10 sessions ont été réalisées par seulement 8 sujets.

### 1.4 Métriques d'évaluation

Comme cela a été mentionné précédemment, les fréquences cardiaques de référence, basées sur l'ECG, sont calculées pour chaque fenêtre de 8 secondes avec un décalage de 2 secondes, et elles sont fournies dans chaque ensemble de données. Les méthodes proposées dans la littérature utilisent la même longueur de fenêtre et le même décalage pour estimer la fréquence cardiaque, afin d'obtenir autant d'estimations de la fréquence cardiaque que de valeurs de référence.

La métrique la plus utilisée dans la littérature pour l'**évaluation des performances** des méthodes est l'erreur absolue moyenne (MAE) en BPM. Elle est calculée comme suit :

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\text{FC}_{\text{est}}(i) - \text{FC}_{\text{ref}}(i)| \quad (1.1)$$

où  $\text{FC}_{\text{ref}}(i)$  représente la fréquence cardiaque de référence à la  $i$ -ème fenêtre temporelle,  $\text{FC}_{\text{est}}(i)$  représente la fréquence cardiaque estimée et  $N$  est le nombre total de fenêtres temporelles.

Pour l'**évaluation de la complexité** des modèles d'apprentissage profond, nous utilisons deux métriques :

- Le nombre de paramètres du modèle.
- Le nombre d'opérations en virgule flottante (FLOPs) par inférence.

## 1.5 Comparaison des méthodes

Méthode	Ensembles de données	Algorithme	MAE (BPM)
<b>Méthodes classiques</b>			
TROIKA, 2014 [17]	IEEE SPC 2015 - train	Décomposition du signal SSR SPT	<b>2.34</b>
JOSS, 2015 [4]	IEEE SPC 2015 - train	MMV Soustraction spectrale SPT	<b>1.28</b>
SpaMa, 2016 [20]	IEEE SPC 2015 - train IEEE SPC 2015 - test PPG-DaLiA	Filtrage spectral basé sur la PSD	<b>0.89</b> <b>3.35</b> <b>11.06</b>
WFPV, 2017 [5]	IEEE SPC 2015 - train IEEE SPC 2015 - test	Filtre de Wiener Vocodeur de phase	<b>1.02</b> <b>2.96</b>
CurToSS, 2020 [21]	IEEE SPC 2015 - train IEEE SPC 2015 - test PPG-DaLiA	MMV Soustraction spectrale Suivi de la courbe	<b>2.2</b> <b>4.5</b> <b>5.0</b>
<b>Méthodes d'Apprentissage Profond</b>			
DeepPPG, 2019 [22]	IEEE SPC 2015 - train IEEE SPC 2015 - test PPG-DaLiA	CNN	<b>4</b> <b>16.51</b> <b>7.65</b>
CorNET, 2019 [23]	IEEE SPC 2015 - train IEEE SPC 2015 - test	CNN + LSTM	<b>4.67</b> <b>5.55</b>
NAS-PPG, 2021 [24]	PPG-DaLiA	CNN + LSTM	<b>6.02</b>
Q-PPG, 2022 [3]	PPG-DaLiA	TCN	<b>4.36</b>

TAB. 1.5 : Tableau de comparaison des méthodes de l'état de l'art. Les différents résultats de MAE (BPM) correspondent à différents ensembles de données

En ce qui concerne les méthodes de l'approche classique, d'après le tableau 1.5, nous pouvons constater que les performances des méthodes TROIKA [17] et JOSS [4] ont été évaluées uniquement sur l'ensemble de données IEEE SPC 2015 - train, sur lequel elles obtiennent de très bons résultats. JOSS a effectivement amélioré les performances de TROIKA, réduisant l'erreur absolue moyenne de 2.34 BPM à 1.28 BPM. Cependant, il convient de noter que le jeu de données IEEE SPC 2015 - train est considéré comme celui contenant les activités les plus simples et présentant le plus de limitations, comme cela a été détaillé dans 1.3.4. De plus, nous pouvons observer que la méthode SpaMa [20], qui affiche le meilleur résultat sur l'ensemble de données IEEE SPC 2015 - train, n'obtient pas la plus faible erreur sur IEEE SPC 2015 - test, et présente une erreur très élevée sur l'ensemble de données PPG-DaLiA (cette méthode, publiée en 2016, a été réévaluée en 2019 dans [22] lors de la publication de PPG-DaLiA).

La méthode WFPV [5] obtient de très bons résultats sur les ensembles de données IEEE SPC 2015 (train et test), mais n'a pas été évaluée sur le jeu de données PPG-DaLiA. Quant à la méthode CurToSS [21], elle affiche de très bons résultats sur les trois ensembles de données. Toutefois, elle présente une complexité de calcul bien plus élevée que toutes les autres méthodes comme cela a été mis en lumière dans [3].

Dans le cadre de ce travail, nous prévoyons d'étendre les évaluations réalisées sur les méthodes JOSS et WFPV, notamment en termes de robustesse, étant donné qu'elles ont obtenu de très bons résultats, mais n'ont pas été explorées plus en détail depuis la publication de l'ensemble de données PPG-DaLiA.

Deep PPG [22] a été la première solution d'apprentissage profond proposée dans la littérature pour la résolution de la problématique et elle s'est avérée plus performante que la méthode SpaMa [20] sur l'ensemble de données PPG-DaLiA mais pas sur les ensembles de données IEEE SPC 2015 - train et IEEE SPC 2015 - test. La méthode CorNET [23], qui a proposé un modèle d'apprentissage profond différent, a obtenu des résultats comparables à ceux des méthodes classiques sur les ensembles de données IEEE SPC 2015 (-train et -test), mais n'a pas été évaluée sur PPG-DaLiA.

Ces deux méthodes sont basées sur des architectures conçues manuellement, ce qui peut être sous-optimal en termes de compromis précision-complexité. Une approche de recherche d'architecture neuronale (NAS) a été appliquée au modèle dans NAS-PPG [24], obtenant un réseau CNN+LSTM qui atteint un bon résultat sur l'ensembe de données PPG-DaLiA, tout en réduisant la complexité du modèle par rapport à Deep PPG. Le modèle reste toutefois trop volumineux pour un déploiement sur des appareils portables (800 000 paramètres en virgule flottante).

Q-PPG a été la première méthode à proposer une méthodologie pour la dérivation automatique de plusieurs modèles qui offrent différents compromis entre précision et complexité (cela permet de sélectionner le modèle le plus adapté en fonction des contraintes de la cible matérielle). De plus, Q-PPG surpasse toutes les solutions précédentes d'apprentissage profond sur l'ensemble de données PPG-DaLiA. Par rapport à Deep PPG, une meilleure MAE est obtenue, même avec le modèle le plus simple (6.07 BPM contre 7.65 BPM), malgré une réduction de la complexité de  $7572\times$ . Par rapport à NAS-PPG, le même modèle Q-PPG obtient une MAE comparable (6.07 BPM contre 6.02 BPM) avec  $88,4\times$  moins d'opérations [3]. De plus, le modèle Q-PPG le plus précis obtient une MAE de 4.36 BPM, ce qui surpasse toutes les méthodes précédentes sur cet ensemble de données. Cela nous motive à explorer davantage la méthode Q-PPG.

## **Conclusion**

Dans ce chapitre, nous avons introduit la problématique de l'estimation de la fréquence cardiaque à partir de signaux PPG affectés par des artefacts de mouvement. Les méthodes existantes dans la littérature, qui abordent cette problématique en intégrant les données PPG avec des données inertielles provenant d'accéléromètres, ont été passées en revue. Les ensembles de données utilisés pour le développement et l'évaluation de ces méthodes ont été présentés, et leurs limites ont été discutées. Enfin, une analyse comparative des méthodes étudiées a été présentée afin de motiver les deux chapitres suivants.

# **Chapitre 2**

## **Approche Classique**

Dans le chapitre précédent, nous avons eu un aperçu des différentes approches et méthodes proposées dans la littérature pour l'estimation de la fréquence cardiaque à partir de signaux PPG fortement affectés par des artefacts de mouvement. Dans ce chapitre, nous explorons deux méthodes de l'approche classique : JOSS et WFPV. Nous commençons d'abord par présenter les méthodologies proposées par chacune d'elles ainsi que les fondements théoriques qui les accompagnent. Par la suite, nous implémentons les deux méthodes sur MATLAB afin de les évaluer en termes de robustesse et de temps d'exécution. Ces deux aspects revêtant une importance primordiale dans le contexte d'un déploiement sur des appareils portables utilisés au quotidien.

## 2.1 Méthode JOint Sparse Spectrum reconstruction (JOSS)

### 2.1.1 Méthodologie proposée

La méthode JOSS a été proposée en 2015 par Z. Zhang [4]. Cette approche consiste en 3 étapes clés :

- L'**estimation conjointe des spectres haute résolution et parcimonieux** des signaux PPG et d'accéléromètre en utilisant le modèle de vecteur de mesures multiples (MMV).
- Une **soustraction spectrale** pour la réduction des pics correspondant aux artefacts de mouvement dans le spectre PPG après les avoir identifiés dans les spectres des signaux d'accéléromètre.
- Le **suivi du pic spectral (SPT)** pour sélectionner le pic associé à la fréquence cardiaque (FC) dans le spectre PPG nettoyé. Cette sortie est ensuite exploitée afin d'estimer la FC dans les fenêtres temporelles suivantes.

L'organigramme de la méthode est illustré dans la figure suivante :

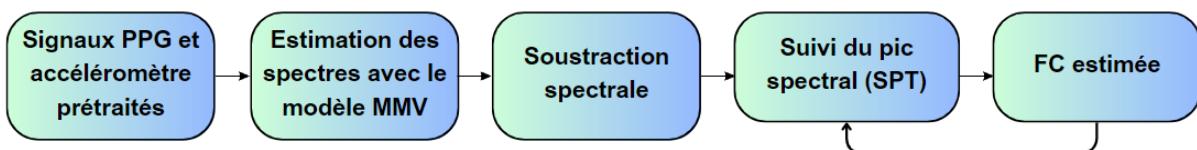


FIG. 2.1 : Organigramme de la méthode JOSS.

Les 3 étapes de la méthode sont discutées en détail dans ce qui suit.

### 2.1.1.1 Reconstruction conjointe parcimonieuse des spectres en utilisant le modèle MMV

Afin d'estimer la FC, le domaine spectral est utilisé. Cependant, la méthode conventionnelle d'estimation du spectre qui est le périodogramme présente une limitation : l'effet de fuite spectrale, où de l'énergie spectrale d'un pic dominant "fuit" dans des bandes de fréquences adjacentes. Cet effet peut rendre difficile la distinction entre un pic spectral faible associé à la FC et un pic dominant associé à un artefact de mouvement s'ils sont proches [17].

Pour surmonter cette limitation, JOSS utilise une estimation conjointe des spectres des signaux PPG et d'accéléromètre basée sur la reconstruction de signaux parcimonieux (SSR) car cette méthode a démontré une plus grande résolution de spectres, ce qui est nécessaire pour la problématique de l'estimation de la FC.

#### Reconstruction de signaux parcimonieux (SSR) et modèle MMV

La reconstruction de signaux parcimonieux (SSR) [19] est utilisée pour trouver la représentation la plus parcimonieuse d'un signal observé, c'est-à-dire qui ne contient que quelques coefficients non nuls, dans une base donnée. Cette technique a de nombreuses applications, notamment dans la compression de données.

Le modèle SSR de base peut être exprimé comme suit :

$$\mathbf{y} = \Phi \mathbf{x} + \mathbf{v} \quad (2.1)$$

Ici,  $\Phi$  est la matrice base connue de taille  $M \times N$ . Le vecteur  $\mathbf{y}$  est un signal observé de taille  $M \times 1$ .  $\mathbf{x}$  est un vecteur de solution inconnu qui est supposé être parcimonieux ou compressif.  $\mathbf{v}$  est un vecteur de bruit inconnu. Le but est de trouver le vecteur le plus parcimonieux  $\mathbf{x}$  à partir de  $\mathbf{y}$  et  $\Phi$ . Cela peut être réalisé en résolvant le problème d'optimisation suivant :

$$\hat{\mathbf{x}} \leftarrow \min_{\mathbf{x}} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \lambda g(\mathbf{x}) \quad (2.2)$$

où  $\lambda$  est un paramètre de régularisation et  $g(\mathbf{x})$  est une fonction de contrainte qui encourage la parcimonie de la solution  $\hat{\mathbf{x}}$ . Les fonctions de contrainte les plus utilisées sont la norme  $\ell_1$ ,  $\|\mathbf{x}\|_1$ , ou la norme  $\ell_p$  ( $0 < p < 1$ ),  $\|\mathbf{x}\|_p$ . Ce modèle est dit modèle SMV ou modèle de Vecteur de Mesure Unique, car le signal  $\mathbf{y} \in \mathbb{R}^{M \times 1}$  est représenté par un seul canal de mesure.

Le modèle de Vecteur de Mesures Multiples (MMV) est une extension du modèle SMV permettant d'estimer des représentations parcimonieuses de signaux multicanaux, en estimant les vecteurs de solution conjointement à partir de plusieurs vecteurs de mesures.

Le modèle est donné par :

$$\mathbf{Y} = \boldsymbol{\Phi} \mathbf{X} + \mathbf{V} \quad (2.3)$$

où  $\mathbf{Y}$  de taille  $M \times L$  est la matrice signal multicanal observé, composée de  $L$  vecteurs de mesure,  $\boldsymbol{\Phi}$  est la matrice base connue de taille  $M \times N$ ,  $\mathbf{X}$  de taille  $N \times L$  est la matrice solution parcimonieuse désirée et  $\mathbf{V}$  de taille  $M \times L$  représente le bruit de mesure ou les erreurs de modèle. Les colonnes de la matrice solution  $\mathbf{X}$  ont une contrainte de parcimonie commune,  $\mathbf{X}$  est parcimonieuse par ligne, c'est-à-dire que seules quelques lignes de  $\mathbf{X}$  sont non nulles alors que la plupart des lignes sont nulles.

### Estimation des spectres avec le modèle MMV

La méthode JOSS utilise le modèle MMV pour estimer conjointement les spectres parcimonieux des signaux PPG et d'accéléromètre.

Pour cela, les colonnes de  $\mathbf{Y} \in \mathbb{R}^{M \times L}$  sont prises comme étant les segments simultanés des signaux PPG et d'accéléromètre dans la fenêtre considérée ( $L = 4$ ), les  $M$  lignes de  $\boldsymbol{\Phi} \in \mathbb{C}^{M \times N}$  ( $M < N$ ) ont comme éléments la base de la transformée de Fourier discrète :

$$\Phi_{m,n} = e^{j \frac{2\pi}{N} mn}, \quad m = 0, \dots, M-1 \quad ; \quad n = 0, \dots, N-1 \quad (2.4)$$

chaque colonne de la matrice solution souhaitée  $\mathbf{X} \in \mathbb{C}^{N \times L}$  contient les coefficients de chaque exponentielle complexe présente dans le spectre du signal correspondant dans  $\mathbf{Y}$ . La multiplication matricielle entre  $\boldsymbol{\Phi}$  et  $\mathbf{X}$  donne une transformée de Fourier inverse de  $\mathbf{X}$ . Les éléments des spectres estimés de chacun des  $L$  signaux ( $L = 4$ ) sont donnés par :

$$s^{(l)}[i] = |\hat{x}^{(l)}[i]|^2, \quad i = 1, \dots, N \quad ; \quad l = 1, \dots, L \quad (2.5)$$

L'algorithme M-FOCUSS (FOCal Underdetermined System Solver) régularisé est adopté pour résoudre itérativement le problème d'optimisation, encourageant la parcimonie de la solution. L'algorithme est décrit dans [33] et se résume comme suit :

$$\mathbf{W}_{k+1} = \text{diag}(c_k[i]^{1-p/2}) \quad \text{où } c_k[i] = \left( \sum_{l=1}^L \left( x_k^{(l)}[i] \right)^2 \right)^{1/2}, \quad p \in [0, 2] \quad (2.6)$$

$$\mathbf{Q}_{k+1} = \boldsymbol{\Phi}_{k+1}^H \left( \boldsymbol{\Phi}_{k+1} \boldsymbol{\Phi}_{k+1}^H + \lambda \mathbf{I} \right)^{-1} \mathbf{Y} \quad \text{où } \boldsymbol{\Phi}_{k+1} = \boldsymbol{\Phi}_k \mathbf{W}_{k+1} \text{ avec } \lambda \geq 0 \quad (2.7)$$

$$\mathbf{X}_{k+1} = \mathbf{W}_{k+1} \mathbf{Q}_{k+1} \quad (2.8)$$

Ici,  $k$  désigne la  $k$ -ème itération,  $l$  la  $l$ -ème colonne et  $i$  la  $i$ -ème ligne de la matrice solution parcimonieuse  $\mathbf{X}$ ,  $\text{diag}$  l'opération de diagonalisation,  $\boldsymbol{\Phi}^H$  est l'hermitien de  $\boldsymbol{\Phi}$  et  $\mathbf{I}$  la matrice identité.

Les paramètres clés de l'algorithme M-FOCUSS régularisé sont :

- $p$ , qui détermine le choix de la norme contrainte encourageant la parcimonie de la solution. Les valeurs de  $p$  inférieures à 1 donnent des solutions parcimonieuses.

L'algorithme converge plus rapidement pour de faibles valeurs de  $p$ . Cependant, pour de faibles valeurs de  $p$ , il a été constaté que l'algorithme se retrouve piégé dans un minimum local. Des valeurs entre 0.8 et 1.0 présentent un bon compromis entre vitesse de convergence et qualité de la solution, [33]

- $max\_iters$ , le nombre maximal d'itérations,
- $\lambda$ , le paramètre de régularisation, une valeur  $\lambda$  plus grande encourage la parcimonie de la solution au détriment de la précision de reconstruction. [33]

Pour l'initialisation de la matrice solution  $\mathbf{X}$ , il a été démontré que n'importe quel point de départ aléatoire permettait à l'algorithme de converger.

### 2.1.1.2 Soustraction Spectrale

La contrainte de parcimonie commune du modèle MMV aide à identifier les pics correspondant aux artefacts de mouvement (MA) dans le spectre PPG en utilisant les spectres des signaux d'accéléromètre étant donné que la contrainte encourage les emplacements de fréquences de MA dans le spectre PPG à être alignés avec ceux dans les spectres  $ACC_{x,y,z}$ . Le spectre du PPG est ainsi nettoyé par soustraction spectrale comme suit :

- A chaque fréquence  $f_i$  ( $i = 1, \dots, N$ ), le maximum des coefficients des 3 spectres des signaux d'accéléromètre,  $C_i$ , est choisi.
- $C_i$  ( $i = 1, \dots, N$ ) est soustrait du coefficient du spectre PPG se trouvant à la fréquence  $f_i$ .
- Tous les coefficients spectraux avec des valeurs  $< \frac{1}{4}$  du pic maximum du spectre PPG obtenu sont mis à zéro.

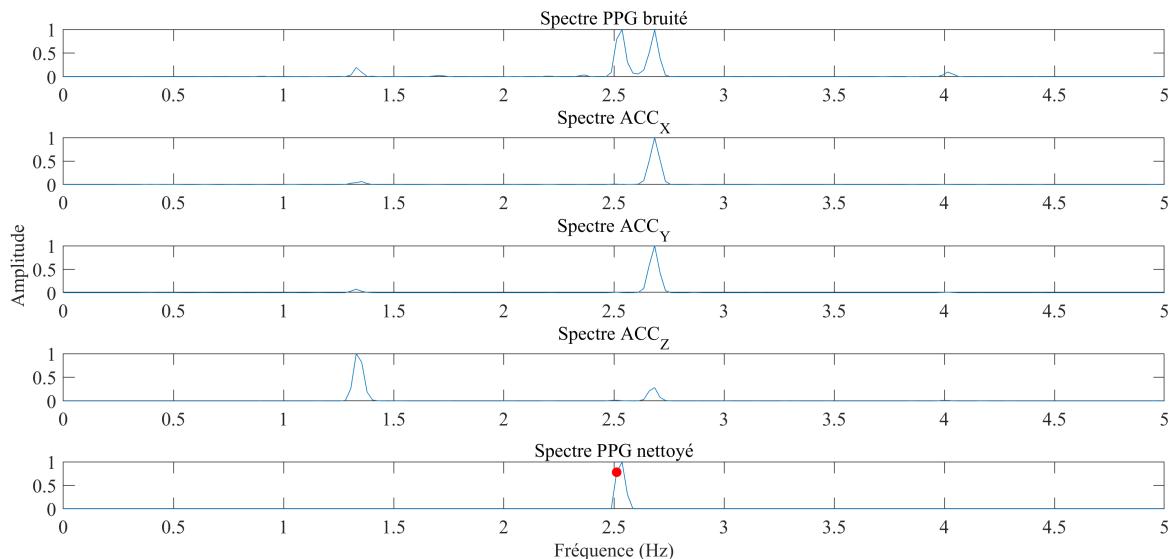


FIG. 2.2 : Illustration de l'étape de soustraction spectrale sur une fenêtre donnée pour la méthode JOSS. Le cercle rouge indique la valeur réelle de la FC en Hz.

### 2.1.1.3 Suivi du pic spectral (SPT)

Cette étape de suivi du pic spectral (SPT) prend en entrée le spectre PPG nettoyé et a pour objectif de choisir correctement le pic correspondant à la fréquence cardiaque (FC). Cette étape repose sur l'observation selon laquelle les valeurs de FC dans deux fenêtres temporelles successives sont proches. La valeur de la FC estimée dans la fenêtre précédente est utilisée pour réduire la plage de recherche du pic correspondant à la FC dans la fenêtre actuelle. Cette partie de l'algorithme permet également de gérer les situations inattendues d'absence de pics liés à la FC.

SPT se compose de 4 étapes :

#### a- Initialisation

Une initialisation correcte est nécessaire pour que l'algorithme puisse suivre efficacement la fréquence cardiaque (FC).

Pendant cette phase d'initialisation, l'utilisateur est censé réduire autant que possible les mouvements de la main pendant plusieurs secondes. Le **kurtosis** du spectre PPG entre 0.8 Hz et 2.5 Hz (correspondant à des fréquences cardiaques de 48 BPM à 150 BPM) est utilisé pour déterminer la qualité du signal. Un kurtosis élevé indique une distribution relativement pointue avec un pic prononcé, tandis qu'un kurtosis faible indique une distribution relativement plate. Un signal PPG avec un spectre de kurtosis élevé présente des pics plus visibles dans le spectre.

- Si la valeur du kurtosis est supérieure à 10, la qualité du signal sur la fenêtre temporelle actuelle est jugée suffisante, une initialisation peut être faite pour le suivi de la fréquence cardiaque (FC) en prenant le pic le plus dominant dans le spectre entre 0.8 Hz et 2.5 Hz et l'algorithme passe alors à l'étape suivante dans la fenêtre temporelle ultérieure.
- Si la valeur du kurtosis est inférieure à 10, cela indique que les mouvements de la main ne sont pas assez réduits, et que la qualité du signal ne permet pas une initialisation dans la fenêtre temporelle actuelle. Dans ce cas, aucune estimation de FC n'est produite. L'algorithme continue de vérifier si une initialisation peut être faite dans la fenêtre temporelle suivante.

### b- Sélection du pic

Une fois l'initialisation faite, la sélection du pic correspondant à la FC pour chaque fenêtre se fait avec la connaissance des valeurs FC estimées dans les fenêtres temporelles précédentes.

Dans ce qui suit, la valeur de la FC dans la fenêtre antérieure est notée **pervBPM**, et correspond dans le spectre à une fréquence d'indice  $i$  tel que  $i = 1, \dots, N$ , cet emplacement de fréquence est noté **prevLoc**. On note aussi **curLoc** et **curBPM**, respectivement, l'emplacement de fréquence et la valeur en BPM du pic spectral sélectionné pour la FC sur la fenêtre actuelle.

- Tout d'abord, une plage de recherche  $R1$  est définie, avec  $\Delta_1$  un entier positif, tel que :  $R1 = [\text{prevLoc} - \Delta_1, \text{prevLoc} + \Delta_1]$ . On cherche au plus les **3 plus grands pics** spectraux. S'il y a des pics présents dans cette plage, l'emplacement de fréquence correspondant au pic le plus **proche** de prevLoc est choisi comme l'indice de fréquence de la valeur FC sur la fenêtre actuelle, curLoc.

Lorsqu'on cherche à sélectionner le pic le plus proche de prevLoc, il peut arriver que deux pics soient à une distance égale de prevLoc (l'un avec une fréquence inférieure et l'autre avec une fréquence supérieure à prevLoc). Afin de déterminer lequel de ces pics doit être choisi, il est nécessaire de prédire si la valeur actuelle de la FC augmente ou diminue par rapport à la valeur précédemment estimée. Pour ce faire, l'algorithme Smoother [34] est utilisé.

Cet algorithme de lissage permet d'éliminer le bruit ou les fluctuations d'une série de données. Une courbe lissée est ajustée aux données à l'aide d'une méthode de moindres carrés sous contrainte. Le degré de lissage est contrôlé par un paramètre  $\lambda$ , qui détermine le compromis entre la fidélité à la série d'origine et le degré de lissage dans la courbe finale. Une valeur  $\lambda$  plus élevée donnera une courbe plus lisse, tandis qu'une valeur  $\lambda$  plus faible permettra plus de détails et de variations dans la courbe finale.

Ici l'algorithme Smoother est utilisé pour révéler la tendance de la série FC sur les  $H$  valeurs précédemment estimées ( $H \geq 10$ ). Le paramètre de lissage,  $\lambda$ , est réglé sur une petite valeur afin de prédire le changement local des valeurs de FC. Si la tendance est croissante, la valeur curLoc renvoyée est celle de la fréquence supérieure à prevLoc, sinon c'est la fréquence inférieure qui est retenue.

- Si aucun pic n'est présent dans  $R1$ , la plage de recherche est étendue à  $R2 = [\text{prevLoc} - \Delta_2, \text{prevLoc} + \Delta_2]$  avec  $\Delta_2 > \Delta_1$ . Si des pics sont trouvés, le pic **maximal** est sélectionné. Si aucun pic n'est trouvé, l'emplacement de fréquence de la valeur FC précédente, prevLoc est choisi comme emplacement de la FC actuelle, curLoc.

L'organigramme de cette phase de SPT est illustré dans la figure suivante :

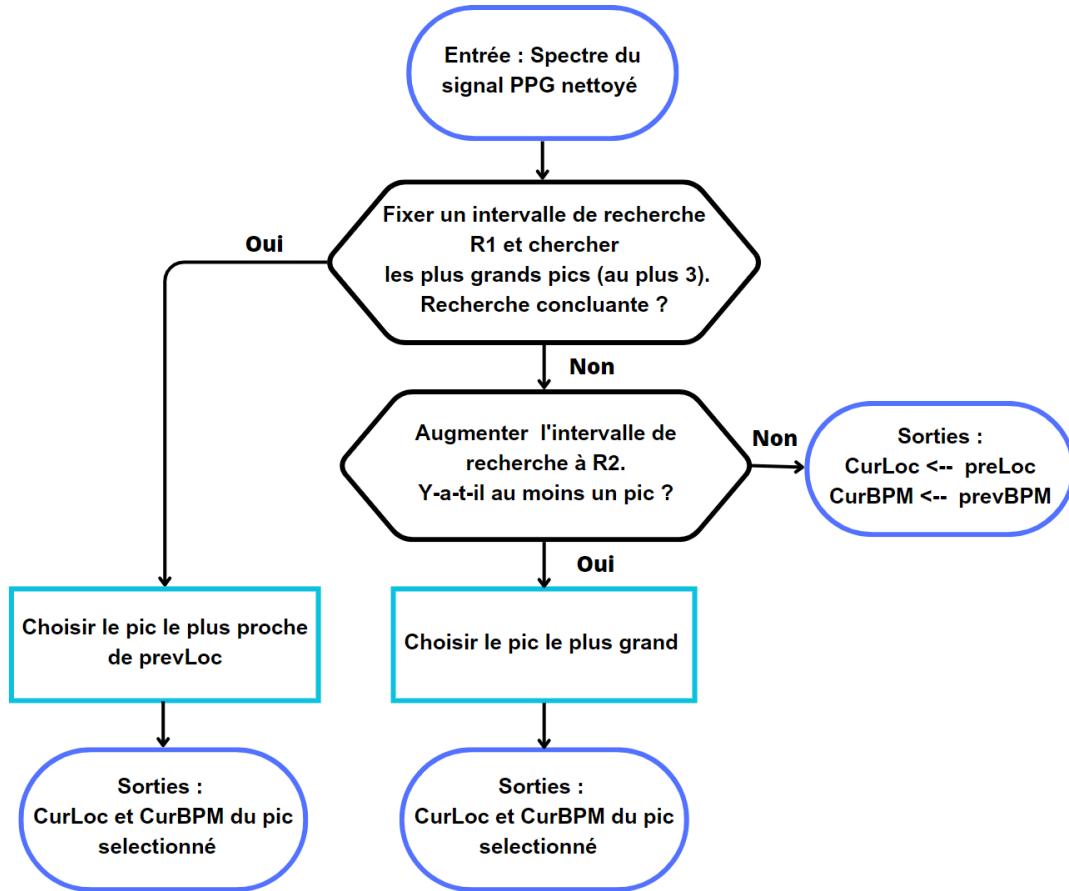


FIG. 2.3 : Organigramme de la phase de sélection du pic de la méthode JOSS.

### c- Vérification du pic

L'étape de sélection du pic peut renvoyer un pic spectral associé à un artefact de mouvement (MA). Ainsi, une étape de vérification est nécessaire.

La vérification repose sur l'observation selon laquelle le changement de la FC entre deux fenêtres temporelles successives ne dépasse généralement pas 12 BPM. Si la différence entre la valeur estimée pour la fenêtre actuelle, curBPM, et la valeur précédente, prevBPM, est supérieure à **12 BPM**, il est très probable que le pic spectral sélectionné soit incorrect. Dans ce cas, les valeurs de curLoc et curBPM sont réinitialisées respectivement à prevLoc et prevBPM. Le nombre de fois où la FC précédente a été choisie comme la FC actuelle est noté **Trap\_count**. Si Trap\_count est trop grand ( $> 3$ ), il y a une forte probabilité qu'en gardant la FC constante dans plusieurs fenêtres temporelles successives, le pic spectral cible a été perdu et se retrouve en dehors de la plage de recherche. Un mécanisme de découverte est déclenché pour redécouvrir l'emplacement du pic FC. Ceci est illustré dans la figure suivante :

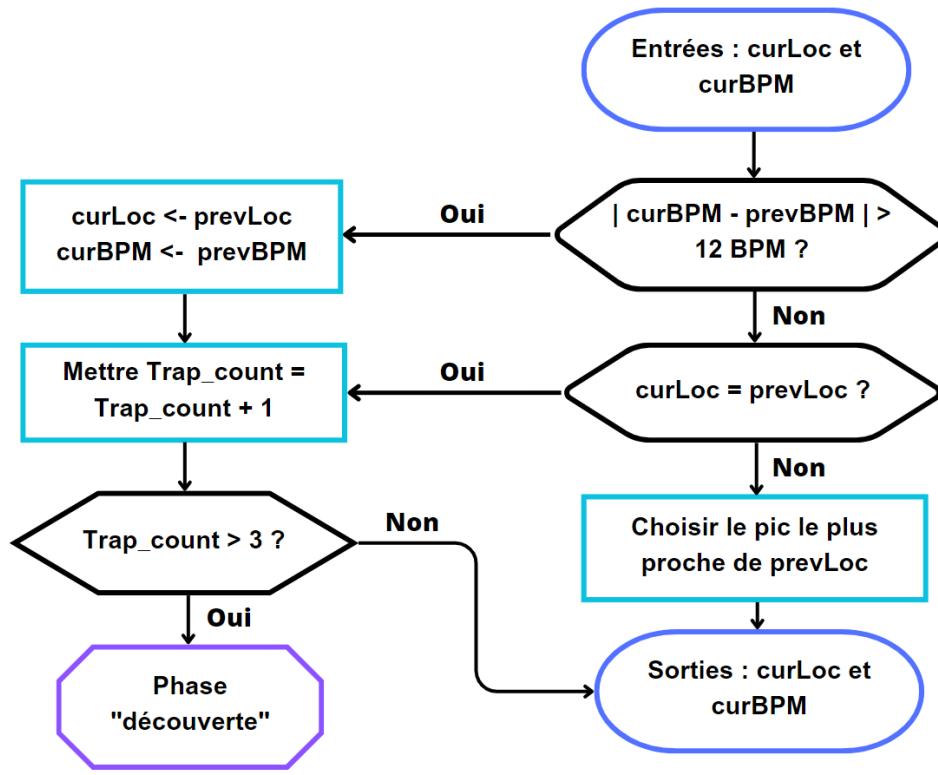


FIG. 2.4 : Organigramme de la phase de vérification du pic de la méthode JOSS.

#### d- Découverte du pic

Afin de découvrir l'emplacement spectral du pic FC perdu dans la fenêtre de temps actuelle, une prédiction est réalisée.

L'algorithme Smoother est appliqué aux emplacements spectraux des **K** précédentes valeurs estimées de FC.  $K$  doit être plus grand que  $H$  et le paramètre de lissage,  $\lambda$ , est fixé à une valeur élevée, étant donné que le but ici est de prédire une tendance globale de l'évolution de la FC. Une fois avoir obtenu une série lissée ajustée aux données, une interpolation linéaire renvoie une prédiction pour la valeur de l'emplacement spectral de la FC dans la fenêtre temporelle actuelle, notée **predictLoc**. Dans une plage de recherche assez grande ( $\Delta_3 > \Delta_2$ ) :  $R3 = [ \text{prevLoc} - \Delta_3, \text{prevLoc} + \Delta_3 ]$ , l'emplacement du pic maximal dans le spectre est renvoyé comme valeur pour curLoc, si aucun pic n'est trouvé, c'est la valeur predictLoc qui est retournée. Cette phase est illustrée par la figure suivante :

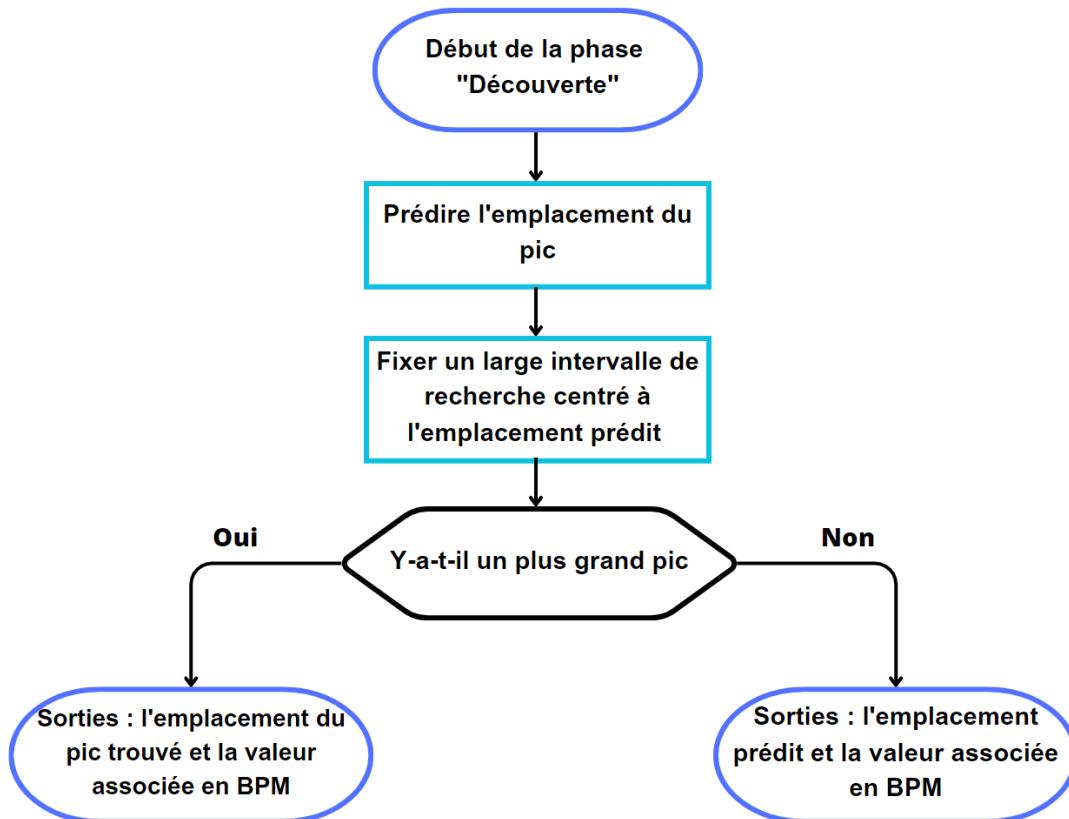


FIG. 2.5 : Organigramme de la phase de découverte du pic de la méthode JOSS.

### 2.1.2 Paramètres de la méthode

L'approche proposée par JOSS implique un nombre important de paramètres à régler, au moins 15 au total. C'est l'étape du suivi du pic spectral avec ses nombreuses règles et seuils heuristiques qui comporte le plus de paramètres. Les recommandations pour le réglage des paramètres de l'algorithme M-FOCUSS régularisé ont été discutées précédemment.

Étape	Sous-étape	Paramètre	Signification	Valeur dans l'article [4]
Estimation des spectres	M-FOCUSS régularisé	N	nombre de points des spectres	1024
		p	norme choisie pour la contrainte de parcimonie	0.8
		$\lambda$	paramètre de régularisation	$10^{-10}$
		max_iters	nombre maximal d'itérations	4
Soustraction spectrale		$\alpha_{th}$	tous les pics $<\alpha_{th} \times$ amplitude du pic maximal du spectre nettoyé sont mis à zéro	1/4
Suivi du pic spectral	Initialisation	kurt	seuil minimal acceptable de kurtosis du spectre permettant une initialisation	10
	Sélection du pic	$\Delta_1$	définie la 1ère plage de recherche d'emplacement de pic R1	15
		H	le nombre de FC précédemment estimées considérées par l'algorithme de lissage	10
		$\lambda_1$	le paramètre de régularisation de l'algorithme de lissage	5
		$\Delta_2$	définie la 2ème plage de recherche d'emplacement de pic R2	25
	Vérification du pic	diff_BPM	différence maximale tolérée entre la FC actuelle estimée et la précédente en BPM	12
		Trap_count	le nombre maximal toléré pour garder la FC précédente comme estimation actuelle	3
	Découverte du pic	K	le nombre de FC précédentes considérées pour la prédiction de l'emplacement du pic perdu	30
		$\lambda_2$	le paramètre de régularisation de l'algorithme de lissage, ici utilisé pour la prédiction	20
		$\Delta_3$	définie la plage de recherche R3 d'emplacement de pic centrée sur la valeur prédite	30

TAB. 2.1 : Paramètres réglables par l'utilisateur dans la méthode JOSS.

## 2.2 Méthode Wiener Filtering Phase Vocoder (WFPV)

La méthode WFPV a été proposée en 2017 par A.Temko [5]. Elle se base sur l'utilisation du filtre de Wiener et du vocodeur de phase. Le filtre de Wiener est utilisé pour atténuer les composantes de bruit présentes dans le signal PPG en utilisant l'estimation de la signature du bruit obtenue à partir des signaux d'accéléromètre. Le vocodeur de phase, quant à lui, est exploité pour surmonter la résolution limitée de la transformée de Fourier discrète et affiner l'estimation initiale de la fréquence dominante. Contrairement aux méthodes proposées précédemment, la méthode WFPV nécessite très peu de paramètres et seuils à régler. La méthode se compose de 4 étapes principales : prétraitement des signaux (PPG et ACC), débruitage du signal PPG afin d'atténuer les artefacts de mouvement, estimation de la fréquence cardiaque (FC) et finalement une étape de post-traitement. L'organigramme de la méthode est illustré dans la figure suivante :



FIG. 2.6 : Organigramme de la méthode WFPV.

### 2.2.1 Méthodologie proposée

#### 2.2.1.1 Prétraitement

Lors de l'étape de prétraitement, les deux signaux PPG et les trois signaux d'accéléromètre présents dans l'ensemble de données IEEE SPC 2015 sont filtrés avec un filtre passe-bande Butterworth d'ordre 4 (0.4 - 4 Hz). Les deux signaux PPG sont normalisés pour avoir une moyenne de zéro et une variance unitaire avant d'être moyennés. Le signal PPG moyen et les 3 signaux d'accéléromètre sont sous-échantillonnés de 125 Hz à 25 Hz et sont ensuite soumis à la transformée de Fourier discrète (DFT) en utilisant 1024 points.

#### 2.2.1.2 Débruitage du signal PPG en utilisant le filtre de Wiener

Le filtre de Wiener est un outil couramment utilisé en traitement du signal afin d'estimer un signal désiré à partir d'un processus bruité observé [35]. En supposant connus les spectres du signal stationnaire et du bruit additif, le filtre de Wiener est conçu pour minimiser l'erreur quadratique moyenne entre le signal d'origine et le signal restauré, en utilisant une fonction de transfert qui est optimisée pour les caractéristiques de la distorsion présente dans le signal. Dans ce cas, le filtrage de Wiener est appliqué pour estimer le signal PPG nettoyé à partir du signal PPG observé.

En supposant que le signal PPG bruité est corrompu par du bruit additif représentant les artefacts de mouvements, nous avons l'équation suivante :

$$X(f) = S(f) + N(f); \quad (2.9)$$

où  $X(f)$ ,  $S(f)$  et  $N(f)$  représentent respectivement le spectre du signal PPG bruité, celui du signal PPG nettoyé et celui du bruit (artefacts de mouvements). L'estimation du signal PPG nettoyé peut donc être obtenue comme suit :

$$\tilde{S}(f) = X(f) - N(f) = \left(1 - \frac{N(f)}{X(f)}\right) X(f) = W(f)X(f) \quad (2.10)$$

Pour un signal observé dans un bruit aléatoire additif et non corrélé, le filtre de Wiener dans le domaine fréquentiel est donné par la formule suivante [35] :

$$W(f) = \frac{P_{XX}(f) - P_{NN}(f)}{P_{XX}(f)} = \frac{P_{SS}(f)}{P_{SS}(f) + P_{NN}(f)} \quad (2.11)$$

où  $P_{SS}(f)$ ,  $P_{XX}(f)$  et  $P_{NN}(f)$  sont respectivement les spectres de puissance du signal nettoyé, du signal observé et du bruit.

La convolution du filtre dans le domaine temporel est équivalente à une multiplication dans le domaine fréquentiel. L'objectif est donc de calculer les coefficients du filtre dans le domaine fréquentiel. Ainsi, le filtre de Wiener agit comme un atténuateur où les fréquences les plus affectées par le bruit sont atténuerées.

Le filtre 2.11 nécessite des estimations séparées des spectres de puissance du signal nettoyé  $P_{SS}(f)$  et de celui du bruit  $P_{NN}(f)$ . Le spectre de puissance du bruit peut être directement estimé à partir des signaux de l'accéléromètre, en faisant la moyenne des spectres des trois signaux d'accéléromètre. Le spectre de puissance du signal PPG nettoyé,  $P_{SS}(f)$  peut, quant à lui, être estimé de deux manières différentes :

1- En soustrayant le bruit du signal observé,  $P_{SS}(f) = P_{XX}(f) - P_{NN}(f)$ , on obtient les coefficients suivants dans le domaine fréquentiel :

$$w_1(t, k) = 1 - \frac{P_{NN}(t, k)}{\frac{1}{c} \sum_{i=t-C+1}^t P_{XX}(i, k)} \quad (2.12)$$

2- De manière récursive en moyennant les résultats antérieurs du filtrage sur C fenêtres  $P_{SS}(f) = \frac{1}{c} \sum_{i=t-C}^{t-1} w_2(i, k)P_{XX}(i, k)$ , on obtient les coefficients suivants dans le domaine fréquentiel :

$$w_2(t, k) = \frac{\frac{1}{C} \sum_{i=t-C}^{t-1} w_2(i, k)P_{XX}(i, k)}{\frac{1}{C} \sum_{i=t-C}^{t-1} w_2(i, k)P_{XX}(i, k) + P_{NN}(t, k)} \quad (2.13)$$

Dans les équations 2.12 et 2.13,  $w(t, k)$  est le poids de la  $k$ -ième fréquence dans la fenêtre temporelle d'indice  $t$ , les spectres de puissance du bruit et du signal PPG sont normalisés en les divisant par leurs valeurs maximales et le spectre de puissance du signal PPG observé est moyenné sur les  $C$  dernières fenêtres  $\frac{1}{C} \sum_{i=t-C+1}^t P_{XX}(i, k)$ .

L'enveloppe spectrale du signal PPG nettoyé est obtenue en multipliant l'enveloppe spectrale du signal observé,  $P_{XX}(f)$ , par les coefficients de filtre calculés  $w(t, k)$ . Dans ce cas, les enveloppes spectrales traitées avec les deux filtres conçus sont mises à l'échelle par leurs écarts types et les signaux résultants sont ensuite moyennés pour donner une enveloppe spectrale finale du signal PPG nettoyé.

La fréquence dominante, c'est-à-dire celle ayant l'amplitude la plus élevée dans le spectre du signal PPG nettoyé, est convertie en estimation de la fréquence cardiaque (en BPM).

Ces étapes sont illustrées dans la figure suivante :

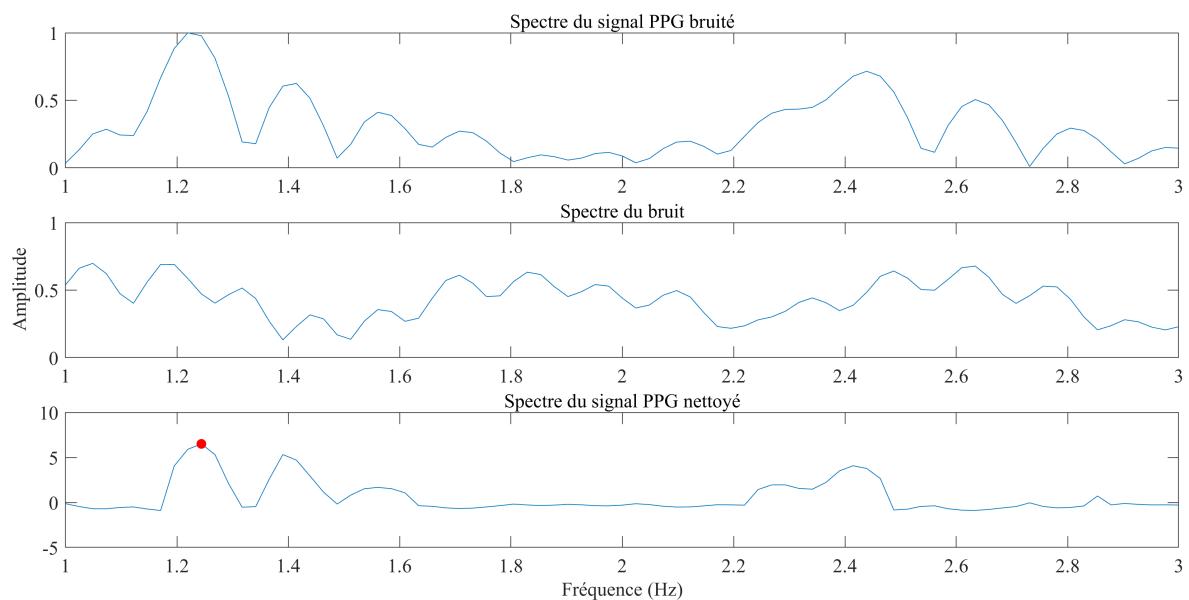


FIG. 2.7 : Débruitage du signal PPG et estimation initiale de la FC pour la méthode WFPV. Le point rouge correspond à l'estimation initiale de la FC.

### 2.2.1.3 Affinement de la fréquence cardiaque estimée en utilisant le vocodeur de phase

Dans la méthode WFPV, la technique du vocodeur de phase [36] est utilisée pour affiner l'estimation initiale de la fréquence cardiaque (fréquence dominante dans le spectre du signal PPG nettoyé).

#### Principe du vocodeur de phase

Le vocodeur de phase est une variation de la transformée de Fourier à court terme (STFT) qui utilise des informations de phase pour améliorer les estimations de fréquence.

La transformée de Fourier à court terme (STFT) est une méthode d'analyse qui commence par découper un signal en courts segments. La transformée de Fourier rapide (FFT) est ensuite appliquée à chaque segment séparément et le spectre résultant de chaque segment peut être manipulé de diverses manières. Les différentes étapes de la (STFT) sont illustrées dans la figure suivante :

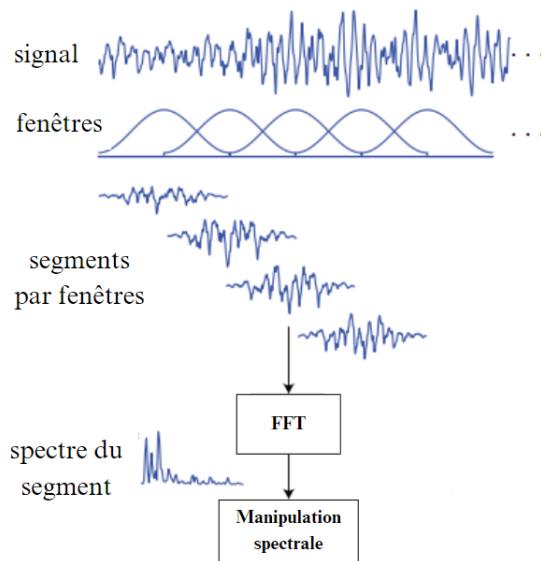


FIG. 2.8 : Étapes de la transformée de Fourier à court terme (STFT).

La résolution en fréquence de la FFT est :

$$\text{résolution en Hz} = (\text{taux d'échantillonnage}) / (\text{taille de la fenêtre})$$

Le vocodeur de phase permet d'améliorer la résolution en fréquence de la STFT sans nuire excessivement à la résolution temporelle. En effet, il permet d'obtenir de meilleures estimations de fréquence en utilisant des informations de phase que la STFT ignore.

Supposons qu'il y ait une sinusoïde de fréquence inconnue mais avec des phases connues : au temps  $t_1$  la sinusoïde a une phase  $\theta_1$  et à  $t_2$  elle a une phase  $\theta_2$ . La sinusoïde peut avoir une fréquence qui la déplace directement de  $\theta_1$  à  $\theta_2$  en un temps  $\Delta t = t_2 - t_1$ , elle peut également commencer à  $\theta_1$ , faire le tour complet du cercle et finir à  $\theta_2$  ou bien elle peut faire deux tours ou même n tours comme illustré dans la figure suivante :

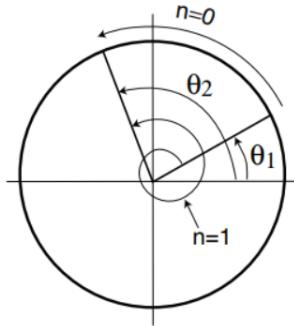


FIG. 2.9 : Illustration du changement de phase d'un signal sinusoïdal.

Ainsi la fréquence doit être :

$$f_n = \frac{(\theta_2 - \theta_1) + 2\pi n}{2\pi\Delta t} \quad (2.14)$$

où  $n$  est un entier positif.

### Utilisation du vocodeur de phase

Remarque préliminaire : la résolution fréquentielle dans ce cas est limitée par la taille de la fenêtre des données analysées (8 secondes) et est égale à  $1/8 = 0.125$  Hz équivalent à 7.5 BPM. Le zéro-padding avant la DFT est utilisé pour interpoler l'enveloppe spectrale à d'autres fréquences améliorant ainsi cette résolution. En effet, 1024 points sont utilisés pour la DFT, cela ne crée pas de nouvelle information, mais permet une meilleure révélation de l'information existante dans le signal.

Après avoir fait, dans l'étape précédente, une première estimation de la fréquence cardiaque notée  $f$  (fréquence dominante dans le spectre du signal PPG nettoyé), les phases  $\theta_2$  et  $\theta_1$  des fenêtres actuelle et précédente sont exploitées pour affiner l'estimation initiale de la fréquence en utilisant la formule suivante :

$$\arg \min_n (\tilde{f}(n) - f); \tilde{f}(n) = \frac{(\theta_2 - \theta_1 + 2\pi n)}{(2\pi(t_2 - t_1))}, \forall n \in \mathbf{N} \quad (2.15)$$

où  $n$  est un entier positif,  $t_2 - t_1 = 2s$  ce qui correspond au décalage entre deux fenêtres. Nous calculons  $\tilde{f}(n)$  pour plusieurs valeurs de  $n$  en utilisant l'équation 2.15 et la valeur de  $n$  pour laquelle  $\tilde{f}(n)$  est la plus proche de l'estimation initiale de la FC est choisie. Par conséquent, la valeur  $f$  est affinée en une nouvelle valeur  $\tilde{f}(n) \leftarrow f$ .

### 2.2.1.4 Post-traitement

Les étapes de post-traitement comprennent le suivi de l'historique des estimations précédentes de la FC et le lissage.

L'historique des estimations précédentes de la FC est utilisé pour guider la recherche de l'amplitude maximale dans le spectre du signal PPG nettoyé dans la fenêtre actuelle. En effet, l'estimation de la FC actuelle est censée être dans une certaine plage autour de l'estimation précédente. La plage de recherche  $\pm\tau_i$  est initialisée avec une amplitude suffisamment grande ( $\pm 25$  BPM) pour les 30 à 60 premières secondes de chaque enregistrement. Ensuite, elle s'adapte aux spécificités de la physiologie de chaque utilisateur en prenant la différence maximale en valeur absolue des estimations consécutives de la fréquence cardiaque,  $f(t)$ , observées jusqu'à présent comme suit :

$$\tau_i = \max\{|f(t) - f(t-1)| : k < t < i-1\} \quad (2.16)$$

Si la différence entre les estimations de fréquence cardiaque actuelle et précédente est supérieure à 5 BPM, une étape finale de lissage est effectuée en faisant la moyenne pondérée entre l'estimation actuelle et sa prédiction à l'aide de la régression linéaire :

$$\tilde{f} = \alpha f + (1 - \alpha)f_{\text{LinReg}} \quad (2.17)$$

où  $\alpha = 0.8$  et  $f_{\text{LinReg}}$  est une prédiction obtenue grâce à une régression linéaire ajustée sur les 6 dernières estimations de la fréquence cardiaque en utilisant une erreur des moindres carrés.

## 2.2.2 Paramètres de la méthode

Comme il a été mentionné précédemment, la méthode WFPV repose sur très peu de paramètres et seuils heuristiques à régler.

- **Dans la partie débruitage et estimation de la FC**, nous avons 2 paramètres :
  - $C$  : le nombre de fenêtres précédentes sur lesquelles sont moyennés les spectres de puissance  $P_{XX}(f)$  et  $P_{SS}(f)$ . Fixé à 15 dans le papier.
  - $N$  : le nombre de points utilisés pour la DFT. Fixé à 1024.
- **Dans la partie post-traitement**, nous avons 4 paramètres :
  - $\tau_1$  : la valeur initiale de la plage de recherche. Fixée à 25 BPM.
  - $Diff_{max}$  : la différence maximale acceptable entre deux valeurs successives de FC. Fixé à 5 BPM.
  - $\alpha$  : paramètre de la somme pondérée. Fixé à 0.8.
  - $N_{regression}$  : nombre de points utilisés pour la régression linéaire. Fixé à 6.

## 2.3 Implémentation et évaluation de la méthode JOSS

La méthode JOSS a démontré dans [4] une erreur absolue moyenne d'estimation de **1.28 BPM** lorsqu'elle a été évaluée sur l'ensemble de données IEEE SPC 2015 - train. Dans cette section, nous proposons d'étendre cette évaluation et d'investiguer les performances de la méthode en termes de robustesse et de temps d'exécution.

### 2.3.1 Implémentation sur MATLAB

L'algorithme JOSS fait partie des travaux fondateurs dans la problématique de l'estimation de la FC à partir de signaux PPG présentant des artefacts de mouvement. Malheureusement, pour JOSS, comme pour la plupart des autres solutions proposées dans la littérature, les implémentations ne sont pas disponibles en ligne, ce qui limite la possibilité de réévaluer et d'améliorer les algorithmes proposés. Par exemple, l'évaluation des méthodes existantes sur de nouveaux jeux de données est une étape importante pour valider leur capacité de généralisation et leur robustesse. La première étape de ce travail a donc été d'élaborer une implémentation de l'algorithme JOSS tel que présenté dans [4]. Par la suite, cette implémentation nous a permis d'évaluer l'approche.

Le code a été développé en MATLAB et a été testé en utilisant l'ensemble de données et les paramètres expérimentaux de l'article de référence [4]. Les résultats obtenus ont ensuite été comparés à ceux présentés dans l'article, comme indiqué dans le tableau 2.2.

#### Remarques préliminaires :

- Pour l'implémentation, une fenêtre temporelle de 8 secondes et un shift de 2 secondes ont été utilisés pour traiter simultanément le signal PPG et les signaux d'accéléromètre. La fréquence cardiaque (FC) a été estimée sur chaque fenêtre temporelle.
- Avant d'être introduits comme entrées de l'algorithme, les deux signaux PPG et les trois signaux d'accéléromètre présents dans les ensembles de données IEEE SPC 2015 (train et test) sont filtrés avec un filtre passe-bande Butterworth d'ordre 4 (0.4 - 4 Hz) pour éliminer les valeurs irréalistes pour une FC humaine. Ensuite, les signaux sont normalisés pour avoir une moyenne de zéro et une variance unitaire et les deux signaux PPG sont moyennés. Le signal PPG moyen et les 3 signaux d'accéléromètre sont enfin sous-échantillonnés de 125 Hz à 25 Hz.

Sujet	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Moyenne (std)
Papier JOSS [4]	1.33	1.75	1.47	1.48	0.69	1.32	0.71	0.56	0.49	3.81	0.78	1.04	1.28 (2.61 )
Notre implémentation de JOSS	1.33	1.88	0.68	1.83	0.91	2.89	0.79	0.98	0.48	2.38	3.76	0.97	1.57 (3.14)

TAB. 2.2 : Comparaison des résultats en MAE (BPM) de notre implémentation de JOSS avec ceux de l'article de référence [4] sur l'ensemble de données IEEE SPC 2015 – train.

Nous pouvons constater que les résultats obtenus avec notre implémentation sont très proches de ceux présentés dans [4]. La différence peut être expliquée par d'éventuels détails sur l'implémentation qui n'ont pas été explicitement donnés dans l'article.

### 2.3.2 Évaluation de la robustesse de la méthode

#### 2.3.2.1 Impact du réglage des paramètres sur les performances

Une caractéristique fondamentale de la méthode JOSS est qu'elle repose sur un nombre très important de paramètres ajustables manuellement, ce qui nous amène à penser que cela pourrait conduire à un manque de généralisation lorsqu'elle est évaluée sur différents ensembles de données sans réglage des paramètres. Nous abordons cette question en évaluant les performances de cette dernière sur l'ensemble de données IEEE SPC 2015 - test. Ce jeu de données a été choisi car il présente des durées d'enregistrements similaires à celles du jeu de données IEEE SPC 2015 - train mais implique l'utilisation de nouvelles données sur lesquelles la méthode n'a pas été ajustée.

Nous avons les résultats suivants :

Session	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Moyenne (std)
Avec paramètres de [4]	32.97	14.65	4.87	106.22	3.74	11.50	3.65	4.83	4.88	0.87	<b>18.82 (13.66)</b>
Avec paramètres optimisés	13.55	2.09	2.12	3.41	2.98	3.22	3.59	6.86	3.53	0.80	<b>4.21 (6.02)</b>

TAB. 2.3 : Résultats en MAE (BPM) de la méthode JOSS avec et sans optimisation des paramètres sur l'ensemble de données IEEE SPC 2015 – test.

La première ligne du tableau présente les erreurs absolues moyennes obtenues sur l'ensemble de données IEEE SPC 2015 – test en utilisant directement les paramètres réglés sur l'ensemble de données IEEE SPC 2015 – train. La deuxième ligne, quant à elle, présente les erreurs obtenues en utilisant la meilleure combinaison de paramètres parmi plusieurs combinaisons testées. Nous pouvons remarquer que la performance a été considérablement améliorée ainsi, passant de **18.82 BPM** sans réglage des paramètres à **4.21 BPM** avec optimisation de ces derniers sur le jeu de données considéré.

À partir de ces résultats, nous pouvons confirmer notre hypothèse selon laquelle la méthode JOSS présente une faible capacité de généralisation lorsqu'elle est utilisée directement avec les valeurs des paramètres fournies dans l'article de référence [4] (sans adaptation au nouvel ensemble de données). En effet, JOSS repose sur un nombre élevé de paramètres qui entraînent un surajustement (overfitting) aux données sur lesquelles ils sont réglés. Bien qu'il soit possible d'obtenir de bons résultats en les **ajustant** sur un ensemble de données de test spécifique, il est important de souligner que cela n'est pas pratique dans des situations réelles, car cela nécessiterait de le faire chaque fois que de nouvelles données sont traitées.

Les figures 2.10 et 2.11 illustrent respectivement l'estimation pour le sujet ayant obtenu la meilleure performance (sujet 8 correspondant à la session 10) ainsi que pour celui ayant obtenu la moins bonne performance (sujet 1 correspondant à la session 1) de l'ensemble de données IEEE SPC 2015 – test.

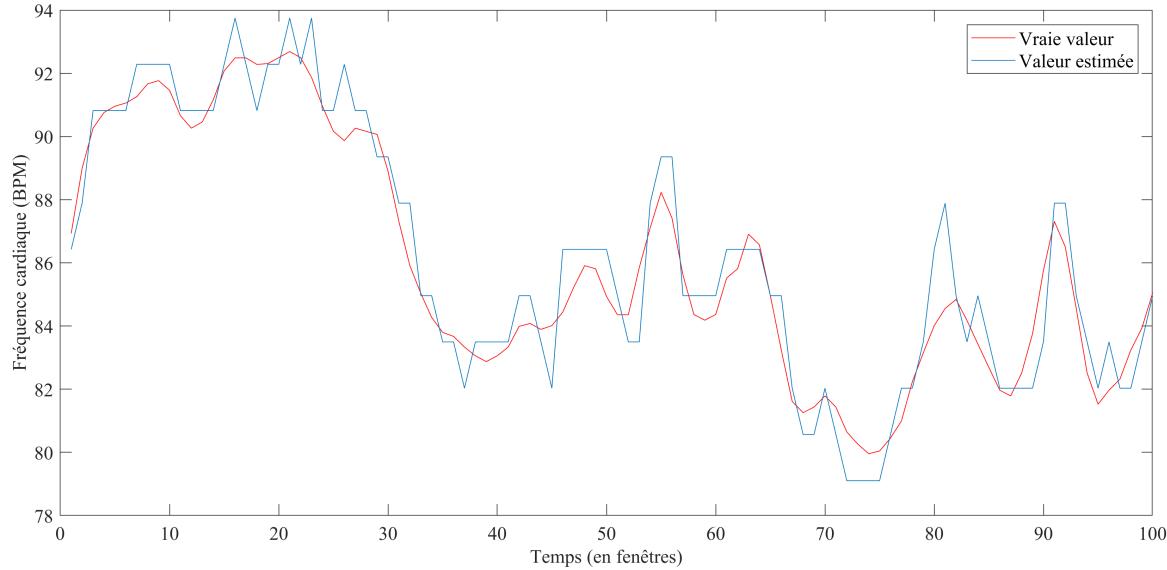


FIG. 2.10 : Estimation faite sur le meilleur sujet (Sujet 8) de l'ensemble de données IEEE SPC 2015 – test avec réglage des paramètres de la méthode JOSS.

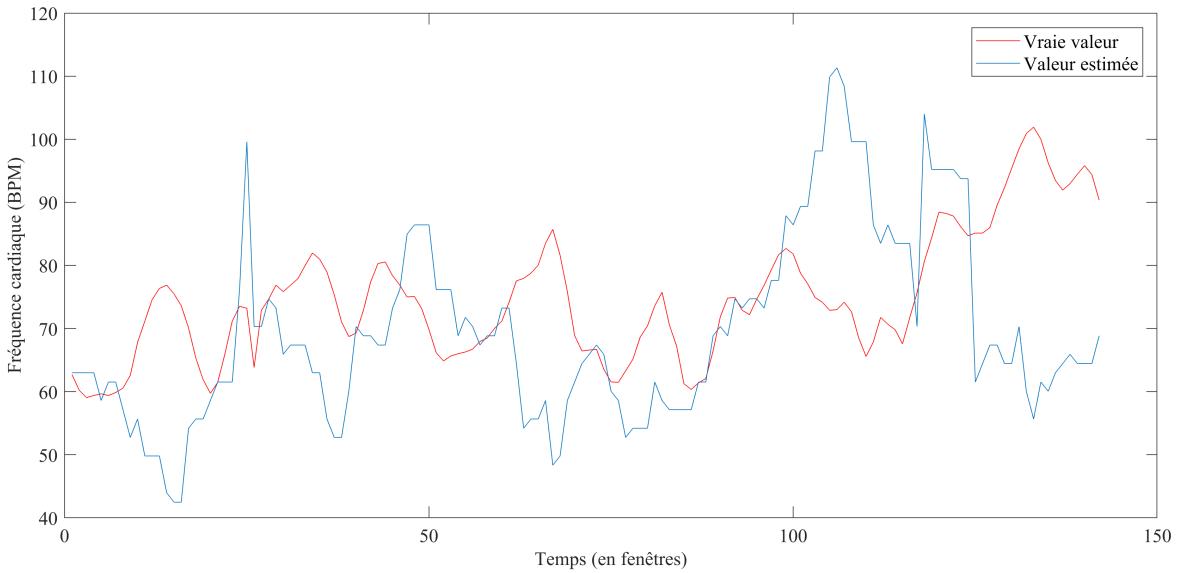


FIG. 2.11 : Estimation faite sur le pire sujet (Sujet 1) de l'ensemble de données IEEE SPC 2015 – test avec réglage des paramètres de la méthode JOSS.

### 2.3.2.2 Capacité de généralisation à de nouvelles activités

Le jeu de données IEEE SPC 2015 - train utilisé lors de l'évaluation dans l'article de référence [4] présente plusieurs limitations qui ont été discutées dans 1.3.4. Le jeu de données est restreint, avec moins d'une heure d'enregistrement au total. Il se compose d'un seul type d'activité, à savoir la course sur tapis, et les données ont été collectées dans des conditions contrôlées, dans un environnement de type laboratoire utilisant une machine d'exercice.

L'implémentation de la méthode nous permet, dans cette section, d'évaluer sa capacité de généralisation à de nouvelles activités. Pour ce faire, nous utilisons le jeu de données PPG-DaLiA proposé en 2019 [22]. Cet ensemble de données comprend au total plus de 36 heures d'enregistrements sur 15 sujets. Elle couvre 8 types différents d'activités (typiquement effectuées dans la vie quotidienne). Les sujets ont effectué ces activités dans un environnement non contrôlé. Tous ces éléments font de PPG-DaLiA un choix particulièrement adapté pour évaluer la robustesse de la méthode.

Les résultats obtenus par JOSS sur l'ensemble de données PPG-DaLiA sont présentés dans la figure 2.12.



FIG. 2.12 : Résultats en MAE (BPM) de JOSS par sujet et par activité sur PPG-DaLiA.

Dans la Figure 2.12, la présence de “–” indique que l'algorithme n'a pas pu estimer la fréquence cardiaque (FC) sur l'intégralité de la durée de l'activité. Cela signifie que l'algorithme a pu estimer la FC sur une partie de l'activité, mais qu'il a divergé à un certain moment.

L'algorithme JOSS, comme de nombreux autres algorithmes, est très dépendant des estimations précédentes. Ainsi, lorsqu'une grande erreur est commise lors d'une estimation sur une fenêtre, cela peut se répercuter sur toutes les fenêtres suivantes. L'algorithme, tel qu'il est présenté, est censé prendre en compte cette limitation. En effet, le processus "Découverte" est normalement prévu pour ça, il intervient lorsqu'une certaine estimation de la FC se répète trop de fois ce qui pourrait indiquer que l'algorithme s'éloigne de la vraie valeur. Néanmoins, le processus "Découverte" peut lui-même contribuer à la divergence de l'algorithme. Si ce dernier est trop souvent sollicité, cela amène à un enchaînement d'estimations totalement aberrantes (l'algorithme diverge). Ce problème peut intervenir même pour une activité dite simple si la durée d'enregistrement est trop longue. Il peut également intervenir si l'activité est courte mais qu'elle présente des artefacts de mouvements trop forts.

- Nous pouvons donc remarquer d'après la figure précédente que l'algorithme n'arrive pas à estimer la FC sur plusieurs activités et ce sur la plupart des sujets. Il s'agit surtout des activités présentant les durées d'enregistrements les plus longues : déjeuner (30 min), travailler (20 min), s'asseoir (10 min). Les activités pour lesquelles l'algorithme arrive à faire l'estimation sur toutes les fenêtres d'enregistrement pour la plupart des sujets sont celles qui ont une durée similaire à celle des enregistrements de l'ensemble de données IEEE SPC 2015 - train : babyfoot (5 min), vélo (8 min).
- Les activités qui sont les moins sujettes aux mouvements brusques (s'asseoir, déjeuner) ont les erreurs les plus faibles sur les sujets pour qui l'estimation a pu se faire. Les activités intenses physiquement (FC moyenne élevée sur l'activité, tableau 1.4) présentent des erreurs très élevées (babyfoot et escaliers).

### 2.3.3 Évaluation du temps d'exécution

L'estimation de la FC sur les 12 enregistrements de l'ensemble de données IEEE SPC 2015 - train prend **754.9 secondes** sur MATLAB R2017b avec un processeur Intel Core i7-8550U @ 1.80GHz, soit en moyenne une minute par session d'enregistrement de 5 minutes. Ce temps d'exécution peut être considéré comme élevé dans ce contexte et est en grande partie attribué à l'algorithme M-FOCUSS régularisé.

Le temps d'exécution obtenu sur MATLAB ne donne qu'une idée approximative du temps d'exécution réel, car celui-ci dépend entièrement du dispositif cible utilisé. Différentes configurations matérielles peuvent avoir un impact significatif sur les performances.

### 2.3.4 Conclusions de l'évaluation

Pour conclure, l'évaluation de la méthode JOSS a montré qu'elle présentait des limitations. Elle dépend de nombreux paramètres et seuils heuristiques, ce qui limite considérablement sa capacité de généralisation si ces paramètres ne sont pas ajustés spécifiquement pour chaque jeu de données. De plus, ses performances sont sensibles à la durée des en-

registrements et à la présence d'importants artefacts de mouvements, ce qui peut parfois entraîner une divergence de l'algorithme. Enfin, elle présente un temps d'exécution élevé.

## 2.4 Implémentation et évaluation de la méthode WFPV

L'objectif de cette partie est d'évaluer la méthode WFPV en termes de robustesse et de temps d'exécution. La robustesse impliquant l'impact du réglage des paramètres sur les performances et la capacité de généralisation à des données présentant des artefacts de mouvement différents.

Nous présentons tout d'abord les résultats du papier de référence [5] obtenus sur les jeux de données IEEE SPC 2015 (train et test) :

**Remarque :** dans le cadre de l'utilisation des ensembles de données IEEE SPC 2015, les deux signaux PPG ont été exploités et la plage de fréquences cardiaques d'intérêt a été fixée de 60 à 180 BPM, correspondant à 1 et 3 Hz respectivement. Cet intervalle a été fixé selon la plage de fréquences cardiaques présentes dans le jeu de données utilisé dans l'étude.

- Sur l'ensemble de données IEEE SPC 2015 – train, la méthode a obtenu les résultats suivants :

Sujet	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Moyenne (std)
MAE (BPM)	1.25	1.41	0.71	0.97	0.75	0.92	0.65	0.97	0.55	2.06	1.03	0.99	1.02 (1.25)

TAB. 2.4 : Résultats en MAE (BPM) de la méthode WFPV sur l'ensemble de données IEEE SPC 2015 – train.

- Sur l'ensemble de données IEEE SPC 2015 – test, la méthode a obtenu les résultats suivants :

Session	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Moyenne (std)
MAE (BPM)	9.59	2.57	2.25	3.01	2.73	1.57	2.10	3.44	1.61	0.75	2.96 (3.71)

TAB. 2.5 : Résultats en MAE (BPM) de la méthode WFPV sur l'ensemble de données IEEE SPC 2015 – test.

Nous pouvons déjà constater que l'erreur dans l'estimation de la fréquence cardiaque dépend du type d'activité physique. En effet, dans le cas de l'ensemble de données IEEE SPC 2015 – train, il a été demandé aux sujets de courir sur un tapis de course, cette activité est caractérisée par le fait qu'elle sollicite très peu l'utilisation du bras et de la main (emplacement du capteur PPG), ce qui minimise les artefacts de mouvement, et qu'elle implique un mouvement périodique facilement identifiable dans les spectres des signaux PPG et d'accéléromètre. Cela résulte dans une erreur absolue moyenne de **1.02 BPM**, ce qui peut être considéré comme une très bonne performance. D'autre part, pour l'ensemble de données IEEE SPC 2015 – test, il a été demandé aux sujets de mener des activités sollicitant l'avant-bras et la partie supérieure du bras (par exemple, se serrer la main, s'étirer, courir, sauter et faire des pompes) ce qui a impliqué des artefacts de mouvement plus forts et en conséquence une erreur absolue moyenne de **2.96 BPM**, plus élevée que la première erreur. Cette erreur peut tout de même être elle aussi considérée comme une bonne performance.

### 2.4.1 Implémentation sur MATLAB

Afin d'étendre l'évaluation de la méthode WFPV, nous utilisons l'ensemble de données PPG-DaLiA, proposée en 2019 [22], qui a l'avantage d'avoir des enregistrements de plus longue durée et qui inclut des activités diversifiées impliquant différents niveaux d'intensité des artefacts de mouvement comme cela a été expliqué plus tôt dans la partie 2.3.2.2. Pour ce faire, nous avons adapté le code publié par l'auteur du papier [37] pour pouvoir l'utiliser sur ce jeu de données.

**Remarque :** Dans le cadre de l'utilisation de l'ensemble de données PPG-DaLiA, **un seul** signal PPG a été exploité et la plage de fréquences cardiaques d'intérêt a été fixée de 39 à 192 BPM, correspondant à 0.65 et 3.2 Hz respectivement. Cet intervalle a été fixé selon la plage de fréquences cardiaques présentes dans l'ensemble de données.

### 2.4.2 Évaluation de la robustesse de la méthode

#### 2.4.2.1 Impact du réglage des paramètres sur les performances

Nous avons dans un premier temps gardé les mêmes valeurs des paramètres proposées dans l'article de référence [5] :

- $C = 15$
- $N = 1024$
- $\tau_1 = 25 \text{ BPM}$
- $Diff_{\max} = 5 \text{ BPM}$
- $\alpha = 0.8$
- $N_{\text{regression}} = 6$

Nous avons obtenu les résultats suivants :

Sujet	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	Moyenne (std)
MAE (BPM)	9.18	9.39	10.23	15.41	31.08	15.31	4.70	12.17	20.43	9.38	21.96	13.40	7.48	9.02	11.16	13.35 (18.84)

TAB. 2.6 : Résultats en MAE (BPM) de la méthode WFPV sur l'ensemble de données PPG-DaLiA.

Nous avons par la suite essayé différentes combinaisons de valeurs de paramètres, nous avons essayé les valeurs suivantes pour chaque paramètre :

- $C = 5, 10, 13, 15, 17, 20, 30, 40$
- $N = 256, 512, 1024, 2048$
- $\tau_1 = 15, 25, 30 \text{ BPM}$
- $Diff_{\max} = 5, 10, 15, 25 \text{ BPM}$
- $\alpha = 0.6, 0.7, 0.8, 0.9$
- $N_{\text{regression}} = 5, 6, 7$

Les différentes combinaisons essayées ont toutes obtenu relativement les mêmes résultats que la combinaison du papier. La plus grande erreur obtenue a été de 15.39 BPM. Les paramètres étant fixés à  $C = 5$ ,  $N = 256$ ,  $\tau_1 = 25 \text{ BPM}$ ,  $Diff_{\max} = 25 \text{ BPM}$ ,  $\alpha = 0.6$ ,  $N_{\text{regression}} = 7$

Cette erreur peut être justifiée par différents éléments : le nombre de fenêtres précédentes sur lesquelles sont moyennés les spectres de puissances  $P_{XX}(f)$  et  $P_{SS}(f)$  ( $C = 5$ ) n'est pas suffisant, le nombre de points utilisés pour la DFT ( $N = 256$ ) est aussi insuffisant pour avoir une bonne résolution (rappelons que plus le nombre de points est élevé, plus la résolution est grande), le paramètre de la somme pondérée ( $\alpha = 0.6$ ) n'est pas adapté car il donne trop d'importance à la prédiction faite par la régression linéaire.

L'erreur la plus faible a été obtenue en utilisant la combinaison de valeurs proposée dans le papier.

Nous pouvons donc conclure que les performances de la méthode WFPV dépendent très peu du réglage des paramètres, cela peut être expliqué par leur nombre très réduit, ce qui évite le risque d'un surajustement sur les jeux de données utilisées dans l'article de référence (IEEE SPC 2015). Il est important de noter que ceci est un avantage majeur par rapport aux méthodes proposées antérieurement car cette propriété est essentielle pour des applications réelles, l'objectif étant de proposer une méthode implantable sur des appareils portables pouvant se généraliser à des données sur lesquelles elle n'a pas été ajustée.

Pour ce qui est des méthodes précédemment proposées, les connaissances extraites (en termes de règles et de valeurs seuils) ne sont valables que sur les données utilisées. Augmenter le nombre de degrés de liberté d'une méthode entraîne généralement une meilleure performance, mais cela se fait au prix d'un risque accru de mauvaise généralisation sur des données inconnues, surtout si l'ajustement est effectué en post-traitement pour corriger les erreurs de l'algorithme principal.

#### 2.4.2.2 Capacité de généralisation à de nouvelles activités

Un autre aspect très important à considérer lors de l'évaluation de la robustesse est la capacité de généralisation de la méthode à de nouvelles données collectées lors de différents types d'activités présentant des artefacts de mouvement plus ou moins prononcés. L'ensemble de données PPG-DaLiA [22] nous offre l'opportunité d'effectuer cette évaluation.

La figure 2.13 permet de visualiser l'erreur obtenue par sujet et par activité et l'erreur moyenne par activité.

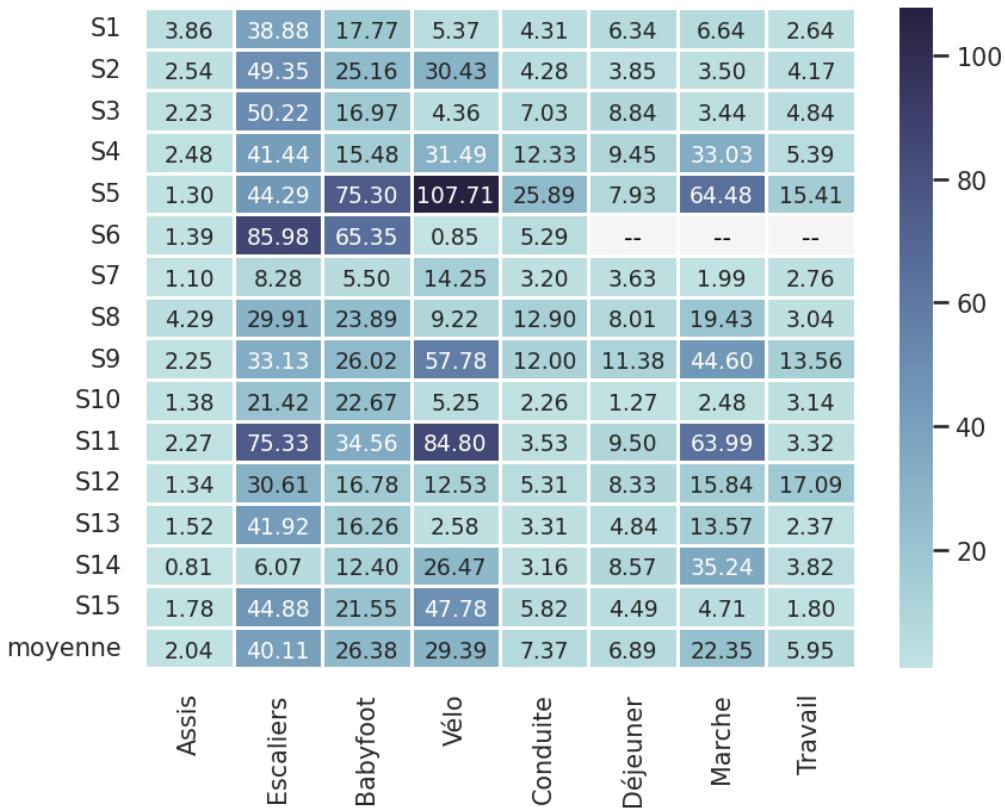


FIG. 2.13 : Résultats en MAE (BPM) de la méthode WFPV par sujet et par activité pour l'ensemble de données PPG-DaLiA.

A partir de ces résultats, nous pouvons confirmer l'observation faite sur les deux ensembles de données IEEE SPC 2015 (train et test) selon laquelle l'erreur dans l'estimation de la fréquence cardiaque dépend du type d'activité physique. En effet, de manière générale, nous constatons que pour des activités simples ne nécessitant pas beaucoup d'efforts physiques (FC moyenne faible, tableau 1.4) et donc sans grands artefacts de mouvement, comme s'asseoir, conduire, déjeuner et travailler, l'estimation de la FC est plutôt précise. Bien que les erreurs ne soient pas aussi faibles que celles observées sur les ensembles de données IEEE SPC 2015, cela peut s'expliquer par la durée plus longue des enregistrements et les conditions réelles de collecte des données.

D'autre part, pour les activités plus complexes exigeant davantage d'efforts physiques (FC moyenne plus élevée, tableau 1.4) et impliquant donc plus d'artefacts de mouvement, comme jouer au babyfoot, faire du vélo, monter des escaliers, les erreurs deviennent significativement plus grandes, rendant la méthode totalement inutilisable.

Il y a une corrélation claire entre FC moyenne sur l'activité (tableau 1.4) et erreur absolue moyenne obtenue.

Les figures 2.15 et 2.14 montrent les résultats obtenus pour le meilleur sujet (sujet 7) et le pire sujet (sujet 5) de l'ensemble de données PPG-DaLiA. Ce sont les mêmes sujets qui ont obtenu respectivement la meilleure et la pire erreur lorsqu'ils ont été testés avec d'autres méthodes.

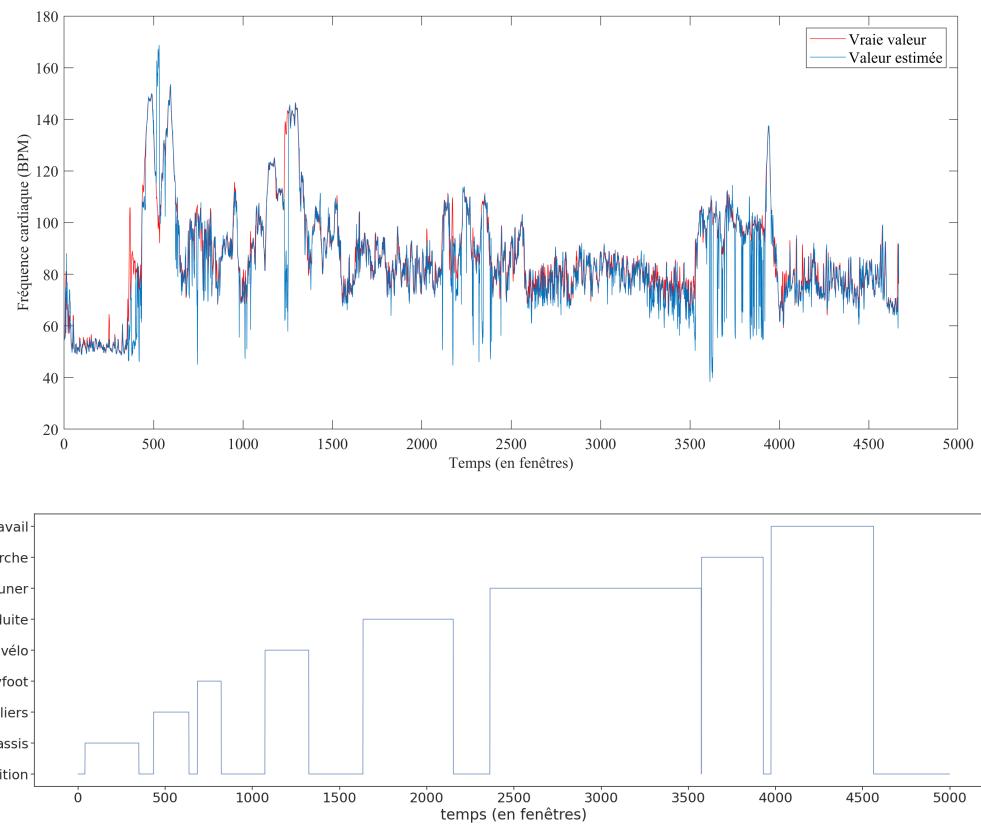


FIG. 2.14 : Estimation faite sur le meilleur sujet (S7) de PPG-DaLiA avec WFPV.

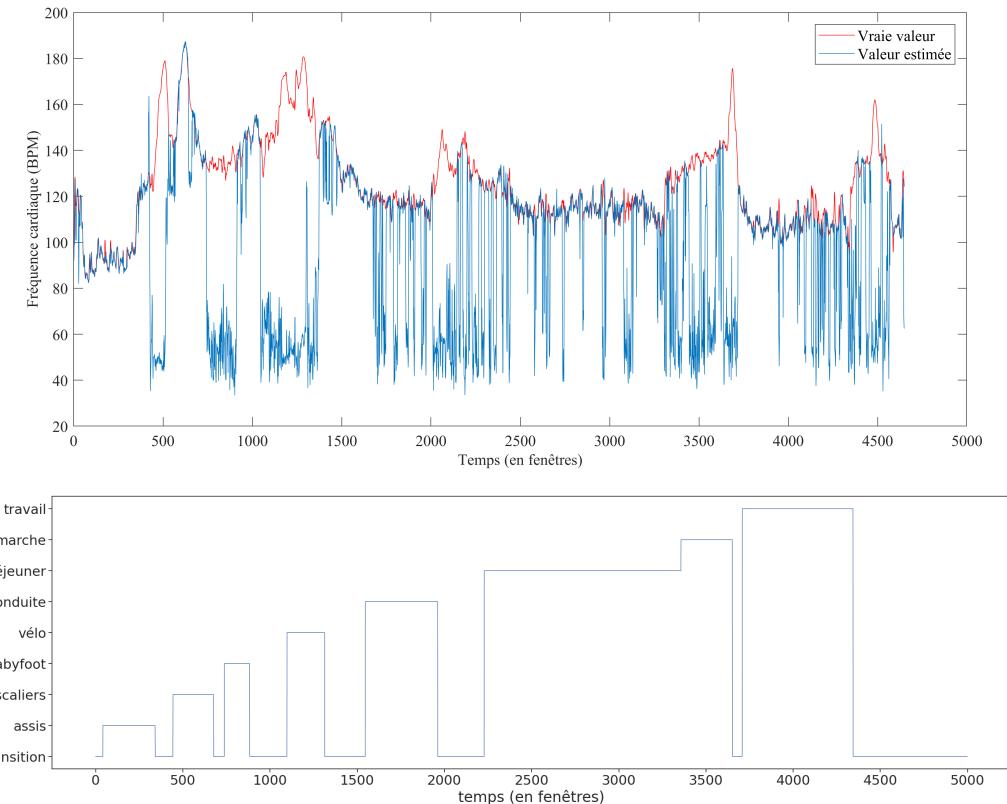


FIG. 2.15 : Estimation faite sur le pire sujet (S5) de PPG-DaLiA avec WFPV.

### 2.4.3 Évaluation du temps d'exécution

Un autre avantage considérable de la méthode WFPV par rapport aux autres méthodes est son temps d'exécution très réduit. Il lui faut 1.676 secondes sur Matlab 2017b avec un processeur Intel Core i7-8550U @ 1.80GHz pour traiter **tout** l'ensemble de données IEEE SPC 2015 - train alors que JOSS, à titre d'exemple, prend 754.9 secondes pour traiter le même jeu de données.

Ainsi, l'algorithme a un temps d'exécution faible, bien adapté pour le suivi de la fréquence cardiaque sur des appareils portables.

### 2.4.4 Conclusion de l'évaluation

Pour conclure, la méthode WFPV présente plusieurs avantages, tel que son nombre très réduit de paramètres, ce qui évite d'avoir à les régler manuellement pour chaque ensemble de données. De plus, elle présente un faible temps d'exécution et offre de bonnes performances pour des activités simples. Cependant, la méthode se révèle totalement inadaptée pour des activités physiques intenses comportant de forts artefacts de mouvement. Tous ces points nous amènent à envisager l'utilisation de WFPV uniquement dans le cas d'activités simples.

## Conclusion

Dans ce chapitre, nous avons présenté l'évaluation de deux méthodes classiques, JOSS et WFPV, suite à leur implémentation sur MATLAB. Nous avons pu identifier les avantages ainsi que les limitations de chaque méthode. D'une part, nous avons mis en lumière le manque de généralisation des deux méthodes à de nouvelles activités plus longues et complexes. D'autre part, nous avons démontré que la méthode JOSS est fortement dépendante du réglage de ses nombreux paramètres et seuils heuristiques, ce qui impacte ses performances, contrairement à la méthode WFPV qui maintient de bonnes performances indépendamment du réglage des paramètres. Nous avons finalement obtenu une idée sur les temps d'exécution de chaque méthode. WFPV se distingue par sa rapidité d'exécution, tandis que JOSS présente un temps d'exécution significativement plus long.

# **Chapitre 3**

## **Approche par apprentissage profond**

Dans ce chapitre, nous explorons l'approche par apprentissage profond. Cette dernière se distingue de la plupart des méthodes de l'approche classique par sa capacité de généralisation et sa plus grande applicabilité aux données réelles. Cependant, elle reste relativement moins explorée du fait qu'elle est moins adaptée au déploiement sur des appareils portables à ressources limitées. Récemment, A. Burrello et al. ont proposé la méthode Q-PPG [3], qui présente la première méthodologie pour optimiser automatiquement les modèles d'apprentissage profond pour le suivi de la fréquence cardiaque à partir de signaux PPG fortement affectés par des artefacts de mouvement. Nous commençons par présenter les notions nécessaires à la compréhension de cette méthode, puis nous menons plusieurs expérimentations visant à investiguer ses forces et ses limitations.

## 3.1 Notions préliminaires

Dans cette section, nous abordons les notions essentielles à la compréhension de la méthode Q-PPG [3]. Certaines généralités ont été omises afin de favoriser la clarté de l'exposé.

### 3.1.1 Les Réseaux de Neurones Convolutifs (CNN)

Les réseaux de neurones convolutifs (CNN), également connus sous le nom de ConvNet, sont une classe de réseaux de neurones artificiels profonds, à propagation avant (non récurrents), qui sont utilisés pour analyser des données ayant une topologie en grille connue. Cela inclut les données 1D comme les séries temporelles, les données 2D comme les images et les données 3D comme les vidéos. Le concept moderne des réseaux de neurones convolutifs provient des travaux de Yann LeCun et al., publiés en 1998 [38]. Depuis lors, les CNN ont démontré une très bonne performance dans des tâches complexes de vision par ordinateur [39] et de traitement du langage naturel [40].

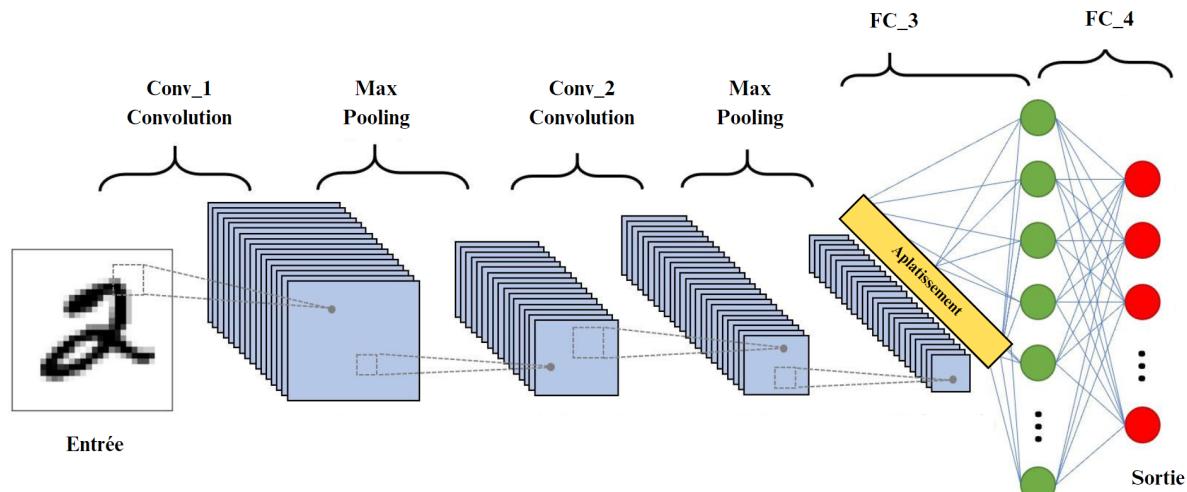


FIG. 3.1 : Illustration de l'arrangement des couches dans un CNN.

Les CNN sont composés de deux principaux blocs. Le premier bloc prend une image en entrée et extrait de manière hiérarchique ses caractéristiques. Le second bloc reçoit le vecteur hiérarchique produit par le premier et le fait passer à travers un réseau de neurones entièrement connectés pour produire les prédictions finales. La figure 3.1 montre une illustration de l'arrangement des couches dans un CNN.

→ Afin d'expliquer les notions propres aux CNN, nous considérons dans cette sous-section l'exemple des CNN-2D.

## Bloc d'extraction de caractéristiques

Ce premier bloc agit comme un extracteur automatique de caractéristiques. Il se compose de plusieurs parties décrites dans ce qui suit :

### Les couches de convolution

Les couches de convolution sont la composante clé des réseaux de neurones convolutifs. Elles sont composées de deux parties, une partie linéaire (opération de convolution) suivie d'une partie non linéaire (fonction d'activation).

En pratique, une couche de convolution notée  $L$  reçoit en entrée une image de dimensions  $w_L \times x_L \times I_L$  (où  $w_L$  est la largeur de l'image,  $x_L$  sa hauteur et  $I_L$  le nombre de canaux).  $O_L$  convolutions sont appliquées sur cette entrée,  $O_L$  étant le nombre de filtres constituant la couche de convolution. Chaque convolution est effectuée avec un filtre de dimensions  $f_L \times g_L \times O_L$  (où  $f_L$  est la largeur du filtre,  $g_L$  sa hauteur et  $O_L$  le nombre de canaux), généralement  $f_L = g_L = f$  et est appelé taille du filtre [41]. Une illustration de l'étape de convolution est donnée dans la figure 3.2

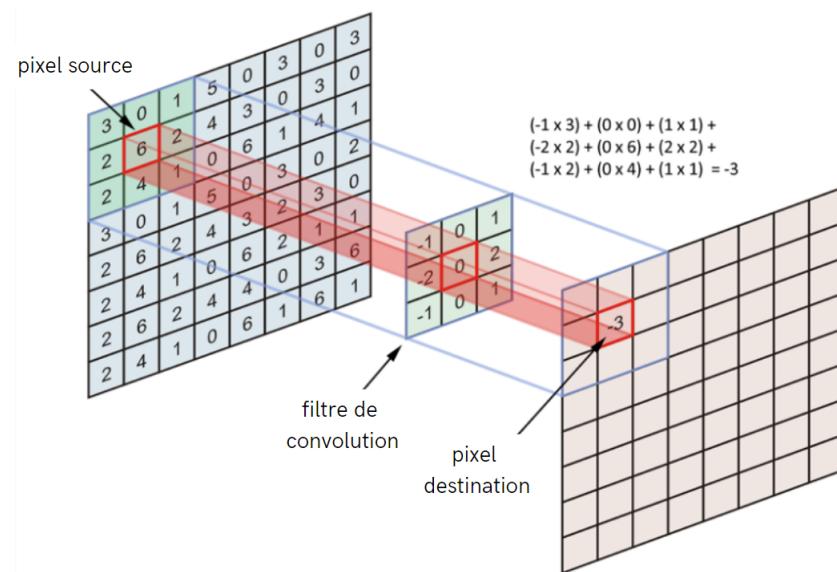


FIG. 3.2 : Illustration de l'étape de convolution.

Outre le nombre et la taille des filtres, la couche de convolution possède également deux autres hyperparamètres :

- Le stride  $s$  : il spécifie la distance de déplacement du filtre après chaque opération de convolution.
- Le padding  $p$  : il détermine le nombre de zéros ajoutés autour de l'image. Le padding simple garantit que la sortie de la convolution a la même dimension que l'entrée.

A la sortie de la couche de convolution  $L$ , une carte de caractéristiques “feature map” est produite, elle est de dimensions  $y_L \times z_L \times O_L$  où  $y_L = \lfloor \frac{w_L - f + 2p}{s} \rfloor + 1$ ,  $z_L = \lfloor \frac{x_L - f + 2p}{s} \rfloor + 1$ . [38]

En plus des transformations linéaires, des transformations non linéaires, appelées fonctions d’activation, sont utilisées. Elles sont appliquées élément par élément sur la sortie de chaque opération de convolution. Il existe différentes fonctions d’activation, nous pouvons citer : la fonction ReLU, la fonction sigmoïde, la fonction tangente hyperbolique, etc. [42]

### Les couches de Pooling

Ce type de couches est fréquemment utilisé entre deux couches de convolution. La couche de pooling prend en entrée plusieurs cartes de caractéristiques et applique l’opération de pooling à chacune d’entre elles. L’opération de pooling a pour objectif de réduire la taille des images tout en préservant les caractéristiques essentielles. D’abord, l’image est subdivisée en cellules régulières, puis on peut choisir de conserver soit la valeur maximale de chaque cellule, soit la moyenne des valeurs. En effet, il existe deux principaux types de Pooling : le max-Pooling et le average-Pooling. Le max-Pooling renvoie la valeur maximale de la partie de l’image couverte par le filtre. L’average-Pooling renvoie, quant à lui, la moyenne de toutes les valeurs de la partie de l’image couverte par le filtre [43]. Ces deux types sont illustrés dans la figure 3.3

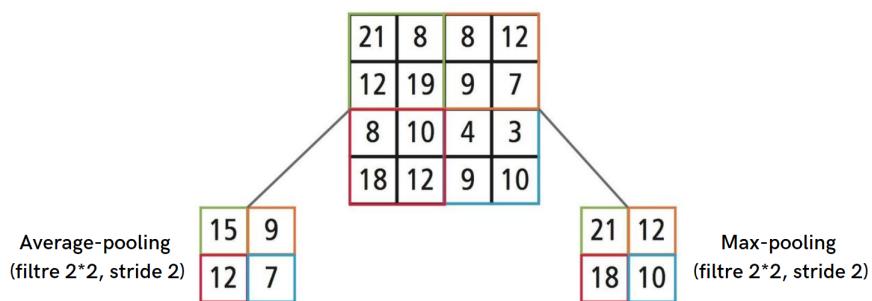


FIG. 3.3 : Illustration du Average-pooling et du Max-Pooling.

### L’aplatissement (flattening)

L’aplatissement est la dernière étape du bloc d’extraction de caractéristiques, où les lignes des matrices sont concaténées pour former un vecteur unidimensionnel. Ce vecteur peut ensuite être connecté au bloc de prédiction.

## Bloc de prédiction

Ce bloc consiste essentiellement en un Réseau de Neurones Simple Entièrement Connecté (couches FC), composé de deux ou trois couches cachées et une couche de sortie. On peut le considérer comme la phase finale d'apprentissage, qui relie les caractéristiques extraites par le bloc précédent aux sorties souhaitées. Les couches entièrement connectées sont utilisées pour créer des combinaisons non linéaires finales des caractéristiques et pour effectuer des prédictions par le réseau.

### 3.1.2 Normalisation par lots (Batch Normalization)

La normalisation par lots (Batch-Normalization ou BN) a été introduite pour la première fois par S.Ioffe et al. [44]. C'est une méthode algorithmique qui accélère et stabilise l'entraînement des réseaux de neurones profonds. Elle consiste à normaliser les vecteurs d'activation des couches cachées en utilisant la moyenne et la variance du lot courant (les données d'entraînement sont divisées en petits lots). Cette étape de normalisation est appliquée juste avant (ou juste après) la fonction non linéaire.

La normalisation par lots est effectuée différemment lors de la phase d'entraînement et celle de test.

#### Phase d'entraînement

À chaque couche cachée, la normalisation par lots transforme les vecteurs d'activation de la manière suivante :

$B$  désigne le mini-lot courant de taille  $m$ . La moyenne et la variance sur  $B$  peuvent donc être obtenues comme suit :

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i, \quad (3.1)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2, \quad (3.2)$$

pour une couche du réseau avec un vecteur d'activation de dimension  $d$ ,  $x = (x^{(1)}, \dots, x^{(d)})$ . Chaque dimension est ensuite normalisée séparément :

$$\hat{x}_i^{(k)} = \frac{x_i^{(k)} - \mu_B^{(k)}}{\sqrt{\left(\sigma_B^{(k)}\right)^2 + \epsilon}}, \quad (3.3)$$

où  $k \in [1, d]$  et  $i \in [1, m]$ ;  $\mu_B^{(k)}$  et  $\sigma_B^{(k)}$  sont la moyenne et la variance pour chaque dimension respectivement.  $\epsilon$  est ajouté au dénominateur pour assurer la stabilité numérique et est une constante arbitrairement petite. L'activation normalisée résultante  $\hat{x}(k)$  a une moyenne nulle et une variance unitaire si  $\epsilon$  n'est pas pris en compte.

Une étape de transformation est ensuite effectuée :

$$y_i^{(k)} = \gamma^{(k)} \hat{x}_i^{(k)} + \beta^{(k)}, \quad (3.4)$$

où  $\gamma^{(k)}$  et  $\beta^{(k)}$  sont appris avec les autres paramètres du réseau. Cette étape permet au modèle de choisir la distribution optimale pour chaque couche cachée.  $\gamma^{(k)}$  permet d'ajuster l'écart-type ;  $\beta^{(k)}$  permet d'ajuster le biais [45].

### Phase de test

Lors de la phase de test, les statistiques utilisées pour la normalisation ne sont pas calculées sur le lot de données de test individuel. Au lieu de cela, les statistiques globales de la population ( $\mu_{pop}$  et  $\sigma_{pop}$ ) sont calculées en utilisant toutes les ( $\mu_{batch}$ ,  $\sigma_{batch}$ ) déterminées pendant l'entraînement et sont directement utilisées dans l'équation 3.3 lors de l'évaluation (au lieu d'appeler les équations 3.1 et 3.2). Cela permet de maintenir la cohérence avec la normalisation effectuée pendant l'entraînement et assure que le modèle est évalué de manière cohérente [45].

#### 3.1.3 Les Réseaux de Convolution Temporelle (TCN)

Les Réseaux de Convolution Temporelle (TCN) sont une sous-classe de CNN-1D spécialisés dans le traitement de séries temporelles. Récemment, il a été démontré que les TCN surpassent les Réseaux de Neurones Récursifs sur plusieurs tâches, en obtenant une précision plus élevée pour le même nombre de paramètres. La particularité des TCN par rapport aux CNN unidimensionnels standards réside dans l'utilisation de convolutions **causales** et **dilatées** [46].

**Causalité** : constraint la sortie de la convolution  $y_t$  à dépendre uniquement des entrées  $x_{\tilde{t}}$  avec  $\tilde{t} \leq t$ . Si un zéro-padding est appliqué sur le côté gauche du vecteur d'entrée, cela assure une convolution causale. Le principe est illustré dans la figure 3.4.

→ Zero-padding signifie l'ajout de zéros.

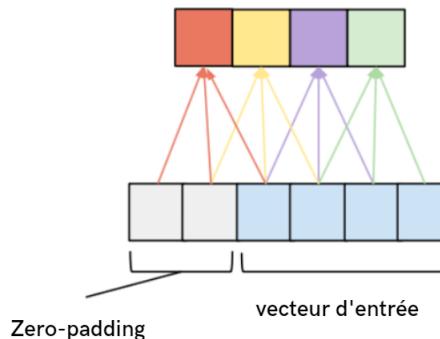


FIG. 3.4 : Illustration du principe de causalité.

**Dilatation :** insère un écart fixe  $d$  entre les pas de temps d'entrée traités par les filtres de convolution. La convolution dilatée permet d'augmenter le champ réceptif des filtres sur l'axe temporel sans augmenter le nombre de paramètres entraînables. (ce point est expliqué plus en détails par la suite). Le principe est illustré dans 3.5 pour différentes valeurs de  $d$ .

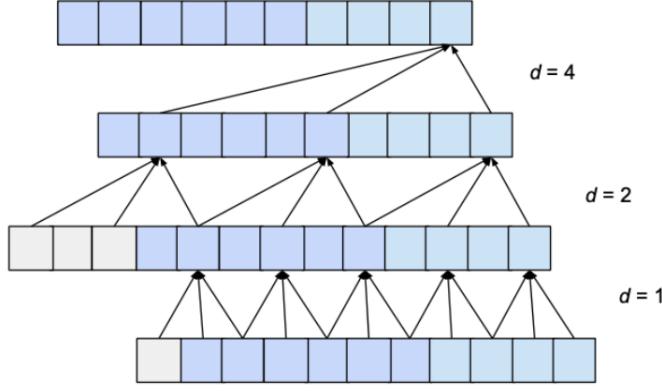


FIG. 3.5 : Illustration du principe de dilatation pour différentes valeurs de  $d$ .

La fonction implémentée par une couche de convolution d'un TCN est :

$$\mathbf{y}_t^m = \sum_{i=0}^{K-1} \sum_{l=0}^{C_{in}-1} \mathbf{x}_{ts-di}^l \cdot \mathbf{W}_i^{l,m}, \quad (3.5)$$

qui est répétée pour tout  $m \in [0, C_{out}-1]$  et pour tout  $t \in [0, T-1]$ . Dans la formule,  $\mathbf{x} \in \mathbb{R}^{C_{in} \times T}$  et  $\mathbf{y} \in \mathbb{R}^{C_{out} \times T/s}$  sont les activations d'entrée et de sortie, respectivement composées de  $C_{in}$  et  $C_{out}$  canaux,  $T$  est la longueur de la sortie sur l'axe temporel,  $\mathbf{W} \in \mathbb{R}^{C_{in} \times C_{out} \times K}$  est la matrice multidimensionnelle des poids des filtres,  $d$  est le facteur de dilatation,  $s$  est le stride, et  $K$  est la taille du filtre.

→ Par la suite, les notations de l'équation 3.5 sont utilisées pour toutes les notions impliquant les TCNs.

### 3.1.4 Outils de recherche d'architectures neuronales (NAS)

Les outils de recherche d'architectures neuronales (Neural Architecture Search, NAS) sont utilisés pour concevoir des architectures de réseaux de neurones pour une tâche donnée en optimisant automatiquement les hyperparamètres tels que les canaux de couches, les dimensions de filtres, etc. qui sont habituellement ajustés manuellement [26]. Le NAS explore automatiquement un vaste espace de conception de configurations d'hyperparamètres possibles, co-optimisant ainsi la performance du réseau résultant avec son coût, mesuré en termes d'empreinte mémoire ou d'opérations en virgule flottante par inférence (FLOPs). Récemment, des approches NAS légères appelées DmaskingNAS ont été proposées. Ces dernières commencent par une architecture de réseau de départ et effectuent un processus d'élagage (pruning), c'est-à-dire la suppression d'éléments du réseau tout

en minimisant l'impact sur sa performance ce qui permet de réduire sa consommation en ressources. Dans cette partie, nous expliquons deux de ces méthodes, utilisées dans la méthode Q-PPG [3], qui sont MorphNet [26] qui propose un élagage au niveau des canaux et Pruning in time (PIT) [27] qui cible l'optimisation des facteurs de dilatation  $d$  dans les TCNs.

#### 3.1.4.1 MorphNet

MorphNet est un outil de recherche d'architectures neuronales (NAS) proposé en 2018 par A.Gordon et al. [26]. Dans cette approche, le réseau est entraîné avec une contrainte qui annule certains canaux de sortie dans chaque couche. La contrainte ajoutée à la fonction de perte (avec un facteur de pondération  $\lambda_{MorphNet}$ ) représente la contribution de chaque canal en termes de paramètres ou d'opérations en virgule flottante (FLOPS) dans le réseau.

#### Remarque :

Dans ce qui suit, les notations des CNN-2D sont utilisées. Les couches du réseau sont notées de  $L = 1$  à  $M + 1$ . Chaque couche  $L$  correspond à une couche de convolution ou une couche entièrement connectée et possède une largeur d'entrée  $I_L$  et une largeur de sortie  $O_L$ . Dans le cas d'une couche de convolution,  $I_L$  et  $O_L$  correspondent respectivement au nombre de canaux d'entrée et de sortie. Pour simplifier la notation, les couches entièrement connectées sont considérées comme un cas particulier de la convolution où les canaux correspondent aux neurones. Ainsi, pour chaque couche  $L = 1, \dots, M + 1$ , sont associées des dimensions spatiales d'entrée  $w_L \times x_L$ , des dimensions spatiales de sortie  $y_L \times z_L$  et des dimensions de filtre  $f_L \times g_L$ . La matrice de poids associée à la couche  $L$  a donc des dimensions  $I_L \times O_L \times f_L \times g_L$  et elle mappe une entrée de dimensions  $w_L \times x_L \times I_L$  à une sortie de dimensions  $y_L \times z_L \times O_L$ . L'ensemble des paramètres à apprendre du réseau est noté  $\theta$  et  $\mathcal{L}(\theta)$  la fonction de perte mesurant la performance de ce dernier.

### Formulation de la solution

MorphNet limite la recherche d'architecture à l'optimisation des largeurs de sortie  $O_{1:M}$  de toutes les couches. Ainsi, la solution part d'une architecture initiale du réseau  $O_{1:M}^o$ , qui comprend non seulement un ensemble initial de largeurs de sortie, mais aussi les dimensions des filtres, la topologie du réseau et d'autres choix d'architecture qui sont considérés comme fixes et cherche uniquement à déterminer le nombre optimal de canaux par couche  $O_{1:M}$ , qui assurerait le compromis entre performance du réseau et contrainte de ressources. Limiter l'optimisation uniquement aux largeurs des couches (nombre de canaux de sortie) a démontré son efficacité tout en restant simple. La contrainte de ressources est notée  $\mathcal{F}(O_{1:M}) \leq \zeta$ .

### Contraintes

Deux types de contraintes sont considérées dans [26] : la taille du modèle (nombre de paramètres) et le nombre d'opérations en virgule flottante par inférence (FLOPs).

Dans le cas d'une optimisation sur la taille du modèle, la contrainte est exprimée comme suit :

$$C = f_L \times g_L \quad (3.6)$$

et dans le cas d'une contrainte sur les FLOPs, nous avons :

$$C = 2 \times y_L \times z_l \times f_L \times g_L \quad (3.7)$$

La fonction de contrainte sur une couche L donnée est donc :

$$\mathcal{F}(\text{couche } L) = C \times I_L \times O_L \quad (3.8)$$

En appliquant une régularisation dans le but de minimiser la contrainte, les neurones ou les canaux qui contribuent le plus à la contrainte seront encouragés à être mis à zéro. Ces neurones ou canaux doivent être exclus de l'équation précédente car un réseau équivalent peut être construit sans les poids menant ou sortant de ces neurones /canaux. Pour refléter cela, l'équation est réécrite comme suit :

$$\mathcal{F}(\text{couche } L) = C \sum_{i=0}^{I_L-1} A_{L,i} \sum_{j=0}^{O_L-1} B_{L,j}, \quad (3.9)$$

où  $A_{L,i}$  ( $B_{L,j}$ ) vaut 1 si l'i-ème entrée (j-ème sortie) de la couche L est active, c'est-à-dire non mise à zéro.

La fonction contrainte totale est obtenue en additionnant l'équation précédente sur toutes les couches du réseau :

$$\mathcal{F}(O_{1:M}) = \sum_{L=1}^{M+1} \mathcal{F}(\text{couche } L) \quad (3.10)$$

## Régularisation

Un terme de régularisation est ajouté à la fonction de perte qui optimise la performance du réseau. Ceci est explicité dans 3.11 :

$$\min_{\theta} \mathcal{L}(\theta) + \lambda_{MorphNet} \mathcal{F}(O_{1:M}) \quad (3.11)$$

Pour rendre l'apprentissage réalisable par descente de gradient, il est nécessaire de remplacer la norme  $\ell_0$  discontinue qui apparaît dans  $\mathcal{F}$  par une norme de substitution continue. Dans MorphNet, c'est la norme  $\ell_1$  sur les variables  $\gamma_L$  de la normalisation par lots (batch normalization) [44] qui est utilisée. Beaucoup de modèles récents appliquent une normalisation par lots à chaque couche, cela en fait une solution simple pour intégrer la régularisation. Chaque neurone ou canal de sortie a un  $\gamma$  particulier qui détermine sa mise en échelle. Il suffit d'éliminer les neurones ou canaux correspondants aux  $\gamma$  mis à zéros.

Le terme de régularisation  $\mathcal{F}$  devient donc :

$$\mathcal{G}(\theta, \text{ couche } L) = C \sum_{i=0}^{I_L-1} |\gamma_{L-1,i}| \sum_{j=0}^{O_L-1} B_{L,j} + C \sum_{i=0}^{I_L-1} A_{L,i} \sum_{j=0}^{O_L-1} |\gamma_{L,j}| \quad (3.12)$$

Le terme de régularisation pour tout le réseau est donné par :

$$\mathcal{G}(\theta) = \sum_{L=1}^{M+1} \mathcal{G}(\theta, \text{ couche } L) \quad (3.13)$$

Et le problème d'apprentissage devient :

$$\theta^* = \operatorname{argmin}_{\theta} \{\mathcal{L}(\theta) + \lambda_{MorphNet} \mathcal{G}(\theta)\} \quad (3.14)$$

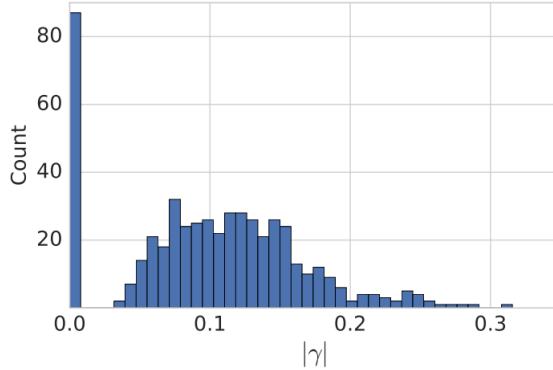


FIG. 3.6 : Histogramme des valeurs de  $\gamma$  pour l'une des couches de ResNet101 entraîné avec une régularisation de type FLOPs. Certains des  $|\gamma|$  ont été mis à zéro et sont séparés clairement des  $|\gamma|$  non nuls [26].

Cette régularisation est très efficace pour induire de la parcimonie dans les valeurs de  $\gamma$  du réseau, comme le montre l'exemple de la figure 3.6. Il existe une séparation claire entre les  $\gamma$  qui ont été annulés et ceux qui continuent de contribuer à la performance du réseau. Il suffit d'éliminer les neurones ou canaux correspondants aux  $\gamma$  mis à zéros pour avoir un nouvel ensemble de largeurs de sortie  $O'_{1:M}$  qui représente un compromis entre l'optimisation de la performance donnée par  $\mathcal{L}$  et la satisfaction de la contrainte donnée par  $\mathcal{G}$ . L'importance relative entre performance et minimisation de la contrainte de ressources peut être choisie en réglant le paramètre de régularisation  $\lambda_{MorphNet}$ .

## Algorithm

L'algorithme MorphNet se résume comme suit :

---

### Algorithm 1 : L'algorithme MorphNet

---

- 1 : Entraîner le réseau pour trouver  $\theta^* = \operatorname{argmin}_{\theta} \{\mathcal{L}(\theta) + \lambda_{MorphNet} \mathcal{G}(\theta)\}$ .
  - 2 : Trouver les nouvelles largeurs  $O'_{1:M}$  induites par  $\theta^*$ .
  - 3 : Retourner  $O'_{1:M}$ .
-

Le réglage du paramètre de régularisation  $\lambda_{MorphNet}$  se fait manuellement. Il est possible de déterminer si une valeur est trop grande ou trop petite très tôt dans l'entraînement en observant si la contrainte s'effondre à zéro ou ne diminue pas du tout.

#### 3.1.4.2 Pruning In Time (PIT) :

La plupart des méthodes NAS se sont concentrées sur l'optimisation des hyperparamètres propres aux CNN standards, tels que le nombre de canaux dans chaque couche et la taille des filtres. Dans le cas des TCN, le facteur de dilatation  $d$  est fondamental pour la précision du réseau, comme cela est détaillé dans [46], car il contrôle la relation entre le champ réceptif dans le temps (c'est-à-dire la plage d'échantillons couverte par chaque étape de convolution) et la taille des filtres (qui détermine la complexité du modèle). Le réglage manuel des facteurs de dilatation peut être très fastidieux et chronophage.

La méthode PIT, de M. Risso et al., publiée en 2022 [27], a été la première à proposer un NAS ciblant l'optimisation automatique des facteurs de dilatation  $d$  dans les TCN. L'algorithme part d'un réseau initial avec des filtres de taille maximale et un facteur de dilatation  $d = 1$  dans toutes les couches et apprend simultanément les poids du TCN et les facteurs de dilatation optimaux pour chaque couche, en fonction d'une contrainte de coût. Dans le cas de PIT, la taille du modèle a été considérée comme la contrainte cible.

#### Formulation de la solution :

Afin de pouvoir optimiser les facteurs de dilatation lors de l'entraînement, il est crucial de les **intégrer** dans un terme **différenciable** ajouté à la fonction de perte. A l'instar des autres méthodes DmaskingNAS, l'approche repose sur l'utilisation de paramètres supplémentaires entraînables, désignés par le terme  $\gamma$ .

Chaque convolution temporelle dans un TCN est caractérisée par une certaine taille de filtre  $k$  et un certain facteur de dilatation  $d$ . Ensemble, ces deux paramètres définissent le champ réceptif  $rf$  de la couche :

$$rf = (k - 1).d + 1 \quad (3.15)$$

PIT limite l'espace de recherche à des solutions où le facteur de dilatation  $d$  est régulier et exprimé sous forme de puissances de 2. Cela permet de conserver un degré de liberté élevé pour augmenter ou réduire le champ réceptif tout en restreignant l'espace des solutions et en permettant une formalisation simple expliquée par la suite.

Pour chaque convolution temporelle, PIT commence par définir un vecteur de paramètres binaires  $\gamma$ , contenant  $L = \lfloor \log_2(rf_{\max} - 1) \rfloor + 1$  éléments, où  $rf_{\max}$  est le  $rf$  maximal pris en charge. En particulier,  $\gamma_0$  est constant et toujours égal à 1 et est ajouté simplement pour simplifier la notation mathématique.

En ce qui concerne les  $\gamma_{1:L-1}$  restants, ils sont les principaux paramètres qui contrôlent la dilatation. Une binarisation est appliquée aux  $\gamma$  en utilisant une approche similaire à celle de BinaryConnect [47].

Pour limiter l'espace de recherche à des facteurs de dilatation réguliers, les paramètres entraînables de  $\gamma$  sont combinés de la manière représentée dans la partie rouge de la Figure 3.7. Les éléments de  $\gamma$  sont multipliés pour former un nouvel ensemble de  $\Gamma$ , qui seront les valeurs utilisées pour effectuer le masquage :

$$\Gamma_i = \prod_{k=0}^{L-1-i} \gamma_k \quad (3.16)$$

Dans la Figure 3.7,  $\gamma_0$  a été remplacé directement par 1. De manière intuitive, les paramètres  $\Gamma$  agissent comme des sélecteurs qui permettent d'activer ou de désactiver des tranches de temps entières du filtre de convolution. Ainsi, ils codent des schémas réguliers de dilatation basés sur des puissances de deux.

Comme le montre la figure,  $d = 1$  est atteint uniquement lorsque tous les paramètres entraînables sont égaux à 1, c'est-à-dire lorsque  $\Gamma_0 = \gamma_0 \cdot \gamma_1 \cdots \gamma_{L-1} = 1$ . Pour avoir des facteurs de dilatation plus grands, à chaque fois que  $d$  est doublé, il faut supprimer la condition sur un  $\gamma_i$ , donc pour  $d = 2$ ,  $\Gamma_1 = \gamma_0 \cdot \gamma_1 \cdots \gamma_{L-2} = 1$ , etc., jusqu'au facteur de dilatation maximal pris en charge  $d = 2^{L-1}$ .

Les éléments de  $\Gamma$  sont ensuite combinés dans un vecteur de masque  $M$ , de longueur  $rf_{max}$ , comme indiqué dans la Figure 3.7. Chaque élément de  $M$  est ensuite multiplié par tous les poids de filtres relatifs à l'étape temporelle correspondante (et à tous les canaux) pour effectuer le masquage réel. Le côté droit de la Figure 3.7 montre comment les différentes alternatives de dilatation sont mappées sur les  $\Gamma$  entraînables grâce à la procédure de masquage décrite ci-dessus, pour le cas où  $rf_{max} = 9$ .

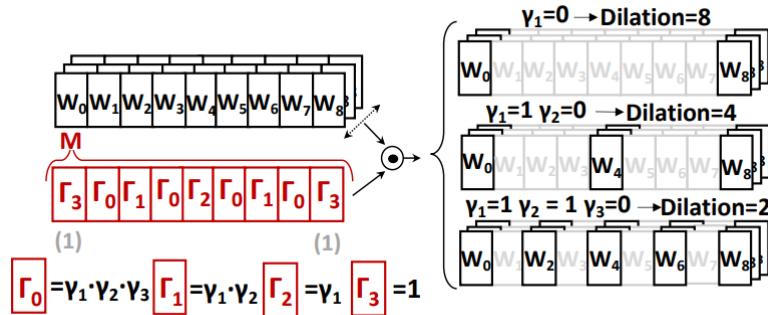


FIG. 3.7 : Combinaison des éléments de  $\gamma$  entre eux et avec les poids du filtre de convolution pour former différents schémas de dilatation. Exemple pour  $rf_{max} = 9$  [27].

Afin d'entraîner  $\gamma$ ,  $M$  doit être exprimée sous une forme différentiable. Pour ce faire, la transformation tensorielle suivante est utilisée :

$$M = \prod_{\text{colonnes}} \{[(\gamma \cdot \mathbb{1}_{1 \times L}) \odot T + (\mathbb{1}_{L \times L} - T)] \cdot K\} \quad (3.17)$$

où  $\prod_{\text{colonnes}}$  indique le produit de tous les éléments de chaque colonne de la matrice finale,  $\mathbb{1}_{i \times j}$  est une matrice composée de 1 de taille  $i \times j$  et  $\odot$  représente le produit de Hadamard.  $T$  et  $K$  sont deux matrices constantes composées de 0 et 1. Un exemple de

ces deux matrices et de l'ensemble de la transformation est présenté dans la Figure 3.8, pour le même  $rf_{max}$  et  $L$  utilisés dans la Figure 3.7. Encore une fois,  $\gamma_0$  a été directement remplacé par 1 pour plus de clarté.  $\mathbf{T}$  est une matrice triangulaire supérieure avec des colonnes inversées, tandis que  $\mathbf{K}$  peut être générée pour n'importe quelle valeur de  $rf_{max}$ .

$$\begin{aligned}
 & \underbrace{\gamma \cdot \mathbb{1}_{1 \times L}}_{\left[ \begin{array}{c} 1 \\ Y_1 \\ Y_2 \\ Y_3 \end{array} \right]} \cdot \underbrace{\left[ \begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array} \right]^t}_{\mathbf{T}} = \left[ \begin{array}{cccc} 1 & 1 & 1 & 1 \\ Y_1 & Y_1 & Y_1 & Y_1 \\ Y_2 & Y_2 & Y_2 & Y_2 \\ Y_3 & Y_3 & Y_3 & Y_3 \end{array} \right] \odot \underbrace{\left[ \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array} \right]}_{\mathbf{T}} = \left[ \begin{array}{cccc} 1 & 1 & 1 & 1 \\ Y_1 & Y_1 & Y_1 & 0 \\ Y_2 & Y_2 & 0 & 0 \\ Y_3 & 0 & 0 & 0 \end{array} \right] + \underbrace{\left[ \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{array} \right]}_{(\mathbb{1}_{L \times L} - \mathbf{T})} = \\
 & = \left[ \begin{array}{cccc} 1 & 1 & 1 & 1 \\ Y_1 & Y_1 & Y_1 & 1 \\ Y_2 & Y_2 & 1 & 1 \\ Y_3 & 1 & 1 & 1 \end{array} \right] \cdot \underbrace{\left[ \begin{array}{cccc} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{array} \right]}_{\mathbf{K}} = \left[ \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & Y_1 & Y_1 & Y_1 \\ 1 & Y_2 & Y_2 & Y_2 \\ 1 & Y_3 & Y_3 & Y_3 \end{array} \right] \stackrel{\prod_{\text{columns}}}{\downarrow} \underbrace{\left[ \begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & Y_1 & Y_1 & Y_1 \\ 1 & Y_2 & Y_2 & Y_2 \\ 1 & Y_3 & Y_3 & Y_3 \end{array} \right]}_{rf_{max}}
 \end{aligned}$$

FIG. 3.8 : Exemple de génération du vecteur de masque  $M$  avec des opérations tensorielles différentiables[27].

Finalement, l'équation complète de la couche pour la convolution dilatée dans PIT devient :

$$\mathbf{y}_t^m = \sum_{i=0}^{rf_{max}-1} \sum_{l=0}^{C_{in}-1} \mathbf{x}_{t-i}^l \cdot (\mathbf{M}_i \odot \mathbf{W}_i^{l,m}) \quad (3.18)$$

## Régularisation

Après avoir modifié toutes les couches de convolution avec des vecteurs  $\gamma$ , PIT effectue un entraînement standard, où la fonction de perte est augmentée d'un terme de régularisation afin de favoriser l'élagage des  $\gamma_i$ , permettant ainsi l'exploration d'architectures avec  $d > 1$ .

Le terme de régularisation prend en compte la contrainte de taille du réseau comme suit :

$$\mathcal{L}_R^{\text{taille}}(\gamma) = \lambda_{PIT} \sum_{l=1}^{\text{couches}} C_{in}^{(l)} \cdot C_{out}^{(l)} \sum_{i=1}^{L-1} \text{round} \left( \frac{rf_{max}-1}{2^{L-i}} \right) |\hat{\gamma}_i^{(l)}|, \quad (3.19)$$

où  $\lambda_{PIT}$  contrôle la force de la régularisation et  $(l)$  fait référence à la  $l$ -ème couche.  $C_{in}^{(l)}$  et  $C_{out}^{(l)}$  sont le nombre de canaux d'entrée /sortie dans la  $l$ -ème couche.  $\text{round} \left( \frac{rf_{max}-1}{2^{L-i}} \right)$  est le nombre de tranches de temps de filtre non masquées ajoutées par chaque  $\gamma_i^{(l)}$  non-nul.

La fonction de perte optimisée lors de la phase d'élagage est la suivante :

$$\mathcal{L}_{PIT}(\mathbf{W}, \gamma) = \mathcal{L}^{\text{perf}}(\mathbf{W}) + \mathcal{L}_R^{\text{taille}}(\gamma), \quad (3.20)$$

où  $\mathbf{W}$  est l'ensemble des paramètres entraînables du réseau et  $\mathcal{L}^{\text{perf}}(\mathbf{W})$  est la fonction de perte liée aux performances du TCN.

### Procédure d'entraînement

L'algorithme 2 résume la procédure d'entraînement mise en œuvre dans PIT. Comme le montre l'algorithme, le processus est composé de trois phases distinctes. Initialement, une phase appelée “warmup” d'une durée de  $Steps_{wu}$  époques d'entraînement est effectuée. Dans cette phase, tous les éléments des vecteurs  $\gamma$  sont initialisés à 1. Par conséquent, les poids du réseau initial sont entraînés avec des filtres de taille maximale et  $d = 1$  dans toutes les couches de convolution en ne considérant que la performance du réseau comme objectif. Ensuite, la boucle principale d'élagage est initiée. Ici, les poids et les vecteurs  $\gamma$  sont initialisés simultanément avec la fonction de perte régularisée. Une fois que la phase d'élagage a atteint la convergence, la troisième phase de fine-tuning est initiée, où tous les  $\gamma$  sont figés à leurs dernières valeurs binarisées et le réseau résultant avec dilatation est affiné en ne considérant encore une fois que la performance.

---

#### Algorithm 2 : L'algorithme PIT

---

```

1 : for  $i \leftarrow 1, \dots, Steps_{wu}$  do # boucle warmup
2 :   Mettre à jour  $\mathbf{W}$  en fonction de  $\nabla_{\mathbf{W}} \mathcal{L}_{\text{perf}}(\mathbf{W})$ 
3 : end for
4 : while convergence pas encore atteinte do #boucle pruning
5 :   Mettre à jour  $\mathbf{W}$  et  $\gamma$  en fonction de  $\nabla_{\mathbf{W}, \gamma} \mathcal{L}_{PIT}(\mathbf{W}, \gamma)$ 
6 : end while
7 : for  $i \leftarrow 1, \dots, Steps_{ft}$  do #boucle fine-tuning
8 :   Mettre à jour  $\mathbf{W}$  en fonction de  $\nabla_{\mathbf{W}} \mathcal{L}_{\text{perf}}(\mathbf{W})$ 
9 : end for

```

---

### 3.1.5 Architecture TEMPONet

Dans la méthode Q-PPG [3] que nous explorons dans ce chapitre, le réseau initial utilisé est une version légèrement modifiée de TEMPONet (Temporal Embedded Muscular Processing Online Network) [48], une architecture basée sur un réseau de convolution temporelle (TCN) qui a été proposée pour une autre application : la reconnaissance de gestes à partir de signaux EMG. Nous présentons ici la topologie du réseau TEMPONet telle que proposée dans [48]. Nous aborderons dans la section 3.2.1 les modifications apportées à celle-ci pour l'adapter à la problématique d'estimation de la FC à partir des signaux PPG et d'accéléromètre.

L'architecture est représentée dans la figure 3.9. Elle prend en entrée des fenêtres temporelles des signaux des 14 électrodes EMG de 150 ms échantillonnés à 2 kHz ( $300 \times 1 \times 14$ ).

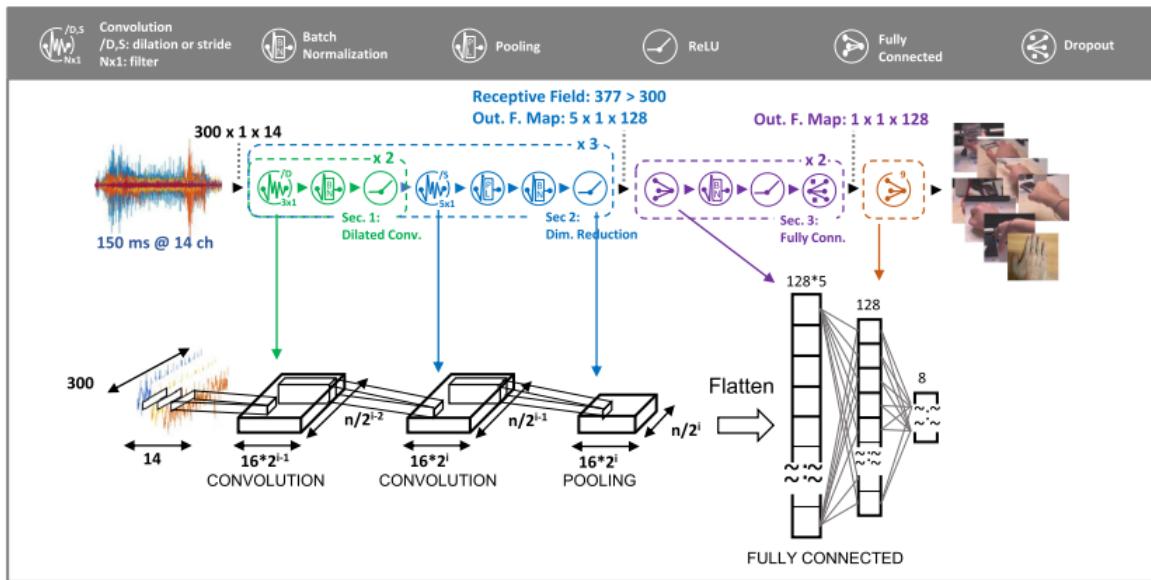


FIG. 3.9 : L'architecture TEMPONet [48].

La topologie comprend un extracteur de caractéristiques composé de 3 blocs de convolution successifs, chacun d'eux contenant :

- 2 couches de convolution temporelle avec une taille de filtre de  $3 \times 1$ , un facteur de dilatation variable,
- 1 couche de convolution 1D avec une taille de filtre de  $5 \times 1$ , un stride et un padding variables, suivie d'une couche de average pooling avec un filtre  $2 \times 1$ .

Les blocs 1, 2 et 3 ont respectivement un facteur de dilatation  $d = 2, 4, 8$ , et un stride pour les convolutions 1D  $s = 1, 2, 4$ . La figure 3.10 présente l'exemple du bloc 2.

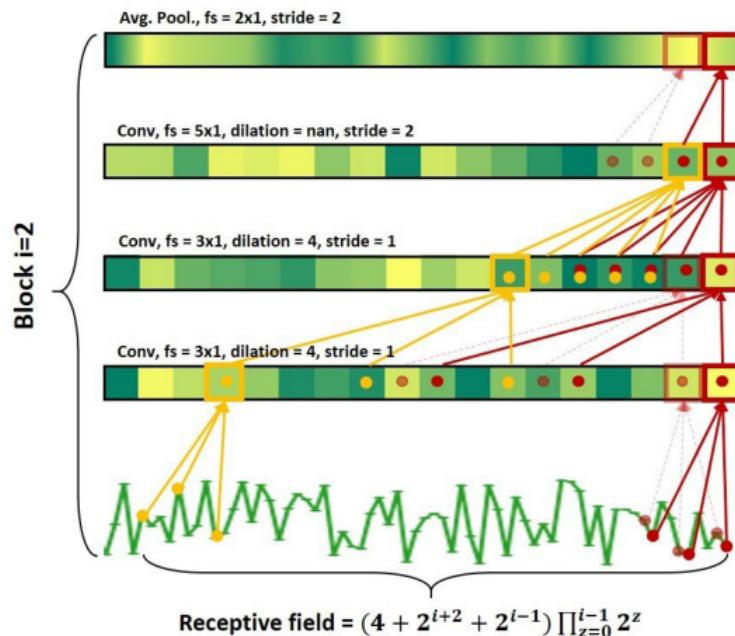


FIG. 3.10 : Structure du bloc 2 de convolution du réseau TEMPONet [48].

La couche de convolution 1D avec stride des blocs 1, 2, 3 augmente le nombre de canaux de sortie à 32, 64 et 128 respectivement, tandis que chaque couche de average pooling divise par deux la longueur de la séquence. Cela permet de construire un réseau où la dimension temporelle est consommée lentement et le nombre de vecteur caractéristiques (les canaux) extraits augmente. D'autre part, les couches TCN avec dilatation permettent d'augmenter progressivement le champ réceptif sur les séquences temporelles. Les facteurs de dilatation sont choisis de manière à ce que le champ réceptif à la fin du réseau couvre une fenêtre temporelle entière. Ainsi, une fenêtre d'entrée plus large peut être analysée en empilant davantage de blocs au lieu d'augmenter la taille des filtres, cela a pour effet de rendre le réseau plus profond plutôt que plus large, limitant ainsi le nombre de paramètres.

L'extracteur de caractéristiques est suivi du bloc de classification composé de 2 couches entièrement connectées. La première aplatis les 128 vecteurs caractéristiques obtenus à la sortie du dernier bloc de convolution et l'architecture se termine par une couche SoftMax qui produit une distribution de probabilité sur les classes de sortie (les gestes de la main).

Toutes les couches utilisent une fonction d'activation ReLU et sont équipées d'une normalisation par lots (BN).

## 3.2 Méthode Q-PPG

Dans cette section, nous présentons Q-PPG [3], une méthodologie qui permet de générer automatiquement une famille de Réseaux de Convolution Temporelle (TCNs) profonds pour l'estimation de la FC à partir de signaux PPG et d'accéléromètre, tous dérivés d'un unique réseau initial. Le processus implique une cascade de deux outils de recherche d'architectures neuronales (NAS) appliqués au réseau initial pour l'alléger et une étape de post-traitement effectuée sur les résultats du modèle allégé. L'organigramme de la méthode est présenté dans la figure 3.11.

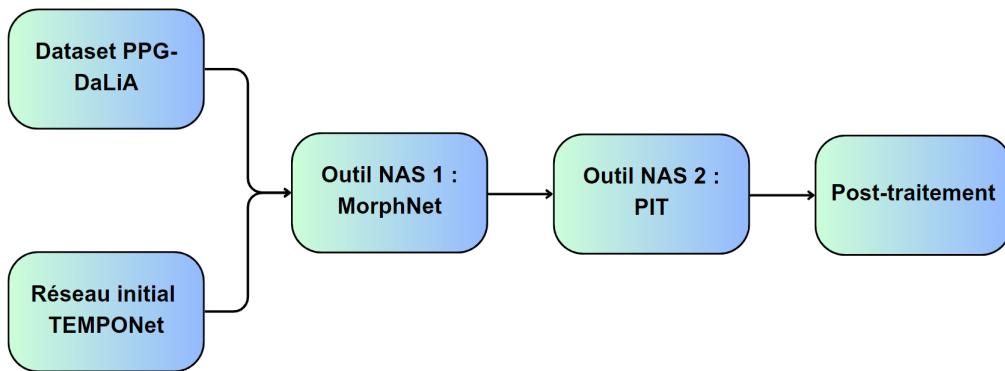


FIG. 3.11 : Organigramme de la méthode Q-PPG.

### 3.2.1 Réseau initial

Les entrées de Q-PPG sont :

- Un ensemble de données d'entraînement : le jeu de données PPG-DaLiA [22] contenant des données PPG et d'accéléromètre et les étiquettes de fréquence cardiaque correspondant à chaque fenêtre de temps (ces dernières sont issues de mesures de capteurs ECG).
- Un réseau TCN initial qui sert de modèle de référence à partir duquel tous les autres modèles de sortie sont générés. Le choix d'un Réseau de convolution temporelle (TCN) est motivé par les performances démontrées par les TCN dans le traitement de séries temporelles, ainsi que par leur avantage d'être plus adaptés à une implémentation matérielle à faible coût en comparaison avec les solutions précédemment proposées dans la littérature qui utilisent des architectures CNN-2D [22] [49].

Le processus consiste ensuite à optimiser l'architecture initiale en utilisant deux outils de recherche d'architectures neuronales (NAS) en cascade pour réduire le réseau initial tout en maintenant l'erreur MAE aussi faible que possible.

Pour couvrir un large espace de recherche d'architectures, le point de départ doit être un TCN relativement grand et précis. Le réseau initial choisi est une version adaptée de TEMPONet [48] présentée précédemment dans la sous-section 3.1.5. La structure de TEMPONet a été légèrement modifiée par rapport à [48] afin de la rendre compatible avec la tâche de l'estimation de la fréquence cardiaque et pour élargir l'espace exploré par Q-PPG. Une première modification pour assurer la compatibilité avec l'application consiste à modifier la première couche en y incluant 4 canaux, ce qui correspond au nombre de séries d'entrée (1 signal PPG et 3 signaux d'accéléromètre). De plus, la dernière couche entièrement connectée a été modifiée, en changeant le nombre de neurones à 1, comme requis pour une tâche de régression scalaire. Enfin, le facteur de dilatation est fixé à  $d = 1$  dans tout le modèle, ce qui laisse une liberté maximale à l'algorithme pour explorer toutes les valeurs possibles de ce facteur en l'augmentant pendant la phase d'optimisation de l'architecture.

Mis à part ces modifications, le réseau initial est identique à la version originale de TEMPONet.

### 3.2.2 Optimisation de l'architecture

Dans cette partie, nous décrivons la méthodologie utilisée pour générer différentes architectures TCN pour le suivi de la fréquence cardiaque dans l'espace **Précision de l'estimation vs. Complexité**. Comme indiqué dans la figure 3.11, une cascade de deux outils de recherche d'architectures neuronales (NAS) différents, appelés MorphNet [26] et Pruning-In-Time (PIT) [27], ont été exploités pour cette exploration. Ces deux outils ont été expliqués en détails dans la sous-section 3.1.4.

Une vue d'ensemble de la manière dont les deux outils sont utilisés est présentée dans la figure 3.12. Avant de commencer la recherche, les couches du réseau initial sont modifiées en ajoutant un nouvel ensemble de paramètres entraînables supplémentaires appelés masques ( $\alpha_i$  et  $\beta_i$  dans la figure) qui multiplient chacun un sous-ensemble des poids de la couche. Dans ce cas, les masques sont ajoutés aux couches de convolution du TEMPONet modifié. Les masques  $\alpha_i$  sont également appliqués aux couches de neurones entièrement connectés (sauf la dernière) contrairement aux  $\beta_i$  qui ne peuvent pas être appliqués à ces dernières. Ces masques sont ensuite entraînés avec les paramètres normaux du réseau, encourageant l'algorithme d'entraînement à réduire leurs valeurs.

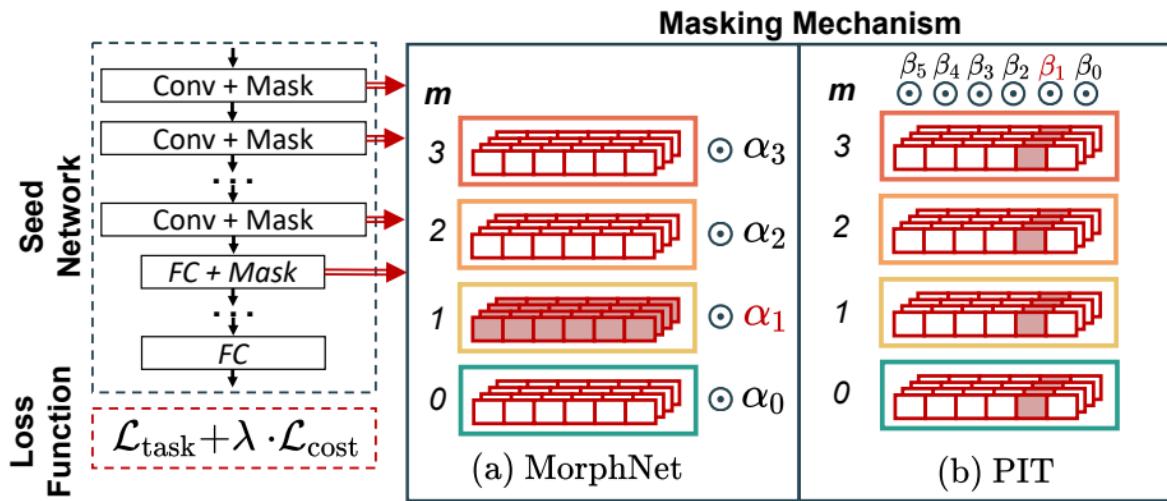


FIG. 3.12 : Schéma expliquant la fonctionnalité des deux algorithmes NAS utilisés pour l'optimisation de l'architecture. Les couches de pooling et autres ne sont pas illustrées pour des raisons de simplicité [3].

Après l'entraînement, les architectures optimisées sont obtenues simplement en éliminant les parties des couches correspondant à des masques de faible valeur. Comme le montre la figure 3.12, les masques sont contraints à de petites valeurs pendant l'entraînement en ajoutant à la fonction de perte normale  $\mathcal{L}_{task}$  un terme de régularisation supplémentaire  $\mathcal{L}_{cost}$ . Ce dernier calcule le coût attendu de toutes les couches de l'architecture (nombre de paramètres ou FLOPs) en fonction des valeurs des masques. Différentes architectures sont obtenues dans l'espace **Précision de l'estimation vs. Complexité** en modifiant l'importance relative des deux termes de la fonction de perte grâce aux paramètres de régularisation.

Comme il a été mentionné dans la sous-section 3.1.4 les deux outils utilisés diffèrent principalement par le mécanisme de masquage. MorphNet [26] masque tous les poids relatifs au même canal de sortie de convolution avec un  $\alpha_i$  (correspondant aux  $\gamma_i$  de BN). Par conséquent, cet outil peut être utilisé pour optimiser automatiquement le nombre de canaux de sortie  $C_{out}$  dans chaque couche de convolution. PIT [27], quant à lui, masque tous les poids correspondant au même pas de temps (et à tous les canaux de sortie) avec un  $\beta_i$  (correspondant aux  $\Gamma_i$  dans la figure 3.7), avec effet d'insérer des "trous" dans les filtres de convolution. Ainsi, cet outil peut être utilisé pour rechercher automatiquement le facteur de dilatation  $d$  optimal d'un TCN.

Il a été constaté empiriquement que l'exécution de MorphNet en premier, suivie de PIT, donne de bien meilleurs résultats que l'ordre inverse. Intuitivement, cela se produit parce que MorphNet opère dans un espace de recherche plus large et plus précis car les combinaisons possibles de canaux sont plus nombreuses que les valeurs de dilatation régulières possibles dans une couche de convolution typique.

#### 3.2.3 Post-traitement

Une étape de post-traitement est appliquée aux estimations  $FC_i$  produites par le modèle obtenu après l'étape d'optimisation de l'architecture. Elle est appliquée comme suit :

- La dernière prédiction du modèle,  $FC_n$ , est comparée à la moyenne des  $N$  précédentes valeurs,  $E_{n,N} = E[FC_{n-1}, \dots, FC_{n-N}]$ .
- Si la différence entre ces deux valeurs est supérieure à un seuil  $P_{th}$ , l'estimation renvoyée est  $FC_n = E_{n,N} \pm P_{th}$ .

Dans [3],  $N$  est fixé à 10 et  $P_{th} = E_{n,N}/10$ .

### 3.3 Expérimentations et évaluation

Toutes les expérimentations de cette section ont été réalisées sur Google Colab en utilisant Python 3.10 et TensorFlow 2.10. Elles se basent sur les implémentations proposées sur GitHub par les auteurs de Q-PPG [50] et MorphNet [51]. Les expérimentations ont été réalisées sur trois ensembles de données : PPG-DaLiA, IEEE SPC 2015 - train et IEEE SPC 2015 - test. Les détails de l'implémentation sont comme suit :

- Une fenêtre temporelle de 8 secondes et un shift de 2 secondes ont été utilisés pour traiter les signaux d'entrée qui sont le signal PPG et les trois signaux d'accéléromètre. La fréquence cardiaque (FC) a été estimée sur chaque fenêtre temporelle.
- Tous les entraînements ont été réalisés en mini-lots (batch) avec un optimiseur Adam et un taux d'apprentissage de  $10^{-3}$  :
  - Pour l'ensemble de données PPG-DaLiA, la taille d'un mini-lot a été fixée à 128. 350 époques d'entraînement ont été utilisées pour MorphNet, et 500 époques pour PIT.
  - Pour les ensembles de données IEEE SPC 2015, la taille du mini-lot a été adaptée aux tailles des jeux de données, elle a été fixée à 10. 80 époques d'entraînement ont été utilisées pour MorphNet, et 100 époques pour PIT.

- Pour l'évaluation des performances des modèles appris, le protocole de validation croisée Leave-One-Session-Out (LOSO), tel que proposé dans [22], a été utilisé. Ce protocole implique l'organisation des sujets en quatre ensembles de données choisis de manière aléatoire. Trois ensembles ont été utilisés comme ensemble d'entraînement, tandis que le quatrième a été divisé pour créer l'ensemble de test, composé d'un seul sujet, et l'ensemble de validation. Ainsi, l'entraînement est réalisé en tout 15 fois pour PPG-DaLiA, 12 fois pour IEEE SPC 2015 - train et 10 fois pour IEEE SPC 2015 - test (correspondant au nombre de sujets de chaque ensemble de données), de sorte que chaque sujet serve exactement une fois de données de test. 500 époques d'entraînement ont été utilisées pour PPG-DaLiA et 100 pour IEEE SPC 2015.

### 3.3.1 Expérimentations sur la méthodologie de recherche d'architectures proposée

Dans cette partie, nous explorons la méthodologie proposée pour la recherche d'architectures, telle que décrite dans [3], qui inclut deux outils NAS, comme expliqué précédemment. Nous commençons par étudier l'effet du choix d'un des deux types de contraintes de ressources proposées par MorphNet, à savoir la taille du modèle (c'est-à-dire le nombre de paramètres) et le nombre d'opérations en virgule flottante par inférence (FLOPs). Ensuite, nous évaluons l'impact du réglage des paramètres de régularisation de MorphNet et PIT, désignés respectivement par  $\lambda_{MorphNet}$  et  $\lambda_{PIT}$ , qui sont utilisés pour ajuster l'importance relative entre performance et minimisation des contraintes de ressources. Enfin, nous analysons l'effet de l'étape de "warmup" sur le réseau obtenu. Ces expérimentations ont été réalisées en utilisant l'ensemble de données PPG-DaLiA.

#### 3.3.1.1 Effet du choix de la contrainte de ressources

Dans cette partie, nous étudions l'effet du choix d'un des deux types de contraintes de ressources proposées par MorphNet [26]. Le tableau 3.1 présente le nombre de canaux par couche dans les réseaux élagués obtenus en ciblant soit l'optimisation des FLOPs ou du nombre de paramètres du modèle à l'aide de MorphNet. Les deux architectures présentées dans le tableau obtiennent une MAE similaire de 5.25 BPM et 5.29 BPM bien qu'elles soient très différentes. En effet, nous notons que le régularisateur FLOPs supprime principalement les canaux de filtres des premières couches du réseau situées près de l'entrée, tandis que le régularisateur de taille du modèle a tendance à supprimer les canaux /neurones des couches près de la sortie de l'architecture. Cela peut être expliqué par le fait que les premières couches du TCN traitent des séquences temporelles plus longues, ce qui entraîne des convolutions consommant plus de FLOPs. En revanche, les couches près de la sortie ont tendance à avoir un nombre plus élevé de canaux, cela permet au réseau d'apprendre des caractéristiques plus complexes des séquences d'entrée. Les couches plus profondes ont donc des matrices de poids plus grandes, ce qui explique le fait que le régularisateur de taille du modèle se concentre sur la suppression des canaux de ces dernières.

Les modèles induits par le ciblage des FLOPs offrent un meilleur compromis entre consommation de FLOPs et performance, mais un compromis taille du modèle - performance non optimal, et vice versa. La force de l'outil MorphNet réside dans sa capacité à répondre à différentes contraintes de ressources, offrant ainsi une flexibilité en fonction des exigences matérielles.

Couche	Régularisation appliquée		
	Nombre de paramètres (MAE = 5.25 BPM)	FLOPs (MAE = 5.29 BPM)	
Bloc de convolution 1	1 <b>32</b>	<b>17</b>	
	2 <b>30</b>	<b>14</b>	
	3 <b>64</b>	<b>25</b>	
Bloc de convolution 2	4 62	48	
	5 60	41	
	6 69	33	
Bloc de convolution 3	7 16	12	
	8 21	20	
	9 <b>25</b>	<b>32</b>	
Couches FC	10 <b>29</b>	<b>68</b>	
	11 56	52	
Sortie	12 1	1	

TAB. 3.1 : Comparaison du nombre de canaux par couche de deux architectures élaguées avec MorphNet, qui obtiennent des performances similaires, en ciblant l'optimisation soit de la taille du modèle ou du nombre de FLOPs.

Dans les expérimentations qui suivent, nous ne considérerons que la taille du modèle comme contrainte cible.

### 3.3.1.2 Effet du réglage des paramètres de régularisation

Afin d'examiner exclusivement l'effet du réglage des paramètres de régularisation, les étapes de “warmup” et de post-traitement ne sont pas appliquées dans cette partie. Le protocole de recherche suivant est adopté : d'abord MorphNet est appliqué au réseau initial, avec différentes forces de régularisation ( $\lambda_{MorphNet} = 10^{-6}$ ,  $\lambda_{MorphNet} = 10^{-5}$ ,  $\lambda_{MorphNet} = 10^{-4}$ ,  $\lambda_{MorphNet} = 10^{-3}$ ) cela donne un premier ensemble de TCNs (tracé bleu dans la figure 3.13). Ensuite, PIT est appliqué aux TCNs obtenus lors de la première phase, en répétant à nouveau l'entraînement avec différentes forces de régularisation ( $\lambda_{PIT} = 10^{-7}$ ,  $\lambda_{PIT} = 10^{-5}$ ,  $\lambda_{PIT} = 5 \times 10^{-3}$ ) (losange, carré et cercle dans la figure 3.13). La figure 3.13 nous permet d'observer les erreurs absolues moyennes obtenues par différentes architectures, chacune possédant un nombre de paramètres différent.

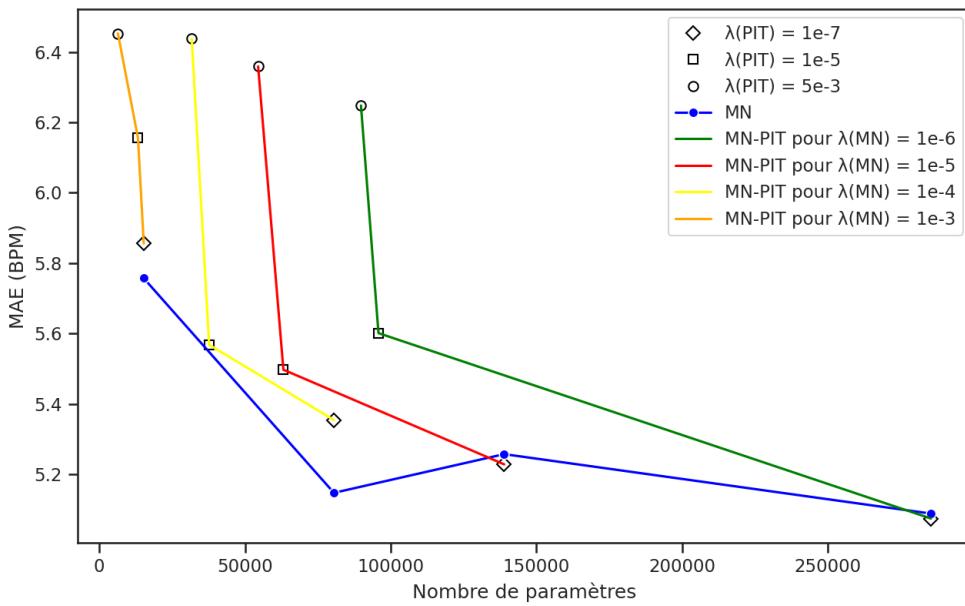


FIG. 3.13 : Résultats en MAE (BPM) obtenus avec des architectures possédant différents nombres de paramètres. L'étiquette MN correspond à l'utilisation de MorphNet uniquement, MN-PIT correspond à l'utilisation de MorphNet suivi de PIT.

D'après les résultats obtenus, il est clair que la méthodologie proposée par Q-PPG offre un vaste choix d'architectures avec des variations significatives du nombre de paramètres (allant d'un peu plus de 6000 à plus de 285000 paramètres), tout en maintenant une erreur absolue moyenne sur l'estimation de la fréquence cardiaque relativement faible (entre 5.08 à 6.45 BPM). Cette caractéristique est une force majeure de la méthodologie, car elle permet de choisir un modèle spécifique en fonction des contraintes imposées par la cible matérielle et ce, en réglant uniquement deux paramètres  $\lambda_{MorphNet}$  et  $\lambda_{PIT}$ . Cependant, il est important de souligner qu'un grand nombre de valeurs doivent être testées pour parvenir à l'architecture souhaitée, ce qui nécessite un entraînement à chaque modification.

### 3.3.1.3 Effet de l'étape de “warmup”

Ici, nous analysons l'impact de l'étape de “warmup”, qui est une étape optionnelle précédant la phase d'élagage dans PIT. Pour ce faire, les paramètres de régularisation sont fixés à  $\lambda_{MorphNet} = 10^{-6}$  et  $\lambda_{PIT} = 10^{-6}$ , le post-traitement n'est pas appliqué et la contrainte de taille du modèle est utilisée. Nous testons différentes valeurs pour le nombre d'époques d'entraînement de la phase de “warmup”, noté  $Steps_{wu}$ . Les résultats sont présentés dans le tableau suivant :

$Steps_{wu}$	1	10	30
Nombre de paramètres du modèle	141 667	190 667	269 259

TAB. 3.2 : Effet de l'étape de “warmup” sur le nombre de paramètres du modèle.

Nous pouvons voir d'après ces résultats que le nombre d'époques d'entraînement consacrées à la phase de “warmup”, noté  $Steps_{wu}$ , a un impact significatif sur le nombre de paramètres du modèle élagué. En raccourcissant cette phase, c'est-à-dire en réduisant le nombre d'époques d'entraînement, nous obtenons un réseau initial moins précis avant le début de l'étape d'élagage, ce qui favorise les simplifications du modèle et réduit le nombre de paramètres. En effet, l'effet des paramètres sur les performances est moins important pour un entraînement plus court.

#### 3.3.2 Évaluation des performances

Dans cette partie, nous évaluons les performances de la méthode en termes d'erreur absolue moyenne (MAE) en utilisant trois ensembles de données : PPG-DaLiA, qui a été utilisé dans le papier de référence, IEEE SPC 2015 - train et IEEE SPC 2015 - test. Le premier ensemble de données est utilisé pour évaluer les performances en fonction du type d'activité, tandis que les deux autres servent à évaluer l'impact de l'utilisation d'ensembles de données restreints lors de l'entraînement.

##### 3.3.2.1 Évaluation sur l'ensemble de données PPG-DaLiA

Afin d'évaluer les performances sur l'ensemble de données PPG-DaLiA, nous utilisons le modèle ayant atteint la meilleure MAE parmi ceux que nous avons obtenus dans la partie 3.3.1 et ce, avec la configuration suivante :  $\lambda_{MorphNet} = 10^{-6}$ ,  $\lambda_{PIT} = 10^{-6}$  et  $Steps_{wu} = 10$ . Le modèle atteint une MAE de **4.56 BPM** avec le post-traitement décrit 3.2.3 et a, en tout, 190667 paramètres. Nous obtenons les résultats suivants :

Sujet	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	Moyenne (std)
MAE (BPM)	3.70	3.43	2.40	5.14	12.19	4.49	2.42	5.91	7.37	2.75	4.51	5.75	2.26	2.74	3.34	4.56 (5.27)

TAB. 3.3 : Résultats en MAE (BPM) de la méthode Q-PPG sur PPG-DaLiA.

Les résultats obtenus par la méthode Q-PPG sur le jeu de données PPG-DaLiA peuvent être considérés comme très bons. Rappelons que la méthode WFPV avait obtenu une MAE de **13.35 BPM** sur le même ensemble de données, tandis que JOSS avait divergé pour la plupart des activités.

Le sujet S5 obtient l'erreur la plus élevée (12.19 BPM), ce qui peut s'expliquer par le fait que ce sujet est un cas atypique dans l'ensemble de données, sa fréquence cardiaque moyenne étant significativement plus élevée que celles de tous les autres sujets. Par conséquent, lorsque le modèle est entraîné sur les données des autres sujets, il ne se généralise pas correctement sur S5. Pour ce qui est du sujet ayant obtenu la plus petite erreur, c'est le sujet S13, avec une erreur absolue moyenne de 2.26 BPM.

### Chapitre 3. Approche par apprentissage profond

---

Les estimations faites sur le meilleur sujet (S13) et le pire sujet (S5) de l'ensemble de données PPG-DaLiA sont illustrées dans les figures 3.14 et 3.15, respectivement.

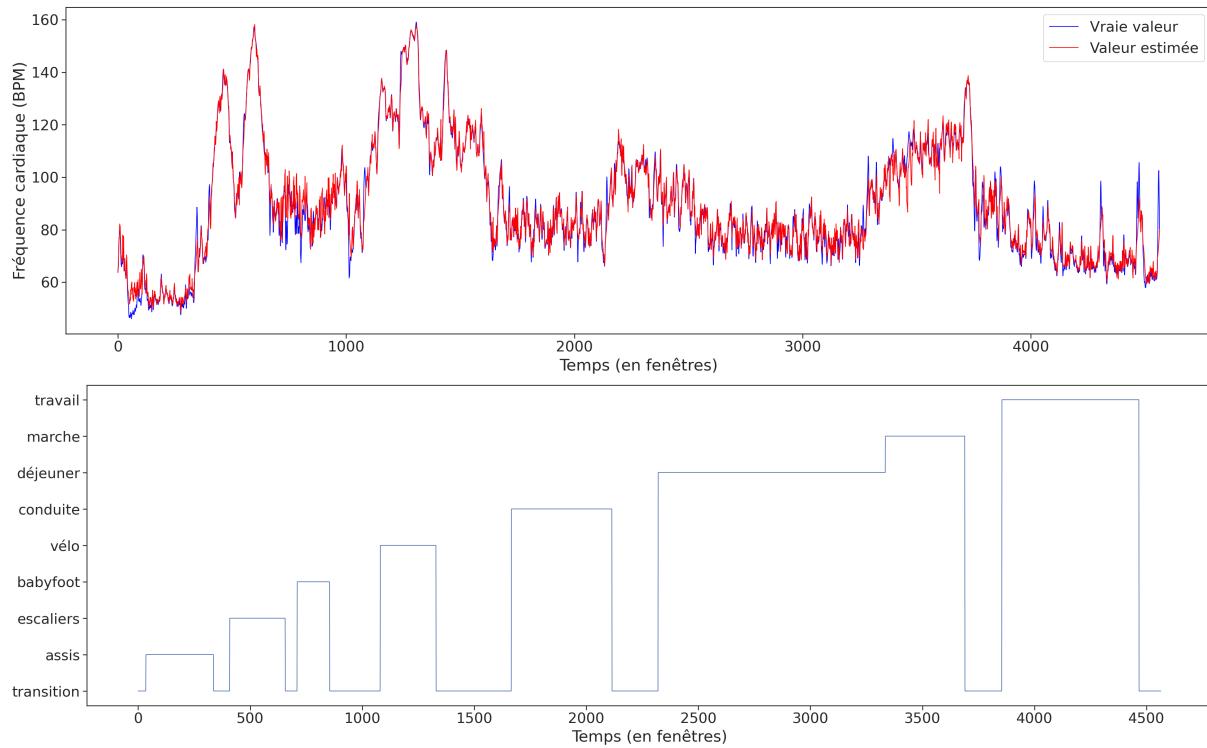


FIG. 3.14 : Estimation faite sur le meilleur sujet (S13) de PPG-DaLiA avec Q-PPG.

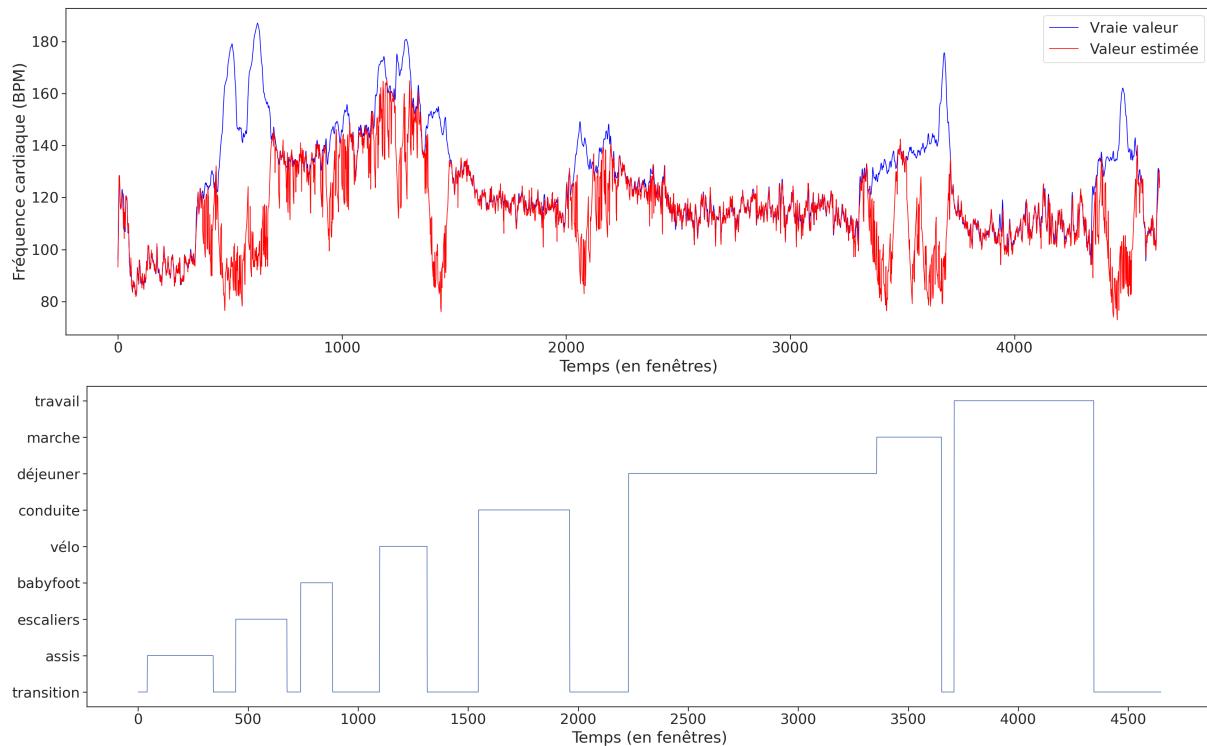


FIG. 3.15 : Estimation faite sur le pire sujet (S5) de PPG-DaLiA avec Q-PPG.

Nous passons maintenant à l'évaluation des performances sur les activités, la figure 3.16 permet de visualiser l'erreur absolue moyenne obtenue par sujet et par activité et l'erreur moyenne par activité.

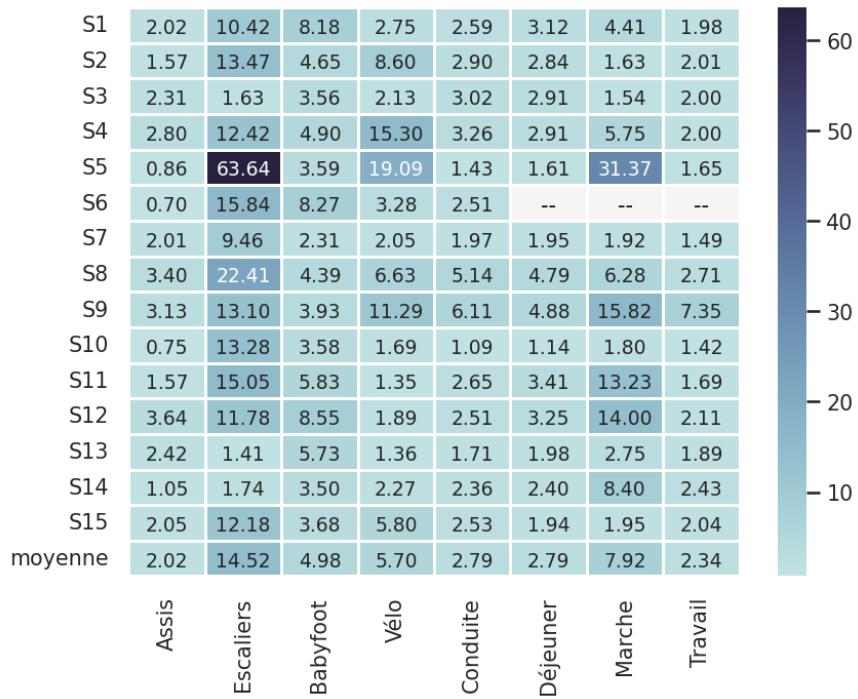


FIG. 3.16 : Résultats en MAE (BPM) de la méthode Q-PPG par sujet et par activité pour l'ensemble de données PPG-DaLiA.

D'après la figure 3.16 nous pouvons voir que la méthode Q-PPG obtient de très bons résultats sur la plupart des activités. Le tableau 3.4 nous permet d'analyser plus en détails ces résultats et les comparer à ceux obtenus par la méthode WFPV.

	MAE Q-PPG (BPM)	MAE WFPV (BPM)	FC_Moyenne (BPM)	Durée de l'activité (min)
Assis	2.02	2.04	61.07	10
Escaliers	14.52	40.11	118.53	5
Babyfoot	4.98	26.38	90.4	5
Vélo	5.70	29.39	123.05	8
Conduite	2.79	7.37	84.7	15
Déjeuner	2.79	6.89	83.58	30
Marche	7.92	22.35	99.18	10
Travail	2.34	5.95	76.4	20

TAB. 3.4 : Analyse des résultats obtenus avec Q-PPG et WFPV pour chaque activité incluse dans l'ensemble de données PPG-DaLiA, en précisant la fréquence cardiaque moyenne et la durée de chaque activité.

Nous pouvons voir que pour la méthode Q-PPG, tout comme ce qui avait été trouvé pour la méthode WFPV, il existe une forte corrélation entre la fréquence cardiaque moyenne d'une activité et l'erreur absolue moyenne (MAE) obtenue pour cette dernière. Les erreurs les plus élevées sont obtenues pour les activités ayant les fréquences cardiaques moyennes les plus grandes, en l'occurrence : escaliers, babyfoot, vélo et marche. Cette observation était attendue, car une fréquence cardiaque élevée correspond généralement à une activité physique plus intense, ce qui entraîne des artefacts de mouvement plus importants. Toutefois, il est important de noter que la méthode Q-PPG obtient de bien meilleurs résultats que la méthode WFPV sur toutes les activités.

Une observation intéressante supplémentaire concerne le fait que les erreurs les plus élevées sont associées aux activités les plus courtes en durée. On peut donc supposer qu'avec des durées d'enregistrement plus longues, le modèle serait entraîné sur un plus grand nombre de données de ce type, ce qui pourrait conduire à de meilleurs résultats. Cette supposition reste tout de même à confirmer.

#### 3.3.2.2 Évaluation sur les ensembles de données IEEE SPC 2015

Afin d'évaluer les performances sur les ensembles de données IEEE SPC 2015, nous utilisons les modèles ayant atteint les meilleures MAEs, obtenus avec les configurations suivantes :  $\lambda_{MorphNet} = 10^{-4}$ ,  $\lambda_{PIT} = 10^{-8}$  et  $Steps_{wu} = 10$  (108553 paramètres) pour IEEE SPC 2015 - train et  $\lambda_{MorphNet} = 10^{-3}$ ,  $\lambda_{PIT} = 10^{-8}$  et  $Steps_{wu} = 10$  (119642 paramètres) pour IEEE SPC 2015 - test.

Nous obtenons les résultats suivants sur le jeu de données IEEE SPC 2015 - train avec post-traitement :

Sujet	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Moyenne (std)
MAE (BPM)	32.20	5.60	2.28	3.17	1.16	2.93	1.72	6.28	2.26	19.24	6.63	3.42	7.24 (5.66)

TAB. 3.5 : Résultats en MAE (BPM) de la méthode Q-PPG sur l'ensemble de données IEEE SPC 2015 - train.

Nous obtenons les résultats suivants sur l'ensemble de données IEEE SPC 2015 - test avec post-traitement :

Session	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Moyenne (std)
MAE (BPM)	3.41	3.64	8.05	38.04	8.77	7.09	13.49	13.47	5.58	2.96	10.45 (7.82)

TAB. 3.6 : Résultats en MAE (BPM) de la méthode Q-PPG sur l'ensemble de données IEEE SPC 2015 - test.

Nous pouvons voir que la méthode Q-PPG obtient de très mauvais résultats lorsque les modèles sont entraînés sur des ensembles de données limitées malgré la simplicité des activités (l'ensemble de données IEEE SPC 2015 - train). Rappelons que les méthodes WFPV et JOSS avaient obtenu respectivement des MAE de **1.02 BPM** et **1.28 BPM** sur IEEE SPC 2015 - train et **2.96 BPM** et **4.21 BPM** sur IEEE SPC 2015 - test.

Ceci peut être facilement expliqué car les performances des modèles de deep learning dépendent fortement des données sur lesquelles ils sont entraînés. En effet, lorsque le modèle est entraîné sur un ensemble de données restreint, cela peut résulter dans de nombreuses limitations comme des risques de surajustement et la possibilité de biais résultant de l'influence disproportionnée de certaines caractéristiques présentes dans l'ensemble de données [52]. Ainsi, nous pouvons voir l'apport positif de l'introduction du jeu de données PPG-DaLiA pour le développement de méthodes de Deep Learning consacrées à l'estimation de la fréquence cardiaque à partir de signaux PPG fortement affectés par des artefacts de mouvement.

### 3.3.3 Expérimentations sur le post-traitement

Bien que les erreurs obtenues sur l'ensemble de données PPG-DaLiA soient très bonnes, des valeurs aberrantes sont parfois observées, cela est illustré par le tracé vert sur la figure 3.17. En effet, une solution basée sur l'apprentissage profond permet en moyenne une très bonne précision d'estimation mais peut parfois produire des erreurs élevées et imprévisibles, surtout lorsque le segment d'entrée traité diffère significativement des données sur lesquelles le modèle a été entraîné. Pour résoudre ce problème, une simple étape de post-traitement appliquée aux estimations du modèle permet de réduire les erreurs ponctuelles.

Le post-traitement est basé sur l'observation selon laquelle les changements brusques de la fréquence cardiaque sont physiologiquement limités. Par conséquent, une estimation du modèle qui diffère significativement des valeurs antérieures est probablement due à une erreur du modèle.

3 post-traitements sont expérimentés :

- Post-traitement 1 : Le post-traitement appliqué dans Q-PPG et décrit dans la sous-section 3.2.3.
- Post-traitement 2 : Celui proposé dans WFPV et présenté dans le paragraphe 2.2.1.4.
- Post-traitement 3 : Une variante du post-traitement Q-PPG où le filtre moyenneur simple est remplacé par une moyenne pondérée, qui donne plus de poids aux observations les plus récentes. Le poids de la  $i$ -ème précédente estimation est  $e^{-\alpha i}$ . Le paramètre  $\alpha$ , détermine l'importance relative accordée à l'estimation récente par rapport aux plus anciennes, ici, il est déterminé empiriquement de sorte à obtenir les meilleurs résultats et il est fixé à 0.2.

Le tableau 3.7 présente les résultats du modèle le plus précis sur PPG-DaLiA. Les résultats sont indiqués à la fois sans et avec les différents post-traitements.

Sujet	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	Moyenne (BPM)
Sans post-traitement	4.19	3.52	2.47	5.94	12.12	5.49	2.20	7.281	9.16	3.15	4.92	7.03	2.11	2.76	3.43	<b>5.05</b>
Avec post-traitement 1	3.70	3.43	2.400	5.14	12.19	4.49	2.41	5.91	7.37	2.75	4.50	5.74	2.25	2.74	3.34	<b>4.56</b>
Avec post-traitement 2	4.14	3.53	2.48	5.72	12.16	5.58	2.21	7.14	8.95	3.12	4.94	7.00	2.13	2.80	3.39	<b>5.02</b>
Avec post-traitement 3	3.67	3.42	2.42	5.05	12.06	4.36	2.44	5.81	7.24	2.70	4.46	5.72	2.28	2.75	3.32	<b>4.51</b>

TAB. 3.7 : Résultats obtenus en MAE (BPM) par le modèle le plus précis sur l'ensemble de données PPG-DaLiA sans et avec différents post-traitements.

Un exemple de la sortie produite par le modèle sans et avec les différents post-traitements est présenté dans la figure 3.17.

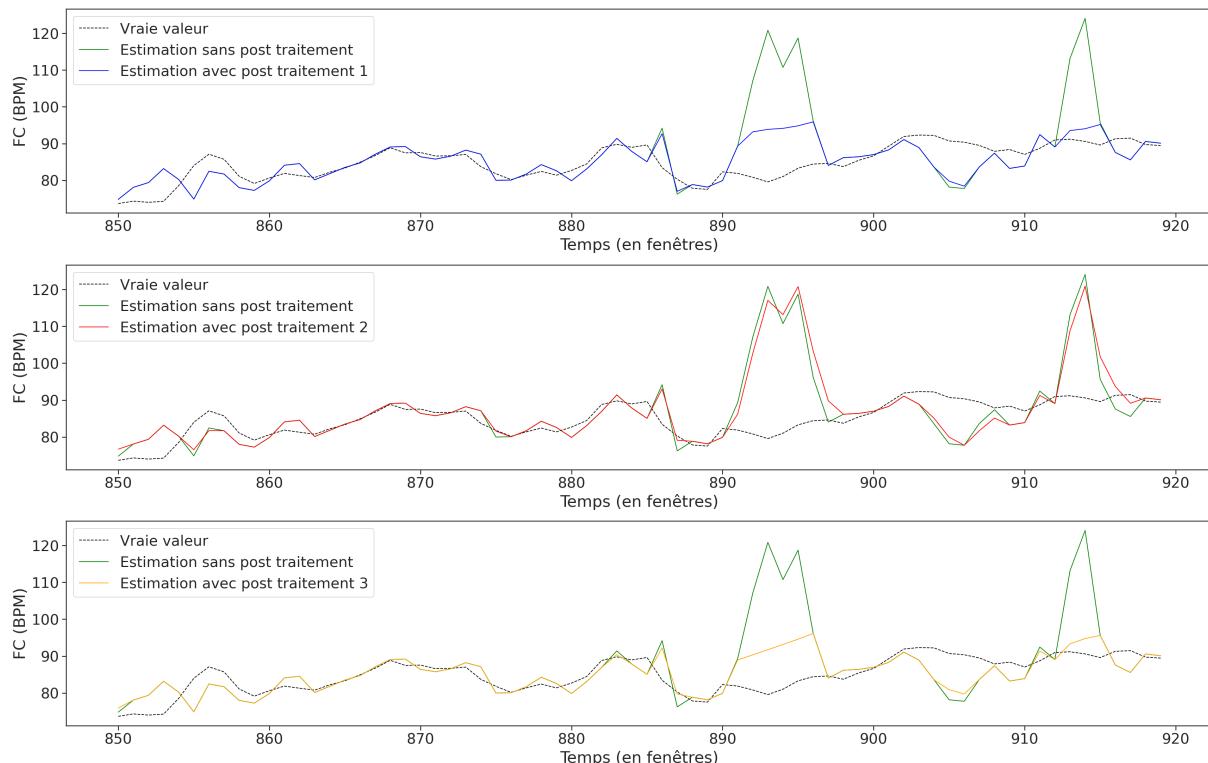


FIG. 3.17 : Estimations produites par le modèle le plus précis sur une partie de l'enregistrement du sujet S13 sans et avec différents post-traitements.

D'après les résultats présentés dans le tableau 3.7 et la figure 3.17, nous faisons les constatations suivantes :

- Le post-traitement 2 ne parvient pas à corriger les erreurs ponctuelles. Cela peut être expliqué par le poids trop important donné à l'estimation erronée dans la somme pondérée. Cependant, accorder plus de poids à l'estimation faite par la régression linéaire produit une divergence dans les estimations de la FC lorsque le post-traitement est sollicité sur plusieurs segments successifs.

- Le post-traitement 1 permet de réduire les erreurs ponctuelles, ce qui fait passer la MAE de **5.05 BPM** à **4.56 BPM** sur l'ensemble de données PPG-DaLiA. Le filtre moyenneur qui considère les 10 dernières estimations garantit une meilleure robustesse de l'estimation de la FC et permet de lisser davantage la sortie du modèle.
- Le post-traitement 3 permet d'améliorer légèrement les résultats obtenus par le post traitement 1 (**4.51 BPM**), car il permet à la fois un lissage robuste en considérant la moyenne sur les 10 dernières valeurs tout en donnant plus de poids aux segments plus récents, ce qui permet de conserver la dynamique de progression de la FC.

Le sujet S1 de l'ensemble de données IEEE SPC 2015 - train a obtenu une MAE de **16.83 BPM** sans post-traitement, tandis qu'avec le post-traitement Q-PPG, censé améliorer les résultats, la MAE est passée à **32.20 BPM**. La figure 3.18 permet d'illustrer cette situation :

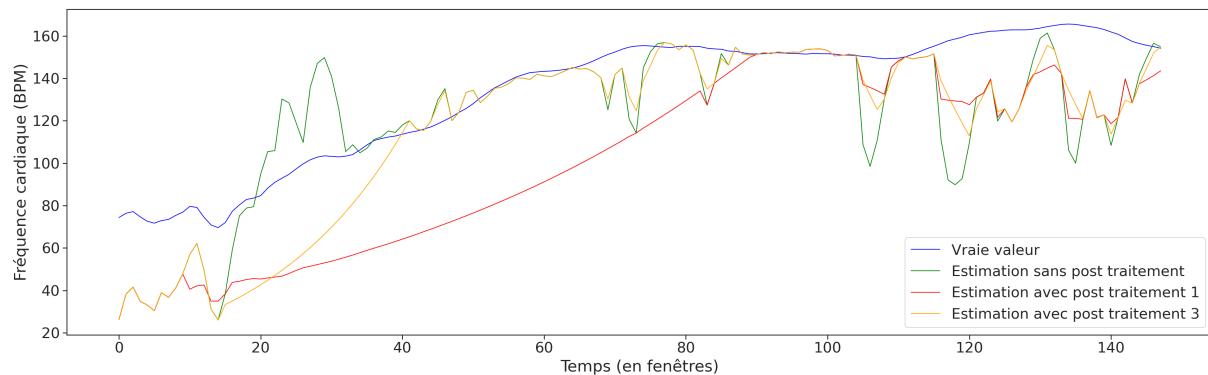


FIG. 3.18 : Estimations faites avec et sans post-traitement sur le sujet S1 de l'ensemble de données IEEE SPC 2015 - train avec la méthode Q-PPG.

Nous pouvons voir d'après la figure que les valeurs de la FC ont été mal estimées dès le départ (tracé vert). Par la suite, les valeurs estimées par le modèle augmentent considérablement pour approcher les vraies valeurs (tracé bleu) et corriger l'allure initiale.

- Le post-traitement 1 (tracé rouge) ne permet pas cette augmentation rapide et modifie les estimations pour qu'elles coïncident avec l'allure de départ. Nous voyons ici, une limite de ce post-traitement qui est fortement dépendant des premières estimations.
- Le post-traitement 3 permet de corriger en partie ce problème (tracé orange). Ceci est dû au fait qu'il accorde une plus grande importance aux segments les plus récents, ce qui maintient l'élan de progression de la FC.

Ainsi, nous voyons la nécessité d'une investigation plus approfondie des méthodes de post-traitement utilisées dans le contexte de cette problématique.

## Conclusion

Dans ce chapitre, nous avons exploré en détail la méthode Q-PPG [3]. Nous avons effectué différentes expérimentations qui nous ont permis de confirmer les forces de la méthode, notamment sa capacité à répondre à différentes contraintes de ressources, offrant ainsi une flexibilité en fonction des exigences matérielles, tout en obtenant des MAE sur les estimations de la fréquence cardiaque meilleures que celles des autres méthodes. De plus, nous avons constaté sa meilleure capacité de généralisation.

Toutefois, nous avons également mis en évidence certaines limitations, telle que la forte dépendance des performances aux données sur lesquelles les modèles sont entraînés. Aussi, bien que les erreurs obtenues sur le jeu de données PPG-DaLiA soient bonnes, des valeurs aberrantes sont parfois observées. Ces valeurs aberrantes peuvent être corrigées avec le post-traitement, mais il convient de noter que le post-traitement lui-même présente également des limitations et doit être choisi avec précaution.

# **Conclusion et perspectives**

## Conclusion et perspectives

---

Dans ce travail, nous avons traité la problématique de l'estimation de la fréquence cardiaque à partir de signaux PPG affectés par des artefacts de mouvement. Nous avons exploré trois méthodes proposées dans la littérature pour la résolution de cette problématique, deux méthodes de l'approche classique, à savoir JOSS et WFPV, ainsi qu'une méthode de l'approche par apprentissage profond qui est Q-PPG. La principale contribution de ce travail a été de mener différentes expérimentations sur les méthodologies proposées et d'avoir évalué les méthodes sur des ensembles de données sur lesquels elles n'ont jamais été évaluées.

En ce qui concerne les méthodes de l'approche classique, nous avons fait plusieurs constatations. D'abord, nous avons constaté que les deux méthodes avaient de faibles capacités de généralisation lorsqu'elles étaient appliquées à de nouvelles activités plus longues et complexes mais qu'elles présentaient de bons résultats sur les ensembles de données simples. De plus, nous avons observé que la méthode JOSS était fortement dépendante de la configuration de ses nombreux paramètres et seuils heuristiques, ce qui avait un impact sur ses performances, contrairement à la méthode WFPV qui conservait de bonnes performances indépendamment de la configuration des paramètres. Enfin, nous avons évalué les temps d'exécution de chaque méthode. La méthode WFPV s'est avérée rapide, tandis que la méthode JOSS était significativement plus lente en termes de temps d'exécution.

Pour ce qui est de la méthode Q-PPG, nous avons réalisé diverses expérimentations qui nous ont permis de confirmer sa capacité d'adaptation à différentes contraintes de ressources, offrant ainsi une flexibilité en fonction des exigences matérielles. De plus, nous avons constaté qu'elle présentait une meilleure précision et une meilleure capacité de généralisation dans l'estimation de la fréquence cardiaque en comparaison avec les deux autres méthodes évaluées. Cependant, nous avons également identifié certaines limitations, telles que la forte dépendance des performances par rapport aux données sur lesquelles les modèles sont entraînés.

Ces résultats nous incitent à envisager différentes perspectives de travaux futurs :

- Pour ce qui est des méthodes de l'approche classique, nous avons remarqué que la plupart des paramètres et seuils heuristiques ajustés manuellement se trouvent dans la partie SPT des algorithmes, ce qui limite leur capacité de généralisation. Il serait donc envisageable d'automatiser cette partie en utilisant des méthodes d'apprentissage automatique.
- L'utilisation d'outils de recherche d'architectures neuronales (NAS) dans Q-PPG a donné des résultats très prometteurs pour l'optimisation des modèles en vue de futures implémentations. Il serait donc intéressant d'explorer davantage cette voie en utilisant d'autres outils NAS proposés dans la littérature.
- Étant donné que les performances des méthodes d'apprentissage profond dépendent fortement des données sur lesquelles les modèles sont entraînés, il est nécessaire de collecter un ensemble de données plus vaste et plus diversifié. Comme nous l'avons vu dans nos résultats, l'ensemble de données le plus large disponible à ce jour, PPG-DaLiA, a montré ses limites.

## **Conclusion et perspectives**

---

- Les solutions basées sur l'apprentissage profond offrent généralement une précision d'estimation très élevée, mais elles peuvent parfois générer des erreurs imprévisibles et importantes. Ces valeurs aberrantes peuvent être corrigées par un processus de post-traitement, cependant, il est important de souligner que le post-traitement lui-même comporte également certaines limites. Par conséquent, une étude plus approfondie des méthodes de post-traitement est nécessaire.

# Bibliographie

1. CASTANEDA, Denisse ; ESPARZA, Aibhlin ; GHAMARI, Mohammad ; SOLTAN-PUR, Cinna ; NAZERAN, Homer. A review on wearable photoplethysmography sensors and their potential future applications in health care. *International journal of biosensors & bioelectronics*. 2018, t. 4, n° 4, p. 195.
2. MEJIA-MEJIA, Elisa ; ALLEN, John ; BUDIDHA, Karthik ; EL-HAJJ, Chadi ; KY-RIACOU, Panicos A ; CHARLTON, Peter H. Photoplethysmography signal processing and synthesis. In : *Photoplethysmography*. Elsevier, 2022, p. 69-146.
3. BURRELLO, Alessio ; PAGLIARI, Daniele Jahier ; RISSO, Matteo ; BENATTI, Simone ; MACII, Enrico ; BENINI, Luca ; PONCINO, Massimo. Q-ppg : Energy-efficient ppg-based heart rate monitoring on wearable devices. *IEEE Transactions on Biomedical Circuits and Systems*. 2021, t. 15, n° 6, p. 1196-1209.
4. ZHANG, Zhilin. Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction. *IEEE transactions on biomedical engineering*. 2015, t. 62, n° 8, p. 1902-1910.
5. TEMKO, Andriy. Accurate heart rate monitoring during physical exercises using PPG. *IEEE Transactions on Biomedical Engineering*. 2017, t. 64, n° 9, p. 2016-2024.
7. ZHANG, Gus ; ZHANG, Weiguo. Heart rate, lifespan, and mortality risk. *Ageing Research Reviews*. 2009, t. 8, p. 52-60. Disp. à l'adr. DOI : [10.1016/j.arr.2008.10.001](https://doi.org/10.1016/j.arr.2008.10.001).
10. ALLEN, John. Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement*. 2007, t. 28, n° 3, R1.
11. ALMARSHAD, Malak Abdullah ; ISLAM, Md Saiful ; AL-AHMADI, Saad ; BA-HAMMAM, Ahmed S. Diagnostic Features and Potential Applications of PPG Signal in Healthcare : A Systematic Review. In : *Healthcare*. MDPI, 2022, t. 10, p. 547. N° 3.
12. MORAES, Jermana L ; ROCHA, Matheus X ; VASCONCELOS, Glauber G ; VASCONCELOS FILHO, José E ; DE ALBUQUERQUE, Victor Hugo C ; ALEXANDRIA, Auzuir R. Advances in photoplethysmography signal analysis for biomedical applications. *Sensors*. 2018, t. 18, n° 6, p. 1894.
13. SUN, Yu ; THAKOR, Nitish. Photoplethysmography revisited : from contact to non-contact, from point to imaging. *IEEE transactions on biomedical engineering*. 2015, t. 63, n° 3, p. 463-477.

## Bibliographie

---

14. TAMURA, Toshiyo ; MAEDA, Yuka ; SEKINE, Masaki ; YOSHIDA, Masaki. Wearable photoplethysmographic sensors—past and present. *Electronics*. 2014, t. 3, n° 2, p. 282-302.
15. BISWAS, Dwaipayan ; SIMÕES-CAPELA, Neide ; VAN HOOF, Chris ; VAN HELLEPUTTE, Nick. Heart rate estimation from wrist-worn photoplethysmography : A review. *IEEE Sensors Journal*. 2019, t. 19, n° 16, p. 6560-6570.
16. ISMAIL, Shahid ; AKRAM, Usman ; SIDDIQI, Imran. Heart rate tracking in photoplethysmography signals affected by motion artifacts : a review. *EURASIP Journal on Advances in Signal Processing*. 2021, t. 2021, n° 1, p. 1-27.
17. ZHANG, Zhilin ; PI, Zhouyue ; LIU, Benyuan. TROIKA : A General Framework for Heart Rate Monitoring Using Wrist-Type Photoplethysmographic Signals During Intensive Physical Exercise. *IEEE Transactions on Biomedical Engineering*. 2015, t. 62, n° 2, p. 522-531. Disp. à l'adr. DOI : [10.1109/TBME.2014.2359372](https://doi.org/10.1109/TBME.2014.2359372).
18. BISWAS, Dwaipayan ; SIMÕES-CAPELA, Neide ; VAN HOOF, Chris ; VAN HELLEPUTTE, Nick. Heart Rate Estimation From Wrist-Worn Photoplethysmography : A Review. *IEEE Sensors Journal*. 2019, t. 19, n° 16, p. 6560-6570. Disp. à l'adr. DOI : [10.1109/JSEN.2019.2914166](https://doi.org/10.1109/JSEN.2019.2914166).
19. GORODNITSKY, Irina F ; RAO, Bhaskar D. Sparse signal reconstruction from limited data using FOCUSS : A re-weighted minimum norm algorithm. *IEEE Transactions on signal processing*. 1997, t. 45, n° 3, p. 600-616.
20. SALEHIZADEH, Seyed MA ; DAO, Duy ; BOLKHOVSKY, Jeffrey ; CHO, Chae ; MENDELSON, Yitzhak ; CHON, Ki H. A novel time-varying spectral filtering algorithm for reconstruction of motion artifact corrupted heart rate signals during intense physical activities using a wearable photoplethysmogram sensor. *Sensors*. 2015, t. 16, n° 1, p. 10.
21. ZHOU, Menglian ; SELVARAJ, Nandakumar. Heart rate monitoring using sparse spectral curve tracing. In : *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, p. 5347-5352.
22. REISS, Attila ; INDLEKOFER, Ina ; SCHMIDT, Philip ; VAN LAERHOVEN, Kristof. Deep PPG : large-scale heart rate estimation with convolutional neural networks. *Sensors*. 2019, t. 19, n° 14, p. 3079.
23. BISWAS, Dwaipayan ; EVERSON, Luke ; LIU, Muqing ; PANWAR, Madhuri ; VERHOEF, Bram-Ernst ; PATKI, Shrishail ; KIM, Chris H. ; ACHARYYA, Amit ; VAN HOOF, Chris ; KONIJNENBURG, Mario ; VAN HELLEPUTTE, Nick. CorNET : Deep Learning Framework for PPG-Based Heart Rate Estimation and Biometric Identification in Ambulant Environment. *IEEE Transactions on Biomedical Circuits and Systems*. 2019, t. 13, n° 2, p. 282-291. Disp. à l'adr. DOI : [10.1109/TBCAS.2019.2892297](https://doi.org/10.1109/TBCAS.2019.2892297).
24. SONG, Seok Bin ; NAM, Jung Woo ; KIM, Jin Heon. NAS-PPG : PPG-Based Heart Rate Estimation Using Neural Architecture Search. *IEEE Sensors Journal*. 2021, t. 21, n° 13, p. 14941-14949. Disp. à l'adr. DOI : [10.1109/JSEN.2021.3073047](https://doi.org/10.1109/JSEN.2021.3073047).

26. GORDON, Ariel ; EBAN, Elad ; NACHUM, Ofir ; CHEN, Bo ; WU, Hao ; YANG, Tien-Ju ; CHOI, Edward. MorphNet : Fast Simple Resource-Constrained Structure Learning of Deep Networks. In : *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2018.
27. RISSO, Matteo ; BURRELLO, Alessio ; PAGLIARI, Daniele Jahier ; CONTI, Francesco ; LAMBERTI, Lorenzo ; MACII, Enrico ; BENINI, Luca ; PONCINO, Massimo. Pruning In Time (PIT) : A Lightweight Network Architecture Optimizer for Temporal Convolutional Networks. In : *2021 58th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2021. Disp. à l'adr. DOI : [10.1109/dac18074.2021.9586187](https://doi.org/10.1109/dac18074.2021.9586187).
28. TAN, Chang Wei ; BERGMEIR, Christoph ; PETITJEAN, Francois ; WEBB, Geoffrey I. *IEEEPPG Dataset*. Zenodo, 2020. Disp. à l'adr. DOI : [10.5281/zenodo.3902710](https://doi.org/10.5281/zenodo.3902710).
32. HAMILTON, P. Open source ECG analysis. In : *Computers in Cardiology*. 2002, p. 101-104. Disp. à l'adr. DOI : [10.1109/CIC.2002.1166717](https://doi.org/10.1109/CIC.2002.1166717).
33. COTTER, Shane F ; RAO, Bhaskar D ; ENGAN, Kjersti ; KREUTZ-DELGADO, Kenneth. Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Transactions on Signal Processing*. 2005, t. 53, n° 7, p. 2477-2488.
34. EILERS, Paul HC. A perfect smoother. *Analytical chemistry*. 2003, t. 75, n° 14, p. 3631-3636.
35. VASEGHI, Saeed V. *Advanced digital signal processing and noise reduction*. John Wiley & Sons, 2008.
36. SETHARES, William A ; BAÑUELOS, Diego. *Rhythm and transforms*. T. 1. Springer, 2007.
38. LECUN, Y. ; BOTTOU, L. ; BENGIO, Y. ; HAFFNER, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. 1998, t. 86, n° 11, p. 2278-2324. Disp. à l'adr. DOI : [10.1109/5.726791](https://doi.org/10.1109/5.726791).
39. BHATT, Dulari ; PATEL, Chirag ; TALSANIA, Hardik ; PATEL, Jigar ; VAGHELA, Rasmika ; PANDYA, Sharnil ; MODI, Kirit ; GHAYVAT, Hemant. CNN variants for computer vision : history, architecture, application, challenges and future scope. *Electronics*. 2021, t. 10, n° 20, p. 2470.
42. SZANDALA, Tomasz. Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. *CoRR*. 2020, t. abs/2010.09458. Disp. à l'adr. arXiv : [2010.09458](https://arxiv.org/abs/2010.09458).
44. IOFFE, Sergey ; SZEGEDY, Christian. Batch Normalization : Accelerating Deep Network Training by Reducing Internal Covariate Shift. *CoRR*. 2015, t. abs/1502.03167. Disp. à l'adr. arXiv : [1502.03167](https://arxiv.org/abs/1502.03167).
46. BAI, Shaojie ; KOLTER, J Zico ; KOLTUN, Vladlen. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv :1803.01271*. 2018.

## Bibliographie

---

47. COURBARIAUX, Matthieu ; HUBARA, Itay ; SOUDRY, Daniel ; EL-YANIV, Ran ; BENGIO, Yoshua. Binarized neural networks : Training deep neural networks with weights and activations constrained to + 1 or -1. *arXiv preprint arXiv :1602.02830*. 2016.
48. ZANGHIERI, Marcello ; BENATTI, Simone ; BURRELLO, Alessio ; KARTSCH, Victor ; CONTI, Francesco ; BENINI, Luca. Robust Real-Time Embedded EMG Recognition Framework Using Temporal Convolutional Networks on a Multicore IoT Processor. *IEEE Transactions on Biomedical Circuits and Systems*. 2020, t. 14, n° 2, p. 244-256. Disp. à l'adr. DOI : [10.1109/TBCAS.2019.2959160](https://doi.org/10.1109/TBCAS.2019.2959160).
49. CHUNG, Heewon ; KO, Hoon ; LEE, Hooseok ; LEE, Jinseok. Deep learning for heart rate estimation from reflectance photoplethysmography with acceleration power spectrum and acceleration intensity. *Ieee Access*. 2020, t. 8, p. 63390-63402.
52. LECUN, Yann ; BENGIO, Yoshua ; HINTON, Geoffrey. Deep learning. *nature*. 2015, t. 521, n° 7553, p. 436-444.

# Webographie

6. BRITISH HEART FOUNDATION. *Pulse Rate* [<https://www.bhf.org.uk/informationsupport/heart-matters-magazine/medical/ask-the-experts/pulse-rate>]. [s. d.]. Consulté le 20/06/2023.
8. SCRIPPS HEALTH. *What Heart Monitor is Best at Detecting Irregular Heartbeats?* [[https://www.scripps.org/news\\_items/6973-what-heart-monitor-is-best-at-detecting-irregular-heartbeats](https://www.scripps.org/news_items/6973-what-heart-monitor-is-best-at-detecting-irregular-heartbeats)]. [s. d.]. Consulté le 10/06/2023.
9. STAFF, Mayo Clinic. *Exercise intensity : How to measure it* [<https://www.mayoclinic.org/healthy-lifestyle/fitness/in-depth/exercise-intensity/art-20046887?pg=2>]. 2021. Consulté le 18/06/2023.
25. MICROSOFT CORPORATION. *NNI : Neural Network Intelligence* [<https://github.com/microsoft/nni>]. [s. d.]. Consulté le 10/06/2023.
29. REISS, Attila et al. *PPG-DaLiA* [<https://archive.ics.uci.edu/dataset/495/ppg+dalia>]. 2019. Consulté le 05/04/2023.
30. EMPATICA. *Empatica E4* [<https://www.empatica.com/en-eu/research/e4/>]. Consulté le 16/05/2023.
31. BIOSIGNALSPLUX. *RespiBAN Professional* [<https://www.biosignalsplex.com/en/respiban-professional>]. Consulté le 16/05/2023.
37. ANDTEM2000. *PPG\_WFPV\_TBME2017.m* [[https://github.com/andtem2000/PPG/blob/master/PPG\\_WFPV\\_TBME2017.m](https://github.com/andtem2000/PPG/blob/master/PPG_WFPV_TBME2017.m)]. 2017. Consulté le 05/04/2023.
40. SONI, Sanskar ; CHOUHAN, Satyendra Singh ; RATHORE, Santosh Singh. *Text-ConvoNet :A Convolutional Neural Network based Architecture for Text Classification*. 2022. Disp. à l'adr. arXiv : [2203.05173 \[cs.CL\]](https://arxiv.org/abs/2203.05173). Consulté le 08/05/2023.
41. NG, Andrew. *Convolutional Neural Networks* [<https://www.coursera.org/learn/convolutional-neural-networks>]. [s. d.]. Consulté le 04/05/2023.
43. SAHA, Sumit. *A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way* [<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>]. 2022. Consulté le 12/05/2023.
45. HUBER, Johann. *Batch normalization in 3 levels of understanding* [<https://towardsdatascience.com/batch-normalization-in-3-levels-of-understanding-14c2da90a338>]. 2022. Consulté le 12/05/2023.
50. EML-EDA. *Q-PPG* [<https://github.com/eml-eda/q-ppg>]. 2021. Consulté le 10/05/2023.

## Webographie

51. GOOGLE-RESEARCH. *morph-net* [<https://github.com/google-research/morph-net.git>]. 2019. Consulté le 10/05/2023.