

Une approche basée apprentissage profond et informée par la physiologie pour la segmentation de l'ECG

A. SALAMA¹ A. KACHENOURA¹ S. ALMUHAMMAD ALALI¹ G. CARRAULT¹ L. SENHADJI¹ A. KARFOUL¹

¹Univ Rennes, INSERM, LTSI - UMR 1099, F-35000 Rennes, France

Résumé – Cet article présente un pipeline supervisé pour la segmentation automatique des ondes ECG, comprenant une étape de prétraitement adaptée pour le traitement en temps réel, un modèle CNN-LSTM avec un mécanisme d'attention multi-têtes pour la segmentation, et une étape de post-traitement informée par la physiologie pour réduire le taux de faux positifs. Évaluée sur la base de données QT de PhysioNet, l'approche proposée se révèle supérieure aux méthodes existantes, en particulier pour la segmentation des ondes P et T, mettant en évidence les avantages des mécanismes d'attention et du post-traitement. Cette étude souligne également l'importance d'incorporer des critères d'évaluation supplémentaires, tels que le nombre d'ondes ECG segmentées, en plus des métriques standards de classification, afin de garantir une évaluation non biaisée de la qualité de segmentation.

Abstract – This paper presents a three-stage supervised pipeline for automatic ECG wave segmentation, consisting of a shallow preprocessing stage for real-time analysis, a CNN-LSTM model with multi-head attention for segmentation, and a physiology-driven postprocessing step to reduce false positives. Evaluated on the PhysioNet QT database, the proposed approach outperforms existing methods, particularly for P and T wave segmentation, highlighting the benefits of attention mechanisms and postprocessing. This study also emphasizes the importance of incorporating additional evaluation criteria, such as the number of segmented ECG waves, alongside standard classification metrics to ensure an unbiased segmentation assessment.

1 Introduction

La segmentation des différentes ondes ECG (P, QRS et T) et des segments associés (PR, QT, ST, etc.) joue un rôle crucial dans la détection des anomalies cardiaques et la prédiction des événements indésirables [1]. Bien que de nombreuses méthodes automatiques aient été proposées, la segmentation de l'ECG reste un défi majeur en raison de la variabilité morphologique du signal ECG et de la présence de divers types de bruits. Ces derniers peuvent altérer la précision de segmentation conduisant ainsi à un taux élevé de fausses détections. Deux familles de méthodes de segmentation de l'ECG peuvent être identifiées : les méthodes non supervisées et les méthodes supervisées. Les approches non-supervisées reposent sur des techniques classiques telles que l'algorithme de Pan-Tompkins [2], la Transformée en Ondelettes [3], la Transformée de Fourier [4], ainsi que la Décomposition en Mode Empirique (EMD) [5]. Les approches supervisées, notamment celles basées sur l'apprentissage profond, ont récemment montré un grand potentiel pour la segmentation de l'ECG. Par exemple, l'approche ECG-SegNet basée sur un modèle de réseaux de neurones type BiLSTM (Bidirectional Long Short Term Memory) a été proposée dans [6]. De plus, l'approche ConvLSTM_{SA} combinant à la fois des réseaux neuronaux de type CNN (Convolutional Neural Network), BiLSTM et une couche d'auto-attention (Self-Attention, SA) a été introduite dans [7]. Cette dernière a permis d'obtenir des performances nettement supérieures comparées aux approches précédentes. Néanmoins, même si ces méthodes récentes montrent des résultats prometteurs en termes de précision, elles comportent plusieurs limites. Certaines se focalisent uniquement sur la segmentation du complexe QRS, d'autres nécessitent des étapes de prétraitement assez coûteuses, telles que le débruitage, ce qui limite leur utilisation en temps réel. De plus, ces méthodes

considèrent souvent la segmentation comme une classification binaire des échantillons, sans prendre en compte le nombre d'ondes détectées. Ceci peut engendrer un taux de faux positifs élevé, malgré une précision quasi-parfaite dans la classification des échantillons. Afin de pallier à ces insuffisances, un nouveau pipeline en trois étapes pour la segmentation de l'ECG est proposé dans cet article : (i) une étape de prétraitement peu coûteuse adaptée pour un traitement en temps réel, (ii) une étape de segmentation basée sur un modèle CNN-LSTM qui intègre des couches d'attention multi-têtes (Multihead Attention, MA), et (iii) un post-traitement informé par les durées physiologiques des ondes P, QRS et T. Une étude comparative du pipeline proposé, noté ici, ConvLSTM_{MA}, avec une approche classique non-supervisée, ECGdeli [8] et une approche supervisée ConvLSTM_{SA} [7], en utilisant la base de données QT disponible depuis la plateforme PhysioNet est également proposée.

2 Méthodologie

2.1 Base de données

La base de données QT [9] de PhysioNet est exploitée ici pour évaluer le pipeline proposé. Cette base de données comprend 105 enregistrements ECG de quinze minutes à deux dérivations (un enregistrement par patient), échantillonnés à 250 Hz. Chaque enregistrement inclut les annotations du début, du pic et de la fin des ondes P, QRS et T. Dans cette étude, chaque enregistrement ECG a été divisé en segments de 2000 échantillons (soit une durée de 8 secondes). Le début et la fin de chaque segment ont été choisis comme appartenant à la ligne de base de l'ECG, segment reliant la fin de l'onde T d'un cycle ECG au début de l'onde P du cycle suivant. En ce qui concerne les méthodes supervisées, ConvLSTM_{SA} et

ConvLSTM_{MA}, 70 % des segments obtenus ont été utilisés pour l'entraînement des modèles, tandis que les 30 % restants ont été utilisés pour la phase de test. Une validation croisée à 5 blocs a été utilisée pour évaluer les performances. De plus, une stratégie patient indépendant a été adoptée dans un souci de généralisation des modèles. L'approche non-supervisée ECGdeli a été évaluée uniquement sur les mêmes données test, que les méthodes supervisées, afin de garantir une équité entre toutes les méthodes.

2.2 Pipeline proposé

Comme indiqué précédemment, le pipeline proposé comprend trois étapes successives :

Le pré-traitement : l'objectif de cette étape est de préparer les données afin de garantir un bon conditionnement du processus d'apprentissage du modèle de segmentation ConvLSTM_{MA}. Plus précisément, chaque segment ECG (issu de la dérivation I) est normalisé par sa valeur absolue afin de corriger les variations d'amplitude entre les segments. Ensuite, deux autres entrées sont dérivées de l'ECG (comme proposé dans [7]) : (i) une entrée obtenue par lissage cubique de chaque segment ECG. Ce lissage consiste à ajuster un polynôme cubique aux données du segment [10], et (ii) la deuxième entrée est une simple dérivée du segment ECG afin de fournir plus d'informations sur les variations lentes et rapides du signal. Il est important de noter que cette étape de pré-traitement est simple et permet un traitement quasi en temps réel, avec un temps de traitement inférieur à 0.008 secondes pour une fenêtre d'analyse de 8 secondes.

La segmentation par le modèle ConvLSTM_{MA} : l'architecture du modèle ConvLSTM_{MA} est l'association de plusieurs modèles de réseaux de neurones, comme illustré dans la Figure 1, afin de capturer efficacement les différentes caractéristiques des signaux d'entrée. Plus précisément, ce modèle intègre successivement : (i) une couche d'entrée à quatre signaux, où les deux premières entrées correspondent, respectivement, à la première et à la deuxième voies des ECG traités. La troisième et la quatrième entrées représentent, respectivement, la version lissée du premier signal ECG (issue de la dérivation I) et sa dérivée première ; (ii) une couche convolutive de 64 filtres de taille 3, permettant d'extraire les caractéristiques locales et d'améliorer la détection des instants associés aux débuts et fins des ondes P, QRS et T ; (iii) une couche BiLSTM avec 128 cellules, utilisée pour capturer les dépendances temporelles inter et intra-segments ECG, optimisant ainsi la précision de la segmentation ; (iv) une couche d'attention multi-têtes, composée de trois têtes et d'un paramètre clé de taille 128, afin de mettre en évidence les parties les plus informatives du signal ; (v) une couche BiLSTM avec 128 cellules, utilisant les poids de la couche précédente d'attention multi-têtes pour affiner la représentation des données ECG, et (vi) une couche de sortie de trois voies correspondant aux trois classes possibles (ondes P, QRS et T) qui peuvent être attribuées à chaque échantillon temporel du segment ECG.

Le post-traitement : la majorité des méthodes de segmentation de l'ECG, proposées dans la littérature, sont uniquement évaluées au travers des métriques statistiques usuelles, telles que la justesse, le rappel, la précision, la spécificité et le score F1). Cependant, bien qu'elles soient fréquemment employées, de bonnes valeurs de ces métriques ne garantissent pas né-

cessairement une bonne performance de segmentation, car elles se concentrent uniquement sur une classification binaire des échantillons (correctement ou incorrectement classés). En conséquence, elles ne renseignent pas sur le taux de fausses détections des ondes P, QRS et T. Il est important de noter que les fausses détections sont souvent dues à des artefacts isolés ou des artefacts qui masquent partiellement les ondes cibles, entraînant des segmentations incorrectes de plusieurs ondes, comme illustré dans la Figure 2b et Figure 2e. Pour minimiser le Taux de Fausses Détections (TFD), cette étude propose également un algorithme de post-traitement qui exploite les connaissances physiologiques, telles que les durées des ondes P (≤ 120 ms), QRS (70-120 ms) et T, ainsi que celles des intervalles ECG (RR, ST, etc.). Cet algorithme comporte deux étapes : (i) la première étape dite de Suppression, cherche à minimiser les surdétections de certains artefacts aigus qui sont labellisés comme événements d'intérêts (P, QRS ou T). Concrètement, dans cette étape, les événements segmentés dont la durée, notée $\mathcal{L}$, est inférieure à un seuil prédéfini $\theta_{\text{Suppression}}^{(\beta)}$, $\beta \in \{P, QRS, T\}$ sont rejetés, éliminant ainsi les artefacts susceptibles d'être classés à tort comme une des ondes ECG comme illustré dans la Figure 2c ; et (ii) la deuxième étape dite de Fusion, cherche à éliminer les fausses détections qui sont dues au fait que certaines ondes sont détectées de manière fragmentée. Dans ce cas, si deux événements segmentés successifs sont séparés par un intervalle de temps ($\mathcal{J}$) inférieur à un seuil défini $\theta_{\text{Fusion}}^{(\beta)}$, ils sont fusionnés en un seul événement (voir la Figure 2f).

Critères d'évaluation : Pour assurer une analyse exhaustive des performances des méthodes de segmentation de l'ECG, nous proposons, en plus des métriques statistiques classiques (Critère 1) de tenir compte du nombre d'ondes faussement détectées (Critère 2) :

Critère 1 : calcule les métriques standards de classification, à savoir la Justesse, le Rappel, la Précision, la Spécificité, et le score-F1. Ici on parle bien de la classification des échantillons temporels comme appartenant ou non à une onde.

Critère 2 : évalue le nombre d'ondes P, QRS et T qui ont été surdétections ou non détectées.

2.3 Résultats

Cette section présente une comparaison des performances du pipeline proposé avec deux méthodes : l'une non-supervisée (EcgDeli) et l'autre supervisée (ConvLSTM_{SA}). Afin d'assurer une comparaison équitable entre les différentes approches et de mettre en évidence l'apport de l'étape de post-traitement proposée, toutes les méthodes de segmentation considérées dans cette étude sont évaluées avec et sans cette étape de post-traitement.

En termes de classification (Critère 1), les résultats figurant dans le Tableau 1 montrent que les méthodes supervisées, à savoir ConvLSTM_{SA} et ConvLSTM_{MA}, offrent une segmentation de meilleure qualité par rapport à la méthode non-supervisée ECGDeli, sans ou avec post-traitement. Il est également à noter que le pipeline proposé ConvLSTM_{MA} est globalement plus performant que celui basé sur ConvLSTM_{SA}, en particulier pour la segmentation des ondes P et T, qui présentent des Rapports Signal sur Bruit (RSB) plus faibles par rapport à l'onde QRS. Une analyse détaillée des résultats révèle que l'amélioration apportée par le pipeline ConvLSTM_{MA}

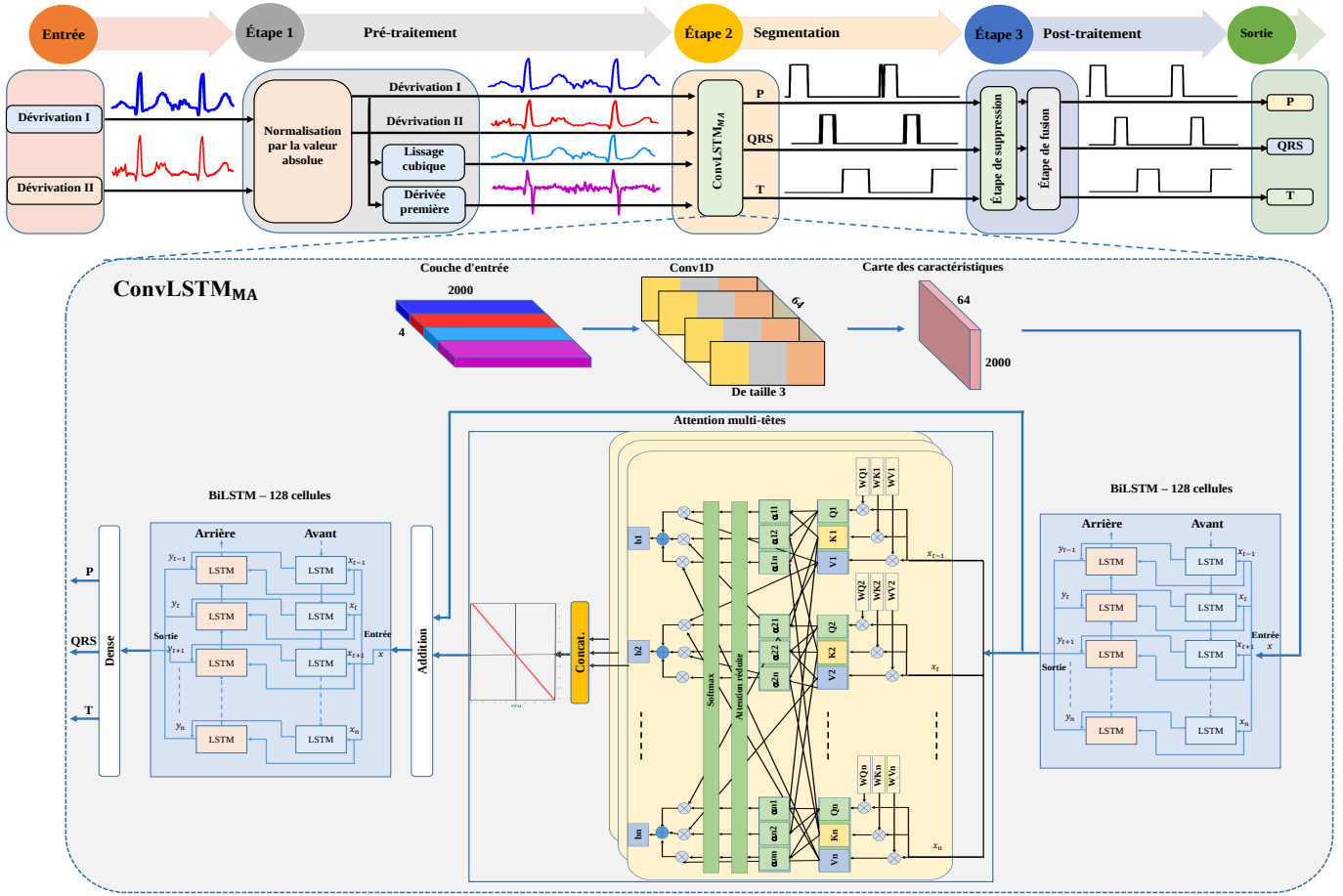


FIGURE 1 : Le pipeline à trois étapes proposé pour la segmentation de l'ECG.

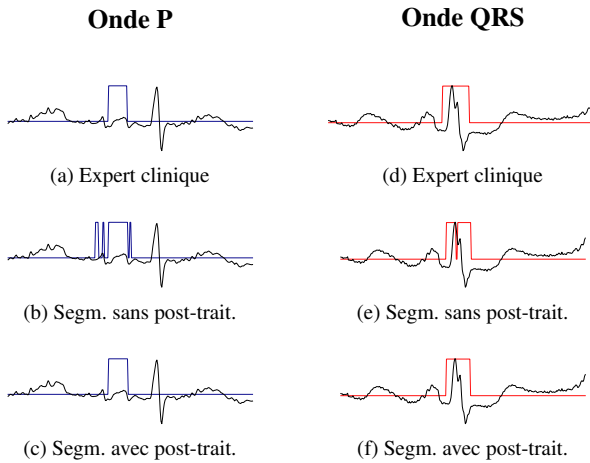


FIGURE 2 : Exemple de segmentation des ondes P et QRS : (a), (d) annotation de l'expert ; (b), (e) sans l'étape de post-traitement proposée (post-trait.) ; (c), (f) avec l'étape de post-traitement proposée.

est particulièrement marquée lorsqu'elle est évaluée à l'aide du score F1. Cette métrique est définie comme la moyenne harmonique de la Sensibilité et de la Précision. Elle est généralement utilisée pour évaluer le meilleur compromis entre ces deux indicateurs, notamment lorsque les classes sont non-balancées. Plus précisément, bien que les méthodes ECGDeli et ConvLSTM_{SA} affichent, en l'absence et en présence de l'étape de post-traitement proposée, un score F1 approximatif de 70% et 89% pour l'onde P, de 79% et 93% pour l'onde QRS, et de 75% et 88% pour l'onde T, l'approche proposée atteint des scores F1 de 91%, 94% et 91% pour les ondes P, QRS et T respectivement.

En termes du nombre d'ondes détectées (Critère 2), les méthodes supervisées ConvLSTM_{SA} et ConvLSTM_{MA} présentent, en l'absence de l'étape de post-traitement proposée, une surdéttection des ondes P, QRS et T. En comparaison à la méthode non-supervisée ECGDeli, cette surdéttection est plus marquée. Cependant, ConvLSTM_{MA} montre systématiquement une surdéttection moins importante que ConvLSTM_{SA}. En effet, pour un total de 16 368 ondes P annotées dans la base de données, ECGDeli présente une surdéttection de 11.56%, ConvLSTM_{SA} de 17.82% et ConvLSTM_{MA} de 15.64%. Pour les ondes QRS et T, dont le nombre total dans la base de données est, respectivement, de 17786 et 17709, ECGDeli montre une surdéttection de 3.3% et 3.2% pour l'onde QRS et T. En revanche, ConvLSTM_{SA} et ConvLSTM_{MA} présentent une surdéttection d'environ 20.24% et 12.85% pour l'onde QRS, et de 32.94% et 27.26% pour l'onde T. Ce phénomène de surdét-

TABLE 1 : Performances des trois pipelines en termes du Critère 1 et du Critère 2, moyennés sur 5 blocs de validation croisée.

Critère	Métrique	Modèle	Sans Post-traitement			Avec Post-traitement		
			P	QRS	T	P	QRS	T
Critère 1	Justesse	ECGDeli	0.9345	0.9511	0.8988	0.9345	0.9511	0.8989
		ConvLSTM _{SA}	0.9830	0.9850	0.9555	0.9829	0.9849	0.9561
		ConvLSTM _{MA}	0.9854	0.9879	0.9637	0.9855	0.9881	0.9641
	Rappel	ECGDeli	0.9129	0.8750	0.7614	0.9129	0.8750	0.7614
		ConvLSTM _{SA}	0.8761	0.9193	0.8590	0.8732	0.9187	0.8594
		ConvLSTM _{MA}	0.9091	0.9445	0.8986	0.9074	0.9450	0.9001
	Précision	ECGDeli	0.5760	0.7289	0.7439	0.5758	0.7289	0.7440
		ConvLSTM _{SA}	0.9175	0.9433	0.9169	0.9183	0.9433	0.9195
		ConvLSTM _{MA}	0.9147	0.9445	0.9260	0.9159	0.9454	0.9269
	Spécificité	ECGDeli	0.9364	0.9603	0.9340	0.9364	0.9603	0.9341
		ConvLSTM _{SA}	0.9925	0.9930	0.9796	0.9926	0.9930	0.9802
		ConvLSTM _{MA}	0.9924	0.9933	0.9799	0.9925	0.9934	0.9801
	Score F1	ECGDeli	0.70586	0.79438	0.75212	0.7057	0.7943	0.7521
		ConvLSTM _{SA}	0.8949	0.9302	0.8858	0.8936	0.9299	0.8872
		ConvLSTM _{MA}	0.9116	0.9443	0.9116	0.9113	0.9450	0.9128
Critère 2	Nombre d'ondes	Annotation de l'expert	16368	17786	17709	16368	17786	17709
		ECGDeli	18261	18374	18279	18060	18374	18243
		ConvLSTM _{SA}	19286	21385	23543	16521	18192	18449
		ConvLSTM _{MA}	18928	20071	22537	16889	18188	18562

tection est largement atténué grâce à l'utilisation de l'étape de post-traitement proposée, comme le montre le Tableau 1, et ce, d'autant plus pour les approches supervisées qui présentent, au final, une surdéttection moins importante que celle de la méthode non-supervisée ECGDeli. Plus précisément, après le post-traitement, la méthode ECGDeli affiche un taux de surdéttection de 10.34%, 3.3% et 3% pour l'onde P, QRS et T, respectivement. L'approche supervisée ConvLSTM_{SA}, quant à elle, présente un taux de surdéttection de 3.2%, 2.26% et 4.8% pour des ondes P, QRS et T, respectivement. Enfin, notre approche, ConvLSTM_{MA}, affiche des taux de surdéttection, globalement plus faibles, de 0.93%, 2.28% et 4.1% pour le modèle proposé. Les résultats obtenus confirment l'utilité de l'étape de post-traitement proposée pour pallier au problème des fausses détections des méthodes basées sur l'apprentissage profond, et ce, malgré leur très bonne performance en termes de classification. Notons aussi, que le pipeline proposé (ConvLSTM_{MA} + post-traitement) offre des performances significativement plus élevées pour les deux critères évalués.

3 Conclusion

Dans cet article, un nouveau pipeline en trois étapes basé sur l'apprentissage profond a été proposé pour la segmentation automatique de l'ECG. Ce modèle comprend une étape de pré-traitement suivie par une étape de segmentation en utilisant le modèle ConvLSTM_{MA} proposé, et enfin une étape de post-traitement informée par la physiologie afin de pallier au problème des faux positifs. Les résultats obtenus ont montré que le modèle ConvLSTM_{MA} offre de meilleures performances comparées aux méthodes ECGDeli [8] et ConvLSTM_{SA} [7], en particulier pour la segmentation des ondes P et T, qui sont généralement caractérisées par un faible RSB. De plus, l'approche proposée met en évidence l'intérêt du mécanisme d'attention multi-têtes pour attribuer différents niveaux d'importance aux caractéristiques du signal ECG. Par ailleurs, cette étude a montré les limites de l'utilisation seule des métriques classiques de classification, qui peuvent conduire à des conclusions biaisées sur l'efficacité des méthodes de segmentation

évaluées. Cela est particulièrement vrai pour les modèles d'apprentissage profond, qui souffrent généralement d'un taux non négligeable de surdéttections. Ainsi, l'intégration de critères d'évaluation complémentaires, tels que le nombre d'ondes segmentées, s'avère essentielle pour une évaluation plus robuste. Les résultats ont également montré que l'intégration de l'étape de post-traitement proposée, informée par des connaissances physiologiques, permet de réduire significativement le taux de fausses détections. Les travaux futurs se concentreront sur l'évaluation du pipeline proposé sur des bases de données réelles de plus grande taille, ainsi que sur une exploration approfondie de l'explicabilité du modèle d'apprentissage profond proposé.

Remerciements

Ce travail a été financé en partie par le PEPR Santé Numérique ; Projet DIIP-HEART : ANR-22-PESN-0018 et en partie par les projets PrepRisC et DeMUG du programme ARED de la Région Bretagne, France.

Références

- [1] A. L. Goldberger, Z. D. Goldberger et A. Shvilkin, *Clinical electrocardiography : a simplified approach*, 8^e éd. Elsevier Health Sciences, 2012.
- [2] J. Pan et W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, n° 3, pp. 230–236, 1985.
- [3] G. Lenis, N. Pilia, T. Oesterlein, A. Luik, C. Schmitt et O. Dössel, "P wave detection and delineation in the ECG based on the phase free stationary wavelet transform and using intracardiac atrial electrograms as reference," *Biomed. Eng. Biomed. Tech.*, vol. 61, n° 1, pp. 37–56, 2016.
- [4] I. S. N. Murthy et U. C. Niranjan, "Component wave delineation of ECG by filtering in the fourier domain," *Med. Biol. Eng. Comput.*, vol. 30, pp. 169–176, 1992.
- [5] H. N. Abderahman, H. R. Dajani et V. Z. Groza, "Adaptive r-peak detector in extreme noise using emd selective analyzer," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, 2022, pp. 1–6.
- [6] H. Abrishami, C. Han, X. Zhou, M. Campbell et R. Czonek, "Supervised ECG interval segmentation using LSTM neural network," in *Proc. Int. Conf. Bioinformatics Comput. Biol. (BIOCOMP)*, 2018, pp. 71–77.
- [7] A. Malali, S. Hiriannaiah, G. M. Siddesh, K. G. Srinivasa et N. T. Sanjay, "Supervised ECG wave segmentation using convolutional LSTM," *ICT Express*, vol. 6, n° 3, pp. 166–169, 2020.
- [8] N. Pilia, C. Nagel, G. Lenis, S. Becker, O. Dössel et A. Loewe, "ECGdeli-an open source ECG delineation toolbox for MATLAB," *SoftwareX*, vol. 13, p. 100639, 2021.
- [9] P. Laguna, R. G. Mark, A. L. Goldberger et G. B. Moody, "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG," *Comput. Cardiol.*, vol. 24, pp. 673–676, 1997.
- [10] D. S. G. Pollock, "Smoothing with cubic splines," Queen Mary and Westfield College, Department of Economics, 1993.