

# Un filtrage basé apprentissage profond pour le débruitage des signaux de vibrations cardiaques

Salman ALMUHAMMAD ALALI<sup>1</sup> Amar KACHENOURA<sup>1</sup> Lotfi SENHADJI<sup>1</sup> Alfredo I. HERNANDEZ<sup>1</sup> Cindy MICHEL<sup>2</sup>  
Laurent ALBERA<sup>1</sup> Ahmad KARFOUL<sup>1</sup>

<sup>1</sup>Univ. Rennes, Inserm, LTSI-UMR 1099, F-35000 Rennes, France

<sup>2</sup>CardiaMetrics, La Tronche, Grenoble, France

**Résumé** – Cette étude présente une approche efficace pour le débruitage des signaux sonores cardiaques obtenus par un implant innovant situé au niveau du fundus gastrique, destiné à la surveillance longitudinale dans le contexte de l’insuffisance cardiaque. Cette méthode exploite la capacité des réseaux de neurones convolutifs à filtrer efficacement le bruit en utilisant des noyaux capables de capturer des informations utiles, à la fois des signaux d’intérêts et des artefacts. L’idée originale est la réutilisation d’un des noyaux du modèle pré-entraîné, initialement conçu pour la classification de l’insuffisance cardiaque, comme filtre pour le débruitage des signaux de vibrations cardiaques. Les résultats obtenus, sur une base de données acquise lors d’une expérimentation préclinique, démontrent l’efficacité de l’approche proposée en comparaison avec des techniques de débruitage standards, telles que les ondelettes et la décomposition en modes empiriques.

**Abstract** – This study presents an efficient approach for denoising heart sound signals obtained by an innovative implant located at the gastric fundus, intended for longitudinal heart monitoring in the context of heart failure. This method takes advantage of the ability of CNN filters to distinguish noise from meaningful signals, ensuring the preservation of relevant information. Specifically, it repurposes one of the kernels of a pre-trained 1D-CNN model, originally designed for heart failure classification, as a filter to denoise cardiac vibration signals. The obtained results validated the effectiveness of the proposed approach in comparison to standard denoising techniques, such as wavelet and empirical mode decomposition.

## 1 Introduction

l’Insuffisance Cardiaque (IC) est un syndrome clinique provoqué par des anomalies structurelles ou fonctionnelles du cœur, entraînant un débit cardiaque insuffisant (1). L’IC touche plus de 15 millions de vies en Europe (1; 2). La prévalence et l’impact de l’IC devraient aussi s’aggraver en raison du vieillissement de la population mondiale (3). Le suivi à distance et quotidien des principaux biomarqueurs cardiaques apparaît comme une solution adéquate pour les personnes à risque et pour les patients atteints d’IC (4). La fonction cardiaque repose sur deux mécanismes interconnectés : électrique et mécanique. l’ElectroCardioGramme (ECG), couramment utilisé pour diagnostiquer les anomalies cardiaques, demeure insuffisant pour détecter certains syndromes lorsqu’il est utilisé seul. Par exemple, les maladies causées par des défauts mécaniques, tels que des valves cardiaques défectueuses, nécessitent l’analyse des signaux de PhonoCardioGramme (PCG) pour un diagnostic efficace (5). Le PCG capte les sons du cœur pendant la contraction systolique et la relaxation diastolique, générant deux composantes principales : le premier bruit du cœur (son cardiaque S1), résulte de la fermeture des valves mitrale et tricuspide, et le deuxième bruit du cœur (son cardiaque S2) est produit par la fermeture des valves aortique et pulmonaire. L’analyse de ces deux sons est essentielle pour le diagnostic de l’IC. Traditionnellement, les signaux PCG sont enregistrés de manière non invasive via un microphone placé sur la poitrine. Cependant, bien que le PCG soit efficace pour la surveillance et le diagnostic à court terme, son mode d’acquisition non invasif présente des limites pour une surveillance continue à long terme de l’IC. Dans ce contexte, les

signaux des vibrations cardiaques, en particulier les signaux d’ACCélérométrie cardiaque (ACC), ont émergé comme une source d’information précieuse (6).

Récemment, notre équipe de recherche a contribué au développement d’un dispositif implantable innovant, minimalement invasif, au niveau du fundus gastrique (6). Une chaîne de traitement des signaux ACC 3D acquis par l’implant a démontré la faisabilité d’une surveillance à long terme de la fonction cardiaque et d’une détection précoce de l’IC en utilisant un tel dispositif implantable (6). Cependant, même si les résultats préliminaires sont encourageants, les artefacts et les bruits liés au site gastrique ont eu un impact significatif sur l’efficacité de l’analyse des signaux ACC et constituent une limitation de cette étude (6). En effet, les fortes amplitudes de ces bruits et artefacts peuvent complètement noyer les événements cardiaques d’intérêt, S1 et S2, et rendre leur analyse très difficile. Pour une exploitation optimale du nouveau dispositif implantable, avoir des méthodes de débruitage des signaux ACC capables de rehausser au mieux les événements d’intérêts, S1 et S2 est plus que nécessaire.

Plusieurs méthodes de débruitage de signaux monovoie ont été proposées dans la littérature pour le débruitage des signaux de vibrations cardiaques. Parmi les plus utilisées, on peut citer celles basées sur les ondelettes (7; 8) et les techniques qui exploitent la Décomposition Modale Empirique(EMD) (9). Cependant, pour les approches ondelettes, l’ondelette mère, la méthode de seuillage, et l’échelle appropriée sont des paramètres qui peuvent fortement être dépendants des données analysées (10). De même, pour l’EMD, le choix du niveau de décomposition (nombre des composantes AM-FM, appelées IMFs) et des IMFs à retenir n’est pas toujours évident (11).

Dans cette étude, nous proposons une approche, exploitant la capacité des Réseaux de Neurones Convolutifs (CNN) à filtrer efficacement le bruit en utilisant divers noyaux capables de capturer des informations utiles, à la fois des signaux d'intérêts et des bruits. Les performances de cette dernière sont évaluées sur des données réelles de sept porcs (quatre sains et trois atteints d'IC), acquises par le nouvel implant, et sont comparées à celles obtenues par les ondelettes et l'EMD.

## 2 Base de données

L'évaluation de la méthode proposée a été menée sur des signaux ACC acquis lors d'une expérimentation préclinique à l'aide de l'implant gastrique sur sept porcs (quatre sains et trois atteints d'IC) (6). L'implant permet de recueillir des données ACC sur trois axes orthogonaux ( $ACC_x$ ,  $ACC_y$ ,  $ACC_z$ ). Pendant 14 jours, un enregistrement d'une durée de 30 secondes par heure a été réalisé avec une fréquence d'échantillonnage de 4 KHz. En raison de problèmes techniques ayant perturbé les acquisitions programmées, seulement 999 enregistrements (de 30 secondes) ont été retenus au final. De plus, ce jeu de données est inégalement réparti entre les animaux : le nombre de segments retenus pour les porcs sains est d'environ 4 fois plus grand que celui des porcs malades. Plus d'information sur le nombre de segments retenus par porc sont données dans le Tableau 1.

## 3 Méthodologie

Comme mentionné dans l'introduction, les réseaux de neurones convolutifs, CNNs, sont reconnus pour leur capacité, lors de la phase d'entraînement, à éliminer efficacement le bruit grâce aux divers noyaux capables de capturer des informations contextuelles liées à la fois aux signaux d'intérêts, aux bruits et aux autres artefacts (12). Partant de ce principe, nous proposons une méthode de débruitage qui comprenant deux étapes principales. La première étape consiste à construire un modèle CNN uni-dimensionnelle (1D-CNN), le moins complexe possible, qui offre la meilleure performance en termes de classification entre les animaux sains et ceux présentant une IC. Dans la deuxième étape, une recherche exhaustive est faite pour la sélection du noyau pré-entraîné qui satisfait deux critères bien définis : i) avoir un taux d'acceptation le plus élevé des données enregistrées, et ii) maximiser le Rapport Signal sur Bruit (RSB) des ondes S1 et S2 en sortie. Plus de détails sur le calcul de ces deux critères seront donnés par la suite.

### 3.1 Étape 1 : apprentissage du modèle 1D-CNN

Afin de définir l'architecture du modèle 1D-CNN proposé, plusieurs architectures ont été évaluées. Celle offrant le meilleur compromis classification / complexité a été retenue (voir Figure 1). Cette dernière est composée de deux couches 1D-CNN avec 16 et 32 noyaux chacun est de taille 3, respectivement. Les deux couches sont suivies d'activation unité de rectification linéaire (ReLU, pour Rectified Linear Unit) complétées par une couche de normalisation. Quant à la couche finale, elle est composée d'une couche de mise en commun moyenne

(average pooling), d'une couche entièrement connectée, terminée par une activation exponentielle normalisée (Softmax). Comme indiqué dans le Tableau 1, le jeu de données est limité et inégalement distribué entre les porcs (par exemple, 50 enregistrements pour le porc 7 contre 316 pour le porc 4). Pour remédier à ce problème de classes déséquilibrées, une augmentation des données a été appliquée. Plus précisément, une fenêtre glissante de 20 secondes avec un pas de décalage d'une seconde a été appliquée sur chaque segment de 30 secondes. Ainsi, 10 segments distincts de 20 secondes sont générés à partir de chaque enregistrement original, multipliant par dix la taille du jeu de données. De plus, et afin de garantir un apprentissage robuste, fiable et généralisable, un schéma de validation croisée de type Leave-One-pig-Out (LOO) est adopté. Plus précisément, le modèle 1D-CNN est entraîné en utilisant des données de six porcs, le porc restant servant d'ensemble de test indépendant. Ce processus est répété de manière à ce que chaque porc soit utilisé une fois comme ensemble de test, ce qui donne un total de sept modèles entraînés. Les performances de classification du modèle 1D-CNN proposé, évaluées en termes de Sensibilité (Se), de Spécificité (Sp), de F-score, et de Précision, sont respectivement de 0,9813, 0,9935, 0,9844 et 0,9893. Ces résultats démontrent la capacité du modèle proposé à extraire des caractéristiques robustes et fiables à partir de signaux ACC très bruyants, afin de distinguer précisément entre les classes saines et HF.

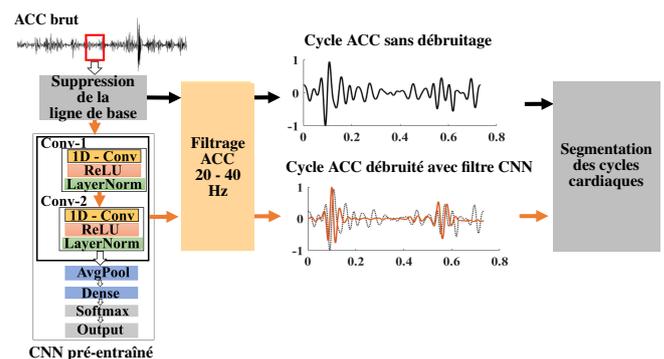


FIGURE 1 : Une visualisation comparative illustre l'impact du débruitage sur un signal ACC, se concentrant sur un seul cycle cardiaque. Le signal original est représenté en noir, tandis que la version débruitée, traitée à l'aide de filtres CNN pré-entraînés, est affichée en orange.

### 3.2 Étape 2 : sélection du meilleur noyau pré-entraîné

Comme mentionné précédemment, l'idée clé est d'utiliser le modèle 1D-CNN pré-entraîné comme filtre de débruitage tout en préservant la majeure partie des informations relatives aux ondes S1 et S2. Pour ce faire, une identification d'un noyau optimal au sein de ses deux couches convolutives 1D est nécessaire. Une stratégie de recherche gloutonne, sur l'ensemble des noyaux pré-entraînés du modèle, est adoptée dans notre cas : le modèle CNN pré-entraîné est appliqué sur toute la base de données ACC. Le RSB de S1 et S2, et le nombre de données retenus (définis dans la section 4) sont calculés à la sortie de chaque noyau. Le filtre offrant le meilleur compromis entre la préservation du plus grand nombre possible d'enregistrements

et le meilleur RSB est alors choisi comme filtre de débruitage optimal (Figure 1).

## 4 Critères d'évaluation

Les performances de la méthode proposée ont été évaluées au travers de deux critères qui visent à sélectionner uniquement les enregistrements présentant un RSB acceptable pour une segmentation robuste de S1 et S2. **Critère 1** Le premier critère est le taux d'acceptation des données enregistrées. Un enregistrement de 30 secondes est dit "accepté" si i) le rapport entre les amplitudes maximales des ondes S1 et S2 et l'amplitude maximale du bruit de fond (figure 2) est supérieur ou égal à 2, ii) au moins deux cycles ACC ont une corrélation supérieure à 0.6 (voir (6) pour plus de détails). **Critère 2** Le deuxième critère calcule l'écart du RSB de S1 et S2 avant et après débruitage. Soient  $RSB_{bruit}^\theta$  et  $RSB_{débruité}^\theta$  les valeurs de RSB du segment  $\theta \in \{S1, S2\}$  avant et après débruitage. Le RSB pour le  $n$ -ième cycle est défini comme suit :

$$RSB_{\xi,n}^\theta = 10 \log_{10} \left( \frac{\max((z_{\xi,n}^\theta)^{\odot 2})}{\frac{1}{T_{bruit}} \sum_{i=1}^{T_{bruit}} |u_{\xi,n}[i]|^2} \right) \quad (1)$$

où  $\xi \in \{\text{bruit, débruité}\}$ ,  $z_{\xi,n}^\theta$  est l'événement  $\theta$  sous forme de vecteur, et  $T_{bruit}$  représente la durée du bruit (figure 2). Pour chaque enregistrement,  $RSB_{\xi}^\theta$  est moyenné sur tous les cycles, et le RSB global du jeu de données est calculé comme la moyenne sur tous les enregistrements acceptés. L'indice d'amélioration du RSB  $\eta^\theta$  est calculé comme suit :

$$\eta^\theta = \frac{RSB_{débruité}^\theta - RSB_{bruit}^\theta}{|RSB_{bruit}^\theta|} \times 100 \quad (2)$$

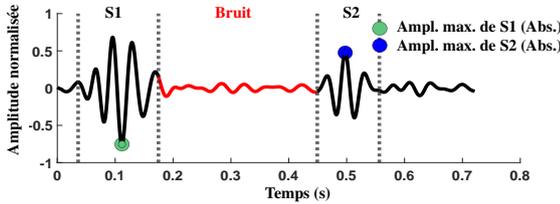


FIGURE 2 : Définition du bruit de fond (segment en rouge) utilisé pour le calcul des deux critères d'évaluation.

## 5 Résultats

Dans cette section l'impact de la méthode proposée sur le taux de données acceptées et sur l'amélioration des RSB des ondes S1 et S2 est évalué. De plus, une étude comparative avec deux méthodes classiques, une basée ondelette et l'autre exploitant l'EMD (largement utilisées dans la littérature pour le rehaussement des ondes S1 et S2) est aussi proposée. Pour une étude équitable, les paramètres de ces deux méthodes ont été fixés (de manière exhaustive) comme ceux offrant le meilleur compromis des deux critères évalués. Ainsi, l'ondelette utilisée ici est une ondelette biorthogonale avec cinq niveaux de décomposition. Un seuillage dit de l'estimation du risque sans biais de Stein (13) couplé à une stratégie de seuillage doux a aussi

été appliqué. Concernant l'EMD, huit IMFs ont été initialement extraites, les signaux débruités sont reconstruits par une somme d'un certain nombre d'IMFs (trois IMFs parmi les huit dans notre cas) qui offrent les meilleures performances.

Une inspection visuelle de chaque méthode de débruitage est présentée dans la Figure 3. Cette figure illustre un exemple de débruitage d'un ensemble de cycles cardiaques de signaux ACC, segmentés d'un enregistrement de 30 secondes de l'axe Y, ( $ACC_y$ ). Pour montrer tout l'enregistrement de 30 secondes, les cycles cardiaques ont été segmentés, synchronisés et empilés. Clairement, la méthode proposée offre un meilleur rehaussement des sons S1 et S2, même si le résultat obtenu par ondelettes semble satisfaisant. De plus, le bruit de fond, entre S1 et S2 est beaucoup plus atténué par la méthode basée CNN, ce qui permet une meilleure segmentation des événements cardiaques d'intérêt S1 et S2.

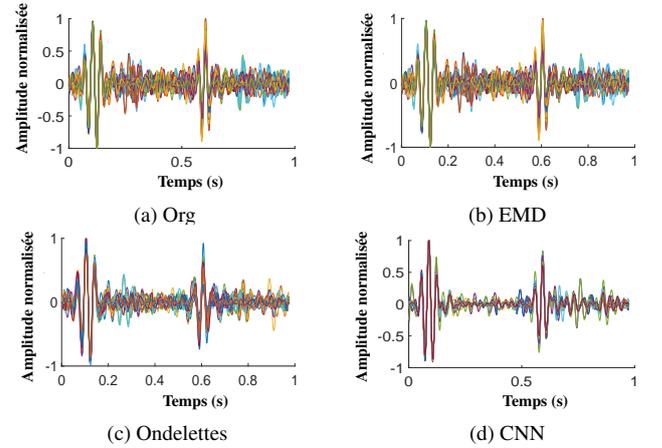


FIGURE 3 : Exemple de débruitage d'un ensemble de cycles cardiaques ACC extraits d'un enregistrement de 29 secondes sur l'axe Y.

### 5.1 Impact du débruitage sur le taux d'acceptation des données

Le Tableau 1 résume les taux d'acceptation des enregistrements (critère 1) avant et après débruitage. Avant débruitage (Org) le taux d'acceptation est de 85,98 %. Les méthodes EMD et ondelettes n'améliorent pas ce taux d'acceptation (84,78 % et 81,98 %, respectivement). En revanche, la méthode de débruitage basée CNN atteint un taux d'acceptation significativement plus élevé de 90,69 %. Ceci est essentiellement dû au fait que la méthode basée CNN atténue drastiquement le bruit de fond (comme nous l'avons montré dans la Figure 3). Pour les deux autres méthodes, même si elles arrivent à débruiter les ondes S1 et S2, elles semblent être moins efficaces dans le filtrage du bruit de fond entre ces deux ondes.

### 5.2 Impact du débruitage sur le RSB

Le Tableau 2 montre la moyenne et l'écart-type des valeurs de RSB (Critère 2) pour les événements S1 et S2 sur les trois axes ACC, avant et après débruitage. La stratégie de débruitage proposée surpasse significativement les deux méthodes classiques (ondelettes et EMD), comme l'indiquent les valeurs

TABLE 1 : Distribution des enregistrements acceptés par porc avec et sans débruitage.

Classe	ID Porc	Total enreg.	Enregistrements acceptés par méthode			
			Org	EMD	Ondelettes	CNN
Sain	1	95	81	81	81	<b>87</b>
	2	163	132	135	135	<b>144</b>
	3	232	189	191	188	<b>217</b>
	4	316	288	280	283	<b>311</b>
IC	5	70	57	52	47	<b>46</b>
	6	73	66	64	56	<b>53</b>
	7	50	46	44	35	<b>48</b>
<b>Total</b>	<b>Tous</b>	<b>999</b>	<b>859</b>	<b>847</b>	<b>819</b>	<b>906</b>
<b>Accept (%)</b>	<b>-</b>	<b>-</b>	<b>85,98</b>	<b>84,78</b>	<b>81,98</b>	<b>90,69</b>

du RSB moyennes et le taux d'amélioration des données acceptées, calculées pour S1 et S2 et pour les trois axes d'ACC. Par exemple, sans débruitage sur l'axe Y, les valeurs moyennes du RSB pour S1 et S2 sont respectivement de 15,4 dB et 11,4 dB. Après débruitage par la méthode basée CNN, le RSB (taux d'amélioration) est de 18,8 dB (21,6 %) pour S1 et de 13,2 dB (16 %) pour S2. Concernant les approches Ondelettes et EMD, l'amélioration du RSB n'est pas significative avec des taux allant de 1,2 % à 4,6 %.

TABLE 2 : Moyenne  $\pm$  écart-type du RSB et taux d'amélioration ( $\eta$ ) pour les axes ACC.

Méthode	ACC S1			ACC S2		
	X	Y	Z	X	Y	Z
<b>Sans débruitage</b>	15,2 $\pm$ 5,9	15,4 $\pm$ 5,6	15,6 $\pm$ 5,9	10,8 $\pm$ 5,7	11,4 $\pm$ 6,1	11,2 $\pm$ 5,8
<b>EMD</b>	15,4 $\pm$ 5,4	15,7 $\pm$ 5,5	16,0 $\pm$ 5,6	11,1 $\pm$ 5,8	11,7 $\pm$ 5,8	11,6 $\pm$ 5,7
$\eta$ (%)	+1,2	+2,0	+3,0	+3,0	+2,6	+3,2
<b>Ondelettes</b>	15,6 $\pm$ 5,6	15,8 $\pm$ 5,8	15,9 $\pm$ 5,8	11,3 $\pm$ 5,9	11,8 $\pm$ 6,1	11,5 $\pm$ 6,0
$\eta$ (%)	+2,5	+2,9	+2,0	+4,6	+3,5	+2,9
<b>CNN</b>	<b>17,7 <math>\pm</math> 8,0</b>	<b>18,8 <math>\pm</math> 8,5</b>	<b>17,4 <math>\pm</math> 7,6</b>	<b>12,3 <math>\pm</math> 8,1</b>	<b>13,2 <math>\pm</math> 9,1</b>	<b>12,2 <math>\pm</math> 7,6</b>
$\eta$ (%)	<b>+17,1</b>	<b>+21,6</b>	<b>+11,3</b>	<b>+13,5</b>	<b>+16,0</b>	<b>+8,7</b>

## 6 Conclusion

Une nouvelle méthodologie, probablement la première du genre, exploitant des modèles CNN initialement entraînés à des fins de classification, a été proposée afin de relever les défis de débruitage dans le contexte de la surveillance longitudinale des maladies cardiaques chroniques. Plus précisément, un filtre, dit optimal du modèle CNN pré-entraîné, est réexploité pour le débruitage des signaux ACC acquis par un dispositif innovant, minimalement invasif, implanté au niveau du fundus gastrique. Les résultats obtenus ont montré que la méthode de débruitage proposée améliore significativement le RSB des deux ondes d'intérêt S1 et S2. D'un point de vue clinique, cette amélioration présente un double intérêt. Le premier est de permettre une exploitation optimale du nouveau implant en minimisant le pourcentage de données non exploitables. Le deuxième consiste à une segmentation plus précise des ondes S1 et S2 qui permettra un meilleur diagnostic des pathologies étudiées.

## Remerciements

Cette étude a été financée en partie par le projet ANR-DiGS, ANR-18-CE19-0008, et par le projet DeMUG du programme ARED de la Région Bretagne, France.

## Déclaration éthique

Le protocole d'étude animale a été examiné et approuvé par un comité d'éthique conformément aux directives du Ministère Français de la Recherche.

## Références

- [1] F. P. Hessel, "Overview of the socio-economic consequences of heart failure," *Cardiovasc. Diagn. Ther.*, vol. 11, n° 1, p. 254, 2021.
- [2] B. Shahim, C. J. Kapelios, G. Savarese et L. H. Lund, "Global public health burden of heart failure : an updated review," *Card. Fail. Rev.*, vol. 9, p. e11, 2023.
- [3] G. M. C. Rosano, P. Seferovic, G. Savarese, I. Spoletini, Y. Lopatin, F. Gustafsson, A. Bayes-Genis, T. Jaarsma, M. Abdelhamid, A. G. Miqueo *et al.*, "Impact analysis of heart failure across European countries : an ESC-HFA position paper," *ESC Heart Fail.*, vol. 9, n° 5, pp. 2767–2778, 2022.
- [4] H. Saner, N. Schuetz, P. Buluscek, G. Du Pasquier, G. Ribaudo, P. Urwyler et T. Nef, "Case report : ambient sensor signals as digital biomarkers for early signs of heart failure decompensation," *Front. Cardiovasc. Med.*, vol. 8, p. 617682, 2021.
- [5] K. Sørensen, "Seismocardiography : Interpretation and clinical application," Ph.D. Thesis, 2021.
- [6] H. Areiza-Laverde, C. Dopierala, L. Senhadji, F. Boucher, P. Y. Gumery et A. Hernández, "Analysis of cardiac vibration signals acquired from a novel implant placed on the gastric fundus," *Front. Physiol.*, vol. 12, p. 748367, 2021.
- [7] P. Varady, "Wavelet-based adaptive denoising of phonocardiographic records," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2, pp. 1846–1849, 2001.
- [8] P. K. Jain et A. K. Tiwari, "An algorithm for automatic segmentation of heart sound signal acquired using seismocardiography," in *Proc. Int. Conf. Syst. Med. Biol.*, 2016, pp. 157–161.
- [9] J. Fleureau, J.-C. Nunes, A. Kachenoura, L. Albera et L. Senhadji, "Turning tangent empirical mode decomposition : a framework for mono-and multivariate signals," *IEEE Trans. Signal Process.*, vol. 59, n° 3, pp. 1309–1316, 2010.
- [10] G. R. Sahoo, J. H. Freed et M. Srivastava, "Optimal wavelet selection for signal denoising," *IEEE Access*, 2024.
- [11] J. Zheng et H. Pan, "Mean-optimized mode decomposition : An improved EMD approach for non-stationary signal processing," *ISA Trans.*, vol. 106, pp. 392–401, 2020.
- [12] N. V. N. Rodrigues, L. R. Abramo et N. S. T. Hirata, "The information of attribute uncertainties : what convolutional neural networks can learn about errors in input data," *Mach. Learn. Sci. Technol.*, vol. 4, n° 4, p. 045019, 2023.
- [13] D. L. Donoho et I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Am. Stat. Assoc.*, vol. 90, n° 432, pp. 1200–1224, 1995.