Deep learning-based approach for denoising heart vibration signals

Salman Almuhammad Alali[1], Amar Kachenoura[1], Lotfi Senhadji[1], Alfredo I. Hernandez[1], Cindy Michel[2], Laurent Albera[1] and Ahmad Karfoul[1]

[1] Univ. Rennes, Inserm, LTSI-UMR 1099, F-35000 Rennes, France
[2] CardiaMetrics, La Tronche, Grenoble, France

Abstract

This study proposes a novel approach for the denoising of heart sound signals obtained by an implant located at the gastric fundus site. The implant is intended for longitudinal heart monitoring in the context of heart failure. This method leverages the capability of CNN filters to differentiate between noise and meaningful signals, thereby ensuring the preservation of pertinent information. Specifically, one of the kernels of a pretrained 1D-CNN model, originally designed for heart failure classification, is repurposed as a filter to denoise cardiac vibration signals. The obtained results confirm the effectiveness of the proposed approach in comparison to standard denoising techniques, such as wavelet and empirical mode decomposition.

1. Introduction

Heart Failure (HF) is a clinical syndrome caused by the heart's inability to pump sufficient blood due to structural or functional cardiac abnormalities [1]. HF affects over 64 million people worldwide [2]. It is expected to worsen in the coming years with the aging population [3]. Regular remote monitoring of key cardiac biomarkers could be helpful for at-risk individuals and HF patients for early detection or prevention of complications [4]. Heart function relies on two interconnected mechanisms, namely the electrical and the mechanical. Electrocardiography, commonly used as a cardiac diagnostic tool, remains insufficient to detect certain pathological conditions. For instance, mechanical dysfunction, such as valvular disorders, requires PhonoCardioGram (PCG) signal analysis for effective diagnosis [5]. PCG captures heart sounds during systolic contraction and diastolic relaxation, thereby generating two key sounds. The first heart sound (S1) is produced by the closure of the atrioventricular valves, while the subsequent sound (S2) is generated by the closure of the semilunar valves. Analysis of these sounds is important to HF diagnosis [6]. Conventionally, PCG is acquired noninvasively through chest-mounted microphones. However, while PCG is effective in short-term monitoring and diagnosis, its non-invasive acquisition method imposes limitations for long-term monitoring of HF. In this context, cardiac vibration signals, particularly cardiac ACCeleration (ACC) signals, have emerged as a valuable source of information for cardiac diagnosis [6, 7].

Recently, our research team contributed to the development of a pioneering mini-invasive cardiac implantable device in the gastric fundus [6]. Furthermore, a processing pipeline for 3D ACC signals acquired by this implant demonstrated the feasibility of long-term monitoring of heart function and early detection of HF using such a device. Nonetheless, despite the promising preliminary results, artifacts and noises related to the gastric site had a significant impact on the efficiency of ACC signal analysis and constituted the main limitation of this study [6]. Indeed, the high amplitudes of these noises and artifacts have the potential to completely obscure the cardiac events of interest, S1 and S2, thereby making their analysis very difficult. To improve the implant's clinical potential, developing effective ACC signal denoising methods that can effectively enhance the clarity of events of interest, S1 and S2, is mandatory.

A plethora of signal denoising methods have been proposed in the literature. For example, wavelet-based approaches [8] and techniques that exploit Empirical Mode Decomposition (EMD) [9] are among the most widely used. Nevertheless, such methods still present several limitations. For wavelet approaches, the mother wavelet, the thresholding method, and the appropriate scale are parameters that can be highly dependent on the data analyzed [8]. Similarly, for EMD, the choice of decomposition level (number of AM-FM components, termed Intrinsic Mode Functions (IMFs)) and which IMFs to retain is not always evident [9]. In this study, an approach exploiting Deep Neural Networks (DNNs) was proposed for the denoising of ACC signals acquired by the implant. More precisely, the approach leverages the ability of Convolutional Neural Networks (CNNs) to effectively filter out noise using various kernels able to capture useful information patterns about both signals of interest and noise artifacts. The performance of the proposed approach is evaluated on real data from seven pigs (4 healthy and 3 with HF) acquired via the novel implant, and compared to results obtained using wavelets and EMD.

2. Dataset

The proposed method evaluation was conducted on ACC signals acquired during a preclinical experiment using the gastric implant on seven pigs (4 healthy and 3 with HF) [6]. The implant gathers ACC data along three orthogonal axes (ACC_x, ACC_y, ACC_z). Over a period of 14 days, 30-second recordings were captured hourly. The first second of each recording is skipped due to amplifier effects (see [6] for more details). In addition, due to technical issues that disrupted scheduled acquisitions, only 999 recordings were retained in the final analysis. Note also that this dataset shows unequal distribution among animals, where healthy pigs have about 4 times more recordings than HF pigs. The details of the number of recordings retained for each pig is specified in Table 1. The sampling frequency is 4 kHz.

3. Methodology

CNN-based models are well-known for their ability, during the training phase, to effectively filter noise through various kernels capable of capturing contextual information related to signals of interest, noise, and other artifacts [10]. Building on this, a two-step denoising method is proposed in this paper. In the first step, a One-Dimensional CNN (1D-CNN) model is constructed to optimize the balance between the complexity/performance of the model in terms of classification between healthy animals and those with HF. In the second step, a greedy search is conducted to select the best pre-trained kernel to be used as a denosing filter that satisfies two well-defined criteria: i) having the highest acceptance rate of the recorded data, and ii) maximizing the Signal-to-Noise Ratio (SNR) of the output S1 and S2 waves. The calculation of these two criteria will be presented subsequently.

3.1. Step 1: Training 1D-CNN model

Different CNN architectures were evaluated. The one offering the optimal classification/complexity compromise, in terms of distinguishing between healthy and HF pigs, was selected (Figure 1). The employed architecture includes two 1D convolutional layers, comprising 16 and 32 kernels (of size 3), respectively. Both are followed by Rectified Linear Unit (ReLU) activation and a normalization layer. Then, an average pooling layer, a fully connected layer, a softmax layer, and an output classification layer are used. As shown in Table 1, the dataset is limited and unevenly distributed among pigs (e.g., 50 records for pig 7 versus 316 for pig 4). To address this imbalance, enhance the representation of underrepresented classes, and ensure robust model training, data augmentation was applied. Specifically, a 20-second sliding window with a 1-second step shift was utilized, where each 29-second recording generated 10 overlapping segments, thereby increasing the dataset size tenfold. In addition, to ensure robust, reliable, and generalizable training, a Leave-One-pig-Out (LOO) cross-validation scheme is adopted. Specifically, the model is trained using data from six pigs, while the recording for the remaining 7^{th} pig serves as an independent test set. This process was repeated seven times, with each pig used once as a test set, resulting in seven models. The classification performances of the proposed 1D-CNN model, evaluated in terms of Sensitivity, Specificity, F-score, and Accuracy, are 0.9813, 0.9935, 0.9844, and 0.9893, respectively. The findings demonstrate the efficacy of the proposed model in extracting robust and reliable features from highly noisy ACC signals, enabling accurate distinction between healthy and HF classes.

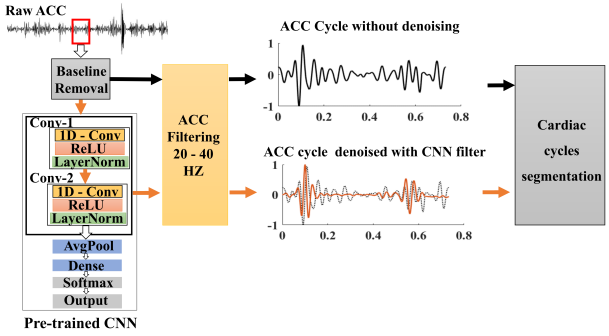


Figure 1: Comparative visualization shows denoising impact on an ACC signal, focusing on a single cardiac cycle: original signal (black) vs. CNN-based denoised (orange).

3.2. Step 2: Denoising kernel selection

CCNN models learn hierarchical features, with early layers detecting simple patterns and deeper layers identifying complex ones [10]. The key idea in this study is the use of a pretrained 1D-CNN model to select the best-suited kernel denoising while preserving the majority of the information related to the S1 and S2 waves. To do this, a greedy search strategy is employed on the set of pretrained kernels of the model. More precisely, the pretrained CNN model is applied to the entire ACC database. Then the SNR of S1 and S2, and the number of retained data (defined in section 4) are calculated at the output of each kernel. The kernel offering the best compromise between preserving the largest possible number of records and the best SNR is then chosen as the optimal denoising filter, Figure 1.

4. Evaluation criteria

The performance of the proposed method was evaluated through two criteria that aim to select only the recordings presenting an acceptable SNR for a robust segmentation of S1 and S2.

Criterion 1 is the acceptance rate of the recorded data. A recording is said to be "accepted" if i) the ratio between the maximum amplitudes of the S1 and S2 waves and the maximum amplitude of the background noise (Figure 2) is greater than or equal to 2 in the mean cycle of coherent ACC cycles, ii) at least two ACC cycles have a correlation greater than 0.6 (see [6] for more details).

Criterion 2 is the SNR improvement for S1 and S2 before and after denoising. Let $\text{SNR}^{\theta}_{\text{noisy}}$ and $\text{SNR}^{\theta}_{\text{denoised}}$ represent the SNR values of the segment $\theta \in \{\text{S1}, \text{S2}\}$ before and after denoising. The SNR for the n-th cycle is defined as:

$$\text{SNR}^{\theta}_{\xi,n} = 10\log_{10}\left(\frac{\max\left((\mathbf{z}^{\theta}_{\xi,\mathbf{n}})^{\odot\mathbf{2}}\right)}{\frac{1}{T_{\text{noise}}}\sum_{i=1}^{T_{\text{noise}}}|u_{\xi,n}[i]|^2}\right) \quad (1)$$

where $\xi \in \{\text{noisy, denoised}\}$, $\mathbf{z}^{\theta}_{\xi,\mathbf{n}}$ is the event θ as a vector, and T_{noise} represents noise duration (Figure 2). For each recording, $\text{SNR}^{\theta}_{\xi}$ is averaged across all cycles, and the overall dataset SNR is calculated as the average across all accepted recordings. The SNR enhancement index η^{θ} is computed as:

$$\eta^{\theta} = \frac{\text{SNR}^{\theta}_{\text{denoised}} - \text{SNR}^{\theta}_{\text{noisy}}}{|\text{SNR}^{\theta}_{\text{noisy}}|} \times 100 \quad (2)$$

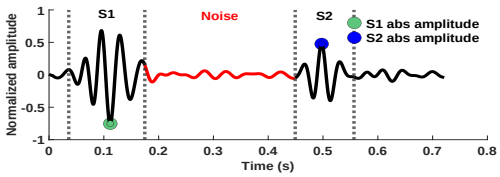
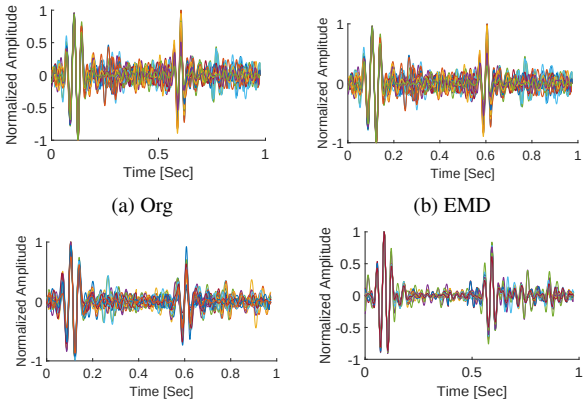


Figure 2: Segmented S1 and S2 heart events within the ACC heart cycle, with a noise/baseline segment in red.

5. Results

This section presents the evaluation an evaluation of the performance of the proposed method in terms of data acceptance rate and SNR improvement for heart sounds S1 and S2 (Criteria 1 and 2). A comparative analysis is performed with two conventional denoising techniques that are widely used in the literature: Wavelet and EMD-based methods. For a fair comparison, the parameters of the conventional methods were tuned to achieve the best trade-off between the two evaluation criteria. The wavelet-based approach used a biorthogonal wavelet with five decomposition levels, combined with Stein's Unbiased Risk Estimate (SURE) thresholding [11] and a soft thresholding strategy. For the EMD method, eight intrinsic IMFs were initially extracted, and the denoised signals were reconstructed using the sum of selected IMFs (three out of the eight in our case) that yielded the best performance. The visual inspection of each denoising method is presented in Figure 3. The figure presents an example of denoising applied to a set of cardiac cycles extracted from a 29-second ACC recording along the Y-axis (ACC_y). To display the full recording, the cardiac cycles were segmented, aligned, and stacked. The proposed method clearly provides superior enhancement of the S1 and S2 heart sounds, although the wavelet-based result also appears satisfactory. Furthermore, the CNN-based approach achieves greater suppression of background noise between S1 and S2, enabling more accurate segmentation of the key cardiac events.



(a) Org (b) EMD

(c) Wavelet (d) CNN

Figure 3: Example of denoising a set of ACC heart cycles extracted from a 29-second Y-axis accelerometer recording.

5.1. Impact of denoising on acceptance rate

Table 1 summarizes the data acceptance rates (Criterion 1) before and after denoising. Before denoising (Org), the acceptance rate is 85.98%. The EMD and wavelet-based methods do not improve this rate, yielding 84.78% and 81.98%, respectively. In contrast, the CNN-based denoising method achieves a significantly higher acceptance rate of 90.69%. This improvement is mainly due to the CNN's stronger suppression of background noise, as shown in Figure 3, unlike EMD and wavelet methods, which are less effective in reducing background noise between S1 and S2.

5.2. Impact of denoising on SNR

Table 2 shows the mean and standard deviation of SNR values (Criterion 2) for S1 and S2 events across the three ACC axes, before and after denoising. The proposed denoising strategy significantly outperforms the two classical methods (wavelets and EMD), as indicated by the higher average SNR values and improvement rates for S1 and S2 on all axes. For instance, on the Y-axis without denoising,

the mean SNRs for S1 and S2 are 15.4 dB and 11.4 dB, respectively. After CNN-based denoising, SNRs increase to 18.8 dB (21.6% improvement) for S1 and 13.2 dB (16% improvement) for S2. Conversely, wavelet and EMD methods yield only modest improvements, with rates ranging from 1.2% to 4.6%.

Table 1: Distribution of accepted records per pig with and without denoising.

Class	Pig ID	Total records	Accepted records by method			
			Org	EMD	Wavelet	CNN
Healthy	1	95	81	81	81	**87**
	2	163	132	135	135	**144**
	3	232	189	191	188	**217**
	4	316	288	280	283	**311**
HF	5	70	57	52	47	**46**
	6	73	66	64	56	**53**
	7	50	46	44	35	**48**
Total	**All**	**999**	**859**	**847**	**819**	**906**
Accept (%)	**–**	**–**	**85.98**	**84.78**	**81.98**	**90.69**

Table 2: Mean $\pm$ standard deviation of SNR and improvement ratio (η) for ACC axes.

Method	ACC S1			ACC S2		
	X	Y	Z	X	Y	Z
No denoising	15.2 ± 5.9	15.4 ± 5.6	15.6 ± 5.9	10.8 ± 5.7	11.4 ± 6.1	11.2 ± 5.8
EMD	15.4 ± 5.4	15.7 ± 5.5	16.0 ± 5.6	11.1 ± 5.8	11.7 ± 5.8	11.6 ± 5.7
η (%)	+1.2	+2.0	+3.0	+3.0	+2.6	+3.2
Wavelet	15.6 ± 5.6	15.8 ± 5.8	15.9 ± 5.8	11.3 ± 5.9	11.8 ± 6.1	11.5 ± 6.0
η (%)	+2.5	+2.9	+2.0	+4.6	+3.5	+2.9
CNN	$\mathbf{17.7 \pm 8.0}$	$\mathbf{18.8 \pm 8.5}$	$\mathbf{17.4 \pm 7.6}$	$\mathbf{12.3 \pm 8.1}$	$\mathbf{13.2 \pm 9.1}$	$\mathbf{12.2 \pm 7.6}$
η (%)	**+17.1**	**+21.6**	**+11.3**	**+13.5**	**+16.0**	**+8.7**

6. Conclusion

A novel approach leverages CNN models initially trained for classification to address denoising challenges in the long-term monitoring of chronic heart diseases. This method involves the selection of the most effective kernel from a low-cost, pretrained CNN, based on its ability to achieve the best trade-off between recording acceptance rate and SNR improvement. The selected kernel is then repurposed for denoising ACC signals acquired by an innovative, minimally invasive implantable device placed at the gastric fundus. The experimental results demonstrate that the proposed CNN-based denoising method significantly enhances the SNR of the two key heart sounds, S1 and S2. From a clinical perspective, this improvement has two main benefits. First, it enables more reliable use of the implanted device by reducing the percentage of unusable recordings. Second, it allows for more precise segmentation of S1 and S2, which is imperative for improving the diagnosis and monitoring of cardiac pathologies.

Acknowledgments

This work was supported in part by the ANR-DiGS project Project-ANR-18-CE19-0008, in part by the PEPR Santé Numérique; Projet DIIP-HEART: ANR-22-PESN-0018, and DeMUG projects of the ARED program of Région Bretagne, France.

References

[1] Sara JD, Toya T, Taher R, Lerman A, Gersh B, Anavekar NS. Asymptomatic left ventricle systolic dysfunction. European Cardiology Review 2020;15.

[2] Shahim B, Kapelios CJ, Savarese G, Lund LH. Global public health burden of heart failure: an updated review. Cardiac Failure Review 2023;9.

[3] Rosano GM, Seferovic P, Savarese G, Spoletini I, Lopatin Y, Gustafsson F, Bayes-Genis A, Jaarsma T, Abdelhamid M, Miqueo AG, et al. Impact analysis of heart failure across european countries: an esc-hfa position paper. ESC heart failure 2022;9(5):2767–2778.

[4] Saner H, Schuetz N, Buluschek P, Du Pasquier G, Ribaudo G, Urwyler P, Nef T. Case report: Ambient sensor signals as digital biomarkers for early signs of heart failure decompensation. Frontiers in cardiovascular medicine 2021;8:617682.

[5] Sørensen K. Seismocardiography: Interpretation and clinical application 2021;.

[6] Areiza-Laverde H, Dopierala C, Senhadji L, Boucher F, Gumery PY, Hernández A. Analysis of cardiac vibration signals acquired from a novel implant placed on the gastric fundus. Frontiers in Physiology 2021;12:748367.

[7] Giorgis L, Frogerais P, Amblard A, Donal E, Mabo P, Senhadji L, Hernández AI. Optimal algorithm switching for the estimation of systole period from cardiac microacceleration signals (sonr). IEEE Transactions on Biomedical Engineering 2012;59(11):3009–3015.

[8] Sahoo GR, Freed JH, Srivastava M. Optimal wavelet selection for signal denoising. IEEE Access 2024;.

[9] Fleureau J, Nunes JC, Kachenoura A, Albera L, Senhadji L. Turning tangent empirical mode decomposition: a framework for mono-and multivariate signals. IEEE Transactions on signal Processing 2010;59(3):1309–1316.

[10] Rodrigues NV, Abramo LR, Hirata NS. The information of attribute uncertainties: what convolutional neural networks can learn about errors in input data. Machine Learning Science and Technology 2023;4(4):045019.

[11] Donoho DL, Johnstone IM. Adapting to unknown smoothness via wavelet shrinkage. J Am Stat Assoc 1995; 90(432):1200–1224.

Address for correspondence:

Salman Almuhammad Alali
Universite de Rennes, LTSI, Rennes, 35042, France
salman.almuhammad-alali@univ-rennes.fr