# Noise-Robust Heart Sound Analysis for Wearable Monitoring Using Cascaded Dual-Channel ANC and Wavelet Denoising

Jiaqi Zhang[1], Chenghao Yang[1], Chen Shen[1], Tiancheng Cao[1,2], Hen-Wei Huang[1,3]
[1]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore
[2]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, USA
[3]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore
jiaqi022@e.ntu.edu.sg

*Abstract*— **Phonocardiogram (PCG) signals are widely used for automated cardiovascular abnormality screening and wearable health monitoring. However, during real-world acquisition, PCG signals are often corrupted by environmental noise, which masks critical transient structures such as S1 and S2 and degrades automatic analysis and classification performance. To address the challenge of jointly achieving effective noise suppression and transient structure fidelity under complex noise conditions, this paper proposes a cascaded PCG front-end framework that integrates dual-channel subband adaptive noise cancellation with subband-selective wavelet denoising. The framework first suppresses environmental noise highly correlated with the reference channel at the subband level, and then applies selective wavelet-domain thresholding to further attenuate residual noise while preserving heart sound transients.A self-trained lightweight heart sound classification model is used for evaluation. Under low-SNR conditions, the proposed cascaded approach achieves classification performance gains of 9.3%−12.7% over noisy inputs, significantly outperforming single-stage wavelet denoising (6.3%−9.1%) and ANC (4.4%−5.0%). Stable positive gains are also maintained at medium-to-high SNRs, whereas single-stage wavelet denoising exhibits performance degradation. These results demonstrate that the proposed front-end jointly suppresses noise and preserves key transient structures at the signal level, relieving low-memory classifiers from learning additional noise robustness and enabling efficient collaboration for embedded and wearable heart sound analysis systems.**

*Index Terms*— **Phonocardiogram (PCG); Front-end Signal Processing; Dual-channel Adaptive Noise Cancellation; Subband Processing; Noise Robustness**

## I. INTRODUCTION

Phonocardiogram (PCG) signals provide a low-cost and non-invasive means for measuring cardiac mechanical activity and have been widely used for the automated screening of cardiovascular abnormalities [1]. In recent years, advances in machine learning have further promoted the development of PCG-based classification systems for clinical decision support and wearable monitoring applications [2]. However, due to their low amplitude, PCG signals are highly susceptible to environmental noise, body motion, and respiratory interference during real-world acquisition. These disturbances often exhibit substantial overlap with heart sounds in both the time and frequency domains, potentially obscuring transient heart sound components that are critical for reliable classification and diagnosis [3]–[5]. Consequently, effectively suppressing noise while preserving heart sound structures and discriminative information under complex noise conditions remains a core challenge in PCG-based automatic analysis and is essential for improving system robustness.

To mitigate noise interference, various front-end denoising and enhancement methods have been proposed. Time–frequency analysis approaches, such as wavelet denoising, improve the separability of transient heart sound components through multi-band representations [6]. However, their energy-based thresholding mechanisms may suppress fine-grained temporal structures when noise overlaps with discriminative heart sound components, particularly under high signal-to-noise ratio (SNR) conditions [7]–[9]. On the other hand, multi-channel noise suppression methods exploit noise reference observations to estimate and cancel environmental interference. Dual-channel approaches, such as active noise control (ANC), can effectively reduce overall noise energy when a high-quality noise reference is available [10], but their optimization objectives typically focus on error energy minimization without explicitly preserving heart sound transient structures [11], [12]. Consequently, existing methods tend to emphasize either noise suppression or structural preservation, while rarely addressing both within a unified framework.

To address the aforementioned challenges, this paper proposes a cascaded front-end processing framework that integrates dual-microphone noise suppression with wavelet denoising for robust PCG signal enhancement. The framework first suppresses non-cardiac interference at the signal level through dual-microphone processing, improving the separability between heart sounds and noise. Subsequently, classification-relevant heart sound transients are enhanced via subband time–frequency analysis in the wavelet domain. By decoupling noise source suppression and temporal structure preservation into two complementary stages, the proposed method improves both noise robustness and signal fidelity across diverse acoustic conditions.

The main contributions of this work are summarized as follows: 1) A cascaded PCG front-end enhancement framework that combines dual-microphone noise suppression with wavelet denoising to jointly address noise robustness

*This work was supported by the Nanyang Professorship, the MOE Tier 1 grant RG71/24, and the MTC MedTech Programmatic Fund M24N9b0130.

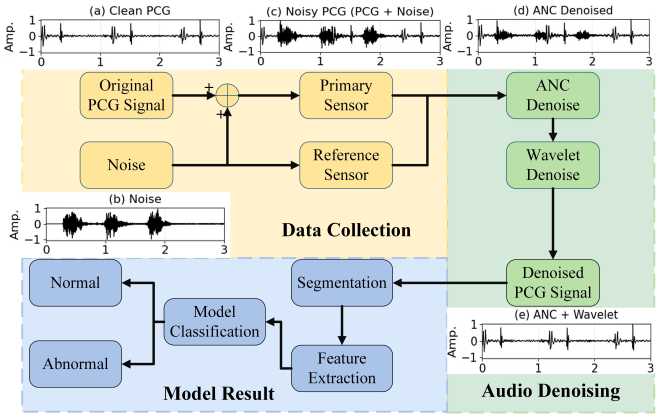


Fig. 1. Overall framework of the proposed cascaded heart sound analysis system.

TABLE I

Subband configuration for ANC and wavelet denoising

| Method | Band | Range (Hz) | Strategy |
|---|---|---|---|
| ANC | Low | 0–250 | Pass-through |
| | High | $> 250$ | Subband NLMS |
| Wavelet | $cD_1$ | 500–1000 | Thresholding |
| | $cD_2$ | 250–500 | Thresholding |
| | $cA_2$ | 0–250 | Pass-through |

and temporal structure fidelity; 2) A signal-level analysis of the complementary roles of dual-microphone noise suppression and wavelet denoising at different processing stages, validating the effectiveness of the adopted cascade order.

## II. METHODOLOGY

As illustrated in Fig. 1, this paper proposes a cascaded real-time heart sound analysis system designed for complex noise environments. The core processing pipeline consists of three sequential stages: (1) dual-microphone subband adaptive noise cancellation (Subband ANC), (2) wavelet-domain denoising, and (3) heart sound classification. The following subsections describe each stage in detail.

It is worth noting that, in the front-end design, this work follows a widely adopted frequency-band assumption in prior studies, namely that the primary discriminative information of heart sounds is concentrated in the low-frequency range (typically below 250 Hz), and that aggressive noise suppression in this band may distort critical cardiac sound structures [13], [14]. Accordingly, a protective pass-through strategy is applied to the low-frequency, heart-sound-dominant band, while adaptive noise cancellation and threshold-based processing are restricted to higher-frequency bands. The corresponding band configuration and processing strategies employed in the dual-channel ANC and wavelet denoising stages are summarized in Table I.

### A. Dual-Microphone Subband Adaptive Noise Cancellation

As shown in Fig. 2, this work adopts a dual-microphone Subband ANC front-end structure. The primary channel

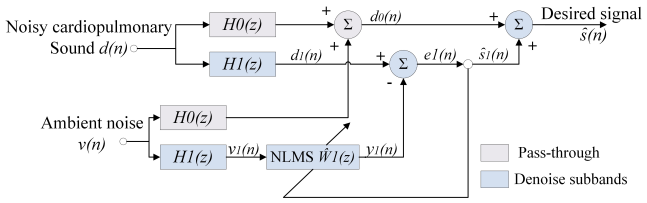


Fig. 2. Block diagram of the proposed dual-microphone subband ANC algorithm, where $H_n(\cdot)$ denotes FFT-based subband decomposition with frequency-band selection and IFFT reconstruction.

captures noisy heart sounds, while the auxiliary channel provides an environmental noise reference. To enable frequency-dependent processing, both signals are first decomposed into low- and high-frequency subbands using analysis filter banks $H_0(z)$ and $H_1(z)$. In implementation, subband decomposition is realized via FFT-based frequency masking: the signals are transformed into the frequency domain, the target bands are selected using predefined masks, and the corresponding time-domain subband signals are reconstructed by inverse FFT (IFFT). The low-frequency subband, which mainly contains structural heart sound components, is directly preserved, whereas the high-frequency subband exhibits a higher noise proportion.

Based on this decomposition, ANC is applied only to the high-frequency subband. In this band, the primary-channel signal can be regarded as a superposition of heart sound and noise components, while the auxiliary-channel subband is strongly correlated with the noise component in the primary channel and thus serves as the reference input for adaptive filtering.

The reference subband signal is fed into an adaptive filter based on the normalized least mean squares (NLMS) criterion to estimate the interference components in the primary channel that are correlated with the reference noise. Let the adaptive filter coefficient vector be $\mathbf{w}(n)$ and the reference input vector be $\mathbf{v}(n)$. The filter output is given by

$$y_1(n) = \mathbf{w}^{\mathsf{T}}(n)\mathbf{v}(n), \tag{1}$$

which corresponds to an estimate of the high-frequency noise component in the primary channel.

After subtracting the estimated noise $y_1(n)$ from the primary-channel subband, the residual signal $e_1(n)$ is obtained and used as the error signal to update the filter coefficients. The NLMS update rule is expressed as

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \frac{e_1(n)\mathbf{v}(n)}{\|\mathbf{v}(n)\|_2^2 + \epsilon}, \tag{2}$$

where $\mu$ is the step size and $\epsilon$ is a regularization term for numerical stability.

After the adaptive filter converges, the residual signal $e_1(n)$ mainly contains the heart sound components within the high-frequency band and can therefore be regarded as the estimated heart sound output of the high-frequency subband, denoted as $\hat{s}_1(n)$. The final denoised heart sound signal is reconstructed in the time domain by superimposing the

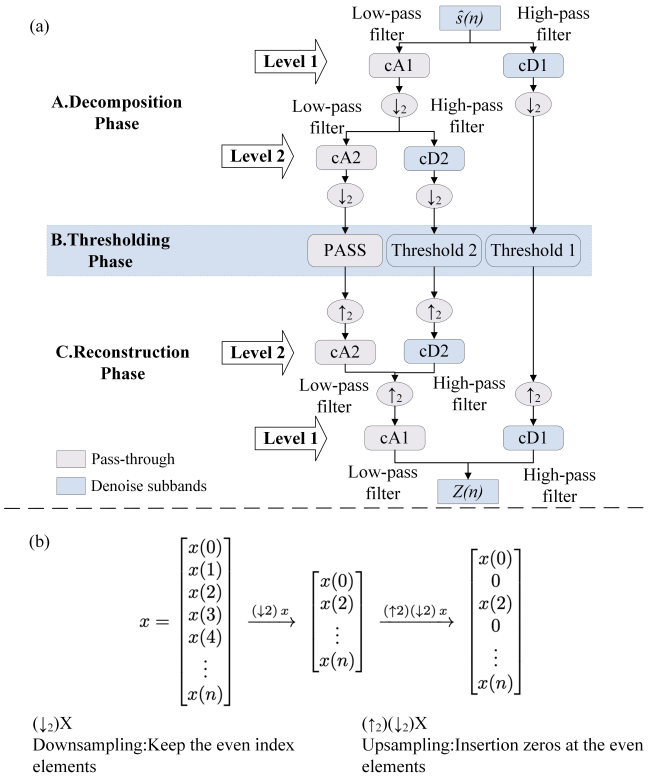


Fig. 3. Subband-selective wavelet denoising following ANC. (a) Two-level DWT with selective thresholding applied to detail subbands. (b) Equivalent downsampling (↓ 2) and upsampling (↑ 2) operations in the wavelet transform.

low-frequency pass-through component $d_0(n)$ and the high-frequency residual component $\hat{s}_1(n)$:

$$\hat{s}(n) = d_0(n) + \hat{s}_1(n). \tag{3}$$

The output signal $\hat{s}(n)$ obtained from the subband ANC stage is subsequently processed by wavelet-domain denoising, as described in the following subsection.

### B. Subband-Selective Wavelet Denoising

Following the subband ANC stage, a subband-selective wavelet denoising module is applied to further suppress residual noise, as illustrated in Fig. 3(a).

To further suppress residual noise while preserving the structural characteristics of heart sounds, wavelet-domain denoising is introduced after the ANC front end. Let the output of the subband ANC be denoted as $X(n)$. A two-level discrete wavelet transform (DWT) is applied to $X(n)$. At each decomposition level, the signal is successively processed by low-pass and high-pass analysis filters, followed by a factor-of-two downsampling (↓ 2), thereby forming a multiresolution wavelet subband representation. In this context, downsampling (↓ 2) is equivalent to retaining samples at even indices; correspondingly, during the reconstruction stage, upsampling (↑ 2) restores the sampling rate by inserting zeros at even positions. The equivalent sampling mechanism is illustrated in Fig. 3(b).

Through the two-level decomposition, the signal is separated into one approximation subband $cA_2$ and two detail subbands $cD_2$ and $cD_1$, which correspond to different frequency ranges, as summarized in Table I. Considering the distinct distributions of noise and heart sound components across different scales, threshold suppression is applied only to the mid- and high-frequency detail subbands involved in denoising, while the lowest-frequency approximation subband is preserved without modification.

For each detail subband $cD_j$ involved in denoising ($j \in \{1, 2\}$), the noise scale is first estimated using the median absolute deviation (MAD) method:

$$\hat{\sigma}_j = \frac{\text{median}(|cD_j - \text{median}(cD_j)|)}{0.6745}. \tag{4}$$

The constant $0.6745$ represents the proportionality factor between the MAD and the standard deviation under a Gaussian distribution, and is used to normalize the MAD estimate into an unbiased estimate of the noise standard deviation.

Based on this estimate, a subband-dependent threshold is computed according to the universal threshold rule:

$$\lambda_j = \hat{\sigma}_j \sqrt{2 \ln N_j}, \tag{5}$$

where $N_j$ denotes the number of coefficients in the subband $cD_j$.

Subsequently, soft-thresholding is applied to the selected detail subbands $cD_1$ and $cD_2$. Specifically, coefficients with magnitudes below the corresponding threshold are set to zero, while those with magnitudes exceeding the threshold are shrunk by subtracting the threshold value from their magnitudes while preserving their original signs, thereby achieving a continuous amplitude shrinkage. This operation suppresses noise-dominated small coefficients while retaining significant coefficients associated with heart sound transient structures. The subband coefficients that are not involved in the denoising process remain unchanged.

Finally, the processed subband coefficients are reconstructed into the time domain via the inverse discrete wavelet transform, yielding the wavelet-denoised output signal $Z(n)$. Only the selected detail subbands are modified by thresholding, whereas the remaining subbands are directly used for reconstruction.

Through this subband-selective wavelet denoising strategy, residual high-frequency noise is effectively attenuated while the low-frequency structure and discriminative information of heart sounds are largely preserved, providing a more robust input for subsequent feature extraction and classification.

### C. Lightweight Classification Backend for Evaluation

To provide a task-level and reproducible evaluation of the proposed denoising framework, a lightweight convolutional neural network, referred to as *MicroHeartCNN*, is employed as the classification backend. Similar robustness-oriented evaluation strategies using auxiliary classifiers have been adopted in biomedical signal analysis under perturbations [15]. It should be emphasized that the classifier itself

TABLE II

MicroHeartCNN architecture and hyper-parameters (input: $1 \times 64 \times 94$).

| Stage | Feature map (H×W) | Layer specification | MicroHeartCNN |
|---|---|---|---|
| Input: $x$ in $R^{1\times64\times94}$ (C×H×W) | | | |
| 0 | $64 \times 94$ | Input | in_ch=1 |
| 1 | $32 \times 47$ | Downsample Conv | Conv $3 \times 3$, $s = 2$, out=4 |
| 2 | $16 \times 24$ | Downsample Conv | Conv $3 \times 3$, $s = 2$, out=6 |
| 3 | $8 \times 12$ | Downsample Conv | Conv $3 \times 3$, $s = 2$, out=6 |
| | $1 \times 1$ | Global Pooling | AdaptiveAvgPool2d(1) |
| | | Classifier | Linear ($6 \rightarrow 2$) |
| **Parameters** | | | 622 (≈0.00062M) |
| **MACs (Conv+FC)** | | | 168,204 (≈0.168M) |
| **FP32 size** | | | 2.43 KB |

is not the primary methodological contribution of this work. Instead, it serves as a fixed and compact evaluation tool that translates signal-level improvements into classification-level outcomes under consistent conditions.

The input to the classifier is a single-channel time–frequency representation of heart sounds with a fixed dimensionality of $1 \times 64 \times 94$, corresponding to the log-Mel spectrogram extracted from the denoised signal. A fixed input resolution is used throughout all experiments to ensure fair and repeatable evaluation across different denoising configurations. The network outputs logits for binary classification, distinguishing between normal and abnormal heart sounds.

The MicroHeartCNN architecture is intentionally designed to be extremely compact while retaining sufficient representational capability for heart sound classification [16]. The network consists of three convolutional blocks, followed by a global average pooling layer and a fully connected classifier. Each convolutional block comprises a $3 \times 3$ convolution with stride 2, batch normalization, and ReLU activation. Strided convolutions are adopted to progressively reduce the spatial resolution of feature maps, allowing hierarchical feature abstraction with minimal computational overhead.

Global average pooling is applied after the final convolutional block to aggregate spatial information and eliminate the need for large fully connected layers. This design choice significantly reduces the number of trainable parameters and mitigates overfitting, which is particularly important given the limited model capacity and the relatively small size of biomedical datasets.

The complete network contains only 622 trainable parameters, corresponding to an FP32 model size of approximately 2.43 KB. Table II summarizes the architecture configuration and the resulting feature map sizes. Such a distribution reflects a deliberate design choice that prioritizes feature extraction over classifier complexity, making the network well suited for deployment on memory- and power-constrained platforms [17]. This design philosophy is consistent with recent low-complexity heart sound classification models targeting real-time and resource-constrained platforms [18].

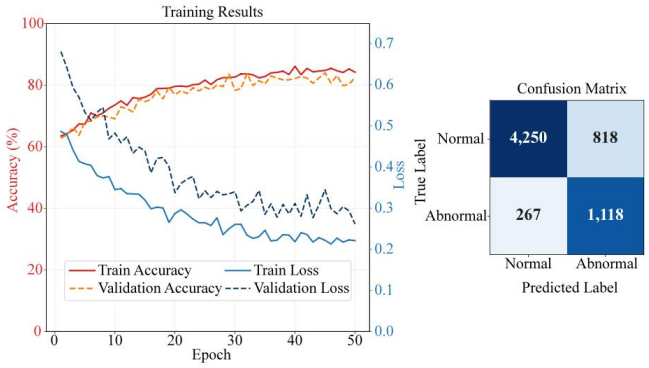


Fig. 4. Training and validation accuracy and loss curves, and confusion matrix of the proposed MicroHeartCNN.

## III. RESULT

### *A. Auxiliary Classification Results Using a Lightweight CNN*

The MicroHeartCNN classifier was trained for 50 epochs following the training protocol described in Section II-C. Model selection was performed based on validation accuracy to avoid overfitting, and the best-performing checkpoint was evaluated on a held-out test set. All experiments were conducted using the same network architecture and training configuration to ensure consistent evaluation.

Fig. 4 presents the training and validation accuracy and loss curves over the training process, together with the confusion matrix obtained on the test set. The learning curves indicate stable convergence behavior, with both training and validation losses decreasing steadily as training progresses. No severe divergence between training and validation performance is observed, suggesting that the lightweight model can be effectively optimized despite its limited capacity.

The final training accuracy reaches 84.24%, while the best validation accuracy is 84.00%, indicating consistent generalization across the training, validation, and test sets data splits. On the test set, the classifier achieves an overall accuracy of 83.12% and a macro-averaged F1-score of 80.89%, while maintaining an extremely compact FP32 model size of approximately 2.43 KB. These results demonstrate that the MicroHeartCNN provides a reliable baseline level of classification performance under strict memory constraints, consistent with prior lightweight CNN-based diagnostic systems for remote and telemedicine applications [19].

The confusion matrix in Fig. 4 further illustrates the classification behavior of the model. A balanced trade-off between sensitivity and specificity is observed, with an abnormal heart sound recall of 80.72%. This characteristic is desirable for screening-oriented applications, where the priority is to minimize missed pathological cases, even at the cost of an increased false alarm rate. Importantly, the classifier performance is sufficiently stable and interpretable to serve its intended role as an auxiliary evaluation module, without dominating the system-level contributions of the proposed denoising approach.

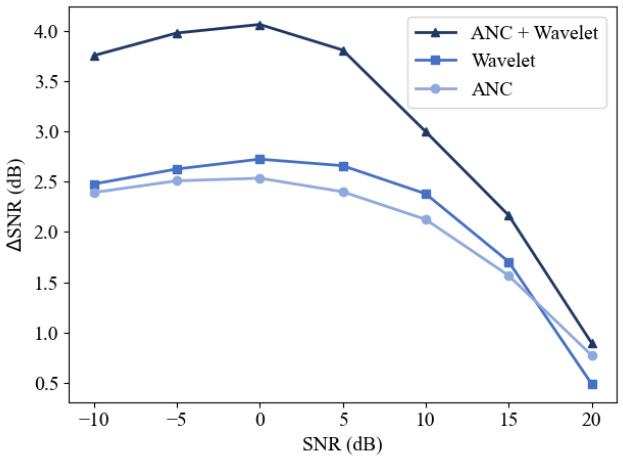


Fig. 5. Comparison of ΔSNR achieved by different denoising pipelines under varying injected SNR conditions.

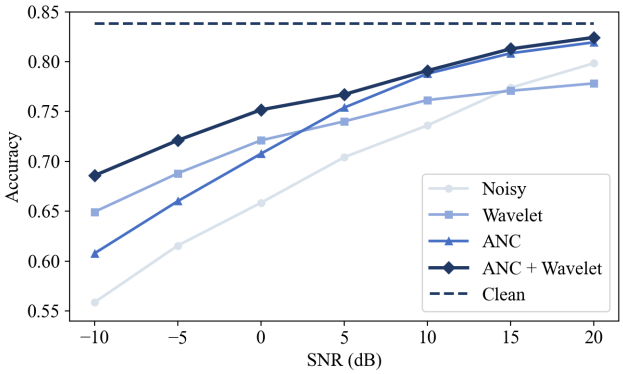


Fig. 6. Classification accuracy under different SNR conditions for four processing pipelines. The dashed line denotes the Clean reference upper bound.

### B. Denoising Performance Evaluation

Fig. 5 compares the denoising performance of three front-end processing strategies under different injected SNR conditions, including adaptive noise cancellation (ANC), wavelet-based denoising (Wavelet), and their cascaded combination (ANC+Wavelet). The horizontal axis denotes the injected SNR, while the vertical axis represents the denoising gain ΔSNR, defined as

$$\Delta\text{SNR} = \text{SNR}_{\text{out}} - \text{SNR}_{\text{in}}, \tag{6}$$

where $\text{SNR}_{\text{in}}$ and $\text{SNR}_{\text{out}}$ denote the measured signal-to-noise ratios before and after denoising, respectively.

At low injected SNR levels ($-10$ dB to 0 dB), the cascaded ANC+Wavelet approach consistently achieves the largest SNR gain, with ΔSNR of approximately 3.8–4.1 dB, outperforming ANC-only and Wavelet-only methods (about 2.4–2.7 dB). This indicates that the two-stage front-end is more effective at suppressing noise energy and improving the overall SNR under severe noise conditions. As the injected SNR increases, the ΔSNR of all three methods gradually decreases, reflecting the reduced headroom for further SNR improvement when the input signal quality is already high. At 20 dB, the ΔSNR of ANC+Wavelet is approximately 0.9 dB, while the gains of the single-stage methods are even smaller, indicating limited SNR improvement at high injected SNR levels.

Overall, all three methods yield positive ΔSNR across the entire tested SNR range, indicating that the front-end processing does not degrade signal quality. Compared with single-stage denoising strategies, the proposed ANC+Wavelet cascade exhibits more consistent and larger SNR gains over a wide range of noise conditions, providing a more robust input for subsequent frame-based heart sound classification experiments.

### C. Model-Based Evaluation of Denoising Methods

To evaluate the practical benefit of front-end denoising for downstream heart sound classification, four processing pipelines are compared under a unified evaluation protocol:

1) **Noisy**: noisy PCG signals directly used as input;
2) **Wavelet**: wavelet-threshold denoising applied to noisy PCG;
3) **ANC**: dual-microphone ANC-based adaptive noise cancellation only;
4) **ANC + Wavelet**: cascaded processing with ANC followed by wavelet denoising.

To ensure a fair comparison, all four processing pipelines adopt the same data split, identical framing and feature extraction procedures, and are evaluated using the same fixed binary classification model. Except for the front-end denoising stage, all training and inference configurations remain identical. Experiments are conducted over a wide input SNR range from $-10$ dB to 20 dB.

Fig. 6 shows the classification accuracy of the four pipelines under different SNR conditions. Overall, classification accuracy increases monotonically with SNR for all methods, indicating that noise intensity is a critical factor affecting discriminative performance. Compared with the Noisy baseline, introducing front-end denoising consistently improves classification accuracy across the entire SNR range, with more pronounced gains observed under low-SNR conditions. This suggests that front-end processing effectively mitigates the masking effect of strong noise on discriminative features.

Among all pipelines, the cascaded ANC + Wavelet approach achieves the highest classification accuracy over the full SNR range from $-10$ dB to 20 dB, demonstrating superior and stable robustness. At high SNR levels, the performance of all methods gradually approaches the upper bound obtained with clean signals (Clean), while the cascaded scheme remains consistently superior. This indicates that the proposed front-end cascade not only enhances robustness under severe noise but also avoids performance degradation in high-SNR scenarios.

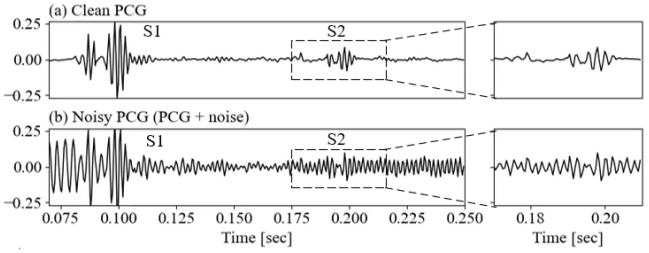


Fig. 7. Qualitative comparison between clean and noisy PCG signals. (a) Clean PCG waveform with clearly distinguishable S1 and S2 events. (b) Noisy PCG (PCG + noise), where strong periodic background oscillations obscure low-amplitude transient details, particularly around S2 (highlighted in the zoomed-in regions).

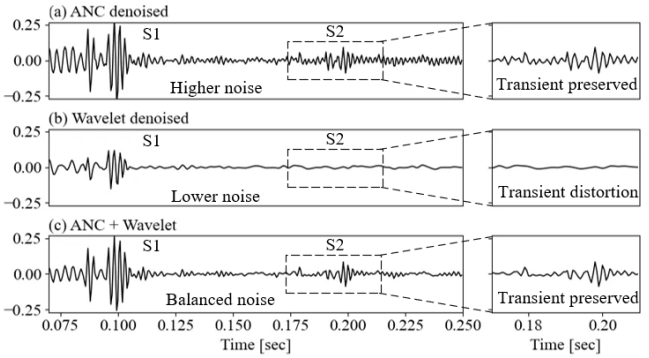


Fig. 8. Qualitative comparison of time-domain PCG signals after different front-end denoising methods. (a) ANC denoised signal, where transient structures (S1/S2) are largely preserved but noticeable residual noise remains. (b) Wavelet denoised signal, showing stronger background noise suppression at the cost of transient smoothing and distortion. (c) Cascaded ANC+Wavelet output, achieving a more balanced trade-off with reduced noise and preserved transient morphology (highlighted in the zoomed-in regions).

To visually illustrate the impact of noise on the temporal structure of heart sounds, Fig. 7 compares clean PCG signals with their noisy counterparts. Under strong noise conditions, pronounced periodic background oscillations emerge, partially obscuring low-amplitude transient details around S1 and S2. In particular, the transient contour of S2 becomes less distinguishable, as highlighted in the zoomed-in regions. Such degradation of transient structures directly impairs the model's ability to discriminate between normal and abnormal heart sounds, accounting for the degraded performance of the Noisy pipeline under low-SNR conditions.

Fig. 8 further compares the time-domain behavior of ANC, Wavelet, and the cascaded ANC+Wavelet scheme, providing a structural explanation for the performance trends observed in Fig. 6. ANC is able to largely preserve the transient morphology and peak structures of S1/S2, with relatively clear transient contours in the zoomed regions; however, noticeable residual noise remains in non-transient intervals, limiting its overall denoising capability. In contrast, Wavelet denoising achieves stronger background noise suppression but tends to smooth low-amplitude transient components, leading to peak attenuation or even transient distortion.

Benefiting from the complementary characteristics of the two methods, the cascaded ANC+Wavelet approach exhibits a more balanced behavior. ANC first exploits inter-channel noise correlation to suppress dominant interference while preserving transient structures, after which Wavelet denoising further attenuates residual background noise. As a result, the output simultaneously achieves reduced noise levels and improved transient fidelity. This balanced enhancement enables the cascaded strategy to consistently improve feature separability over a wide SNR range, ultimately leading to more stable and superior classification performance.

## IV. CONCLUSIONS

This paper presents a cascaded front-end processing framework for heart sound analysis in complex noise environments, which integrates dual-microphone subband adaptive noise cancellation (ANC) with subband-selective wavelet denoising. Across a wide SNR range from $-10$ dB to 20 dB, the proposed cascaded scheme consistently achieves the highest classification accuracy at all tested conditions. Under low-SNR scenarios ($-10$ dB, $-5$ dB, and 0 dB), the accuracy improvements reach 12.7%, 10.6%, and 9.3%, respectively, clearly outperforming the single-stage wavelet or ANC approaches. In the mid-to-high SNR range (5–20 dB), the cascaded method maintains stable positive gains of 2.6%–6.3% and avoids the performance degradation observed in single-stage wavelet denoising at high SNRs. These results demonstrate that the proposed front-end effectively balances noise robustness and preservation of discriminative heart sound features, providing a practical and generalizable solution for wearable heart sound analysis under resource-constrained conditions.

## REFERENCES

[1] A. Gudigar, U. Raghavendra, M. Maithri, J. Samanth, M. A. Inamdar, V. Vidhya, J. Vicnesh, M. A. Prabhu, R. Tan, C. H. Yeong, F. Molinari, and U. R. Acharya, "Automated system for the detection of heart anomalies using phonocardiograms: A systematic review," *IEEE Access*, vol. 12, pp. 138 399–138 428, 2024.

[2] A. Mahmood, H. Dhahri, M. Alhajlah, and A. Almaslukh, "Enhanced classification of phonocardiograms using modified deep learning," *IEEE Access*, vol. 12, pp. 178 909–178 916, 2024.

[3] D. S. Panah, A. Hines, and S. Mckeever, "Exploring the impact of noise and degradations on heart sound classification models," *Biomed. Signal Process. Control.*, vol. 85, p. 104932, 2022.

[4] Y. Arjoune, T. N. Nguyen, R. W. Doroshow, and R. Shekhar, "A noise-robust heart sound segmentation algorithm based on shannon energy," *IEEE access : practical innovations, open solutions*, vol. 12, pp. 7747 – 7761, 2024.

[5] F. Azam, M. I. Ansari, S. I. S. K. Nuhash, I. McLane, and T. Hasan, "Cardiac anomaly detection considering an additive noise and convolutional distortion model of heart sound recordings," *Artificial intelligence in medicine*, vol. 133, p. 102417, 2022.

[6] P. Chen and Q. Zhang, "Classification of heart sounds using discrete time-frequency energy feature based on s transform and the wavelet threshold denoising," *Biomed. Signal Process. Control.*, vol. 57, 2020.

[7] G. Baldazzi, E. Sulas, M. Urru, R. Tumbarello, L. Raffo, and D. Pani, "Wavelet denoising as a post-processing enhancement method for non-invasive foetal electrocardiography," *Computer methods and programs in biomedicine*, vol. 195, p. 105558, 2020.

[8] S. Debbal, "Heart cardiac's sounds signals segmentation by using the discrete wavelet transform (dwt)," *Biomedical Research and Clinical Reviews*, 2021.

[9] M. Zhang, H. Ma, Y. Wu, and W. Cui, "An improved wavelet threshold function denoising method based on iga optimization," *Vibroengineering Procedia*, vol. 55, p. 267–272, Sep. 2024. [Online]. Available: https://doi.org/10.21595/vp.2024.24412

[10] X. Shen, J. Ji, D. Shi, Z. wu Luo, and W.-S. Gan, "The principle underlying the wireless reference microphone enhancing noise reduction performance in multi-channel active noise control windows," *Mechanical Systems and Signal Processing*, 2024.

[11] X. Shen, D. Shi, S. Peksi, and W. Gan, "A multi-channel wireless active noise control headphone with coherence-based weight determination algorithm," *Journal of Signal Processing Systems*, vol. 94, pp. 811 – 819, 2022.

[12] M. Fynn, S. Nordholm, and Y. Rong, "Coherence function and adaptive noise cancellation performance of an acoustic sensor system for use in detecting coronary artery disease," *Sensors*, vol. 22, no. 17, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/17/6591

[13] T. H. Chowdhury, K. N. Poudel, and Y. Hu, "Time-frequency analysis, denoising, compression, segmentation, and classification of pcg signals," *IEEE Access*, vol. 8, pp. 160 882–160 890, 2020.

[14] G. K. Soni, H. Singh, H. Arora, and A. Soni, "Ultra low power cmos low pass filter for biomedical ecg/eeg application," in *2020 Fourth International Conference on Inventive Systems and Control (ICISC)*, 2020, pp. 558–561.

[15] C. A. Ellis, R. L. Miller, and V. D. Calhoun, "Evaluating augmentation approaches for deep learning-based major depressive disorder diagnosis with raw electroencephalogram data*," in *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2024, pp. 1–5.

[16] T. Cao, Z. Zhang, W. Soon Ng, W. Ling Goh, and Y. Gao, "Fpga implementation of poolformer network using python-driven high-level synthesis framework for edge-aiot speech recognition," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 34, no. 1, pp. 317–321, 2026.

[17] T. Cao, W. S. Ng, W. L. Goh, and Y. Gao, "Dwt-poolformer: Discrete wavelet transform-based quantized parallel poolformer network implemented in fpga for wearable ecg monitoring," in *2024 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2024, pp. 1–5.

[18] J. Ji, L. Zhu, H. Zhang, K. Qian, K. Xu, Z. Song, B. Hu, B. W. Schuller, and Y. Yamamoto, "Weight light, hear right: Heart sound classification with a low-complexity model," in *2024 32nd European Signal Processing Conference (EUSIPCO)*, 2024, pp. 326–330.

[19] S.-A. Chiu, C.-Y. Yeh, and W.-C. Fang, "An effective lightweight cnn-based system for respiratory disease diagnosis in telemedicine," in *2025 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)*, 2025, pp. 213–214.