



ARTICLE

# Confidence-Regulated Heart Murmur Classification via Joint Representation Learning and Decision Optimization

HyeSun Chang<sup>1</sup> and Sangjun Lee<sup>1\*</sup>

Department of AI/SW Convergence, Soongsil University, Seoul, Republic of Korea

\*Corresponding Author: Sangjun Lee. Email: [sangjun@ssu.ac.kr](mailto:sangjun@ssu.ac.kr)

Received: 21 March 2026; Accepted: 14 May 2026; Published: 15 June 2026

**ABSTRACT:** Accurate identification of heart murmurs from auscultation recordings is essential for early cardiovascular screening and diagnosis. While deep learning offers strong potential for automated heart murmur classification, existing models often exhibit overconfident, incorrect predictions and limited generalization due to dataset bias and class imbalance. To address these challenges, this study proposes a two-stage confidence-regulated learning framework that jointly optimizes feature representation and decision reliability. Rather than focusing solely on improving classification performance, this work emphasizes enhancing prediction reliability through confidence-aware decision-making. The proposed framework integrates supervised contrastive learning (SCL) to strengthen the discriminative structure of feature embeddings and reward-based optimization (RBO) to regulate prediction confidence under uncertainty. In this framework, a convolutional neural network-based encoder first extracts acoustic representations, and a long short-term memory-based classifier refines the learned embeddings before final prediction. SCL improves intra-class compactness and inter-class separability, while the proposed confidence-regulated mechanism enables the model to adaptively accept or defer predictions based on a dynamically adjusted threshold. This approach allows the model to balance prediction accuracy and decision reliability by reducing overconfident errors in uncertain cases. The proposed method is evaluated on the PhysioNet 2022 heart murmur dataset. Experimental results show that the proposed framework improves the validation Score from 0.8064 to 0.8233, where the Score is defined as the mean of sensitivity and specificity. More importantly, these results demonstrate improved reliability and more balanced decision-making under uncertainty, beyond a marginal increase in aggregate performance. These findings demonstrate that jointly optimizing representation learning and confidence-regulated decision-making provides an effective and clinically relevant approach for robust heart murmur classification.

**KEYWORDS:** Heart murmur classification; phonocardiogram; supervised contrastive learning; confidence-aware decision-making; reinforcement learning

## 1 Introduction

Cardiovascular diseases (CVDs) remain one of the leading causes of mortality worldwide, making early detection essential for timely intervention and improved clinical outcomes [1]. Auscultation has long served as a fundamental and non-invasive tool for cardiac assessment; however, its diagnostic accuracy is often influenced by clinician expertise, environmental noise, and the inherent variability of heart sounds [2]. These factors can lead to an inconsistent interpretation of heart murmurs in clinical practice. With the increasing adoption of digital stethoscopes and computer-aided diagnostic technologies, automated analysis of phonocardiogram recordings has emerged as a promising approach to improve the objectivity and accessibility of cardiac screening.

Recent advances in deep learning have significantly enhanced the performance of automated heart sound analysis. By learning representations from time-frequency inputs such as spectrograms, deep learning models can capture both spectral characteristics and temporal dynamics of cardiac cycles [3]. These developments have enabled substantial improvements in heart murmur classification accuracy. Nevertheless, achieving clinically reliable decision support remains challenging. In real-world scenarios, diagnostic models must not only accurately distinguish between murmur categories but also remain robust to data heterogeneity and provide reliable predictions when confronted with uncertain or ambiguous inputs.

Despite their promising performance, existing artificial intelligence-based heart sound classification models exhibit important limitations. A key concern is that such models often produce highly confident yet incorrect predictions, which can be detrimental in safety-critical medical applications. Furthermore, model performance frequently degrades when evaluated on data collected from different populations, devices, or recording conditions, reflecting the effects of dataset bias and limited generalization [4,5]. These challenges indicate that improving classification accuracy alone is insufficient; instead, there is a need for learning frameworks that jointly address feature representation quality and decision reliability.

To address these challenges, this study proposes a learning framework that integrates discriminative representation learning with confidence-aware decision regulation. Rather than focusing solely on improving classification accuracy, the proposed approach emphasizes reliable decision behavior under uncertainty. Specifically, supervised contrastive learning is employed to structure the embedding space and enhance class separability, thereby improving the robustness of learned acoustic representations [6]. In addition, a reward-driven decision mechanism is introduced to regulate predictive confidence and enable adaptive deferral of uncertain cases through an expert query strategy. By incorporating decision cost into the optimization process, the framework encourages more reliable classification behavior when model uncertainty is high [7].

The primary contribution of this study lies in the development of a confidence-regulated learning framework that jointly combines supervised contrastive representation learning with reward-based decision optimization for heart murmur classification. Unlike conventional approaches that focus primarily on improving classification accuracy or feature extraction alone, the proposed method explicitly connects representation refinement and confidence-aware decision regulation within a unified framework. Through this integrated design, the framework aims to improve not only classification performance but also the reliability and clinical applicability of automated heart murmur classification systems.

To provide a clear assessment of the proposed framework, this study focuses on controlled comparisons within a unified experimental setting, thereby isolating the contributions of representation learning and confidence-regulated decision-making.

## 2 Related Work

Research on automated heart murmur classification has advanced significantly with the development of deep learning techniques, aiming to overcome the subjectivity and variability inherent in traditional auscultation [2,8]. Deep learning-based approaches analyze phonocardiogram (PCG) recordings to automatically extract discriminative features and improve diagnostic accuracy and consistency [3,9–11]. This section reviews three key areas relevant to this study: deep learning for PCG analysis, contrastive learning for feature enhancement, and reward-based approaches for confidence-aware decision regulation.

### 2.1 Deep Learning for PCG Analysis and Its Limitations

Deep learning has significantly improved heart murmur classification by enabling automatic feature extraction from heart sound signals [3,9–11]. Earlier approaches relied on handcrafted features such as

mel-frequency cepstral coefficients (MFCCs) and wavelet transformations. These traditional methods often struggled to generalize across datasets and required domain expertise [12–14].

More recent methods leverage Convolutional Neural Networks (CNNs) to learn discriminative spectral representations from PCG spectrograms, while Long Short-Term Memory (LSTM) networks and related architectures capture temporal dependencies in cardiac cycles [3,9,11,15]. Recent studies have further explored advanced murmur-classification architectures, including transformer-based and multi-scale designs. These models jointly exploit frequency-domain and time-domain characteristics to improve classification performance. However, differences in task formulation and evaluation settings make direct comparisons across studies difficult.

Despite these advancements, several limitations remain. Deep learning models often produce overconfident yet incorrect predictions. Such overconfidence poses a significant risk in clinical applications. These models can also exhibit degraded performance when applied to data from different populations, devices, or recording environments, reflecting the effects of dataset bias and limited generalization [5,16]. The lack of mechanisms to account for prediction uncertainty further limits their reliability in real-world deployment. These challenges highlight the need for approaches that go beyond accuracy and explicitly address decision reliability.

## **2.2 Contrastive Learning for Feature Enhancement**

Contrastive learning has emerged as an effective technique for improving representation quality in deep learning models. By encouraging semantically similar samples to be embedded closer together while pushing dissimilar samples apart, contrastive learning structures the feature space and enhances discriminability and generalization [6,17].

Supervised contrastive learning (SCL), which incorporates label information into this process, has demonstrated effectiveness in medical applications, including cardiac sound analysis [17,18]. By improving class separability in the embedding space, SCL facilitates the detection of subtle acoustic variations in heart sounds, thereby improving classification performance.

However, enhancing feature representations alone does not fully address the problem of unreliable predictions. Models that rely solely on improved feature separability may still exhibit overconfidence in uncertain cases, leading to incorrect yet highly confident outputs in safety-critical settings [10,19]. This limitation motivates the need for complementary mechanisms that regulate decision behavior in addition to improving representation quality.

## **2.3 Reward-Based Approaches for Confidence-Aware Decision Regulation**

Beyond representation learning, recent studies have explored decision-level strategies that incorporate feedback signals to improve reliability in medical prediction tasks [7,20]. These approaches introduce adaptive mechanisms that adjust model behavior based on prediction outcomes, enabling more flexible and context-aware decision-making.

In particular, reward-based optimization provides a framework for incorporating decision cost and prediction confidence into the learning process. By associating different outcomes with corresponding utilities, such approaches enable models to balance classification accuracy with the cost of incorrect or uncertain predictions. This perspective is especially relevant in medical applications, where the consequences of incorrect decisions can be significant.

Despite these developments, most existing heart murmur classification models primarily focus on improving classification accuracy and do not explicitly incorporate mechanisms for confidence-aware

decision regulation. To the best of our knowledge, reward-based strategies have not been effectively applied to static heart murmur classification tasks to enable adaptive deferral of uncertain predictions. This gap motivates the proposed approach, which integrates discriminative feature learning with confidence-regulated decision optimization in a unified framework.

### 3 Methodology

This study proposes a two-stage framework for heart murmur classification that jointly optimizes feature representation and decision reliability. The framework combines supervised contrastive learning (SCL) for structured representation learning with a reward-based optimization (RBO) mechanism for confidence-regulated decision-making.

#### 3.1 Dataset

This study utilizes the CirCor DigiScope heart sound dataset from the PhysioNet Challenge 2022 [21,22]. The dataset contains 3163 phonocardiogram (PCG) recordings from 963 pediatric patients across four auscultation sites. The recordings are labeled as Murmur Absent (73.8%), Murmur Present (19.0%), and Unknown (7.2%). Recordings range from 5 to 65 s in duration and were originally sampled at 4000 Hz. The dataset captures substantial physiological variability across patients, auscultation locations, and recording conditions. Murmur annotations are provided by experienced cardiac physiologists based on a combination of auditory assessment and visual inspection of phonocardiogram signals. This annotation process reflects real-world clinical practice, where both acoustic patterns and waveform characteristics are considered during diagnosis. In this study, the task is formulated as a binary classification problem by excluding the Unknown class to reduce label ambiguity. To address class imbalance, a class-weighted loss is applied.

A strict patient-wise split is used. The training set contains 9495 segments from 611 patients (79.7% Absent, 20.3% Present), and the validation set contains 4012 segments from 263 patients (81.0% Absent, 19.0% Present). This patient-wise split prevents data leakage and ensures that the model is evaluated on previously unseen individuals, which is critical for reliable clinical deployment.

#### 3.2 Preprocessing

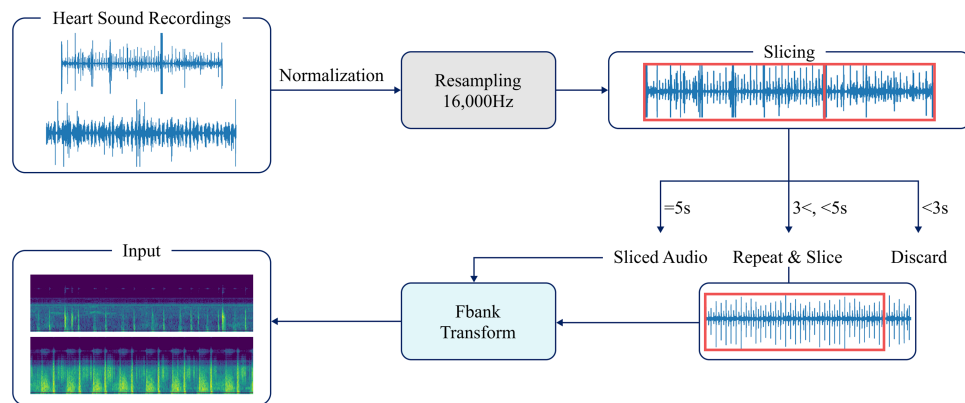
All recordings are resampled to 16 kHz and normalized to the range  $[-1, 1]$  to ensure a consistent preprocessing pipeline and adequate temporal resolution. Although murmur-related frequency components are primarily below 900 Hz, a higher sampling rate improves waveform representation and supports stable time–frequency feature extraction using Mel filter bank representations. This choice preserves diagnostically relevant information while maintaining computational efficiency.

Each recording is segmented into non-overlapping 5-s clips. Segments shorter than 3 s are discarded, while longer residual segments are repeated and truncated. This segmentation strategy is designed to balance temporal coverage and data efficiency. Fixed-length segments allow stable batch processing while preserving sufficient cardiac-cycle information for reliable murmur detection. In particular, each segment typically contains multiple cardiac cycles, allowing the model to learn robust acoustic representations from variable and non-stationary heart sound patterns within a standardized input length. Discarding very short segments prevents the introduction of incomplete or noisy acoustic patterns that may degrade model performance.

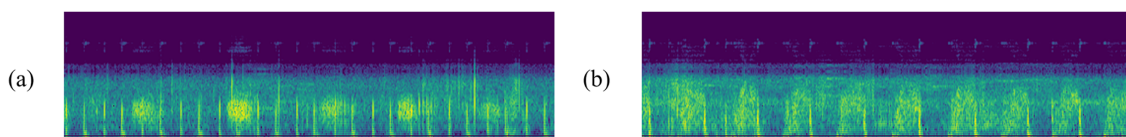
Following segmentation, Mel filter bank (Fbank) features are extracted from each audio clip to obtain a time–frequency representation of heart sounds. These features are computed using the Kaldi implementation, in which each waveform is divided into overlapping frames using a Hanning window with a 10 ms

frame shift. A set of 64 Mel-scaled triangular filters is applied to estimate the energy distribution across frequency bands.

Unlike conventional pipelines that rely on Short-Time Fourier Transform (STFT)-based spectrograms, this approach directly estimates Mel-scale energy features from the waveform. This results in a more compact representation while retaining perceptually relevant frequency information. Prior studies have shown that Mel filter bank representations are robust to noise and effective at capturing subtle acoustic variations in heart sounds [23]. This robustness is particularly important in real-world auscultation scenarios, where recordings may be affected by environmental noise, sensor variability, and patient movement. Overall, the preprocessing pipeline is designed to standardize input representations while preserving clinically relevant acoustic characteristics. This design aims to improve both model stability and generalization performance. An overview of the preprocessing pipeline is illustrated in Fig. 1. Representative examples of the resulting Fbank features are shown in Fig. 2.



**Figure 1:** Preprocessing pipeline for heart sound recordings. Raw audio signals are normalized and resampled to 16,000 Hz, segmented into fixed-length clips, and transformed into Mel filter bank (Fbank) representations. Segments shorter than 3 s are discarded, while residual segments are repeated and truncated to maintain a consistent input length.



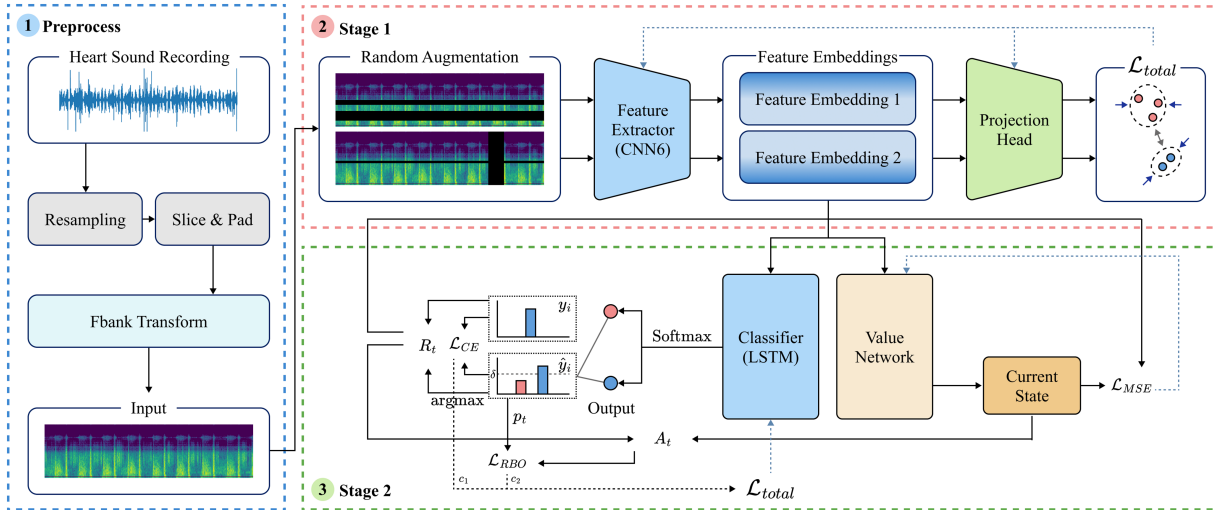
**Figure 2:** Examples of Mel filter bank representations of heart sounds. (a) Murmur Absent sample showing regular cardiac cycles; (b) Murmur Present sample exhibiting irregular spectral patterns and additional acoustic components associated with pathological murmurs.

### 3.3 Overall Framework

This study presents a structured two-stage training framework for heart murmur classification, comprising a representation learning stage based on supervised contrastive learning (Stage 1) and a confidence-regulated decision-making stage using reward-based optimization (Stage 2). The framework is designed not only to extract discriminative acoustic features from heart sound recordings but also to regulate prediction confidence under uncertainty, which is essential for reliable clinical deployment.

As illustrated in Fig. 3, the pipeline begins with preprocessing steps that convert variable-length recordings into standardized time–frequency representations. These inputs are then used to learn a structured embedding space with improved intra-class compactness and inter-class separability. In the second stage,

the model learns to regulate its own predictions through a confidence-based decision mechanism, enabling it to either make autonomous predictions or defer uncertain cases.



**Figure 3:** Overall architecture of the proposed framework. The model first learns discriminative feature embeddings using supervised contrastive learning (Stage 1), followed by confidence-regulated classification with reward-based optimization (Stage 2), including an auxiliary value network for decision evaluation.

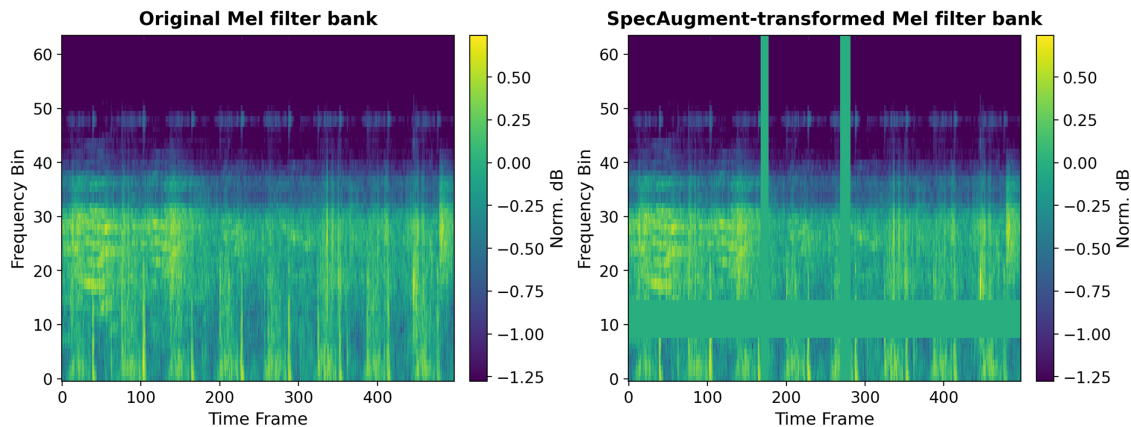
The training follows a sequential design. Stage 1 focuses on constructing a well-separated feature space, while Stage 2 learns decision behavior conditioned on confidence. This separation allows the model to first establish stable representations before introducing decision-level optimization, improving both training stability and interpretability.

### 3.4 Stage 1: Supervised Contrastive Representation Learning

The first stage of training focuses on learning robust and discriminative feature representations through supervised contrastive learning (SCL). Given an input segment  $X_i \in \mathbb{R}^{T \times F}$ , where  $T$  and  $F$  denote the time and frequency dimensions of the Mel filter bank representation, two augmented views  $\tilde{X}_i^{(1)}$  and  $\tilde{X}_i^{(2)}$  are generated using time and frequency masking as proposed in SpecAugment [24]. These augmentations randomly suppress limited regions along the temporal and spectral axes, simulating variability in recording conditions. Fig. 4 illustrates a comparison between the original Mel filter bank and its SpecAugment-transformed version. This strategy encourages the model to focus on invariant, class-relevant acoustic patterns while improving robustness to noise and localized distortions without systematically eliminating diagnostically relevant murmur information.

The augmented views are processed through a shared feature extractor  $f(\cdot)$ , implemented using CNN6 from the Pretrained Audio Neural Networks (PANNs) framework [25]. CNN6 consists of four convolutional blocks, each comprising a  $5 \times 5$  convolution, batch normalization, ReLU activation, and pooling operations, enabling hierarchical extraction of time–frequency features. A layer-wise summary of the architecture is provided in Table 1. After the final convolutional block, the feature map is aggregated by temporal mean pooling followed by frequency-domain max and mean pooling, and the resulting representations are combined to form the final embedding vector. The network is initialized with pretrained weights from AudioSet, a large-scale audio classification dataset containing over 5000 h of labeled audio across 527 sound

classes [26]. Rather than being used as a fixed feature extractor, the pretrained CNN6 encoder is fully fine-tuned during Stage 1 on the target PCG dataset. This design allows the model to benefit from a stable acoustic initialization while adapting its representations to the temporal and spectral characteristics of heart sound signals. Although AudioSet and PCG recordings differ in domain, the pretrained initialization provides general acoustic priors that can be refined through task-specific training.



**Figure 4:** Comparison between the original Mel filter bank and the SpecAugment-transformed representation of the same heart sound segment, illustrating localized time and frequency masking applied to partial regions of the input.

**Table 1:** Layer-wise configuration of the CNN6.

Stage	Layer Type	Configuration	Channels
Input	Mel filter bank	Single-channel input	1
Block 1	ConvBlock $5 \times 5$	Conv $5 \times 5$ + BN + ReLU + AvgPool	64
Block 2	ConvBlock $5 \times 5$	Conv $5 \times 5$ + BN + ReLU + AvgPool	128
Block 3	ConvBlock $5 \times 5$	Conv $5 \times 5$ + BN + ReLU + AvgPool	256
Block 4	ConvBlock $5 \times 5$	Conv $5 \times 5$ + BN + ReLU + AvgPool	512
Pooling	Aggregation	Mean (time) + Max/Mean (freq)	512
Output	Embedding	Final feature vector	512

The feature extractor outputs high-dimensional embeddings  $Z_i^{(1)}$  and  $Z_i^{(2)}$ , which are passed through a projection head  $g(\cdot)$ , modeled as a two-layer multilayer perceptron (MLP) with ReLU activation. The projection head maps the encoder output to a 128-dimensional projected representation for contrastive learning, yielding projected representations  $P_i^{(1)}$  and  $P_i^{(2)}$ . Supervised contrastive learning is applied to these projected representations, while the encoder output  $Z$  is retained for downstream classification. Although the contrastive objective is optimized on  $P$ , the encoder and projection head are trained jointly during Stage 1. Consequently, the discriminative structure encouraged in the projected representations is propagated back to the encoder through gradient updates, improving the quality of the encoder representation  $Z$ . This design enables contrastive optimization through the projection head while preserving  $Z$  as the primary representation for subsequent classification.

Supervised contrastive learning is applied to structure the embedding space by leveraging label information to define semantically meaningful positive pairs [27]. This formulation encourages samples from the same class to cluster together while separating samples from different classes. To address class imbalance in

the dataset, class weights are incorporated into the supervised contrastive loss, increasing the contribution of underrepresented classes during training.

Inspired by the contrastive repulsion mechanism proposed in [28], an additional repulsion term  $\mathcal{L}_{repulsion}$  is incorporated to explicitly penalize high similarity between samples from different classes. The repulsion loss is defined as:

$$\mathcal{L}_{repulsion} = -\frac{1}{|N(i)|} \sum_{j \in N(i)} sim(p_i, p_j), \quad (1)$$

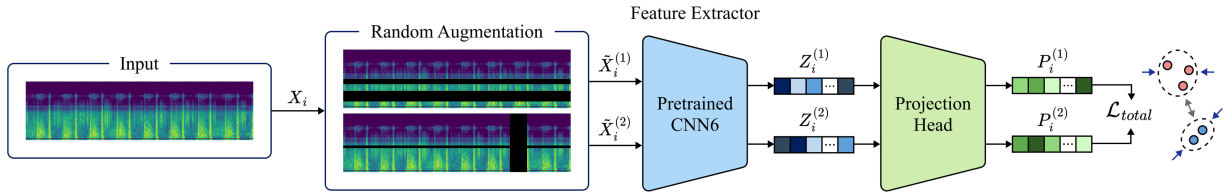
where  $N(i)$  denotes the set of negative pairs from different classes. This term further enforces inter-class separation by reducing similarity among dissimilar samples.

The total Stage 1 objective is defined as:

$$\mathcal{L}_{Stage1} = \mathcal{L}_{SCL} + \lambda_{repulsion} \mathcal{L}_{repulsion}. \quad (2)$$

The balancing coefficient  $\lambda_{repulsion}$  controls the relative contribution of the repulsion term. The combination of class-aware attraction and contrastive repulsion results in a well-structured embedding space with improved intra-class compactness and inter-class separability.

The overall architecture and training flow for this stage are illustrated in Fig. 5. The learned representation serves as the foundation for Stage 2, where classification decisions are further refined through confidence-regulated optimization.



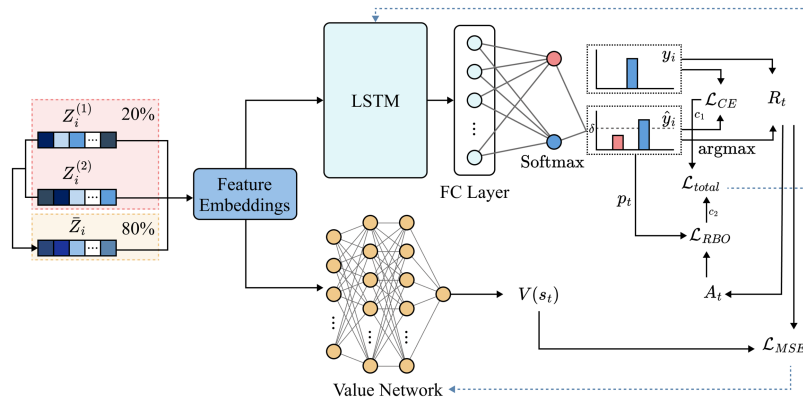
**Figure 5:** Stage 1 pipeline. Two augmented views are processed by a shared feature extractor (CNN6) and then passed through a projection head. The model is trained using supervised contrastive learning with an additional repulsion loss to improve inter-class separation.

### 3.5 Stage 2: Confidence-Regulated Decision Optimization

Once feature representations are refined, Stage 2 focuses on classification and confidence-regulated decision-making. A bidirectional LSTM-based classifier is employed to process the encoder-derived embeddings. In this framework, the classifier operates on learned embeddings rather than on raw PCG waveforms. Accordingly, the LSTM serves as a trainable transformation and classification module built on top of the learned representation. This design provides a flexible classifier head while preserving the structured representations learned in Stage 1. The classifier consists of stacked bidirectional LSTM layers followed by a fully connected (FC) layer. The output logits are then transformed into class probabilities using a Softmax function. A dropout layer with a rate of 0.3 is applied to reduce overfitting and improve generalization. The overall architecture and training process are illustrated in Fig. 6.

The classifier receives input embeddings constructed from the encoder outputs after Stage 1 training. Since two augmented views are generated for each original input during Stage 1, the encoder produces two embeddings per input, denoted as  $Z_i^{(1)}$  and  $Z_i^{(2)}$ . Using both embeddings directly in Stage 2 would double the number of training instances. To preserve the original number of inputs, the two embeddings

are consolidated into a single embedding for each input before classification. Specifically, 80% of the inputs are formed using the averaged embedding  $\bar{Z}_i$ , while the remaining 20% use a randomly selected embedding from either  $Z_i^{(1)}$  or  $Z_i^{(2)}$ . The 80:20 ratio was fixed throughout all experiments and introduced as a heuristic regularization choice to combine representation stability with modest stochastic variation. This design provides a stable representation for most inputs while retaining limited stochastic variation derived from augmentation. The resulting embeddings are processed by the LSTM to produce prediction logits  $\hat{y}_i$ , which are compared with the ground truth  $y_i$  to compute the classification loss  $\mathcal{L}_{CE}$  using the standard cross-entropy objective. To address class imbalance, class weights are incorporated into the cross-entropy loss for more balanced learning across classes. The encoder is frozen during Stage 2 to preserve the structured feature representations learned in Stage 1. This prevents changes to the embedding space and ensures consistent feature distributions during classifier training.



**Figure 6:** Stage 2 pipeline. Feature embeddings are passed to an LSTM classifier. Predictions are optimized using cross-entropy and reward-based objectives, while a value network is trained separately using mean squared error.

To regulate prediction reliability, a confidence threshold  $\delta$  is applied to the classifier's Softmax output. If the maximum predicted probability satisfies  $\max(\hat{y}_i) \geq \delta$ , the prediction is accepted; otherwise, the model defers the decision. This confidence-based gating mechanism reduces unreliable predictions and improves decision trustworthiness during training.

To further refine decision behavior under uncertainty, reward-based optimization (RBO) is applied. The classifier output is treated as a categorical distribution over the prediction classes, and the confidence-regulated decision objective is optimized using a stabilized strategy inspired by Proximal Policy Optimization (PPO) [29]. The probability ratio is defined as

$$p_t = \frac{\pi_\theta(a_t | s_t)}{\pi_{\theta_{\text{old}}}(a_t | s_t)}, \quad (3)$$

where  $\pi_\theta(a_t | s_t)$  and  $\pi_{\theta_{\text{old}}}(a_t | s_t)$  denote the current and previous probabilities assigned to decision  $a_t$  for input  $s_t$ , respectively. The advantage is defined as

$$A_t = R_t - V(s_t), \quad (4)$$

where  $R_t$  is the reward assigned based on the classification outcome, and  $V(s_t)$  is the value estimated by the auxiliary value network. The value network is implemented as a multilayer perceptron (MLP) that takes the encoder-derived embedding as input and outputs a scalar value estimate. It consists of three fully connected layers with dimensions  $512 \rightarrow 256 \rightarrow 128 \rightarrow 1$ , with ReLU activation functions applied after the

first two layers. The final output represents the estimated value associated with the input embedding and is used to compute the advantage term in Stage 2. Since heart murmur classification is formulated as a single-step static decision problem, the learning process does not involve sequential state transitions. As a result, the advantage formulation simplifies by excluding future return estimation, allowing direct optimization based on immediate outcomes.

The reward-based decision optimization objective for the classifier is defined as

$$\mathcal{L}_{\text{RBO}} = -\mathbb{E}_t [\min(p_t A_t, \text{clip}(p_t, 1-\epsilon, 1+\epsilon)A_t)] - \lambda_e \mathcal{L}_{\text{entropy}}, \quad (5)$$

where  $\epsilon$  is the clipping parameter,  $\lambda_e$  controls the contribution of entropy regularization, and  $\mathcal{L}_{\text{entropy}}$  is defined as

$$\mathcal{L}_{\text{entropy}} = \mathbb{E}_t [\mathcal{H}(\pi_\theta(\cdot | s_t))]. \quad (6)$$

The overall classifier objective in Stage 2 is defined as

$$\mathcal{L}_{\text{Stage2}} = c_1 \mathcal{L}_{\text{CE}} + c_2 \mathcal{L}_{\text{RBO}}, \quad (7)$$

where  $c_1$  and  $c_2$  control the balance between supervised classification and confidence-regulated decision optimization.

The value network is optimized separately using a mean squared error loss,

$$\mathcal{L}_{\text{value}} = \text{MSE}(V(s_t), R_t), \quad (8)$$

which aligns the predicted value with the observed reward signal. This separation improves optimization stability by decoupling confidence-regulated decision optimization from value estimation.

### 3.6 Reward Design and Adaptive Threshold

To enable confidence-regulated decision-making, the RBO framework incorporates a dynamically adjusted confidence threshold and a structured reward design.

The confidence threshold  $\delta$  is updated using a decreasing sigmoid function over training epochs:

$$\delta(\text{epoch}) = \frac{1}{1 + e^{k(\text{epoch}-\text{shift})}} + \delta_{\min}. \quad (9)$$

This schedule begins with a high, conservative threshold, encouraging frequent deferral of uncertain predictions. As training progresses, the threshold gradually decreases, allowing the model to make more autonomous decisions as its predictive confidence improves. The parameters  $k$ ,  $\text{shift}$ , and  $\delta_{\min}$  control the decay rate, transition point, and minimum threshold value, respectively.

The reward signal is designed to jointly account for prediction correctness, confidence, and class imbalance:

$$R_t = R_t^{\text{base}} + 0.1(1 - c_t), \quad (10)$$

where  $c_t$  denotes the prediction confidence. The base reward is defined according to prediction correctness and confidence level:

$$R_t^{\text{base}} = \begin{cases} w_{y_t} + 0.5, & \text{if } \hat{y}_t = y_t \text{ and } c_t \geq \delta \\ w_{y_t}, & \text{if } \hat{y}_t = y_t \text{ and } c_t < \delta \\ -w_{y_t} - 2.0, & \text{if } \hat{y}_t \neq y_t \text{ and } c_t \geq \delta \\ -w_{y_t}, & \text{if } \hat{y}_t \neq y_t \text{ and } c_t < \delta \end{cases} \quad (11)$$

Here,  $w_{y_t}$  represents the class-dependent weighting factor used to address class imbalance. The reward structure assigns strong penalties to overconfident incorrect predictions, discouraging unreliable decisions, while rewarding correct predictions with an additional incentive when they are made confidently.

The additional term  $0.1(1 - c_t)$  acts as a confidence-based regularizer. This term provides a small incentive for lower-confidence predictions, encouraging the model to remain cautious near the decision boundary and mitigating the tendency toward uniformly overconfident outputs.

This reward formulation guides the model to balance prediction accuracy with confidence, resulting in more reliable and calibrated decision-making in clinically uncertain scenarios.

## 4 Experiment

To evaluate the effectiveness of the proposed two-stage framework, a comparative analysis is conducted between the baseline model and the proposed method integrating supervised contrastive learning (SCL) and reward-based optimization (RBO). The objective is to assess whether structured representation learning and confidence-regulated decision-making provide complementary benefits in improving both classification performance and decision reliability.

Experiments are performed on the validation split constructed using a strict patient-wise partitioning strategy, ensuring that no patient appears in both training and validation sets. This setup reflects realistic clinical deployment, where models must generalize to previously unseen individuals.

### 4.1 Evaluation Metrics

Model performance is evaluated using sensitivity, specificity, and their arithmetic mean, referred to as the Score. All reported metrics are computed at the segment level on the segmented inputs obtained after the patient-wise train-validation split described above, rather than at the patient level. These metrics are widely used in clinical classification tasks, where both detecting pathological cases and avoiding false alarms are equally critical.

Sensitivity measures the proportion of correctly identified positive cases (murmur present), reflecting the model's ability to detect clinically relevant abnormalities. Specificity measures the proportion of correctly identified negative cases (murmur absent), indicating the model's ability to avoid false positives. Given the inherent class imbalance in the dataset, reliance on a single metric can be misleading. Therefore, the Score, defined as the arithmetic mean of sensitivity and specificity, is used as the primary evaluation criterion to provide a balanced assessment of performance. We note that the official George B. Moody PhysioNet Challenge 2022 adopted a different task formulation and evaluation protocol. In contrast to the official challenge setting, this study formulates murmur classification as a binary task and evaluates performance using sensitivity, specificity, and their mean Score. Accordingly, the reported results are intended to support controlled comparison within the proposed framework rather than direct comparison with official challenge rankings.

## 4.2 Computational Setup and Hyperparameters

The implementation is carried out in Python 3.9 using PyTorch 2.6.0, with Torchaudio 2.6.0 and NumPy 1.24.3. All experiments are conducted on a system equipped with dual NVIDIA GeForce RTX 4090 GPUs (48 GB total VRAM), an Intel<sup>®</sup> Core™ i9-13900KF CPU, and 64 GB of RAM, running Ubuntu 22.04.2 LTS. A batch size of 8 is used for training.

The training process follows the two-stage framework described in Section 3. In Stage 1, the CNN6 encoder and projection head are fine-tuned using the Adam optimizer with a learning rate of 0.0005. The total loss combines supervised contrastive learning and the repulsion term, with the balancing coefficient  $\lambda_{\text{repulsion}}$  set to 0.3.

The bidirectional LSTM classifier and the value network are trained using the Adam optimizer with a learning rate of 0.0005. A dropout rate of 0.3 is applied to the LSTM classifier to mitigate overfitting. The total classification loss combines the cross-entropy loss and the reward-based optimization objective, with coefficients  $c_1 = 0.3$  and  $c_2 = 0.7$ , respectively, assigning greater weight to the RBO objective.

The clipping parameter  $\epsilon$  for stabilized updates is set to 0.2. The confidence threshold  $\delta$  is dynamically adjusted during training using a decreasing sigmoid function, allowing the model to transition from conservative to more autonomous decision-making behavior as training progresses.

## 4.3 Performance Comparison

The performance comparison between the baseline and the proposed method variants is presented in Table 2. The baseline model consists of a CNN6 encoder and an LSTM classifier trained using standard supervised learning.

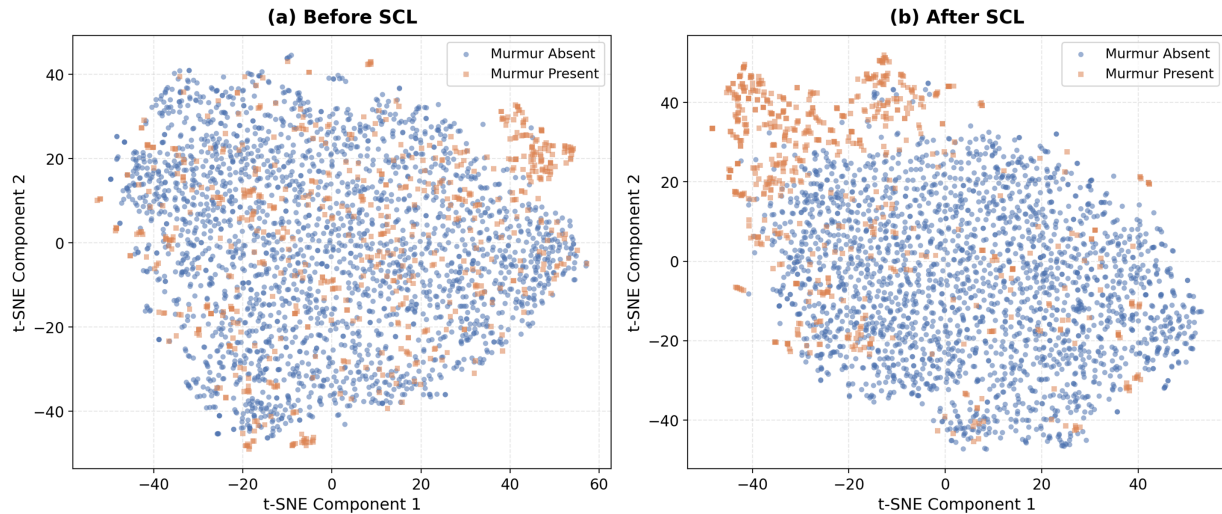
**Table 2:** Comparative performance of baseline and proposed methods on the validation set in terms of sensitivity, specificity, and score.

Methods	Sensitivity	Specificity	Score
Baseline (CNN6 + LSTM)	0.6745	0.9382	0.8064
Baseline + SCL (Feature Learning)	0.6509	<b>0.9563</b>	0.8036
Baseline + RBO (Decision Regulation)	0.7021	0.9317	0.8170
<b>Proposed Method (SCL + RBO)</b>	<b>0.7284</b>	0.9182	<b>0.8233</b>

The PhysioNet 2022 dataset is a widely used benchmark for heart murmur classification. In this study, the evaluation focuses on controlled comparisons between the baseline and proposed variants to isolate the impact of supervised contrastive learning and confidence-regulated decision optimization. Rather than benchmarking existing methods across varying experimental settings, this design enables a clearer assessment of each component's contribution within a consistent framework.

The baseline model achieves a Score of 0.8064, serving as a reference for comparison. Introducing SCL alone improves feature representation by enhancing inter-class separation, yielding the highest specificity of 0.9563. To further analyze this effect, we visualize the embedding space before and after applying SCL using t-SNE, as shown in Fig. 7. The baseline embeddings exhibit substantial overlap between classes, indicating limited discriminative structure. In contrast, the SCL-enhanced embeddings demonstrate improved intra-class compactness and clearer inter-class separation, despite inherent acoustic similarity between heart sound patterns, suggesting that SCL effectively structures the feature space for discriminative learning. This behavior is consistent with the observed reduction in false positives. However, the improvement in specificity

does not translate into a higher overall Score, suggesting that representation learning alone is insufficient to optimize decision behavior.



**Figure 7:** t-SNE visualization of feature embeddings before and after supervised contrastive learning (SCL). (a) Before SCL, the embeddings exhibit substantial overlap between murmur-absent and murmur-present samples, indicating limited discriminative structure. (b) After SCL, the embeddings show improved intra-class compactness and clearer inter-class separation, despite the inherent acoustic similarity of heart sound signals.

In contrast, applying RBO alone significantly increases sensitivity from 0.6745 to 0.7021, thereby improving the detection of murmur cases. This improvement highlights the effectiveness of confidence-regulated decision optimization in encouraging the model to identify positive cases more aggressively. However, this is accompanied by a slight reduction in specificity, indicating a trade-off between sensitivity and false-positive control.

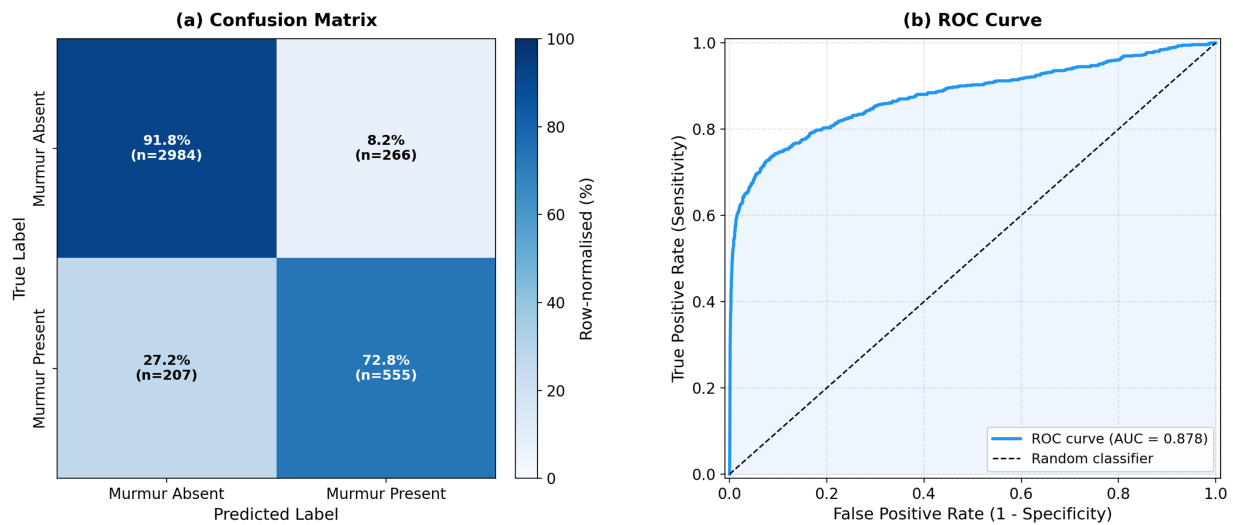
The proposed method, which integrates both SCL and RBO, achieves the highest Score of 0.8233, representing a relative improvement of 2.1% over the baseline. This result demonstrates that feature representation learning and decision regulation provide complementary benefits. SCL enhances the quality of the learned feature space, while RBO guides the model toward more reliable decision-making under uncertainty. The results confirm that jointly optimizing representation and decision behavior leads to a more balanced and clinically reliable classification system. To further visualize the classification performance of the proposed method, Fig. 8 presents the confusion matrix and ROC curve on the validation set. The confusion matrix shows strong class-wise performance for both murmur-absent and murmur-present cases, while the ROC curve further confirms the model's discriminative capability across decision thresholds.

#### 4.4 Discussion

The experimental results demonstrate that jointly optimizing feature representation and confidence-regulated decision-making leads to consistent improvements in heart murmur classification performance. The proposed framework not only enhances overall classification performance but also achieves a better balance between sensitivity and specificity, which is critical in clinical diagnostic settings.

**Impact of Feature Representation Learning (SCL):** The Baseline + SCL variant achieved the highest specificity among all methods. This indicates that supervised contrastive learning effectively improves inter-class separability in the embedding space by clustering semantically similar samples while pushing dissimilar

ones apart. As a result, the model becomes more conservative in its predictions, reducing false positives and improving reliability in identifying murmur-absent cases. This behavior is particularly desirable in clinical environments, where minimizing unnecessary alarms is important to avoid overdiagnosis and additional testing.



**Figure 8:** Segment-level confusion matrix and ROC curve of the proposed method on the validation set. (a) Row-normalized confusion matrix. (b) ROC curve with an AUC of 0.878.

However, the improvement in specificity does not directly translate to a higher overall Score. This suggests that while SCL enhances feature quality, it does not explicitly regulate how these features are used during decision-making. In particular, the increased separation can lead to more conservative decision boundaries, making the model less sensitive to ambiguous or borderline murmur cases, which may limit sensitivity. Consequently, the model remains limited in its ability to adapt its predictions under uncertainty, highlighting the need for an additional mechanism that governs decision behavior.

Although supervised contrastive learning improves representation quality within the labeled dataset used in this study, its generalization benefits may be further enhanced through large-scale unlabeled pretraining. In particular, self-supervised objectives such as masked prediction or related representation learning strategies on larger heart sound datasets could provide a stronger initialization for Stage 1 and further improve the effectiveness of the confidence-regulated decision optimization applied in Stage 2. Exploring such pretraining strategies remains an important direction for future work.

**Impact of Confidence-Regulated Decision Optimization (RBO):** The Baseline + RBO variant shows a significant increase in sensitivity, indicating improved detection of murmur-present cases. This improvement arises from the confidence-regulated decision mechanism, which dynamically adjusts the acceptance threshold during training. Initially, the model adopts a conservative strategy by deferring uncertain predictions, effectively reducing overconfident errors. As training progresses and the model becomes more competent, the threshold gradually decreases, allowing the model to make more autonomous predictions.

This adaptive process encourages the model to explore decision boundaries more effectively and reduces the tendency to ignore difficult positive cases. As a result, the model becomes more sensitive to subtle murmur patterns, improving its ability to detect clinically relevant abnormalities. However, this gain in sensitivity is accompanied by a slight reduction in specificity, reflecting the inherent trade-off between detecting positive cases and avoiding false positives.

**Complementary Effects of SCL and RBO:** The full proposed method, which combines SCL and RBO, achieves the highest overall Score, demonstrating that the two components provide complementary benefits. SCL enhances the structure and separability of the feature space, ensuring that representations are discriminative and robust. RBO, on the other hand, regulates how these features are utilized during prediction, guiding the model to make more reliable decisions under uncertainty. By integrating these two components, the framework effectively addresses both representation-level and decision-level limitations present in conventional deep learning approaches. The improved sensitivity indicates better detection of murmur cases, while the maintained specificity ensures that false positives remain controlled. This balance is essential for real-world deployment, where both missed diagnoses and false alarms carry significant clinical implications. Furthermore, the proposed framework is not constrained to a specific backbone architecture. Incorporating more advanced feature extraction or classification components is expected to yield further performance improvements, which remains a promising direction for future work.

It should also be noted that the encoder is initialized using AudioSet pretraining, which may introduce a domain mismatch with PCG signals. Although the encoder is fully fine-tuned on the target dataset, a direct comparison with random initialization was beyond the scope of this study and remains an important direction for future investigation.

These findings highlight that improving feature representation alone is insufficient for achieving optimal performance. Instead, incorporating a confidence-aware decision-making mechanism is crucial for developing reliable and clinically applicable models.

## 5 Conclusion

This study proposes a Joint Feature-Decision Learning Framework that integrates supervised contrastive learning (SCL) and reward-based optimization (RBO) to enhance feature discriminability and improve classification reliability under clinical uncertainty. Experimental results demonstrate that the proposed method outperforms the baseline, achieving a Score of 0.8233. The findings confirm the complementary roles of the two components: SCL improves class separability, leading to higher specificity, while RBO regulates prediction confidence, enhancing sensitivity and enabling more reliable decision-making. By jointly optimizing feature representation and decision behavior, the proposed framework provides a balanced and effective approach for heart murmur classification. This work highlights the importance of integrating structured representation learning with confidence-aware decision mechanisms for developing clinically applicable models. Future work will focus on refining the reward formulation with class-aware penalties, exploring uncertainty-aware value estimation, and extending the framework to more complex settings, such as multi-label and open-set classification.

**Acknowledgement:** Not applicable.

**Funding Statement:** This work was supported by Innovative Human Resource Development for Local Intellectualization Program through the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (IITP-2026-RS-2022-00156360).

**Author Contributions:** The authors confirm their contributions to the paper as follows: HyeSun Chang: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing—Original Draft Preparation, Writing—Review and Editing, Visualization. Sangjun Lee: Supervision, Writing—Review and Editing, Project Administration, Funding Acquisition. All authors reviewed and approved the final version of the manuscript.

**Availability of Data and Materials:** The data that support the findings of this study are publicly available in the PhysioNet repository (<https://physionet.org/>), specifically the CirCor DigiScope dataset from the PhysioNet Challenge 2022.

**Ethics Approval:** This study utilizes publicly available anonymized data and does not involve direct human or animal subject interaction. Therefore, ethical approval is not required.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

PCG	Phonocardiogram
SCL	Supervised Contrastive Learning
RBO	Reward-Based Optimization
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
GAP	Global Average Pooling
FC	Fully Connected
CE	Cross-Entropy

## References

1. Ameen A, Fattoh IE, Abd El-Hafeez T, Ahmed K. Advances in ECG and PCG-based cardiovascular disease classification: a review of deep learning and machine learning methods. *J Big Data*. 2024;11(1):159. doi:10.1186/s40537-024-01011-7.
2. Omarov B, Tuimebayev A, Abdrakhmanov R, Eskarayeva B, Sultan D, Aidarov K. Digital stethoscope for early detection of heart disease on phonocardiography data. *Int J Adv Comput Sci Appl*. 2023;14(9):716–24.
3. Kamson AP, Crecsilla Lewis M, Vishnu Sunil BN, Jeevannavar SS, Sawant A, Ghosh PK. E2E multi-scale CNN with LSTM for murmur detection in PCG or noise identification. In: 2023 International Conference on Electrical, Communication and Computer Engineering (ICECCE). Piscataway, NJ, USA: IEEE; 2023. p. 1–6.
4. Goetz L, Seedat N, Vandersluis R, van der Schaar M. Generalization—a key challenge for responsible AI in patient-facing clinical applications. *npj Digit Med*. 2024;7:126. doi:10.1038/s41746-024-01127-3.
5. Norori N, Hu Q, Aellen FM, Faraci FD, Tzovara A. Addressing bias in big data and AI for health care: a call for open science. *Patterns*. 2021;2(10):100347. doi:10.1016/j.patter.2021.100347.
6. Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. In: Proceedings of the 37th International Conference on Machine Learning, ICML'20; 2020 Jul 12–18; Vienna, Austria. p. 1597–607.
7. Gayathri R, Sangeetha SKB, Mathivanan SK, Rajadurai H, Benjula Anbu Malar MB, Mallik S, et al. Enhancing heart disease prediction with reinforcement learning and data augmentation. *Syst Soft Comput*. 2024;6:200129. doi:10.1016/j.sasc.2024.200129.
8. Chorba JS, Shapiro AM, Le L, Maidens J, Prince J, Pham S, et al. Deep learning algorithm for automated cardiac murmur detection via a digital stethoscope platform. *J Am Heart Assoc*. 2021;10:e019905. doi:10.1101/2020.04.01.20050518.
9. Alkhodari M, Hadjileontiadis LJ, Khandoker AH. Identification of congenital valvular murmurs in young patients using deep learning-based attention transformers and phonocardiograms. *IEEE J Biomed Health Inform*. 2024;28(4):1803–14. doi:10.1109/jbhi.2024.3357506.
10. Manshadi OD, Mihandoost S. Murmur identification and outcome prediction in phonocardiograms using deep features based on Stockwell transform. *Sci Rep*. 2024;14:7592. doi:10.1038/s41598-024-58274-6.
11. Lu H, Yip JB, Steigleder T, Griefshammer S, Heckel M, Jami NVSJ, et al. A lightweight robust approach for automatic heart murmurs and clinical outcomes classification from phonocardiogram recordings. In: 2022 Computing in Cardiology (CinC); 2022 Sep 4–7; Tampere, Finland. p. 1–4.

12. Noman F, Salleh SH, Ting CM, Samdin SB, Ombao H, Hussain H. A Markov-switching model approach to heart sound segmentation and classification. *IEEE J Biomed Health Inform.* 2020;24(3):705–16. doi:10.1109/jbhi.2019.2925036.
13. Nogueira DM, Ferreira CA, Gomes EF, Jorge AM. Classifying heart sounds using images of motifs, MFCC and temporal features. *J Med Syst.* 2019;43:168. doi:10.1007/s10916-019-1286-5.
14. Khan FA, Abid A, Khan MS. Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features. *Physiol Meas.* 2020;41:055006. doi:10.1088/1361-6579/ab8770.
15. Das S, Dandapat S. Heart murmur severity stages classification using Multikernel residual CNN. *IEEE Sens J.* 2024;24:13019–27. doi:10.1109/jsen.2024.3373226.
16. Dawood T, Chen C, Sidhu BS, Ruijsink B, Gould J, Porter B, et al. Uncertainty aware training to improve deep learning model calibration for classification of cardiac MR images. *Med Image Anal.* 2023;88:102861. doi:10.1016/j.media.2023.102861.
17. Antoni L, Bruoth E, Bugata P, Bugata P, Gajdoš D, Hudák D, et al. Murmur identification using supervised contrastive learning. In: *2022 Computing in Cardiology (CinC); 2022 Sep 4–7; Tampere, Finland.* p. 1–4.
18. Le D, Truong S, Brijesh P, Adjeroh DA, Le N. sCL-ST: supervised contrastive learning with semantic transformations for multiple lead ECG arrhythmia classification. *IEEE J Biomed Health Inform.* 2023;27(6):2818–28. doi:10.1109/jbhi.2023.3246241.
19. Kompa B, Snoek J, Beam AL. Second opinion needed: communicating uncertainty in medical machine learning. *npj Digit Med.* 2021;4(1):4. doi:10.1038/s41746-020-00367-3.
20. Jayaraman P, Desman J, Sabounchi M, Nadkarni GN, Sakhuja A. A primer on reinforcement learning in medicine for clinicians. *npj Digit Med.* 2024;7:337. doi:10.1038/s41746-024-01316-0.
21. Reyna MA, Kiarashi Y, Elola A, Oliveira J, Renna F, Gu A, et al. Heart murmur detection from phonocardiogram recordings: the George B. Moody PhysioNet Challenge 2022. *PLoS Digit Health.* 2023;2:e0000324.
22. Oliveira J, Renna F, Costa PD, Nogueira M, Oliveira C, Ferreira C, et al. The CirCor DigiScope dataset: from murmur detection to murmur classification. *IEEE J Biomed Health Inform.* 2022;26(6):2524–35.
23. Azam FB, Ansari MI, Nuhash SISK, McLane I, Hasan T. Cardiac anomaly detection considering an additive noise and convolutional distortion model of heart sound recordings. *Artif Intell Med.* 2022;133:102417. doi:10.1016/j.artmed.2022.102417.
24. Park DS, Chan W, Zhang Y, Chiu CC, Zoph B, Cubuk ED, et al. SpecAugment: a simple data augmentation method for automatic speech recognition. *arXiv:1904.08779.* 2019.
25. Kong Q, Cao Y, Iqbal T, Wang Y, Wang W, Plumbley MD. PANNs: large-scale pretrained audio neural networks for audio pattern recognition. *IEEE/ACM Trans Audio Speech Lang Process.* 2020;28:2880–94.
26. Gemmeke JF, Ellis DPW, Freedman D, Jansen A, Lawrence W, Moore RC, et al. Audio set: an ontology and human-labeled dataset for audio events. In: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).* Piscataway, NJ, USA: IEEE; 2017. p. 776–80.
27. Khosla P, Teterwak P, Wang C, Sarna A, Tian Y, Isola P, et al. Supervised contrastive learning. *arXiv:2004.11362.* 2021.
28. Zheng H, Chen X, Yao J, Yang H, Li C, Zhang Y, et al. Contrastive attraction and contrastive repulsion for representation learning. *arXiv:2105.03746.* 2023.
29. Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. Proximal policy optimization algorithms. *arXiv:1707.06347.* 2017.