IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Instrumentation Mesure Métrologie
Vol. 25, No. 4, August, 2026, pp. 189-198
Journal homepage: http://iieta.org/journals/i2m

Patient-Independent Heart Rate Trajectory Classification for Personalized Cardiac Rehabilitation Using a Hybrid CNN–TCN–GTN–SE Network

Steven Chin Su Leong[1], Nor Maniha Abdul Ghani[1*], Mohd Ridzuan Mohd Said[2], Nur Zahirah Mohd Ali[1], Syamimi Mardiah Shaharum[1], Nafrizuan Mat Yahya[3]

[1] Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan 26600, Malaysia
[2] Kulliyyah of Medicine, International Islamic University Malaysia, Kuantan 25200, Malaysia
[3] Faculty of Manufacturing and Mechatronic Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan 26600, Malaysia

Corresponding Author Email: normaniha@umpsa.edu.my

https://doi.org/10.18280/i2m.250404

Received: 1 June 2026
Revised: 22 July 2026
Accepted: 5 August 2026
Available online: 26 August 2026

Keywords:
cardiac rehabilitation, deep learning, heart rate trajectory, patient-independent validation, photoplethysmography, remote monitoring, Temporal Convolutional Network, wearable device

ABSTRACT

Cardiovascular rehabilitation requires continuous heart rate (HR) monitoring to evaluate physiological adaptation during exercise and recovery. This study proposes a hybrid framework combining Convolutional Neural Network (CNN), Temporal Convolutional Network (TCN), Gated Transformer Network (GTN), and Squeeze-and-Excitation (SE) blocks (CNN–TCN–GTN–SE) for classifying rehabilitation HR trajectories using wearable-derived HR measurements estimated by wrist-worn photoplethysmography (PPG) devices. HR trajectories were collected from 38 participants with ischemic heart disease (IHD) during a structured 480-second rehabilitation protocol. A total of 1,827 rehabilitation sessions were analysed, with each session represented by a 480-point HR trajectory sampled at 1 Hz. The proposed framework integrates convolutional feature extraction, temporal convolutional modelling, transformer-based attention, and SE channel recalibration. Model generalization across unseen participants was assessed using a leave-one-patient-out (LOPO) validation protocol. The proposed framework achieved participant-level accuracy of 0.9336 ± 0.0661 and AUC of 0.9837 ± 0.0286 across 38 held-out participants. Pooled evaluation achieved an accuracy of 0.9376 and an AUC of 0.9851. From a computational perspective, the network consisted of 218,185 trainable parameters, occupied approximately 852.29 KiB of memory, and demonstrated a mean prediction latency of 1.59 ms per sample on an NVIDIA Tesla T4 GPU. These findings support the feasibility of wearable-derived HR trajectory analysis for personalized cardiac rehabilitation monitoring.

1. INTRODUCTION

Cardiovascular rehabilitation plays an important role in improving recovery, exercise tolerance, and long-term cardiovascular outcomes among patients with ischemic heart disease (IHD). Accurate heart rate (HR) monitoring is central to rehabilitation programmes because HR dynamics provide important insights into cardiovascular adaptation during exercise and recovery [1]. Conventional rehabilitation approaches frequently rely on generalized exercise prescriptions based on age-predicted HR zones or metabolic equivalent (MET)-based thresholds. However, substantial inter-individual variability in physiological response may limit the effectiveness of generalized rehabilitation strategies [2, 3].

Recent advancements in wearable sensing technologies have enabled continuous and non-invasive cardiovascular monitoring during rehabilitation and daily activity. Wrist-worn photoplethysmography (PPG) devices integrated with smart wearable systems have demonstrated promising capabilities for HR estimation in ambulatory environments [4-6]. These technologies have increasingly been incorporated into rehabilitation monitoring systems because they improve accessibility and support remote and home-based rehabilitation programmes [7-14]. Recent studies have emphasized the importance of standardized validation and reporting for wrist-worn PPG heart-rate devices, particularly because measurement accuracy may vary with exercise intensity, skin tone, and motion artefacts [10, 15, 16].

Machine learning and deep learning techniques have further improved physiological time-series analysis and wearable cardiovascular monitoring. Convolutional neural networks (CNNs) have demonstrated effectiveness in extracting localized temporal features from physiological signals [17, 18], while recurrent and temporal architectures such as long short-term memory (LSTM), gated recurrent unit (GRU), and temporal convolutional networks (TCN) improve sequential pattern learning from time-series data [19-21]. More recently, transformer-based attention mechanisms and hybrid architectures have shown strong capability for modelling long-range temporal dependencies and dynamic feature interactions in biomedical signals [22-25].

Despite these advancements, several limitations remain in existing studies. First, many previous studies focus primarily on arrhythmia detection, stress monitoring, or generalized cardiovascular prediction rather than rehabilitation-specific HR trajectory classification [24, 25]. Second, many studies rely on multimodal physiological inputs or controlled laboratory datasets, whereas practical rehabilitation systems frequently operate using wearable-derived HR trajectories obtained from wrist-worn PPG devices. Third, rigorous patient-independent evaluation remains limited in rehabilitation monitoring studies, particularly when multiple rehabilitation sessions are collected from the same participant.

Therefore, this study proposes a hybrid CNN–TCN–GTN–SE (Squeeze-and-Excitation) architecture for rehabilitation HR trajectory classification using wearable-derived HR measurements obtained from wrist-worn PPG devices. The proposed framework integrates convolutional, temporal, channel-attention, and transformer-based attention mechanisms to model localized and long-range physiological patterns from rehabilitation HR trajectories.

The contributions of this study are summarized as follows:

- Development of a wearable rehabilitation-specific HR trajectory classification framework using wrist-worn smart devices.
- Proposal of a hybrid CNN–TCN–GTN–SE architecture for modelling localized, temporal, and channel-wise physiological features.
- Evaluation using strict patient-independent leave-one-patient-out (LOPO) validation.

The subsequent sections are arranged as follows. Section 2 reviews relevant literature on wearable cardiovascular monitoring and deep learning-based physiological signal analysis. Section 3 details the research methodology, encompassing dataset preparation, preprocessing techniques, model architecture, and performance evaluation. Section 4 reports the experimental findings. Section 5 interprets the results and addresses the limitations of the proposed approach. Section 6 concludes the paper and highlights opportunities for future investigation.

2. LITERATURE REVIEW

Wearable physiological monitoring has become increasingly important in cardiovascular rehabilitation because it enables continuous assessment of exercise response and cardiovascular adaptation outside conventional clinical environments [26, 27]. Wrist-worn PPG devices are particularly attractive because they provide non-invasive HR estimation while supporting ambulatory monitoring during rehabilitation activities [9, 10]. Previous studies have demonstrated acceptable agreement between wearable-derived HR measurements and reference electrocardiography (ECG) systems during structured exercise conditions [13, 14]. Recent validation studies have further emphasized the importance of wearable HR measurement accuracy during physical activity and rehabilitation monitoring, particularly in exercise-based cardiovascular populations [28-30].

Deep learning approaches have significantly improved physiological signal analysis and cardiovascular monitoring. CNN architectures have demonstrated promising performance in HR estimation, arrhythmia detection, and physiological feature extraction [17, 18, 24]. Recurrent architectures such as LSTM and GRU are commonly used for modelling sequential cardiovascular signals and temporal physiological dynamics [19, 20]. The TCN architecture further improves long-range temporal modelling using dilated convolutions and causal temporal filtering [21].

Recent studies have also explored hybrid and attention-based architectures for physiological time-series analysis [22, 23]. Transformer-based attention mechanisms improve temporal representation learning by modelling interactions across varying time scales, while SE blocks improve feature recalibration by emphasizing informative feature channels.

Despite these advances, several important gaps remain. First, most previous studies focus on arrhythmia detection, stress monitoring, or generalized cardiovascular prediction rather than rehabilitation-specific HR trajectory classification. Second, many studies rely on multimodal physiological signals or controlled laboratory datasets, whereas rehabilitation systems commonly operate using wearable-derived HR trajectories from consumer wearable devices. Third, few studies have evaluated rehabilitation HR trajectories using rigorous patient-independent validation strategies that account for repeated rehabilitation sessions collected from the same participant.

Therefore, this study investigates whether a hybrid CNN–TCN–GTN–SE framework can improve the classification of adaptive and maladaptive rehabilitation HR trajectories using wearable-derived HR measurements collected during structured rehabilitation.

3. METHODS

Authorization to conduct the study was granted under approval number IREC 2023-194 by the Institutional Research Ethics Committee of IIUM. Participation in the study was voluntary, and written informed consent was obtained from every participant before any study-related activities were performed. Ethical oversight was maintained throughout the project in accordance with recognized international guidelines for biomedical research involving human participants.

3.1 Data acquisition and preparation

The data acquisition platform comprised wrist-worn wearable sensors, a smartphone application, and Firebase cloud storage services. HR measurements derived from PPG signals were automatically synchronized with the mobile application and stored within the cloud environment for subsequent analysis. The HR values, rather than the raw PPG waveforms, were used as inputs to the proposed framework.

Because the proposed framework operates on wearable-derived HR measurements rather than raw PPG signals, the measurement characteristics of the wearable device were considered. The wearable internally estimates HR from the acquired PPG signals using proprietary signal processing algorithms before exporting HR measurements at 1 Hz (one HR measurement per second). Consequently, the recorded HR values may be affected by measurement uncertainty arising from motion artefacts, variations in skin contact pressure, ambient illumination, physiological variability, and the proprietary filtering and smoothing processes implemented by the wearable device. These factors may introduce short-term fluctuations, delayed HR responses, or occasional missing measurements. Therefore, evaluating the robustness of the proposed framework under realistic HR measurement uncertainty is important for practical wearable cardiac rehabilitation monitoring.

The investigated population consisted of 38 participants with IHD enrolled in a supervised rehabilitation program. The participants had a mean age of 43.6 years and included 29 males and 9 females.

Participants completed a structured 480-second rehabilitation protocol consisting of warm-up, moderate-intensity exercise, and higher-intensity exercise phases. HR measurements were sampled at 1 Hz using wearable smart devices integrated with the Orcus rehabilitation monitoring application [31, 32]. Therefore, each rehabilitation session corresponded to a complete 480-second wearable-derived HR trajectory containing 480 sequential HR measurements.

A total of 1,827 rehabilitation HR trajectories were collected and labelled as adaptive or maladaptive rehabilitation responses (Table 1).

Table 1. Distribution of rehabilitation heart rate (HR) trajectories

Label	Class	Trajectories	HR Patterns
0	Adaptive	1,186	Progressive increase in HR corresponding to exercise intensity
1	Maladaptive	641	Exaggerated, blunted, or irregular HR response

Each participant contributed multiple rehabilitation sessions, resulting in repeated HR trajectories per participant. Session counts varied across participants, ranging from 6 to 63 sessions per participant, with a mean of 48.08 ± 15.68 sessions (Table 2). Because session counts were not balanced across participants, participant-level LOPO summaries were used as the primary evaluation strategy to reduce bias from unequal session weighting.

Table 2. Session distribution across participants

Statistics	Sessions per Participant
Number of participants	38
Total sessions	1,827
Mean	48.08
Standard deviation (SD)	15.68
Minimum	6
Maximum	63

Rehabilitation trajectories were labelled according to predefined clinical criteria established by an experienced cardiologist based on HR response characteristics observed throughout the 480-second rehabilitation protocol. The assessment considered the overall temporal evolution of each HR trajectory rather than isolated HR measurements. The labeling followed a standardized assessment in which adaptive trajectories (Class 0) were defined by a smooth, physiologically expected HR elevation during the transition from warm-up to high-speed exercise, followed by stabilization. In contrast, maladaptive trajectories (Class 1) were identified by specific indicators including blunted (flat) responses despite increased exercise intensity, exaggerated spikes, or irregular fluctuations inconsistent with standard cardiac rehabilitation adaptation patterns for patients with IHD. The predefined clinical criteria were established before model development and remained unchanged throughout the study to ensure consistent annotation across all rehabilitation sessions.

During preprocessing, rehabilitation sessions containing incomplete records or missing HR measurements were excluded to ensure data quality and consistency. Each participant was assigned a unique participant identifier, which was used for participant-wise LOPO partitioning to ensure that all rehabilitation sessions from the same participant were allocated exclusively to either the training, validation, or held-out test set. Z-score normalization was applied within each LOPO fold, with normalization parameters computed using only training participants.

3.2 Proposed CNN–TCN–GTN–SE framework

The proposed framework integrates CNN, TCN, GTN, and SE modules for rehabilitation HR trajectory classification. The architecture was designed to capture complementary aspects of physiological temporal dynamics from wearable-derived HR trajectories, as shown in Figure 1. The sequence of procedures involved in the proposed classification methodology is depicted in Figure 2.

This hybrid configuration was adopted because standalone transformer-based architectures may be sensitive to localized fluctuations and motion-related noise inherent in 1 Hz wearable HR signals. Local temporal characteristics were extracted while reducing the influence of short-term fluctuations in the input sequence. The resulting feature representations were then processed through dilated temporal convolutional layers to capture dependencies spanning extended time horizons. Subsequently, multi-head self-attention combined with residual normalization was employed to model interactions among temporal features across different stages of the rehabilitation trajectory.

The attention-enhanced features were then adaptively fused with the original TCN features through a sigmoid gating mechanism. The gating mechanism dynamically regulated the contribution of local temporal information and global contextual representations before residual learning and layer normalization were applied. Finally, the recalibrated feature representations were forwarded to the SE module for channel-wise feature refinement prior to binary classification. Table 3 and Table 4 present the training configuration and computational characteristics of the model.

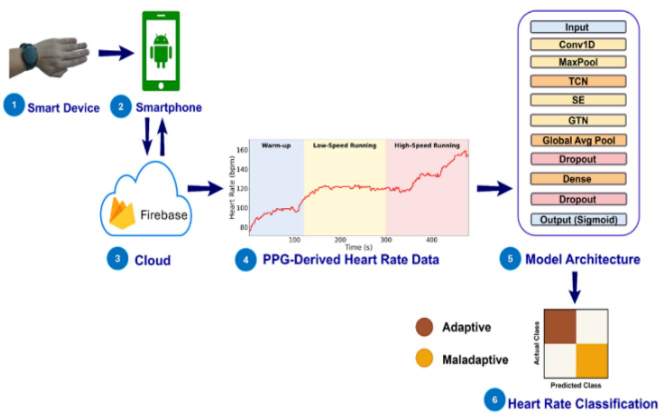

Figure 1. Processing pipeline and architectural design of the proposed CNN–TCN–GTN–SE network

Note: Convolutional Neural Network (CNN), Temporal Convolutional Network (TCN), Gated Transformer Network (GTN), Squeeze-and-Excitation (SE).

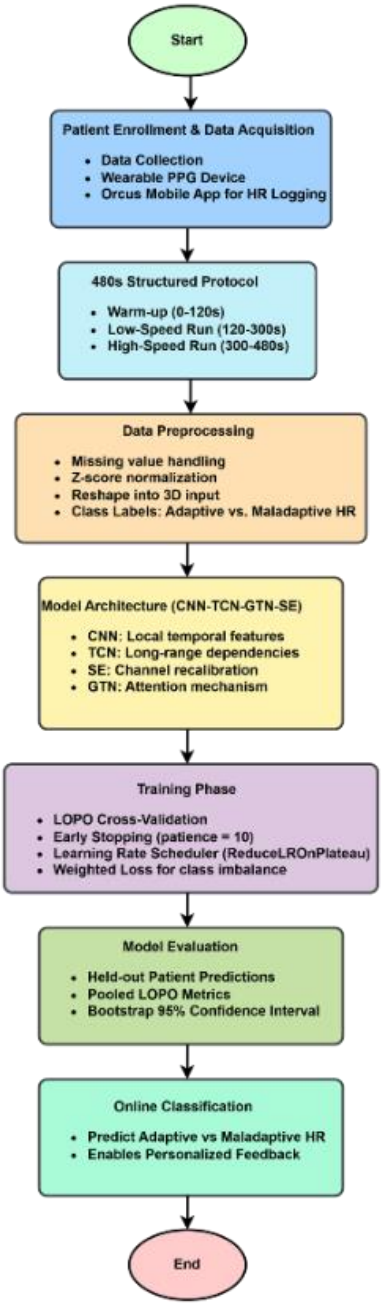

Figure 2. End-to-end pipeline for heart rate (HR) trajectory classification

The SE operation is defined in Eq. (1):

$$s = \sigma\left(W_2 \cdot ReLU\left(W_1 \cdot GAP(X)\right)\right), X' = X \odot s \qquad (1)$$

where, X is the incoming feature tensor, while GAP(·) corresponds to the global average pooling operation. The trainable parameter matrices are denoted by W_1 and W_2 and σ denotes the sigmoid activation function. The symbol $\odot$ indicates element-wise scaling between the feature map and the attention weights.

The self-attention process adopted in the proposed architecture is represented by Eq. (2):

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad (2)$$

The query, key, and value embeddings are denoted by Q, K, and V, respectively. The scaling factor d_k corresponds to the dimensionality of the key embeddings and is introduced to improve numerical stability during attention score computation.

Following the multi-head self-attention operation, a sigmoid gating mechanism was employed to generate adaptive gating coefficients for feature fusion. The gate dynamically regulates the contribution of the original TCN features and the attention-enhanced representations. This mechanism preserves informative local temporal characteristics while selectively incorporating long-range contextual information before residual learning and layer normalization.

Table 3. Network design specifications and training hyperparameters

Component	Parameter	Settings
Input	Timesteps	480
Input	Channels	1
Conv1D Layer	Filters	64
Conv1D Layer	Kernel size	3
Conv1D Layer	Activation	ReLU
SE Block	Reduction ratio	8
MaxPooling Layer	Pool size	2
TCN Module	Filters	64
GTN Module	Attention heads	4
GTN Module	Key dimension	64
GTN Module	Gating activation	Sigmoid
GTN Module	Feature fusion	Adaptive gated fusion
Dropout layer	Rate	0.3
Dense Layer	Hidden units	32
Output layer	Activation	Sigmoid
Optimization	Optimizer	Adam
Optimization	Initial Learning Rate	0.001
Training	Mini-batch Size	32
Training	Maximum Epochs	100
Training	Early-Stopping Criterion	Patience = 10
Execution Environment	Platform	Google Colab (T4 GPU)
Validation Strategy	Protocol	LOPO

Note: Temporal Convolutional Network (TCN), Gated Transformer Network (GTN), leave-one-patient-out (LOPO).

Table 4. Computational characteristics of the proposed framework

Metric	Value
Total parameters	218,185
Trainable parameters	218,185
Approximate model size	852.29 KiB
Average inference time	1.59 ms per sample
Hardware platform	NVIDIA Tesla T4 GPU

The gating operation adopted in the GTN module is defined in Eq. (3):

$$G = \sigma\left(W_g[F_{TCN}; F_{att}] + b_g\right) \qquad (3)$$

where, F_{TCN} denotes the temporal feature representation extracted by the TCN, while F_{att} represents the attention-enhanced feature representation generated by the multi-head self-attention module. The notation $[;]$ corresponds to feature concatenation. The trainable weight matrix and bias vector are denoted by W_g and b_g, respectively, whereas $\sigma(\cdot)$ corresponds to the sigmoid activation function. The output G represents the adaptive gating coefficient used to regulate feature fusion.

The gated feature fusion process is defined in Eq. (4):

$$F_{GTN} = G \odot F_{att} + (1 - G) \odot F_{TCN} \quad (4)$$

where, F_{GTN} denotes the fused feature representation produced by the GTN module. The symbol $\odot$ indicates element-wise multiplication. The adaptive gating coefficient G dynamically regulates the relative contribution of the attention-enhanced features and the original temporal features extracted by the TCN before residual learning and layer normalization.

3.3 Training and evaluation strategy

To examine performance under patient-independent conditions, a LOPO evaluation scheme was employed. In each fold, the dataset of a single participant was isolated as the test set, and another participant was assigned to the validation set. The model was subsequently trained using data from the remaining participants.

The dataset exhibited an unequal distribution between adaptive and maladaptive rehabilitation trajectories. To mitigate potential bias during model optimization, balanced class weights were computed independently from the training data within each LOPO fold and incorporated into the binary cross-entropy loss function. No oversampling, undersampling, or synthetic data generation technique was applied. A fixed classification threshold of 0.5 was adopted for all experiments to convert predicted probabilities into binary class labels. In addition, the precision–recall area under the curve (PR-AUC) was computed from the pooled held-out predictions as a complementary measure of classification performance under class imbalance.

Binary cross-entropy loss was minimized during training as defined in Eq. (5):

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log p_i + (1 - y_i \log(1 - p_i))] \quad (5)$$

Performance was evaluated using two complementary strategies. The primary evaluation employed participant-level macro averaging, where classification metrics were first computed independently for each held-out participant and subsequently averaged as mean ± SD across all 38 LOPO folds. This approach assigns equal importance to every participant regardless of the number of rehabilitation sessions contributed, thereby reducing bias arising from unequal session counts in the estimation of patient-independent performance.

In addition, pooled (micro) evaluation was performed by aggregating all held-out rehabilitation sessions from every LOPO fold into a single evaluation set. Participants contributing a larger number of rehabilitation sessions have a proportionally greater influence on the pooled metrics. Consequently, the pooled results primarily reflect session-level classification performance and are reported as secondary descriptive estimates. Therefore, the participant-level macro-averaged LOPO results are regarded as the principal findings of this study, whereas the pooled metrics provide complementary information regarding overall classification performance across all rehabilitation sessions.

To evaluate the robustness of the proposed framework against wearable HR measurement uncertainty, controlled perturbations were introduced into the held-out rehabilitation HR trajectories. Gaussian measurement noise with SDs of 1, 3, and 5 bpm was injected to simulate increasing levels of wearable HR measurement uncertainty. In addition, robustness against missing measurements was evaluated by randomly removing 20% of the HR measurements and by introducing a consecutive 20-second missing segment. Participant-level macro performance was computed independently for each held-out participant and summarized as the mean ± SD across all 38 LOPO folds.

4. RESULTS

4.1 Participant-level leave-one-patient-out performance

The participant-level LOPO results presented in this section represent the primary evaluation of the proposed framework using macro averaging. Each held-out participant contributes equally to the reported performance regardless of the number of rehabilitation sessions available, thereby minimizing bias arising from unequal session counts across participants.

The proposed CNN–TCN–GTN–SE framework demonstrated strong patient-independent classification performance across rehabilitation HR trajectories. Participant-level LOPO evaluation yielded a mean accuracy of 0.9336 ± 0.0661 and a mean AUC of 0.9837 ± 0.0286 across 38 held-out participants. Table 5 summarizes the participant-level LOPO classification performance across 38 held-out participants.

Table 5. Leave-one-patient-out (LOPO) performance summary across 38 held-out participants (macro)

Metric	Mean ± SD
Accuracy	0.9336 ± 0.0661
AUC	0.9837 ± 0.0286
Precision	0.8308 ± 0.2035
Sensitivity	0.9074 ± 0.1641
Specificity	0.9008 ± 0.1894
F1-score	0.8468 ± 0.1592
MCC	0.7844 ± 0.1851

Performance variability was observed across the 38 held-out participants, as reflected by the SDs reported in Table 5. The proposed framework achieved a high mean AUC of 0.9837 ± 0.0286 and an MCC of 0.7844 ± 0.1851. However, larger variability was observed for precision, sensitivity, specificity, and MCC, indicating that classification performance differed among individual participants. This variability is likely attributable to participant-specific rehabilitation HR trajectory characteristics, unequal numbers of rehabilitation sessions, and overlap between adaptive and maladaptive physiological response patterns.

4.2 Pooled leave-one-patient-out performance

To complement the participant-level analysis, pooled (micro) evaluation was performed by aggregating all held-out rehabilitation sessions across the LOPO folds. Because pooled evaluation weights participants according to their number of sessions, these results are reported as secondary session-level estimates. The participant-level LOPO results remain the primary findings of this study.

The pooled evaluation results indicated that the proposed framework effectively captured discriminative temporal patterns associated with adaptive and maladaptive rehabilitation responses using wearable-derived HR trajectories. Table 6 presents the pooled LOPO evaluation results across all rehabilitation HR trajectories.

Table 6. Pooled held-out leave-one-patient-out (LOPO) performance across all rehabilitation trajectories

Metric	Pooled Value
Accuracy	0.9376
AUC	0.9851
Precision	0.8962
Sensitivity	0.9298
Specificity	0.9418
F1-score	0.9127
MCC	0.8645

4.3 Ablation study of architectural components

A systematic ablation study was conducted using the same LOPO validation protocol adopted in the main experiments to ensure a fair comparison among all network configurations. Performance metrics were first computed independently for each held-out participant and subsequently summarized as mean ± SD across all 38 LOPO folds. Table 7 presents the participant-level macro-averaged performance of each architecture.

Table 7. Participant-level macro-averaged leave-one-patient-out (LOPO) ablation results

Model	Accuracy	AUC	MCC
CNN	0.8344 ± 0.1174	0.9448 ± 0.0718	0.5634 ± 0.2214
CNN–TCN	0.8881 ± 0.1096	0.9746 ± 0.0498	0.6686 ± 0.2332
CNN–TCN–SE	0.9005 ± 0.0959	0.9757 ± 0.0490	0.6926 ± 0.2312
CNN–TCN–GTN	0.9144 ± 0.1001	0.9799 ± 0.0437	0.7501 ± 0.2180
CNN–TCN–GTN–SE	**0.9336 ± 0.0661**	**0.9837 ± 0.0286**	**0.7844 ± 0.1851**

Note: Convolutional Neural Network (CNN), Temporal Convolutional Network (TCN), Gated Transformer Network (GTN), Squeeze-and-Excitation (SE).

The ablation results demonstrated that each architectural component contributed to progressive performance improvement under the same LOPO validation protocol. Compared with the CNN baseline, the inclusion of the TCN substantially improved temporal feature learning. Incorporating the GTN further enhanced long-range temporal dependency modeling through adaptive feature fusion, while the SE module provided additional channel-wise feature recalibration. Thus, the complete CNN–TCN–GTN–SE framework achieved the highest participant-level accuracy, AUC, and MCC among all evaluated architectures.

The pooled confusion matrix in Figure 3 shows the classification distribution across all held-out LOPO predictions, indicating that most adaptive and maladaptive rehabilitation HR trajectories were correctly classified. The ROC curve in Figure 4 further demonstrates strong discrimination between the two response classes, with a pooled AUC of 0.9851. In addition, the precision–recall curve in Figure 5 shows stable classification performance under class imbalance, with a PR-AUC of approximately 0.97. Since the PR-AUC was computed using pooled held-out predictions, it is reported as a complementary session-level metric alongside the primary participant-level LOPO evaluation.

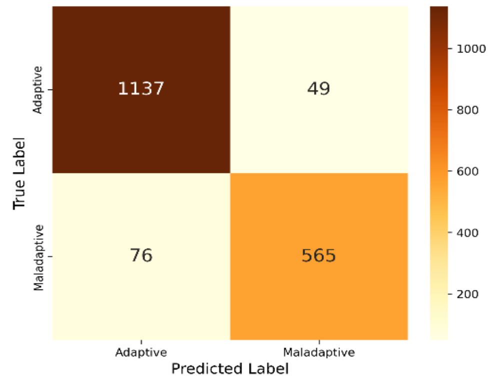

Figure 3. Confusion matrix of pooled leave-one-patient-out (LOPO) predictions

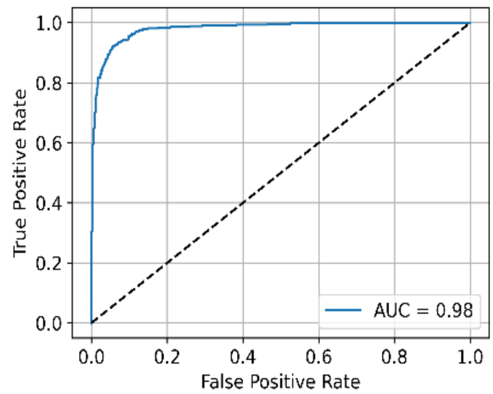

Figure 4. Receiver operating characteristic (ROC) curve of the proposed CNN–TCN–GTN–SE framework

Note: Convolutional Neural Network (CNN), Temporal Convolutional Network (TCN), Gated Transformer Network (GTN), Squeeze-and-Excitation (SE).

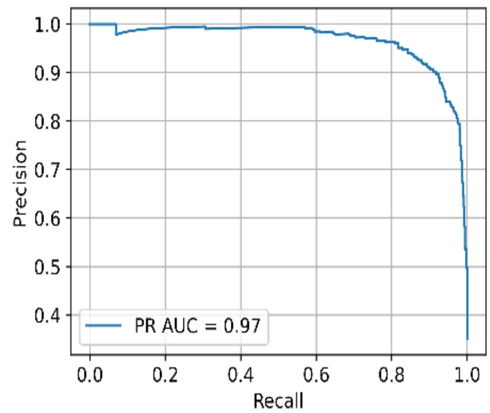

Figure 5. Precision–recall curve of the proposed model

4.4 Robustness under heart rate measurement uncertainty

Table 8 presents the participant-level macro classification performance under representative wearable HR measurement uncertainties, including additive Gaussian measurement noise and simulated missing measurements. The proposed CNN–TCN–GTN–SE framework demonstrated a gradual reduction in classification accuracy as Gaussian HR measurement noise increased. Accuracy decreased from 0.9336 ± 0.0661 under clean conditions to 0.8807 ± 0.1044 under 5 bpm noise. In contrast, only negligible performance variation was observed under simulated missing HR measurements. Accuracy was 0.9238 ± 0.0865 with 20% random missing measurements and 0.9269 ± 0.0844 for a consecutive 20-second missing segment. Compared with the CNN–TCN baseline, the proposed CNN–TCN–GTN–SE framework consistently achieved higher participant-level macro classification accuracy under all evaluated measurement uncertainty conditions, demonstrating improved robustness to realistic wearable HR measurement errors.

Table 8. Robustness evaluation under representative heart rate (HR) measurement perturbations

Perturbation	CNN–TCN–GTN–SE Accuracy	CNN–TCN Accuracy
Clean	0.9336 ± 0.0661	0.8881 ± 0.1096
Gaussian noise, 1 bpm	0.9147 ± 0.0946	0.8846 ± 0.1073
Gaussian noise, 3 bpm	0.8984 ± 0.0919	0.8784 ± 0.1160
Gaussian noise, 5 bpm	0.8807 ± 0.1044	0.8720 ± 0.1372
Random missing measurements, 20%	0.9238 ± 0.0865	0.8881 ± 0.1096
Consecutive missing segment, 20 s	0.9269 ± 0.0844	0.8896 ± 0.1086

Note: Convolutional Neural Network (CNN), Temporal Convolutional Network (TCN), Gated Transformer Network (GTN), Squeeze-and-Excitation (SE).

4.5 Training stability across leave-one-patient-out folds

Table 9 summarizes the training stability across the 38 LOPO folds. The proposed CNN–TCN–GTN–SE framework achieved its best validation performance at a median epoch of 36 (interquartile range: 29–42), whereas the CNN–TCN baseline reached its best validation performance at a median epoch of 41 (interquartile range: 33–50). Both architectures achieved their best validation performance before the maximum training limit under the adopted early-stopping strategy. This pattern indicates stable optimization across the LOPO folds despite differences in rehabilitation HR trajectory characteristics among held-out participants.

Table 9. Training stability across the 38 leave-one-patient-out (LOPO) folds

Model	Median Best Validation Epoch	Interquartile Range (Q1–Q3)
CNN–TCN–GTN–SE	36	29–42
CNN–TCN	41	33–50

Note: Convolutional Neural Network (CNN), Temporal Convolutional Network (TCN), Gated Transformer Network (GTN), Squeeze-and-Excitation (SE).

5. DISCUSSION

The proposed framework integrates convolutional feature extraction, temporal sequence modelling, channel-wise recalibration, and transformer-based attention mechanisms to capture localized and long-range physiological patterns from rehabilitation HR trajectories.

Compared with standalone recurrent and convolutional architectures, the hybrid framework demonstrated improved classification capability under patient-independent LOPO evaluation. The integration of TCN and GTN modules improved temporal representation learning, while SE blocks enhanced feature recalibration across informative channels. These complementary mechanisms contributed to stronger discrimination between adaptive and maladaptive rehabilitation trajectories.

The robustness evaluation further demonstrated that the proposed framework maintained stable classification performance under representative wearable HR measurement uncertainties, including Gaussian measurement noise and simulated missing HR measurements. Although classification accuracy gradually decreased with increasing measurement noise, only marginal performance variation was observed under missing HR measurements. These findings suggest that the learned temporal representations are relatively resilient to moderate wearable HR measurement uncertainty and intermittent data loss commonly encountered in wrist-worn PPG-based cardiac rehabilitation monitoring.

The proposed framework also demonstrated promising performance using wearable-derived HR trajectories without requiring multimodal physiological signals or invasive monitoring systems. This approach supports the feasibility of remote rehabilitation monitoring using consumer wearable devices and facilitates the future development of intelligent home-based rehabilitation systems. However, the present study evaluated wearable-derived HR trajectories rather than raw PPG waveforms directly.

Unlike many prior studies on arrhythmia detection or generalized cardiovascular monitoring, this work specifically investigated rehabilitation-oriented HR trajectory classification using patient-independent validation. Direct comparison remains challenging, as most prior work emphasizes stress monitoring and multimodal physiological analysis rather than this task under strict patient-independent evaluation.

The proposed CNN–TCN–GTN–SE framework contained 218,185 trainable parameters with an approximate model size of 852.29 KiB. Inference benchmarking demonstrated an average prediction latency of 1.59 ms per sample on an NVIDIA Tesla T4 GPU platform. These findings suggest potential suitability for lightweight rehabilitation monitoring applications using wearable physiological data. However, direct embedded-device deployment and energy-consumption evaluation were not investigated in the present study.

Interpretation of the current findings should consider the characteristics of the study cohort. The sample size was relatively limited, and the participant distribution was not demographically balanced. These factors may affect external validity and population-level generalizability. In addition, session counts varied across participants, resulting in unequal numbers of rehabilitation trajectories per participant. To reduce bias from unequal session weighting, participant-level LOPO summaries were used as the primary evaluation strategy. External validation using independent multicentre rehabilitation datasets was not performed in the present study.

One limitation of the present study is that the annotation protocol was based on predefined clinical criteria established by a single experienced cardiologist. Consequently, inter-rater agreement could not be evaluated, and some degree of subjective variability may exist. Future work will involve multiple independent cardiologists to assess annotation reproducibility and inter-rater reliability.

Another limitation of the present study is that synchronized reference ECG measurements were not available during wearable HR acquisition. Consequently, the proposed framework evaluates rehabilitation trajectory classification using wearable-derived HR measurements rather than the absolute accuracy of HR estimation against a clinical reference standard. Although strict LOPO validation and robustness evaluation under representative HR measurement uncertainties reduce the likelihood that the model relies on participant-specific characteristics or isolated measurement artefacts, the possibility that wearable device characteristics or annotation bias influenced the learned representations cannot be completely excluded. Therefore, future studies will incorporate synchronized ECG measurements together with multi-expert clinical annotation to further validate the physiological relevance of the learned HR trajectory patterns.

Future work will also focus on multicentre validation, larger and more balanced rehabilitation cohorts, subgroup analysis, and lightweight deployment for wearable rehabilitation monitoring systems.

6. CONCLUSIONS

The proposed framework demonstrated strong patient-independent classification performance under strict LOPO evaluation, achieving high discrimination capability between adaptive and maladaptive rehabilitation responses.

The integration of CNN, TCN, GTN, and SE modules enabled effective modelling of localized temporal features, long-range dependencies, channel-wise feature recalibration, and temporal attention-based interactions from rehabilitation HR trajectories. The framework also demonstrated the feasibility of wearable-derived HR trajectory analysis for remote and personalized cardiovascular rehabilitation monitoring.

The proposed methodology contributes to wearable physiological monitoring and rehabilitation-oriented time-series analysis by incorporating patient-independent evaluation, participant-level statistical analysis, computational performance analysis, and hybrid temporal deep learning mechanisms. Future work will focus on multicentre validation, subgroup evaluation, multimodal physiological integration, and lightweight deployment for intelligent rehabilitation monitoring systems.

ACKNOWLEDGMENT

This research received financial support from the Distinguished Research Grant Scheme administered by Universiti Malaysia Pahang Al-Sultan Abdullah under grant number RDU233012.

REFERENCES

[1] Warner, A., Vanicek, N., Benson, A., Myers, T., Abt, G. (2022). Agreement and relationship between measures of absolute and relative intensity during walking: A systematic review with meta-regression. PLoS ONE, 17(11): e0277031. https://doi.org/10.1371/JOURNAL.PONE.0277031

[2] Meixner, B., Filipas, L., Holmberg, H.C., Sperlich, B. (2025). Zone 2 intensity: A critical comparison of individual variability in different submaximal exercise intensity boundaries. Translational Sports Medicine, 2025(1): 2008291. https://doi.org/10.1155/TSM2/2008291

[3] Van Iterson, E.H., Laffin, L.J., Svensson, L.G., Cho, L. (2022). Individualized exercise prescription and cardiac rehabilitation following a spontaneous coronary artery dissection or aortic dissection. European Heart Journal Open, 2(6): 1-12. https://doi.org/10.1093/EHJOPEN/OEAC075

[4] Wang, R.J., Veera, S.C.M., Asan, O., Liao, T. (2024). A systematic review on the use of consumer-based ECG wearables on cardiac health monitoring. IEEE Journal of Biomedical and Health Informatics, 28(11): 6525-6537. https://doi.org/10.1109/JBHI.2024.3456028

[5] Ding, C., Xiao, R., Wang, W.J., Holdsworth, E., Hu, X. (2024). Photoplethysmography based atrial fibrillation detection: A continually growing field. Physiological Measurement, 45(4): 04TR01. https://doi.org/10.1088/1361-6579/AD37EE

[6] Fioravanti, V.B.O., Freitas, P.G., Rodrigues, P.G., et al. (2024). Machine learning framework for Inter-Beat Interval estimation using wearable Photoplethysmography sensors. Biomedical Signal Processing and Control, 88: 105689. https://doi.org/10.1016/J.BSPC.2023.105689

[7] Stock, R., Gaarden, A.P., Langørgen, E. (2024). The potential of wearable technology to support stroke survivors' motivation for home exercise – Focus group discussions with stroke survivors and physiotherapists. Physiotherapy Theory and Practice, 40(8): 1779-1790. https://doi.org/10.1080/09593985.2023.2217987

[8] Sher, J., Lewis, C.W., Lin, C. (2024). Using wearable technologies to monitor physical activity and exercise in patients: A narrative review. Current Sports Medicine Reports, 23(8): 284-289. https://doi.org/10.1249/JSR.0000000000001187

[9] Wang, T.L., Wu, H.Y., Wang, W.Y., et al. (2023). Assessment of heart rate monitoring during exercise with smart wristbands and a heart rhythm patch: Validation and comparison study. JMIR Formative Research, 7: e52519. https://doi.org/10.2196/52519

[10] Kim, C., Song, J.H., Kim, S.H. (2023). Validation of wearable digital devices for heart rate measurement during exercise test in patients with coronary artery disease. Annals of Rehabilitation Medicine, 47(4): 261-271. https://doi.org/10.5535/ARM.23019

[11] Shea, M.G., Headley, S., Mullin, E.M., Brawner, C.A., Schilling, P., Pack, Q.R. (2022). Comparison of ratings of perceived exertion and target heart rate–based exercise prescription in cardiac rehabilitation. Journal of Cardiopulmonary Rehabilitation and Prevention, 42(5): 352-358. https://doi.org/10.1097/HCR.0000000000000682

[12] Aharon, K.B., Gershfeld-Litvin, A., Amir, O., Nabutovsky, I., Klempfner, R. (2022). Improving cardiac rehabilitation patient adherence via personalized interventions. PLoS ONE, 17(8): e0273815. https://doi.org/10.1371/JOURNAL.PONE.0273815

[13] Ho, W.T., Yang, Y.J., Li, T.C. (2022). Accuracy of wrist-

worn wearable devices for determining exercise intensity. Digital Health, 8: 20552076221124393. https://doi.org/10.1177/20552076221124393
[14] Martín-Escudero, P., Cabanas, A.M., Dotor-Castilla, M.L., et al. (2023). Are activity wrist-worn devices accurate for determining heart rate during intense exercise? Bioengineering, 10(2): 254. https://doi.org/10.3390/BIOENGINEERING10020254
[15] Hung, S.H., Serwa, K., Rosenthal, G., Eng, J.J. (2025). Validity of heart rate measurements in wrist-based monitors across skin tones during exercise. PLoS ONE, 20(2): e0318724. https://doi.org/10.1371/JOURNAL.PONE.0318724
[16] Mühlen, J.M., Stang, J., Skovgaard, E.L., et al. (2021). Recommendations for determining the validity of consumer wearable heart rate devices: Expert statement and checklist of the INTERLIVE Network. British Journal of Sports Medicine, 55(14): 767-779. https://doi.org/10.1136/BJSPORTS-2020-103148
[17] Moreno, J.P., Sepúlveda, M.A., Pino, E.J. (2024). 1D convolutional neural network impact on heart rate metrics for ECG and BCG signals. Journal of Medical and Biological Engineering, 44: 437-447. https://doi.org/10.1007/S40846-024-00872-W
[18] Gullapalli, B.T., Nathan, V., Rahman, M.M., Kuang, J., Gao, J.A. (2024). A framework for extracting heart rate variability features from earbud-PPG for stress detection. In 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, pp. 1-5. https://doi.org/10.1109/EMBC53108.2024.10782088
[19] Pham, T.D. (2021). Time–frequency time–space LSTM for robust classification of physiological signals. Scientific Reports, 11: 6936. https://doi.org/10.1038/s41598-021-86432-7
[20] Zhong, J., Liu, Y.F., Cheng, X.K., Cai, L.M., Cui, W.D., Hai, D. (2022). Gated recurrent unit network for psychological stress classification using electrocardiograms from wearable devices. Sensors, 22(22): 8664. https://doi.org/10.3390/S22228664
[21] Dudukcu, H.V., Taskiran, M., Taskiran, Z.G.C., Yildirim, T. (2023). Temporal Convolutional Networks with RNN approach for chaotic time series prediction. Applied Soft Computing, 133: 109945. https://doi.org/10.1016/J.ASOC.2022.109945
[22] Kasnesis, P., Toumanidis, L., Burrello, A., Chatzigeorgiou, C., Patrikakis, C.Z. (2023). Feature-level cross-attentional PPG and motion signal fusion for heart rate estimation. In 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC), Torino, Italy, pp. 1731-1736. https://doi.org/10.1109/COMPSAC57700.2023.00267
[23] Li, J. (2025). Heart rate estimation using deep learning and hybrid artifact removal for optical volume tracing signals in motion states. International Journal of High Speed Electronics and Systems, 2540321. https://doi.org/10.1142/S0129156425403213
[24] Bulut, M.G., Unal, S., Hammad, M., Pławiak, P. (2025). Deep CNN-based detection of cardiac rhythm disorders using PPG signals from wearable devices. PLoS ONE, 20(2): e0314154. https://doi.org/10.1371/JOURNAL.PONE.0314154
[25] Dickson, E.L., Ding, E.Y., Saczynski, J.S., et al. (2021). Smartwatch monitoring for atrial fibrillation after stroke—The Pulsewatch Study: Protocol for a multiphase randomized controlled trial. Cardiovascular Digital Health Journal, 2(4): 231-241. https://doi.org/10.1016/J.CVDHJ.2021.07.002
[26] Hughes, A., Shandhi, M.M.H., Master, H., Dunn, J., Brittain, E. (2023). Wearable devices in cardiovascular medicine. Circulation Research, 132(5): 652-670. https://doi.org/10.1161/CIRCRESAHA.122.322389
[27] De Fazio, R., Mastronardi, V.M., De Vittorio, M., Visconti, P. (2023). Wearable sensors and smart devices to monitor rehabilitation parameters and sports performance: An overview. Sensors, 23(4): 1856. https://doi.org/10.3390/S23041856
[28] Loro, F.L., Martins, R., Ferreira, J.B., et al. (2024). Validation of a wearable sensor prototype for measuring heart rate to prescribe physical activity: Cross-sectional exploratory study. JMIR Biomedical Engineering, 9: e57373. https://doi.org/10.2196/57373
[29] Hu, T.H., Zhang, X.B., Millham, R.C., Xu, L., Wu, W.Q. (2025). Implementation of wearable technology for remote heart rate variability biofeedback in cardiac rehabilitation. Sensors, 25(3): 690. https://doi.org/10.3390/S25030690
[30] Kitagaki, K., Hongo, Y., Futai, R., Hasegawa, T., Morikawa, H., Shimoyama, H. (2025). Validity of heart rate measurement using wearable devices during cardiopulmonary exercise testing in patients with cardiovascular disease: Prospective pilot validation study. JMIR Cardio, 9: e77911. https://doi.org/10.2196/77911
[31] Ghani, N.M.A., Lim, W.J., Hoh, W.S., Hassan, S.A.A. (2024). Rehabilitation and home health monitoring based-AI scheduling application for coronary artery disease and cardiovascular patients. Instrumentation Mesure Métrologie, 23(2): 161-173. https://doi.org/10.18280/i2m.230207
[32] Said, M.R.M., Abd Razak, N.A., Siang, H.W., Shaharum, S.M., Abd Ghani, N. (2025). APCU 36 Preliminary finding of ORCUS: A home-based cardiac rehabilitation application for stable ischaemic heart disease. Open Heart, 12(Suppl 1): A16. https://doi.org/10.1136/OPENHRT-2024-APCU.36

NOMENCLATURE

AUC	area under the curve
CNN	convolutional neural network
ECG	electrocardiography
GRU	gated recurrent unit
GTN	gated transformer network
HR	heart rate
IHD	ischemic heart disease
IoT	internet of things
LOPO	leave-one-patient-out
LSTM	long short-term memory
MCC	Matthews correlation coefficient
PPG	photoplethysmography
PR-AUC	precision–recall area under the curve
ReLU	rectified linear unit
SE	squeeze-and-excitation
TCN	temporal convolutional network

Symbols

σ	sigmoid activation function
d_k	attention key dimension
K	key matrix
L	binary cross-entropy loss
Q	query matrix
s	channel attention scaling vector
V	value matrix
X	input feature map
y	ground-truth label
$\hat{y}$	predicted probability output