Proceedings of CONF-SPML 2026 Symposium: The 2nd Neural Computing and Applications Workshop 2025
DOI: 10.54254/2755-2721/2026.TJ30929

Research on Motion Heart Rate Detection Method Based on Photoplethysmography and Human Acceleration

Jiarui He

Aulin College, Northeast Forestry University, Harbin, China
h3294248221@outlook.com

Abstract. Heart rate is an important indicator reflecting the health status of the human body and has significant value in sports health monitoring. In order to solve the problem of motion artifacts interfering with the optical pulse wave recording (PPG) signal during motion, this paper describes PPG and proposes a deep convolutional attention network (DCan) method based on signal and human acceleration (ACC) signal fusion. This model uses multi-scale convolution and attention mechanism to extract features from PPG signal ACC and combines the complementary information of these two signal types to improve the accuracy of heart rate prediction. According to the experimental results, the average absolute error (MAE) of DCAN using the model to predict heart rate was reduced by 23% and 32% compared to the C-RNN model and NAS-PPG model, respectively, and showed higher stability and accuracy in various sports scenarios. This study provides reliable technical support for heart rate monitoring during exercise.

Keywords: Photoelectric pulse wave, human acceleration, heart rate estimation, convolution, attention mechanism.

1. Introduction

Heart rate (HR) is an important indicator of human health. Heart rate research is an important method for analyzing the cardiovascular health status in daily life [1]. High heart rate may lead to various diseases such as heart failure, myocardial infarction, and sudden cardiac death [2]. On the contrary, if the heart rate is too low, the blood supply will be insufficient, and symptoms such as dizziness, fatigue, blurred vision, and fainting may occur. Accurate heart rate monitoring is crucial for early detection of many diseases. Sports heart rate refers to the heart rate during physical activity, and monitoring sports heart rate is crucial for evaluating athletes' health status and controlling training load. On the other hand, this reflects the athlete's health condition and guides you to adjust the training load so that the athlete can better adapt to the training objectives [3]. With the popularity of wearable devices, users' demand for device functionality continues to increase. The approach that is suggested in this paper gives wearable devices better heart rate monitoring functionality, increases their usefulness and competitiveness in the market, and facilitates the proliferation of wearable devices in the health monitoring sector.

Electrocardiography (ECG) is a very accurate technique of the measuring heart rate, but it is inconvenient to monitor ECG in the real world. As such, wrist-worn heart rate monitors have

DOI: 10.54254/2755-2721/2026.TJ30929

become popular recently. The sensors used in these devices are mostly photoplethysmography (PPG) signals because they are non-invasive and portable. At present, there exist different commercial products (wearable devices) with PPG sensors including Apple Watch, Fitbit Charge, Samsung Simband etc. Nevertheless, the estimation of heart rate using PPG signals is more difficult than using traditional ECG data. Motion artifacts (MA) in particular play a major role in reducing the quality of PPG signal especially during exercise [4]. Thus, the enhancement of the precision and consistency of the PPG-based heart rate estimation is highly important. Accelerometers (ACC) are capable of measuring triaxial acceleration (x, y, z) of human body which indicates the current motion state. Integration of PPG and ACC signals will be useful in enhancing the accuracy of estimating heart rate. Researchers have improved the estimation of heart rate during physical activity by combining PPG with ACC. Among these, the VADAF method, which integrates data from multiple sensors, has significantly enhanced accuracy, although it entails higher computational complexity [5-8].

The purpose of this study is to investigate algorithms for detecting the speed of commercial centers from three perspectives:

The signal fusion signal of low-frequency PPG is easily controlled by operational artifacts, resulting in inaccurate pulse recognition. This study utilizes the operational information provided by the acceleration signal PPG to effectively mix the signal with the acceleration signal PPG, suppress signal noise, and improve detection accuracy

Algorithm optimization and real-time performance PPG: We are developing an efficient deep learning algorithm framework to achieve real-time heartbeat recognition by quickly processing acceleration and acceleration signals. This algorithm takes into account the low power consumption of the wear device and the limited computing resources required to meet low computational demands

Adaptability and Durability: Study the adaptability and persistence of algorithms under different types and intensities of motion to improve the stability and adaptability of individual differences in different business scenarios.

2. The main research content

The system diagram of this study is provided, which also serves as the overall framework of this article. Chapter 2 introduces the DCAN model for data processing, and Chapter 3 explains the evaluation index.

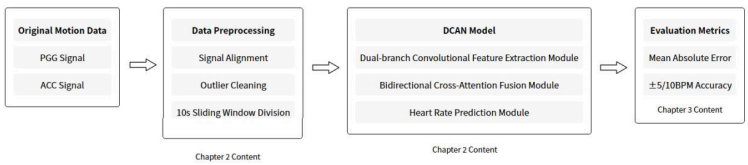

Figure 1. The overall framework of this article

From the functional scheme of the system, it can be seen that the signals used in this article are PPG and ACC signals (Figure 1). The data preprocessing step consists of three main steps: Adjusting the signal; Clearance of aberrant values; 10 seconds divided into sliding windows. The proposed model (DCAN) comprises three modules. The model is evaluated on the basis of three key measures which include the mean absolute error (MAE), accuracy of 5 BPM, and accuracy of 10 BPM.

DOI: 10.54254/2755-2721/2026.TJ30929

This paper seeks to perform a systematic and innovative research in order to address these significant challenges. In particular, the training and model tests are conducted independently with respect to different intensity levels of exercise (low, medium, high) and individuals. The goal is to develop highly customized models that effectively improve the accuracy of heart rate estimation. At the same time, in the design of model architectures, we propose and develop attention mechanisms, which have unique advantages in terms of feature extraction and merging [9,10]. This study introduces a mechanism of attention. This mechanism allows the model to automatically focus on the important information field when processing PPG and ACC signals, dynamically adjust its attention to the different characteristics of the signals, deeply analyze the characteristics of the two signals, and achieve deeper and more efficient fusion. This results in a more accurate and universally applicable estimate of heart rate. This provides innovative ideas and methods to promote the widespread application of heart rate estimation techniques based on the fusion of PPG and ACC in areas such as medical health surveillance and sports sciences.

3. Research on heart rate estimation method based on fusion of PPG and ACC signals

3.1. Dataset introduction and data preprocessing

3.1.1. Dataset

All data used in this experiment is provided by Chipsea Technologies. It includes 12 exercise scenarios (sprint course, outdoor walking, outdoor course, outdoor cycling, rest, dart jump, stair climb, indoor walking, indoor course, indoor cycling, jump code, free activities). Each scenario includes data collected from two individuals, in particular PPG and ACC signals sampled at 32 Hz and a reference heart rate signal at 1 Hz. Exercise scenarios cover low, moderate and high intensity activities (as seen in Table 1).

Table 1. Test scenarios, methods, and number of participants

Testing Scenario	Testing Method	Number of Testers
Rest	Sit quietly for 3min	2 people
Indoor Walking	Rest for 2min + Brisk walking (3km/h) for 10min	2 people
Indoor Running	Rest for 2min + Running for 15min (heart rate shows two peaks exceeding 160)	2 people
Indoor Cycling	Rest for 2min + Cycling for 15min	2 people
Free Activity	Free activity: playing table tennis for 3min + moving hands left and right for 2min + others (roadblock obstacle course for 3min + juggling, etc.) for 5 minutes	2 people
Outdoor Walking	Rest for 2min + Walking outdoors for 10min	2 people
Outdoor Running	Rest for 2min + Running outdoors for 15min	2 people
Outdoor Cycling	Rest for 2min + Cycling outdoors for 10min	2 people
Stair Climbing Up and Down	Stair climbing up and down for 3min	2 people
Jumping Jacks	Rest for 1min + Jumping jacks for 2min + Rest for 2min + Jumping jacks for 1min + Rest for 1min	2 people
Sprint Running	Rest for 1min + Treadmill at 6km/h + maximum speed + Treadmill at 8km/h + maximum speed + Treadmill at 6km/h	2 people
Skipping Rope	Rest for 1min + Skipping rope for 4min (you can take a break to adjust the skipping rope in the middle)	2 people

DOI: 10.54254/2755-2721/2026.TJ30929

3.1.2. Data preprocessing

The data processing stage is a key step in ensuring the accuracy and reliability of analytical results. Primitive data contain many outliers, especially cardiac frequency data, which can appear very high or low. These external factors, which, if untreated, can be caused by sensor dysfunction or exceptional interfaces, can introduce deviations into later analyses and highlight the reliability of the study results. Therefore, this study uses an outlier suppression approach during pre-processing to ensure data quality.

As to the data structure, PPG and ACC signals are sampled at 32 Hz, whereas cardiac frequency signal is sampled at 1 Hz. Time order differences between signals may be a result of significant variations in sampling frequencies or interference during data collection. To ensure effective alignment, this study first performs strict time alignment to suppress outliers in the PPG and ACC signals and maintain agreement with the cardiac frequency signal in the time domain.

After matching the PPG and ACC signals, additional correspondence with the cardiac frequency signals is required. Considering the differences in sampling frequencies, in this study, all 32 PPG and ACC samples were mapped to cardiac frequency samples using a 1 Hz cardiac frequency signal as a reference. This proportional correspondence approach exhibits a significant information loss when arranging data with different sampling rates on a unified time scale, thereby providing a reliable basis for later feature extraction and model construction.

3.2. Dual-convolutional attention network based on fusion of PPG and ACC signals

To estimate cardiac frequency during exercise, this article proposes a dual motor attention network (DCAN) to predict cardiac frequency during exercise. As shown in Figure 2, the model adopts a traditional coder-decoder architecture and consults four branches and three main modules: (1) a double-branch entrained feature extraction module; (2) a cross-attention fusion module; (3) a fusion and feature prediction module. The model takes PPG and ACC signals as input, extracts features at multiple levels, models the interactions between the signals, and finally generates predicted cardiac frequency values. The module is introduced in detail below.

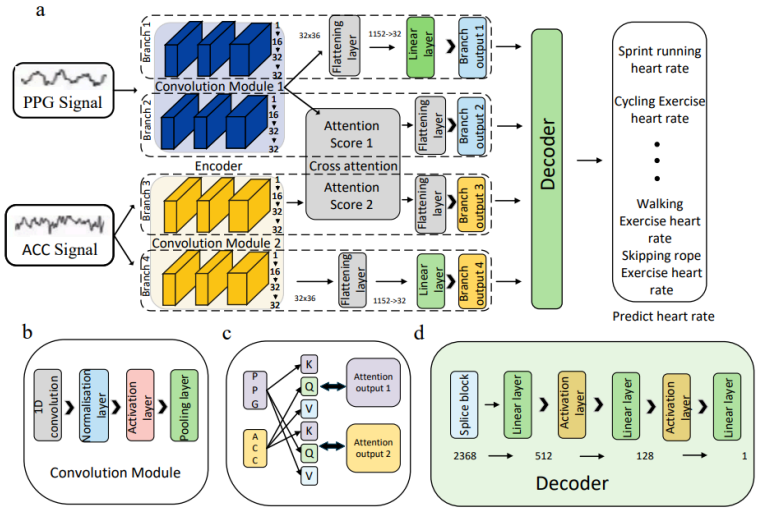

Figure 2. DCAN model architecture diagram. a) Architecture DCAN; b) The internal structure of one entrained layer; c) Multi-head bidirectional cross-attention architecture; d) Architecture of fusion and feature prediction module (decoder)

DOI: 10.54254/2755-2721/2026.TJ30929

3.2.1. Double branch entrainment characteristic extraction module

The PPG signal and the ACC signal have different signal characteristics and a confidence signal confidence signal model, but since both are signals with a 1 D time hierarchy 1 D, the entrainment operation is adapted to the characteristic extraction. This model uses two different entrainment layers (entrainment layer 1 and entrainment layer 2) to process these signals separately. To capture intermodal information, each signal is processed through two branches: the first branch extracts features specific to the signal itself, and the second branch promotes intermodal fusion to extract intermodal features.

• PPG Signal Processing Branch

The PPG signal primarily reflects changes in vessel volume and contains a wealth of information about pulse fluctuations. The PPG feature extraction process designed in this study consists of two stages:

Feature preservation sub-branch: Local features are extracted through a three-layer special feature preservation structure. The entrainment block has an entrainment layer (5x1 core size), a lot-wise normalization layer, and an activation function ReLU, followed by a grouping layer (2x1 core size). This makes it possible to extract and compress the temporal signal, giving a stable representation of the features.

Dimensionality reduction sub-branch: The dimensionality of the extracted features is reduced using fully connected layers and unnecessary information is removed without losing important temporal characteristics. This gives a small, discriminatory feature output in 32 dimensions.

• ACC Signal Processing Branch

The ACC signal (X, Y, Z axis) reflects the user's acceleration changes during movement and provides a reference for suppressing motion artifacts and correcting the PPG signal. The ACC Processing Division also includes two steps:

Sub-branch of characteristic preservation: The temporal characteristics of the multi-channel acceleration signal are extracted through three entrainment blocks. Each block contains an entrainment layer, a lot-wise normalization layer, and an activation function to effectively capture motion models in different directions.

Dimensionality-reducing sub-branch: Fully connected layers compress multidimensional features into a compact 32-dimensional display.

Furthermore, the Dropout mechanism is integrated into the network to reduce the risk of over-adjustment and improve the generalization ability of the model.

3.2.2. Cross-attention fusion module

To establish a dynamic relationship between the characteristics of PPG and ACC signals, this model introduces an innovative bidirectional multi-head cross-attention mechanism. Before discussing the bidirectional multi-head cross-attention mechanism, we first develop the concept of an attention mechanism.

• Attention Computation Process

The traditional mechanism of attention (ego attention) was proposed in a 2017 Transformers article. Attention is calculated based on the sequence itself, and the model can dynamically assign different attention weights (attention scores) to each element of the sequence when processing related sequences. This allows the model to capture the dependencies between different positions in the sequence.

The basic formula for the traditional attention mechanism appears to be:

DOI: 10.54254/2755-2721/2026.TJ30929

$$\text{Attention}(Q,K,V)=\text{Softmax}(\frac{Q^T}{\sqrt{\alpha}})V \tag{1}$$

Of these, Q, K, and V in this expression represent the quest, key, and value, respectively. Q, K, and V are derived from input X by a linear transformation, as shown in Figure 2. The specific formula seems to be:

$$Q=XW^Q \tag{2}$$

$$K=XW^K \tag{3}$$

$$V=XW^V \tag{4}$$

Here, W^Q, W^K, and W^V re the three drivable parameter matrices used to obtain Q, K, and V. The use of three drivable parameter matrices gives the model dynamic tuning capabilities, improving the model's tuning capabilities. To further improve the model's tuning ability, Transformer expands the automatic attention mechanism and proposes a multi-head attention mechanism.The term "multihead" actually refers to multiple mechanisms, and the multihead attention mechanism is actually a multiple ego attention mechanism, as shown in Figure 2. where n is the number of channels, or "heads."

In the case of the cross-attention mechanism, the difference from the automatic attention mechanism is that in automatic attention, quests, keys, and values all come from one sequence, while in cross-attention, Q comes from the first sequence and K and V come from the second sequence. The cross-attention module in this study employs the standard Query-Key-Value computation framework.

• Two-way attention design

The model integrates two sets of cross-attention calculations:

PPG Guided Attention: use PPG features as query and ACC features as key and value, allowing the model to actively retrieve relevant movement information based on heart rate features.

ACC Guided Attention: Conversely, use ACC features as query and PPG features as key and value, analyzing the impact of motion features on PPG signal from the point of view of motion characteristics

This bi-directional design establishes a mechanism of “dialogue” between the signals, capturing more effectively the interactive relationship between the two signals compared to unidirectional attention or simple concatenation of characteristics. Using multi-headed attention further improves the model's learning ability, as several heads can compensate for each other's weaknesses, thereby improving the model's generalization ability.

3.3. Evaluation indicators

This bi-directional design establishes a mechanism of “dialogue” between the signals, capturing more effectively the interactive relationship between the two signals compared to unidirectional attention or simple concatenation of characteristics. Using multi-headed attention further improves the model's learning ability, as several heads can compensate for each other's weaknesses, thereby improving the model's generalization ability:

$$\mathcal{M}AE=\frac{1}{n}\sum_{i=1}^{n}|y_i-\widehat{y_i}| \tag{5}$$

DOI: 10.54254/2755-2721/2026.TJ30929

In the formula, yi represents the actual heart rate value of the i-th sample (unit: bpm), hat{yi is the corresponding predicted value and n is the sample size. This is a dimensional consistent indicator and its value directly defines the physical meaning of the prediction error. It serves as the primary guiding factor in optimizing the model in this research.

The accuracy of the threshold means the percentage of forecasted values within a specific range of the actual value. This research utilizes a two – level evaluation system for thresholds, with a strict accuracy standard (5 bpm) and a loose one (10 bpm). The strict accuracy standard can reflect the accuracy of the model's heart rate prediction, while the flexible accuracy standard can reflect the stability of the model's heart rate prediction.

Strict accuracy standard (±5 bpm): Define that the predicted value falling within the true value range of ±5 bpm is an effective prediction and calculate the proportion of actual predictions:

$$Acc_j = \frac{1}{n} \sum_{i=1}^{n} \Pi\left(|y_i - \hat{y}_i| \leq 5\right) \tag{6}$$

Loose accuracy standard (±10 bpm): Define that the predicted value falling within the true value range of ±10 bpm is an effective prediction and calculate the proportion of effective predictions:

$$Acc_{10} = \frac{1}{n} \sum_{i=1}^{n} \Pi\left(|y_i - \hat{y}_i| \leq 10\right) \tag{7}$$

Where II(.) is an indicator function.

$$\Pi(x) = \begin{cases} 1, & whencondition \\ 0, & otherwise \end{cases} \tag{8}$$

To reflect the generalization ability of the model, this study uses five-fold cross-validation to train and test the model. The specific operations are as follows:

1) Divide the dataset into five mutually exclusive subsets of equal size $\{D_1, ..., D_5\}$.

2) Take as the test set in turn (k=1,…,5) and the remaining subsets as the training set.

3) Calculate the mean and standard deviation of the test results of each fold and finally express them in the form of $\mu \pm \sigma$.

Using five-fold cross-validation can make full use of limited data and avoid the contingency of a single division. At the same time, it uses standard deviation to quantify can reflect the stability and reliability of model performance.

4. Exercise heart rate detection results and analysis

4.1. Experimental parameter settings

The study employs the PyTorch framework in Python to construct the model. To simulate real-world wearable devices, model training is conducted using the built-in CPU (AMD Ryzen 5 5625U) of a standard computer, without utilizing a GPU. The training settings are as follows (see Table 2): the Adam optimizer is selected, the loss function is Mean Squared Error Loss (MSE loss), the initial learning rate is set to 0.001, the batch size is 32, and the maximum number of training epochs is 100.

DOI: 10.54254/2755-2721/2026.TJ30929

Table 2. Model training parameter settings

Parameter	Adam Value
Optimizer	Adam
Loss Function	MSE loss
Learning Rate	0.001
Batch Size	128
Max Epochs	100

4.2. Experimental results

4.2.1. Individual experimental results

In the experiments involving model training and testing for individual subjects, the MAE results are shown in Table 3, and the accuracy results are presented in Table 4. Training on a single person allows testing the models ability to learn personalized physiological indicators and specific exercise types.

Table 3. Comparison of model MAE performance (unit: BPM)

Activity Scene	DCAN	C-RNN	NAS-PPG
Sprint Run 1	3.06±1.05	4.46±2.51	4.08±0.35
Sprint Run 2	4.65±0.69	8.62±2.25	9.65±1.82
Outdoor Walking 1	1.72±0.13	2.84±0.12	2.80±0.41
Outdoor Walking 2	1.15±0.22	2.07±0.20	2.06±0.12
Outdoor Running 1	2.64±0.20	2.86±0.47	3.03±0.54
Outdoor Running 2	2.33±0.72	2.70±0.40	3.23±0.47
Outdoor Cycling 1	2.36±0.86	2.99±0.63	4.04±0.42
Outdoor Cycling 2	1.08±0.12	1.41±0.17	1.68±0.23
Rest 1	1.48±0.34	1.99±0.36	1.96±0.41
Rest 2	1.13±0.25	0.68±0.15	1.58±0.28
Jumping Jacks 1	5.24±1.28	6.61±0.72	9.17±1.18
Jumping Jacks 2	7.01±1.72	8.87±4.09	8.31±1.91
Stair Climbing 1	1.98±0.39	3.65±0.73	5.50±1.61
Stair Climbing 2	2.26±0.79	4.12±0.86	5.29±1.01
Indoor Walking 1	0.83±0.14	1.41±0.07	1.53±0.19
Indoor Walking 2	1.02±0.07	0.99±0.07	1.84±0.27
Indoor Running 1	2.50±0.13	2.98±0.44	4.22±0.22
Indoor Running 2	2.27±0.25	3.74±0.46	2.95±0.29
Indoor Cycling 1	1.72±0.39	2.69±0.27	2.84±0.33
Indoor Cycling 2	1.60±0.46	2.56±0.31	2.93±0.35
Skipping Rope 1	5.75±1.20	10.10±1.68	8.57±3.01
Skipping Rope 2	5.65±0.72	6.72±6.42	7.37±0.92
Free Activity 1	1.16±0.12	1.48±0.16	2.07±0.17
Free Activity 2	2.12±0.19	3.98±0.64	3.68±0.46
Average MAE	2.51	2.26	3.71

DOI: 10.54254/2755-2721/2026.TJ30929

Table 4. Comparison of model accuracy performance

ctivity Scene	±5 BPM Accuracy(%)			±10 BPM Accuracy(%)		
	DCAN	C-RNN	NAS-PPG	DCAN	C-RNN	NAS-PPG
Sprint Rum 1	81.88±7.88	77.43±13.63	71.80±3.29	96.65±3.60	93.33±9.05	95.68±2.97
Sprint Rum 2	61.41±11.81	48.31±13.43	70.51±12.29	86.45±2.05	69.63±9.59	94.98±8.97
Outdoor Walking 1	98.89±0.84	82.77±4.08	83.33±5.31	100.00±0.00	98.98±0.83	98.75±1.02
Outdoor Walking 2	97.70±1.94	88.26±1.92	93.09±4.00	99.85±0.81	99.36±1.27	97.95±0.31
Outdoor Rumning 1	82.74±7.50	94.86±5.35	84.04±2.90	97.44±0.45	99.50±0.33	97.93±0.69
Outdoor Rumning 2	92.26±8.51	88.13±9.06	90.33±6.78	98.13±1.80	96.60±1.22	97.47±0.52
Outdoor Cycling 1	90.38±8.19	85.19±4.76	80.70±5.08	99.73±0.03	96.97±2.36	93.73±1.04
Outdoor Cycling 2	98.14±1.19	97.33±0.70	95.93±1.64	99.98±0.23	100.00±0.00	99.64±0.70
Rest 1	99.46±1.02	100.00±2.76	92.23±4.65	97.00±1.71	99.92±1.32	100.00±0.00
Rest 2	95.14±2.08	95.68±0.00	97.47±3.44	100.30±0.00	100.00±0.00	98.88±2.16
Jumping Jacks 1	59.72±6.89	61.94±6.30	39.79±4.58	86.67±9.99	72.38±12.04	70.28±7.64
Jumping Jacks 2	67.64±13.90	49.16±24.35	59.24±11.70	77.41±5.02	88.79±3.17	72.13±5.76
Stair Climbing 1	84.59±7.43	74.47±10.58	62.63±11.97	96.96±3.18	91.58±3.81	83.71±5.64
Stair Climbing 2	92.87±6.80	83.58±8.98	65.73±9.07	97.43±1.05	94.54±1.52	85.85±4.34
Indoor Walking 1	99.31±0.62	98.06±1.60	97.93±1.24	100.00±0.00	100.00±0.00	99.72±0.34
Indoor Walking 2	99.16±0.87	97.02±0.71	96.11±4.79	99.86±0.82	99.36±0.85	99.83±0.65
Indoor Rumning 1	86.80±11.90	86.56±5.00	67.34±1.58	93.53±0.28	93.86±3.85	95.18±1.65
Indoor Rumning 2	99.14±0.61	97.02±5.07	95.74±1.77	99.86±0.28	99.36±0.85	99.83±0.65
Indoor Cycling 1	88.03±8.27	82.73±4.81	82.24±3.78	99.61±0.09	97.93±1.04	97.74±1.28
Indoor Cycling 2	94.71±3.53	85.73±5.46	85.35±4.15	98.11±1.10	97.05±2.60	97.17±0.67
Skipping Rope 1	60.31±2.74	88.38±3.67	81.74±2.98	99.41±0.55	98.53±0.93	98.97±1.00
Skipping Rope 2	65.62±8.11	69.33±22.33	59.49±6.25	88.07±7.86	74.13±26.41	73.18±9.10
Free Activity 1	60.44±9.70	50.16±5.03	58.94±4.03	84.40±4.38	87.73±3.26	74.74±5.13
Free Activity 2	99.18±0.74	82.80±1.24	93.24±0.98	100.00±0.00	100.00±0.00	95.90±0.66
Average	91.17±9.20	82.25±6.53	73.34±8.85	99.83±0.33	98.83±1.25	95.17±2.91

The DCAN model proposed in this study demonstrates significant advantages across 24 test scenarios. In terms of mean absolute error (MAE), DCAN achieves an average MAE of 2.51 BPM, representing a 23% reduction compared to C-RNN (3.26 BPM) and a 32% reduction compared to NAS-PPG (3.71 BPM). The primary reason DCAN's average MAE is lower than that reported in some existing studies lies in differences in datasets—variations in dataset size and the granularity of exercise type classification leads to differences in model learning difficulty. In terms of ±5 BPM and ±10 BPM accuracy metrics, DCAN generally outperforms the baseline models (Table 4). Although not the best in the resting scenario, it is very close to the optimal performance and nearly reaches 100%, with only a negligible gap. Notably, DCAN achieves the best MAE performance in the vast majority of scenarios, such as:

In the Sprint Run 1 scenario, the MAE of DCAN (3.06 ± 1.05) is only 68.6% of that of C-RNN, and the precision ±10 BPM (96.65%) is 1.67 points higher than that of the NAS-PPG. In the Jump Rope 2 scenario, the MAE of DCAN (5.65 ± 0.72) is 15.9% lower than that of C-RNN, and its precision difference ±5 BPM (±8.09%) is significantly lower than that of C-RNN ±15.03%, indicating that its output is more stable.

Comparing the typical differences for each model, we find that DCAN works more stably in most scenarios.:

DOI: 10.54254/2755-2721/2026.TJ30929

– The MAE type difference for indoor driving 1 (-0.14) is only 75.4% of that of NAS-PPG.

– The precision difference ±5 BPM for free activity 1 (± 0.74%) is much lower than that of C-RNN (±1.24%).

This indicates that DCAN offers a capacity for generation between stronger scenarios and is particularly suited to monitoring cardiac frequency in complex real environments.

4.2.2. Results under different exercise intensities

In order to minimize the influence of physiological differences and to focus on examining the influence of exercise type (intensity) on cardiac frequency prediction, this study formed a model that uses exercise intensity classification to analyze how exercise intensity influences cardiac frequency estimation.

The intensity of the exercise is classified as follows:

Exercise intensity is categorized as:

– High intensity: sprints, jumps with a lan, jump chords, climbing stairs

– Average intensity: Outdoor course, outdoor course, outdoor cycling, outdoor cycling

– Low intensity: running forward, running forward, rest, free activity

As will become clear from Table 5, the DCAN model showed significant performance benefits at low- and high-intensity practice skinarios, demonstrating the effectiveness of the two-channel attention architecture.

Table 5. Comprehensive performance comparison of models under different exercise intensities

Model	Intensity	MAE(BPM)	±5 BPM(%)	±10 BPM(%)
DCAN	Low	1.79±0.34	96.18±1.53	99.36±0.42
	Medium	2.97±1.52	82.97±9.16	96.38±3.07
	High	3.74±0.81	76.45±2.84	93.83±5.75
C-RNN	Low	2.26±0.23	91.46±2.51	98.22±0.60
	Medium	2.16±0.75	91.23±5.83	96.65±1.18
	High	5.67±0.69	63.98±9.44	87.49±4.79
NAS-PPG	Low	2.15±0.11	92.07±3.08	98.62±0.19
	Medium	3.22±0.23	80.58±7.04	96.89±0.37
	High	5.15±0.59	63.97±4.79	89.04±1.80

In the low-intensity exercise scenario, DCAN achieved an MAE of 1.79±0.34 BPM and an accuracy of 96.18±1.53%±5 BPM, outperforming all comparison models. ±The typical difference in accuracy at 10 BPM is only 0.42%, indicating a high ability to capture physiological basic rhythms. This advantage comes from the advantage of efficiently harvesting PPG periodicity in the cross-fair module and the subtle behavior of the business branch.

In the mean intensity exercise scenario, C-RNN was able to achieve a slightly better MAE than DCAN (2.97 ± 1.52 BPM), which may be why the CNN-RNN combination is better at obtaining continuous characteristics in mean intensity exercise and moderate exercise with stronger time dependence. NAS-PPG is also a sequence-based model, but its generalization ability may be weak. Overall, the performance of DCAN in the mean intensity scenario is comparable to C-RNN.

In high-intensity training skinarios, DCAN has the most important advantages of MAE (3.74 ± 0.81 BPM) 34.0% lower than C-RNN and accuracy ±5 BPM (76.97%) 13.0 points higher than NAS-PPG.

DOI: 10.54254/2755-2721/2026.TJ30929

Overall, DCAN works perfectly to predict heart rate during low- and high-intensity exercise. Future work could focus on multi-model fusion or adding data to further improve the accuracy of the mean intensity scenario.

4.3. Abratiocoe

The DCAN model consists of four branches. To confirm the effectiveness of each branch, a cutting test was performed by removing each branch individually. Models were created and evaluated in individual frames under the same experimental conditions. Detailed experimental results are shown in Tables 6-8.

Table 6. Performance comparison of ablation experiment MAE (BPM)

Scene	DCAN	withoutb1	without b2	without b3	without b4
Sprint Run 1	3.06±1.05	3.70±1.31	3.59±1.24	2.86±0.66	2.43±0.28
Sprint Run 2	4.65±0.69	5.19±0.69	6.70±1.00	5.82±2.68	4.89±1.00
Outdoor Walking 1	1.72±0.13	1.82±0.49	1.27±0.13	1.50±0.17	1.61±0.30
Outdoor Walking 2	1.15±0.22	1.30±0.04	1.64±0.27	1.48±0.20	1.34±0.18
Outdoor Running 1	2.64±0.22	3.28±1.20	3.02±0.42	2.59±0.23	2.29±0.45
Outdoor Running 2	2.33±0.72	3.18±2.11	2.47±0.71	3.05±1.28	2.35±0.38
Outdoor Cycling 1	2.36±0.86	1.58±0.18	2.36±0.51	1.42±0.19	1.50±0.46
Outdoor Cycling 2	1.08±0.12	1.29±0.10	1.27±0.10	1.21±0.13	1.15±0.14
Rest 1	1.48±0.34	1.67±0.30	1.72±0.26	1.66±0.28	1.65±0.44
Rest 2	1.13±0.25	1.03±0.11	1.15±0.26	1.23±0.09	1.10±0.15
Jumping Jacks 1	5.24±1.28	6.67±0.81	5.99±1.32	6.88±1.61	5.85±0.94
Jumping Jacks 2	7.01±1.72	5.17±0.88	6.88±1.62	6.58±1.74	6.75±1.28
Stair Climbing 1	1.98±0.39	2.65±0.25	2.26±0.65	2.38±0.50	1.98±0.38
Stair Climbing 2	2.26±0.79	1.93±0.36	2.85±0.86	2.79±0.73	2.19±0.28
Indoor Walking 1	0.83±0.14	0.92±0.18	1.20±0.06	0.99±0.13	0.93±0.07
Indoor Walking 2	1.02±0.07	1.06±0.17	1.09±0.15	1.18±0.44	0.90±0.23
Indoor Running 1	2.50±0.13	2.37±0.13	3.14±0.29	3.26±0.69	3.32±0.88
Indoor Running 2	2.27±025	2.44±0.29	2.97±1.35	2.37±0.18	2.86±0.30
Indoor Cycling 1	1.72±0.39	1.75±0.25	1.69±0.51	1.57±0.42	1.28±0.41
Indoor Cycling 2	1.60±0.46	1.58±0.35	1.68±1.29	1.62±0.94	1.12±0.23
Skipping Rope 1	5.75±1.20	5.88±022	6.60±0.34	4.61±0.21	7.42±1.23
Skipping Rope 2	5.65±0.72	5.03±0.46	5.94±0.64	5.13±0.63	5.12±0.55
Free Activity 1	1.16±0.12	1.26±0.11	1.43±0.13	1.33±0.06	101±0.06
Free Activity 2	2.12±0.19	2.32±0.31	2.83±0.33	2.55±0.28	2.27±0.17

Table 7. Performance comparison of ablation experiment ±5 BPM accuracy (%)

Scene	DCAN	without b1	without b2	without b3	without b4
Sprint Run 1	81.08±11.88	78.53±20.44	76.18±17.69	80.10±14.52	90.22±5.63
Sprint Run 2	67.47±7.01	79.31±9.14	54.62±5.90	62.07±9.65	95.45±3.63
Outdoor Walking 1	98.90±1.64	97.36±2.21	97.23±2.00	98.63±0.70	97.80±1.42
Outdoor Walking 2	97.97±0.94	94.83±1.31	97.07±1.70	97.31±2.36	98.67±0.94
Outdoor Running 1	93.08±7.90	89.30±27.34	91.29±9.07	90.01±17.65	92.21±4.75
Outdoor Running 2	87.61±7.19	78.30±19.61	84.23±10.05	80.46±12.25	91.48±2.05

DOI: 10.54254/2755-2721/2026.TJ30929

Table 7. (continued)

Outdoor Cycling 1	90.28±8.51	97.28±1.42	91.25±9.83	99.69±1.38	97.26±1.24
Outdoor Cycling 2	99.13±1.51	97.03±1.85	94.85±4.00	97.05±1.03	98.63±1.01
Rest 1	99.46±1.02	98.56±1.08	96.92±2.04	92.50±1.83	100.00±0.00
Rest 2	97.24±6.09	99.46±1.04	98.97±0.16	99.46±1.08	100.00±0.00
Jumping Jacks 1	69.79±8.09	60.56±14.68	66.92±9.16	55.40±12.83	60.22±5.92
Jumping Jacks 2	67.64±13.09	97.10±7.85	99.06±7.55	80.67±12.17	93.16±10.45
Stair Climbing 1	95.94±1.38	99.02±1.35	95.07±8.67	96.67±6.24	95.06±4.64
Stair Climbing 2	99.31±0.62	99.10±2.15	99.06±7.55	98.83±5.64	92.95±8.34
Indoor Walking 1	89.70±8.10	89.40±7.10	89.85±10.68	96.31±0.62	95.06±1.43
Indoor Walking 2	99.31±0.62	99.17±2.02	93.98±1.06	98.83±5.64	99.09±0.63
Indoor Running 1	89.80±1.61	89.72±0.34	90.08±4.01	98.61±1.08	99.12±1.62
Indoor Running 2	99.16±0.87	99.62±3.77	98.09±0.79	97.75±0.68	97.73±0.67
Indoor Cycling 1	93.89±3.23	87.79±1.04	93.85±1.17	98.42±1.39	97.17±3.08
Indoor Cycling 2	98.07±5.57	95.23±3.03	91.56±3.42	92.42±5.81	93.54±2.74
Skipping Rope 1	65.42±8.91	64.98±7.08	57.27±5.72	66.30±8.58	69.06±8.53
Skipping Rope 2	60.61±8.04	59.60±7.07	52.06±5.86	62.33±8.87	41.00±5.13
Free Activity 1	99.18±0.74	99.18±0.26	98.68±1.12	99.68±0.84	99.67±0.66
Free Activity 2	91.17±2.04	90.17±3.92	83.69±4.23	87.85±4.52	91.34±2.91

Table 8. Performance comparison of ablation experiment ±10 BPM accuracy (%)

Scene	DCAN	without b1	without b2	without b3	without b4
Sprint Run 1	96.65±3.60	97.95±4.69	95.81±5.67	98.05±1.80	92.85±5.04
Sprint Run 2	86.45±2.05	84.99±8.22	82.77±4.79	82.47±7.14	95.08±1.44
Outdoor Walking 1	100.00±0.00	99.98±0.28	99.98±0.28	100.00±0.00	99.87±0.25
Outdoor Walking 2	99.75±0.51	99.86±0.25	98.97±0.25	99.62±0.51	100.00±0.00
Outdoor Running 1	98.44±1.48	98.51±1.61	98.34±1.62	97.26±1.25	98.94±2.91
Outdoor Running 2	98.13±1.80	95.18±6.61	97.57±0.18	97.77±1.62	98.08±1.01
Outdoor Cycling 1	97.93±0.33	97.93±0.33	97.93±0.34	100.00±0.00	97.84±1.08
Outdoor Cycling 2	99.70±1.71	99.75±2.02	99.84±1.08	99.84±1.08	100.00±0.00
Rest 1	97.00±1.71	97.84±1.08	97.84±1.08	97.84±1.08	100.00±0.00
Rest 2	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00
Jumping Jacks 1	86.67±9.99	84.17±7.64	87.88±4.51	81.67±9.05	84.17±4.70
Jumping Jacks 2	80.71±5.00	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00
Stair Climbing 1	97.41±2.22	100.00±0.00	100.00±0.00	89.19±1.14	79.93±6.78
Stair Climbing 2	99.96±1.08	98.43±5.84	98.00±7.82	98.92±1.32	90.00±0.00
Indoor Walking 1	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	99.12±1.78
Indoor Walking 2	99.63±0.82	99.98±0.28	99.98±0.28	99.62±0.51	100.00±0.00
Indoor Running 1	99.80±0.28	99.96±0.14	99.98±0.28	100.00±0.00	97.36±1.99
Indoor Running 2	99.16±0.87	99.92±0.28	99.68±2.53	99.02±1.07	99.00±1.00
Indoor Cycling 1	99.66±1.09	97.51±1.71	98.67±5.17	99.25±1.43	95.02±1.03
Indoor Cycling 2	98.11±1.99	99.92±0.32	99.68±3.97	99.02±1.07	99.85±1.02
Skipping Rope 1	99.41±0.55	98.95±1.06	99.71±3.04	99.71±3.04	97.05±3.29
Skipping Rope 2	88.07±7.86	99.71±3.32	99.71±3.02	99.28±2.46	98.80±1.03
Free Activity 1	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00
Free Activity 2	99.43±0.33	99.67±0.67	99.33±0.62	99.83±0.33	99.67±0.41

DOI: 10.54254/2755-2721/2026.TJ30929

Ablation results show that removing Branches 1–4 can lead to performance improvements in certain motion scenarios, possibly because some branches interfere with heart rate prediction under specific conditions. Therefore, averaged results are used for analysis, as shown in Table 6-8.

Analysis of each individual branch:

Branch 1 (b1): Branch 1 uses a three-layer CNN to extract the fundamental characteristics of the PPG signal. The results of the resection indicate that this branch is important in intense movement scenarios (e.g., leg elevation or arm vibration). Suppression causes MAE to rise by 0.48 BPM and accuracy to fall by 7.16% at 5 BPM. This implies that the support captures signal function mainly in situations of being stationary and moving slowly. In comparison with other areas, the extracted time horizon features focus on the low - frequency component, which is good for the overall stability of the model. In summary, Branch 1 successfully enhances functional extraction during low - intensity exercise, and deeper pull - in structures can make solidity better in future work.

Branch 2 (b2): Branch 2, which is based on Branch 1, incorporates an attention-based request mechanism to model temporal characteristics that influence motion. The ablation results showed that suppression of this support decreased accuracy by 8.39% at 10 BPM and 6.58% at 5 BPM, suggesting its significance in mean-intensity exercise scenarios. The internal attention module assists the model in concentrating on key information in complex motion settings and improves its capability to differentiate relevant motion features.

Branch 3 (b3): Branch 3 takes in the accelerometer auxiliary signal and fine - tunes the PPG modeling with spatial motion features. It's especially useful for rhythmic activities, such as jogging or cycling. Excision experiments have shown that this suppression reduces the accuracy of 10 BPM by 9.36% and 5 BPM by 8.97%, capturing a periodic motion model that makes heart rate estimation more stable and accurate. Overall, Branch 3 is a necessary complement to model acceleration-based features, especially for types of lead rhythm activity...

Branch 4 (b4): Branch 4 combines global properties through a triaxial acceleration signal through an attention mechanism and acts as a "stabilizer" for the model. The results of ablation show that it is essential in complex situations (e.g., free activity, sprint): its removal increases MAE at 0.63 BPM in sprint training 1 and decreases accuracy at 0.16% ±5 BPM in free activity 2. We confirm that the influence on the static environment (e.g., rest) is minimal and that this branch focuses on multiaxial motion property fusion. Stability proposals showing that the product is acting as a comparative pressure on other sectors based on the characteristic standards selected for the dynamic adaptation of sector 1/block 3 inputs.

In summary, the DCAN model is able to achieve highly precise cardiac manipulation by combining a large collection of characteristics (local + global), behavioral artifacts (interaction of dynamic characteristics), and speed of activity (branching complementarity). Ablation experiments confirm that each branch participates meaningfully and the entire model achieves the best overall performance, confirming the efficiency of the DCAN architecture.

5. Conclusion

The proposed Dual Combustion Attention Network (DCAN) has shown strong performance advantages in exercise-based cardiac velocity assessment. Compared to traditional C-RNN and NAS-PPG models, DCAN significantly reduces the mean absolute error (MAE) by 2.51 BPM compared to C-RNN, achieves a 23% reduction compared to NAS-PPG, and achieves excellent accuracy of ±5 BPM and ±10 BPM. These results indicate that DCAN effectively adapts to the intensity and scenario of diversity education and has strong individual benefits in low- and high-intensity environments.

DOI: 10.54254/2755-2721/2026.TJ30929

Compared to C-RNNs, DCAN not only removes long-term dependencies in sequences, but also uses Bose DCAN to improve the display of local properties. Compared to NAS-PPG, DCAN companies offer multimodal signal fusion, PPG cardiovascular data exchange, and business function ACC. This complementary relationship is important in scenarios with strong behavioral artifacts. If the PPG signal is strongly blocked, ACC provides valuable business context for removing ACC, and if ACC alone cannot well represent the business space, the blood dynamic properties of PPG provide reliable compensation.

Furthermore, the design of a two-way crossover detection mechanism is important to improve the performance of DCANs. By enabling two lines of attention between PPG and ACC, the model generates dynamic intermodal interactions, overcoming the limitations of static chaining used in traditional methods. This mechanism allows the model to adaptively adjust the concentration on the signal characteristics depending on the usage conditions, improving the accuracy and stability of the prediction. ablation study further confirms the effectiveness of the corresponding modules. Real estate preservation, dimensionality reduction, attention mechanisms, and multilevel trait fusion have all the effects.

DCAN Excellent performance arises from the following three main factors: With the double wound branch design, the extract of the characteristics specific to modularity is guaranteed; Bidirectional mutual attention promotes deeply dynamic complementarity between signals; Fusing hierarchical characteristics maintains detailed local details while incorporating high-level semantics. By combining these elements, DCAN You can make more accurate and robust heart rate predictions in operating scenarios.

The physiological monitoring systems that were developed in this study were tested in a controlled laboratory setting, yet they are not without limitations. Due to the small sample size, the generalizability and representativeness of the findings might be restricted. Moreover, test settings are comparatively optimal and signal interference is unlikely to capture all the nuances of complex real-life and clinical scenarios overall. Hardware performance restricts some sensors, which enhances acquisition accuracy. The next step can involve extending datasets, enhancing algorithms, optimizing hardware design, and enhancing system stability and utility.

This PPG ACC study is methodological and enhances the detection of heart rhythm based on fusion and gives a new perspective to multimodal physiological signaling research. The focus on complementarity and dynamic interactions may greatly influence the creation of wearable devices in health monitoring and sports science.

References

[1] Kamal A, Harness J, Irving G, et al. Skin photoplethysmography—A review. Computer Methods and Programs in Biomedicine, 1989, 28(4): 257–269.

[2] Fernandes R A, Vaz Ronque E R, Venturini D, et al. Resting heart rate: Its correlations and potential for screening metabolic dysfunctions in adolescents. BMC Pediatrics, 2013, 13: 1–7.

[3] Garber C E, Blissmer B, Deschenes M R, et al. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: Guidance for prescribing exercise. Medicine & Science in Sports & Exercise, 2011, 43(7): 1334–1359.

[4] Biswas D, Simões-Capela N, Van Hoof C, et al. Heart rate estimation from wrist-worn photoplethysmography: A review. IEEE Sensors Journal, 2019, 19(16): 6560–6570.

[5] Zhang Z. Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction. IEEE Transactions on Biomedical Engineering, 2015, 62(8): 1902–1910.

[6] Zhang Z, Pi Z, Liu B. TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. IEEE Transactions on Biomedical Engineering, 2014, 62(2): 522–531.

DOI: 10.54254/2755-2721/2026.TJ30929

[7] Cotter S F, Rao B D, Engan K, et al. Sparse solutions to linear inverse problems with multiple measurement vectors. IEEE Transactions on Signal Processing, 2005, 53(7): 2477–2488.
[8] Hasan K, Chowdhury M H, Pathan N S, et al. Heart rate estimation from wrist PPG signal during intense physical exercise. SN Computer Science, 2023, 4(5): 684.
[9] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. 2014.
[10] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. Advances in Neural Information Processing Systems, 2017, 30.