electronics

MDPI

Article

Exercise ECG Classification Based on Novel R-Peak Detection Using BILSTM-CNN and Multi-Feature Fusion Method

Xinhua Su, Xuxuan Wang * and Huanmin Ge

School of Sports Engineering, Beijing Sport University, Beijing 100084, China; suxinhua@bsu.edu.cn (X.S.); gehuanmin@bsu.edu.cn (H.G.)
* Correspondence: wxx0419@bsu.edu.cn

Abstract: Excessive exercise is a primary cause of sports injuries and sudden death. Therefore, it is vital to develop an effective monitoring technology for exercise intensity. Based on the noninvasiveness and real-time nature of an electrocardiogram (ECG), exercise ECG classification based on ECG features could be used for detecting exercise intensity. However, current R-peak detection algorithms still have limitations, especially in high-intensity exercise scenarios and in the presence of noise interference. Additionally, the features utilized for exercise ECG classification are not comprehensive. To address these issues, the following tasks have been accomplished: (1) a hybrid time–frequency-domain model, BILSTM-CNN, is proposed for R-peak detection by utilizing BILSTM, multi-scale convolution, and an attention mechanism; (2) to enhance the robustness of the detector, a preprocessing data generator and a post-processing adaptive filter technique are proposed; (3) to improve the reliability of exercise intensity detection, the accurate heart rate variability (HRV) features derived from the proposed BILSTM-CNN and comprehensive features are constructed, which include various descriptive features (wavelets, local binary patterns (LBP), and higher-order statistics (HOS)) tested by the feasibility experiments and optimized deep learning features extracted from the continuous wavelet transform (CWT) of exercise ECG signals. The proposed system is evaluated by real ECG datasets, and it shows remarkable effectiveness in classifying five types of motion states, with an accuracy of 99.1%, a recall of 99.1%, and an F1 score of 99.1%.

Keywords: R-peak detection; exercise ECG signals; feature fusion; deep learning; exercise intensity

check for updates

Academic Editor: Michael Winter

Received: 30 November 2024
Revised: 4 January 2025
Accepted: 8 January 2025
Published: 12 January 2025

Citation: Su, X.; Wang, X.; Ge, H. Exercise ECG Classification Based on Novel R-Peak Detection Using BILSTM-CNN and Multi-Feature Fusion Method. *Electronics* **2025**, *14*, 281. https://doi.org/10.3390/electronics14020281

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Excessive exercise is a primary cause of sports injuries and sudden death. Therefore, fatigue detection is of vital importance [1–3]. As a vital sign monitoring technology, an ECG is extensively utilized in modern healthcare devices, owing to its noninvasiveness and convenience. The HRV features of ECG signals are measured by the corresponding R-peak location of ECG signals, providing key insights into heart health, stress response, and exercise performance. Sivanantham used linear and nonlinear HRV features for cardiac arrhythmia detection and achieved an accuracy of 90.26% [4]. Jos Vicente focused on the detection of driver's drowsiness by means of HRV analysis and achieved a sensitivity of 85% for drowsy minutes classification [5]. Meng proposed a sleep stages classification method based on HRV features, achieving an average accuracy of 88.67% for subject-specific classifiers [6]. Mustafa Radha used Long Short-Term Memory (LSTM) neural networks to classify sleep stages based on HRV features, achieving an accuracy of 77.00 ± 8.90%

across the entire database [7]. Marta Vigier used HRV features for cancer classification, achieving a classification accuracy of 86% for the ensemble model [8]. To enhance the accuracy of classification, methods based on fusion features have been adopted. Chen integrated HRV features and ECG image features for exercise fatigue classification and the accuracy value improved to 94.32% [9]. Cheol Ho Song developed a multi-dimensional feature fusion-based stress classification method by combining LSTM and Xception, which overcame the limitations of traditional HRV-based stress classification and achieved an accuracy of 99.51% [10]. Ahmed S. Eltrass extracted ECG features by integrating HRV features, other linear and nonlinear features extracted from ECG, and features of the deep neural network method, and the corresponding proposed diagnostic system achieved an accuracy of 98.75% [11]. Prakash put forward a modified 1-D U-Net architecture to identify electrocardiogram fiducial points and segment the signals. Subsequently, features were extracted and classified using the Random Forest algorithm. The proposed method was validated on two publicly available databases, namely LUDB and MIT-BIH. Experimental results demonstrate that the performance of the proposed method was superior to that of state-of-the-art techniques [12]. Although all the aforementioned methods have introduced effective features and achieved good results, none of them have focused on enhancing the reliability of HRV features, which may limit further improvements in performance.

As mentioned above, HRV features are important physiological indicators for analyzing ECG signals. Note that, the correct R-peak detection is the base of the accurate HRV features. Recently, R-peak detection techniques have been predominantly focused on the medical field, where the ECG signals are relatively stable and the noise interference is small and controlled. So far, the R-peak detection research can be mainly divided into two categories: the classical threshold-based methods and deep learning methods. The classical R-peak detection algorithms primarily enhance the energy of the QRS complex, followed by the amplitude-threshold-based R-peak detection [13]. For example, JP Martínez et al. developed a wavelet-based detector that remains outstanding in standard databases [14]. These algorithms could rapidly detect R-peaks, but few consider robustness to noise ECG signals. Deep learning methods have shown remarkable performance in R-peak detection. Wang proposed two parallel 1D residual neural networks to capture the time-domain characteristics of QRS waveform and attained a sensitivity value of 99.92% on the MIT-BIH Arrhythmia dataset (MITDB) [15]. Laitala applied LSTM networks to construct the R-peak detector [16]. Vijayrangan combined U-Net, Inception, and Residual blocks to detect R-peak on a combined dataset and achieved nearly an accuracy of 98.37% [17]. Mehri proposed an R-peak detection architecture for 3D vectorcardiogram (VCG) that requires no preprocessing or post-processing steps. Experiments were conducted on four different public databases. Using the proposed approach, high F1 scores of 99.80% and 99.64% were achieved in the leave-one-out cross-validation and cross-database validation protocols [18]. Wang proposed an end-to-end electrocardiogram waveform detection method (ECT-net) based on a convolutional neural network (CNN) and transformer. It performed best among comparison methods and achieved F1 scores of 94.27% for the P wave, 97.32% for the QRS wave, and 93.92% for the T wave [19].

At present, most of the research focuses on feature analysis and the classification of medical ECG signals. Unfortunately, it is unreasonable to directly apply these methods to analyze exercise ECG signals. The reason is that exercise ECG signals fluctuate greatly under different exercise intensities, and strong noises are generated during exercise. Thus, for exercise ECG signals, the accuracy of HRV features derived from R-peak detection and the effectiveness of the resulting ECG classification all face the great challenges. Therefore, in order to improve the reliability of HRV features and explore more abundant ECG features for exercise fatigue classification, we have developed an adaptive R-peak detection

algorithm suitable for high-intensity exercise and introduced more suitable descriptor features and more subtle deep learning features.

The main goal of this work is to devise an effective solution for exercise fatigue detection based on ECG. Compared with previous studies, in order to enhance the reliability of HRV, we bolster the reliability of R-peak detection. Specifically, we introduce an innovative ECG data generator. This device is purposefully engineered to generate noisy ECG signals, which act as essential training materials for nurturing a robust R-peak detector. At the same time, to overcome the difficulties caused by the scarcity of publicly available exercise ECG data and the lack of precise labels, we carry out intra-class segmentation. This methodology significantly increases the number of samples in our dataset, laying a more stable groundwork for further analysis. Based on the above-mentioned generated training data, a novel R-peak detection method based on BILSTM and multi-scale convolution (called BILSTM-CNN) is proposed, whose inputs are hybrid time–frequency signals. To the best of our knowledge, this is the first attempt to detect R-peaks using time–frequency features. And, an adaptive filtering function is proposed to further filter out unnatural R-peaks during high-intensity exercise. Up to this point, the algorithm specifically designed for detecting R-peaks in ECG signals during high-intensity exercise has been completed. Based on the proposed R-peak detector BILSTM-CNN, accurate HRV features are obtained and then the accuracy of exercise ECG classification is improved. Furthermore, by integrating HRV features with statistical descriptor features and deep learning features extracted from ECG, the proposed ECG classification framework has an accuracy of up to 99.1%, which provides a strong base for exercise fatigue monitoring based on ECG signals. This method provides a certain basis for sports fatigue detection.

The rest of this article is organized as follows: Section 2 describes the datasets. In Section 3, pre-processing, the proposed robust and adaptive R-peak detection model BILSTM-CNN and the resulting multi-feature fusion model for exercise ECG classification are introduced. Section 4 introduces experiments and results, which demonstrate the effectiveness of the proposed R-peak detection model BILSTM-CNN and the novel ECG classification method based on multi-feature fusion. Finally, conclusions and future work are provided in Section 5.

2. Datasets

In this work, two prominent available databases, the GUDB database [20] and the EPFL database (https://zenodo.org/records/5727800; accessed on 25 November 2021) [21], are used to train and test our R-peak detection model. Clearly, the more accurate the R-peak detection, the higher the accuracy of ECG classification. Thus, the experiment results of ECG classification could further validate the effectiveness of the proposed R-peak detection method. Here, the EPFL database is employed in exercise ECG classification for fatigue detection due to the standard classification of fatigue levels shown in Figure 1.



Figure 1. The segments corresponding to the levels of exercise fatigue [21] are seg1: [VT2 − 50, VT2 − 30], seg2: [VT2 + 60, VT2 + 80], seg3: [VO2max − 50, VO2max − 30], seg4: [VO2max − 10, VO2max + 10], and seg5: [VO2max + 60, VO2max + 80].

(1) GUDB: The GUDB database includes 24 ECG records of 120 s duration from 24 subjects. The ECG signals were recorded using an Attys Bluetooth data acquisition

board. This board has a sampling rate of 250 Hz and a resolution of 24-bit over a range of ±2.42 V. All the records are collected during 3 physical activities: treadmill walking, hand-bike usage, and jogging. Since the GUDB dataset has the R-labeled dataset, it is used to train and test R-peak detection for exercise ECG data.

(2) EPFL: The EPFL database contains 100 ECG records from 20 subjects and 5 different segments, each record has a 20s duration. The ECG signal was sampled at 500 Hz and then downsampled at 250 Hz. The original ECG signals were measured at a maximum of 10 mV. Then, they were scaled down by a factor of 1000; hence, the data are represented in uV. In the EPEL dataset, exercise intensities are divided into 5 segments based on the ventilation threshold and VO2max, including before and after the so-called second ventilatory threshold or VT2, before and in the middle of VO2max, and during the recovery after exhaustion shown in Figure 1. Thus, since the EPFL database is R-labeled and exercise fatigue-labeled, the EPFL database is applied in the R-peak detection and ECG classification estimation during exercise.

3. Proposed R-Peak Detection Algorithm and Multi-Feature Fusion Classification

In this paper, a novel multi-feature fusion-based ECG classification framework is developed (Figure 2), which aims to construct accurate and comprehensive ECG features by combining more accurate HRV features, more suitable descriptor features, and more subtle deep learning features.

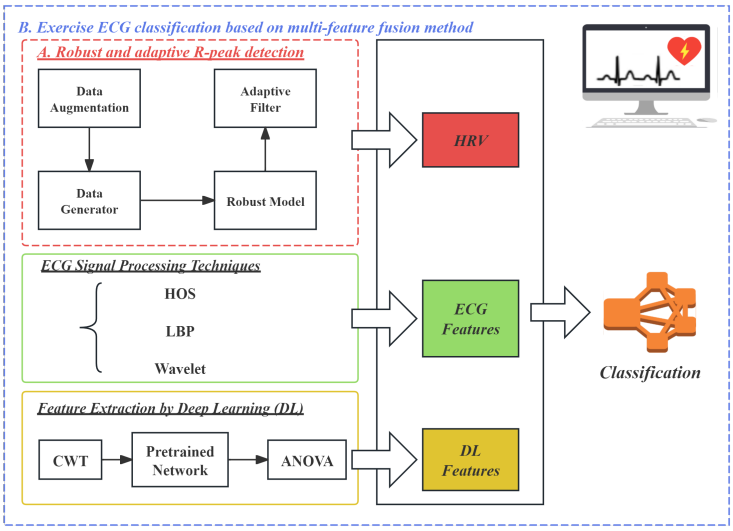


Figure 2. Overview of exercise ECG classification based on multi-feature fusion method.

Specifically, the more accurate HRV features are derived from our R-peak detection model BILSTM-CNN (Figure 3), which is based on BILSTM, multi-scale convolution, and attention mechanism. BILSTM networks and 1D multi-scale convolution enhance R-peak detection by capturing long-term dependencies and analyzing features across multiple scales, respectively. Additionally, CNN is equipped with an attention mechanism that allows for dynamically adjusting to different segments and flexibility, capturing key features in the sequence. To enhance the robustness of the proposed BILSTM-CNN, an ECG data generator as preprocessing is developed to augment training data. In this process, intra-class segmentation and recombination are employed to augment the training data. To further improve the accuracy of R-peak correction and avoid false detection of R-peaks, we propose an adaptive filter as post-processing of the proposed BILSTM-CNN, which accommodates the R-peak detection during high-intensity exercise. Then,

different descriptor features with physiological meanings (such as wavelets and local binary patterns) are utilized, and significant differences are verified by experiments in ECG signals of different exercise intensities. These features can further extract more rich information from ECG signals. Finally, optimally selected deep learning features based on Inception V3 networks and ANOVA analysis are extracted from the CWT of exercise ECG signals.

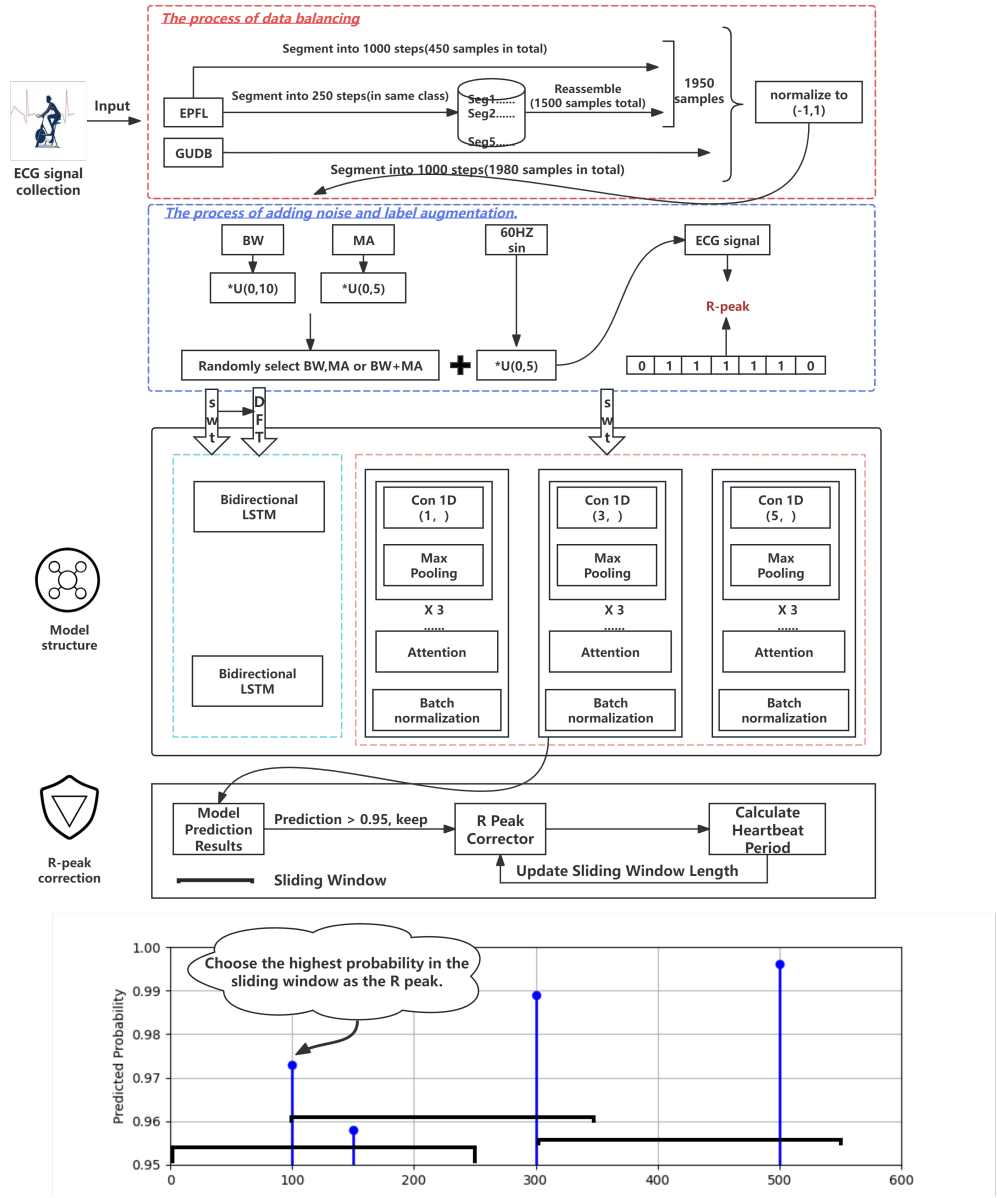


Figure 3. The proposed robust and adaptive R-peak detection. BW = baseline wander; MA = muscle artifact. U indicates uniform distribution where a random multiplier is drawn. SWT represents wavelet transform and DFT represents Discrete Fourier Transform. The sliding window represents the set cardiac cycle. *: multiplication.

3.1. Robust and Adaptive R-Peak Detection

This section mainly introduces the proposed R-peak detector BILSTM-CNN designed for exercise ECG signals. As shown in Figure 3, our R-peak detector primarily consists of three parts: hybrid time–frequency inputs (SWT represents the time-domain representation of the ECG signals whose R-peak morphology is enhanced by means of wavelet transform; in contrast, DFT represents the conversion that turns time-domain signals into frequency-domain signals), the BILSTM-CNN model, and post-processing.

3.1.1. Signal Preprocessing for R-Peak Detection

As we illustrate in Section 2, the GUDB and EPFL datasets are applied in R-peak detection for exercise ECG signals. Each database is resampled at 250 Hz to ensure uniformity across two databases. To ensure the objectivity of the experiment and prevent data leakage, the datasets are split based on subject ID, shown in Table 1. All ECG signals are segmented into 1000-sample intervals, generating 1980 samples for the GUDB database and 450 samples for the EPFL database in the training phase.

Table 1. Train and Test Data Split.

Dataset	Train	Test
EPFL (subject ID)	1–18	19–20
GUDB (subject ID)	1–22	23–24

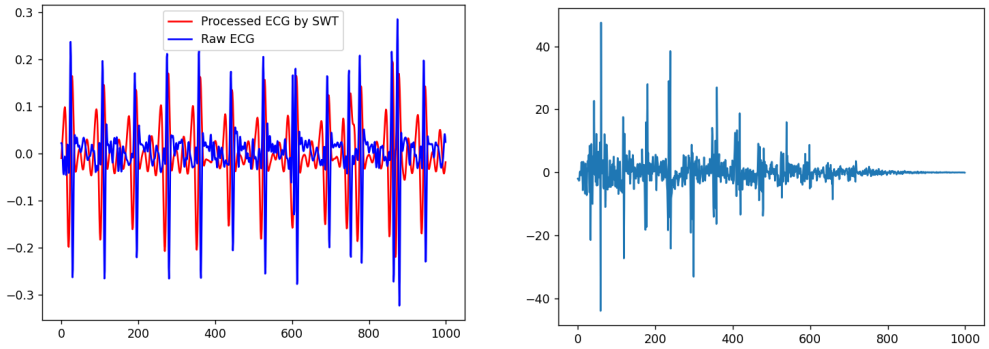
To balance the sample sizes of the training data, the samples of EPFL dateset should be increased. The data of the five different fatigue categories in the EPFL are all divided by 250 steps, and then any four 250-step-long segments within the same category are recombined; each category generates 300 samples, as shown in Figure 3. Consequently, the total number of samples of the EPFL dateset for training is 1950 ($450 + 300 \times 5$). This method is called intra-class segmentation. All samples are then normalized to fall within the range of $[-1, 1]$.

During the exercise process, an amount of noise is induced, such as baseline wander and muscle artifact. To enhance the robustness of the R-peak detector for exercise ECG signals, a data generator induced by exercise ECG signals is developed. The ECG generator is utilized to generate both training data and testing data. The generation of the training set is aimed at enhancing the model's robustness to noise, while the generation of the testing set is intended to further evaluate the model's robustness to noise. The generator mixes ECG signals used in the train and test phases with baseline wander and muscle artifact, both of which are derived from the MBNSTD(MIT-BIH Noise Stress Test database) [21]. Besides, for the power-line interference noise sources, we simulate them using a 60 Hz sine wave. In addition, to ensure the objectivity of the experiments and prevent data leakage, the noise data from the MBNSTD used for training and testing are also split. Specifically, the noises in training are randomly chosen from the range of 1–7,000,000 samples, and the noises in testing are chosen from 7,000,000–9,000,000 samples in the MBNSTD. Finally, since the obvious nonstationarity of the ECG signals, improved binary labeling is employed to label each time step within the window, which addresses the issue of the highly imbalanced R-class samples distribution in training [16]. The R-peak occurrence will be considered around ±2 samples from the original R-peak label. The schematic diagram can be found in the lower right corner of the input part in Figure 3.

Though the constructed ECG data generator is inspired by [16], the differences are obvious below. Firstly, our data generator is developed based on the publicly available exercise ECG signals. To address the scarcity of publicly available exercise ECG data and the lack of labels, the intra-class segmentation in the EPFL dataset is proposed to increase the samples. Secondly, to enhance the robustness of the R-peak model against noise, the model is trained by generating noisy data. Moreover, to ensure the objectivity of the experiments, the noise data from the MIT-BIH Noise Stress Test database and Gaussian noise are added for testing data, rather than only the Gaussian noise considered in [16]. Besides that, unfiltered real ECG data are also used for testing. The corresponding experiment analysis is indicated in Section 4.

3.1.2. Hybrid Time–Frequency Inputs

The inputs of the proposed neural network directly affect the R-peak detection performance. In this case, the hybrid time–frequency signals are constructed as inputs. First, the signals generated by the proposed data generator are enhanced by SWT, which amplifies the most dominant peaks. Then, DFT is applied to deal with the signals transformed by SWT. The combination of DFT and SWT establishes an effective time–frequency analysis technique. SWT allows one to analyze the signal at different scales (or resolutions), which captures the features of signals at various scales. DFT can further analyze the frequency content of the resulting coefficients of SWT Figure 4.



(**a**) (**b**)

Figure 4. (**a**) SWT enhances the local features at the positions of R-peaks; (**b**) DFT extracts the frequency-domain representation of the signal.

Thus, the hybrid time–frequency inputs provide a comprehensive analysis of exercise ECG signals in both the time and frequency domains, which is essential for feature extraction.

(a) R-peak Enhancement by SWT:

As a peak enhancement method, SWT is applied to filter out noise and further enhance the morphology of R-peaks. Since the energy of ECG signals typically concentrates within the frequency range of 3 to 40 Hz [22], a four-level decomposition with the Symlet 4 wavelet is applied (Figure 5).

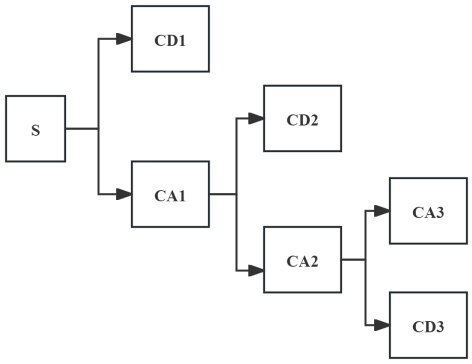


Figure 5. Tree-structured decomposition process of SWT: “S” represents the signal itself, “CD” represents the detail coefficients, and “CA” represents the approximate coefficients.

(b) Frequency-Domain Analysis by DFT:

For ECG signals, the R-peak typically manifests as a spike within a specific frequency range and frequency analysis can accurately capture the features of R-peak. However, the original ECG signals possess complexity, containing numerous frequency components and various types of noise interference at the same time. These noises and redundant frequency components would mask the characteristics of the R-peak, making it extremely difficult to accurately locate the R-peak directly from the original signals and thus having a negative impact on the detection accuracy. Fortunately, the DFT can precisely solve this problem. It is capable of transforming the ECG signals in the time domain into those in the frequency domain. In this way, the complicated frequency components in the signals become distinguishable. This enables the R-peak to be clearly differentiated from other waveform interference factors, such as P-waves, T-waves, and various artifacts.

3.1.3. BILSTM-CNN Model

To enable the R-peak detection model to capture long-term dependencies features and local features simultaneously, the BILSTM-CNN model is proposed by combining bidirectional LSTM with multi-scale 1D convolution. The specific details are as follows.

(a) Bidirectional LSTM:

For one-dimensional ECG signals, R-peak detection is a time-sequential problem, and the position of the R-peak is influenced by the preceding and succeeding signals [23–25]. As is well known, LSTM is well suited for handling time-sequential problems. Nevertheless, LSTM only considers information before the current moment and fails to fully utilize subsequent information, as shown in the following:

$$C_t = f_t \odot C_{t-1} + i_t \odot g_t, \tag{1}$$

where C_{t-1} and g_t are the information state of the previous moment and the information state of the current input, respectively. And f_t and i_t represent the forget gate and input gate parameters for information retention, respectively.

Compared with LSTM, BILSTM [23] considers both past and future information at the current moment, thereby achieving a more comprehensive understanding and analysis of the temporal features of ECG signals. The formulation of BILSTM is as follows:

$$\begin{aligned} C_t^{(f)} &= f_t \odot C_{t+1} + i_t \odot g_t \\ h_t &= C_t \oplus C_t^{(f)} \end{aligned}, \tag{2}$$

where C_{t+1} is the information state of the next moment, $\oplus$ denotes the concatenation operation, and h_t is the combination of the preceding and following information.

(b) Attention Mechanism-based Multi-Scale Convolution:

When dealing with complex signals, single-scale convolution has limitations in capturing features due to its fixed kernel size, resulting in poorer performance [26]. Here, multi-scale convolution is adopted to capture local features at different scales in R-peak detection, and the sizes of the 1D convolutional kernels are 1, 3, and 5. Moreover, the attention mechanism is integrated into CNN to adjust dynamically different parts and capture the key features in the sequence flexibly [27], improving the accuracy of features and enhancing the robustness of the proposed BILSTM-CNN.

The proposed BILSTM-CNN model includes 204,529 parameters, of which 204,433 parameters are trainable. Input and output shapes are defined as (batch size, number of time samples, features), with time samples fixed at 1000 and features fixed at 1. Binary

cross-entropy and Adam are selected as the loss function and optimizer. To achieve better training performance, the learning rate is variable, determined by the following

$$\text{Learning_rate} = \frac{\text{Initial_learning_rate}}{1 + \text{decay_rate} \times \text{epoch}}, \tag{3}$$

where Initial_learning_rate is 0.001 and decay_rate is 0.9.

3.1.4. Post-Processing

Various noise occurs in exercise ECG signals, and then some peaks are incorrectly detected as R-peaks. It is well known that multiple R-peaks cannot occur within a single cardiac cycle. To filter out these false peaks, the R-peak with the highest probability remains the true one when multiple R-peaks are detected in one cardiac cycle. In general, the cardiac cycle used in the R-peak filter is a fixed value. However, the cardiac cycle of exercise ECG signals shows significant variations during high-intensity exercise. Hence, we propose an adaptive rule to update the cardiac cycle of the exercise ECG signals with the sliding window shown in the R-peak correction part in Figure 3. The results shown in Figure 6 validate the effectiveness of the adaptive filtering with the proposed adaptive cardiac cycle as a post-processing step for R-peak detection. In Figure 6a, the red outline indicates the false R-peak. After using adaptive filtering, the erroneous R-peak is filtered out in Figure 6b. Specifically, the update process of the size of the sliding window is provided as follows (Algorithm 1):

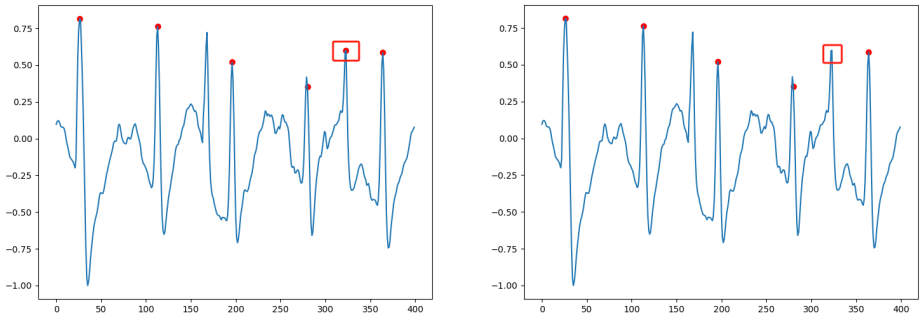
Algorithm 1: Update length of the sliding window

1. Initialization: $k = 1$, $L_{(1)} = 250$, corresponding to the normal heartbeat cycle step (for normal adults, the cardiac cycle is around 0.8–1 s, and we choose 1 s as the benchmark; at a sampling frequency of 250 Hz, it corresponds to 250 in length. Therefore, the initial size of the sliding window is set to 250). To adapt to the changes in the cardiac cycle during high-intensity exercise, this $L_{(1)}$ will be updated once every 4 s.
2. Repeat
   - (1) Within a window of 1000 samples, all R-peak suspicious points $R_1, R_2, \cdots, R_l$ are detected by the proposed prediction model BILSTM-CNN. Here, we select 0.95 as the user-defined probability threshold.
   - (2) When multiple R-peaks are detected within the sliding window with $L_{(k)}$, the R-peak with the highest probability is selected as the true R-peak named as R_1^*.
   - (3) The next sliding window starts from the detected R-peak R_1^*, and the size of the sliding window remains $L_{(k)}$.
   - (4) Repeat Step (2) and Step (3) until all suspicious R-peaks within the first 1000 steps are filtered out and the true R-peaks $R_1^*, R_2^*, \cdots, R_t^* (t \leq l)$ are obtained.
   - (5) Update the size of the sliding window by

$$L_{(k+1)} = \frac{1}{t-1} \sum_{i=1}^{t-1} (R_i^* - R_{i-1}^*), \tag{4}$$

     where R_i^* and R_{i-1}^* represent the positions of the preceding and succeeding R-peaks respectively, and t represents the number of true R-peaks occurring in 1000 steps.
   - (6) Set $k \longleftarrow k + 1$

   End Do.



(**a**) (**b**)

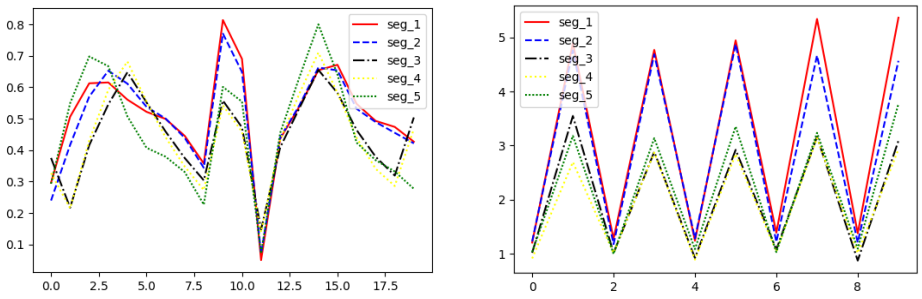
Figure 6. The comparison results of R-peak detection without the proposed adaptive filter and one with the proposed adaptive filter: (**a**) due to noise and high-intensity movement, false detections of R-peaks occur; (**b**) the falsely detected R-peak is filtered out by the proposed adaptive filter.

In summary, based on the characteristics of exercise ECG signals, we propose an adaptive and robust R-peak detection method BILSTM-CNN. The effectiveness and robustness of the proposed R-peak detection have been verified by experiments in Sections 4.2.1 and 4.3. Accurate R-peak locations are the foundation for accurate HRV features, which are important features for the subsequent exercise ECG classification.

3.2. Exercise ECG Classification Based on Multi-Feature Fusion Method

3.2.1. Feature Selection

Exercise ECG classification is of great significance for exercise fatigue detection. The precise and comprehensive features guarantee the accurate results of ECG classification. HRV features, known as RR interval variability, are important indicators for exercise ECG classification. Based on our proposed R-peak detection method, the more accurate HRV features in Table A1 could be extracted. In addition, various features are commonly used for medical ECG classification, such as wavelet features, higher-order statistics (HOS) features, 1-D LBP features, and multiple R-peak-based features. To verify the applicability of these features for exercise ECG classification, each descriptor feature is tested on the classification experiments for five classes of exercise ECG signals in the EPFL dataset. Figure 7a,b demonstrates the feature variability of wavelet features and HOS features across five classes of exercise ECG signals, respectively. Finally, the descriptor features in Table A2 are employed for exercise ECG classification since they show good performance as wavelet features, indicated in Figure 7a.



(**a**) (**b**)

Figure 7. Average beats from the EPFL database are grouped by the five exercise fatigue classes (Seg_1, Seg_2, Seg_3, Seg_4, and Seg_5): (**a**) represents the comparison of the extracted wavelet features under five different fatigue levels, and (**b**) represents the comparison of the HOS features under five different fatigue levels (best seen in color). The abscissa represents the number of features, and the ordinate represents the calculated values.

The above features are predefined and may not fully capture all useful information. In contrast, deep learning features can automatically learn complex feature representations. Therefore, we transformed the original signal into 2D images using CWT. The reason for choosing CWT is that CWT provides both time and frequency resolution of the signal and could capture finer-grained dynamic changes [28]. Compared with Short-Time Fourier Transform (STFT) and Gabor Transform, CWT shows greater adaptability to nonstationary signals. In the training phase, to avoid the challenges of training a deep CNN model, transfer learning technology is adopted by using pretrained neural networks trained with ample data and transferring this knowledge to the targeted classification system [29]. How to select an optimal pretrained model? Compared with traditional architectures such as VGGNet [30] and GoogleNet [31], Inception V3 networks could extract multi-scale features by learning complex patterns of the data. Then, a wider range of features at different resolutions can be captured by Inception V3. The input of the model is an ECG time–frequency graph with 3 channels and a size of 224 × 224. The feature extraction is achieved through the Inception modules in the Inception V3 network. Finally, the dense layer with 1024 nodes is added before the fully connected layer of the original Inception model. From the pretrained Inception V3 architecture, 1024 features are extracted, and ANOVA analysis [32] is employed to select the optimal features.

3.2.2. SVM Classifier

Support Vector Machine (SVM) is chosen as the primary classifier in this study due to its strong performance in previous ECG classification tasks. SVM maximizes the margin between classes by finding a hyperplane that separates the data points of different classes as cleanly as possible, and they are particularly effective for high-dimensional data. In addition, SVMs are robust to overfitting data and can handle nonlinear decision boundaries due to their regularization parameters and kernel functions. The good generalization ability further contributes to their effectiveness in the classification of nonlinear physiological signals. The results also verify that SVM has the best classification performance among four classifiers: SVM, Random Forest, LDA, and KNN.

4. Experiments and Results

4.1. Experiments and Implementation Details

Our experimental settings are categorized into the following three classes:

- Compare R-peak detection performance for exercise ECG signals between the proposed BILSTM-CNN and existing R-peak detection methods: LSTM [16], Hamilton [33], Christov [34], Engzee [35], and Pan–Tompkins [36]. These experiments are conducted on the EPFL dataset and the GUDB dataset.
- Compare exercise ECG classification performance based on HRV features induced by different R-peak detection methods, including the proposed BILSTM-CNN and the existing R-peak detection methods: LSTM [16], Hamilton [33], Christov [34], Engzee [35], and Pan–Tompkins [36]. These experiments are conducted on EPFL dataset.
- Evaluate exercise ECG classification performance based on the proposed multi-feature fusion method, which combines HRV features extracted from the proposed BILSTM-CNN detection, the verified descriptor features, and optimal deep learning features of exercise ECG signals. These experiments are conducted on the EPFL dataset.

We adopt four metrics to evaluate the experiment performance: precision (PR), recall (RE), accuracy (AC), and F1 score (F1), and the formulas are expressed as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{5}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{6}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{7}$$

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{8}$$

where TP (true positive) denotes the count of samples correctly predicted as positive classes, FN (false negative) represents the count of samples incorrectly predicted as negative classes, FP (false positive) indicates the count of samples incorrectly predicted as positive classes, and TN (true negative) signifies the count of samples correctly predicted as negative classes.

4.2. Results

4.2.1. Comparison Results of R-Peak Detection Performance

To validate the performance of the proposed R-peak detection BILSTM-CNN, we apply different R-peak detection methods on the EPFL dataset and GUDB dataset. Since the provided original EPFL dataset has been filtered, three datasets are constructed for the comprehensive analysis of R-peak detection performance. Specifically, the test datasets based on the EPFL dataset include three cases: the original EPFL dataset, the EPFL dataset with MIT-BIH noise, and the EPFL dataset with Gaussian noise. In addition, the experiments are conducted on unfiltered real data from GUDB, and results on these unprocessed real signals enhance the practical applicability of our proposed R-peak detection method.

Predicted R-peak is considered true positive if it falls within one-tenth of the sampling rate (in this study, 25 samples at 250 Hz sampling rate) from the ground truth. Additionally, to measure its robustness against noise, we add combinations of bandwidth (BW), amplitude modulation (AM), and motion artifact (ME) noise to the test data, as well as Gaussian noise with varying signal-to-noise ratios (SNR).

4.2.2. R-Peak Detection in the EPFL Dataset

Table 2 presents a comparison for five exercise intensities based on Figure 1. Here, Seg1, Seg2, and Seg5 are categorized as low-intensity exercises, and Seg3 and Seg4 are categorized as high-intensity exercises. We observed that our BILSTM-CNN method outperforms classical methods in terms of PR, RE, and F1. Under lower exercise intensities (Seg1, Seg2, and Seg5), our algorithm achieves an F1 value of 99.0%, demonstrating remarkable performance. For high-intensity exercise (Seg3 and Seg4), our algorithm still maintains an F1 value above 98.0%, while the performance of traditional algorithms declined significantly. Overall, the experimental results show that our proposed algorithm is more adaptable to exercise-induced changes in ECG signals. In addition, to demonstrate the importance of adaptive filtering mentioned in Section 3.1.4 for the proposed BILSTM-CNN, we also analyze the R-peak detection performance of the proposed R-peak detection method without adaptive filtering. Experiment results indicate that the adaptive filter further improves the performance of R-peak detection by dynamically updating the cardiac cycle to accommodate high-intensity exercise.

Table 2. Comparison of R-peak detection performance conducted by all R-peak detection algorithms on the EPFL dataset without noise. The bold fonts represent the optimal performance.

Type	Metric	OURS (Adaptive)	OURS	LSTM	Christov	Engzee	Hamilton	Pan–Tompkins
Seg_1	PR	99.9%	99.8%	99.3%	100%	99.1%	99.8%	**100%**
	RE	**99.9%**	99.2%	98.4%	95.7%	98.1%	99.7%	98.7%
	F1	**99.8%**	99.5%	**99.8%**	97.4%	98.6%	99.7%	99.3%
Seg_2	PR	99.9%	99.9%	99.0%	**100%**	98.2%	99.4%	**100%**
	RE	**99.9%**	98.7%	97.8%	97.0%	96.4%	98.7%	98.2%
	F1	**99.9%**	99.3%	98.3%	98.2%	97.2%	99.1%	99.1%
Seg_3	PR	**99.7%**	99.6%	97.5%	98.1%	90.3%	96.4%	99.1%
	RE	**97.2%**	96.7%	96.8%	89.1%	84.2%	94.1%	82.3%
	F1	**98.3%**	98.0%	97.1%	92.4%	86.7%	95.3%	88.8%
Seg_4	PR	**99.8%**	99.5%	97.2%	98.5%	88.6%	97.3%	99.5%
	RE	**98.4%**	96.6%	96.1%	86.7%	82.0%	95.5%	85.8%
	F1	**99.1%**	98.1%	96.6%	90.1%	84.7%	96.4%	91.2%
Seg_5	PR	**100%**	**100%**	99.5%	99.8%	96.4%	99.6%	99.9%
	RE	**99.9%**	99.4%	99.2%	96.1%	93.6%	98.5%	98.2%
	F1	**99.9%**	99.6%	99.3%	97.6%	94.8%	99.0%	99.0%

4.2.3. R-Peak Detection in the EPFL Dataset with MIT-BIH Noise

To verify the robustness of the R-peak detector to noise, noises from the MIT-BIH Noise Stress Test database are added to the ECG signal to simulate the real-world conditions mentioned in Section 3.1.1. As shown in Table 3, the R-peak detection performance of our proposed method achieves F1 values over 97.3% for five different exercise intensities. Compared with the noise-free results in Table 2, the performance of all methods is degraded, while our method has the least performance degradation, which demonstrates the robustness of the proposed method on the EPFL dataset with MIT-BIH noise. Similarly, the effectiveness of R-peak post-processing is verified by testing the proposed method without dynamic post-processing.

Table 3. Comparison of R-peak detection performance conducted by all R-peak detection algorithms on the EPFL dataset with the MIT-BIH Noise Stress Test database. The bold fonts represent the optimal performance.

Type	Metric	OURS (Adaptive)	OURS	LSTM	Christov	Engzee	Hamilton	Pan–Tompkins
Seg_1	PR	**99.7%**	**99.7%**	99.1%	98.8%	95.0%	91.6%	**99.7%**
	RE	**98.5%**	98.2%	98.0%	91.8%	89.4%	91.0%	91.7%
	F1	**99.0%**	**99.0%**	98.5%	94.9%	91.9%	91.3%	91.7%
Seg_2	PR	**100%**	99.9%	98.4%	98.5%	95.4%	94.8%	**100%**
	RE	**99.4%**	98.5%	96.2%	91.8%	90.4%	93.6%	84.7%
	F1	**99.7%**	99.1%	97.2%	94.9%	92.8%	94.2%	91.1%
Seg_3	PR	**99.4%**	99.3%	97.2%	94.9%	83.9%	88.6%	98.2%
	RE	**95.3%**	94.8%	94.4%	78.6%	72.9%	85.7%	63.5%
	F1	**97.3%**	96.9%	95.7%	85.1%	77.5%	87.1%	74.9%
Seg_4	PR	**99.6%**	**99.6%**	96.8%	94.4%	80.5%	86.8%	98.3%
	RE	**95.5%**	94.6%	93.2%	75.9%	68.2%	83.6%	65.4%
	F1	**97.5%**	97.0%	94.9%	82.0%	73.2%	85.1%	76.1%

Table 3. *Cont.*

Type	Metric	OURS (Adaptive)	OURS	LSTM	Christov	Engzee	Hamilton	Pan–Tompkins
Seg_5	PR	**100%**	**100%**	99.3%	98.6%	84.5%	91.5%	99.7%
	RE	**99.1%**	98.7%	98.7%	91.7%	76.1%	90.2%	92.3%
	F1	**99.5%**	99.3%	98.9%	94.8%	79.8%	90.9%	95.7%

4.2.4. R-Peak Detection in the EPFL Dataset with Gaussian Noise

To further evaluate the performance of our method at different noise levels, we generate test datasets by adding Gaussian noise with different signal-to-noise ratios (SNR) to the original EPFL dataset. Here, SNRs are set to be 20, 10, 5, 1, 0.5, 0.4, 0.3, 0.2, and 0.1. Subsequently, R-peaks of the generated noisy ECG signals are detected by all methods, and the results are shown in Figure 8. It is evident that our R-peak detection method BILSTM-CNN outperforms other R-peak methods at all noise levels. In particular, under high noise conditions, the performance of traditional methods decreases significantly, while the performance of our R-peak detection method decreases slightly. The experiment results further illustrate the robustness of the proposed BILSTM-CNN.

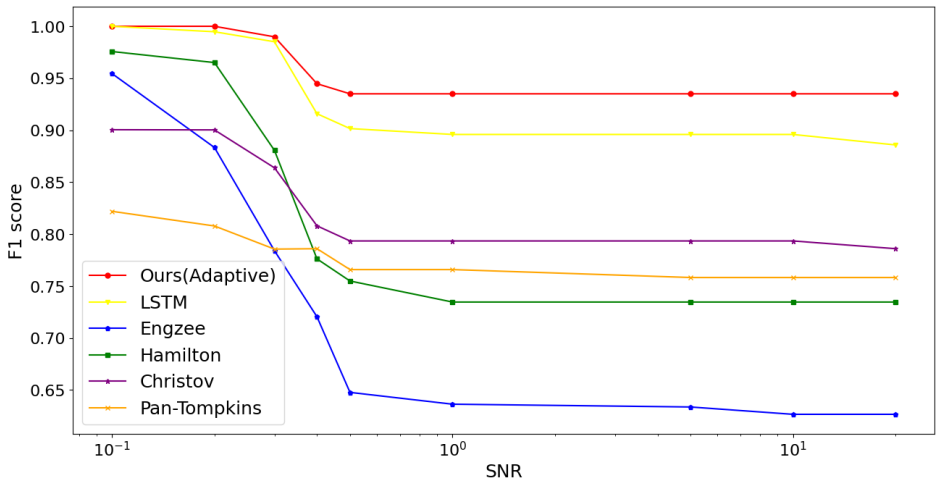


Figure 8. Comparison of the F1 score conducted by all R-peak detection algorithms on the EPFL dataset with various Gaussian noise (different SNRs). The results indicate that our R-peak detection method achieves the best robustness.

4.2.5. R-Peak Detection in the Real Data from GUDB

In this part, unfiltered real exercise ECG signals from GUDB are chosen as the test data to verify the effectiveness of our R-peak detection method, BILSTM-CNN. As shown in Table 4, the proposed R-peak detection BILSTM-CNN outperforms traditional algorithms under various motion conditions. For low exercise intensities (hand cycling and walking), the R-peak detection performance obtained by our algorithm achieves an F1 score of nearly 1, which indicates that our R-peak detection algorithm can be effectively applied to the R-peak detection of real exercise ECG signals.

Table 4. Comparison of R-peak detection performance conducted by all R-peak detection algorithms on the GUDB dataset. The bold fonts represent the optimal performance.

Type	Metric	OURS (Adaptive)	OURS	LSTM	Christov	Engzee	Hamilton	Pan–Tompkins
hand_bike	PR	**99.7%**	99.4%	98.3%	97.1%	99.2%	99.7%	99.4%
	RE	**100%**	**100%**	98.2 %	96.8%	99.4 %	99.4%	99.1%
	F1	**99.8%**	99.7%	98.1%	96.9%	99.2%	99.5%	99.2%

Table 4. *Cont.*

Type	Metric	OURS (Adaptive)	OURS	LSTM	Christov	Engzee	Hamilton	Pan–Tompkins
jogging	PR	**98.0%**	97.0%	96.2%	88.9%	97.3%	82.1%	92.1%
	RE	**99.7%**	99.3%	95.3%	84.9%	92.4%	75.4%	83.5%
	F1	**98.8%**	98.4%	95.7%	86.8%	94.7%	78.6%	87.6%
walking	PR	99.8%	99.8%	98.6%	97%	**100%**	**100%**	97.5%
	RE	**99.9%**	99.6%	98.6%	96.8%	99.5%	99.4%	98.5%
	F1	**99.8%**	99.6%	98.6%	96.8%	99.7%	99.6%	98.0%

4.3. Exercise ECG Classification Based on HRV Features

HRV features are crucial for exercise ECG classification, which can be used to analyze fatigue levels. Since HRV features are derived from the R-peak position, the accurate detection of the R-peak position is the basis of HRV features. Conversely, the higher the classification accuracy of ECG signals based on HRV features, the better the performance of the corresponding R-peak detection method To verify the effectiveness of the proposed R-peak detection method BILSTM-CNN, we compared the performance of ECG classification only based on HRV features induced by different R-peak detection methods. As shown in Table 5, the AC, RE, and F1 values based on the proposed R-peak detection algorithm are 63.3%, 62.5%, and 62.9%, respectively, all of which outperform other methods based on the classical R-peak algorithms. Namely, the experimental results demonstrate that our proposed BILSTM-CNN can achieve more accurate R-peak detection.

Table 5. The exercise ECG classification performance(average five class) based on HRV features induced by different R-peak detection methods. The bold fonts represent the optimal performance.

	BILSTM-CNN	LSTM	Hamilton	Christov	Engzee	Pan–Tompkins
AC%	**63.3%**	62.0%	58.0%	56.2%	54.0%	57.2%
RE%	**62.5%**	62.1%	58.3%	56.1%	53.8%	57.1%
F1%	**62.9%**	61.8%	58.1%	56.2%	53.9%	57.1%

4.4. Exercise ECG Classification Based on Multi-Feature Fusion

Here, we combine HRV features, descriptor features, and deep learning features to construct more accurate and comprehensive ECG features, and SVM is selected as a classifier to achieve better classification performance of exercise ECG signals. In terms of the constructed features, the effectiveness of three types of features has been validated and analyzed. Specifically, HRV features derived from the proposed R-peak detection method are more accurate, which is validated in Section 4.3. Descriptor features are selected by analyzing the significant difference in these features in different classes of exercise ECG signals, and the effectiveness of such features for exercise ECG classification is demonstrated in Section 3.2.1 and Figure 7. For deep learning features, Inception V3 is selected to extract the more refined features. The experiments are conducted by comparing five pretrained neural networks under four classification algorithms: SVM, Random Forest, LDA, and KNN, and the results are shown in Table 6. Here, 1024 features of the pretrained Inception V3 are extracted, and ANOVA analysis [32] is employed to select the optimal features. The total number of selected optimal deep learning features is 100, with 924 features discarded. The experiment results in Table 6 indicate that the Inception V3 network shows superior performance in feature extraction, and the SVM presents the best classification performance.

Then, to demonstrate the effectiveness of feature fusion for exercise ECG classification, we compare the classification results based on HRV features, HRV features fused with ECG descriptor features, and all features in Figures A1–A3. As shown in Table 7, the proposed

multi-feature fusion method greatly enhances the accuracy of fatigue detection, achieving AC, RE, and F1 of 99.1%.

Table 6. Comparison of exercise ECG signal classification performance based on different pretrained neural networks and different classifiers. The bold fonts represent the optimal performance.

Model	Classifier	Performance Metrics		
		Accuracy	Recall	F1 Score
Inception V3	SVM	**95.5%**	**95.5%**	**95.4%**
	Random Forest	91.5%	91.5%	91.0%
	KNN	89.7%	89.6%	88.9%
	LDA	79.2%	79.2%	73.7%
VGG16	SVM	76.0%	80.5%	73.2%
	Random Forest	76.0%	80.5%	73.2%
	KNN	76.0%	80.5%	73.2%
	LDA	75.3%	80.0%	72.0%
VGG19	SVM	76.7%	80.9%	75.1%
	Random Forest	77.3%	81.6%	75.4%
	KNN	76.0%	80.5%	73.2%
	LDA	75.3%	80.0%	72.0%
ResNet50	SVM	81.3%	84.7%	81.5%
	Random Forest	83.3%	86.5%	83.6%
	KNN	85.3%	84.3%	82.5%
	LDA	81.3%	80.0%	74.5%
ResNet101	SVM	78.7%	82.5%	78.0%
	Random Forest	80.7%	84.2%	80.7%
	KNN	84.0%	83.0%	80.9%
	LDA	81.3%	80.0%	74.5%

Table 7. The impact of feature fusion on sports fatigue detection. The bold fonts represent the optimal performance.

Features	Accuracy (%)	Recall (%)	F1 Score (%)
HRV features	63.3%	62.5%	62.9%
HRV and ECG features	81.3%	81.4%	81.6%
All features	**99.1%**	**99.1%**	**99.1%**

Additionally, we also compare classification results obtained by our proposed algorithm with existing feature fusion methods, which are tested on ECG datasets. Although the ECG datasets are used in different applications, the exercise ECG dataset for fatigue detection in our method is more variable and noisy than other ECG datasets used in different applications. In addition, the classification based on motion ECG signals tested in this paper is a five-classification problem, while other methods are mostly binary classification and three-classification problems. In this case, our proposed multi-feature fusion approach still achieves high-quality performance, as shown in Table 8, which further demonstrates that the proposed fusion features can better capture more comprehensive features of ECG signals.

Table 8. Comparison of ECG classification performance based on different feature fusion methods.

Author	Dataset	Feature Extraction	Feature Extraction Network	AC (%)
Ahmed S. Eltrass et al. [11]	MIT-BIH ARR & NSR BIDMC	HRV, CQ-NSGT	AlexNet	98.74%
Chen et al. [9]	EPHNOGRAM [37]	HRV, STFT	VGG16	96.90%

Table 8. *Cont.*

Author	Dataset	Feature Extraction	Feature Extraction Network	AC (%)
Cheol et al. [10]	WESAD [38]	HRV, STFT	LSTM and Xception	99.51%
Alan Jovic [39]	PhysioBank	HRV, ECG	None	99.7%
Proposed	EPFL	HRV, CWT, ECG features	Inception V3	99.1%

5. Discussions and Future Work

In this study, we explore a multi-feature fusion method to improve the exercise ECG classification for fatigue monitoring. Here, three contributions guarantee the effectiveness of the proposed method. First, an adaptive and robust R-peak detection model has been proposed, which achieves more accurate R-peak detection, and more precise HRV features are captured. The proposed R-peak detection method includes the ECG data generator, BILSTM-CNN model, and adaptive post-processing. The ECG data generator ensures robustness; the constructed BILSTM-CNN could capture multi-scale features of exercise ECG signals, and post-processing guarantees the correction of misjudged R-peaks. Second, descriptor features of ECG signals have been verified the effectiveness for classifying exercise ECG signals. Finally, multi-feature fusion method are presented by combining HRV features derived from the proposed R-peak detection model, descriptor features, and deep learning features. Experiment results show that our R-peak detection approach achieves better performance on real exercise ECG data and ECG signals with added noise (MIT-BIH noise and Gaussian noise). In terms of classification performance, extensive experiments validate the superiority of our multi-feature fusion approach over state-of-the-art classification algorithms. While we have made significant progress, there are still potential research directions worth exploring further. Future studies could focus on improving the accuracy and robustness of R-peak detection, as well as exploring more complex deep learning models and feature fusion methods to enhance exercise intensity detection performance.

Author Contributions: Conceptualization, X.S. and X.W.; methodology, X.S. and X.W.; software, X.W.; validation, X.S. and X.W.; formal analysis, X.W.; investigation, X.S.; resources, X.S. and H.G.; data curation, X.S. and X.W.; writing—original draft preparation, X.W.; writing—review and editing, X.S. and X.W.; visualization, X.W.; supervision, X.S.; project administration, X.S.; funding acquisition, X.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China under Grant 62301056, the Fundamental Research Funds for Central Universities No. 2024JCYJ005, and the National Natural Science Foundation of China under Grant 12371094.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

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

Appendix A

Table A1. Feature Types and Descriptions with Calculations.

Feature Type	Description
Time-domain feature	
SDSD	Standard deviation of difference between RR intervals. $$SDSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(\Delta RR_i - \overline{\Delta RR})^2}$$
SDNN	Standard deviation of NN (RR) intervals. $$SDNN = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(RR_i - \overline{RR})^2}$$
RMSSD	Root mean square of successive differences. $$RMSSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(\Delta RR_i)^2}$$
Mean (HRV-NNX)	Mean value of RR intervals. $$Mean(HRV-NNX) = \frac{1}{N}\sum_{i=1}^{N} RR_i$$
PNN50	Proportion of NN50 divided by total NN (RR) intervals. $$PNN50 = \frac{NN50}{N}$$ NN50: Number of pairs of successive NN intervals that differ by more than 50 milliseconds.
Frequency-domain feature	
Mean (PLF)	Average percentage of low-frequency components. $$Mean(PLF) = \frac{1}{n}\sum_{i=1}^{n} PLF_i$$
Mean (PHF)	Average percentage of high-frequency components. $$Mean(PHF) = \frac{1}{n}\sum_{i=1}^{n} PHF_i$$
Mean (LFHF ratio)	Average LF/HF ratio. $$Mean(LFHFratio) = \frac{1}{n}\sum_{i=1}^{n} LFHF_i$$

Table A1. *Cont.*

Feature Type	Description
Mean (SD1)	Average short axis (SD1) in the Poincaré plot.
Mean (SD2)	Average long axis (SD2) in the Poincaré plot.
Mean (SD1 SD2 ratio)	Average of SD1 and SD2 ratios.

Table A2. Descriptor Features Used for Exercise ECG Classification.

Features	Description
Wavelets features	Time- and frequency-domain extraction using Daubechies wavelet (db1) for a 20-dimensional descriptor [40,41].
HOS	Morphological description via kurtosis and skewness over 5 intervals per beat [42,43] for a 10-dimensional descriptor.
LBP	Feature extraction from signals using uniform LBP encoding, yielding a 59-dimensional descriptor [44].
Morphological Descriptor	Euclidean distance from the R-peak to four key points in the beat for a 4-dimensional descriptor [45,46].
R-R Interval	Key descriptor in ECG classification, including Pre-RR, Post-RR, Local-RR, and Global-RR for a 4-dimensional descriptor [47].
R-Peak Amplitude	Intensity metrics of cardiac activity: Maximum, Minimum, Median, and Mean for a 4-dimensional descriptor [48].

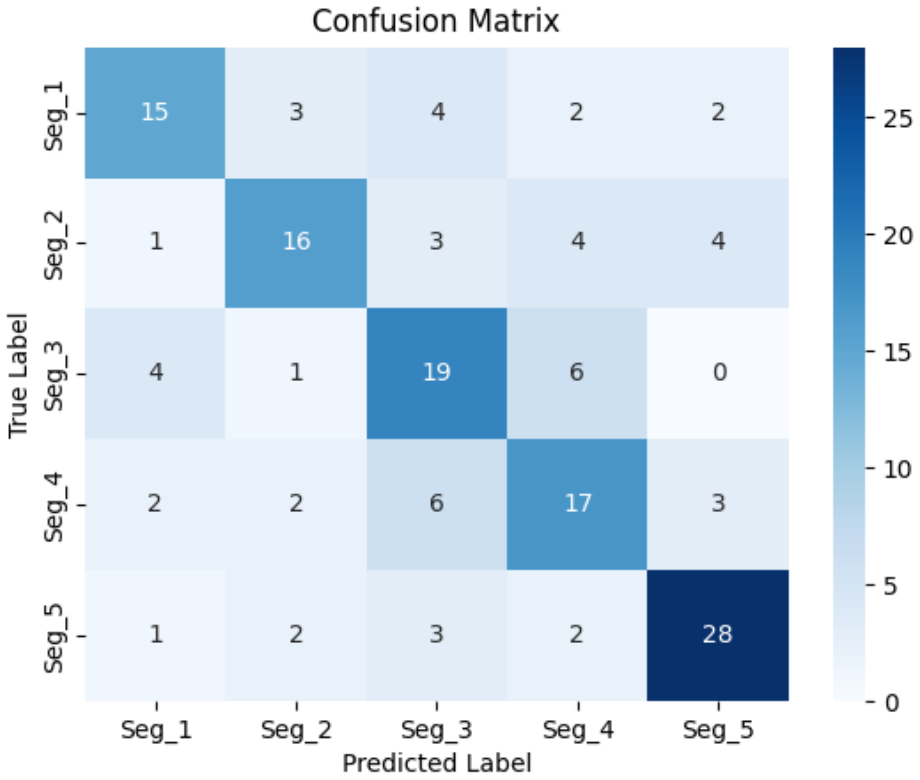


Figure A1. The confusion matrix of using HRV features to classify.

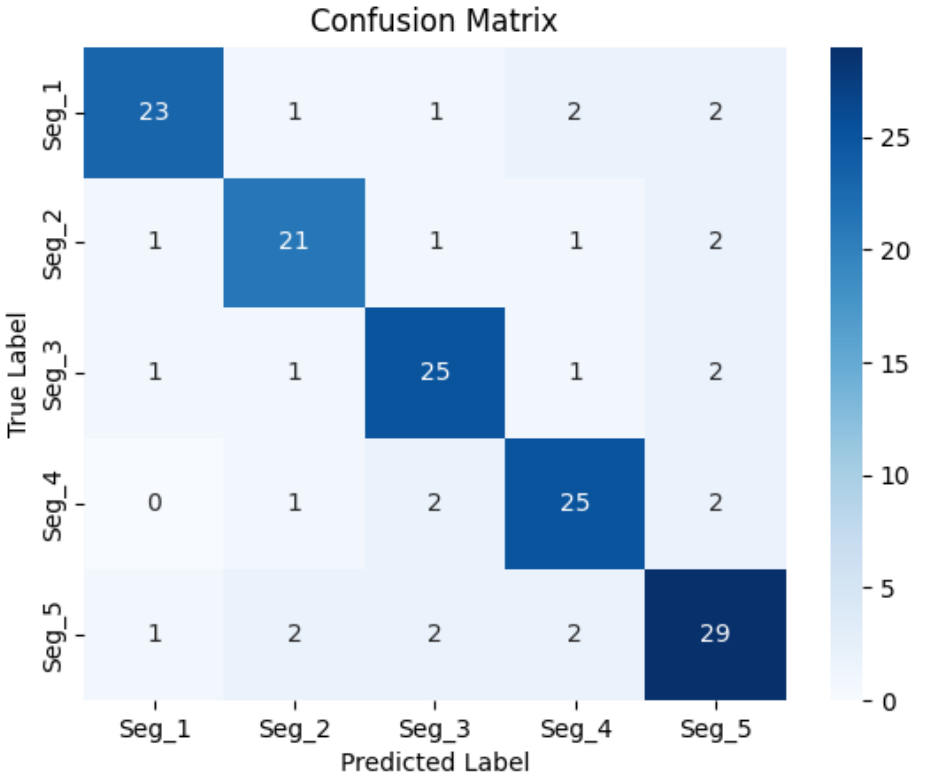


Figure A2. The confusion matrix of using HRV and ECG features to classify.

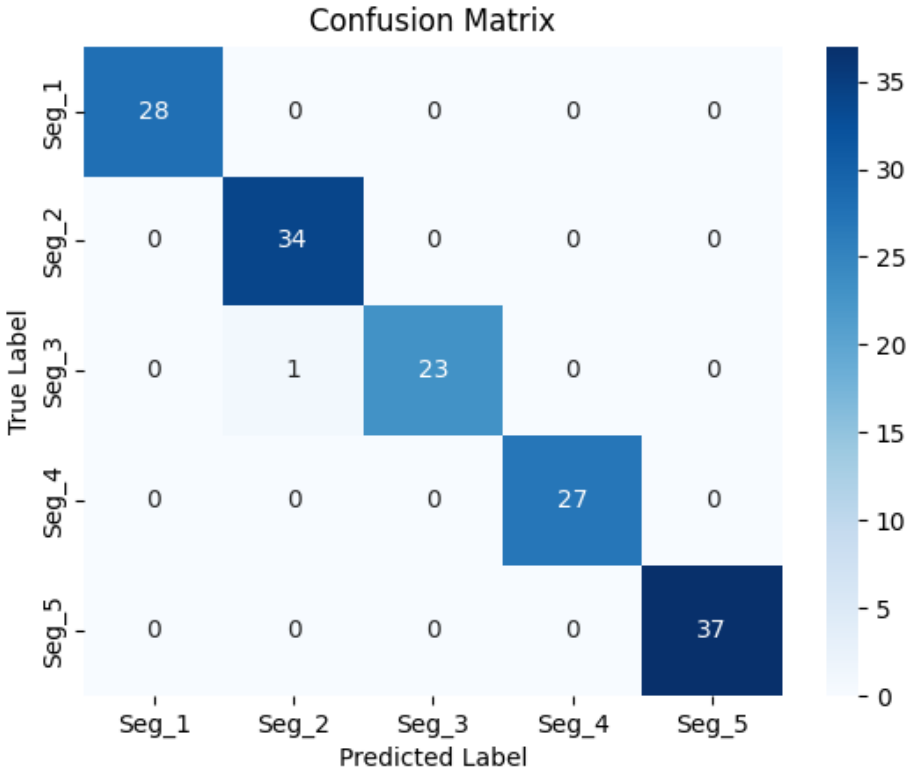


Figure A3. The confusion matrix of using all features to classify.

References

1. Finocchiaro, G.; Papadakis, M.; Robertus, J.L.; Dhutia, H. Etiology of sudden death in sports: Insights from a United Kingdom regional registry. *J. Am. Coll. Cardiol.* **2016**, *67*, 2108–2115.
2. Casa, D.J.; Guskiewicz, K.M.; Anderson, S.A.; Courson, R.W.; Heck, J.F. National Athletic Trainers' Association position statement: preventing sudden death in sports. *J. Athl. Train.* 2012, *47*, 96–118. [CrossRef] [PubMed]
3. Marijon, E.; Tafflet, M.; Celermajer, D.S.; Dumas, F.; Perier, M.C. Sports-related sudden death in the general population. *Circulation* 2011, *124*, 672–681. [CrossRef] [PubMed]
4. Sivanantham, A.; Devi, S.S. Cardiac arrhythmia detection using linear and non-linear features of HRV signal. In Proceedings of the 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies, Ramanathapuram, India, 8–10 May 2014; pp. 795–799.
5. Vicente, J.; Laguna, P.; Bartra, A.; Bailón, R. Detection of driver's drowsiness by means of HRV analysis. In Proceedings of the 2011 Computing in Cardiology, Hangzhou, China, 18–21 September 2011; pp. 89–92.
6. Xiao, M.; Yan, H.; Song, J.; Yang, Y.; Yang, X. Sleep stages classification based on heart rate variability and random forest. *Biomed. Signal Process. Control* **2013**, *8*, 624–633. [CrossRef]
7. Radha, M.; Fonseca, P.; Moreau, A.; Ross, M.; Cerny, A.; Anderer, P.; Long, X.; Aarts, R.M. Sleep stage classification from heart-rate variability using long short-term memory neural networks. *Sci. Rep.* **2019**, *9*, 14149. [CrossRef] [PubMed]
8. Vigier, M.; Vigier, B.; Andritsch, E.; Schwerdtfeger, A.R. Cancer classification using machine learning and HRV analysis: Preliminary evidence from a pilot study. *Sci. Rep.* **2021**, *11*, 22292. [CrossRef]
9. Chen, Y.; Ma, X.; Su, X.; Ge, H. Classification of Exercise Fatigue by Fusion of Image Features and Linear Features of ECG. In *Abstracts of the 13th National Sports Science Conference—Special Reports (Sports Engineering Branch)*; Beijing Sport University: Beijing, China, 2023; pp. 30–32.
10. Song, C.H.; Kim, J.S.; Kim, J.M.; Pan, S. Stress Classification using ECGs based on a Multi-dimensional Feature Fusion of LSTM and Xception. *IEEE Access* **2024**, *12*, 19077–19086. [CrossRef]
11. Eltrass, A.S.; Tayel, M.B.; Ammar, A.I. Automated ECG multi-class classification system based on combining deep learning features with HRV and ECG measures. *Neural Comput. Appl.* **2022**, *34*, 8755–8775. [CrossRef]
12. Prakash, A.J.; Patro, K.K.; Samantray, S.; Sasmal, P.; Kumari, P.L.; Geetamma, T. A new approach of transparent and explainable artificial intelligence technique for patient-specific ecg beat classification. *IEEE Sens. Lett.* **2023**, *7*, 5501604. [CrossRef]
13. Elgendi, M.; Eskofier, B.; Dokos, S.; Abbott, D. Revisiting QRS detection methodologies for portable, wearable, battery-operated, and wireless ECG systems. *PLoS ONE* **2014**, *9*, e84018. [CrossRef] [PubMed]
14. Martínez, J.P.; Almeida, R.; Olmos, S.; Rocha, A.P.; Laguna, P. A wavelet-based ECG delineator: Evaluation on standard databases. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 570–581. [CrossRef] [PubMed]
15. Wang, X.; Zou, Q. QRS detection in ECG signal based on residual network. In Proceedings of the 2019 IEEE 11th International Conference on Communication Software and Networks (ICCSN), Chongqing, China, 12–15 June 2019; pp. 73–77.
16. Laitala, J.; Jiang, M.; Syrjälä, E.; Naeini, E.K.; Airola, A.; Rahmani, A.M.; Dutt, N.D.; Liljeberg, P. Robust ECG R-peak detection using LSTM. In Proceedings of the 35th Annual ACM Symposium on Applied Computing, Brno, Czech Republic, 30 March–3 April 2020; pp. 1104–1111.
17. Zahid, M.U.; Kiranyaz, S.; Ince, T.; Devecioglu, O.C.; Chowdhury, M.E.H.; Khandakar, A.; Tahir, A.; Gabbouj, M. Robust R-peak detection in low-quality Holter ECGs using 1D convolutional neural network. *IEEE Trans. Biomed. Eng.* **2021**, *69*, 119–128. [CrossRef] [PubMed]
18. Mehri, M.; Calmon, G.; Odille, F.; Oster, J. A Deep Learning Architecture Using 3D Vectorcardiogram to Detect R-Peaks in ECG with Enhanced Precision. *Sensors* **2023**, *23*, 2288. [CrossRef]
19. Wang, D.; Qiu, L.; Zhu, W.; Dong, Y.; Zhang, H.; Chen, Y. Inter-patient ECG characteristic wave detection based on convolutional neural network combined with transformer. *Biomed. Signal Process. Control* **2023**, *81*, 104436. [CrossRef]
20. Kazemnejad, A.; Karimi, S.; Gordany, P.; Clifford, G.D.; Sameni, R. An open-access simultaneous electrocardiogram and phonocardiogram database. *Physiol. Meas.* **2021**, *45*, 055005. [CrossRef]
21. Porr, B.; Howell, L. R-peak detector stress test with a new noisy ECG database reveals significant performance differences amongst popular detectors. *BioRxiv* **2019**, 722397.
22. Mahmoodabadi, S.Z.; Ahmadian, A.; Abolhasani, M.D.; Eslami, M.; Bidgoli, J.H. ECG feature extraction based on multiresolution wavelet transform. In Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 31 August–3 September 2005; pp. 3902–3905.
23. Graves, A.; Jaitly, N.; Mohamed, A.-R. Hybrid speech recognition with deep bidirectional LSTM. In Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding, Olomouc, Czech Republic, 8–12 December 2013; pp. 273–278.
24. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* **2005**, *18*, 602–610. [CrossRef]
25. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]

26. Wang, H.; Lu, C.; Zhang, Q.; Hu, Z.; Yuan, X.; Zhang, P.; Liu, W. A novel sleep staging network based on multi-scale dual attention. *Biomed. Signal Process. Control* **2022**, *74*, 103486. [CrossRef]
27. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.
28. Antoine, J.-P.; Carrette, P.; Murenzi, R.; Piette, B. Image analysis with two-dimensional continuous wavelet transform. *Signal Process.* **1993**, *31*, 241–272. [CrossRef]
29. Hermessi, H.; Mourali, O.; Zagrouba, E. Deep feature learning for soft tissue sarcoma classification in MR images via transfer learning. *Expert Syst. Appl.* **2019**, *120*, 116–127. [CrossRef]
30. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
31. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
32. Shlens, J. A tutorial on principal component analysis. *arXiv* **2014**, arXiv:1404.1100.
33. Hamilton, P.S.; Tompkins, W.J. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. *IEEE Trans. Biomed. Eng.* **1986**, *33*, 1157–1165. [CrossRef] [PubMed]
34. Christov, I.I. Real time electrocardiogram QRS detection using combined adaptive threshold. *Biomed. Eng. Online* **2004**, *3*, 28. [CrossRef]
35. Engelse, W.A.H.; Zeelenberg, C. A single scan algorithm for QRS-detection and feature extraction. *Comput. Cardiol.* **1979**, *6*, 37–42.
36. Pan, J.; Tompkins, W.J. A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* **1985**, *3*, 230–236. [CrossRef]
37. Kazemnejad, A.; Gordany, P.; Sameni, R. EPHNOGRAM: A simultaneous electrocardiogram and phonocardiogram database. PhysioNet. 2021. Available online: https://physionet.org/content/ephnogram/1.0.0/ (accessed on 15 November 2021).
38. Schmidt, P.; Reiss, A.; Duerichen, R.; Marberger, C.; Van Laerhoven, K. Introducing WESAD, a multimodal dataset for wearable stress and affect detection. In Proceedings of the 20th ACM International Conference on Multimodal Interaction, Boulder, CO, USA, 16–20 October 2018; pp. 400–408.
39. Jovic, A.; Bogunovic, N. Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. *Artif. Intell. Med.* **2011**, *51*, 175–186. [CrossRef]
40. Mar, T.; Zaunseder, S.; Martínez, J.P.; Llamedo, M.; Poll, R. Optimization of ECG classification by means of feature selection. *IEEE Trans. Biomed. Eng.* **2011**, *58*, 2168–2177. [CrossRef] [PubMed]
41. Al-Fahoum, A.S.; Howitt, I. Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias. *Med. Biol. Eng. Comput.* **1999**, *37*, 566–573. [CrossRef] [PubMed]
42. Osowski, S.; Linh, T.H. ECG beat recognition using fuzzy hybrid neural network. *IEEE Trans. Biomed. Eng.* **2001**, *48*, 1265–1271. [CrossRef]
43. De Lannoy, G.; François, D.; Delbeke, J.; Verleysen, M. Weighted SVMs and feature relevance assessment in supervised heart beat classification. In *Biomedical Engineering Systems and Technologies: Proceedings of the Third International Joint Conference, BIOSTEC 2010, Valencia, Spain, 20–23 January 2010*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 212–223.
44. Kaya, Y.; Uyar, M.; Tekin, R.; Yıldırım, S. 1D-local binary pattern based feature extraction for classification of epileptic EEG signals. *Appl. Math. Comput.* **2014**, *243*, 209–219. [CrossRef]
45. De Chazal, P.; O'Dwyer, M.; Reilly, R.B. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 1196–1206. [CrossRef] [PubMed]
46. Zhang, Z.; Dong, J.; Luo, X.; Choi, K.-S.; Wu, X. Heartbeat classification using disease-specific feature selection. *Comput. Biol. Med.* **2014**, *46*, 79–89. [CrossRef] [PubMed]
47. Luz, E.J.D.S.; Schwartz, W.R.; Cámara-Chávez, G.; Menotti, D. ECG-based heartbeat classification for arrhythmia detection: A survey. *Comput. Methods Programs Biomed.* **2016**, *127*, 144–164. [CrossRef] [PubMed]
48. Xia, W.; Li, Y.; Dong, S. Radar-based high-accuracy cardiac activity sensing. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 4003213. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.