Optimized Photoplethysmography-Based Classification of Calf Muscle Fatigue Using Particle Swarm Optimization with Logistic Regression

Sigit Dani Perkasa [1], Fadli Ama [2,*], Prisma Megantoro [3]
[1,3] Faculty of Advanced Technology and Multidiscipline, Universitas Airlangga, Surabaya, Indonesia
[2] Department of Physics, Universitas Airlangga, Surabaya, Indonesia
Email: [1] sigit.dani.perkasa-2020@ftmm.unair.ac.id, [2] fadliama@fst.unair.ac.id, [3] prisma.megantoro@ftmm.unair.ac.id
*Corresponding Author

Abstract*—This study investigates photoplethysmography (PPG) as a non-invasive, cost-effective alternative for real-time muscle fatigue monitoring, addressing limitations inherent to conventional methods like electromyography (EMG) and blood lactate testing. A PPG-based system was developed to classify fatigued versus non-fatigued states of the calf muscle using a DFRobot SEN0203 sensor at a 1000 Hz sampling rate. The raw PPG signals were segmented into 1-second intervals and processed to compute first and second derivatives—yielding vascular (VPG) and arterial (APG) photoplethysmograms—which enabled extraction of key features including heart rate (HR), heart rate variability (HRV), peak systolic and diastolic voltages, maximum systolic slope (u), minimum diastolic slope (v), and arterial stiffness indicators (b–a and c–a ratios). A Particle Swarm Optimization (PSO) algorithm was employed to optimize both feature selection and hyperparameters within a Logistic Regression (LR) model, achieving perfect classification accuracy (1.0) with training and prediction times of 0.0053 s and 0.0016 s, respectively. Notably, HRV and the minimum diastolic slope—reflecting autonomic regulation and vascular compliance—emerged as the most influential features with weights of 12.3747 and 23.9367. Comparative analyses revealed that although LightGBM matched the PSO-LR accuracy, neural network approaches performed poorly (0.50 accuracy), likely due to overfitting and limited training data. These findings underscore the viability of PPG for muscle fatigue monitoring, with promising applications in sports science, rehabilitation, and occupational health.

Keywords*—*Photoplethysmography (PPG), Muscle Fatigue Monitoring, Particle Swarm Optimization, Logistic Regression

I. INTRODUCTION

Muscle fatigue monitoring is essential in sports science, rehabilitation, occupational health, and chronic disease management to enhance performance and prevent injury. Wearable technologies, such as photoplethysmography (PPG), enable real-time, non-invasive monitoring, benefiting athletes by reducing injury risks [1] and improving well-being in high-risk workers like firefighters [2]. In chronic disease management, low-cost sensors support independent living and prevent health decline [3]. These innovations highlight the need for accessible, non-invasive fatigue assessment tools to improve health outcomes across various domains.

Traditional muscle fatigue detection methods, like electromyography (EMG) and blood lactate testing, are invasive, expensive, and impractical for continuous monitoring. While valuable in research and clinical settings, these techniques are unsuitable for real-time, portable applications. EMG requires precise placement of equipment, limiting long-term monitoring [4], while blood lactate measurement is invasive and not feasible outside clinical environments [5]. Recent advancements in wearable, non-invasive sensors, particularly PPG, provide an alternative for continuous, real-time monitoring, offering valuable insights into muscle fatigue without the limitations of traditional methods [6], [7].

While PPG is widely used in cardiovascular monitoring, its application in muscle fatigue detection remains underexplored. The potential of PPG for assessing localized muscle fatigue during dynamic activities is promising; however, challenges such as motion artifacts and signal noise can obscure the subtle physiological changes associated with fatigue. In this study, we address these challenges by incorporating advanced signal processing techniques to enhance signal quality and mitigate noise.

This study aims to develop a PPG-based system to classify fatigued and non-fatigued states of the calf muscle. To improve the detection sensitivity beyond raw PPG analysis, the system incorporates novel techniques by computing the first and second derivatives of the PPG signal—yielding vascular (VPG) and arterial (APG) photoplethysmograms—which capture essential waveform characteristics such as the maximum systolic slope (u) and the minimum diastolic slope (v). These derivative features provide a more sensitive indicator of muscle fatigue compared to traditional metrics.

Furthermore, to optimize classification performance, the study integrates Particle Swarm Optimization (PSO) with a Logistic Regression (LR) model. PSO is employed for its efficiency in navigating high-dimensional feature spaces and for fine-tuning model hyperparameters, resulting in improved accuracy and computational efficiency. Comparative analyses with other machine learning approaches reveal that while models like LightGBM can match the performance of the PSO-LR model, neural network approaches may underperform due to overfitting and limited training data.

This study contributes to the field of non-invasive muscle fatigue monitoring by proposing an optimized PPG-based system that effectively addresses challenges such as motion artifacts and signal noise, enhances feature extraction through derivative analysis, and demonstrates superior performance

via the integration of PSO with LR. These contributions pave the way for more reliable and accessible real-time fatigue assessment in diverse health and performance settings.

II. METHOD

A. System Configuration

The system configuration used for PPG signal processing, feature extraction, model optimization, and evaluation for fatigue classification is illustrated in the diagram in Fig. 1.

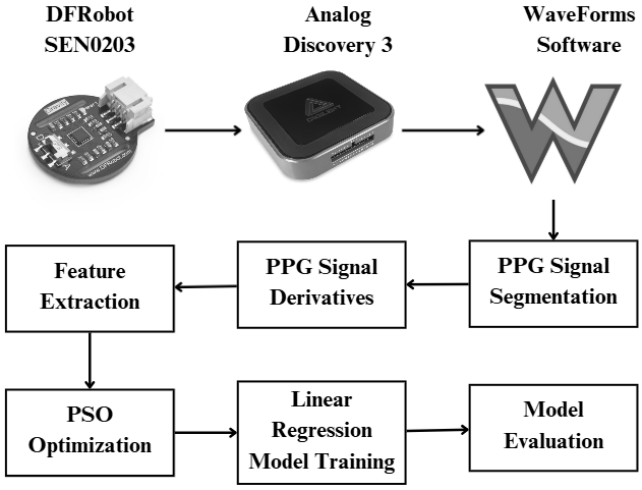


Fig. 1. System configuration diagram

The system configuration, as shown in Fig. 1, comprises a series of steps that ensure robust fatigue classification: the DFRobot SEN0203 sensor first acquires raw PPG signals from the calf muscle, which are then transmitted to the Analog Discovery 3 device for real-time visualization via WaveForms Software, ensuring signal quality; subsequently, key features such as heart rate and variability are extracted, and the first and second derivatives (VPG and APG) are computed to capture subtle waveform characteristics essential for detecting fatigue; the processed signals are segmented into smaller intervals to enable detailed temporal analysis; next, these segments undergo Particle Swarm Optimization (PSO) to fine-tune the parameters and feature weights of a Linear Regression model; finally, the optimized model is trained and its performance is evaluated using metrics like accuracy, precision, recall, and F1-score, thereby creating an efficient workflow that enhances the reliability of real-time muscle fatigue monitoring.

B. Signal Acquisition

The signal acquisition process consisted of PPG sensor placement on the calf and a data collection process. The PPG sensor placement is illustrated in Fig. 2.

Fig. 2 illustrates the PPG sensor placement used in this study, positioned on the gastrocnemius muscle to monitor blood flow and pulse during activities. This placement has been shown to provide accurate heart rate data and insights into cardiovascular health by capturing pulse wave characteristics and heart rate variability [8], [9].The gastrocnemius muscle is a reliable site for measuring pulse signals, reflecting blood flow dynamics during movement [10], [11]. The data collection process was performed in three parts. First, PPG signals are collected for 2 minutes when a set of different movements are performed to capture the non-fatigued state and continued with 30 seconds of rest. After the rest, 120 seconds of data collection is performed again to capture the fatigued state. This approach aligns with studies using PPG signals to distinguish between fatigued and non-fatigued states during physical activity [12]-[14]. The physical activity set is illustrated in Fig. 3.

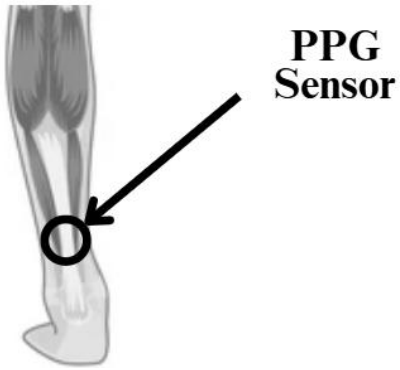


Fig. 2. PPG sensor placement

Fig. 3. Illustration of physical activity exercise for PPG acquisition

Fig. 3 shows a set of exercises designed to record PPG signals during various lower limb movements. These exercises include ankle movements such as plantarflexion (hind foot rising) and dorsiflexion (forefoot rising), knee uplifts, knee extensions with and without plantarflexion, and ankle rotations in both clockwise and counter-clockwise directions. Other exercises involve outward and inward toe movements with a fixed hind foot, as well as knee extensions with dorsiflexion. These exercises target different muscle groups and joints, allowing for the collection of comprehensive PPG data on the calf muscle [15]-[17]. The data is collected with a sampling rate of 1000 Hz for 120 seconds.

C. Signal Preprocessing

1) Signal Segmentation

In this study, the PPG signal was segmented into 1-second intervals, each containing 1000 samples, to facilitate analysis of the physiological changes over time. The segmentation process involves dividing the 120-second raw PPG signal into n segments, where n = 120 segments each containing exactly 1000 samples. The signal segmentation is represented by Eq. (1).

$$x_i = PPG_{(1000\cdot(i-1)+1):1000\cdot i}, \qquad i = 1, 2, 3, \ldots, 120 \tag{1}$$

Where x_i represents the i-th segment of the PPG signal, extracted from the original signal between sample indices $1000(i-1)+1$ to $1000i$. This segmentation approach is consistent with recent works on PPG signal processing, where segmentation into 1-second intervals has been widely adopted for continuous monitoring of heart rate and other cardiovascular features [18]-[21]. While a 1-second interval effectively captures transient physiological variations, exploring alternative segmentation intervals (e.g., shorter segments for higher temporal resolution or longer segments for noise reduction) could influence the sensitivity and accuracy of fatigue detection.

2) PPG Signal Derivatives

In this study, the PPG signal derivatives, specifically the vascular photoplethysmogram (VPG) and the arterial photoplethysmogram (APG), were computed to enhance the analysis of signal features. The VPG and APG are derived from the first and second derivatives of the PPG signal, respectively. Mathematically, the VPG is the first derivative of the PPG signal, and the APG is the second derivative, expressed in Eq. (2) and Eq. (3).

$$VPG(t) = \frac{d}{dt} PPG(t) \tag{2}$$

$$APG(t) = \frac{d^2}{dt^2} PPG(t) \tag{3}$$

These derivatives are particularly useful for detecting key features in the PPG waveform, such as peaks and troughs, which are crucial for cardiovascular analysis and diagnostics [22]-[25].

D. Feature Extraction

1) PPG Features

Heart rate (HR) and heart rate variability (HRV) are essential features extracted from PPG signals for cardiovascular health monitoring. HR can be derived by detecting the peaks of the PPG signal, which correspond to the heartbeats. Mathematically, the heart rate HR can be calculated as the reciprocal of the average time between consecutive peaks (R-peaks) in the PPG signal, as expressed in Eq. (4).

$$HR = \frac{60}{Mean\ RR\ Interval} \tag{4}$$

Where the RR interval is the time difference between two successive peaks in the PPG waveform [26]. For HRV, the variation in the time intervals between successive heartbeats (RR intervals) is analyzed. HRV is a marker of autonomic nervous system activity and can be extracted by computing the standard deviation of RR intervals. The formula for HRV, based on the standard deviation of RR intervals is expressed in Eq. (5).

$$HRV = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(RR_i - \overline{RR})^2} \tag{5}$$

Where RR_i is the i-th RR interval, and $\overline{RR}$ is the mean RR interval [27], [28].

2) VPG Features

The extraction of features such as the maximum slope during systole (u) and the minimum slope during diastole (v) from VPG signals plays a critical role in cardiovascular analysis [22], [29]. The maximum slope during systole, u is typically identified as the steepest rising portion of the VPG signal, while the minimum slope during diastole, v, corresponds to the most negative slope during the relaxation phase of the vascular cycle. Mathematically, the maximum slope during systole u is expressed in Eq. (6).

$$u = max\left(\frac{d}{dt} VPG(t)\right) \tag{6}$$

Similarly, the minimum slope during diastole v is expressed in Eq. (7).

$$v = min\left(\frac{d}{dt} VPG(t)\right) \tag{7}$$

Each character is uniquely represented in the audio domain, allowing accurate transmission and decoding on the receiver side.

3) APG Features

The extracted features of the APG signals are the b-a and c-a ratios. The c-a ratio is defined as the ratio between the amplitude of the c point (representing the peak of the upstroke) and the a point (the maximum value before the dicrotic notch) in the APG waveform. Mathematically, the c-a ratio is expressed in Eq. (8).

$$c - a\ ratio = \frac{A_c}{A_a} \tag{8}$$

where A_c is the amplitude at the c point, and A_a is the amplitude at the "a" point of the APG signal. Similarly, the b-a ratio is calculated as the ratio between the b point (the peak after the dicrotic notch) and the a point, as expressed in Eq. (9).

$$b - a\ ratio = \frac{A_b}{A_a} \tag{9}$$

where A_b is the amplitude at the b point. These ratios are important indicators of arterial compliance and vascular health [24], [30], [31]. Where the a, b, and c points are determined using Eq. (10) to Eq. (12).

$$A_c = max\ (APG(t)), \qquad t \in [t_{start}, t_{end}] \tag{10}$$

$$A_a = max\ (APG(t)), \qquad t \in [t_c, t_{dicrotic\ notch}] \tag{11}$$

$$A_b = max\ (APG(t)), \qquad t \in [t_{dicrotic\ notch}, t_{end}] \tag{12}$$

Where t_{start} and t_{end} are the time intervals corresponding to the systolic upstroke of the APG signal. While t_c is the time of the ***c*** point, and $t_{dicrotic\ notch}$ is the time of the dicrotic notch [22], [24], [32]. The extracted features from PPG, VPG, and APG signals capture these physiological changes—PPG metrics reflect overall cardiac activity, VPG-derived slopes indicate rapid changes in blood flow, and APG ratios provide insights into vascular stiffness—all of which are critical for distinguishing between fatigued and non-fatigued states

E. Logistic Regression

This study uses logistic regression to classify the calf muscle into fatigued and non-fatigued states based on PPG signal features and their derivatives. The model predicts the probability of fatigue using features such as HR, HRV, and characteristics derived from VPG and APG signals, including the maximum slope during systole (*u*), the minimum slope during diastole (*v*), and the c-a and b-a ratios. The logistic regression used in this model is represented by Eq. (13) and Eq. (14).

$$logit(p) = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n \tag{13}$$

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n)}} \tag{14}$$

where p is the probability of being in the fatigued state, β_0 is the intercept of the model, while β_1, $\ldots\beta_n$ is the weight coefficient of features and x_1, x_1, …, x_n is the extracted features from the PPG signals acting as LR inputs. This method is effective for binary classification tasks [33], [34]. Logistic regression assumes a linear relationship between the log-odds of the outcome and the input features, along with independence among observations. In this study, the extracted PPG-derived features satisfy these assumptions, making logistic regression a robust choice for binary classification of fatigued versus non-fatigued states. Its simplicity and interpretability further facilitate the understanding of how each feature contributes to fatigue detection.

F. Particle Swarm Optimization Objective

The optimization process is driven by the mean cross-validation accuracy, which is the objective function. This accuracy is calculated by evaluating the LR model using k-fold cross-validation, ensuring that the model is robust and generalizes well to unseen data. The cross-validation accuracy A is represented by Eq. (15). where *k* represents the number of folds, and X^i_{train} and Y^i_{train} are the training features and labels for the i-th fold, PSO aims to maximize this accuracy by optimizing the weights of the extracted features and the LR hyperparameters. PSO allows for optimizing both the feature weights and the hyperparameters of the logistic regression model [35], [36]. By iteratively adjusting these parameters, PSO enhances the model's ability to classify the calf muscle's fatigue state based on the extracted features from the PPG signals. This approach ensures that the LR model is tailored to the task of fatigue classification, improving its performance and accuracy. The PSO in this research uses 30 particles, with maximum iterations of 100. Along with 0.5 inertia weights (*w*), personal inertia weights coefficient (*c1*) of 2 and global learning coefficients (*c2*) of 3.

$$A = \frac{1}{k}\sum_{i=1}^{k} Accuracy\left(X^i_{train}, Y^i_{train}\right) \tag{15}$$

Particle Swarm Optimization (PSO) was selected for parameter tuning due to its efficient balance between exploration and exploitation, which enables rapid convergence to optimal solutions. Compared to methods like Grid Search or Genetic Algorithms, PSO requires fewer computational resources and iterations, leading to improved model accuracy and reduced training time. This optimization approach effectively fine-tuned both feature weights and hyperparameters, enhancing the overall performance and efficiency of the logistic regression model for fatigue classification.

III. RESULTS

A. Signal Acquisition

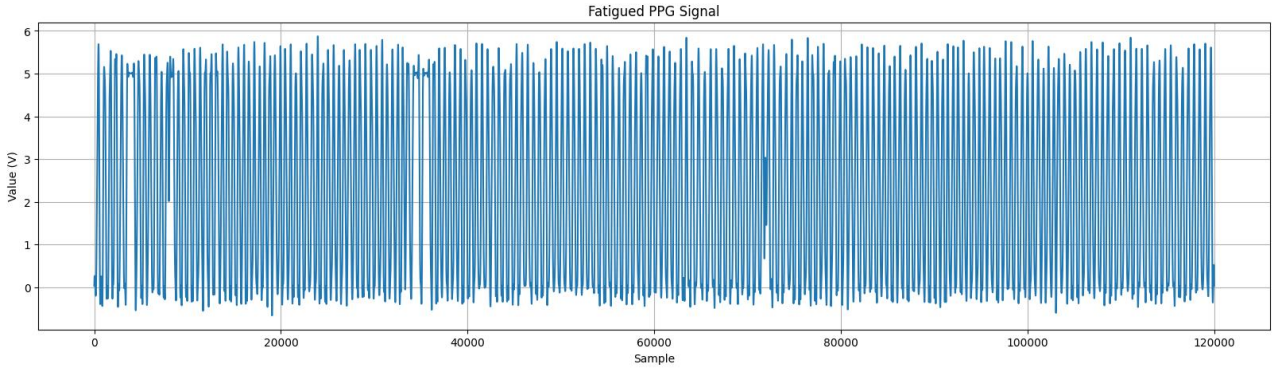
The PPG signal acquisition results using the DFRobot PPG sensor and Analog Discovery 3, with a 120-second acquisition period at a sampling frequency of 1000 Hz. As illustrated in Fig. 4.

Fig. 4 presents the obtained PPG signals. Fig. 4a shows the fatigued signals and Fig. 4b shows the non-fatigued signal. Both signals showed 120000 total samples, aligned with the acquisition time and sampling frequency. The fatigued signal exhibits pronounced irregularities and noise, with amplitude fluctuations. In contrast, the non-fatigued signal remains relatively stable, both signal has similar amplitudes oscillating between approximately -1 and 6 V. Despite the higher noise levels in the fatigued state, the non-fatigued signal is not entirely free from fluctuations, indicating that even under normal conditions, some variability exists.

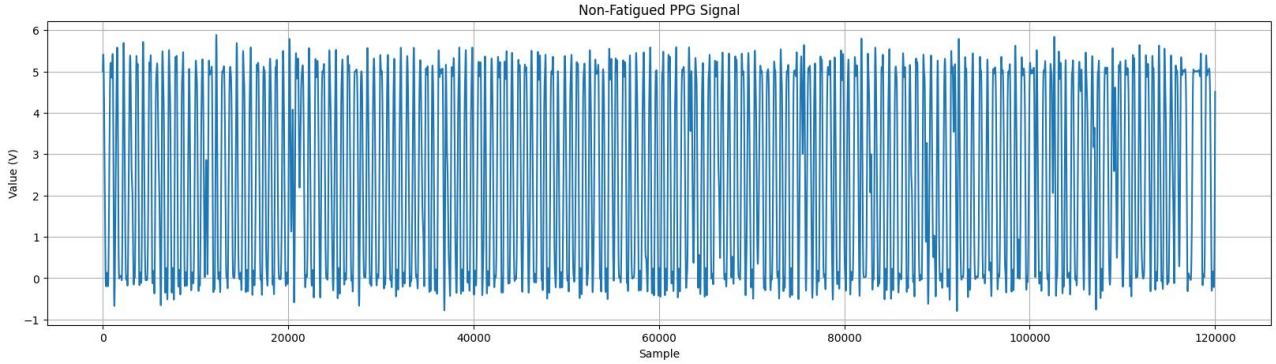
B. Signal Segmentation

After the PPG signals are obtained, both signals are then segmented into 120 parts. Each part consisted of 1000 samples, the segmented samples of both fatigued and non-fatigued signals are illustrated in Fig. 5.

Fig. 5 shows the segmented PPG signals for both fatigued in Fig. 5a and non-fatigued in Fig. 5b, with each signal divided into 120 segments of 1000 samples. The color-coded segments reveal distinct PPG waveforms, each displaying clear systolic and diastolic peaks. In the fatigued signal, the systolic peaks reach up to 6 V, while the diastolic troughs fall to approximately -1 V. The non-fatigued signal exhibits similar amplitude ranges, with peaks fluctuating between 6 and -1 V. This segmentation allows for detailed feature extraction, facilitating the analysis of PPG dynamics.

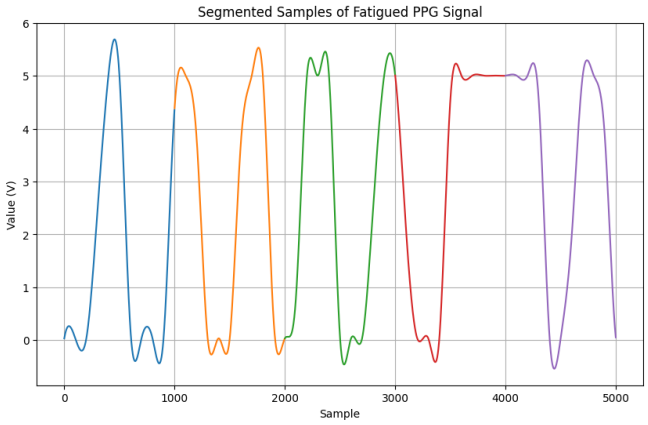


(a)

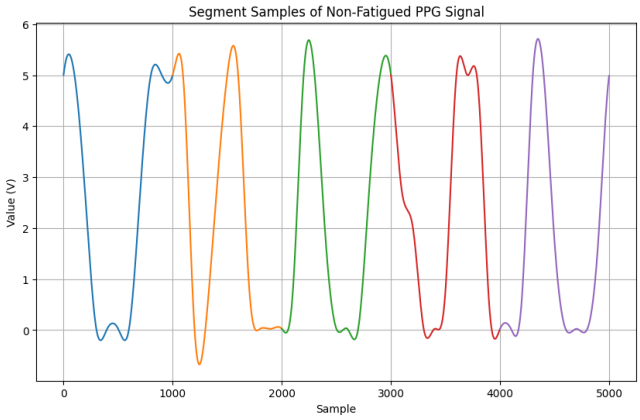


(b)

Fig. 4. Obtained PPG signals, (a) Fatigued signals and (b) Non-fatigued signals



(a)



(b)

Fig. 5. Segmented PPG signals, (a) Fatigued state and (b) Non-fatigued state

C. PPG Signals Derivatives Identification

After the raw PPG signals are segmented into 120 segments, the derivatives of those signals are calculated for further identification intended for finding different peaks of the original signals and their derivatives on the fatigued and non-fatigued states. As presented in Fig. 6.

Fig. 6 shows a sample PPG segment and its derivatives (VPG and APG) for both fatigued and non-fatigued states. In the non-fatigued state (top row), the systolic peak is marked at approximately 5 V, and the diastolic peak is observed around 1 V on the PPG waveform. The corresponding VPG derivative exhibits inflection points *u* and *v*, with values of 0.06 and 0.02, respectively. The APG signal shows peaks *a*, *b*, and *c* with values of 0.0015, -0.0005, and 0.0006, respectively. In the fatigued state (bottom row), the systolic peak remains around 5 V, while the diastolic peak is slightly lower, around 1.5 V. The VPG shows the inflection points *u* and *v* at approximately 0.04 and 0.01, respectively. The APG values for peaks *a*, *b*, and *c* are 0.0013, -0.0003, and 0.0005, showing subtle shifts compared to the non-fatigued condition.

D. Feature Extraction

After the PPG, VPG, and APG signals are identified, feature extraction is then performed on each signal's segment. Fig. 7 presents the Kernel Density Estimate (KDE) plots of extracted features from PPG, VPG, and APG signals for Class 1 (fatigued) and Class 0 (non-fatigued).

Further features such as *v* and *u* show distinctive patterns as well. Class 1 shows a peak around 0.05 for *u*, while Class 0 has a broader range between 0.02 and 0.06. The *v* for Class 1 is concentrated around -0.05, while Class 0 peaks around -0.02, indicating differences in the diastolic slopes.

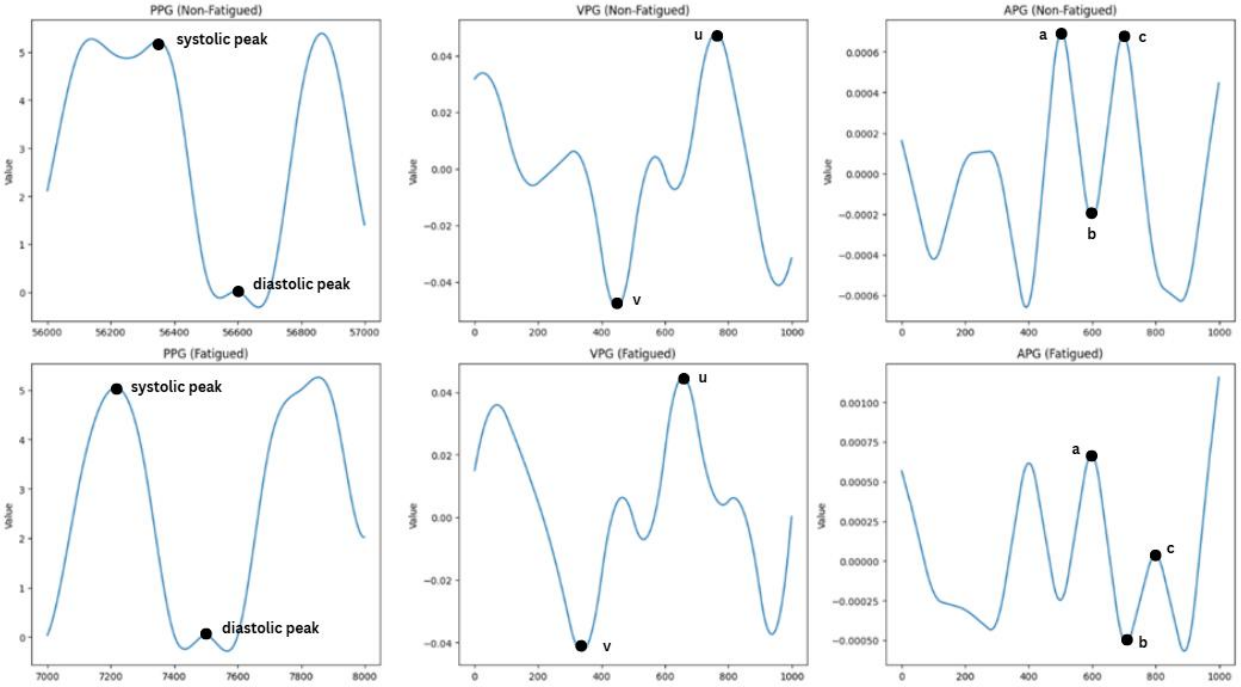


Fig. 6. Identification results of PPG, VPG, and APG signals

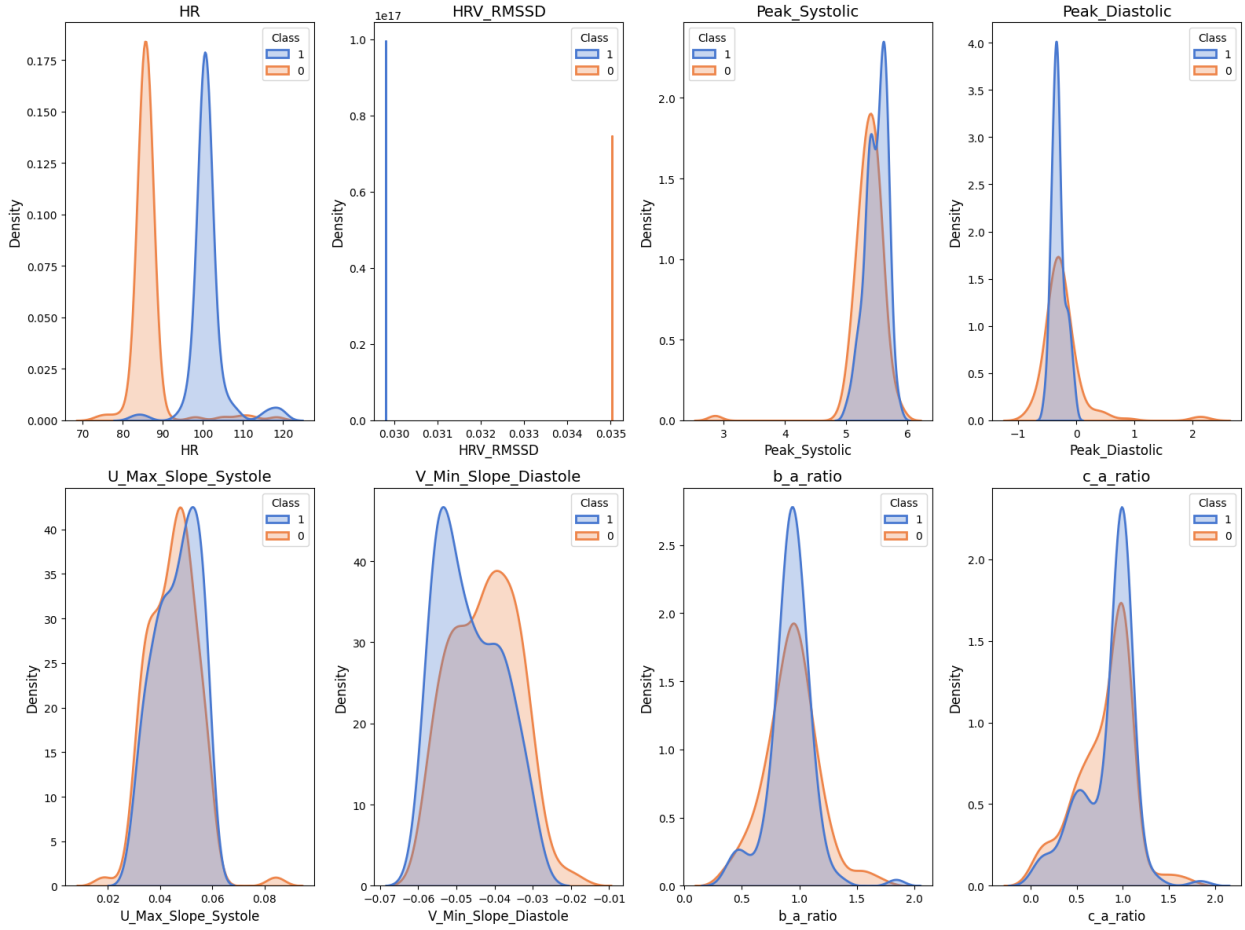


Fig. 7. Kernel density estimates (KDE) plot of extracted feature

The b-a ratio for Class 1 peaks around 1.5, suggesting a higher ratio compared to Class 0, which peaks around 1.2. Similarly, the c-a ratio for Class 1 peaks at 0.6, while Class 0 peaks at 0.4. These differences in feature distributions indicate that PPG, VPG, and APG-derived features can effectively differentiate between the two classes. The statistical description of the extracted features from the combined Class 0 and Class 1 signals is presented in Table 1.

Table 1. Statistical description of combined extracted features

Feature	Count	Mean	Std	Min	25%	50%	75%	Max	Skewness	Kurtosis
HR	240	93.730604	8.803003	74.257426	85.653149	96.696778	100.671141	119.760479	0.404381	-0.566986
HRV	240	0.032412	0.002618	0.0298	0.0298	0.032412	0.035025	0.035025	0	-2.016878
Peak systolic	240	5.425319	0.246275	2.854343	5.317771	5.439169	5.573599	5.884458	-4.811497	48.701109
Peak diastolic	240	-0.274818	0.290815	-0.798075	-0.39699	-0.313972	-0.223681	2.187049	5.205222	39.183023
u	240	0.045637	0.008592	0.018349	0.039075	0.046282	0.052049	0.084246	0.079414	0.689133
v	240	-0.044608	0.008586	-0.059149	-0.05207	-0.044949	-0.037741	-0.019163	0.233061	-0.895398
b-a ratio	240	0.926409	0.210469	0.356793	0.828553	0.934225	1.03306	1.838638	0.271138	2.584006
c-a ratio	240	0.825255	0.291913	0.05524	0.643184	0.986673	1	1.838638	-0.568869	0.903863

Based on Table 1, its known that the data includes 240 samples for each feature, with key statistics such as mean, standard deviation (Std), minimum, 25th percentile (25%), median (50%), 75th percentile (75%), maximum (Max), skewness, and kurtosis. For instance, the mean HR is 93.73 bpm, with a standard deviation of 8.80 bpm, and the values range from 74.26 to 119.76 bpm. The HRV feature shows a narrow range with a mean of 0.0324, and the peak systolic values have a mean of 5.43 V, with a standard deviation of 0.25 V, showing slight variability in systolic peak amplitudes. The peak diastolic values have a mean of -0.27 V, with a high kurtosis value (39.18), indicating a significant concentration around the mean, with some extreme outliers. The *u* and *v* slopes show relatively consistent values, with *u* having a mean of 0.0456 and *v* showing a negative mean of -0.0446. The b-a ratio and c-a ratio features display higher variability, with means of 0.93 and 0.83, respectively, and notable skewness in their distributions.

E. PSO on LR Model

Fig. 8 shows the convergence curve of the PSO algorithm optimizing a linear regression model for classifying fatigue and non-fatigue states. The y-axis represents the negative cross-validation accuracy, where -1 indicates perfect accuracy. Initially, from iteration 0 to 20, the accuracy slightly decreases from about -0.965 to -0.975. After iteration 20, the accuracy improves significantly, reaching -1 at around iteration 30, indicating perfect classification. The curve stabilizes at -1 thereafter, suggesting that the PSO algorithm effectively optimized the model parameters, achieving ideal performance in classifying the two states based on the extracted features. The feature weights for the LR model throughout the optimization process is explained in Fig. 9.

Fig. 9 shows the progress of feature weights for the linear regression (LR) model throughout the PSO iterations. The y-axis represents the feature weights, while the x-axis shows the iteration number. The different colored lines represent individual features, such as HR, HRV, peak systolic, peak diastolic, *u*, *v*, b-a ratio, and c-a ratio. Initially, at the start of the iterations, the feature weights experience significant fluctuations, with some features, such as HR and HRV, showing large jumps in weight values, even reaching values over 2000. These drastic changes are likely due to the PSO exploring different parameter combinations. By iteration 40, the weights stabilize, as shown in the zoomed-in section on the right, where the feature weights are nearly constant with small variations between them. This stability suggests that the PSO has converged, having successfully optimized the feature weights for the linear regression model. The final feature weights are relatively small and consistent, indicating the optimal feature selection and adjustment by the PSO algorithm. Aside from the feature weights, the C parameter of the LR model is also optimized, the progress of the C optimization is illustrated in Fig. 10.

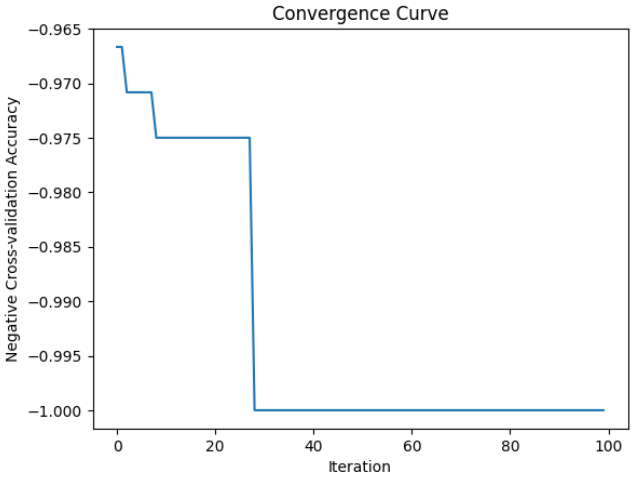


Fig. 8. Convergence curve of PSO optimization

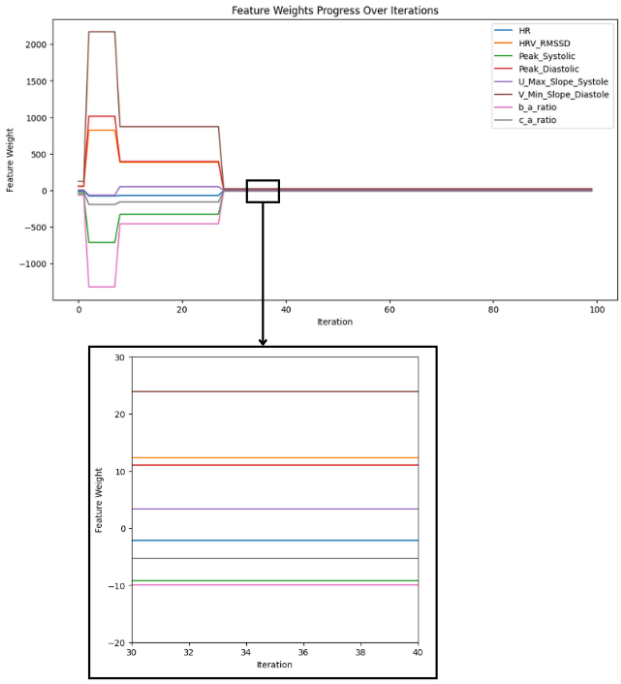


Fig. 9. Feature weights progression for LR model across PSO iterations

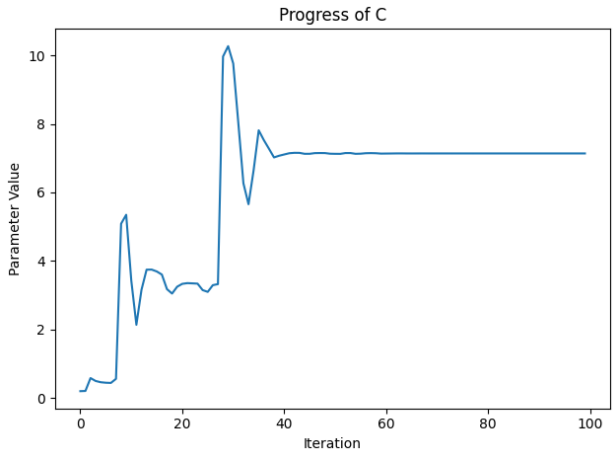


Fig. 10. LR regularization parameter (C) progression across PSO iterations

Fig. 10 shows the progress of the regularization parameter ***C*** for the LR model during the PSO optimization process. The y-axis represents the value of parameter ***C***, while the x-axis shows the iteration number. Initially, the parameter ***C*** fluctuates with small values, then experiences significant jumps at around iteration 20 and again at iteration 40, peaking at a value of about 10. After these sharp changes, the parameter stabilizes and gradually settles around a value of 7. This pattern indicates that the PSO algorithm explored various values for ***C*** before converging to an optimal range, suggesting the model has effectively adjusted the regularization to improve performance. Next, the particle diversity of the PSO algorithm is measured by the mean pairwise distance (diversity) between particles over the iterations. Illustrated in Fig. 11.

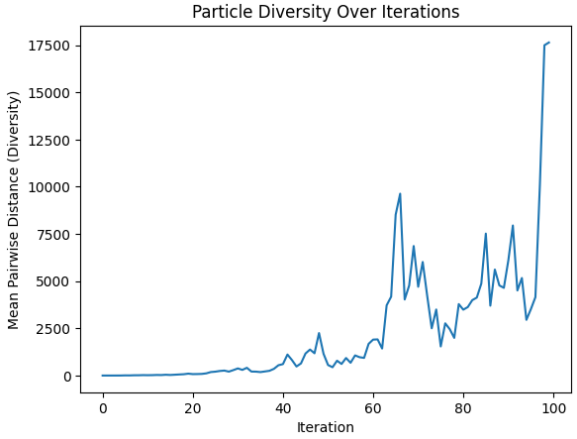


Fig. 11. PSO particle diversity across iterations

Fig. 11 shows the particle diversity of the PSO. The y-axis represents the mean pairwise distance, while the x-axis shows the number of iterations. Initially, the diversity is low, with the mean pairwise distance remaining relatively small in the early iterations, suggesting that the particles are closely clustered around similar solutions. As the iterations progress, the diversity increases significantly, particularly after iteration 60, where the mean pairwise distance experiences substantial fluctuations and sharp peaks. At iteration 100, the diversity reaches a very high value, exceeding 17,500, indicating that the particles have explored a broader range of the solution space. This sharp increase in diversity suggests that the PSO algorithm is diversifying its search, possibly moving toward more distinct solutions as it continues to optimize. The results of the PSO applied to the LR model for classifying fatigue and non-fatigue states using extracted features is explained in Table 2.

Table 2. PSO optimization results on LR model

Parameter	Value
C	7.1366
Penalty	l1
Max Iterations	3000
Tolerance	1,00E-05
Class Weight	None
Feature Weights	[-2.1990, 12.3747, -9.2000, 11.0021, 3.3338, 23.9367, -9.9385, -5.3015]
Obj. Value	1.0
Optimization Time	44s

Based on Table 2, the regularization parameter C was optimized to a value of 7.1366, balancing model complexity and accuracy, while the L1 penalty (Lasso) was used to encourage sparsity in the feature weights. The LR process ran for a maximum of 3000 iterations and stopped when the change in the objective function became negligible, as indicated by the tolerance of 1.00E-05. The class weight was set to None, treating all classes equally without any adjustments for class imbalance. The feature weights, corresponding to the extracted features in the order of HR, HRV, Peak Systolic, Peak Diastolic, *u*, *v*, b-a ratio, and c-a ratio, were optimized as follows: [-2.1990, 12.3747, -9.2000, 11.0021, 3.3338, 23.9367, -9.9385, -5.3015]. These values reflect the importance of each feature, with HRV (12.3747) and *v* (23.9367) having positive weights, suggesting their strong influence on the model, while HR (-2.1990) and b-a ratio (-9.9385) have negative weights, indicating a negative relationship with the classification target. The model achieved perfect classification accuracy (1.0) and completed the optimization in 44 seconds, demonstrating both efficiency and effective parameter tuning by the PSO algorithm.

Fig. 12 shows the training accuracy and error of the LR model over 3000 iterations, using the optimized parameters from the PSO process with the training accuracy in Fig. 9a and error in Fig. 9b. The Training Accuracy Over Iterations graph demonstrates that the model quickly achieves perfect accuracy (1.0) and maintains it throughout the iterations, indicating optimal classification performance early in the training. Similarly, the Training Error Over Iterations graph shows a rapid decrease in error, stabilizing at a very low value near 0, which aligns with the accuracy improvement. These trends suggest that the optimized LR model converged quickly to an ideal solution, achieving both minimal error and perfect accuracy in a short number of iterations.

F. Comparison with Other Models

The PSO-LR model is then compared with other machine learning models using the same signal and extracted features dataset to assess its performance compared to other several popular machine learning models in classifying calf muscle fatigue from PPG signals. The comparison is presented in Table 3.

Table 3 compares the performance of several machine learning models, including the PSO-LR, on the same dataset of signals and features.

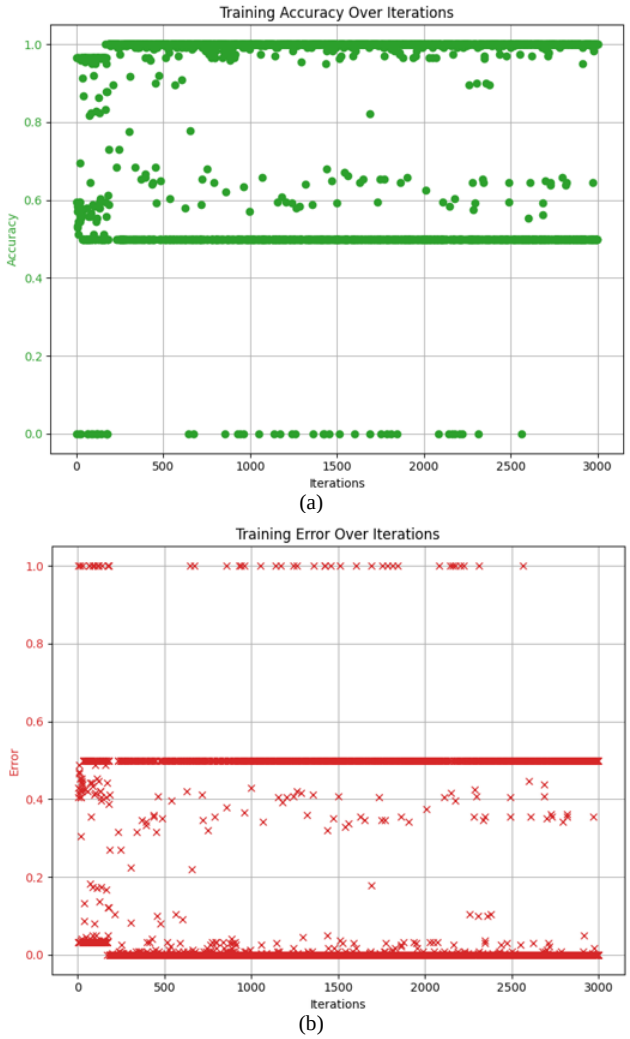


Fig. 12. Optimized LR model training progression, (a) training accuracy and (b) training loss

Linear Regression (LR), K-Nearest Neighbors (KNN), and Support Vector Machine (SVM) all achieved high accuracy of 0.9792, with fast training times and relatively quick prediction times. These models exhibited excellent performance across all metrics, including precision (1.00), recall (0.96), and F1-score (0.98), indicating strong classification capabilities and well-balanced performance. In contrast, Convolutional Neural Network (CNN) and Artificial Neural Network (ANN) performed poorly, both with an accuracy of 0.50, suggesting challenges with the dataset. CNN had the longest training time (90.615 seconds) and slowest prediction time (0.7330 seconds), with a poor cross-validation score (0.1292) and a low F1-score (0.69). Similarly, ANN also struggled with lower performance, achieving an accuracy of 0.50, a moderate F1-score of 0.67, and a higher recall (1.00), but still lagged behind in overall efficiency. On the other hand, both LightGBM and PSO-LR achieved perfect accuracy (1.00), making them the top performers. PSO-LR, optimized by the PSO algorithm, had the fastest training (0.0053 seconds) and prediction time (0.0016 seconds), with perfect precision, recall, and F1-score (1.00). This indicates its superior efficiency and classification capability. LightGBM also reached perfect scores but had slightly longer training (0.0730 seconds) and prediction times (0.0023 seconds) compared to PSO-LR.

VI. Discussion

This study demonstrates the successful application of Particle Swarm Optimization (PSO) in optimizing a Linear Regression (LR) model for calf muscle fatigue classification using Photoplethysmogram (PPG) signals. The fatigued PPG signals exhibited notable irregularities and noise, distinguishing them from the relatively stable non-fatigued signals. These differences provided a clear basis for feature extraction, with physiological parameters such as heart rate (HR), heart rate variability (HRV), and systolic/diastolic peak values showing distinct patterns between the two conditions. HRV and the v feature were identified as crucial for the classification task due to their high feature weights, indicating their strong influence on the model.

The PSO optimization process was critical in fine-tuning the LR model, achieving perfect classification accuracy (1.0) with minimal error. The optimization led to adjusting key parameters, such as the regularization parameter (C), and stabilized feature weights, ensuring the model's ability to classify the fatigued and non-fatigued states accurately. The rapid convergence of the PSO algorithm, reaching optimal performance after just 30 iterations, highlights the effectiveness of the optimization process in enhancing model efficiency. This demonstrates the potential for PSO to improve the performance of machine learning models in fatigue classification tasks.

When compared to other popular machine learning models, the PSO-LR model outperformed K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Convolutional Neural Networks (CNN), and Artificial Neural Networks (ANN) in terms of classification accuracy, training time, and prediction time. While KNN and SVM also achieved high accuracy, the PSO-LR model's speed and efficiency, with training and prediction times of 0.0053 and 0.0016 seconds, respectively, set it apart. In contrast, CNN and ANN struggled with lower accuracy and longer processing times, underscoring the superiority of the PSO-LR approach. These findings suggest that PSO-optimized LR models offer a powerful tool for real-time fatigue classification, with significant potential for applications in health monitoring and other domains requiring efficient signal processing.

Table 3. Current parameter signal transmission and data translation

Model	Accuracy	Training Time (s)	Prediction Time (s)	Mean Cross-Validation Score	Precision	Recall	F1-Score
LR	0.9792	0.0061	0.0028	0.9667	1.00	0.96	0.98
KNN	0.9792	0.0145	0.0305	0.9708	1.00	0.96	0.98
SVM	0.9792	0.0061	0.0032	0.9708	1.00	0.96	0.98
CNN	0.50	90.615	0.7330	0.1292	0.55	0.92	0.69
ANN	0.50	26.893	0.1907	0.2917	0.50	1.00	0.67
LightGBM	1.00	0.0730	0.0023	1.00	1.00	1.00	1.00
PSO-LR	1.00	0.0053	0.0016	1.00	1.00	1.00	1.00

V. CONCLUSION

This study demonstrates that photoplethysmography (PPG), coupled with Particle Swarm Optimization (PSO)-enhanced Logistic Regression (LR), offers an effective and computationally efficient method for classifying calf muscle fatigue. Our findings reveal distinct differences in PPG signals between fatigued and non-fatigued states, with heart rate variability (HRV) and the minimum diastolic slope (v) emerging as the most influential features (weights: 12.3747 and 23.9367, respectively). The optimized PSO-LR model achieved perfect classification accuracy while outperforming traditional models such as LR, KNN, and SVM (accuracy of 0.9792) in terms of efficiency (training time: 0.0053 s, prediction time: 0.0016 s), whereas CNN and ANN performed poorly—likely due to overfitting and limited training data. Statistical analyses confirmed the significance of these differences, reinforcing the reliability of the selected features. Segmenting the signals into 120 one-second intervals and computing VPG and APG derivatives facilitated the detection of subtle waveform shifts, such as changes in systolic and diastolic peaks and APG ratios, which are closely linked to muscle fatigue. These promising results suggest that PPG-based fatigue monitoring can be implemented in real-time wearable devices for applications in sports science, rehabilitation, and occupational health. However, limitations such as a modest sample size and potential variability in sensor placement should be addressed in future studies. Further research is needed to validate these findings across diverse populations, explore additional muscle groups, and tackle practical challenges like motion artifacts and sensor calibration for robust real-world integration.

ACKNOWLEDGMENT

The Authors would like to express our sincere gratitude to the **Research Centre for New and Renewable Energy Engineering, Universitas Airlangga**, for providing the resources and support necessary for the successful completion of this study. We also extend our thanks to the **Biomedical Instrumentation Laboratory, Universitas Airlangga**, for their invaluable assistance in data collection and technical expertise. Special appreciation is given to the faculty and staff from both institutions for their guidance, collaboration, and constructive feedback throughout the course of this research.

REFERENCES

[1] T. T. de Beukelaar and D. Mantini, “Monitoring Resistance Training in Real Time with Wearable Technology: Current Applications and Future Directions,” *Bioengineering*, vol. 10, no. 9, p. 1085, 2023, https://doi.org/10.3390/bioengineering10091085.

[2] M. Meina *et al.*, “Heart Rate Variability and Accelerometry as Classification Tools for Monitoring Perceived Stress Levels—A Pilot Study on Firefighters,” *Sensors*, vol. 20, no. 10, p. 2834, 2020, https://doi.org/10.3390/s20102834.

[3] Z. Liu, V. Cascioli, and P. W. McCarthy, “Healthcare Monitoring Using Low-Cost Sensors to Supplement and Replace Human Sensation: Does It Have Potential to Increase Independent Living and Prevent Disease?,” *Sensors*, vol. 23, no. 4, p. 2139, 2023, https://doi.org/10.3390/s23042139.

[4] M. H. Medagedara, A. Ranasinghe, T. D. Lalitharatne, R. A. R. C. Gopura, and G. K. Nandasiri, “Advancements in Textile-Based sEMG Sensors for Muscle Fatigue Detection: A Journey from Material Evolution to Technological Integration,” *ACS Sensors*, vol. 9, no. 9, pp. 4380–4401, 2024, https://doi.org/10.1021/acssensors.4c00604.

[5] W. Sun *et al.*, “A Review of Recent Advances in Vital Signals Monitoring of Sports and Health via Flexible Wearable Sensors,” *Sensors*, vol. 22, no. 20, p. 7784, 2022, https://doi.org/10.3390/s22207784.

[6] I. Martínez-González-Moro, I. Albertus Cámara, and M.-J. Paredes Ruiz, “Influences of Intense Physical Effort on the Activity of the Autonomous Nervous System and Stress, as Measured with Photoplethysmography,” *International Journal of Environmental Research and Public Health*, vol. 19, no. 23, p. 16066, 2022, https://doi.org/10.3390/ijerph192316066.

[7] D. Kim, Y. Lee, J. Jeong, and S. Kim, “Stimulation method and individual health index study for real-time cardiovascular and autonomic nervous system reactivity analysis using PPG signal,” *Biomedical Signal Processing and Control*, vol. 76, p. 103714, 2022, https://doi.org/10.1016/j.bspc.2022.103714.

[8] M. Metshein *et al.*, “Sensor-Location-Specific Joint Acquisition of Peripheral Artery Bioimpedance and Photoplethysmogram for Wearable Applications,” *Sensors*, vol. 23, no. 16, p. 7111, 2023, https://doi.org/10.3390/s23167111.

[9] J. Liu, M. Liu, Y. Bai, J. Zhang, H. Liu, W. Zhu, “Recent Progress in Flexible Wearable Sensors for Vital Sign Monitoring,” *Sensors*, vol. 20, no. 14, p. 4009, 2020, https://doi.org/10.3390/s20144009.

[10] S. K. Longmore, G. Y. Lui, G. Naik, P. P. Breen, B. Jalaludin, and G. D. Gargiulo, “A Comparison of Reflective Photoplethysmography for Detection of Heart Rate, Blood Oxygen Saturation, and Respiration Rate at Various Anatomical Locations,” *Sensors*, vol. 19, no. 8, p. 1874, 2019, https://doi.org/10.3390/s19081874.

[11] C. K. Vinay, V. Vazhiyal and M. Rao, "Design of a non-invasive pulse rate controlled deep vein thrombosis prophylaxis lower limb device," *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5407-5410, 2019, https://doi.org/10.1109/EMBC.2019.8856441.

[12] W. Feng *et al.*, “Predicting physical fatigue in athletes in rope skipping training using ECG signals,” *Biomedical Signal Processing and Control*, vol. 83, p. 104663, 2023, https://doi.org/10.1016/j.bspc.2023.104663.

[13] N. R. Adão Martins, S. Annaheim, C. M. Spengler, and R. M. Rossi, “Fatigue Monitoring Through Wearables: A State-of-the-Art Review,” *Frontiers in Physiology*, vol. 12, 2021, https://doi.org/10.3389/fphys.2021.790292.

[14] P. Liu, Y. Song, X. Yang, D. Li, and M. Khosravi, “Medical intelligence using PPG signals and hybrid learning at the edge to detect fatigue in physical activities,” *Scientific Reports*, vol. 14, no. 1, p. 16149, 2024, https://doi.org/10.1038/s41598-024-66839-8.

[15] P. Faragó, L. Grama, M.-A. Farago, and S. Hintea, “A Novel Wearable Foot and Ankle Monitoring System for the Assessment of Gait Biomechanics,” *Applied Sciences*, vol. 11, no. 1, p. 268, 2020, https://doi.org/10.3390/app11010268.

[16] K. Vinay, K. Nagaraj, H. R. Arvinda, V. Vikas and M. Rao, "Design of a Device for Lower Limb Prophylaxis and Exercise," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 9, pp. 1-7, 2021, https://doi.org/10.1109/JTEHM.2020.3037018.

[17] L. M. S. do Nascimento, L. V. Bonfati, M. L. B. Freitas, J. J. A. Mendes Junior, H. V. Siqueira, and S. L. Stevan, “Sensors and Systems for Physical Rehabilitation and Health Monitoring—A Review,” *Sensors*, vol. 20, no. 15, p. 4063, 2020, https://doi.org/10.3390/s20154063.

[18] M. Tabbakha *et al.*, "PPG Segmentation Using Deep Learning," *2023 3rd International Conference on Mobile Networks and Wireless Communications (ICMNWC)*, pp. 1-5, 2023, https://doi.org/10.1109/ICMNWC60182.2023.10435787.

[19] T. Zhao, Y. Wang, J. Liu, Y. Chen, J. Cheng and J. Yu, "TrueHeart: Continuous Authentication on Wrist-worn Wearables Using PPG-based Biometrics," *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, pp. 30-39, 2020, https://doi.org/10.1109/INFOCOM41043.2020.9155526.

[20] H. Shin, S. Sun, J. Lee and H. C. Kim, "Complementary Photoplethysmogram Synthesis From Electrocardiogram Using Generative Adversarial Network," *IEEE Access*, vol. 9, pp. 70639-70649, 2021, https://doi.org/10.1109/ACCESS.2021.3078534.

[21] S.-H. Liu, R.-X. Li, J.-J. Wang, W. Chen, and C.-H. Su, “Classification of Photoplethysmographic Signal Quality with Deep Convolution Neural Networks for Accurate Measurement of Cardiac Stroke

Volume," *Applied Sciences*, vol. 10, no. 13, p. 4612, 2020, https://doi.org/10.3390/app10134612.

[22] M. Elgendi, Y. Liang, and R. Ward, "Toward Generating More Diagnostic Features from Photoplethysmogram Waveforms," *Diseases*, vol. 6, no. 1, p. 20, 2018, https://doi.org/10.3390/diseases6010020.

[23] L.-P. Yao and Z. Pan, "Cuff-less blood pressure estimation from photoplethysmography signal and electrocardiogram," *Physical and Engineering Sciences in Medicine*, vol. 44, no. 2, pp. 397–408, 2021, https://doi.org/10.1007/s13246-021-00989-1.

[24] B. N. Krupa, K. Bharathi, M. Gaonkar, S. Karun, S. Nath, and M. A. M. Ali, "Multiclass Classification of APG Signals using ELM for CVD Risk Identification: A Real-Time Application," *The 16th International Conference on Biomedical Engineering*, pp. 32–37, 2017, https://doi.org/10.1007/978-981-10-4220-1_7.

[25] E. J. Argüello-Prada and C. D. Castaño Mosquera, "A Machine Learning Approach to Predict Blood Pressure Risk Levels Using Only Maxima and Minima of Photoplethysmogram (Ppg) First and Second Derivatives," *SSRN*, 2024, https://dx.doi.org/10.2139/ssrn.4797403.

[26] H. Sharma, "Heart rate extraction from PPG signals using variational mode decomposition," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 1, pp. 75–86, 2019, https://doi.org/10.1016/j.bbe.2018.11.001.

[27] M. Bolanos, H. Nazeran and E. Haltiwanger, "Comparison of Heart Rate Variability Signal Features Derived from Electrocardiography and Photoplethysmography in Healthy Individuals," *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4289-4294, 2006, https://doi.org/10.1109/IEMBS.2006.260607.

[28] F. Esgalhado, A. Batista, V. Vassilenko, S. Russo, and M. Ortigueira, "Peak Detection and HRV Feature Evaluation on ECG and PPG Signals," *Symmetry*, vol. 14, no. 6, p. 1139, 2022, https://doi.org/10.3390/sym14061139.

[29] M. Z. Suboh, R. Jaafar, N. A. Nayan, N. H. Harun, and M. S. F. Mohamad, "Analysis on Four Derivative Waveforms of Photoplethysmogram (PPG) for Fiducial Point Detection," *Frontiers in Public Health*, vol. 10, 2022, https://doi.org/10.3389/fpubh.2022.920946.

[30] G. Thambiraj, U. Gandhi, U. Mangalanathan, V. J. M. Jose, and M. Anand, "Investigation on the effect of Womersley number, ECG and PPG features for cuff less blood pressure estimation using machine learning," *Biomedical Signal Processing and Control*, vol. 60, p. 101942, 2020, https://doi.org/10.1016/j.bspc.2020.101942.

[31] M. Ghamari, "A review on wearable photoplethysmography sensors and their potential future applications in health care," *International Journal of Biosensors & Bioelectronics*, vol. 4, no. 4, p. 195, 2018, https://doi.org/10.15406/ijbsbe.2018.04.00125.

[32] P. C. -P. Chao, C. -C. Wu, D. H. Nguyen, B. -S. Nguyen, P. -C. Huang and V. -H. Le, "The Machine Learnings Leading the Cuffless PPG Blood Pressure Sensors Into the Next Stage," *IEEE Sensors Journal*, vol. 21, no. 11, pp. 12498-12510, 2021, https://doi.org/10.1109/JSEN.2021.3073850.

[33] J. Phillips, E. Cripps, J. W. Lau, and M. R. Hodkiewicz, "Classifying machinery condition using oil samples and binary logistic regression," *Mechanical Systems and Signal Processing*, vol. 60–61, pp. 316–325, 2015, https://doi.org/10.1016/j.ymssp.2014.12.020.

[34] G. Charizanos, H. Demirhan, and D. İçen, "Binary classification with fuzzy logistic regression under class imbalance and complete separation in clinical studies," *BMC Medical Research Methodology*, vol. 24, no. 1, p. 145, 2024, https://doi.org/10.1186/s12874-024-02270-x.

[35] X. Cui *et al*., "Adaptive LASSO logistic regression based on particle swarm optimization for Alzheimer's disease early diagnosis," *Chemometrics and Intelligent Laboratory Systems*, vol. 215, p. 104316, 2021, https://doi.org/10.1016/j.chemolab.2021.104316.

[36] S. Alqadhi *et al*., "Combining logistic regression-based hybrid optimized machine learning algorithms with sensitivity analysis to achieve robust landslide susceptibility mapping," *Geocarto International*, vol. 37, no. 25, pp. 9518–9543, 2022, https://doi.org/10.1080/10106049.2021.2022009.