Proceedings of CONF-SPML 2026 Symposium: The 2nd Neural Computing and Applications Workshop 2025
DOI: 10.54254/2755-2721/2026.TJ30570

A Personalized Running Fatigue Monitoring System Based on IMU and PPG

Xuejing Chen

Guangzhou Foreign Language School, Guangzhou, China
1243591810@qq.com

Abstract. Nowadays, exercise-related injuries, especially like knee damage, frequently occur due to improper intensity evalu- ation while doing physical activities. In order to solve this issue, this study presents a wearable system for real-time injury-risk prediction. Inertial measurement units (IMUs) and Photoplethys- mography (PPG) sensors are combined so that running data can be collected such as step frequency and HRV characteristics like Root Mean Square of Successive Differences and Standard Deviation of Normal-to-Normal Intervals (RMSSD/SDNN). These data are able to be used to analyze exercise intensity and to train a machine learning model(LSTM). Exercise intensity is divided into three levels: low, moderate, and high. When potentially risky patterns appear, such as unstable step frequency or reduced RMSSD/SDNN, the model can detect them and alert users (e.g., a red LED) to stop or adjust activity. This approach may help reduce injury risk.

Keywords: Fatigue Monitoring, IMU, PPG, LSTM, running

1. Introduction

Running plays an important role in public health and has become a popular method to enhance fitness. However, over- exercise may lead to accumulating fatigue and increased injury risk. Common monitoring methods rely on subjective reports (e.g. questionnaires about joint function, injury his- tory, and perceived stress [1]) or high-cost laboratory devices such as optical motion capture systems, force plates, and Xsens/Noraxon [2]. Although these devices are accurate, their lab-only applicability limits utility. IMU and PPG sensors have been widely used in wearable fatigue detection systems due to their ability to continuously track biomechanical and physiological signals [3-6]. This work designs a low- cost device suitable for daily environments while maintaining accuracy by using an inertial measurement unit (IMU) and photoplethysmography (PPG) for personalized monitoring. Furthermore, a Long Short-Term Memory (LSTM) model is integrated to enable real-time risk monitoring [7-9].

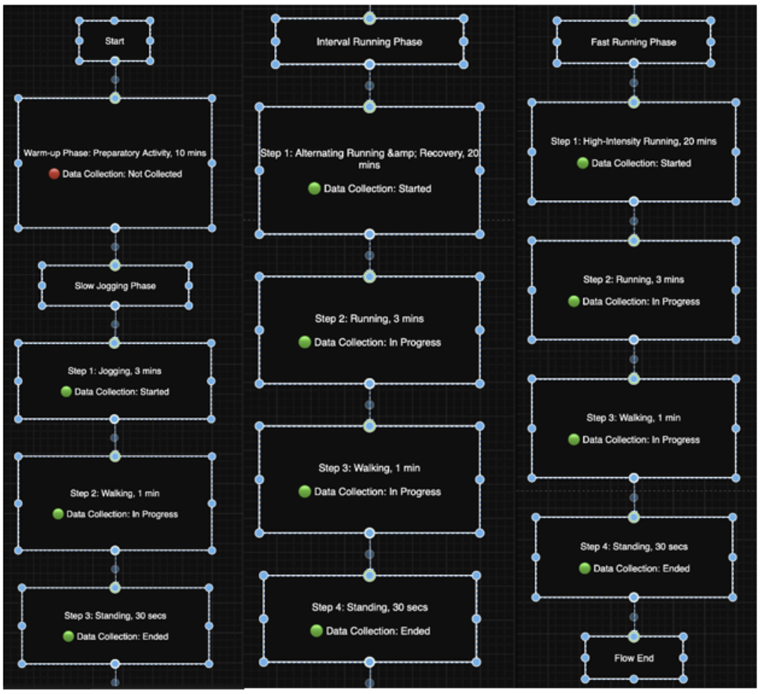
DOI: 10.54254/2755-2721/2026.TJ30570

2. Methodology

2.1. Data acquisition system

This study integrates a Seeed XIAO microcontroller with an LSM6DS3 IMU and a SparkFun Pulse Sensor for PPG acquisition. The IMU records tri-axial acceleration (accelX, accelY, accelZ), and the PPG sensor captures pulse waveforms, synchronized via a common timestamp. The sampling rate is 50 Hz for both modalities. The IMU is positioned slightly be- low the knee, which improves measurement of step frequency and pace [10].

To collect experimental data, three high school students were recruited as participants. Their average ages are 16±1 years old, and their average heights are 170±8 cm, while their average weights are 55±5kg. Each participant performed two complete experimental sessions on separate days. The protocol of each session was designed to induce different fatigue levels and consisted of the following phases. This is a flow diagram of the whole experiment.



(a) Fresh level (b) Middle level (c) Heavy level

Figure 1. Flow diagrams of the whole experiment (three fatigue levels)

Each recording file contained approximately 30,000 rows of synchronized signals, including PPG, accelX, accelY, and accelZ. With six files in total, the dataset comprised roughly 180,000 rows of raw data. This structured protocol ensured that the dataset covered low, moderate, and high fatigue conditions, providing balanced input for subsequent model training and evaluation [8].

2.2. Data preprocessing

The raw signals captured from IMU sensor and PPG sensor might inevitably contain noise that originates from sensor drift and body movement and body difference between various people. Therefore, it is imperative to preprocess before we can train a model [3,11].

DOI: 10.54254/2755-2721/2026.TJ30570

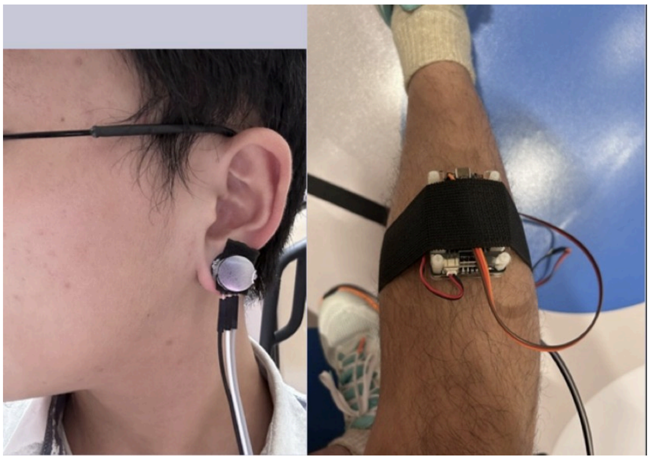


Figure 2. Seeed XIAO integrated with IMU and pulse sensor

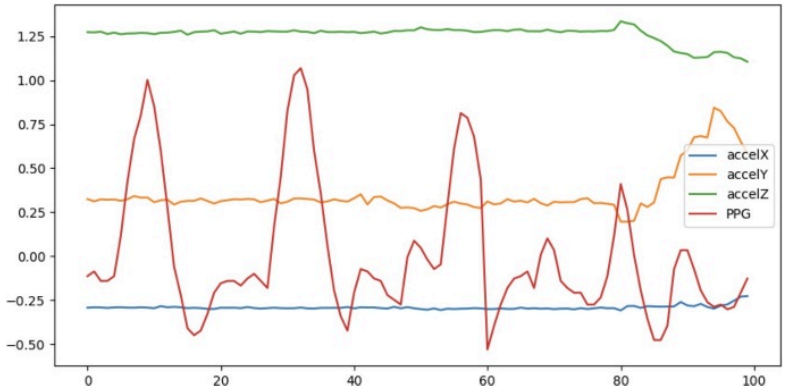


Figure 3. Sample image of raw data

1) Noise Reduction Because signals in PPG and IMU sensor are more likely to be affected by body movement, leading to high frequent noise , which are meaningless and tend to result in misunderstanding of model. Thus, this study use serval metheds including smoothing . More specifically, in the smoothing, this study use moving average filter to reduce the noise. What's more, this study also ensure IMU signals and PPG signals in the same frequency(50HZ) in order to Avoid noise interference caused by time misalignment [12].

2) Normalization Due to differences between individual sensors and physiological characteristics of different subjects, raw signal curves may overemphasize certain features, leading to model misinterpretation. Thus, Z- score normalization is applied to standardize the signals, using the formula:

$$x' = \frac{x - \mu}{\sigma}$$

where x denotes the raw signal value, μ represents the mean of the raw signals, and σ is the standard deviation of the raw signals. This normalization ensures that acceleration and PPG values are on a comparable scale, preventing any single feature from dominating during model training.

3) Segmentation with Sliding Windows This study set window lengths to 2 seconds(100 samples at 50 Hz) so that periods can be covered and maintain the dynamic features. What's more, this study

DOI: 10.54254/2755-2721/2026.TJ30570

also use partial overlap between windows in order to ensure continuity of time series and sample size. The window step size is set to 50, maintaining temporal context while using 50% overlap to ensure continuity and adequate data volume [5].

4) Label Assignment Each sliding window is assigned one of three fatigue level labels, corresponding to different experimental phases: - Fresh level (labeled as 0), - Middle level (labeled as 1), - Heavy level (labeled as 2) [6]. These labels reflect the gradual progression of fatigue, enabling the model to learn the dynamic process of fatigue development over time.

2.3. Model architecture

This study uses Long Short-Term Memory (LSTM) to train a model since this method has advantages to capture dynamic signals in a long period and is able to collect periodic and dynamic sport signals [5,9].

1. Input: Every single input sample corresponds to a win- dow, time length is 2 seconds, and it includes 100 time steps in 50Hz frequency. Every time step has 4 feature channels (accelX, accelY, accelZ, and PPG). Thus, a single input matrix of a sample has a shape of [100, 4]. The label associated with each window corresponds to the fatigue level including fresh (low), middle (moderate), and heavy (high) fatigue.

2. Model training: In order to ensure the balance in three types of samples, this study implements stratified sampling to separate data into training and test subsets with ratio of 8:2. The core of LSTM structure includes a single layer with 64 units, followed by a Dropout layer with a rate of 0.5 to mitigate overfitting. A fully connected layer with 32 units and ReLU activation was then applied for higher-level feature extraction. Finally, a softmax output layer produced the probability dis- tribution over the three fatigue classes. The loss function was

3. categorical cross-entropy, optimized using the Adam optimizer with an initial learning rate of 1×10^{-3}. Accuracy and F1- score were used as the primary evaluation metrics. The batch size is set to 64, and the epochs is 30. What's more, the training process introduces the Early Stopping strategy to stop training when the validation loss did not decrease for several consecutive epochs, thereby preventing overfitting.

4. Output: After training, the test subset can be used to evaluate the model accuracy and F1-score, and the confusion matrix is also drawn in order to analyze fatigue level. In addition, loss curve and accuracy curve are also generated to display the convergence process and the generalization ability of the model.

3. Results and analysis

Training used 30 epochs and batch size 64. Categorical cross-entropy was adopted due to the multi-class nature, and accuracy was the primary metric. Figure 4 shows loss and accuracy over training; validation accuracy stabilizes without evident overfitting.

The final model achieved 67% test accuracy. The dataset comprised 702 windows: 229 (Class 0—Low), 246 (Class 1—Moderate), and 227 (Class 2—High). Each original recording file contained approximately 30,000 rows of syn- chronized PPG and acceleration signals, resulting in six files and a total of about 180,000 rows. The confusion matrix in Fig. 5 indicates better performance on low and high fatigue, with moderate fatigue more frequently misclassified—likely due to transitional physiology and PPG motion artifacts at higher intensity.

DOI: 10.54254/2755-2721/2026.TJ30570

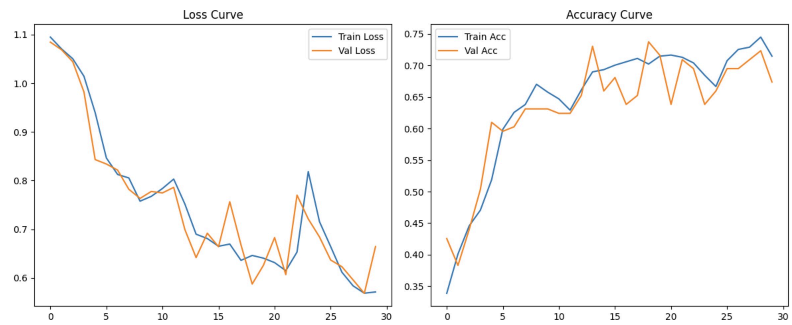


Figure 4. Training and validation loss/accuracy over 30 epochs

```
Confusion Matrix:
[[32 11  3]
 [ 9 37  3]
 [ 4 16 26]]
```

Figure 5. Confusion matrix of test results (3 classes)

4. Future work

This study employed a basic single-layer LSTM architec- ture, which, while capable of capturing temporal dependencies, has limited model capacity. Future work could explore more complex temporal modeling methods, such as multi-layer LSTMs, to further enhance the ability to capture long-term dy- namic dependencies. Additionally, investigating transfer learn- ing and pre-training approaches—leveraging existing large- scale datasets of movement or physiological signals as prior knowledge—could help reduce reliance on large annotated samples.

The dataset used in this study was relatively limited in size and participant diversity, which may affect the model's gener- alizability. Future research should collect data from larger and more diverse populations, covering runners of different ages, genders, and fitness levels to improve model applicability. The experimental design could also be refined by establishing a more continuous and fine-grained fatigue evaluation system, moving beyond the current low-, medium-, and high-level clas- sifications to better reflect the actual progression of exercise- induced fatigue.

This study utilized only PPG and accelerometer signals. Future work could incorporate additional modalities of phys- iological and movement data, such as electrocardiogram (ECG), respiratory rate, gait parameters, and electromyogra- phy (EMG). These signals are closely associated with fatigue states, and their integration may significantly enhance the model's discriminative power. In particular, a multi-modal deep learning framework could enable synergistic modeling of cross-signal features, leading to higher accuracy and ro- bustness under complex exercise conditions.

From an application perspective, future studies may ex- plore integrating this method into wearable devices or sports monitoring platforms to develop real-time fatigue monitoring and early-warning systems. Through edge computing and lightweight model optimization techniques—such as pruning, quantization, and distillation—low-power, low-latency online inference could be achieved. This would provide intelligent support for athlete training, fitness enthusiasts, and exercise prescription in rehabilitation medicine. Furthermore, person- alized modeling approaches could be incorporated to tailor fatigue prediction models to individual runners based on their baseline

DOI: 10.54254/2755-2721/2026.TJ30570

physiological characteristics and exercise habits, thereby improving prediction accuracy and practical utility.

5. Conclusion

Based on a lightweight multi-sensor fusion approach, this study developed a practical system for real-time running fatigue monitoring. The system synchronously captures biome- chanical signals via an inertial measurement unit (IMU) and physiological responses via photoplethysmography (PPG) during running. Experimental results validate that integrat- ing these complementary data sources allows for feasible and interpretable fatigue assessment. Compared to laboratory- based physiological tests or subjective questionnaires, the proposed method offers a cost-effective, continuous, and field- deployable solution.

The current limitations primarily relate to scalability and granularity. The relatively small participant cohort and inherent inter-individual variability may affect generalizability. Further- more, the three-level fatigue classification, while practical, does not fully capture the continuous progression of fatigue.

Future work will focus on expanding the dataset to include runners of diverse fitness levels and demographics to enhance robustness. We also plan to develop a more fine-grained, continuous fatigue evaluation scale. In the long term, we aim to integrate this technology into wearable platforms to provide actionable feedback for runners, coaches, and rehabilitation specialists, supporting personalized training and injury pre- vention.

References

[1] G. B. Wilkerson, M. A. Colston, and C. S. Baker, “A Sport Fit- ness Index for Assessment of Sport-Related Injury Risk, ” Clinical Journal of Sport Medicine, vol. 26, no. 5, pp. 423–428, Dec. 2015. doi: 10.1097/JSM.0000000000000280

[2] A. Arzehgar et al., “Sensor-based technologies for motion analysis in sports injuries: a scoping review, ” BMC Sports Science, Medicine and Rehabilitation, vol. 17, no. 1, Jan. 2025. doi: 10.1186/s13102-025-01063- z

[3] M. H. Md Noor, C. Y. Tan, and M. A. Mohd Ali, “Human activity recognition using inertial measurement unit (IMU) and machine learning, ” Sensors, vol. 20, no. 16, p. 4566, 2020.

[4] A. Reiss and D. Stricker, “Creating subject-independent human activity recognition systems using multiple sensors on different body positions, ” in Proc. IEEE ICDMW, 2012, pp. 1039–1044.

[5] F. Ord´o˜nez and D. Roggen, “Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition, ” Sensors, vol. 16, no. 1, p. 115, 2016.

[6] Y. Wang, J. Liu, and W. Zhang, “Real-time detection of exercise fatigue based on wearable devices and machine learning, ” IEEE Access, vol. 9, pp. 121254–121266, 2021.

[7] A. K. Tripathi, M. Kumar, and D. R. Edla, “A hybrid deep learning model for human fatigue classification using wearable sensor data, ” IEEE Sensors Journal, vol. 22, no. 9, pp. 9033–9041, 2022.

[8] A. K. M. Newaz, M. U. Ahmed, and K. Andersson, “Assessing mental and physical fatigue from facial, speech, and physiological signals: A review, ” Frontiers in Physiology, vol. 13, p. 911255, 2022.

[9] Z. Chen, C. Li, and Z. Sun, “An LSTM-based model for real-time human fatigue detection using multi-sensor data fusion, ” IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1–10, 2022.

[10] R. M. Baig, H. Kim, and M. Lee, “Development of a wearable running gait analysis system using IMU and machine learning, ” IEEE Access, vol. 10, pp. 68472–68484, 2022.

[11] S. Banerjee, R. K. Tripathy, and U. R. Acharya, “Automated detection of fatigue using heart rate variability signals with deep learning models, ” Computers in Biology and Medicine, vol. 141, 105134, 2022.

[12] J. Lee, J. S. Lee, and K. S. Park, “Monitoring of fatigue status using photoplethysmographic signal during dynamic exercise, ” Biomedical Engineering Letters, vol. 10, no. 2, pp. 275–283, 2020.