SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(2), 509.
https://doi.org/10.62617/mcb509

Article

Swimming posture recognition using inertial sensors and CNN-SVM: Unveiling the cellular molecular biomechanics nexus

Chunsheng Xie[1,*], Ling Yin[2], Congying Cui[3]

[1] College of Physical Education, University of Sanya, Sanya 572000, Hainan, China
[2] Sanya Institute of Technology, Sanya 572000, Hainan, China
[3] Faculty of Foreign Trade and Foreign Language, Haikou College of Economics, Haikou 570000, Hainan, China
*** Corresponding author:** Chunsheng Xie, xinhe529@163.com

CITATION

Xie C, Yin L, Cui C. Swimming posture recognition using inertial sensors and CNN-SVM: Unveiling the cellular molecular biomechanics nexus. Molecular & Cellular Biomechanics. 2025; 22(2): 509. https://doi.org/10.62617/mcb509

ARTICLE INFO

Received: 11 October 2024
Accepted: 23 October 2024
Available online: 23 January 2025

COPYRIGHT

Copyright © 2025 by author(s). *Molecular & Cellular Biomechanics* is published by Sin-Chn Scientific Press Pte. Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license. https://creativecommons.org/licenses/by/4.0/

Abstract: The biomechanical mechanisms of swimming involve a number of aspects. The forces exerted by muscles during different swimming postures are crucial. These muscle contractions and relaxations follow specific biomechanical principles. This work aims to develop a swimming posture recognition system based on inertial sensors and a Convolutional Neural Network-Support Vector Machine (CNN-SVM) to improve the accuracy and real-time performance of posture recognition. First, an inertial sensor system to be worn on swimwear is designed to collect three-axis motion data, including acceleration, angular velocity, and magnetometer readings. The collected data are then preprocessed through denoising, normalization, and feature extraction steps to ensure high-quality input data. Next, a Convolutional Neural Network (CNN) is constructed to automatically extract high-level features from the preprocessed sensor data. The CNN model, through multi-layer convolution and pooling operations, effectively captures the spatiotemporal patterns in the motion data, extracting highly distinguishable features for posture recognition. To further improve the model's classification performance, a Support Vector Machine (SVM) classifier is applied based on the CNN model. Specifically, CNN is responsible for feature extraction, while the SVM handles the final posture classification. Cross-validation is used to train and validate the model, assessing its performance. Experimental results show that the model achieves a 95% accuracy rate on the training dataset and maintains an accuracy rate above 93% on the test dataset. The system can accurately and in real-time recognize various swimming postures, including freestyle, breaststroke, backstroke, and butterfly. The recognition accuracy for all four swimming styles exceeds 91%. Understanding these biomechanical mechanisms helps in improving the accuracy of the recognition system. In summary, the proposed method for swimming posture recognition based on inertial sensors and CNN-SVM has significant advantages in accuracy and real-time performance. It allows for better interpretation of the sensor data and more precise identification of different postures. The high accuracy and generalization ability of the proposed system suggest that it can effectively capture and analyze the biomechanical nuances of swimming, providing valuable insights for swimming training and performance evaluation, and opening up new avenues for intelligent sports monitoring. evaluation.

Keywords: biomechanical mechanisms; inertial sensors; convolutional neural network; support vector machine; swimming posture recognition

1. Introduction

Swimming, as a widely participated sport, offers significant health benefits and holds an important place in competitive sports and recreational activities. However, accurately recognizing swimming postures has always been a key challenge in swimming training and performance evaluation [1–3]. Traditional methods of recognizing swimming postures mainly rely on video analysis and the professional judgment of coaches. These methods are not only time-consuming and labor-intensive but also subject to environmental conditions and the subjective factors of observers, making efficient and accurate posture recognition difficult to achieve [4]. With the progress of sensor technology and artificial intelligence, methods based on inertial sensors and deep learning for swimming posture recognition have gradually become a research hotspot [5]. These methods automatically recognize and classify swimming postures by capturing subtle changes in motion data, significantly improving recognition accuracy and real-time performance.

Current swimming posture recognition methods primarily rely on video analysis and manual feature extraction. While these methods have achieved some success in specific scenarios, they often require high computational costs and complex manual configurations. Moreover, they still face challenges in handling complex movement patterns and accommodating individual differences among swimmers. Therefore, the motivation for this work stems from the recognition of the limitations and breakthroughs of existing swimming posture recognition methods. Although current methods have improved posture recognition accuracy to some extent, they still face challenges in handling complex motion patterns and accommodating individual differences among swimmers [6]. Inertial sensors, as portable and efficient data collection tools, can stably and continuously obtain motion data in various environments, providing a new data source for swimming posture recognition. Meanwhile, Convolutional Neural Network (CNN) has demonstrated excellent performance in processing image and spatiotemporal data, making it an ideal choice for feature extraction. Support Vector Machine (SVM), as a powerful classifier, has good generalization ability and can construct optimal classification decision boundaries in high-dimensional spaces. Therefore, combining inertial sensors, CNN, and SVM to construct an efficient swimming posture recognition system is the core objective of this work.

The primary research objective is to develop a swimming posture recognition system based on inertial sensors and CNN-SVM to improve the accuracy and real-time performance of swimming posture recognition. By integrating inertial sensors, CNN, and SVM, a new method for swimming posture recognition is proposed. Compared with traditional methods, this approach has significant advantages in terms of recognition accuracy and real-time performance. Moreover, it exhibits excellent generalization ability, adapting to the recognition needs of different swimmers and various swimming postures. The significance of this work lies in providing a new technological pathway for the application of intelligent sports monitoring devices. With the increasing emphasis on exercise and health, the application prospects of intelligent sports monitoring devices in training and performance evaluation are broad. This work not only provides an efficient method for swimming posture recognition but also offers a reference for posture recognition and monitoring in other sports forms. By improving the accuracy and real-time performance of posture recognition, the proposed system can help swimmers better understand their motion states, promptly adjust training strategies, and enhance performance. Additionally, this system can be applied in swimming teaching and rehabilitation training, providing scientific and reliable data support for coaches and medical personnel.

2. Related work

In recent years, the recognition of sports postures based on inertial sensors has garnered widespread attention in the fields of sports science and intelligent wearable devices. Baniasad et al. studied a running posture recognition method based on an Inertial Measurement Unit (IMU) [7]. They integrated accelerometer and gyroscope data and used a random forest classifier for posture recognition. The findings showed high accuracy across different runners. Cust et al. developed a system for gymnastics movement recognition, and utilized IMU sensor data input into a Long Short Term Memory (LSTM) network to efficiently recognize complex gymnastics movements [8].

In the realm of swimming posture recognition, Xing et al. proposed a multi-sensor fusion-based swimming posture recognition system [9]. While it performed well on static datasets, it faced limitations in dynamic and real-time applications. Similarly, Xu combined accelerometer and magnetometer data with a CNN for swimming posture recognition, achieving high classification accuracy [10], but the real-time application remained a challenge. Hernandez et al. studied motion posture recognition based on deep learning, presenting a model combining CNN and Recurrent Neural Network for real-time recognition of various sports postures [11]. Despite its outstanding performance in multiple sports, its feature extraction capability for swimming posture recognition needed improvement. Taborri et al. explored machine learning algorithms based on inertial sensor data, achieving high recognition accuracy in static environments [12]. Additionally, You et al. addressed data preprocessing issues in posture recognition, proposing a wavelet transform-based denoising method to enhance sensor data quality [13]. Although effective in reducing noise, its high computational complexity limited real-time application. In contrast, Sun et al. used standardization and normalization techniques for preprocessing sensor data, significantly improving data input quality and enhancing model performance [14].

To improve model real-time performance, Zhong et al. developed a lightweight CNN, reducing model parameters and computational complexity for efficient operation on embedded devices [15]. However, this method's generalization ability in practical applications was limited, making it difficult to adapt to different swimmers and various postures.

In summary, although existing research on sports posture recognition based on inertial sensors has achieved certain results across different sports, there remain some shortcomings in the field of swimming posture recognition. First, many methods perform excellently on static datasets but face limitations in actual dynamic and real-time applications. Then, the choice of feature extraction and data preprocessing techniques significantly impacts recognition accuracy and real-time performance, but current research optimization in this regard is insufficient. Additionally, model generalization ability is a pressing issue, with existing methods struggling to adapt to different swimmers and various postures. The innovations of this work are as follows. First, a hybrid model based on inertial sensors and CNN-SVM is proposed. This model leverages the feature extraction capabilities of the CNN and the classification strengths of SVM to improve posture recognition accuracy and real-time performance. Next, an efficient data preprocessing process, including denoising, standardization, and feature extraction, is designed to ensure high-quality data input. Finally, cross-validation is used to train and validate the model, comprehensively evaluating its performance. Experimental results show that this method has promising applications in actual swimming environments. This work not only achieves significant progress in posture recognition accuracy and real-time performance but also proposes a method with excellent generalization ability, adapting to different swimmers and various swimming postures. It offers a new technological pathway for intelligent sports monitoring devices and holds important implications for enhancing swimming training and performance evaluation.

3. Construction of swimming posture recognition model based on inertial sensors and CNN-SVM

3.1. Inertial sensor system design

This work adopts an integrated attitude sensor module with a waterproof treatment suitable for motion data collection in swimming environments [16]. **Figure 1** displays the structure of the sensor module, which internally integrates a high-precision three-axis accelerometer, three-axis gyroscope, and three-axis magnetometer.

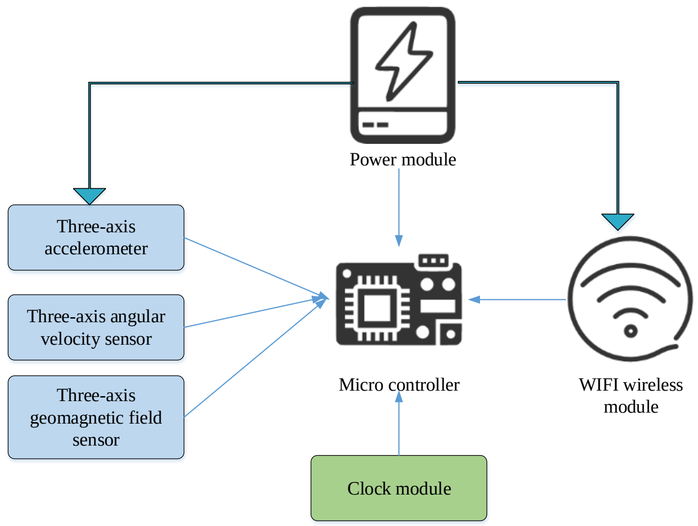

Figure 1. Sensor structure.

Inside the sensor module, a high-performance microcontroller polls the measurement data from each sensor. It combines the data with a dynamic attitude solver and Kalman dynamic filtering algorithm to real-time solve the current motion posture of the sensor module. The posture and time data are uploaded in real-time via a Wireless Fidelity (WIFI) wireless module. When the sensor module remains stationary for more than five minutes, it automatically enters sleep mode to conserve system power; upon detecting motion, it automatically exits sleep mode and returns to normal operation. The entire sensor module is equipped with an external Universal Serial Bus (USB) charging interface and is powered by a built-in power module [17,18].

The sensor module consumes approximately 50 mW of power during normal operation. Equipped with a 1000 mAh lithium-polymer battery, this module can operate continuously for about 20 h, meeting the duration requirements of regular swimming training sessions. To optimize battery efficiency, the sensor module automatically enters sleep mode after five minutes of inactivity and wakes up once movement is detected. This design ensures reliable monitoring during extended training periods without the need for frequent recharging.

To ensure data accuracy and comprehensiveness, this work performs initialization settings on the sensor module. **Table 1** shows the specific parameters.

Table 1. Sensor module parameter settings.

Parameter	Value	Accuracy
Accelerometer Measurement Range	±16 g	0.01 g
Gyroscope Measurement Range	±2000 °/s	0.05 °/s
Data Sampling Frequency	100 Hz	-

The sensor module is installed at the center of the swimmer's waist and back using a strap to ensure the stability and reliability of data collection. To guarantee the real-time accuracy and precision of the data, this work employs a Kalman dynamic filter to process the sensor data. The Kalman filter is a recursive algorithm that optimizes measurement results by estimating the system's state variables and error covariance. The mathematical expressions for the filtering process are shown in Equations (1)–(3):

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(z_k - H\hat{x}_{k|k-1}) \tag{1}$$

$$K_k = P_{k|k-1}H^T(HP_{k|k-1}H^T + R)^{-1} \tag{2}$$

$$P_{k|k} = (I - K_k H)P_{k|k-1} \tag{3}$$

$\hat{x}_{k|k-1}$ represents the predicted state variable, $\hat{x}_{k|k}$ refers to the updated state variable, and K_k is the Kalman gain. $P_{k|k-1}$ is the predicted error covariance, $P_{k|k}$ is the updated error covariance, H is the observation matrix, and R is the measurement noise covariance matrix.

During the data collection process, to obtain high-quality motion labels, this work uses high-speed cameras from Zhongchuang Lianda Company to record synchronized videos. The video frame rate is set to 5000 frames per second. By reviewing the video frame-by-frame synchronized with the sensor data time, the swimmer's actual movements and movement time are determined. The collected data are uploaded in real-time to a server via a WIFI wireless module and stored. Each data collection includes acceleration, angular velocity, and timestamps.

3.2. Data preprocessing

Data preprocessing is an essential step before conducting swimming posture recognition. High-quality data preprocessing can significantly enhance the model's recognition accuracy and stability. The data preprocessing includes denoising, normalization, and feature extraction.

Raw sensor data typically contains various types of noise, such as environmental noise and electronic noise. To eliminate these noises, this work uses a low-pass filter to smooth the data. A cutoff frequency of 10 Hz is selected because it effectively removes high-frequency noise while preserving the key features of swimming motions. The mathematical expression for the filtering process is as follows.

$$y[n] = \frac{1}{N}\sum_{k=0}^{N-1} x[n-k] \tag{4}$$

In Equation (4), $y[n]$ represents the filtered data, $x[n]$ represents the raw data, and N is the filter window size. This work selects a low-pass filter with a window size of 5 to smooth the three-axis acceleration and angular velocity data.

Since the measurement ranges and units of the sensor data differ, standardization is required to eliminate discrepancies between different dimensions. Standardization scales the data to a uniform range, improving the effectiveness of model training. The standardization process is as Equation (5):

$$x' = \frac{x-\mu}{\sigma} \tag{5}$$

x is the raw data, μ is the mean of the data, and σ is the standard deviation of the data. The standardized data has the properties of a mean of 0 and a standard deviation of 1, which facilitates faster convergence of the model.

Feature extraction involves identifying useful feature values from the raw data for posture recognition. To capture the dynamic changes in swimming movements, this work employs a method that combines time-domain and frequency-domain features to extract characteristics from the preprocessed acceleration and angular velocity data. The specific features include Time-domain features: Mean, standard deviation, maximum value, minimum value, and peak value; frequency-domain features: spectral energy, dominant frequency, and band energy extracted through Fast Fourier Transform (FFT). The FFT is represented by Equation (6):

$$X(f) = \sum_{n=0}^{N-1} x[n]\, e^{-jz\pi fn/N} \tag{6}$$

$X(f)$ represents the frequency-domain signal, $x[n]$ is the time-domain signal, and N is the data length. Feature vectors rich in motion information are obtained through feature extraction, providing high-quality input data for subsequent model training and posture recognition.

3.3. CNN construction

CNN has significant advantages in processing time series and image data. To extract high-level features from the preprocessed sensor data, this work designs and trains a CNN-based model for swimming posture recognition [19,20].

The CNN architecture designed here consists of the input layer, convolutional layer, pooling layer, fully connected layer, and output layer. **Figure 2** illustrates the specific structure.

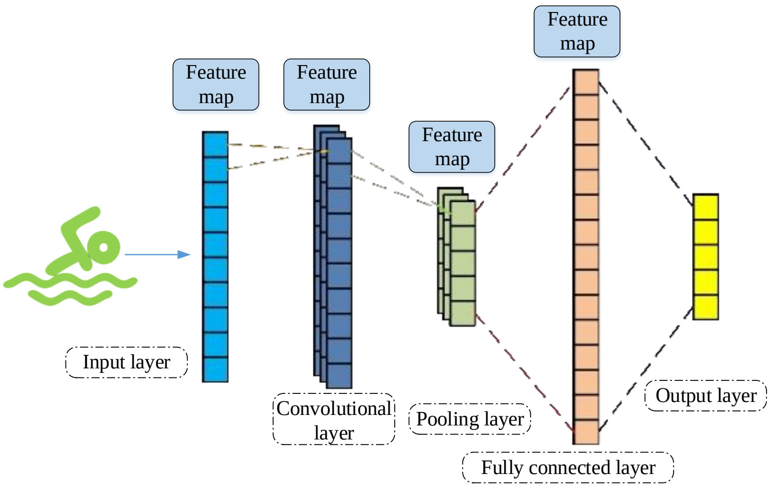

Figure 2. CNN architecture.

The input data consists of multidimensional time series with a shape of (N, T, C). N represents the number of samples, T is the time steps, and C is the number of channels (including data from the three axes of acceleration and angular velocity). Convolution operations are performed to extract the spatiotemporal features from the data. Each convolutional layer contains several convolutional kernels, and the convolution operation is represented by Equation (7):

$$y_{i,j} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x_{i+m,j+n} \times w_{m,n} + b \tag{7}$$

$y_{i,j}$ represents the convolution output, $x_{i+m,j+n}$ is the input data, $w_{m,n}$ is the weight of the convolutional kernel, and b is the bias. Pooling operations are performed to reduce the size of the feature maps while retaining important features and decreasing computational complexity. Common pooling methods include max pooling and average pooling, with max pooling represented by Equation (8):

$$P_{i,j} = \max\left(x_{i+m,j+n}\right) \tag{8}$$

$P_{i,j}$ represents the pooling output. The feature maps obtained from convolution and pooling are flattened into a one-dimensional vector and further processed through fully connected layers for feature extraction. The output of the fully connected layer undergoes a nonlinear transformation via an activation function, commonly using the Rectified Linear Unit (ReLU), represented by Equation (9):

$$y = \max(0, x) \tag{9}$$

The final output layer uses the Softmax activation function to map the feature vector to a probability distribution over the various swimming postures. The Softmax function is as Equation (10).

$$P(y_i) = \frac{e^{z_i}}{\sum_j e^{z_j}} \tag{10}$$

z_i represents the raw output for the i-th class, and $P(y_i)$ denotes the probability for the i-th class.

The model training process consists of three steps: forward propagation, loss calculation, and backward propagation. In forward propagation, the input data passes through the convolutional layers, pooling layers, fully connected layers, and output layers to generate predictions. The loss calculation uses the cross-entropy loss function to compute the difference between the predicted results and the true labels, as shown in the following Equation (11):

$$L = -\sum_{i=1}^{C} y_i \log\left(P(y_i)\right) \tag{11}$$

L represents the loss value, C is the number of classes, y_i denotes the true label, and $P(y_i)$ is the predicted probability. The backpropagation algorithm computes the gradients of the parameters for each layer, and the parameters are updated using the gradient descent method. The parameter update expression is as Equation (12):

$$\theta_{t+1} = \theta_t - \eta \nabla L(\theta_t) \tag{12}$$

θ_t represents the parameters at the t-th iteration, η is the learning rate, and $\nabla L(\theta_t)$ is the gradient of the loss function with respect to the parameters. During model training, the stochastic gradient descent algorithm is employed, along with momentum and learning rate decay strategies, to accelerate model convergence and enhance training stability.

3.4. SVM classifier

This work recruits 20 professional swimmers (age range: 18–30 years old, training experience: 5–10 years) and 15 amateur swimming enthusiasts (age range: 16–40 years old, swimming experience: 2–5 years). These participants possess varying levels of swimming skills to ensure data diversity and representativeness. All participants sign informed consent forms before taking part in the study, ensuring ethical compliance. The data collection process is synchronized using high-speed cameras and worn inertial sensor modules to ensure the accuracy and reliability of the collected data.

SVM is a supervised learning model primarily used for classification and regression analysis. To further enhance the classification performance of swimming posture recognition, this work employs SVM as the final classifier, leveraging the features extracted by CNN [21,22]. **Table 2** outlines the SVM algorithm training process.

Table 2. SVM algorithm training process.

Step	Description
1	Input: Dataset$(x_1,y_1),(x_2,y_2),\cdots,(x_N,y_N)$, where x_i is the feature vector and $y_i \in \{-1,1\}$is the label
2	Initialize: Set regularization parameter CCC and kernel function $K(x_i,x_j)$
3	Formulate the optimization problem:
3.1	$min_\alpha \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j K(x_i,x_j) - \sum_{i=1}^{N}\alpha_i$
3.2	s.t. $0 \le \alpha_i \le C, \sum_{i=1}^{N}\alpha_i\, y_i = 0$
4	Solve for the Lagrange multipliers α: Use an optimization algorithm (e.g., SMO algorithm) to solve for the Lagrange multipliers α_i
5	Calculate weight vector w and bias b:
5.1	$w = \sum_{i=1}^{N}\alpha_i\, y_i \phi(x_i)$
5.2	For any support vectorx_k, compute b
5.3	$b = y_k - \sum_{i=1}^{N}\alpha_i y_i K(x_i,x_j)$
6	Construct the decision function:
6.1	$f(x) = sign(\sum_{i=l}^{N}\alpha_i\, y_i K(x_i,x) + b)$
7	Output: The classification decision function $f(x)$

SVM's fundamental idea is to construct a hyperplane in a high-dimensional space to maximize the margin between different classes, thereby achieving sample classification [23,24]. For linearly separable data, SVM determines the optimal classification hyperplane through the optimization problem described by Equations (13) and (14):

$$min_{w,b}\frac{1}{2}\|w\|^2 \tag{13}$$

$$s.t.\, y_i(w\cdot x_i + b) \ge 1, i = 1,2,\cdots,N \tag{14}$$

w represents the weight vector, b is the bias, y_i is the label for the sample x_i, and N is the number of samples. By solving the Lagrangian dual problem and applying the Karush-Kuhn-Tucker (KKT) conditions, the support vectors and their corresponding decision function can be derived.

For linearly inseparable data, SVM introduces a kernel function to map the data into a high-dimensional space where the data become linearly separable. Commonly used kernel functions include the linear kernel, the radial basis function (RBF) kernel, and the polynomial kernel. This work selects the RBF kernel, and its expression reads:

$$K(x_i,x_j) = \exp\left(-\gamma\left\|x_i - x_j\right\|^2\right) \tag{15}$$

γ is the kernel parameter that controls the mapping effect into the high-dimensional space. In this swimming posture recognition system, the SVM classifier is responsible for classifying the high-level features extracted by the CNN. This work selects a CNN-SVM combined model for swimming posture recognition based on several key considerations. First, CNN demonstrates exceptional performance in extracting spatiotemporal features. Through multiple layers of convolution and pooling operations, local features in images are effectively captured, and higher-level abstract features are progressively built. This is crucial for swimming posture recognition, as swimming movements often involve complex spatial and temporal variations. Through CNN, fine-grained spatiotemporal features are extracted from video sequences, accurately reflecting the athletes' dynamic performance. Second, SVM is particularly effective in classification tasks within high-dimensional spaces [25–27]. Optimal hyperplanes are created by SVM to separate different categories of data, and robust classification capabilities of SVM enable its good performance even when dealing with complex and high-dimensional data [28,29]. This is especially beneficial in scenarios with small sample sizes, which aligns with the limited number of participants here. Therefore, combining the SVM with CNN allows for the utilization of high-dimensional features extracted by CNN, further enhancing classification accuracy. Finally, the CNN-SVM combined model leverages the strengths of both methods to improve the accuracy and real-time performance of swimming posture recognition. Feature extraction from the input data is handled by CNN, while SVM is responsible for the subsequent classification tasks. This division of labor not only boosts the overall performance of the model but also reduces computational complexity to some extent, making real-time applications feasible. The effectiveness of this combined approach provides a solid theoretical foundation and practical support for accurately recognizing swimming postures in this research.

Figure 3 shows the swimming posture recognition model based on the inertial sensor and CNN-SVM.

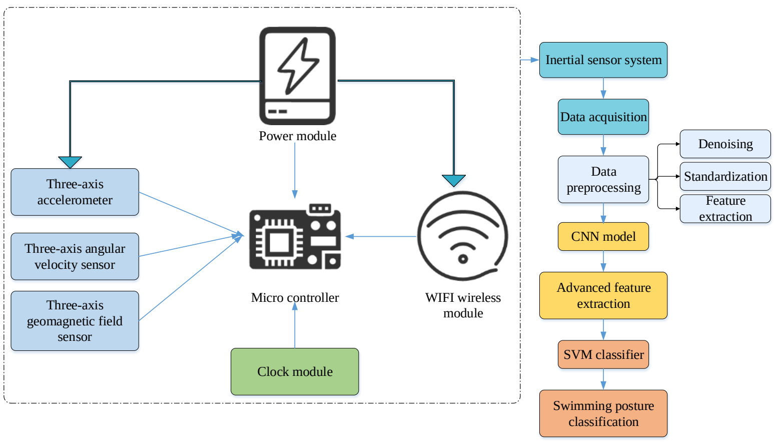

Figure 3. Swimming posture recognition model based on inertial sensors and CNN-SVM.

Figure 4 presents an inertial sensor system designed to be worn on a swimsuit. This system integrates the high-precision three-axis accelerometer, three-axis gyroscope, and three-axis magnetometer, and uploads posture and time data in real-time via a WiFi module. To adapt to the swimming environment, the sensor component is treated for waterproofing, and appropriate measurement ranges and sampling frequencies are selected to ensure the accuracy and reliability of data collection. Next, the collected data are preprocessed, including denoising, standardization, and feature extraction, to guarantee high-quality input data. Specifically, the denoising process uses a Kalman filter to reduce sensor noise, while the standardization step converts the data to a unified scale range. During feature extraction, this work focuses on extracting time-domain and frequency-domain features that reflect the characteristics of swimming actions, enhancing the model's recognition ability. After data preprocessing, a CNN model is constructed to automatically extract high-level features. This model consists of multiple convolutional and pooling layers. By performing layer-by-layer convolution and down-sampling on the input data, the model effectively captures spatiotemporal patterns in the motion data and extracts features with high recognition capability for posture recognition. The filter parameters of the convolutional layers are trained using the backpropagation algorithm, enabling the model to learn optimal feature representations from the data. To further improve classification performance, this work uses an SVM as the classifier on top of the CNN model. Specifically, the CNN model extracts high-level features from the preprocessed sensor data, and the SVM utilizes these features for the final posture classification. By solving a quadratic optimization problem, the SVM finds the optimal classification hyperplane, distinguishing different swimming postures. During training, cross-validation is adopted to evaluate the model's performance, ensuring its generalization ability.

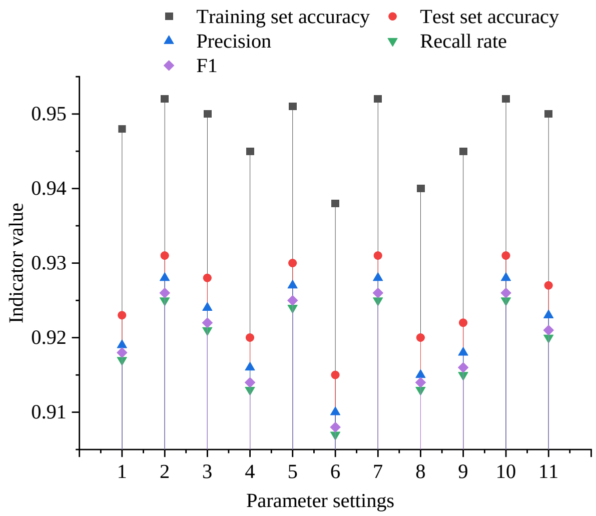

Figure 4. Performance of the model under different parameters.

"1" indicates the number of convolutional layers: 3 layers; "2" indicates the number of convolutional layers: 4 layers; "3" indicates the number of convolutional layers: 5 layers; "4" indicates the number of pooling layers: 2 layers; "5" indicates the number of pooling layers: 3 layers; "6" indicates SVM kernel function: linear kernel; "7" indicates SVM kernel function: RBF; "8" indicates SVM kernel function: polynomial kernel; "9" indicates SVM penalty parameter C: 1; "10" indicates SVM penalty parameter C: 10; "11" indicates SVM penalty parameter C: 100.

4. Experimental results of the swimming posture recognition model based on CNN-SVM

This work conducts a comprehensive evaluation of the performance of the swimming posture recognition system based on inertial sensors and CNN-SVM under different parameter settings. Specifically, it tests the impact of the number of convolutional layers, pooling layers, SVM kernel function types, and penalty parameter C on model performance. **Figure 4** presents the model performance under various parameters. By adjusting the number of convolutional layers, pooling layers, and the SVM kernel function along with the penalty parameter C, variations in the accuracy, precision, recall, and F1 score in the training and testing sets are observed. First, a configuration with three convolutional layers achieves a good balance between accuracy and computational efficiency, resulting in an accuracy of 94.80% on the training set and 92.30% on the test set. This result indicates that three convolutional layers effectively extract features while avoiding overfitting. In contrast, although four convolutional layers provide a higher accuracy on the training set (95.20%), the accuracy on the test set (93.10%) is only slightly higher than that of the three-layer configuration, demonstrating a loss in computational efficiency as model complexity increases. Additionally, the SVM with the RBF kernel function performs best, achieving a test set accuracy of 93.10%. It is found that the RBF kernel handles complex data distributions more effectively than the linear and polynomial kernels, significantly improving the model's generalization capability in classification tasks. Furthermore, when adjusting the SVM's penalty parameter C, a configuration with C set to 10 also achieves good results (test set accuracy of 93.10%), demonstrating its effectiveness in model regularization. In summary, by appropriately selecting the number of convolutional layers and kernel functions, the proposed model strikes a good balance between accuracy and computational efficiency, optimizing overall performance.

To evaluate the system's recognition performance for different swimming strokes, tests are conducted on four swimming styles: freestyle, breaststroke, backstroke, and butterfly stroke. **Figure 5** displays the model's recognition performance for different strokes. It suggests that the model performs best in recognizing breaststroke, achieving both precision and recall rates exceeding 93%, with an F1 score of 93.80%. In contrast, the model's performance for butterfly stroke is slightly lower, with precision and recall rates around 91% and an F1 score of 91.40%. Overall, the model demonstrates ideal recognition performance across different swimming styles, with all metrics exceeding 91%. This indicates that the model has good generalization capabilities and can accurately identify various swimming postures.

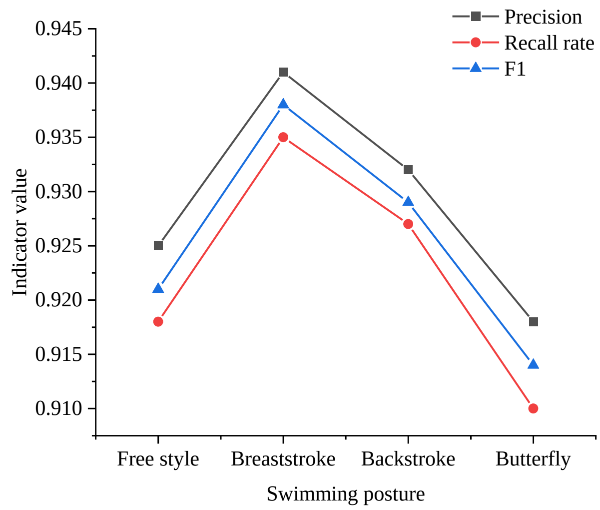

Figure 5. Model recognition performance under different swimming strokes.

This work evaluates the estimation errors for the starting and ending time of movements in the system across different swimming strokes (**Figure 6**). The results show that there are certain differences in time estimation errors for different swimming styles. The average estimation errors for the start and end time of freestyle and breaststroke are both approximately 0.2 seconds, with relatively small standard deviations. This indicates that the time estimation for these two strokes is quite accurate. In contrast, the estimation errors for backstroke and butterfly are relatively larger, especially with higher standard deviations for the end time. This suggests that there is room for further optimization in the time estimation accuracy for these two strokes. Although the time estimation errors for backstroke and butterfly are larger, the average errors are still controlled between 0.1 seconds and 0.35 seconds, which has a limited impact on most training feedback and performance analysis applications. Such a range of errors is sufficient to ensure that coaches and athletes receive timely and practical feedback, allowing for adjustments to technical points during training. Therefore, despite the presence of certain time estimation errors, the system's effectiveness in practical applications remains assured.

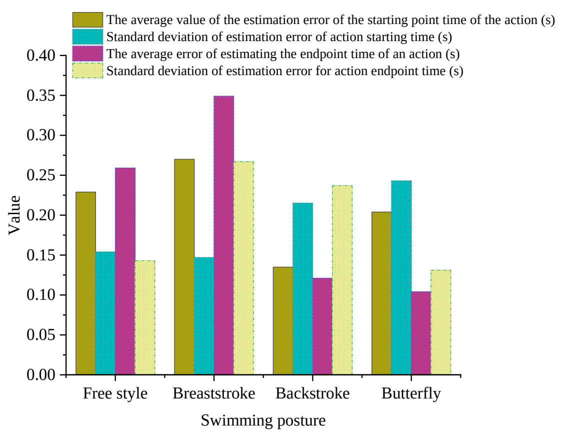

Figure 6. Time estimation errors for the starting and ending movements of the model under different swimming strokes.

This work also tests the system's performance across different populations, evaluating both professional athletes and recreational swimmers to assess its generalization ability. **Figure 7** presents the detailed results. It reveals that the model performs better with athletes, with all metrics exceeding those of recreational swimmers by approximately 1–2 percentage points. This difference may be attributed to athletes having more standardized and stable swimming techniques, which enhances the model's recognition accuracy. Nonetheless, the model's performance with recreational swimmers also reaches a high level, indicating that the system is highly practical and applicable to swimmers of varying skill levels.

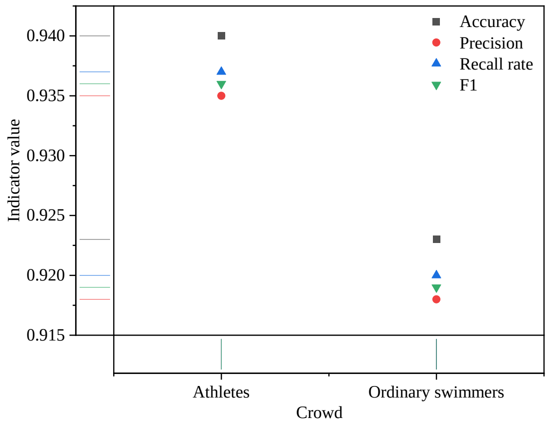

Figure 7. Recognition performance across different populations.

This work also evaluates the system's recognition performance in different environments, including indoor and outdoor swimming pools. **Figure 8** shows the performance across these environments. The data indicate that the model performs slightly better in indoor swimming pools compared to outdoor pools, likely due to the more stable indoor environment and higher quality of sensor data. Nevertheless, the model's performance in outdoor pools is also quite satisfactory, with all metrics exceeding 91%. This suggests that the system is adaptable to different environments, demonstrating strong environmental adaptability and robustness.

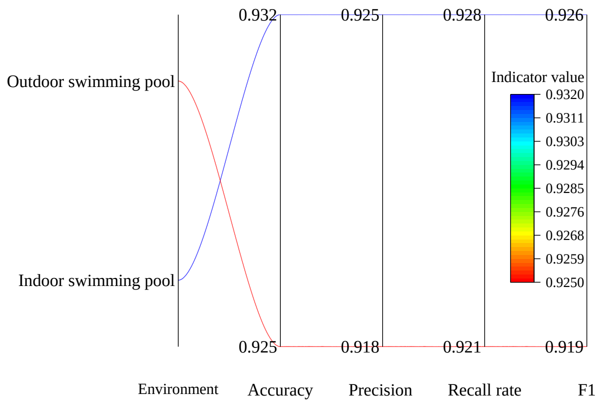

Figure 8. Recognition performance across different environments.

To evaluate the effectiveness of the algorithm presented, this work compares it with other advanced algorithms. **Figure 9** displays the performance of different algorithms in swimming posture recognition. The results indicate that the proposed model outperforms other algorithms across all metrics, particularly in the recognition accuracy for breaststroke and backstroke, where the advantages are especially pronounced. The LSTM model also demonstrates good overall performance, but it is slightly inferior to the proposed model in the recognition accuracy for freestyle and butterfly strokes. Overall, the proposed swimming posture recognition model based on inertial sensors and CNN-SVM exhibits significant advantages in accuracy and real-time performance, making it better suited to meet practical application needs.

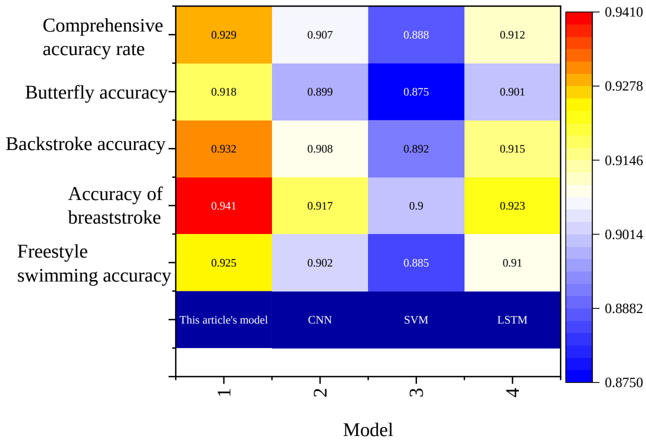

Figure 9. Performance of different algorithms in swimming posture recognition.

Table 3 displays the performance of different algorithms in swimming posture recognition. It can be observed that the method proposed performs exceptionally well across all metrics, particularly in accuracy and precision, surpassing other algorithms. This indicates that the model not only outperforms other methods in terms of recognition accuracy but also exhibits significant advantages in classification stability and reliability.

Table 3. The performance of different algorithms in swimming posture recognition.

Method	**Accuracy (%)**	**Precision (%)**	**Recall rate (%)**
The Proposed Method	92.90%	93.50%	92.80%
LSTM	91.22%	90.00%	91.00%
SVM	88.80%	87.50%	88.00%
CNN	90.66%	89.00%	90.00%

5. Conclusion

The swimming posture recognition method proposed demonstrates significant advantages in accuracy and real-time performance, along with strong generalization capabilities to adapt to various swimmers and different swimming postures. This system provides a new technological pathway for the application of smart sports monitoring devices, playing a crucial role in enhancing swimming training and performance evaluation. The algorithm proposed not only demonstrates its advantages in swimming posture recognition but also holds significant potential for broader applications. In the future, this technology can be extended to other water sports, such as rowing, diving, and water polo, to assist athletes in improving their technical performance. Additionally, in the field of rehabilitation, monitoring patients' postures during hydrotherapy and rehabilitation training can provide physical therapists with important real-time feedback, thereby optimizing personalized rehabilitation programs. These applications will further enhance the social value and practical significance of this work. However, there remains room for improvement regarding the accuracy of time estimation and adaptability to different environments. In particular, the model exhibits relatively larger errors in the estimation of action timing for backstroke and butterfly strokes, necessitating further optimization of the algorithm to enhance precision. Additionally, while the model performs well in both indoor and outdoor environments, the stability of indoor conditions provides higher-quality data for the system. Future work will focus on fine-tuning the algorithm to improve the system's robustness and adaptability under diverse environmental conditions. It will also explore the potential applications of this technology in other sports domains, aiming to achieve broader technological impact and societal benefits.

Author contributions: Conceptualization, CX; methodology, CX; formal analysis, CX and LY; investigation, LY and CC; data curation, CC; writing—original draft preparation, CX; writing—review and editing, CX and LY; supervision, CX. All authors have read and agreed to the published version of the manuscript.

Ethical approval: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Wang Z, Wang J, Zhao H, et al. Using wearable sensors to capture posture of the human lumbar spine in competitive swimming. IEEE Transactions on Human-Machine Systems, 2019, 49(2): 194-205.
2. Cao X, Yan W Q. Pose estimation for swimmers in video surveillance. Multimedia Tools and Applications, 2024, 83(9): 26565-26580.
3. Wang J, Wang Z, Gao F, et al. Swimming stroke phase segmentation based on wearable motion capture technique. IEEE Transactions on Instrumentation and Measurement, 2020, 69(10): 8526-8538.
4. Nandyal S, Dhanyal S S. Video-based yogasan classification for the musculoskeletal disorder using the cervus trail dependent multiclass support vector machine. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 2023, 11(3): 837-855.
5. Hsu Y H, Li C H. Using Machine-Learning Technology to Implement a Nonhardware and Inexpensive Posture Detection System for Analyzing Body Posture Angles in Front Crawl Swimming. Journal of Computers, 2024, 35(1): 131-143.
6. Chen L, Hu D. An effective swimming stroke recognition system utilizing deep learning based on inertial measurement units. Advanced Robotics, 2023, 37(7): 467-479.
7. Baniasad M, Martin R, Crevoisier X, et al. Automatic body segment and side recognition of an inertial measurement unit sensor during gait. Sensors, 2023, 23(7): 3587.
8. Cust E E, Sweeting A J, Ball K, et al. Machine and deep learning for sport-specific movement recognition: A systematic review of model development and performance. Journal of sports sciences, 2019, 37(5): 568-600.

9. Xing H, Liu Y, Guo S, et al. A multi-sensor fusion self-localization system of a miniature underwater robot in structured and GPS-denied environments. IEEE Sensors Journal, 2021, 21(23): 27136-27146.
10. Xu B. Optical image enhancement based on convolutional neural networks for key point detection in swimming posture analysis. Optical and Quantum Electronics, 2024, 56(2): 260.
11. Hernandez V, Suzuki T, Venture G. Convolutional and recurrent neural network for human activity recognition: Application on American sign language. PloS one, 2020, 15(2): e0228869.
12. Taborri J, Palermo E, Rossi S. Automatic detection of faults in race walking: A comparative analysis of machine-learning algorithms fed with inertial sensor data. Sensors, 2019, 19(6): 1461.
13. You N, Han L, Zhu D, et al. Research on image denoising in edge detection based on wavelet transform. Applied Sciences, 2023, 13(3): 1837.
14. Sun J, Xia Y. Pretreating and normalizing metabolomics data for statistical analysis. Genes & Diseases, 2024, 11(3): 100979.
15. Zhong J, Chen J, Mian A. DualConv: Dual convolutional kernels for lightweight deep neural networks. IEEE Transactions on Neural Networks and Learning Systems, 2022, 34(11): 9528-9535.
16. Liu Q. Aerobics posture recognition based on neural network and sensors. Neural Computing and Applications, 2022, 34(5): 3337-3348.
17. Jiang W. Application of motion trajectory recognition based on remote sensing image optical processing in optimizing swimming training schemes. Optical and Quantum Electronics, 2024, 56(2): 264.
18. Powell L, Polsley S, Casey D, et al. The real-time classification of competency swimming activity through machine learning. International journal of aquatic research and education, 2023, 14(1): 6.
19. Giulietti N, Caputo A, Chiariotti P, et al. Swimmernet: Underwater 2d swimmer pose estimation exploiting fully convolutional neural networks. Sensors, 2023, 23(4): 2364.
20. Dong Z, Wang X. An improved deep neural network method for an athlete’s human motion posture recognition. International Journal of Information and Communication Technology, 2023, 22(1): 45-59.
21. Wang H, Shi J, Luo X. Swimmer’s posture recognition and correction method based on embedded depth image skeleton tracking. Wireless Communications and Mobile Computing, 2022, 2022(1): 8775352.
22. Wang G, Zeng X, Lai G, et al. Efficient subject-independent detection of anterior cruciate ligament deficiency based on marine predator algorithm and support vector machine. IEEE Journal of Biomedical and Health Informatics, 2022, 26(10): 4936-4947.
23. Herzog W, Schappacher-Tilp G. Molecular mechanisms of muscle contraction: A historical perspective. J Biomech. 2023;155:111659. doi:10.1016/j.jbiomech.2023.111659
24. Matusovsky OS, Mayans O, Szczesna-Cordary D. Molecular mechanism of muscle contraction: new perspectives and ideas. Biomed Res Int. 2015;2015:694345. doi:10.1155/2015/694345
25. Bu A, Afghah F, Castro N, et al. Actuating Extracellular Matrices Decouple the Mechanical and Biochemical Effects of Muscle Contraction on Motor Neurons. Adv Healthc Mater. Published online November 10, 2024. doi:10.1002/adhm.202403712
26. Johan L. van Leeuwen, Andrew A. Biewener; Biomechanics of swimming and flight. J Exp Biol 15 January 2008; 211 (2): 163. doi: https://doi.org/10.1242/jeb.015651
27. Vilas-Boas, J. P., & Sanders, R. (2023). Swimming biomechanics: an editorial. Sports Biomechanics, 22(12), 1529 – 1531. https://doi.org/10.1080/14763141.2022.2098640
28. Zumbaugh MD, Johnson SE, Shi TH, Gerrard DE. Molecular and biochemical regulation of skeletal muscle metabolism. J Anim Sci. 2022;100(8):skac035. doi:10.1093/jas/skac035
29. Constantin-Teodosiu D, Constantin D. Molecular Mechanisms of Muscle Fatigue. Int J Mol Sci. 2021;22(21):11587. Published 2021 Oct 27. doi:10.3390/ijms222111587