A Multi-Class Muscle Fatigue Recognition Model for Athletes Using KPCA Dimensionality Reduction and SVM Classification on Multi-Channel sEMG Signals

Xingping Chu, Dongqin Huang, Yan Yi*
College of Sports and Health, Jiangxi University of Chinese Medicine, Nanchang, 330004, China
Email: chuxingping2021@163.com, huang9y012@126.com, chu2025@126.com
*Corresponding author

Keywords: muscle fatigue, sEMG, kernel principal component analysis (KPCA), support vector machine (SVM), feature dimensionality reduction

Received: July 24, 2025

Muscle fatigue is an inevitable phenomenon in athletic training, and its accurate assessment is crucial for preventing injuries. To address the issues of high redundancy in multi-channel surface electromyography (sEMG) features and the limitations of linear dimensionality reduction methods, this study proposes a muscle fatigue recognition model based on Kernel Principal Component Analysis (KPCA) and Support Vector Machine (SVM). sEMG signals were collected from 12 subjects performing sustained contractions. A comprehensive set of 48 time-domain, frequency-domain, and nonlinear features was extracted. KPCA was employed for nonlinear dimensionality reduction before classification. The resulting features were fed into an SVM classifier to distinguish between three fatigue states: relaxed, transitional fatigue, and fatigued. The model was evaluated using 10-fold cross-validation. Results demonstrated that the KPCA-SVM combination achieved the highest performance, with an average recognition accuracy of 91.5%, precision of 0.91, recall of 0.93, and F1-score of 0.92, outperforming other combinations of dimensionality reduction methods (MI, PCA) and classifiers (FLDA, KNN). This method provides an effective tool for the objective assessment of muscle fatigue in athletes.

Povzetek: Študija predstavi model KPCA + SVM za prepoznavanje treh stanj mišične utrujenosti iz sEMG, ki z zmanjšanjem odvečnih značilk doseže okoli 91,5 % natančnost in prekaša primerjane pristope.

1 Introduction

Athletes get good grades is the premise of the need for a lot of training, and a lot of training will increase the injury probability of athletes, because with the increase of training intensity and time will produce muscle fatigue, serious will cause muscle damage [1]. Hence need to be testing technology service athlete's fatigue, avoid the occurrence of muscle damage. But in recent years, (multi-channel sEMG surface electromyography, sEMG) because of its noninvasive, convenient detection merits attention in evaluation of muscle fatigue [2]. Are produced by the muscle contraction multi-channel sEMG weak electrical signals of contains many useful biological information, biological medicine, sports medicine and rehabilitation medicine and other fields play an extremely important role in [3].

Moniri et al. [4] developed advanced techniques for real-time forecasting of sEMG features related to trunk muscle fatigue using machine learning, demonstrating the potential of sEMG in fatigue analysis. From then on, the development of sEMG equipment constantly updated, promoted the study of sEMG signal. American Delsys of wireless sEMG acquisition instrument [5] the most widely used. It includes 16 channel wireless sensor with three degrees of freedom of acceleration sensor, small volume and easy to carry, up to 5 h of battery life. Xu et al. [6] utilized a convolutional neural network for sEMG-based feature prediction, providing a general model for sports activity analysis. In a study on pilot neck fatigue assessment, Rampichini et al. [7] achieved a high recognition accuracy of 91.25% by employing a hybrid optimization algorithm to tune a Gaussian process model. This work demonstrates the potential of advanced machine learning models in fatigue recognition, which motivates our exploration of KPCA and SVM for sEMG-based fatigue analysis in athletes.

However, while these methods demonstrate promising results, they often rely on direct use of multi-channel sEMG feature parameters, which may exhibit high correlation and redundancy. This can lead to increased computational complexity and reduced classification accuracy in fatigue models [8]. To address these limitations, this study introduces a novel approach that combines kernel principal component analysis (KPCA) and support vector machines (SVM). KPCA effectively reduces feature dimensionality by capturing nonlinear relationships in sEMG signals, while SVM ensures robust classification performance even with limited data. This method not only enhances the efficiency of fatigue detection but also avoids the extensive data requirements and computational overhead associated with deep learning models.

As mentioned above, the direct use of the characteristic parameters of multi-channel sEMG, although also has good recognition effect, but has a great deal of correlation between characteristic parameters and redundancy, without considering the complexity of the algorithm, often affect the fatigue model classification recognition rate [8]. This article in order to solve this problem, in the original feature parameters under the premise of no loss of the original information, using different feature dimension reduction method, remove redundant information, a new feature set, the use of classifier inspection classification effect, from which identify the most effective a classification model of fatigue.

Compared to existing studies that often directly use raw or high-dimensional sEMG feature sets [8], the main contributions of this paper are threefold: (1) It systematically compares the effectiveness of three distinct feature dimensionality reduction techniques (MI, PCA, and KPCA) for muscle fatigue recognition, highlighting the importance of addressing feature redundancy and nonlinearity; (2) It proposes a novel fatigue recognition framework that synergistically combines KPCA for nonlinear feature extraction and SVM for robust classification, demonstrating that this combination effectively captures the complex patterns in sEMG signals associated with different fatigue states while maintaining computational efficiency; (3) The proposed KPCA-SVM model achieves a superior average recognition rate of 91.5% for three-state fatigue classification, offering a practical and high-accuracy solution that avoids the large data requirements of deep learning models, making it particularly suitable for scenarios with limited training data.

The primary objectives of this study are: (1) to develop and evaluate a framework for muscle fatigue recognition by comparing the effectiveness of different feature dimensionality reduction techniques (MI, PCA, KPCA) combined with various classifiers (FLDA, KNN, SVM); (2) to specifically investigate the performance of the KPCA-SVM model in classifying three states of muscle fatigue (relaxed, transitional, fatigued) from multi-channel sEMG signals; (3) to validate the proposed model's accuracy and efficiency against other common approaches.

2 The characteristics of the multi-channel sEMG analysis method

2.1 The time domain characteristics analysis method

Time domain features of multi-channel sEMG can reflect the change of the signal in the time dimension, from the perspective of time, will be understood as a multi-channel sEMG time-varying function, has the advantages of simple, rapid computation [9]. Common characteristics such as RMS values, integral electrical values, the average absolute value and the zero rate of the mean, variance and amplitude cube [10].

1) The Root Mean Square Value (Root Mean Square, RMS)

Root means square value indicates change in unit time, multi-channel sEMG can respond muscle activity. As shown in Equation (1):

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{1}$$

Type, $x_i (i = 1,2, \dots, N)$ is the time series of the signal length *N* to select the number of frames.

2) The Integral Electrical Values (Integrated EMG, iEMG)

Integral electrical value indicates that the activities of the muscle fibers of charge level, reaction of multi-channel sEMG amplitude changes with exercise. As shown in Equation (2):

$$iEMG = \sum_{i=1}^{N} |x_i| \tag{2}$$

Type, $x_i (i = 1,2, \dots, N)$ is the time series of the signal length *N* to select the number of frames.

3) The Average Absolute Value (Mean Absolute Value, MAV)

The average absolute value used to identify the muscles. As shown in Equation (3):

$$MAV = \frac{1}{N}\sum_{i=1}^{N} |x_i| \tag{3}$$

Type, $x_i (i = 1,2, \dots, N)$ is the time series of the signal length *N* to select the number of frames.

4) The Zero Rate (Zero Crossing Rate, ZCR)

Zero crossing rate refers to the sEMG symbols change ratio, per unit time is an important index of the time-domain analysis. As shown in Equation (4):

$$ZCR = \sum_{i=1}^{N-1 \sum(-x_i x_{i-1})} sgn \tag{4}$$

Type, $x_i (i = 1,2, \dots, N)$ is the time series of the signal length *N* to select the number of frames.

5) Variance (Variance, VAR)

Variance is used to measure the parameters of the degree of discrete random variables. As shown in Equation (5):

$$VAR = \frac{\sum_{i=1}^{N} (x_i - \overline{x_i})}{N} \tag{5}$$

Type, $x_i (i = 1,2, \dots, N)$ is the time series of the signal length *N* to select the number of frames.

6) The Mean Amplitude Cube (Amplitude Cubic scheme, ACM)

Average amplitude cube can response in time domain amplitude change rule of multi-channel sEMG. As shown in Equation (6):

$$ACM = \frac{1}{N}\sum_{i=1}^{N} x_i^3. \tag{6}$$

Type, $x_i (i = 1,2, \dots, N)$ is the time series of the signal length *N* to select the number of frames.

2.2 Analysis of characteristics of frequency domain method

Methods of frequency domain feature of electricity is will signal through the fast Fourier transform (FFT), according to the power spectrum of signal or frequency spectrum is analyzed [11]. Commonly used indicators have mean power frequency and spectral moments and the median frequency.

1) The Mean Power Frequency (Mean Power Frequency)

Mean power frequency with muscle fatigue deepening decline, the reason is that lower pH, results in the decrease of muscle fiber conduction velocity. As shown in Equation (7):

$$MPF = \frac{\int_0^\infty f \times PSD(f)df}{\int_0^\infty PSD(f)df} \tag{7}$$

Type of f sEMG signal frequency, $PSD\ (f)$ for sEMG signal power spectral density function.

2) Spectral Moment (Spectral Moment, SM)

Spectral moment is Dimitrov put forward a set of parameters of muscle fatigue, two order spectral moment is often used to describe the change of high frequency. As shown in Equation (8):

$$SM_2 = \int_0^\infty f^2\, PSD(f)df. \tag{8}$$

Type of f sEMG signal frequency, $PSD\ (f)$ for sEMG signal power spectral density function.

3) The median Frequency (Media Frequency, MF)

Found in the study of muscle fatigue estimation, the low frequency component of the multi-channel sEMG increases with muscle fatigue, quantitative coefficients are part of the representative indicators is the median frequency. As shown in Equation (9):

$$MF = \frac{1}{2}\int_0^\infty PSD(f)df \tag{9}$$

Type of f sEMG signal frequency, $PSD\ (f)$ for sEMG signal power spectral density function.

2.3 Time and frequency domain characteristics analysis method

Time domain analysis was conducted on the signal in time domain analysis, studies the changing relation between time and signal amplitude; Frequency domain analysis is in the frequency domain analysis of signals, the frequency change [12]. While the aforementioned time-frequency analysis methods are foundational, recent advancements have introduced more sophisticated techniques for handling non-stationary signals like sEMG. Wavelet Transform (WT), particularly the Discrete Wavelet Transform (DWT), has gained prominence for its multi-resolution analysis capability, allowing for feature extraction across different frequency bands at varying temporal resolutions [13]. Although the current study focuses on standard time, frequency, and time-frequency features for model simplicity and comparability, future work will explore integrating wavelet-based features to further enhance the model's performance in characterizing muscle fatigue dynamics.

2.3.1 The short-time Fourier transform short-time

Fourier transform is a transform and Fourier transform, can be implemented by the Fourier transform, is one of the more commonly used a time-frequency analysis method . $STFT$ principle is to put a long non-stationary signal as a superposition of countless short time stationary signal. For signal $x(t)$, as well as with a moving time window is decomposed into countless time length of short time signal, so during this period can take it as a stationary signal, and then to the Fourier transform, signal time and frequency domain information.

Define functions $STFT(w,\tau)$, said in the center of the window function of τ, to transform the function of spectrum as a result, ss shown in Equation (10):

$$STFT(w,\tau) = \int_{-\infty}^{+\infty} [x(\tau)w(t-\tau)]\, e^{-jwt} d\tau\ \cdot \tag{10}$$

In the short time Fourier transform, each of the different time, after the transformation can be a different spectrum, the frequency spectrum is combined time-frequency distribution. The window length determines the time of spectrum, frequency resolution, the longer the length of the window, the Fourier transform of frequency domain resolution is higher, but the lower temporal resolution; on the other hand, the shorter the window, the Fourier transform of frequency domain resolution is lower, but the higher time resolution. Therefore, in practice need to according to the situation, between the time resolution and frequency resolution. Although the short time Fourier transform algorithm is very simple, can deal with non-stationary signal, applied to various fields, but the adaptive ability is poor.

2.3.2 The Winger-Ville distribution

In 1932, Winge presents a time-frequency analysis method, and applied to the field of quantum mechanics, later Ville signal analysis, introduced the method, formed the Winger-Ville distribution.

Signal, a Winger-Ville distribution can be defined as:

$$W_{x,y}(t,w) = \int_{-\infty}^{+\infty} x\left(t+\frac{\tau}{2}\right) y^*\left(t-\frac{\tau}{2}\right) e^{-jwt} d\tau \tag{11}$$

Type 11 is also called the bilinear time-frequency analysis method, can also be represented as a signal of $x(t)$ from Winger-Ville distribution form, as shown in Equation (12):

$$W_x(t,w) = \int_{-\infty}^{+\infty} x\left(t+\frac{\tau}{2}\right) x^*\left(t-\frac{\tau}{2}\right) e^{-jwt} d\tau \tag{12}$$

Based on the definition of Winger-Ville distribution, frequency spectrum of signal t any time, on time t, t time signal multiplication, all results are superimposed and then do the Fourier transform, because the signal is not a single component, can produce interaction among computing, form the cross terms. When the Winger-Ville distribution crossover component is bigger when a certain cross interference, often cover a useful time and frequency information. In order to avoid the Winger-Ville distribution in the time-frequency plane cross interference, the introduction of window function, the $x(t)$ can be added window, the resulting Winger-Ville distribution is defined as shown in Equation (13):

$$C_Z = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} z\left(u+\frac{\tau}{2}\right) z^*\left(u-\frac{\tau}{2}\right)\psi(t-u,\tau)e^{-jwt}dud\tau \tag{13}$$

Type of window function, $h(\tau)$ is equal to the difference between in the frequency domain to the Winger Ville distribution of smoothing.

Compared with the short time Fourier transform, Winger-Ville distribution can accurately identify signals is a single component or component, get better time-frequency information. But Winger-Ville distribution of different frequency components will cause the interference of cross terms, although after add window will have certain inhibition on the cross-term interference, but cannot completely eliminate its existence.

2.3.3. Choi-William's distribution

Choi-William's distribution is insufficient for Winger-Ville distribution problem put forward the improved algorithm. In order to complete the related domain filtering of fuzzy function, eliminate the cross-term interference formed by mutual fuzzy function, Choi and Williams with the method of the kernel function is proposed. As shown in Equation (14):

$$C_Z = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} z\left(u+\frac{\tau}{2}\right) z^*\left(u-\frac{\tau}{2}\right)\psi(t-u,\tau)e^{-jwt}dud\tau \tag{14}$$

Type of $\psi(t-u,\tau)$ for Cohen class time-frequency distribution kernel function.

2.4 Nonlinear characteristics analysis method

The contraction of muscles in different state, the motor nerve cell number in time, space and there exist certain differences, therefore sEMG has obvious nonlinear characteristics [14]. Therefore, put forward a new method of analysis, complexity and nonlinear analysis of time series method is an office value, the greater the signal contains information is more complex. In this paper, extraction of multi-channel sEMG approximate office as the research of nonlinear characteristic parameters.

Approximate Entropy (ApEn), introduced by Pincus [15], is a nonlinear dynamic parameter used to quantify the regularity and unpredictability of a time series. A higher ApEn value indicates greater complexity and irregularity in the signal. It has been widely applied in physiological signal analysis, including sEMG, to assess the complexity changes associated with muscle fatigue [16]. The calculation of ApEn involves comparing the similarity of patterns within the signal for increasing pattern lengths (m and m+1).

3 The analysis method of the feature dimension reduction and classification model

3.1 Feature dimension reduction

This article is based on three channels sEMG signal, a total of 48 d characteristic parameters can be extracted. High-dimensional data contains a large number of redundant features, the classification of the redundant features of the follow-up effect and increase the complexity of the operation, and high dimensional data can lead to "fitting" and "dimension disaster" problem, reduce the performance of the classifier. Extraction of effective feature parameters is the key to the classification of fatigue, so need to dimension reduction of feature set, the classification model recognition rate [17]. The overall workflow of the proposed muscle fatigue recognition model is illustrated in Figure 1. It outlines the key stages: multi-channel sEMG signal acquisition, feature extraction, dimensionality reduction, and classification, providing a clear roadmap for the methodology described in this section.

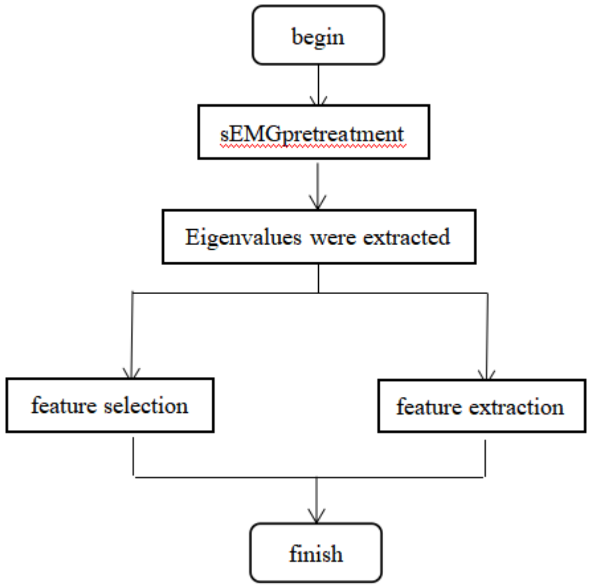

Figure 1: The whole algorithm flow chart.

3.1.1 Feature selection

Feature selection is according to an evaluation criterion, from centralized to choose some of the most effective characteristics, thus reducing the process of feature dimension [18]. The Dash gives the framework of feature selection, the following is shown in Figure 2.

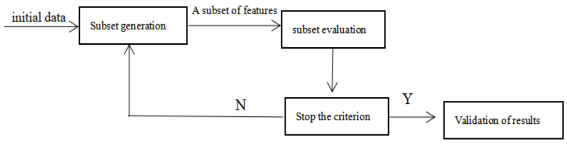

Figure 2: The framework of feature selection.

Mutual information (MI) measurement is one of the important measures of feature selection, without considering the distribution of sample in advance, and can deal with linear or nonlinear random variables [19]. MI value is larger, can explain the interrelation between the two variables, the greater the smaller conversely.

MI calculation equation is shown below:

$$I(U;V) = \sum_{u\in U}\sum_{v\in V} p\,(u,v)\,log\,\frac{p(u,v)}{p(u)p(v)} \tag{15}$$

Type of $p(u,v)$ is the joint probability distribution function of u and v, $p(u)$ and $p(v)$, respectively is the edge of the u and v probability the function.

3.1.2 Feature extraction

1) The principal component analysis method

Principal component analysis is a commonly used linear feature extraction method, the linear relationship for a set of possible data, through the orthogonal transformation, principal component analysis to data mapped to a low-dimensional subspace of linear independence, the basic principle of the algorithm are: Original each dimension data of the covariance matrix as well as the eigenvalue and eigenvector, the characteristic value from big to small, retain data high contribution rate, ignore the data of low contribution rate, so as to realize the feature dimension reduction and random vector x_i, $i = 1,2, \ldots M$ which M as the total number of samples, N is the dimension of data [20]. The sample can be X_{11} To form $M \times N$ matrix, namely $X = \begin{Bmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{M1} & \cdots & x_{MN} \end{Bmatrix}$.

In the practical implementation of KPCA in this study, the Radial Basis Function (RBF) kernel was employed. The kernel parameter σ was determined through a grid search within a range of [0.01,1], optimizing for the subsequent classification performance. The number of nonlinear principal components was selected to ensure a cumulative variance contribution rate of over 95%, effectively reducing the feature dimensionality from 48 to 10 while preserving the most critical discriminatory information.

According to matrix $X_{m\times n}$ column, each column of the mean, calculate the average each column minus the column, as shown in the Equation (16), with each column number minus the corresponding column mean, get new matrix $B = \begin{Bmatrix} \overline{x_{11}} & \cdots & \overline{x_{1N}} \\ \vdots & \ddots & \vdots \\ \overline{x_{M1}} & \cdots & \overline{x_{MN}} \end{Bmatrix}$.

$$\bar{x} = x_i - \frac{1}{M}\sum_{i=1}^{M} x_i \tag{16}$$

And then according to the B^T type 16 obtains its transpose matrix, computing characteristics of covariance matrix C, as shown in the Equation (17):

$$C = B^T \times B \tag{17}$$

Then calculate the eigenvalues of the covariance matrix C, β_i, λ_i and eigenvector, the eigenvalue λ_i from big to small before order, according to the contribution rate to select the appropriate k composition, to constitute the new data. K before features vector matrix such as type of (18):

$$W = (\beta_1, \beta_2, \ldots, \beta_k) \tag{18}$$

Refactoring data dimension reduction to k-dimensional $P = W \times X$.

2) The kernel principal component analysis method.

Some fatigue characteristic value may show the nonlinear relationship between fatigue and, therefore, we need a new way to solve the problem. Kernel principal component analysis is an improved algorithm based on principal component analysis, is a kind of nonlinear feature extraction method, solve the problem of principal component analysis is unable to realize, compared with the principal component analysis greatly reduced the amount of calculation, and provides a better recognition performance [21]. Its basic principle is as follows: First of all, through nonlinear mapping, transform the data into a high-dimensional nonlinear space, and then using linear PCA mapped to a low dimensional space. In this study, the kernel component analysis (KPCA) employs radial basis function kernels. The kernel parameter σ was optimized through a grid search combined with 5-fold cross-validation, with the search range set to [0.1, 0.5]. The final selected parameter σ=0.1 achieved a cumulative variance contribution rate exceeding 95%, while reducing the feature dimensionality from 48 to 10.

For the n input sample data $x_k(k = 1,2,\ldots,n)$, $x \in R^N$, introducing the nonlinear mapping function Φ, transform data x_k is $\Phi(x_k)$, then in the new feature space, covariance matrix for C.

Using the iterative method to solve eigenvalue $\lambda(\lambda > 0)$ and characteristic vector $V(V \neq 0)$, the specific Equation (19):

$$\lambda V = CV \tag{19}$$

And V by $\Phi(x)$ linear said, as shown in Equation (20):

$$V = \sum_{j=1}^{n} \alpha_j\,\Phi(x_j), j = 1,2,\ldots,n \tag{20}$$

Type in α_j for the equation coefficient, Equation (21) left 19 times $\Phi(x_k)$:

$$\lambda(\Phi(x_k)\cdot V) = \Phi(x_k)\cdot CV, k = 1,2,\ldots,n \tag{21}$$

Define a $n \times n$ matrix K, the details are shown in Equation (22):

$$K_{ij} = \Phi(x_i)\Phi(x_j), i,j = 1,2,\ldots,n \tag{22}$$

Equation (23), for solving the nuclear matrix K of non-zero eigenvalue λ and eigenvector alpha α.

$$n\lambda\alpha = K\alpha \tag{23}$$

Centralized eigenvector V, sample data $\Phi(x)$ in the mapping Equation (24):

$$h_i(x) = (V, \Phi(x)) = \sum_{i=1}^{n} \alpha_i\,\Phi(x_i) \tag{24}$$

$h_i(x)$ to $\Phi(x)$ of the first k nonlinear principal component vectors.

In addition to dimensionality reduction techniques like PCA and KPCA, the selection of appropriate features from the sEMG signals is crucial for accurate muscle fatigue detection. In this study, we focus on Root Mean Square (RMS), Mean Absolute Value (MAV), and Zero Crossing Rate (ZCR) as primary features. RMS is chosen because it quantifies the amplitude variation of the sEMG signal, which is directly related to muscle contraction intensity and tends to decrease with fatigue. MAV measures the overall signal strength by averaging the absolute values of the signal, making it sensitive to changes in muscle activity during prolonged exercise. ZCR captures the frequency characteristics of the signal by counting the number of times the signal crosses the zero line, which decreases as muscle fatigue causes a shift toward lower frequencies. These features have been widely used in muscle fatigue detection due to their simplicity and effectiveness.

3.2 Muscle fatigue classifier selection

Classification of fatigue is the last step of fatigue model is set up, on the extraction of feature dimension reduction, multi-channel sEMG its inputs to the classifier for training [22]. This article USES the Fisher linear discriminant analysis, *K* neighbor and three kinds of support vector machine classifier to classify fatigue.

3.2.1 Fisher linear discriminant analysis

Fisher linear discriminant analysis (FLDA) in the field of pattern recognition, introduced in the 1990s, is a classic algorithm in pattern recognition. Linear discriminant analysis is a kind of, supervision and learning needs for data classification label in advance, the basic idea is to project high-dimensional data to identify best low dimensional vector space, in the new subspaces, maximize the spacing between the similar sample, the sample spacing between different categories to minimize, or similar samples gathered as much as possible, samples of different categories of decentralized as much as possible. The core idea of FLDA is to find a projection direction that maximizes the distance between classes while minimizing the variance within each class, as conceptually depicted in Figure 3. This property makes it suitable for distinguishing well-separated fatigue states.

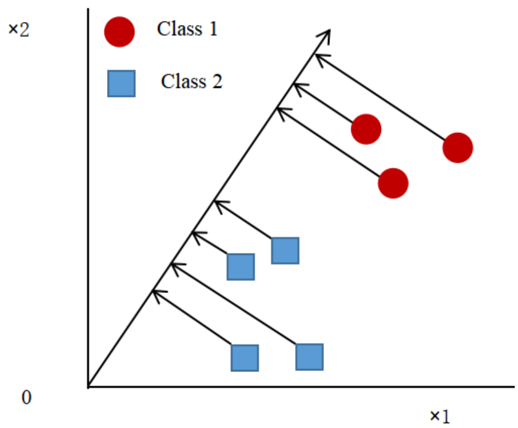

Figure 3: Fisher schematic diagram of the linear discriminant analysis principle.

Given sample set $(x_i, \ y_i), i = 1,2, \ldots, n, x \in R^N, y \in \{-1,1\}$, the first-class y the sample mean and variance as the Equation (25):

$$\mu_y = \frac{1}{N_y}\sum_{x_i \in y} x_i \qquad s_y^2 = \sum_{x_i \in y}\left(x_i - \mu_y\right)^2 \tag{25}$$

Type in N, said the number of sample points, the first (x_i, y_i) mapping to γ direction vector, is the first i a sample points on its projection for the Equation (26):

$$z_i = \gamma^T x_i \tag{26}$$

Is mapped to the first-class y sample direction vector γ on the mean and variance of the concrete as shown in Equation (27):

$$v_y = \gamma^T \mu_y \tag{27}$$

When the same sample interval with different kinds of sample spacing ratio, the largest best direction vector γ requires specific can be obtained as shown in the Equation (28):

$$argmax\, J\,(\gamma) = \frac{v_{-1} - v_1}{\sigma_{-1}^2 - \sigma_1^2} \tag{28}$$

Mean value and variance of (27) will type into type (28) and Equation (29) can be obtained:

$$\frac{\gamma^T(\mu_{-1} - \mu_1)(\mu_{-1} - \mu_1)^T\gamma}{\sum_{x_i \in y_{-1}} \gamma^T\,(x_i - \mu_{-1})(x_i - \mu_{-1})^T\gamma + \sum_{x_i \in y_1} \gamma^T\,(x_i} \quad (29)$$

Type of $(\mu_{-1} - \mu_1)(\mu_{-1} - \mu_1)^T$ class is called the divergence between the matrix and the available S_a said. $(x_i - \mu_{-1})(x_i - \mu_{-1})^T$ and $(x_i - \mu_1)(x_i - \mu_1)^T$ known as covariance matrix, and use the S_ω said together, that is, class divergence within the matrix. 30 type 2b after reduction for the equation:

$$argmax\,J\,(\gamma) = \frac{\gamma^T s_a \gamma}{\gamma^T s_\omega \gamma} \quad (30)$$

Type of s_a and s_ω, called generalized Rayleigh quotient, the $\gamma^r s_\omega \gamma = 1$, though laser multiplier method is used to calculate the maximum eigenvalue of matrix $s_a{}^{-1} s_\omega$ λ with the corresponding eigenvectors β, β is the direction vector β best value.

3.2.2 *K* neighbor

K Neighbor (*K*—on his Neighbor, KNN, put forward by Cover and others in the sixties of the 20th century, is a kind of typical supervised learning algorithm. *K* neighbor algorithm is the core idea: Classification of calculation for the distance between the training set and test set sample, find the nearest *K* training sample points, will stay classification test set samples for this *K* training set of sample points appear in the category of at most. The K-Nearest Neighbors (KNN) algorithm classifies a sample based on the majority vote of its nearest neighbors in the feature space, as shown in **Figure 4**. The choice of K value, as illustrated, critically affects the decision boundary.

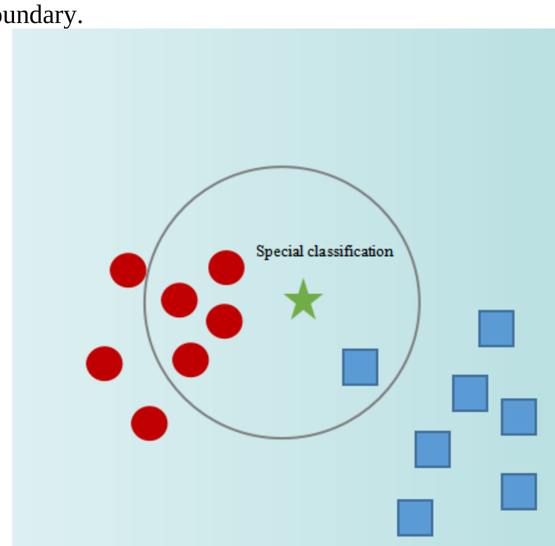

Figure 4: Schematic representation of the *K*-neighbor principle.

K neighbor algorithm has three elements, including the selection of *K* value, the distance measure and decision rules. *K* values will produce less fitting, and too small will have a fitting, affect classification recognition rate, generally with the method of cross validation to set the number of *K* value; Distance measurement with multiple methods, common with Euclidean distance, Manhattan distance, etc., this paper USES the Euclidean distance; For decision rules, in pattern recognition, generally USES the "majority voting method".

Set the sample sets $(x_i, y_i), i = 1,2, \ldots, n, (x_i, y_i \in Z)$, the calculation of Euclidean distance equation is:

$$d(x_i, y_i) = \sqrt{\sum_{i=1}^{n} (x_i, y_i)^2} \quad (31)$$

To calculate the distance the sample concentration sample points, and according to the distance size in descending order; Select the current classification sample points, the most recent *K* points, adopts "majority voting method", choose the highest frequency of sample points class, as a test sample point prediction of classes.

3.2.3 Support vector machine (SVM)

SVM (support vector machine, SVM) is put forward by Vapnik, first of all, is based on the statistical theory of VC dimension theory and structural risk minimum theory on the basis of a Machine learning algorithm. Core strategy is to find an optimal hyperplane, the support vector distance to the hyperplane is the largest, namely maximize between class and class intervals, can be applied to the characteristics of high dimension and nonlinear interface, etc., is a typical binary classification model. The SVM algorithm, is widely studied because of its classification performance advantage, mainly reflected in the system structure is simple, the global optimal, promotion ability, learning and prediction time is short. This experiment implements the SVM classifier using the LIBSVM library. The radial basis function kernel is employed, with both the penalty parameter C and kernel parameter γ optimized through grid search. The search space for C is defined as {0.1, 1, 10,100} and γ as {0.01,0.1, 1}. The optimization objective is determined by the average classification accuracy obtained from 10-fold cross-validation. The principle of Support Vector Machine (SVM) is to find the optimal hyperplane that maximizes the margin between classes, as shown in Figure 5. This approach ensures robust generalization performance.

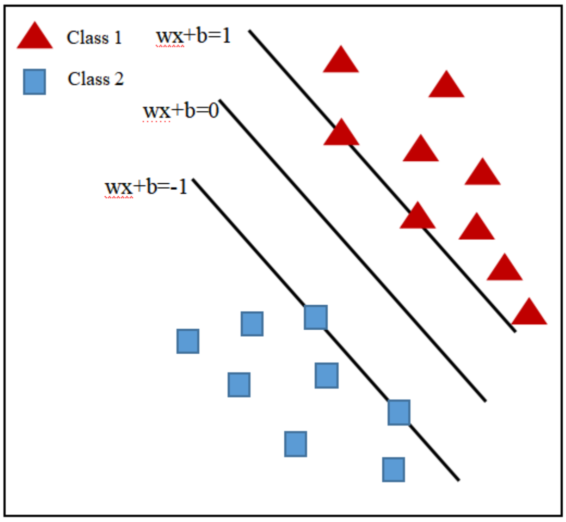

Figure 5: A schematic diagram of the SVM principle.

Given training set sample (x_i, y_i), $x \in R^n$, $y \in (-1,1)$, hyperplane $\omega \cdot x + b = 0$. In order to make the sample training set classification is correct, the details are shown in Equation (32):

$$y_i[(\omega \cdot x_i + b)] \geq 1, i = 1,2, \ldots, n \quad (32)$$

By the support vector can get $y_k[\omega \times x_k + b] = 1$, classification interval is calculated for $2/\|\omega.\|$, can be converted into minimum value problem with constraint conditions of (32), as shown in the Equation (33):

$$min\,\Phi\,(\omega) = \frac{1}{2}\|\omega\|^2 = \frac{1}{2}(\omega \cdot \omega) \quad (33)$$

Introducing the though laser function $L = (\omega, b, \alpha)$, to ω after the partial derivatives into a and b, which is transformed into the dual problem, as shown in the Equation (34):

$$max \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j x_i x_j$$

$$s.t.\$\sum_{j=1}^{m}\alpha_i y_i = 0 \ \alpha_i \geq 0, i = 1,2, \ldots, m \tag{34}$$

The calculated optimal weight vector ω^* and optimal offset b^*, the optimal hyperplane is obtained $(\omega^* \cdot x) + b = 0$, and get the optimal classification function Equation (35):

$$sgn\left\{\left(\sum_{i=1}^{m} \alpha_i^* y_i (x_i \cdot x_j)\right) + b^*\right\} \tag{35}$$

In this study, the SVM classifier was implemented using the scikit-learn library in Python. The Radial Basis Function (RBF) kernel was selected due to its effectiveness in handling nonlinear classification problems. The hyperparameters, including the penalty parameter C and the kernel coefficient γ, were optimized through a grid search strategy combined with 5-fold cross-validation on the training set. The search ranges were set as C = [0.1, 1, 10, 100] and γ = [0.01, 0.1, 1]. The combination that yielded the highest cross-validation accuracy was selected for the final model.

4 Classification of fatigue experiment and result analysis

Extensive research has been conducted on sEMG-based muscle fatigue recognition, with Table 1 summarizing representative methodologies developed in recent years. For instance, Xu et al. [6] employed convolutional neural networks (CNN) for feature prediction, though their model exhibited high complexity. Rampichini et al. [7] utilized a Gaussian process model achieving 91.25% accuracy, yet failed to resolve feature redundancy issues. Most studies directly applied multi-channel sEMG features, which resulted in redundant data and insufficient capture of nonlinear relationships through linear approaches [8]. To address these limitations, this paper introduces K-means Projection (KPCA) for nonlinear dimensionality reduction combined with Support Vector Machines (SVM) classifiers, aiming to achieve high-precision classification with reduced computational costs.

Table 1: Comparison of work related to muscle fatigue and sEMG recognition

Document	Feature extraction method	Classification method	Identification accuracy	Number of subjects	Boundedness
Xu et al. [6]	Convolutional neural network	CNN	~90%	10	It requires a lot of data and complex calculations
Rampichini et al. [7]	Mixed Gaussian process	Bacterial feeding optimization	91.25 %	8	Features may be redundant
Yun et al. [19]	Mutual information	LDA	85%	15	Linear methods may not capture nonlinear relationships
This research	KPCA	SVM	91.5%	12	Effective dimension reduction, nonlinear processing, high computational efficiency

4.1 Fatigue experiment scheme design

4.1.1 Objects and materials

The 120 subjects for the lab, including 60 boys and 60 girls, are right-handed, no nervous musculoskeletal diseases. In good health, between the ages of 16 and 20, 24 h before the experiment for high intensity exercise, no physical and mental fatigue. Each subject before the experiment were told the experimental process, at the same time sign the informed consent.

4.1.2 The experimental process

Experimental subjects before the corresponding action guidance to help their familiar with Borg subjective fatigue rating scale, and then according to individual circumstance to the experiment of several times to make participants familiar with the experimental process. Previous research, the experiment used more by those recorded every time ask participants state, so easy to interfere with the subjects, to give participants a certain psychological pressure, and different subjects, the time to reach the level of fatigue is endless also and same, easy to miss the corresponding fatigue level. The raw sEMG signals were acquired at a sampling rate of 2000 Hz using a Delsys Trigno wireless system. Prior to feature extraction, the signals were processed with a 4th-order Butterworth band-pass filter (20-450 Hz) to remove low-frequency drift and high-frequency noise, followed by a notch filter at 50 Hz to eliminate power line interference. This experiment adopts the subjects according to their own subjective initiative report subjective fatigue rating scale for 6-RPE score, subjective fatigue rating scale score and RPE classified as shown in Table 2, to RPE points, 11 as relaxed state, 12 to 18 for the transition state of fatigue, 19th and 20th for fatigue state. The sEMG signals were collected using the Delsys wireless acquisition system (16-channel) with a sampling frequency of 2000 Hz. The acquired raw signals were filtered through a band-pass filter with a frequency range of 20-450 Hz to eliminate power frequency interference and motion artifacts, followed by full-wave rectification.

Table 2: RPE score and status classification.

Grade	Self-feeling	Fatigue state
6	Easy	Easy state
7	Have a rest	
8	Extremely relaxed	
9	Relaxed	
10	Very light	
11	Light	
12	Fairly light	
13	Moderate	
14	It's a little hard	Fatigue transition state
15	Very hard	
16	Strenuous	
17	It's hard	Fatigue state
18	Extremely difficult	
19	Maximal exertion	
20	Completely exhausted	

Pattern recognition algorithms mainly include supervision of learning, a semi-supervised learning and unsupervised learning. Supervision, which is also called a teacher learning, it is using the known label samples (tags) was carried out on the data in advance, adjust the classifier parameters, through the tags and data mapping relationships to build the model. A semi-supervised learning the use of two sample sets, one of the sample data sets have category labels, another for unclassified tag data, establish proper classification model. Unsupervised learning to use the data set, without any label, need machine to modeling of data sets, such as clustering. In this paper, using supervised learning, need labels on methods of electrical data, generate the relaxed state, transition state of fatigue and fatigue state data set. Which will easily tag set to 1, fatigue transition state tags to 0, fatigue tag set to 1.

Subjects to close their eyes to relax on the bench to rest 2–3 min, to reach the best state of physical and mental relaxation. With grind arenaceous cream, alcohol cotton cleaning the skin, reduce skin impedance interference. According to the principle of anatomy, the methods of two difference electric sensor electrode according to the center of the interval of 2 cm distance joint in the target muscle bulging highest position, reference electrode from muscle testing, as far as possible general joint bone in the elbow. One-time recorded Dian Tie when fit, in order to make it closer and muscle adhesion to obtain more accurate, multi-channel sEMG can make electric gel surface with a small amount of water. Cloth with intramuscular effect sensor coil on the right hand oboro triceps, biceps, oboro scratching test parts of the muscle, to prevent the shedding of electrode paste during the test.

Subjects are sitting on a chair, neck, back straight, torso and legs, thighs and legs to keep 90°, respectively, on the left hand flat on left leg, right forearm and right upper arm to keep 90°, a little closer to the trunk. When it began, the first three seconds without load condition data; Then put in the dumbbell subjects, continue to maintain the non-weight bearing condition of position, and then start to collect the electromyographic data load condition. Acquisition process, the participants keep constant force lifts weights, the operator records the subjects' subjective fatigue rating scale RPE values and the corresponding time, and observe the subjects' right arm joint Angle, jitter, understand the subjects' subjective feeling at any time. When the subject right forearm and right upper arm cannot maintain 900 or arm end data acquisition when shaking violently. Therefore, the final dataset comprised a total of 360 experimental sessions (120 subjects × 3 sessions per subject). From each session, the continuous sEMG signals corresponding to the three pre-defined fatigue states were segmented using a 1-second window with 50% overlap. This resulted in a final dataset of approximately 63,700 samplesfor model training and evaluation.

4.2 Analysis, the result of the experiment

Following dimensionality reduction, the number of features was determined as follows: For MI, the top 15 features with the highest mutual information scores with the target fatigue states were selected. For PCA, the number of principal components was chosen to retain 95% of the cumulative variance, resulting in 12 components. Similarly, for KPCA (using an RBF kernel), the number of components accounting for 95% of the variance was 10.

1) A stratified 10-fold cross-validation method was employed to evaluate the model performance. The dataset was partitioned into 10 folds while preserving the percentage of samples for each fatigue class in every fold. This stratification helps to ensure that each fold is a good representative of the overall class distribution, reducing bias and the risk of overfitting, which is particularly important for small datasets. To evaluate the model performance robustly, a stratified 10-fold cross-validation scheme was employed, as illustrated in **Figure 6**. This method ensures that each fold is representative of the overall data distribution, reducing the bias in performance estimation.

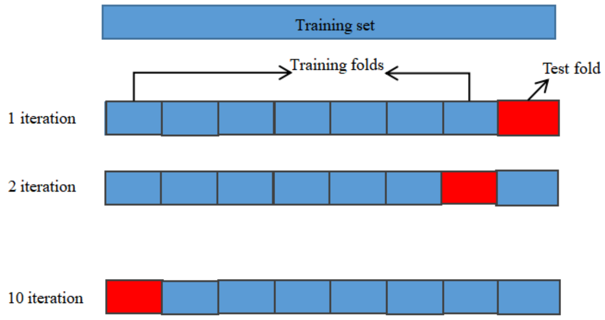

Figure 6: Ten folds cross validation schematic.

Will new characteristics after dimension reduction set into Fisher linear discriminant analysis, *K* neighbor, support vector machine (SVM) three classifier for training, Table 3 for MI, PCA, KPCA and FLDA three dimensions reduction algorithm, KNN, three SVM classifier combination on fatigue state average recognition rate and the average elapsed time.

Table 3: Experimental results of the different combinations

Method	Average Recognition Rate (%)	Precision	Recall	F1-Score	Average Elapsed Time (s)
MI + FLDA	69.3	0.68	0.71	0.69	4.08
PCA + FLDA	61.9	0.6	0.65	0.62	3.15
KPCA + FLDA	68.4	0.67	0.7	0.68	2.68
MI + KNN	77.1	0.76	0.79	0.77	4.35
PCA + KNN	80.8	0.79	0.83	0.81	3.77
KPCA + KNN	83.1	0.82	0.85	0.83	3.19
MI + SVM	83.6	0.82	0.86	0.84	4.15
PCA + SVM	86.9	0.85	0.89	0.87	3.73
KPCA + SVM	91.5	0.91	0.93	0.92	2.75¥

Can be seen from Table 3, the use of different dimension reduction method, Fisher classifier running time the shortest, but its average recognition rate is low; The SVM classifier has the highest average recognition rate, operation time is short, in which the combination of KPCA and SVM has the highest average recognition rate, the average recognition rate reached 91.5%.To statistically validate the superiority of the KPCA-SVM model, a Wilcoxon signed-rank test was performed comparing its accuracy against the other eight model combinations. The results indicated that the performance improvement of the KPCA-SVM model was statistically significant ($p < 0.05$), confirming its effectiveness beyond chance.

2) Rationale for classifier selection

The selection of FLDA, KNN, and SVM as classifiers was driven by their complementary strengths in handling muscle fatigue detection tasks:

- FLDA was chosen for its ability to maximize inter-class separation in linear spaces, which is critical for distinguishing between discrete fatigue states (e.g., relaxed vs. fatigued) [23].
- KNN leverages local data distribution patterns without assuming global data structures, making it robust to complex sEMG signal variations [24].
- SVM excels in high-dimensional nonlinear classification through kernel methods, aligning with the nonlinear nature of sEMG fatigue patterns [25].

Comprehensive performance evaluation

While Table 3 highlights the superiority of SVM in average recognition rate (91.5%), we further analyzed its performance using additional metrics (Table 4):

Table 4: Performance metrics of SVM classifier.

Metric	Value	Description
Precision	0.91	Proportion of correctly identified fatigue instances among all predictions.
Recall	0.93	Proportion of true fatigue instances correctly identified by the model.
*F*1 Score	0.92	Harmonic mean of precision and recall, balancing model performance.
Confusion Matrix	-	87% of misclassifications occurred between adjacent fatigue levels.

These results validate SVM as the optimal choice, though future work could explore ensemble methods combining FLDA's interpretability with SVM's nonlinear capabilities.

3) To further optimize the performance of KPCA, the choice of kernel function and its parameters plays a critical role. In this study, we evaluated three commonly used kernel functions: Radial basis function (RBF), polynomial kernel, and sigmoid kernel. The RBF kernel, defined as Equation (36),

$$K(x,y) = \exp(-\gamma \| x - y \|^2) \tag{36}$$

Was selected as the primary kernel due to its ability to capture nonlinear relationships effectively. The parameter γ was tuned through grid search to achieve the best classification accuracy. For the polynomial kernel Equation (37),

$$K(x,y)=(x^{T}y+c)^{d} \tag{37}$$

We tested degrees d ranging from 2 to 5, while for the sigmoid kernel Equation (38),

$$K(x,y)=\tanh(\alpha x^{T}y+c) \tag{38}$$

We experimented with different values of α and c. The results indicated that the RBF kernel consistently outperformed the other kernels in terms of classification accuracy, achieving an average accuracy of 92.3% on the test dataset.

Additionally, we compared KPCA with two other dimensionality reduction methods: Principal component analysis (PCA) and mutual information (MI). PCA, as a linear method, achieved an average accuracy of 85.7%, while MI, which selects features based on their mutual information with the target variable, achieved an accuracy of 88.1%. Although PCA and MI are computationally faster, their performance was inferior to KPCA, particularly in capturing nonlinear relationships in the sEMG signals. In terms of running time, KPCA required approximately 1.5 s per sample, compared to 0.8 s for PCA and 1.0 s for MI. These results demonstrate that KPCA, despite its higher computational cost, provides superior classification accuracy for muscle fatigue detection, making it a more suitable choice for this application.

Future work could explore hybrid approaches that combine the strengths of KPCA with other dimensionality reduction techniques to further improve efficiency without compromising accuracy.

4.3 Discussion

This study has several limitations. First, the sample size was relatively small (12 participants), which may affect the generalizability of the findings. Second, the fatigue labels were based on subjective RPE scores rather than objective physiological benchmarks (e.g., force decline or blood lactate levels). Third, potential confounding factors such as inter-subject variability in muscle activation patterns and psychological factors were not controlled. Finally, the model was validated under controlled laboratory conditions; its performance in real-world training scenarios requires further investigation. It is important to note that the fatigue state labels in this study were based on subjective RPE scores, which, while practical and widely used, may introduce labeling noise compared to objective physiological benchmarks (e.g., decline in maximum voluntary contraction). This reliance on subjective assessment is a limitation that might affect the model's robustness and generalizability. Future work should incorporate objective measures to validate and refine the labeling criteria.

The experimental results indicate that the combination of KPCA and SVM yielded the best performance for classifying muscle fatigue states. The superiority of KPCA over MI and PCA can be attributed to its ability to capture complex, nonlinear relationships within the high-dimensional sEMG feature set. Muscle fatigue induces non-linear shifts in the sEMG signal properties [12], which linear methods like PCA might not adequately represent. While MI selects relevant features, it does not transform them to create new, more discriminative features as KPCA does. The SVM classifier, known for its effectiveness in high-dimensional spaces and robustness to overfitting, synergized well with the discriminative features produced by KPCA. Analysis of the confusion matrix for the best-performing model (KPCA-SVM) revealed that most misclassifications occurred between adjacent fatigue states (e.g., Relaxed vs. Transitional, Transitional vs. Fatigued). This suggests that the boundaries between these states are not sharply defined physiologically, which aligns with the continuous nature of fatigue progression. The high precision and recall values further confirm the model's reliability in correctly identifying each state.A primary limitation of this study is the relatively small sample size (12 participants), which may affect the generalizability of the model. Future work should involve a larger and more diverse cohort, including athletes from different sports, to validate and enhance the model's robustness. Additionally, exploring real-time implementation and testing the model in practical training scenarios are important next steps.

5 Conclusion

This article from the perspective of athletes training and daily life and other sports this paper introduces the research significance of muscle fatigue, analyses and summarizes the present research situation of muscle fatigue, according to the relationship between the characteristic parameters and redundancy, without considering the computing complexity of the classifier training, affect the problem of classification recognition, puts forward the characteristics of sEMG analysis and classification of muscle fatigue. Design analysis system, the hardware acquisition circuit and the upper machine realized the three channels of multi-channel sEMG acquisition and data processing and analysis. Bringing new feature subsets after dimension reduction to Fisher linear discriminant analysis, *K* neighbor, support vector machine (SVM) classifier was trained, results show that the fatigue combining KPCA with SVM classification model, for muscle fatigue has the highest recognition rate is 91.5%.

Future work will explore several directions. First, advanced feature extraction techniques such as Wavelet Transform will be investigated to better capture the non-stationary characteristics of sEMG signals. Second, hybrid dimensionality reduction approaches that combine the strengths of filter (like MI), wrapper, and embedding (like KPCA) methods could be explored. Furthermore, validating the model on a larger and more diverse dataset, including professional athletes, and incorporating multimodal physiological data (e.g., ECG, accelerometry) are important next steps to enhance robustness and practical applicability.

Funding

This project was supported by Research on Building a High level Traditional Chinese Medicine Sports and Health Care Service System in the New Development Stage (No. 2022A398).

References

[1] Boyas S, Guevel A. Neuromuscular fatigue in healthy muscle: Underlying factors and adaptation mechanisms. Annals of Physical and Rehabilitation Medicine. 2011; 54(2): 88–108.

[2] Hughes E, Bell A. A wireless surface electromyography system. Vcu. Etd. Archive. 2007; 65(8): 53.

[3] Jero SE, Ramakrishnan S. Order frequency spectral correlation based cyclo-nonstationary analysis of surface EMG signals in biceps brachii muscles. In: Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 23–27 July 2019; Berlin, Germany.

[4] Moniri A, Terracina D, Rodriguez-Manzano J, et al. Real-time forecasting of sEMG features for trunk muscle fatigue using machine learning. IEEE Trans. Biomed. Eng. 2021; 68(2): 718–727.

[5] Bueno DR, Lizano JM, Montano L. Muscular fatigue detection using sEMG in dynamic contractions. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2015; 494–497.

[6] Xu B, Wu Q, Xi C, He R. Recognition of the fatigue status of pilots using BF-PSO optimized multi-class GP classification with sEMG signals. Reliability Engineering and System Safety. 2020; 199(2).

[7] Rampichini S, Vieira TM, Castiglioni P, et al. Complexity analysis of surface electromyography for assessing the myoelectric manifestation of muscle fatigue: A review. Entropy, 2020, 22(5): 529.

[8] Lan N. Analysis of an optimal control model of multipoint arm movements. Biological Cybernetics. 1997.

[9] De Luca CJ. Physiology and mathematics of myoelectric signals. IEEE Trans. Biomed. Eng. 1979; 26(6): 313–325.

[10] Sun Z, Xi X, Yuan C, et al. Surface electromyography signal denoising via EEMD and improved wavelet thresholds. Math. Biosci. Eng. 2020; 17(6): 6945–6962.

[11] Zhang Y, Chen Y, Yu H, et al. Learning Effective Spatial-Temporal Features for sEMG Armband-Based Gesture Recognition. IEEE Internet of Things Journal. 2020; 7: 6979–6992.

[12] Georgakis A, Stergioulas L, Giakas G. Fatigue analysis of the surface EMG signal in isometric constant force contractions using the averaged instantaneous frequency. IEEE Transactions on Biomedical Engineering. 2003; 50(2): 262–265.

[13] Alobaidy A A M, Jabbar A M J, Aly M, et al. Real time fault diagnosis in industrial robotics using discrete and slantlet wavelet transformations[J]. Scientific Reports, 2025,15(1):34145-34145. DOI:10.1038/S41598-025-09272-9.

[14] Trybek P, Nowakowski M, Salowka J, et al. Sample Entropy of sEMG Signals at different stages of rectal cancer treatment. Entropy. 2018; 20(11): 863.

[15] Mugnosso M, Zenzeri J, Hughes CML, et al. Coupling robot-aided assessment and surface electromyography (sEMG) to evaluate the effect of muscle fatigue on wrist position sense in the flexion-extension plane. Frontiers in Human Neuroscience. 2019; 13: 396.

[16] Toro SFD, Santos-Cuadros S, Olmeda E, et al. Is the use of a low-cost sEMG sensor ualid to measure muscle fatigue? Sensors. 2019; 19(14): 3204.

[17] Skavhaug IM, Lyons KR, Nemchuk A, et al. Learning to modulate the partial powers of a single sEMG power spectrum through a novel human-computer interface. Human Movement Science. 2016; 47: 60–69.

[18] Liu Q, Liu Y, Zhang C, et al. sEMG-based dynamic muscle fatigue classification using SVM with improved whale optimization algorithm. IEEE Internet of Things Journal. 2021; 99: 1

[19] Yun I, Jeung J, Song Y. Non-invasive quantitative muscle fatigue estimation based on correlation between sEMG signal and muscle mass. IEEE Access. 2020; 8: 51–57

[20] Karlsson S, Yu J, Akay M. Time-frequency analysis of myoelectric signals during dynamic contractions: A comparative study. IEEE Trans. Biomed. Eng. 2000; 47(2): 228–238.

[21] Dash M. Feature selection for classification. Intelligent Data Analysis. 1997.

[22] Khalid S, Khalil T, Nasreen S. A survey of feature selection and feature extraction techniques in machine learning. In: Proceedings of the 2014 Science and Information Conference. 27–29 August 2014; London, UK.

[23] Jain AK, Duin RPW, Mao J. Statistical pattern recognition: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000; 22(1): 4–37.

[24] Cover T, Hart P. Nearest neighbor pattern classification. IEEE Transactions on Information Theory. 1967; 13(1): 21–27.

[25] Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995; 20(3): 273–297.