Research Article

Recognition Technology of Athlete's Limb Movement Combined Based on the Integrated Learning Algorithm

Fei Tan[1] **and Xiaoqing Xie** [2]

[1]*Hunan First Normal University of Physical Education, Changsha, Hunan 410205, China*
[2]*College of Physical Education of Hunan Normal University, Changsha, Hunan 410012, China*

Correspondence should be addressed to Xiaoqing Xie; 16393@hunnu.edu.cn

Human motion recognition based on inertial sensor is a new research direction in the field of pattern recognition. It carries out preprocessing, feature selection, and feature selection by placing inertial sensors on the surface of the human body. Finally, it mainly classifies and recognizes the extracted features of human action. There are many kinds of swing movements in table tennis. Accurately identifying these movement modes is of great significance for swing movement analysis. With the development of artificial intelligence technology, human movement recognition has made many breakthroughs in recent years, from machine learning to deep learning, from wearable sensors to visual sensors. However, there is not much work on movement recognition for table tennis, and the methods are still mainly integrated into the traditional field of machine learning. Therefore, this paper uses an acceleration sensor as a motion recording device for a table tennis disc and explores the three-axis acceleration data of four common swing motions. Traditional machine learning algorithms (decision tree, random forest tree, and support vector) are used to classify the swing motion, and a classification algorithm based on the idea of integration is designed. Experimental results show that the ensemble learning algorithm developed in this paper is better than the traditional machine learning algorithm, and the average recognition accuracy is 91%.

1. Introduction

Human action recognition technology has received attention and importance from research scholars in many fields such as machine learning, distributed computing, situational awareness, security monitoring, and smart home because of its unique research value. For example, the purpose is often through the collection and processing of human movement information to analyze human movement, behavior, and even emotion or use the analysis of human movement and behavior to guide the development and design of intelligent robots [1]. For example, in the field of security monitoring, human motion data collected by surveillance cameras installed on roads, in waiting halls, or in venues of major conferences and events are used to predict certain human behaviors for the purpose of detecting safety hazards [2]. In the field of intelligent elderly care, researchers try to analyze the behavior of the elderly through body sensors and

surveillance cameras for the purpose of safety and in the field of rehabilitation robotics, the study of symbiotic robots helps to make patients and assistive robots work together harmoniously [3]. But there are differences in body postures and conditions among different patients, and studying patients' behaviors is beneficial to make the machine intelligently adapt to patients and improve their experience of using it, thus better helping patients to recover [4]. There are numerous application cases of similar human action recognition and analysis, and it can be seen that human action recognition has a wide range of applications and prospects.

With the continuous development of sensor technology as well as artificial intelligence, the study of human action recognition has brought technological innovation and hope to many fields [5]. As an integral part of social activities, sports have a special status, and the study of movements and behaviors in sports has an important significance for improving traditional sports training, competition, and

competitive sports management and organization. For example, in professional sports training, the degree of mastery of some action elements directly affects whether athletes can approach the limit of performance and also affects whether athletes can effectively avoid sports injuries, so scientific and comprehensive analysis of athletes' technical movements can help athletes quickly improve their performance [6]. Human action recognition based on inertial sensors is an emerging research direction in the field of pattern recognition; in essence, the inertial signal of human action is firstly collected by one or more inertial sensor-integrated devices, and the inertial data is transmitted to the server processing equipment through wireless communication such as ZigBee, and then, the inertial data is preprocessed through online or offline methods such as denoising and feature extraction and feature selection, and finally, the inertial data is analyzed based on the obtained features (as shown in Figure 1). Finally, the action is classified according to the obtained signal features [7]. There are two main types of human motion acquisition methods that have received a lot of attention from researchers: video-based human motion recognition and inertial sensor-based human motion recognition. Although the video-based approach started earlier and has mature algorithms, it still suffers from unsolvable defects such as susceptibility to environmental and other factors and violation of personal privacy [8]. In contrast, inertial sensor-based human motion recognition has significant advantages [9]. When using inertial sensors to collect human daily movement data, it is not limited by the external environment (such as weather and activity range) and does not need to consider external equipment such as cameras [10]. Therefore, when collecting data from the measured object, participants can move freely according to their personal habits, and the data thus obtained is closer to the real situation, and the classification algorithm model trained is more generalizable to the real environment.

In ordinary sports exercise, some sports enthusiasts also expect to be able to analyze their movements with the help of wearable devices or sensors such as video to enhance the experience and reward of sports. In table tennis, for example, the movements of the players are often varied but distinctive. In preparing for a table tennis match, coaches usually analyze opponent's movement characteristics through video materials before the match, hoping to find opponent's weaknesses in attack and defense, so as to find a strategy to overcome the opponent. In addition, coaches will arrange chaperones with similar styles for their players to adapt to opponent's style in advance and to practice the playing style to overcome the opponent. Therefore, the research on how to recognize table tennis swings has at least two implications: first, it can model, recognize, and analyze swings more scientifically and accurately to guide table tennis training and games; second, it can guide the development of smart wearable bracelets and table tennis robots. In this paper, we take the recognition of swings in table tennis as the research goal and study how to recognize swings through wearable acceleration sensors. In addition, traditional machine learning algorithms (decision tree, random forest, and support vector machine) are applied to swing action classification, and a
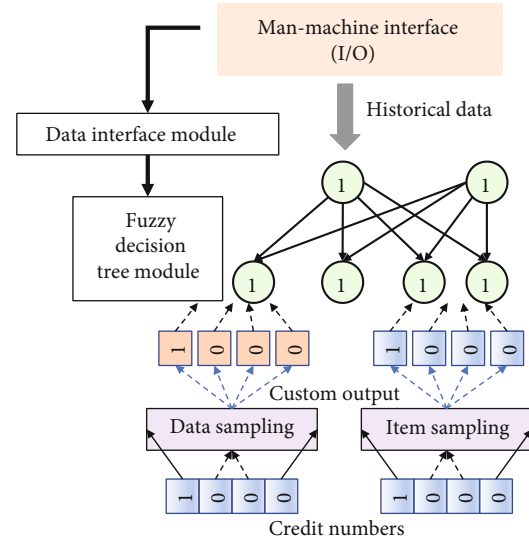

Figure 1: Inertial data denoising process.

classification algorithm based on the integration idea is designed. The integrated learning algorithm can effectively distinguish four types of swings, and the experimental results show that the integrated learning algorithm designed in this paper outperforms the traditional machine learning algorithm and achieves an average recognition accuracy of 91%.

2. Related Work

Athlete limb action recognition belongs to the human action recognition research category, and human action recognition can be traced back to the 1990s. In the past two decades, from vision-based sensors to various wearable body sensors [11], from traditional machine learning-based methods to designing complex deep learning network models [12], researchers in the field of human action recognition have experienced many attempts and made many breakthroughs. Currently, the following three types of research methods are commonly used for human action recognition: vision-based methods, environmental sensor-based methods, and wearable sensor-based methods, [13]. In early research, vision-based approaches dominated the field [14], which implemented human action recognition by one or more cameras installed in the environment. Based on video continuous frame image data, the human contour, bones, or points of interest that can be tracked are extracted, and then, machine learning or deep learning models are trained to recognize human actions. Although video-based approaches can provide rich visual information, articles [15], more attentions should be put into the combination with the wearable sensors to make sure that the classification is accurate. First, privacy is the first issue, and people are wary of cameras, fearing information leakage or being spied on, thus making it difficult to promote vision-based methods in practical applications.

The second is the limitation of the environment usage. Vision-based methods need to be used in a specific

environment containing a camera, and the location of the camera is usually fixed, which makes it difficult for vision-based methods to be used outdoors [16]. Then, again, vision-based methods are more influenced by the environment, such as lighting, shading, weather, and other factors that can affect the performance of vision-based methods. Finally, vision-based methods are more computationally expensive, less scalable, and difficult to integrate in real time and miniaturization. The idea of environment sensor-based methods is generally simpler, and the method is more used in smart home research [17], such as analyzing human actions or behaviors by detecting changes in the location where people or objects are located through distance sensors, for example, by distance, sensors can detect whether people are in a certain place. The position of the object is detected to determine whether the person performs a corresponding action on the object. From the principle, it is clear that such methods can obtain a single piece of information, and the functions that can be accomplished are relatively simple. Designing such systems requires the placement of many sensors in the environment, so the application is also very limited [18].

In addition to the above two methods, wearable sensor-based methods are also the main "position" of human action recognition researchers. Wearable sensors are small in size, widely distributed, easy to carry, and independent of the environment. Researchers can use one or more sensors on the human body to collect human motion data at any time and place, thus freeing them from the constraints of usage and privacy. Compared to the visual approach, wearable sensors can also provide a wealth of data information. This is because wearable sensors are more diverse, and researchers can arrange several multitype sensors on the human body for human motion information collection. In addition, with the development of sensor technology, the cost of wearable devices is gradually decreasing, making it easier to be promoted in practical applications [19]. Wearable sensors, as the name implies, are a general term for all sensors that can be conveniently placed on the human body by wearing or carrying. The human movement recognition method based on wearable sensors collects human movement, physiological, and other information through sensors to complete the analysis and recognition of human movement. With different wearable sensors, researchers can obtain three types of information: environmental information, inertial information, location information, and physiological signal information.

In addition to the study of which sensors to use and which sensor layout to recognize human actions, the classification algorithms to use to classify different kinds of actions are also the main focus of researchers. Current sensor-based action recognition methods can be broadly classified into two categories: one is knowledge-driven and the other is data-driven. In the knowledge-driven approach, an abstract knowledge model of the action to be classified needs to be constructed first, and then, the constructed knowledge model is used for action recognition. A common knowledge model construction method is the ontology modeling method. The ontology modeling method has the features of semantic clarity, logical conciseness, and easy interpretation, but this method cannot deal with human action relationships in the temporal dimension, and this method usually requires designers with rich domain knowledge and experience, so it has not been widely promoted, and there is very little relevant research literature. Compared to knowledge model-based approaches, data-driven approaches are widely adopted and practically applied. Data-driven approaches generally require a large amount of motion data for training, and the approach essentially explores the specificity of the motion itself from the data to identify different motion types. Data-driven approaches do not require tedious mathematical modeling of movements or extensive domain experience of the designer. Currently, data-driven action classification algorithms can be classified into three categories: supervised learning methods, unsupervised learning methods, and semisupervised learning methods. Supervised learning uses training data with labels, unsupervised learning methods do not require training samples to contain label information, and semisupervised learning methods only contain labels in part of the data.

3. Player Swing Recognition

3.1. Data Acquisition. Table tennis motion pattern recognition based on acceleration sensors requires techniques and theories such as feature extraction and pattern recognition, in addition to hardware devices such as sensors. In previous studies, although different recognition systems and frameworks have been formed due to hardware devices, types of recognition movements, application scenarios, etc., these frameworks have many similarities in the basic design process of recognition systems [20]. When designing the system, the collected data is first preprocessed, and the continuous data sequence is segmented into individual data segments using window segmentation techniques after noise elimination. Then, the feature values of each data segment are calculated, and feature vectors are constructed, and if the dimensionality of the feature vectors is too high, feature filtering is required. Finally, these feature vectors are used to make a training data set to train the classifier, and after the training is completed, the classifier can recognize several kinds of action set. The action recognition system is mainly divided into five modules: data acquisition, data preprocessing, feature extraction, feature screening, and recognition and classification (as shown in Figure 2).

With the development of sensor technology, many parameters of human body can be acquired by different types of sensors, such as temperature sensor can measure the temperature of human body, pulse sensor can record the number of pulse beats, and acceleration can collect the motion data of human body. The data acquisition module in this paper uses the acceleration sensor, which integrates a 3-axis acceleration sensor, a 3-axis angular velocity sensor, and a 3-axis magnetometer, all of which support 16-bit precision output. In addition, it can interact with the embedded microprocessor through an integrated circuit bus (IIC) interface with data transfer rates of up to $400 \, k \, Hz/s$ and good dynamic response characteristics. The sensor is set to
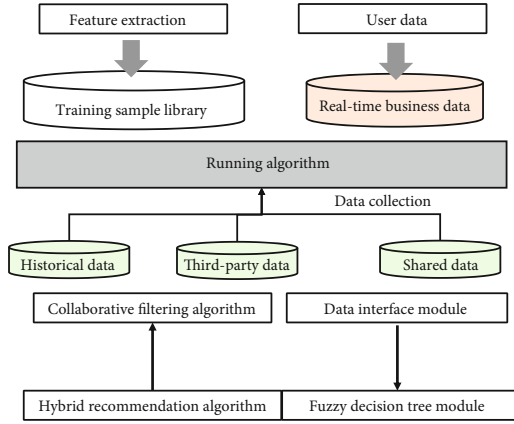
Figure 2: Action recognition flow.

200 Hz for data acquisition, which is not too large and is sufficient to capture the details of the four motion patterns. The data collected during the experiment was stored in text format on a memory card with the sensor and data storage device embedded in the bracelet. The data are stored in columns in the format of:

$$\text{Action} = \{x_i, y_i, z_i\} (i = 1, 2, 3 \cdots), \tag{1}$$

representing x-axis acceleration, y-axis acceleration, z-axis acceleration, x-axis angular velocity, y-axis angular velocity, and z-axis angular velocity, respectively. In this experiment, the x-axis, y-axis, and z-axis acceleration data are taken as the original samples. In order to make the designed classification algorithm have good classification ability, this experiment fully considered the differences of motion data of individuals of different ages and genders. 10 males and 10 females participated in the data collection of this experiment, and their ages were concentrated between 18 and 45 years old. The data were collected using the same bracelet wearing style (e.g., all right-handed, with a vertical grip). After data collection, we labeled the forehand attack, backhand push, forehand roll, and backhand roll samples with the numbers 1, 2, 3, and 4, respectively, in order to distinguish the four movements when training the classifier.

3.2. Data Preprocessing. When collecting four kinds of table tennis sports data by intelligent wearable devices, it may be due to hardware circuits, transmission noise, and human body jitter, so the collected data contains not only effective human motion information but also various noises. If this original data is directly used to extract motion features, the effectiveness, reliability, and recognition accuracy of the system will be greatly reduced. The specific parameters can be from the raw data collected by sensors, which are continuous data sequences generated by user movements over a period of time, and generally, the data length is up to hundreds of thousands or even millions, and direct feature extraction will greatly increase the computational cost.

Translated with http://www.DeepL.com/Translator (free version) in order to reduce the influence of noise to obtain effective data sequences, the original data needs to be fil-
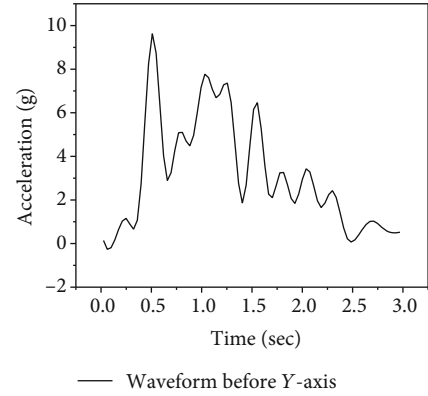
Figure 3: Waveform before y-axis acceleration filtering for forehand attack.
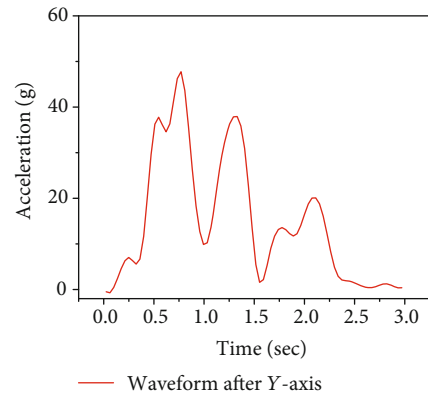
Figure 4: Waveform after y-axis acceleration filtering for forehand attack.

tered. Filtering is an important part of signal processing; it is the use of mathematical methods to extract the useful signal from the received signal containing interference noise of a technique [21]. There are many filtering methods in practical applications, such as mean filtering, low-pass filtering, Gaussian filtering, and Kalman filtering.

In this paper, a one-dimensional Gaussian filtering method is used. Gaussian filtering is a linear smooth filtering method that can effectively eliminate various Gaussian noises widely present in the signal, and the algorithm is shown in

$$\text{Move}(x) = af(x-1) + bf(x) + cf(x+1), \tag{2}$$

$$f(x) = \frac{e^{-x/\delta^2}}{\sqrt{2\pi}\delta}, \tag{3}$$

where movement is the current filter output value; $f(x-1)$, $f(x)$, $f(x+1)$ are the temporally adjacent three sampled values; $f(x)$ is the Gaussian filter coefficient solver function; δ is the standard deviation. Gaussian filtering can not only remove high frequency interference signals but also make the waveform of motion data smoother. Figures 3 and 4 show the waveforms before and after the filtering of any data segment in the Y direction of forehand attack, and it can be

found that the filtering effect of Gaussian filtering algorithm is remarkable.

The raw data collected by sensors are continuous data sequences generated by user movements over a period of time, and generally, the data length is up to hundreds of thousands or even millions, and direct feature extraction will greatly increase the computational cost. To solve this problem, it is a common practice to perform a windowing operation before calculating various feature values to split the longer sensor data signals into many data segments using windows. According to previous research experience, there are two main types of windowing methods commonly used in human motion recognition systems: sliding window segmentation and action-based window segmentation. After the collected swing action data are filtered and processed, the invalid data due to dropped balls or abnormal jitter and other situations may not be removed yet, and in addition, considering the data sequence is too long, this paper uses sliding windows to process the data.

Sliding window segmentation uses a fixed-length window to segment the motion signal acquired over a period of time into several equal-length data segments, and the overlap rate of the window can be set according to specific needs when the window slides backward on the signal sequence in the order of acquisition time. Considering the periodicity and coherence of the four motion modes and the requirement of minimizing the computing cost, the amount of data acquired within 4 seconds is chosen as the length standard. There are various table tennis swings, and only four of them are studied in this paper: (a) forehand attack, (b) backhand push, (c) forehand rub, and (d) backhand rub. In order to ensure the integrity of the window start/stop edge motion data, the window overlap rate is set to 50%. The motion data generated by the same movement of different subjects due to their body size, age, gender, and other factors are different in terms of movement amplitude and data length, while the window length is fixed by sliding window segmentation, which can maintain the consistency of data length and provide convenience for subsequent calculation of feature values. The sliding window segmentation has no specific starting and ending time points, which reduces the complexity of the algorithm of the system. However, this window segmentation technique also has the disadvantage that two types and more action data cannot appear in one window; otherwise, the labels labeling the action types will diverge when making the training dataset.

3.3. Feature Extraction and Selection. After filtering and adding windows to the collected data, it is found that even if the same person does the same action, the window data formed are not exactly the same, so it is necessary to perform feature extraction of the data within the window to form quantitative information that can be used to compare the variability of signals among windows. Currently, there are two general types of methods for feature extraction based on time-domain sequence data, one is statistical and the other is structured. Statistical methods, such as Fourier transform and wavelet transform, use quantitative information of the data as features of the window; in contrast, structured

methods focus on considering the interconnections between the data within the window. Therefore, the choice of which one to choose depends on the actual needs, i.e., the type of signal. For the signals obtained from accelerometers, the fluctuations and oscillations are large, so it is difficult to perform action recognition directly based on the raw data. Existing studies on human action recognition based on acceleration sensors generally use statistical methods for feature extraction, and these methods mainly consider the time-domain and frequency-domain features of the window signal. Among them, time domain features mainly include mean, standard deviation, quartile difference, mean absolute variance (MAD), correlation coefficients between different axes, wave peaks, and wave troughs, whose algorithms are less complex and easy to implement in wearable devices; frequency domain analysis method first uses fast Fourier transform to transform from time domain to frequency domain, and frequency domain features mainly include FFT coefficients, frequency domain entropy, energy spectral density, and energy. FFT coefficients are complicated to calculate and are not suitable for wearable devices. Different swings present different amplitude characteristics and change patterns in the time domain, so this paper mainly adopts the time domain analysis method to calculate the characteristic values of motion data. Twenty-two characteristic quantities are extracted from the acceleration data in X, Y, and Z directions, which include the mean, variance, maximum, minimum, peak, and trough values of acceleration in the three directions, the correlation coefficients between two accelerations in the three directions, equivalent motion energy, and period.

In order to eliminate the direction factor, the equivalent motion energy was obtained by synthesizing the acceleration of the x-, y-, and z-axes, calculated as follows:

$$a_k = \left(a_x^2 + a_y^2 + a_z^2\right)^{-1/2}, \qquad (4)$$

where a_x, a_y, and a_z are the acceleration data in the x-, y-, and z-axes, and a_k is the equivalent motion energy. The mean value of the equivalent motion energy is the quantity of the trend in the data set, which measures the integrated intensity of a class of motion. The formula is as follows:

$$\overline{a_k} = \lim_{n \longrightarrow \infty} \frac{\sum_1^n \left(a_x^2 + a_y^2 + a_z^2\right)^{-1/2}}{n}. \qquad (5)$$

The motion period characterizes the average time interval between two adjacent swings. Generally, the period is obtained based on the signal waveform, and for more regular waveforms such as sine waves and cosine waves, the period can be found by simply calculating the average interval between two adjacent peaks or troughs. Although the motion waveform is periodic, it is difficult to directly find all the peaks and troughs in a waveform that conform to the periodic variation, so this paper further introduces the signal processing technique—autocorrelation analysis method. In theory, the length m of the template should be greater than one complete

swing period. For general considerations, the length m can be taken as 1/4 times the total length of the sample. Starting from the 1st, 2nd, 3rd, ..., $n - m + 1$ data in the sample, we construct subsample data of the same length as the template and perform autocorrelation with the template. When a segment of the waveform in the sample is similar to the template waveform, a "resonance" similar to that in acoustics will occur, and the result of the operation will produce a local maximum, and since the sample is a periodic waveform, the template will produce a "resonance" after every motion cycle during the backward operation. Therefore, the resulting autocorrelation sequence is also periodic, and the period is equal to the desired motion period. The equation for the autocorrelation operation is as follows:

$$\text{move}(\overline{a_k}) = \lim_{n \longrightarrow \infty} \frac{\sum_0^n \left(\text{move}(a_x)^2 + \text{move}(a_y)^2 + \text{move}(a_z)^2 \right)^{-1/2}}{n},$$

(6)

where movement represents the actual signal value at moment a_k in sample; n represents the signal value of i in the autocorrelation template function; m represents the length of the template. In order to obtain the peaks more robustly and eliminate the interference of subpeaks on the cycle solution, the data in the sample should be demeaned and normalized: before each autocorrelation analysis, the data in the sample are differenced from the sample mean by a factor of k ($0 < = k < = 1$), and the data larger than the mean are kept, and the data smaller than the mean are set to zero. The demeaning formula is as follows:

$$\text{action} = \sqrt{\text{move}^2(x) + \text{move}^2(y) + \text{move}^2(z)}. \quad (7)$$

Take 2000 forehand attacks as an example, and compare the waveform of equivalent motion energy with the waveform of equivalent motion energy after deaveraging (as shown in Figure 5).

4. Action Classification and Discussion of Results

When features are extracted, there are some features in the constructed feature set that may be irrelevant or even irrelevant. These components are redundant features, which can also be called noise, and their presence not only does not improve the performance of the action classification algorithm but also reduces the recognition accuracy. Therefore, feature selection is often necessary before doing model training. According to the sequence of feature selection and training classifiers, feature selection methods can be broadly classified into three categories: filter, wrapper, and embedding. The filtering method first selects the features and then trains the classifier, and the performance of the classifier is not considered in the feature selection process, and the two are independent; the wrapper method directly takes the performance of the final classifier to be used as the evaluation criterion of the feature subset and improves the performance of the algorithm by training the classifier several times, which has a large computational overhead, while the embed-
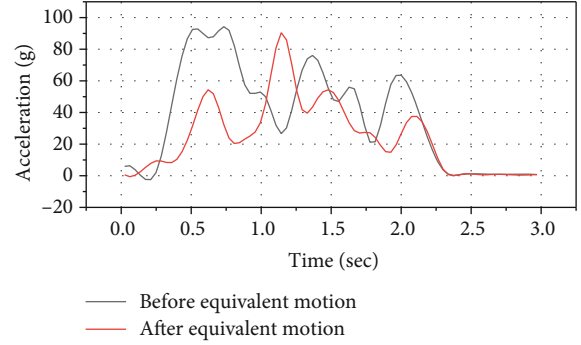

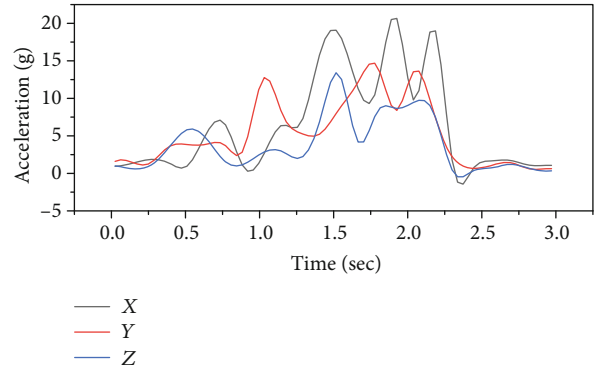FIGURE 5: Comparison of before and after equivalent motion energy reduction rms.

FIGURE 6: Acceleration analysis of the swing motion.

ded method integrates feature selection and classifier training into the same optimization process. The embedded approach integrates feature selection and classifier training into the same optimization process, which makes the algorithm robust to noise and therefore can solve the problems of the first two approaches. There are various table tennis swings, and only four of them are studied in this paper: (a) forehand attack, (b) backhand push, (c) forehand rub, and (d) backhand rub. The acceleration data of the four movements over a period of time are shown in Figure 6, where the horizontal axis represents the time (s) and the vertical axis represents the acceleration magnitude (g).

Crossvalidation and evaluation methods based on statistical methods are generally used in human action recognition studies to assess the performance of classification algorithms. The classification effectiveness of a particular algorithm can be presented in the form of an obfuscation matrix, which reveals the degree to which different kinds of actions may be confused with each other. If there are n kinds of actions to be classified, then the confusion matrix is an nn table. Taking the most common binary classification problem as an example, the confusion matrix is shown in the table. In the table, TP is true positive, which means the number of samples predicted positive class as positive class, i.e., correct prediction; FN is false negative, which means the number of samples predicted positive class as negative class, i.e., wrong prediction; FP is false positive, which means the number of samples predicted negative class as positive class, i.e., wrong prediction; TN is true negative, which means the
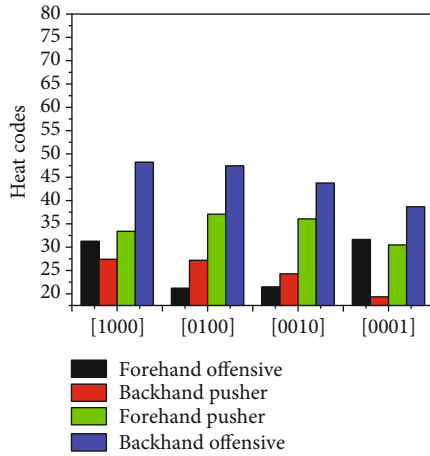
Figure 7: Solo heat codes for the four swings.

Figure 8: Convergence of the objective function.

number of samples predicted negative class as negative class, i.e., correct prediction. Based on the acceleration sensor data, this paper designs an integrated learning strong classifier for ping pong swing action recognition. From the whole design process of integrated learning, it is easy to find that integrated learning has the following characteristics: first, the process is tedious, and each subclassifier needs to be designed before building the integrated learning classifier; second, the integrated learning method requires feature extraction and selection based on human experience, and the advantages and disadvantages of feature extraction often require a lot of experimental verification in practical applications; third, the association between the classifiers of integrated learning. The performance of the classifiers cannot be absolutely guaranteed because of the relatively weak association between the classifiers. The time width of the sliding window in this experiment is set to 4 seconds, and the sampling frequency of the device is 200 Hz, so each sample consists of 800 sampling points, i.e., the t value is 800. Each action sample is an $800 * 3$ matrix. Among all the obtained data samples, 80% of them are used as training samples by random method, totaling 4000 (1000 for each of the four types of actions), and the remaining 20% are used as test samples, totaling 1000 (250 for each of the four types of actions). Unlike the integrated learning where the labels of the four swings are set to 1, 2, 3, and 4, the training data are labeled using discrete one-hot encoding, which is suitable for the error function defined in the form of crossentropy, and this definition method makes the distance between labels more reasonable. Calculate the loss, which is more conducive to parameter solving. The solo-heat coding methods for the four swings are shown in Figure 7 below.

For the table tennis swing recognition task, the input data are 3D acceleration data of certain sequence length. The goal of this chapter is to form a table tennis swing recognition network by continuously training equal length samples of triaxial acceleration data, and finally, given a sequence length of 3D acceleration data as input, the designed network can predict whether it belongs to a certain swing action and the confidence of the response. To achieve this goal, the sequence length is assumed to be T. The idea of
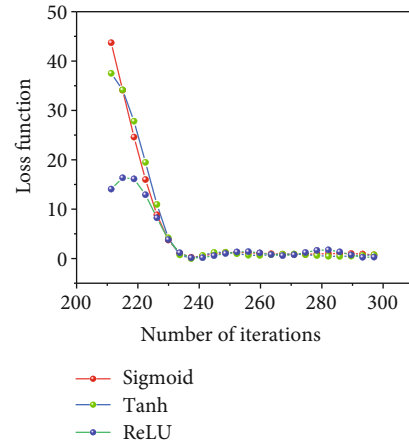
sequence convolution requires the network to have a perceptual field of size T, that is, each element of the output sequence of the network is the inference result of the input sequence of length T. The causal and null convolution design of the TCN model provides a possibility for this idea, where causal convolution is the basis of sequence perception, and null perception can accomplish the task of perceiving inputs with large sequence lengths.

In addition, the design of residual convolution keeps the time series network from degrading performance and guarantees the network training results. In the experiments, the activation function is firstly determined, and the Sigmoid, Tanh, and Re LU functions are used as the activation functions, and the convergence of the objective function is shown in Figure 8. By comparing the experimental results, we find that the objective function converges to about 0.40 when the sigmoid function is used as the activation function, about 0.32 when the tanh function is used, and about 0.23 when the Re LU function is used as the activation function. It can be seen that for the action recognition task in this paper, the convergence of the Re LU function as the activation function is better. Since the initialization value of the network model affects the performance of the network, for example, the learning rate of the network model will affect the convergence speed of the target function and whether it can converge to the minimum value. When the learning rate is too small, the convergence of the target function will be slow, making the network training too inefficient and consuming too many resources; when the learning rate is set too large, it will cause the target function to fail to converge to the set threshold, but oscillate around it. Pushing the ball was the lowest. In terms of overall classification performance, the average recognition accuracy of the integrated classifier reached 91%, which is the best performance above the other three classification models.

5. Conclusion

This paper uses an acceleration sensor as a motion recording device for a table tennis disc and explores the three-axis acceleration data of four common swing motions. The field

of human motion recognition has benefited from the rise and development of sensor technology, and while vision-based approaches are dominant, there is still a lot of room for sensor-based approaches. Compared with vision-based methods, sensor-based methods are more flexible and have good performance in the recognition of daily movements such as walking, running, and walking up and down stairs. Therefore, this paper applies the sensor-based human action recognition method to daily table tennis sports and designs recognition models for classifying four types of actions in table tennis: forehand attack, backhand push, forehand rub, and backhand rub. In this paper, the acceleration sensor is used as a table tennis swing acquisition device, and the three-axis acceleration data of the four common swings are investigated. Traditional machine learning algorithms (decision tree, random forest, and support vector machine) are applied to swing action classification, and a classification algorithm based on the integration idea is designed. The integrated learning algorithm can effectively distinguish the four swings, and the experimental results show that the integrated learning algorithm designed in this paper outperforms the traditional machine learning algorithm and achieves an average recognition accuracy of 91%.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] M. Y.-C. Jiang, M. S.-Y. Jong, W. W.-F. Lau, C. S. Chai, and N. Wu, "Using automatic speech recognition technology to enhance EFL learners' oral language complexity in a flipped classroom," *Australasian journal of educational technology*, vol. 37, no. 2, pp. 110–131, 2021.

[2] T. Sakulchit, B. Kuzeljevic, and R. D. Goldman, "Evaluation of digital face recognition technology for pain assessment in young children," *the clinical journal of pain*, vol. 35, no. 1, pp. 18–22, 2019.

[3] M. Li, X. Yu, K. H. Ryu, S. Lee, and N. Theera-Umpon, "Face recognition technology development with Gabor, PCA and SVM methodology under illumination normalization condition," *Cluster Computing*, vol. 21, no. 1, pp. 1117–1126, 2018.

[4] Y. Meng, C. Ye, S. Yu, J. Qin, J. Zhang, and D. Shen, "Sugarcane node recognition technology based on wavelet analysis," *Computers and electronics in agriculture*, vol. 158, pp. 68–78, 2019.

[5] K. Zhao, T. He, S. Wu et al., "Application research of image recognition technology based on CNN in image location of environmental monitoring UAV," *eurasip journal on image and video processing*, vol. 2018, no. 1, Article ID 150, 11 pages, 2018.

[6] D. Allegra, S. Battiato, A. Ortis, S. Urso, and R. Polosa, "A review on food recognition technology for health applications," *Health Psychology Research*, vol. 8, no. 3, pp. 9287–9297, 2020.

[7] L. Li, X. Mu, S. Li, and H. Peng, "A review of face recognition technology," *IEEE access*, vol. 8, pp. 139110–139120, 2020.

[8] B. Gao, K. Cai, T. Qu, Y. Hu, and H. Chen, "Personalized adaptive cruise control based on online driving style recognition technology and model predictive control," *IEEE transactions on vehicular technology*, vol. 69, no. 11, pp. 12482–12496, 2020.

[9] Y. Li, Z. Chu, and Y. Xin, "Posture recognition technology based on kinect," *IEICE transactions on information and systems*, vol. E103.D, no. 3, pp. 621–630, 2020.

[10] S. Yin, H. Fang, and X. Hou, "Athlete's respiratory frequency and physical energy consumption model based on speech recognition technology," *International journal of speech technology*, vol. 23, no. 2, pp. 389–397, 2020.

[11] Z. Zhang, Z. Li, H. Liu, T. Cao, and S. Liu, "Data-driven online learning engagement detection via facial expression and mouse behavior recognition technology," *Journal of Educational Computing Research*, vol. 58, no. 1, pp. 63–86, 2020.

[12] L. M. Merola, C. Lum, and R. P. Murphy, "The impact of license plate recognition technology (LPR) on trust in law enforcement: a survey-experiment," *Journal of experimental criminology*, vol. 15, no. 1, pp. 55–66, 2018.

[13] C. Rui, "Application of face recognition technology based on CA algorithm in intelligent residential property management," *Journal of intelligent and fuzzy systems*, vol. 35, no. 3, pp. 2909–2919, 2018.

[14] L. Stark, A. C. Stanhaus, and D. L. Anthony, ""I don't want someone to watch me while I'm working": gendered views of facial recognition technology in workplace surveillance," *Journal of the Association for Information Science and Technology*, vol. 71, no. 9, pp. 1074–1088, 2020.

[15] W. Cai, X. Wen, Q. Tu, and X. Guo, "Research on image processing of intelligent building environment based on pattern recognition technology," *Journal of Visual Communication and Image Representation*, vol. 61, pp. 141–148, 2019.

[16] M. Andrejevic and N. Selwyn, "Facial recognition technology in schools: critical questions and concerns," *Learning media and technology*, vol. 45, no. 2, pp. 115–128, 2020.

[17] K. Crawford, "Halt the use of facial-recognition technology until it is regulated," *Nature*, vol. 572, no. 7771, pp. 565–565, 2019.

[18] X.-D. Huang, C.-Y. Wang, X.-M. Fan, J.-L. Zhang, C. Yang, and Z.-D. Wang, "Oil source recognition technology using concentration-synchronous-matrix- fluorescence spectroscopy combined with 2D wavelet packet and probabilistic neural network," *Science of the total environment*, vol. 616-617, pp. 632–638, 2018.

[19] L. Qin, N. Yu, and D. Zhao, "Applying the convolutional neural network deep learning technology to behavioural recognition in intelligent video," *Tehnicki vjesnik technical gazette*, vol. 25, no. 2, pp. 528–535, 2018.

[20] V. Cantino, E. Giacosa, S. Alfiero, S. M. R. Shams, and A. Ferraris, "Introduction: smart tourism businesses (sustainability, measurability, awareness, recognition, & technology)," *Tourism analysis*, vol. 24, no. 3, pp. 261–263, 2019.

[21] F. Cheng, H. Zhang, W. Fan, and B. Harris, "Image recognition technology based on deep learning," *Image Recognition Technology Based on Deep Learning. wireless personal communications*, vol. 102, no. 2, pp. 1917–1933, 2018.