Annals of Emerging Technologies in Computing (AETiC)
Vol. 9, No. 3, 2025

Research Article

Optimized Lightweight CNN for Error Action Recognition in Physical Education Teaching

Shu Zhang[1] and Jacklyn Anne D. Toldoya[2,*]

[1]Hanjian Normal University, China
zs84107341@163.com
[2]Pamantasan ng Lungsod ng Maynila, Philippines
jadtoldoya@163.com
*Correspondence: jadtoldoya@163.com

Received: 16th September 2024; Accepted: 29th Month 2025; Published: 1st July 2025
Abstract: With the improvement of living standards and the enhancement of people's health awareness, participation in sports activities has received widespread attention. Traditional sports equipment, especially in educational environments, lacks the necessary technological advancements to provide precise guidance. The gap between this demand and available resources often leads to incorrect teaching methods, which may hurt students' sports training. To address these challenges, this paper proposes a motion action recognition system utilizing lightweight convolutional neural networks (CNN). This method effectively reduces the noise of sensor data, improves the accuracy and reliability of data, and lays a solid foundation for model training through one-dimensional median filtering and Z-score standardization. Optimize the CNN architecture by adjusting key parameters such as network structure, convolution kernel size, and convolution stride, which are fine-tuned based on training data to maximize the model's recognition ability. The research results provide valuable insights into the effectiveness of teaching techniques and targeted feedback for improving sports training. After sufficient training, the system performed excellently on test data, accurately identifying erroneous movements across various sports actions, particularly in critical areas such as stroke movements, with an accuracy rate of up to 97.82% and an RMSE as low as 1.71%. These results demonstrate the model's high precision and robustness. The system has shown great potential in addressing the current shortage of professional coaches by providing automatic, real-time feedback on motion accuracy.

Keywords: ***CNN; Error Action Recognition; Lightweight Deep Learning; Sports Teaching; Sports Training***

---

1. Introduction

Strengthening physical fitness through physical exercise has been a general concern of society and the masses [1-2]. Since national sports have been widely promoted and popularized, sports can not only help people release pressure but also achieve the purpose of physical fitness, making them loved by the masses. Firstly, this kind of sport requires a relatively simple venue, with relatively low requirements for the physical quality of the athletes, and there is no physical contact during the exercise, which can help avoid injuries caused by violent body collisions. Additionally, racket sports can effectively reduce cholesterol, prevent heart disease, and relieve anxiety, depression, and life pressure. Therefore, racket sports are widely loved by people. Sports teaching is an important way to improve students' physical fitness and sports skills. However, in actual teaching, the occurrence of incorrect actions is inevitable due to factors such as individual differences among students, teacher level, and teaching methods. Incorrect movements may cause injuries to students, affecting their mastery of motor skills and physical health. Therefore, detecting and correcting incorrect movements timely has become an important issue in sports teaching. Deep learning technology has achieved significant results in fields such as image classification, speech recognition, and

Shu Zhang and Jacklyn Anne D. Toldoya, "Optimized Lightweight CNN for Error Action Recognition in Physical Education Teaching", *Annals of Emerging Technologies in Computing (AETiC)*, Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 85-101, Vol. 9, No. 3, 1st July 2025, Published by International Association for Educators and Researchers (IAER), DOI: 10.33166/AETiC.2025.03.006, Available: http://aetic.theiaer.org/archive/v9/v9n3/p6.html.

natural language processing. Among them, Convolutional Neural Networks (CNN), as a common deep learning model, have been widely used in image processing and computer vision tasks. The lightweight CNN model is an improvement on the traditional CNN model to reduce model complexity and computational costs while maintaining high classification accuracy [3].

In recent years, with China's remarkable achievements in the field of science and technology, a variety of high-tech products have flooded into people's daily lives. On the one hand, they bring convenience to people's lives, and on the other hand, they also affect people's health. Especially with the increasing popularity of electronic products among college students, the impact on their health is concerning. According to a recent survey, college students spend an average of 7.49 hours a day on electronic devices, and too much screen time is common, with students spending 75 percent of their time in dormitories. This directly leads to very little average daily exercise time for college students, so the physical fitness of college students in China continues to decline [4]. College students are the advanced force of the country's youth and shoulder the development. China's economic performance is satisfactory, and it has become the second-largest economy in the world. Nowadays, when people are rich in material life, China's sports industry develops slowly. For example, the sports teaching system is not perfect, and the industrial structure is not coordinated, among other problems that need to be solved. In the new era, China's sports undertakings should speed up the construction of sports power and give full play to the important role of science and technology elements in the construction [5]. Sports powers such as America, Japan, and Russia, using science and technology to promote the development of sports, in the downward trend in Brazil's Olympic Games gold medals, China's national team at the Tokyo Olympic Games will be faced with a severe situation, to improve the scientific training of the leading role, requiring the development of sports training science. With the increasingly fierce competition in the field of competitive sports, the role of science and technology in promoting competitive sports is becoming more and more obvious. Relevant research is in full swing. The application of cutting-edge technologies such as computer technology, virtual reality technology, and motion sensing networks has seen explosive growth in sports and promoted the prosperity of the relevant smart sports industry [6].

Once a training plan is established, sticking to execution is the key to success. Conduct regular training according to the plan and record the progress of each training session for subsequent evaluation and adjustment. Regularly evaluate your progress and adjust your training plan based on actual circumstances. This may include increasing or decreasing training intensity, changing training frequency or type, etc. If your training plan is not effective or you feel confused, seeking advice from a professional coach or exercise doctor can be helpful. They can provide personalized guidance and advice to help you optimize your training plan. In addition to the training plan, ensure that the body is given sufficient rest time to recover and adapt to the training stimulus. Reasonable nutrient intake is also crucial, ensuring that the body receives sufficient energy and nutrients to support training and recovery. As interest in sports continues to soar, so does the demand for specialized coaching [7]. In addition to its application in the field of education, lightweight CNN models also have broad commercial feasibility. With the increasing demand for health and exercise among people, the sports training and fitness industry is facing unprecedented development opportunities. Applying lightweight CNN models to these industries can provide users with more professional and personalized training guidance services, thereby attracting more user groups. At the same time, the model can also provide data support for sports equipment manufacturers, helping them develop sports equipment that better meets user needs. In summary, the application of lightweight CNN models in physical education teaching has broad prospects and enormous potential. In sports teaching, an effective training plan is crucial for improving students' sports skills and physical fitness. The traditional way of formulating training plans usually relies on the coach's on-site observation and experience, but this will consume a lot of time and energy. In addition, due to the limited resources of coaches, many sports enthusiasts are unable to receive excellent training guidance. The lightweight CNN model is an improved CNN model aimed at reducing model complexity and computational costs while maintaining high classification accuracy. Applying it to sports teaching can effectively identify and correct incorrect movements, thereby improving teaching quality and reducing sports injuries. The athlete's jumping ability, reaction time, and other deep-seated information by the naked eye. Finally, due to the limited energy and observation angle of coaches, they may misjudge the movements of athletes in training, especially in competitive sports where the difference between foul movements and standard movements is very slight

[8]. Human motion recognition is a crucial research direction in the field of computer vision. The advantage of a human motion recognition method based on video and image data is that it can obtain rich human motion information, such as position, posture, action, etc. At the same time, it can understand the human motion trajectory and environment by analyzing video or image sequences. However, this method also has some drawbacks, such as being easily affected by factors such as lighting and occlusion and requiring a large amount of video or image data for training to achieve good recognition results [9-10].

Motion recognition based on sensor data refers to the use of sensor data acquisition equipment to collect the inertial sensor data generated by the movement of the athlete. When collecting such data, the acquisition equipment with integrated inertial sensors is usually fixed at the key moving parts of the athlete or installed on the sports equipment. The acquisition device will collect the motion sensor data to the PC terminal or mobile phone terminal through wired or wireless transmission. Researchers are also increasingly using inertial sensors to collect data from athletes for motion recognition, mainly including motion tracking, health monitoring, smart home, and other fields. Sensors are embedded in mobile devices to collect the wearer's activity data. The most common sensors are accelerometers and gyroscopes [11]. On this basis, the collected sensor data can be used for human movement recognition. The traditional recognition method is to use machine learning technology, which generally involves manual feature extraction. This method relies heavily on experts with professional knowledge to select features. System performance will be affected by the differences in features selected by different experts, so this method has strong subjectivity. It takes a long time to extract a valid feature manually, and it usually takes five to ten years to extract a feature that everyone agrees on, resulting in inadequate information. At the same time, machine learning methods generally use linear discriminant functions, which are not suitable for dealing with data with complex characteristics [12].

Deep learning methods can use multi-layer nonlinear information processing to quickly and automatically learn effective feature representations from complex real data, which has excellent generalization ability and robustness [13-14]. More and more researchers are paying attention to deep learning, and it has been applied in the field of human movement recognition. Both convolutional neural networks (CNN) and deep ConvLSTM networks have good recognition accuracy when deep learning is used. CNN can extract features from the original data by continuously abstracting the features of the input data through the method of stacked convolutional neural networks. The deep ConvLSTM network cleverly combines the feature extractor, and the LSTM recurrent neural network acts as a classifier. LSTM is a recurrent neural network (RNN), which differs from ordinary recurrent neural networks mainly in that it uses a unique threshold structure to learn and store long-term information. As a result, LSTM can learn the temporal correlation between data. Although these models perform well in identifying incorrect movements in sports, they are designed based on the rich prior knowledge accumulated by the designers and their continuous attempts. The topology structure of the neural network can easily affect the performance of the network. In the current field, designers need to adjust a large number of models' hyperparameters to obtain a well-behaved architecture, and at the same time, they need to manually try different topologies until they find an appropriate network structure [15].

To sum up, due to the continuous popularity of mobile intelligent devices, human motion recognition using sensor data has become an important research field. Due to their ability of automatically extract features extraction, deep neural network algorithms have been successfully applied in the field of sports teaching action recognition. However, the existing neural network algorithms still have some defects, mainly due to the training difficulty of CNN [16] and the dependence of neural network structure design on the prior knowledge of experts. Therefore, research on false action recognition based on the CNN model has important practical significance and research value.

2. Related work

With the research of video technology, domestic video analysis platforms have developed rapidly. The intelligent cloud platform and intelligent game system for football match data will analyze and visualize player data. The system covers all Olympic events and provides athlete profiles, historical data, and schedule information. However, it lacks analysis of athlete trajectory tracking and standardized movements. To solve this problem, Dongrui Software has developed a golf-assisted training system.

Heilmann and Witte [17] tested athletes' technical movements by analyzing images of their velocity angles at specific moments. And compare and analyze the data of professional athletes in the database, identify gaps, and strengthen training. The development of sports video analysis systems has promoted the advancement of video analysis technology. To solve the problems of low image resolution and non-static videos in sports videos, SVM can extract descriptor operators for motion recognition. Understanding the temporal structure in broadcast baseball videos is highly valuable for fine-grained activity recognition. By using fine-grained activity detection to predict pitch velocity and pitch type, a new method is provided to solve the problem of coarse-grained and fine-grained motion recognition in video retrieval. Tian *et al.* [18] used CNN to extract images, preprocessed video images, and used LSTM based bone recognition algorithm to detect body key points. It has developed a deep big data motion object detection system for motion image recognition. According to the development and technological research of sports analysis platforms, the mainstream commercial systems in basketball games currently mainly use foreign analysis systems. The domestic competition analysis system is mainly focused on football matches, golf matches, and other competitions, which require a lot of manual assistance work and have a low degree of intelligence. Therefore, designing and developing a basketball auxiliary training system has a promoting effect on the auxiliary training of domestic basketball teams.

The HCN model directly reshapes bone point data into a 2D matrix and uses CNN to extract features and recognize actions. However, graph neural networks are highly adept at handling non-Euclidean data and modeling relationships between nodes, thus rapidly surpassing previous methods in the field of bone point action recognition. Cheng *et al.* [19] represented human joints as graph nodes and skeletal connections between nodes as graph edges. The use of graph convolution operations to update node features benefits from the ability to model relationships between nodes and improves the accuracy of motion action recognition. The adjacency matrix is immutable during the training process and cannot establish semantic connections between nodes. To overcome this limitation, Xue *et al.* [20] proposed using self-attention mechanisms to capture semantic relationships between nodes. Another advantage of the self-attention mechanism is that it can calculate the attention weights between the target node and all potential nodes, thereby expanding the spatial receptive field of the nodes. Therefore, it is necessary to select corresponding action features for description based on the differences between scenes. The following is a summary of the current research status of action characteristics.Click or tap here to enter text..

Based on human body model features. This method mainly describes states and represents human behavior with constantly changing postures. This function can more accurately describe human motion and avoid errors caused by scene changes in motion recognition by collecting and calculating the angles formed between human body parts and bone joint trajectories. Based on the basic features of images. Local features and global features are two aspects based on the bottom features of an image. Considering the significant temporal and spatial changes in the local structure of the detected image, Zhang [21] used the Laplacian operator for spatiotemporal scale normalization calculation, and classified by calculating local descriptors and scales. Global features are used to represent the global region of interest, and the global features encoded to represent human motion information include spatiotemporal features, contour features, motion energy maps, etc. Although feature extraction methods for low-level images are widely used, there are significant errors in the same features of corresponding actions due to scene changes. Therefore, it is usually recommended to use methods based on human body construction model features. Body area networks can be used to identify various simple and complex movements. Nadeem *et al.* [22] estimated people's movement status and energy consumption, helping them to engage in more scientific exercise. Motion recognition can be combined with videos for automatic annotation and generating highlights by detecting key events in the video. The principle is to use inertial sensors to identify key actions and obtain their time labels, which are used to automatically mark and synchronously record player actions in videos, and generate highlights for video retrieval. The ultimate goal of motion recognition is to apply it to motion assisted training, that is, to use body area networks to monitor athletes' motion characteristics, such as limb posture, limb rotation angle, rotation speed, explosive power, etc.

Currently, some research is using smart wearable devices (mainly wristbands) to recognize movement movements. Zhang *et al.* [23] designed a wearable device for adolescent basketball dribbling to test excessive dribbling time and dribbling palm posture. Hussain *et al.* [24] established criteria for determining parameters related to poor dribbling attitude through a series of experiments. And using these parameter standards to determine the poor dribbling attitude of young basketball players, resulting in the inability to accurately recognize turning movements when only using wristbands, and the inability to accurately understand deep information such as the reaction speed of athletes when turning. However, due to the intense physical activity during basketball games, Android phones may detach and cannot ensure that the direction of the phone remains unchanged during basketball movements, thereby affecting the accuracy of basketball action recognition. Liu *et al.* [25] are dedicated to wearable sensor-based motion recognition, which plays an important role in improving motor skills. Therefore, this article mainly studies the recognition and correction of athletes' incorrect movements to improve the quality of athletes' training and practice [26]. The intelligent recognition of incorrect movements through computer vision technology during athlete training is of great significance for improving training efficiency and reducing sports injuries. Du [27] conducted computer vision simulation analysis on an artificial intelligence-based athlete error action recognition model to explore its feasibility and effectiveness in practical applications. In practical applications, the presence of random noise often interferes with the accuracy of action recognition. Therefore, improving the accuracy of aerobic gymnastics movement recognition in a random noise environment has become an urgent problem to be solved. Chen and Guo [28] analyzed the combination of Lv-3DCNN algorithm in random noise environment and its application in aerobic gymnastics action recognition. Gaussian Mixture Model (GMM) and Convolutional Neural Network (CNN) are two commonly used algorithms that play an important role in human activity detection and recognition. Karunya and Kumar [29] used GMM and CNN algorithms to detect and recognize human activity analysis in motion. Using GMM to preprocess video sequences, remove background interference, and more accurately detect and recognize human activities in motion. Afsar *et al.* [30] analyzed deep learning models for identifying sports activities in sports games using wearable sensors. Real time tracking and evaluation of badminton players' posture during exercise based on deep neural network posture estimation technology. He and Zhang [31] provided the recognition results and matching degree of athlete postures by comparing them with the standard posture library. So as to help coaches and athletes understand the standardization and accuracy of athletes' movements, and then carry out targeted training and adjustments.

In the field of action feature research, global features such as spatiotemporal features, contour features, and motion energy maps are widely used to represent human motion information. However, due to the issue of feature errors caused by scene changes, methods based on human body construction model features have gradually become mainstream. Models such as body area networks can recognize various simple and complex movements, estimate people's movement states and energy consumption, and provide strong support for scientific sports. In this context, this article proposes a method for identifying incorrect actions in physical education teaching based on a lightweight CNN optimization architecture. This method combines the efficiency and optimization strategies of lightweight CNN architecture, aiming to reduce computational complexity, improve recognition accuracy and real-time performance. Compared with existing methods, this method has the following advantages: it adopts a lightweight CNN architecture, reduces model parameters and computational complexity, and improves recognition speed and resource utilization. By introducing optimization strategies such as batch normalization and residual connections, the model's generalization ability and stability have been enhanced. Identifying incorrect movements in physical education teaching provides timely feedback and correction guidance for athletes, which helps improve the quality of training and practice.

3. Lightweight CNN for error action recognition in sports teaching

3.1. Lightweight deep CNN model

Lightweight deep CNN models typically use fewer layers and smaller filter sizes to reduce the number of model parameters and computational complexity. By using low-bit deep neural networks (such as 16-bit or 32-bit floating-point numbers), lightweight deep CNN models can further reduce the storage requirements and computational costs of the model. Lightweight deep CNN models can also be compressed through techniques such as pruning and knowledge distillation to reduce the number of model parameters and computational complexity. Lightweight deep CNN models can improve their performance by using various optimization techniques such as batch normalization and residual connections, as depicted in Figure 1.

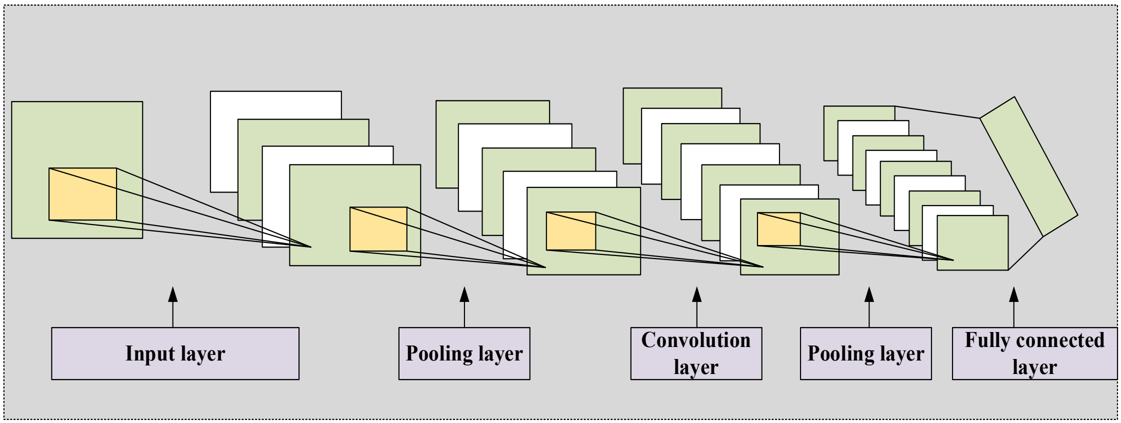

Figure 1. The typical schematic diagram of the CNN model

The core of lightweight deep CNN models lies in their streamlined architecture and optimized parameter settings. Firstly, we reduced the number of layers in the model by carefully selecting key layers to preserve necessary feature extraction capabilities while avoiding unnecessary computational overhead. Specifically, our model employs an appropriate number of convolutional layers that capture local features with smaller filter sizes (such as 3x3 or 5x5), thereby reducing the number of parameters and computational complexity. In addition, we adopted low bit deep neural network technology to represent the weights and activation values in the model as lower precision floating-point numbers (such as 16 bit or 32-bit floating-point numbers), which further reduces the storage requirements and computational costs of the model. This technology not only reduces memory usage, but also accelerates the inference speed of the model, making it more suitable for real-time applications.

In order to further compress the model and reduce computational complexity, we also employed techniques such as pruning and knowledge distillation. Trimming techniques simplify the model structure by removing weights and neurons that contribute less to model performance. Knowledge distillation utilizes a large and complex teacher model to guide the learning process of a simplified student model, enabling the student model to maintain high recognition accuracy while having fewer parameters and faster inference speed. In terms of optimizing model performance, we have adopted techniques such as batch normalization and residual connections. Batch normalization accelerates the convergence speed and improves the stability of the model by normalizing input data. Residual connections alleviate the gradient vanishing problem in deep networks by introducing skip connections, thereby improving the training efficiency and recognition performance of the model. The computational steps involved in convolution are displayed below:

$$CONV_{(ij)} = \sum_{i}^{m-1} \sum_{j}^{n-1} u_{ij} \times w + b \cdots (i = 1,2 \cdots m-1; j = 1,2 \cdots n-1) \tag{1}$$

In CNN, adding activation function layers is necessary because they introduce nonlinearity to the model, enabling it to better learn and understand the complex features of the data. The sigmoid function maps any value between 0 and 1, often used in the output layer to represent probability. However, when the input is very large or very small, there may be a problem of gradient vanishing. The Tanh function is similar to the sigmoid, mapping any value between -1 and 1. Its advantage is that when the absolute value of the input is very large, the derivative still maintains a relatively large value, so there is no problem of gradient disappearance. The full name of the ReLU function is the Rectified Linear Unit, which is very suitable for deep learning networks, especially when there are many layers in the network. ReLU outputs the input value when it is greater than 0, and 0 when it is less than 0. Its advantages are fast calculation speed and effective prevention of gradient disappearance.

$$f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

The tanh function is

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{3}$$

The ReLu function is

$$f(x) = max(0, x) \tag{4}$$

The shortcomings of step 4 in the calculation process also need to be noted. When the input value is less than 0, the output of the ReLU function is 0, which is called a "dead zone". This may lead to some problems when training deep neural networks. For example, if the input of a neuron remains negative, the neuron will remain in a "dead" state and cannot be activated, which can affect the performance of the network. Additionally, although ReLU has computational advantages in certain situations, the presence of non-zero thresholds during training may lead to gradient vanishing or exploding problems.

$$f(x) = \begin{cases} x, x \geq 0 \\ \alpha x, x < 0 \end{cases} \tag{5}$$

The equations for Sig and Tanh are displayed below:

$$\begin{cases} sig(x) = \dfrac{1}{1 + exp(-x)} \\ tanh(x) = \dfrac{exp(x) - exp(-x)}{exp(x) + exp(-x)} \end{cases} \tag{6}$$

Where x is the input.

$$h_{w,b}(x_i) = \begin{bmatrix} p(y_i = 1 \mid x_i; w, b) \\ p(y_i = 2 \mid x_i; w, b) \\ p(y_i = 3 \mid x_i; w, b) \\ \cdots \\ p(y_i = n \mid x_i; w, b) \end{bmatrix} = \frac{1}{\sum_{j=1}^{n} e^{w_j x_i + b_j}} \begin{bmatrix} e^{w_1 x_i + b_1} \\ e^{w_2 x_i + b_2} \\ e^{w_3 x_i + b_3} \\ \cdots \\ e^{w_n x_i + b_n} \end{bmatrix} \tag{7}$$

The w is the weights, and b is the bias. The cross-entropy (CE) formula is as follows

$$loss = -\frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{n} y_{ji} \, log(\hat{y}_{ji}) \tag{8}$$

$$\theta := \theta - \alpha \frac{\partial}{\partial \theta} J(\theta) \tag{9}$$

Where θ is the parameter. The Adam optimizer is defined as:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{10}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \tag{11}$$

Where g_t stands for the object function and m_t and v_t display the parameters. β_1 and β_2 display the regularization parameters. The wrong actions identification accuracy can be measured *RMSE* index as:

$$RMSE = \sqrt{\frac{1}{l}\sum_{i=1}^{l} (y_i - \hat{y}_i)^2} \tag{12}$$

Among them, l is the number of samples, y_i is the true value of the ith sample, and $\hat{y}_i$ is the predicted value of the ith sample. The smaller the RMSE value, the higher the prediction accuracy of the model.

The final structure in the CNN network can be divided into four categories: the overall structure, cell-based structure, hierarchical structure, and network morphism structure. The overall structure is similar to the structure that is usually designed manually. When the number of network layers deepens, if the overall structure is to be directly searched, the search space will become large, and a lot of time and computing resources must be invested to search for the optimal structure. As depicted in Figure 2, the normal cell is on the left, the reduced cell is in the middle, and the final network structure, stacked in sequence, is on the right. Cells can also be stacked in more complex ways, such as by replacing layers with cells in a multi-branch structure. Since the advent of cell-based structures, cell-based search spaces have been successfully applied in many recent works. However, the problem with choosing the right macro architecture is how many cells should be used and how to connect them to build the final structure. In theory, cells can be combined arbitrarily, for example, in the multi-branch structure mentioned earlier, simply replacing layers with cells. Ideally, the macro network structure and the micro network structure (cell structure) should be optimized jointly, rather than just the micro network structure. Otherwise, a lightweight CNN model cannot be constructed. People also need to design the macrostructure manually, which increases the design and operation burden of the model. In addition, some researchers also use network morphism, which is based on the existing network structure and obtained through continuous updating, rather than designing a new network from scratch, to learn lightweight neural networks.

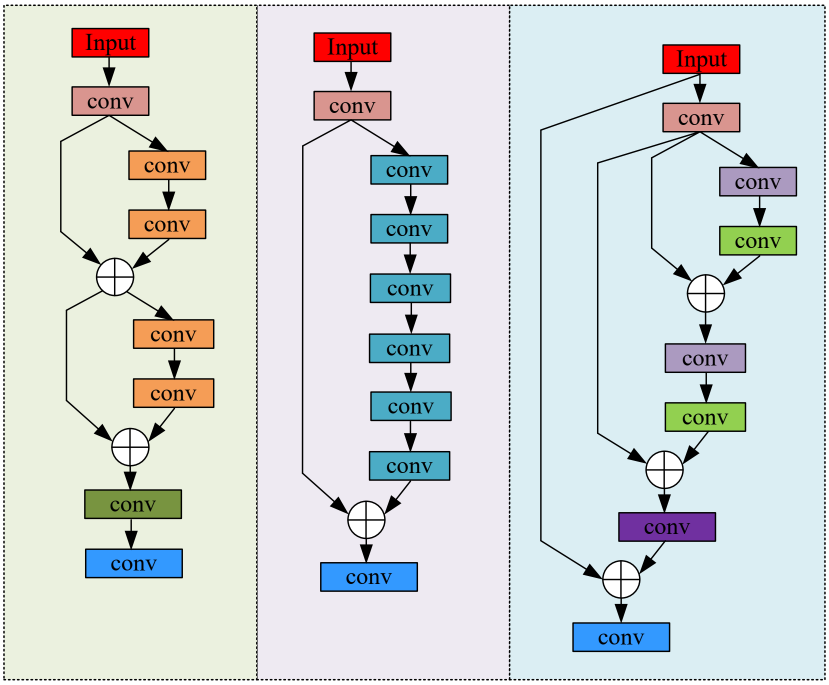

Figure 2. Cell-based structure description

Based on the above discussions, the lightweight CNN model and its application in error action recognition in sports teaching are depicted in Figure 3.

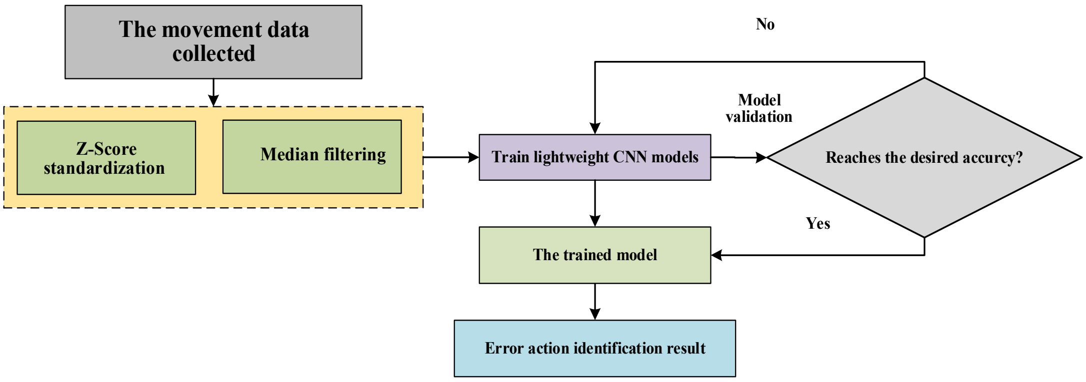

Figure 3. The framework of the CNN-based error action recognition in sports teaching

4. Experimental results and analysis

4.1. Experimental data introduction

Currently, most researchers use human-based bone images to describe changes in human nodes between video frames, which can be quickly and accurately estimated based on deep data. However, bone data is represented as vector sequences or two-dimensional grids, and the dependencies between associated joints cannot be fully expressed. The experiment used a high-performance computing server equipped with an Intel Xeon E5-2680 v4 processor (2.4GHz, 14 cores), NVIDIA Tesla P100 GPU (16GB of video memory), and sufficient memory (128GB DDR4 ECC) to support large-scale data processing and model training. The software environment operating system is Ubuntu 18.04 LTS, the deep learning framework is TensorFlow 2. x, and the Python version is 3.7. These software environments provide a stable and efficient platform for model construction, training, and evaluation. During the model training process, we fully utilized the parallel computing power of GPU to accelerate the convergence speed of the model. Meanwhile, through reasonable memory management and multi-threaded technology, we ensure real-time and efficient data processing.

In this study, we used multiple widely recognized sports datasets to validate the effectiveness of our proposed method, including UCF Sport, Sport LM, NCAA basketball dataset, and volleyball dataset. These datasets not only contain rich human motion information, but are also known for their diversity and practical application value. The UCF Sport dataset covers various sports scenarios such as running, jumping, swimming, etc. Each video clip is labeled with the corresponding sports category. We selected action segments that have certain similarities with cheerleading to better simulate and analyze the energy metabolism mechanism of cheerleaders. The Sport LM dataset focuses on long-term sports activity recognition and includes various indoor and outdoor sports events. We use this dataset to evaluate the performance of the model in handling continuous and complex motion sequences. The NCAA basketball dataset records real-time sensor data from college basketball games, including player position, speed, acceleration, and other information. We selected some data related to cheerleading sports, such as fast movements, jumps, and other movements, to further enrich our experimental sample. Generally, the collected original sensor data should be divided into sequences of the same length through sliding windows, with each equal-length sequence treated as a sample and corresponding labels marked for each sample. Then, data preprocessing should be carried out for the segmented data. The results are mapped to [0,1]. Standardization refers to data normalization based on the mean and standard deviation of the original data.

4.2. Experimental Results Analysis

Figure 4 illustrates the changes in body pressure of the subject. As can be seen from the figure, the pressure on athletes varies with the position of the human body in a certain sport. The total pressure increases from small to large and then decreases, and the bone pore pressure shows the same changing trend, while the effective pressure shows the opposite changing trend. When athletes perform the correct movements, these pressure values will follow a specific pattern of changes; And when the action is wrong, this pressure pattern will be broken. The correlation between pressure changes and action correctness provides effective feature input for lightweight CNN models, thereby improving the accuracy of error action recognition. Although the three pressures exhibit different trends, the total pressure is equal to the effective pressure plus the gap pressure. If the sports action is wrong, the pressure will not follow such a pattern, which lays a foundation for the effective identification of incorrect actions in subsequent sports teaching.

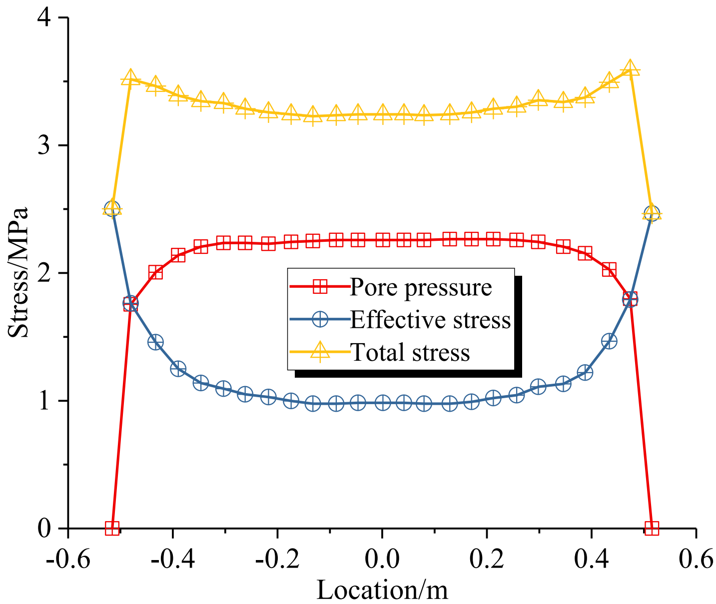

Figure 4. Stress change diagram

Figure 5 shows the impact of the number of hidden layer nodes on model performance. As the number of nodes increases, the RMSE value first decreases and then increases, indicating the existence of an optimal number of nodes that maximizes the model's error recognition accuracy. This discovery emphasizes the importance of parameter selection in model design. By carefully adjusting these parameters, we can optimize the performance of the model to perform well in different motion scenarios. From the figure, it is evident that as the number of nodes in the hidden layer increases, the value of RMSE on 10, 11, and 12 first increases and then decreases. In other words, the accuracy of error action identification of the lightweight CNN model improves initially and then declines. The optimal numbers of hidden nodes (the highest error identification accuracy) are 60, 75, and 135, respectively, which are marked with red squares in Figure 5.

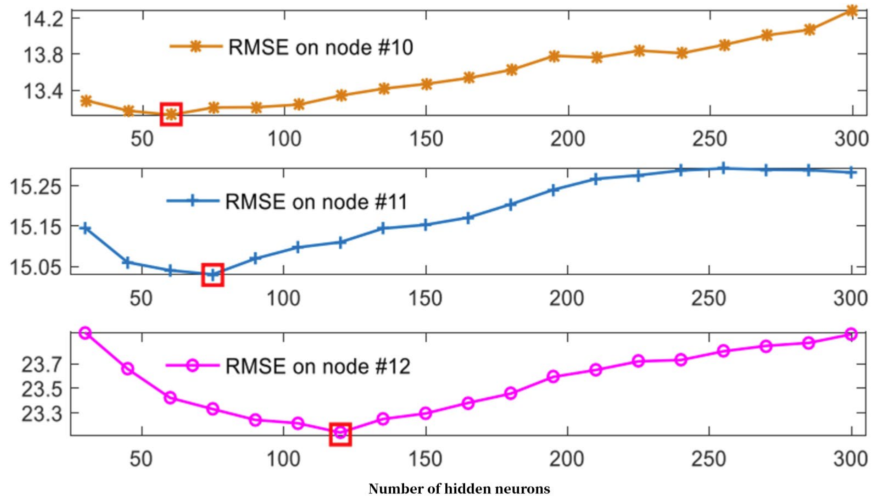

Figure 5. The parameter selection results regarding the RMSE value

Figure 6 shows the graphical results of the model feature optimization matrix under basketball motion. Among them, 1-6 classifies the dribbling and shooting of basketball into datasets. To address the issue of inconsistent standardization of actions, a standardized set of actions can be defined, and the actions of each player and referee can be standardized. This can be achieved through motion capture techniques and machine learning algorithms. When training and optimizing models, a large amount of data can be used for training, and techniques such as cross-validation can be used to evaluate the performance of the model. Meanwhile, the performance of the model can be optimized by adjusting its parameters and structure. Ten percent of the shots were mistakenly identified as having no balls. This indicates that the model may be influenced by background or other factors when dealing with shooting actions, leading to misjudgments. Twelve percent of the no-ball movements were mistakenly identified as running dribbles. This may be due to the similarity in certain aspects between no-ball and dribbling movements. These data indicate that there is a certain degree of misidentification rate in basketball action recognition, especially in distinguishing between actions such as shooting, layup, and no-ball. This may be related to factors such as the diversity of model training data, feature selection, and the complexity of the model. To improve the accuracy of the model, methods such as increasing the amount of training data, optimizing feature extraction methods, and improving the model structure can be considered. At the same time, it is also possible to consider introducing human experts to provide more accurate action labels and guide model improvement. To enhance the reproducibility of the research, we will elaborate on the experimental setup and evaluation criteria in detail. We used large-scale datasets for training and optimizing the model, and comprehensively evaluated its performance using techniques such as cross-validation. By adjusting the parameters and structure of the model, we successfully optimized its performance and improved its accuracy in basketball action recognition tasks. To improve the accuracy of the model, we plan to take the following measures: increase the amount of training data to cover more diverse action scenarios. Optimize feature extraction methods to extract more discriminative action features. Improve the model structure to enhance its generalization ability and robustness.

To verify the validity of the proposed method, this study tested the eight types of badminton swing movements (A1-A8: backhand floater, backhand jump shot, forehand floater, forehand lob, forehand jump shot, backhand smash, forehand smash and forehand net drop) on different methods (BP neural network, support vector machine (SVM), lightweight CNN). The recognition results of different methods for different actions are compared in Figure 7. Even for a backhand jump shot, it is 0.2% higher than that of the BP neural network. For the remaining 7 types of actions, the lightweight CNN model outperforms the comparison algorithm significantly. Especially for the jump shot, the proposed method reached the highest 97.82%. The main reason is that there is a great difference between the jump ball movement and other movements, and the proposed method is the depth model, which has a stronger recognition ability.

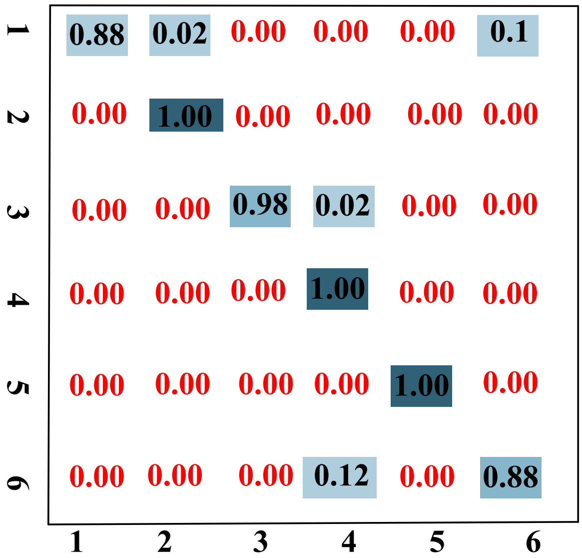

Figure 6. The identification accuracy of different wrong actions in the NCAA basketball dataset

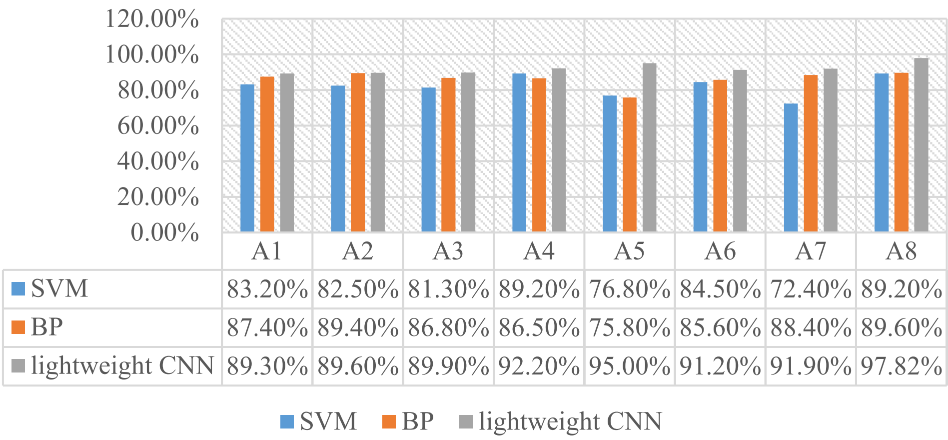

Figure 7. The identification accuracy by (a) SVM, (b) BP, (c) lightweight CNN

In addition to recognition accuracy, recognition time is also an important indicator of model performance. The average running time of different models for ten trials is shown in Figure 8. The figure reveals that the running time of SVM is the shortest in both the training and test sets, mainly because the model is shallow. The running speed is slightly slower, mainly because it produces deep network results. However, the running time of the proposed model is completely within the acceptable range under the current computing level and computing power, which proves its superior performance.

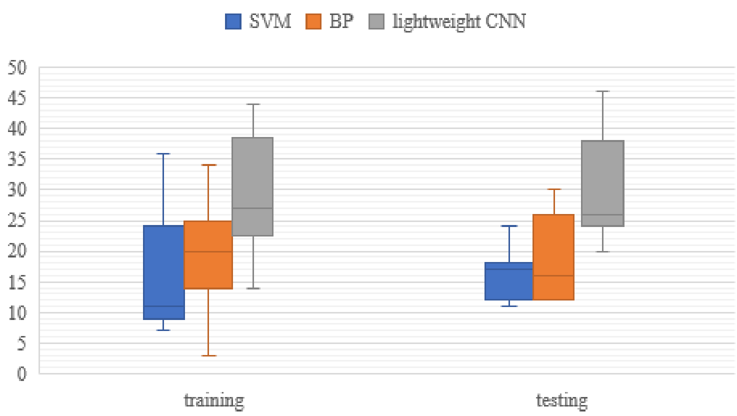

Figure 8. The identification time by (a) SVM, (b) BP, (c) lightweight CNN

The above results are all recognized in the same kind of actions. If there are different actions. Figure 9 shows the recognition results of similar movements in three different sports (A represents the scrum in basketball, B represents the jump in badminton, and C represents the serve in volleyball). As can be seen from the figure, the RMSE values of A, B, and C are 1.71%, 1.92%, and 2.21% respectively. The experiment has a good recognition effect on similar wrong actions in different movements.

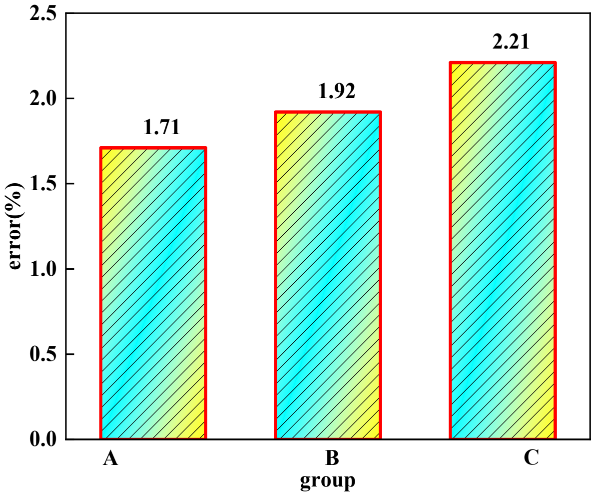

Figure 9. The recognition accuracy of three different kinds of motion

The lightweight CNN model adopts a deep network structure, which enables the model to learn more complex feature representations. Compared to traditional BP neural networks and SVM, deep networks can capture nonlinear relationships in data, thereby improving recognition accuracy. Especially in the recognition of badminton swing movements, due to the complexity and diversity of movements, deep networks can better learn the subtle differences between different movements. The lightweight CNN model has been optimized in parameter selection, such as adjusting the number of hidden layer nodes. These optimization measures enable the model to maintain high recognition accuracy while reducing computational complexity, thereby improving operational efficiency. Lightweight CNN models have powerful feature extraction capabilities. Through structures such as convolutional layers and pooling layers, the model can automatically extract useful features for recognition tasks from raw data. These features are not only representative, but also can reduce the interference of noise, thereby improving the accuracy of recognition. The diversity of the dataset has a significant impact on the performance of the model in the recognition task of badminton swing movements. A dataset containing multiple actions and scenarios can enable the model to learn more comprehensive feature representations, thereby improving the generalization ability of recognition. In this study, we used a dataset containing eight types of badminton swing movements, which enabled the model to learn the differences and connections between different movements.

5. Conclusions

This study proposes a sports teaching error action recognition system based on lightweight convolutional neural network (CNN) to address the shortage of professional coach resources and the lack of precise guidance from traditional sports equipment in the current sports teaching environment. The system significantly reduces the noise of sensor data and improves the accuracy and reliability of the data through effective data preprocessing methods, including one-dimensional median filtering and Z-Score standardization. By optimizing the CNN architecture, including adjusting key parameters such as network structure, convolution kernel size, and convolution stride, we have successfully constructed a model that can efficiently identify erroneous actions. The research results show that the system performs well on test data, especially in identifying erroneous actions in key areas such as stroke movements, demonstrating extremely high accuracy and robustness. This achievement provides valuable insights for improving the teaching techniques of sports training, providing targeted feedback, and enhancing the learning efficiency of sports skills. In addition, the system provides an effective alternative solution to the current shortage of professional coaches by providing automatic feedback on motion accuracy, demonstrating enormous potential for application.

However, this study also has certain limitations. Firstly, although we have achieved significant results on the test data, the performance of the system may be affected by various factors such as sensor type, data quality, motion type, and individual differences. Therefore, in future research, we need to further validate the system's generalization ability and explore how to adapt to the needs of different scenarios and individuals.

Acknowledgement

2021 Hubei Provincial Teaching Research Project of Higher Education Institutions "Reform and Practice of Physical Education Professional Talent Cultivation Mode Based on Core Literacy Orientation", Project No: 2021476

CRediT Author Contribution Statement

Jacklyn Anne D. Toldoya: Writing – original draft, Conceptualization, Supervision, Project administration; Shu Zhang: Methodology, Software, Validation.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] Eva Ellmer, Steven Rynne and Eimear Enright, "Learning in action sports: A scoping review", *European Physical Education Review*, Print ISSN: 1356-336X, Online ISSN: 1741-2749, pp. 263–283, Vol. 26, No. 1, 17 March 2020, Published by Sage Publications, DOI: 10.1177/1356336X19851535, Available: https://journals.sagepub.com/doi/abs/10.1177/1356336X19851535.

[2] Catherine Palmer and Zack Dwyer, "Good running?: The rise of fitness philanthropy and sports-based charity events", *Leisure Sciences*, Print ISSN: 0149-0400, Online ISSN: 1521-0588, pp. 609–623, Vol. 42, No. 5–6, 9 September 2019, Published by Taylor & Francis, DOI: 10.1080/01490400.2019.1656122, Available: https://www.tandfonline.com/doi/abs/10.1080/01490400.2019.1656122.

[3] Azizxon Aloxonovich Tillyaxodjayev, "The importance of marketing communications in the development of kind of sports in Uzbekistan", *South Asian Journal of Marketing & Management Research*, Online ISSN: 2249-877X, pp. 18–29, Vol. 10, No. 12, 2020, Published by Indian Journals, DOI: 10.5958/2249-877X.2020.00095.8, Available: https://www.indianjournals.com/ijor.aspx?target=ijor:sajmmr&volume=10&issue=12&article=002.

[4] Lior Dar, Shmuel Tiosano, Abdulla Watad, Nicola Luigi Bragazzi, Devi Zisman *et al.*, "Are obesity and rheumatoid arthritis interrelated?", *International Journal of Clinical Practice*, Online ISSN: 1742-1241, p. e13045, Vol. 72, No. 1, January 2018, Published by WILEY Online Library, DOI: 10.1111/ijcp.13045, Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/ijcp.13045.

[5] Zhengfeng Zhang, Mengwei Wang, Ze Xu, Yumin Ye, Sining Chen *et al.*, "The influence of Community Sports Parks on residents' subjective well-being: A case study of Zhuhai City, China", *Habitat International*, Print ISSN: 0197-3975, Online ISSN: 1873-5428, pp. 102439, Vol. 117, No. 1, November 2021, Elsevier, DOI: 10.1016/j.habitatint.2021.102439, Available: https://www.sciencedirect.com/science/article/abs/pii/S0197397521001284.

[6] Lin Zhuo, Xiangfeng Guan and Songzhong Ye, "Quantitative evaluation and prediction analysis of the healthy and sustainable development of China's sports industry", *Sustainability*, Online ISSN: 2071-1050, p. 2184, Vol. 12, No. 6, 11 March 2020, Published by MDPI, DOI: 10.3390/su12062184, Available: https://www.mdpi.com/2071-1050/12/6/2184.

[7] Libby Moberg, Jordan Fritch, Danielle Westmark, Daniel Santa Mina, Crystal Krause *et al.*, "Effect of physical activity on fatigue in childhood cancer survivors: a systematic review", *Supportive Care in Cancer*, Print ISSN: 0941-4355, Online ISSN: 1433-7339, pp. 6441-6449, Vol. 30, No. 8, 19 March 2022, Published by Springer Nature, DOI: 10.1007/s00520-022-06960-4, Available: https://link.springer.com/article/10.1007/s00520-022-06960-4.

[8] Tejal Sarika Patel, Alex McGregor, Louise Fawcett, Sheree Bekker, Sean Williams *et al.*, "Coach awareness, knowledge and practice in relation to growth and maturation and training load in competitive, young gymnasts", *International Journal of Sports Science & Coaching*, Print ISSN: 1747-9541, Online ISSN: 2048-397X, pp. 528–543, Vol. 16, No. 3, 1 June 2021, Published by Sage Publications, DOI: 10.1177/1747954120978486, Available: https://journals.sagepub.com/doi/full/10.1177/1747954120978486.

[9] Ying Sun, Yaoqing Weng, Bowen Luo, Gongfa Li, Bo Tao *et al.*, "Gesture recognition algorithm based on multi-scale feature fusion in RGB-D images", *IET Image Processing*, Print ISSN: 1751-9659, Online ISSN: 1751-9667, pp. 1280-1290, Vol. 17, No. 4, April 2023, Published by Wiley Online Library, DOI: 10.1049/ipr2.12712, Available: https://ietresearch.onlinelibrary.wiley.com/doi/full/10.1049/ipr2.12712.

[10] Xinyu Li, Yuan He, Francesco Fioranelli, Xiaojun Jing, Alexander Yarovoy *et al.*, "Human motion recognition with limited radar micro-Doppler signatures", *IEEE Transactions on Geoscience and Remote Sensing*, Print ISSN: 0196-2892, Online ISSN: 1558-0644, pp. 6586–6599, Vol. 59, No. 8, 13 October 2020, Published by IEEE, DOI: 10.1109/TGRS.2020.3028223, Available: https://ieeexplore.ieee.org/abstract/document/9222330.

[11] Jianghui Geng, Qiang Wen, Qijin Chen and Hua Chang, "Six-degree-of-freedom broadband seismogeodesy by combining collocated high-rate GNSS, accelerometers, and gyroscopes", *Geophysical Research Letters*, Print ISSN: 0094-8276, Online ISSN: 1944-8007, pp. 708–716, Vol. 46, No. 2, 18 January 2019, Published by Wiley Online Library, DOI: 10.1029/2018GL081398, Available: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018GL081398.

[12] Xinyu Zhang and Jihua Chen, "A tennis training action analysis model based on graph convolutional neural network", *IEEE Access*, Online ISSN: 2169-3536, pp. 113264-113271, Vol. 11, No. 1, 13 October 2023, Published by IEEE, DOI: 10.1109/ACCESS.2023.3324425, Available: https://ieeexplore.ieee.org/abstract/document/10285097.

[13] Yang Liu, Peng Sun, Nicklolas Wergeles and Yi Shang "A survey and performance evaluation of deep learning methods for small object detection", *Expert Systems with Applications*, Print ISSN: 0957-4174, Online ISSN: 1873-6793, p. 114602, Vol. 172, No. 1, 15 June 2021, Published by Elsevier, DOI: 10.1016/j.eswa.2021.114602, Available: https://www.sciencedirect.com/science/article/abs/pii/S0957417421000439.

[14] Sheraz Aslam, Herodotos Herodotou, Syed Muhammad Mohsin, Nadeem Javaid, Nouman Ashraf *et al.*, "A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids", *Renewable and Sustainable Energy Reviews*, Print ISSN: 1364-0321, Online ISSN: 1879-0690, p. 110992, Vol. 144, No. 1, July 2021, Published by Elsevier, DOI: 10.1016/j.rser.2021.110992, Available: https://www.sciencedirect.com/science/article/abs/pii/S1364032121002847.

[15] Shuting Jin, Xiangxiang Zeng, Feng Xia, Wei Huang and Xiangrong Liu, "Application of deep learning methods in biological networks", *Briefings in bioinformatics*, Online ISSN: 1477-4054, pp. 1902–1917, Vol. 22, No. 2, 2 May 2021, Published by Oxford Academic, DOI: 10.1093/bib/bbaa043, Available: https://academic.oup.com/bib/article/22/2/1902/5826499.

[16] Wen-Bing CHEN, Yu-Lin LI and Yun-Jie CHEN, "An age and gender recognition model based on CNN-SE-ELM", *Computer Engineering & Science*, Print ISSN: 1007-130X, p. 872, Vol. 43, No. 05, 19 June 2021, Published by JOCES, DOI: 10.3969/j.issn.1007-130X.2021.05.014, Available: https://d.wanfangdata.com.cn/periodical/ChVQZXJpb2RpY2FsQ0hJMjAyNTA2MjISEWpzamdjeWt4MjAyMTA1MDE0GggyNTQyejRteg%253D%253D.

[17] Florian Heilmann and Kerstin Witte, "Perception and action under different stimulus presentations: A review of eye-tracking studies with an extended view on possibilities of virtual reality", *Applied Sciences*, Online ISSN: 2076-3417, p. 5546, Vol. 11, No. 12, 15 June 2021, Published by MDPI, DOI: 10.3390/app11125546, Available: https://www.mdpi.com/2076-3417/11/12/5546.

[18] Chunwei Tian, Yong Xu, Wangmeng Zuo, Bo Du, Chia-Wen Lin *et al.*, "Designing and training of a dual CNN for image denoising", *Knowledge-Based Systems*, Print ISSN: 0950-7051, Online ISSN: 1872-7409, p. 1069491, Vol. 226, No. 1, 17 August 2021, Published by IEEE, DOI: 10.1016/j.knosys.2021.106949, Available: https://www.sciencedirect.com/science/article/abs/pii/S0950705121002124.

[19] Xiao Cheng, Haoyu Wang, Jiayi Hua, Guoai Xu and Yulei Sui, "Deepwukong: Statically detecting software vulnerabilities using deep graph neural network", *ACM Transactions on Software Engineering and Methodology (TOSEM)*, Print ISSN: 1049-331X, Online ISSN: 1557-7392, pp. 1–33, Vol. 30, No. 3, 23 April 2021, Published by ACM Digital Library, DOI: 10.1145/3436877, Available: https://dl.acm.org/doi/abs/10.1145/3436877.

[20] Yaxu Xue, Yadong Yu, Kaiyang Yin, Pengfei Li, Shuangxi Xie *et al.*, "Human in-hand motion recognition based on multi-modal perception information fusion", *IEEE Sensors journal*, Print ISSN: 1530-437X, Online ISSN: 1558-1748, pp. 6793–6805, Vol. 22, No. 7, 4 February 2022, Published by IEEE, DOI: 10.1109/JSEN.2022.3148992, Available: https://ieeexplore.ieee.org/abstract/document/9703362.

[21] Xiaojun Zhang, "Application of human motion recognition utilizing deep learning and smart wearable device in sports", *International Journal of System Assurance Engineering and Management*, Print ISSN: 0975-6809, Online ISSN: 0976-4348, pp. 835–843, Vol. 12, No. 4, 27 April 2021, Published by Springer Nature, DOI: 10.1007/s13198-021-01118-7, Available: https://link.springer.com/article/10.1007/s13198-021-01118-7.

[22] Amir Nadeem, Ahmad Jalal and Kibum Kim,"Automatic human posture estimation for sport activity recognition with robust body parts detection and entropy markov model", *Multimedia Tools and Applications*, Online ISSN: 1573-7721, pp. 21465-21498, Vol. 80, No. 1, 16 March 2021, Published by Springer Nature, DOI: 10.1007/s11042-021-10687-5, Available: https://link.springer.com/article/10.1007/s11042-021-10687-5.

[23] Shubing Zhang, Jianhua Zhang and Xiaomin Zhou, "Design and Development of Smart Wearable Products for Basketball Dribble Teaching Training Posture Monitoring", *Wireless Communications and Mobile Computing*, Online ISSN: 1530-8677, p. 2222081, Vol. 2022, No. 1, 31 August 2022, Published by WILEY Online Library, DOI: 10.1155/2022/2222081, Available: https://onlinelibrary.wiley.com/doi/full/10.1155/2022/2222081.

[24] Faisal Hussain, Fawad Hussain, Muhammad Ehatisham-ul-Haq and Muhammad Awais Azam, "Activity-aware fall detection and recognition based on wearable sensors", *IEEE Sensors Journal*, Print ISSN: 1530-437X, Online ISSN: 1558-1748, pp. 4528–4536, Vol. 19, No. 12, 13 February 2019, Published by IEEE, DOI: 10.1109/JSEN.2019.2898891, Available: https://ieeexplore.ieee.org/abstract/document/8641383.

[25] Yi-Xing Liu, Ruoli Wang and Elena M. Gutierrez-Farewik, "A muscle synergy-inspired method of detecting human movement intentions based on wearable sensor fusion", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Print ISSN: 1534-4320, Online ISSN: 1558-0210, pp. 1089–1098, Vol. 29, No. 1, 7 June 2021, Published by IEEE, DOI: 10.1109/TNSRE.2021.3087135, Available: https://ieeexplore.ieee.org/abstract/document/9448148.

[26] Zhengyuan Song and Zhonghai Chen, "Sports action detection and counting algorithm based on pose estimation and its application in physical education teaching", *Informatica*, Print ISSN: 0350-5596, Online ISSN: 1854-3871, p. 1, Vol. 48, No. 10, 10 June 2024, Published by Slovenian Society Informatika, DOI: 10.31449/inf.v48i10.5918, Available: https://informatica.si/index.php/informatica/article/view/5918.

[27] Wenxin Du, "The computer vision simulation of athlete's wrong actions recognition model based on artificial intelligence", *IEEE Access*, Online ISSN: 2169-3536, pp. 6560-6568, Vol. 12, No. 1, 1 January 2024, Published by IEEE, DOI: 10.1109/ACCESS.2023.3349020, Available: https://ieeexplore.ieee.org/abstract/document/10379082.

[28] Zhijun Chen and Shuang Guo, "Combination of Lv-3DCNN algorithm in random noise environment and its application in aerobic gymnastics action recognition", *International Journal of Web Engineering and Technology*, Print ISSN: 1476-1289, Online ISSN: 1741-9212, pp. 213-231, Vol. 19, No. 3, Published by Inderscience, DOI: 10.1504/IJWET.2024.142225, Available: https://www.inderscienceonline.com/doi/abs/10.1504/IJWET.2024.142225.

[29] S. Dhivya Karunya and Krishna Kumar, "Development of detection and recognition of human activity in sports using GMM and CNN algorithms", *International Journal of System of Systems Engineering*, Print ISSN: 1748-0671, Online ISSN: 1748-068X, pp. 520-538, Vol. 14, No. 5, 2 September 2024, Published by Inderscience, DOI: 10.1504/IJSSE.2024.140756, Available: https://www.inderscienceonline.com/doi/abs/10.1504/IJSSE.2024.140756.

[30] Mir Mushhood Afsar, Shizza Saqib, Mohammad Aladfaj, Mohammed Hamed Alatiyyah, Khaled Alnowaiser *et al.*, "Body-worn sensors for recognizing physical sports activities in exergaming via deep learning model", *IEEE Access*, Online ISSN: 2169-3536, Vol. 11, No. 1, pp. 12460-12473, 25 January 2023, Published by IEEE, DOI: 10.1109/ACCESS.2023.3239692, Available: https://ieeexplore.ieee.org/abstract/document/10025735.

[31] Chuanbao He and Min Zhang, "Deep learning neural network-assisted badminton movement recognition and physical fitness training optimization", *Heliyon*, Online ISSN 2405-8440, p. 1, Vol. 10, No. 20, 30th October 2024, Published by Cell Symposia, DOI: 10.1016/j.heliyon.2024.e38865, Available: https://www.cell.com/heliyon/fulltext/S2405-8440(24)14896-7.