International Journal of Information Technology and Web Engineering
Volume 18 • Issue 1

Basketball Footwork and Application Supported by Deep Learning Unsupervised Transfer Method

Yu Feng, Dongguan City College, China
Hui Sun, Dongguan City College, China*

ABSTRACT

The combination of traditional basketball footwork mobile teaching and AI will become a hot spot in basketball footwork research. This article used a deep learning (DL) unsupervised transfer method: Convolutional neural networks are used to extract source and target domain samples for transfer learning. Feature extraction is performed on the data, and the impending action of a basketball player is predicted. Meanwhile, the unsupervised human action transfer method is studied to provide new ideas for basketball footwork action series data modeling. Finally, the theoretical framework of DL unsupervised transfer learning is reviewed. Its principle is explored and applied in the teaching of basketball footwork. The results show that convolutional neural networks can predict players' movement trajectories, unsupervised training using network data dramatically increases the variety of actions during training. The classification accuracy of the transfer learning method is high, and it can be used for the different basketball footwork in the corresponding stage of the court.

KEYWORDS

Deep Learning, Unsupervised Transfer Methods, Basketball Footwork, Convolutional Neural Networks, Human Action Transfer

INTRODUCTION

Basketball, as a collective, competitive, and entertaining sport, is loved by the majority of students, both boys and girls (Ma & Li, 2023). However, the problems in teaching and training restrict the improvement of teaching effect. Especially in the basketball game, the footwork stability and flexible footwork performance between players is very important for the defensive effect and confrontation ability. With the rapid development of artificial intelligence technology, transfer learning technology provides a new method to solve the practical problem of inconsistent data distribution. Unsupervised transfer learning technology has important application value, especially in the case of lack of labeled

DOI: 10.4018/IJITWE.334365
*Corresponding Author

data. In this context, this paper first proposes an unsupervised transfer method of deep learning to make up for the shortcomings of existing technologies. By using a CNN to extract features and predict basketball players' upcoming actions, this paper studies the unsupervised human movement migration method, which provides a new idea for basketball footwork movement sequence data modeling. At the same time, it also discusses the principle of DL unsupervised transfer learning theoretical framework and applies it to the teaching of basketball footwork. Through this research, we explore the application of unsupervised transfer learning technology in the research of basketball footwork, improve the efficiency and effect of teaching work, and provide a new idea and teaching method for Basketball Classroom and practical training.

LITERATURE REVIEW

Zhao and his team studied the routine teaching of footwork among students in the school's basketball elective course (Zhao et al., 2023). They found that in basketball teaching, whether it was the beginner or the proficiency period, the advance and defense movement exercises should be closely combined so the overall technique could be improved fast (Guo et al., 2023). Liu and Zhao pointed out a method to promote the transfer of skill motivation in their research on applying motor skills transfer theory in basketball teaching (Liu & Zhao, 2023). The teaching schedule should be rationally arranged. Physical fitness should be comprehensively developed. Student's ability to analyze and generalize needed to be improved (X. Liang, 2023). Teachers should fully use the transfer law between sports and basketball combination techniques. Individual techniques should be scientifically combined, and students could master these movements through repeated practice (Bu, 2023). When Gong and Srivastava studied the footwork training in basketball teaching, it was concluded that the three paramount footwork patterns are the forward, backward, and lateral sliding steps (Gong & Srivastava, 2022). To sum up, the studies surmised there are relevant theoretical foundations for the research on basketball footwork teaching. Scholars have discussed the composition of footwork, practice awareness, conventional teaching of footwork, practice principles, and issues that should be paid attention to in the routine teaching of basketball footwork using the method of comparison, literature, and logical reasoning.

In recent years, artificial intelligence technology has developed rapidly. It plays a decisive role in various practical application scenarios, and its core is algorithm learning and prediction. In practical applications, the data often do not entirely obey the same distribution for various reasons, such as different times and regions (Naik et al., 2023). This requires transfer learning techniques to apply the learned knowledge from the source domain to the target domain (Duan et al., 2023). A complex and valuable practical problem is that there is a difference in the distribution of data in the source and target domains, and there is no labeled data in the target domain (Zhou et al., 2023). This is the unsupervised transfer learning problem. This technology combines the two features of model discrimination ability and knowledge transfer ability and can be used in challenging unsupervised transfer learning scenarios. It has practical applications for classification learning and data labeling. Chol and his team study person recognition in video (Choi et al., 2023). They propose an unsupervised saliency learning method (Wang & Shi, 2023). This method deals with mismatches between pictures due to significant visual angle changes and person pose changes. Li and team et al. propose an unsupervised human action transfer method, which provides a new idea for modeling action sequence data (B. Li et al., 2023). They present a design for a new loss function based on invariance, thus endowing the network with the ability to unsupervised decoupling action feature representation. Lamsiyah et al. proposes an unsupervised transfer learning method and process based on graph convolutional networks (Lamsiyah et al., 2021). They put sample features and relational graphs into graph convolutional networks and applied unsupervised transfer learning methods based on graph convolutional networks to overcome the shortcomings of existing techniques (Figueiredo et al., 2023).

MATERIALS AND METHODS

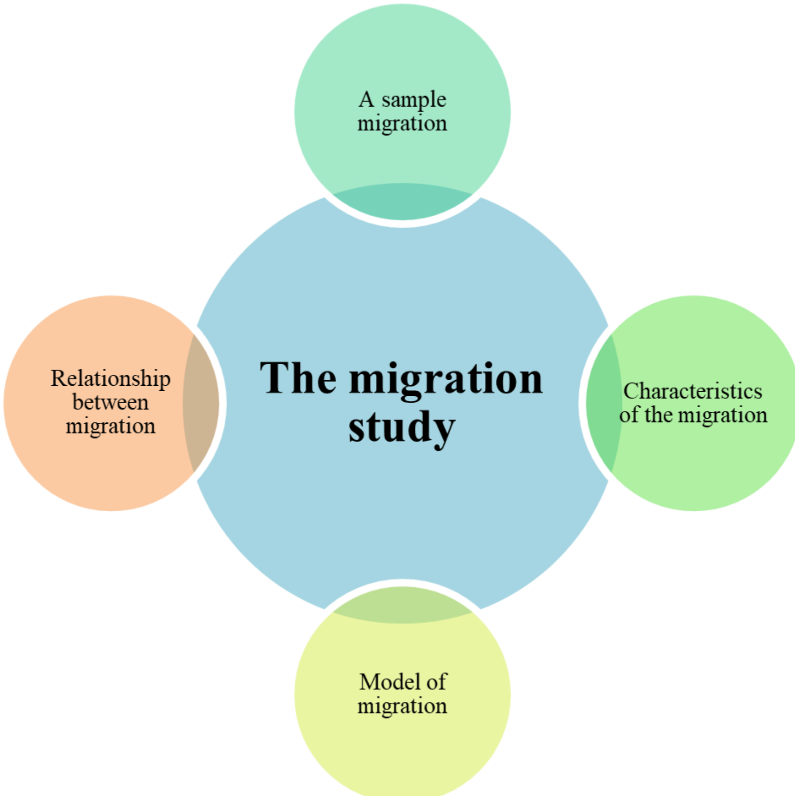
Action Feature Extraction Based on Transfer Learning

Feature extraction refers to the extraction of representative information from raw data that can describe the essential features of the data (Ren et al., 2023). In the field of machine learning and pattern recognition, feature extraction is a very important task because it can transform raw data into higher-level and more abstract representations, which are more suitable for processing and analyzing machine learning algorithms (Su et al., 2023). When extracting features, various techniques and methods are usually used to extract the most representative and useful features from the original data. In the field of action recognition and behavior analysis, feature extraction is a crucial step that can help us better understand the essence and characteristics of actions. In deep learning-based feature extraction, commonly used methods include CNNs\ and recurrent neural networks (RNNs), which can automatically learn and extract features from input data. For example, in video action recognition, 3D convolutional neural networks (3D-CNNs) can be used to extract spatial and temporal features from video sequences. In action recognition and behavior analysis, feature extraction can focus on extracting information that can describe action features from data such as motion trajectory, posture angle, and acceleration. These features can help us understand and distinguish different actions, which can be used to identify and analyze specific types of actions. For the study of basketball footwork, feature extraction involves information such as players' body posture, movement trajectory, stride frequency, stride length, etc. In addition, it is also possible to consider using data obtained from sensor devices, such as accelerometers and gyroscopes, to extract features related to basketball movements. Overall, feature extraction is the process of transforming raw data into higher-level, more abstract, and easier to understand representations. In action feature extraction, selecting appropriate feature extraction methods and representation forms has a significant impact on the final action recognition and analysis results.

Action feature extraction based on transfer learning refers to the process of extracting and analyzing action features by applying knowledge learned from one domain to another using transfer learning technology (H. Li et al., 2023). When conducting transfer learning, it is necessary to select a source domain that is highly correlated with the target domain and rich in data. At the same time, it is necessary to select a model or feature extractor that has already performed well in the source domain, such as the ImageNet pre-training model commonly used in the field of computer vision. In this process, models or feature extractors that have already been trained in the source domain are usually used to adapt to specific tasks in the target domain through fine-tuning or other methods (Shen et al., 2023). Transfer learning can play an important role in fields such as action recognition, behavior analysis, and motion pattern recognition. For example, by pre-training deep neural networks on large-scale source datasets, action feature extraction and classification can be faster in the target domain, thereby accelerating the training process of the model and improving its performance. The key to using transfer learning for action feature extraction lies in selecting appropriate source domain data and models, as well as designing effective transfer learning strategies to ensure full utilization of source domain knowledge in the target domain. The advantage of this method is that it can overcome the problem of insufficient data in the target domain, while improving the generalization ability and performance of the model. In the research of basketball footwork, action feature extraction based on transfer learning can help coaches and athletes better understand and analyze the action features in basketball matches, improve training effectiveness, and enhance game performance. There are generally four ways to transfer learning, as shown in Figure 1.

In image classification, the current application of transfer learning is very successful. There are many pre-trained image classification models on the ImageNet dataset, and these pre-trained models can be transferred to the target task. UCF-101 is a video dataset widely used for action recognition tasks, containing 101 different action categories (H. Liang, 2023). It is published by Khurram Soomro, Amir Roshan Zamir, and Mubarak Shah from the Viterbi Department of Computer Science at the

Figure 1. The way of transfer learning



University of Southern California (Wang, 2023). This dataset is sourced from YouTube, with video clips ranging in length from 4 to 13 frames. The UCF-101 dataset contains approximately 13,320 short video clips, totaling approximately 13,320,000 frames of images. Each category contains an average of approximately 130 video clips. The videos in the UCF-101 dataset have undergone preprocessing, such as video capture and annotation of action categories, video editing and segmentation, and video format conversion to specific encoding and decoding formats. The UCF-101 dataset has become one of the standard benchmark datasets for evaluating the performance of action recognition algorithms and is of great significance for conducting research and algorithm validation related to action recognition. Image classification models pre-trained on ImageNet generally achieve high accuracy on other image datasets. Here, an Inflated 3D Convolutional Network (I3D) is used as the basic feature extraction network. First, the I3D network is pre-trained on the large-scale video dataset of Kinetics. The Kinetics400 dataset corresponds to 400 action classes, so the I3D network must be slightly changed when using the pre-trained model for the action recognition dataset (UCF-101). In addition, the data volume of the Kinetics dataset is much larger than that of the UCF-101 dataset, so transfer learning will increase the model's generalization. The I3D network, after transfer learning, can quickly converge, greatly reducing calculation and training time. The transfer learning effect of the features generated by the DNN model in different layers is different. The high-level abstract features generated by the later network layers are suitable for transfer learning. Therefore, the features generated by the last layer of the Inception module are selected to be input into the subsequent time series modeling network.

Unsupervised Human Action Transfer Methods

Nowadays researchers have proposed an unsupervised human action transfer method, which provides a new idea for modeling action sequence data (Liu, 2023). This novel action redirection network design can be trained end-to-end from unlabeled network data in a 2D key point space. The researchers design a new loss function based on invariance to endow the network with the ability to decouple action feature representations unsupervised. Applying the above action redirection network and invariance-based loss function to the human action transfer task, outperforms the original state-of-the-art methods in both qualitative and quantitative metrics, especially on complex real-world actions. Recently, the cost of obtaining human motion information has been dramatically reduced with the popularity of mobile computing and the application of DL in computer vision. The action transfer process is divided into three stages to deal with the large difference in structure and perspective between the basketball player's footwork movement video and the target movement video, as shown in Figure 2.

The original and target motion videos have large structural and perspective differences, so it is difficult to establish the source-target mapping at the pixel level. Especially when the initial object performs complex actions or when the structure difference between the initial object and the target object is relatively large, the accuracy of the traditional action transfer method is low. The action transfer process is divided into three stages: human key point detection, action redirection, and video rendering. It is only necessary to focus on the problem of action redirection by decomposing the tasks. The input and output of this problem are both 2D human key point sequences. Figure 3 displays the overall framework of action transfer.

Finding paired action and character data in the real world is generally tricky as effective supervision signals for human action transfer tasks. Human motion exhibits complex nonlinearity. It is difficult to establish accurate models and parameters to characterize the process of human action transfer. The invariance of features in three dimensions in human motion data is exploited to deal with these difficulties. The first is motion, which refers to the semantic information of the movement of

Figure 2. Different stages of action transfer

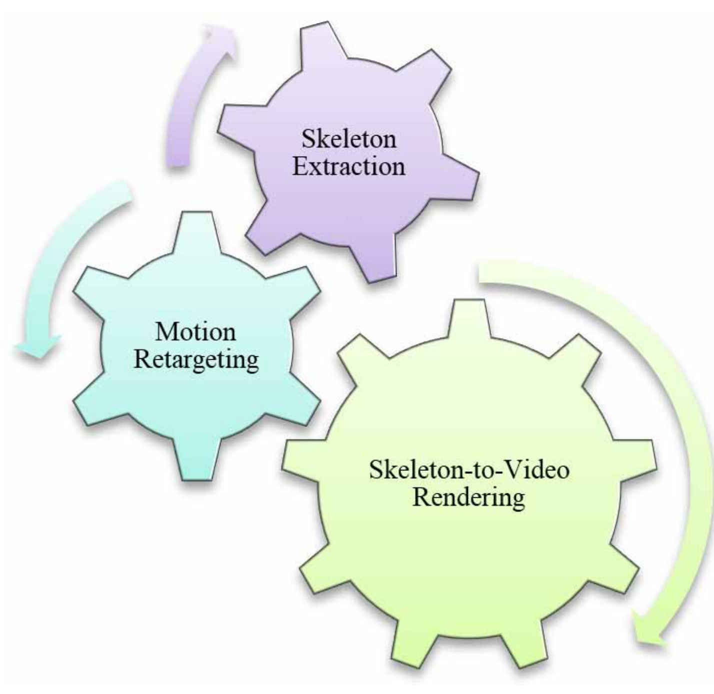
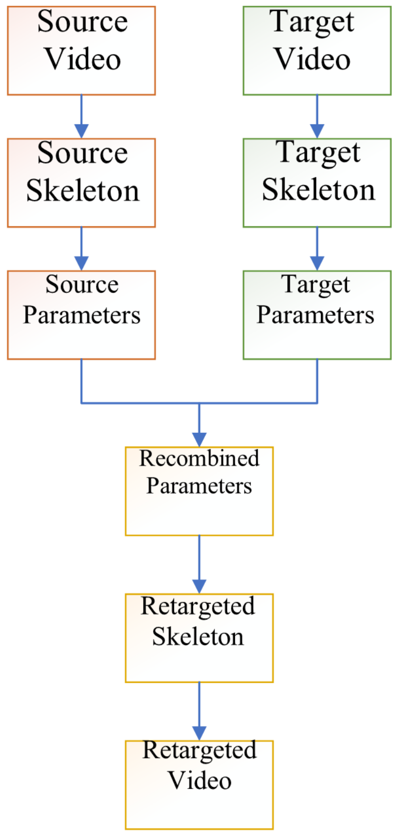


Figure 3. Overall framework of action transfer



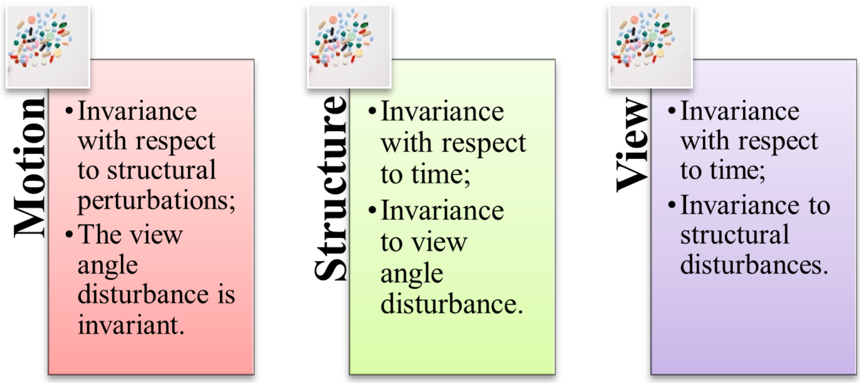
various body parts. The second is structure, which refers to the proportions of the body. The third is the view, which refers to the relative orientation information of the body and the camera. In theory, the overall motion can be reconstructed from these three pieces of information. These three parts of information are independent of each other, and any information is invariant to the disturbance of the additional two. Specifically, they have the properties shown in Figure 4.

In the training implementation, the rotation of the 3D human body is used as the perturbation of the perspective information. Limb scaling is the perturbation of structural information. Motion information does not need to be perturbed explicitly, as it changes over time. Based on these perturbations, the features that are required to be re-encoded by the network have the invariances mentioned above. Then, a series of completely unsupervised loss functions can be derived. The human key point sequence information is decoupled into three mutually orthogonal components of motion, structure, and perspective by training an auto-encoder.

Multi-Source Isomorphic Transfer Method Based on Graph Convolution

A graph convolutional network (GCN) is a deep learning model designed for graph structured data. Its main advantage is that it can effectively deal with the relationship and connection between nodes (Bu, 2023). GCNs can fully consider the topology of graph data and use the connection information between nodes for information transmission and feature extraction, which is different from the traditional data based on grid structure (such as images) (Li et al., 2023). A GCN is more suitable for solving the specific problems of graph data (Baye & Yusuf, 2023). In the multi-source

Figure 4. Invariance of features in three dimensions



isomorphic transfer learning, the source domain and the target domain may come from different data distribution, while a GCN has certain advantages in dealing with heterogeneous data and can effectively learn the differences and similarities between different data domains, so as to adapt to the situation of multi-source data. The GCN framework is relatively flexible. It can adapt to different graph structure data by adjusting the network structure and parameters. At the same time, it can also be combined with other traditional graph analysis methods. The multi-source isomorphism method based on graph convolution aims to solve the transfer problem of graph-structured data in multiple source domains and the situation of unlabeled data in the target domain, which belongs to the unsupervised transfer learning problem. The ultimate goal is to mine the spatial features in the graph structure by reducing the distribution difference between the source domain and the target domain to solve the adaptation problem of the source domain and the target domain. Finally, the source domain label data are utilized to classify the target domain. Figure 5 shows the overall framework of the transfer model.

The idea of the graph CNN model is derived from the spectral decomposition of the graph Laplacian matrix, which is the feature decomposition. Laplacian matrices are constructed from graph structures and are often used in graph theory. This model can be regarded as a linear transformation, which acts the same as the Laplacian operator in mathematical analysis. Laplacian matrices can be called Laplacian operators or discrete Laplacian operators. The Laplacian matrix is described below by taking Figure 6 as an example.

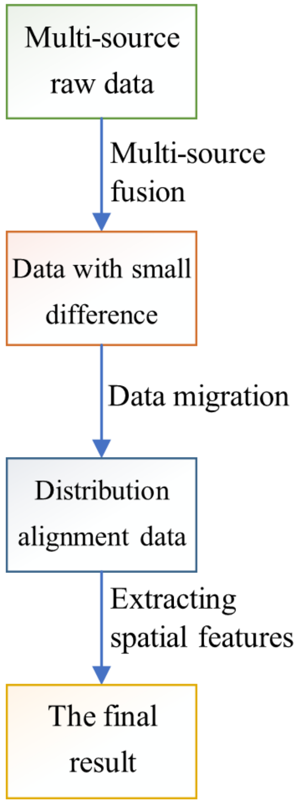
It is assumed that graph G has N nodes, and the defined function is an N-dimensional vector (f1,…, fi,…, fn). The symbol fi represents the function value at node i in the graph. Assuming that a perturbation is added to the i-node, it may become any adjacent node j, j∈Ni. Ni represents the set of adjacent nodes of the i-node. Then, the gain brought by the change of any node j to node i is expressed as:

$$\Delta f_i = \sum_{j \in N_i} f_i - f_j \tag{1}$$

Suppose the weight of each edge is wij. Besides, when wij=0, node i and node j have no edge. After substituting the weight of the edge, (1) can be transformed into:

$$\Delta f_i = \sum_{j \in N} w_{ij} \left(f_i - f_j \right) \tag{2}$$

Figure 5. Overall framework of the transfer model



After expanding (2), (3) can be deduced:

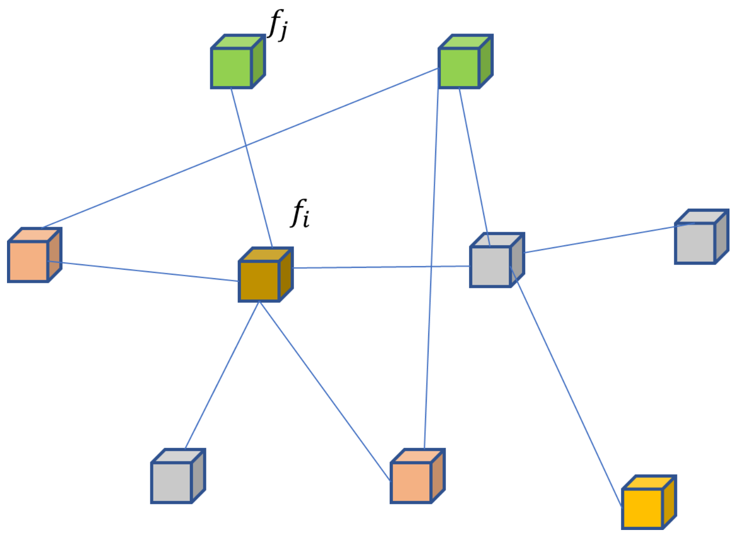
$$\Delta f_i = \sum_{j \in N} w_{ij} f_i - \sum_{j \in N} w_{ij} f_j = d_i f_i - w_i : f \tag{3}$$

As seen in (3), di represents the degree of vertex I and generalizes to all nodes to get the change gain. It is expressed as:

$$\Delta f = \begin{pmatrix} \Delta f_i \\ \dots \\ \Delta f_n \end{pmatrix} = \begin{pmatrix} d_1 f_1 - wi : f \\ \dots \\ d_n f_n - wn : f \end{pmatrix} = Df - Wf = \left(D - W\right) f \tag{4}$$

As seen in (4), D-W is the Laplacian matrix, denoted as L. Laplacian matrices are often used in spectral clustering algorithms. First, the similarity matrix W between data points is defined using the k-nearest neighbor algorithm according to the distance between them. Then, the Laplacian matrix L is obtained according to the similarity matrix, and spectral decomposition is performed. Finally, spectral clustering is performed on the original data points using spectral decomposition to obtain eigenvectors. D is a diagonal matrix and is symmetric. W represents the graph adjacency matrix, which is also symmetric. Therefore, the Laplacian matrix (D-W) is a positive semi-definite

Figure 6. Schematic diagram of laplacian structure



symmetric matrix. It can perform feature decomposition, which is called spectral decomposition, which is expressed as follows:

$$L = U\begin{pmatrix} \lambda_1 & & \\ & \ldots & \\ & & \lambda_n \end{pmatrix} U^T \tag{5}$$

As seen in (5), U indicates that the column vector is a matrix composed of eigenvectors, and λ_i indicates the corresponding eigenvalue. This model corresponds to the traditional Fourier transform. The eigenvalue corresponds to the frequency, the eigenvector corresponds to the basic function, and the Fourier transform is obtained. Therefore, the Fourier transform of f under the eigenvalue λ_i is the multiplication of f and the eigenvector Ui corresponding to the eigenvalue λ_i. So, the matrix form of f in the Fourier transform is obtained as:

$$\begin{pmatrix} \hat{f}\left(\lambda_1\right) \\ \ldots \\ \hat{f}\left(\lambda_n\right) \end{pmatrix} = U^T f \tag{6}$$

Similarly, the inverse Fourier transform becomes the summation of the corresponding eigenvalues λ_i, which is expressed as:

$$\begin{pmatrix} f\left(\lambda_1\right) \\ \ldots \\ f\left(\lambda_n\right) \end{pmatrix} = U\hat{f} \tag{7}$$

The ultimate goal of the graph CNN model is to introduce variable parameters. Therefore, a convolution kernel h is defined. According to the convolution theorem, the Fourier transform of the function convolution is equal to the product of the Fourier transform. Multiply the graph Fourier transform of the convolution of h and f by U to obtain the convolution of h and f in the original domain, which is expressed as:

$$\left(f{*}h\right)_G = U\begin{pmatrix} \hat{h}\left(\lambda_1\right) & & \\ & \cdots & \\ & & \hat{h}\left(\lambda_n\right) \end{pmatrix} U^T f \tag{8}$$

In this way, a multi-source isomorphic transfer algorithm based on graph convolution is obtained. The angle of spectral decomposition is to use the theory of spectral decomposition to perform convolution operations. The spatial angle is a convolution operation based on the node's neighbors. The above sequences represent graph convolution from the perspective of spectral decomposition.

RESULTS AND ANALYSIS

Analysis and Comparison of Exit Training Strategies

During the experiment, practical training techniques, such as pre-training and random Dropout, are tried to improve the model's performance. Pre-training on the Kinetics dataset when the transfer learning method will bring a significant improvement in the accuracy of the behavior recognition model. The deep model may have overfitting problems, so the Dropout training strategy is used. The performance of the I3D-Long Short-Term Memory (LSTM) model and the I3D-Gated Recurrent Unit (GRU) model under different training strategies on the UCF-101 dataset is shown in Figure 7.

Figure 7. Performance of I3D models under different training strategies

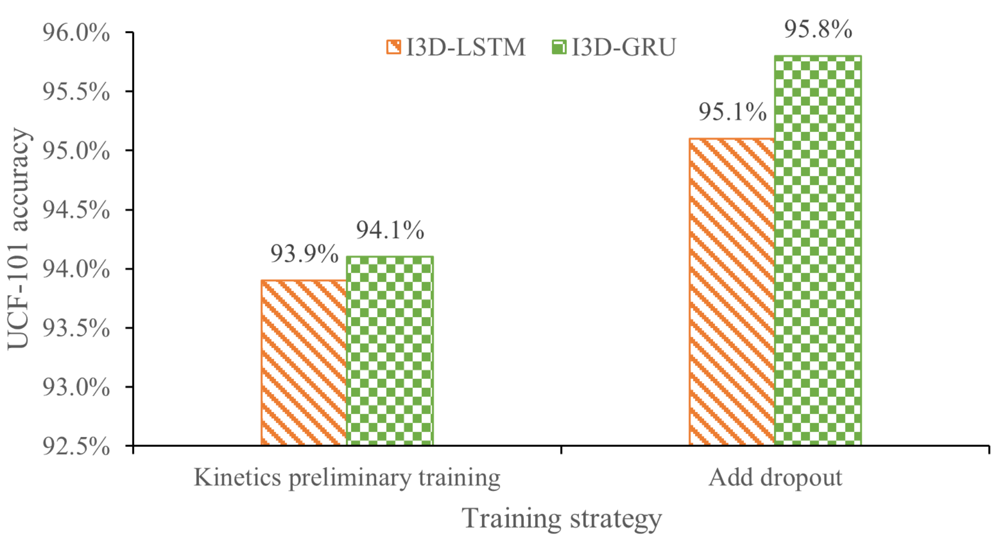


Figure 7 indicates that the I3D-LSTM and the I3D-GRU networks proposed here have high recognition accuracy on the UCF-101 dataset. Different training strategies are also essential. The accuracy of the I3D-LSTM network with Dropout is about 1.2% higher than that without Dropout. Similarly, the accuracy of the I3D-GRU network with Dropout is about 1.7% higher than that without Dropout. The recognition accuracy of the I3D-GRU network is higher than that of the I3D-LSTM network. It is speculated that this may be because GRU has fewer parameters than the network and can avoid overfitting well.

The UCF-101 data set is compared with other mainstream algorithm models after referring to a large number of literatures to prove that the I3D-LSTM and I3D-GRU networks proposed here are more advanced than the current mainstream network models for behavior recognition. The details are shown in Table 1.

From Table 1, the I3D-LSTM and I3D-GRU networks proposed here outperform the current mainstream algorithms on the UCF-101 dataset. This fully demonstrates that pre-training on the large-scale video dataset of Kinetics can effectively increase the performance of action recognition networks. The I3D-LSTM network is about 7.1% higher than the two-stream network, which shows that the LSTM network is better at modeling time series features than the RNN. In summary, the I3D-LSTM and I3D-GRU networks proposed here combine the advantages of the current mainstream network models and are efficient behavioral recognition models. Table 2 reveals the recognition results for the single-modal data.

From Table 2, action recognition based on skeleton node data achieves the highest recognition accuracy among the three data modalities. This shows that among all data modalities, the feature of the skeleton node can best reflect human behavior. Moreover, it has strong anti-interference for occlusion, noise, and other factors. The Kinetics dataset is similar to the UCF-101 dataset. Therefore, the recognition effect based on Red-Green-Blue (RGB) image sequence is better than that based on the depth image sequence.

Analysis of the Training Effect of Turn Back Running

SPSS(Statistical Package for the Social Sciences) software is used to test and analyze the effect of school basketball players' 4*10m reentry running. The experimental group and the control group are arranged. Through the independent sample T-test, it is concluded that the P value of the experimental and control groups' ages is 0.77 (P>0.05). The P value of height is 0.72 (P>0.05), the P value of

Table 1. Comparison with mainstream algorithm models

Algorithm model	Accuracy
Two-Stream	87.96%
I3D	94.28%
I3D-LSTM	95.08%
I3D-GRU	95.77%

Table 2. Recognition performance based on unimodal data

Network model	Data modal	Accuracy
I3D-GRU	Depth image sequence	88.37%
	RGB image sequence	89.14%
AlexNet	Skeleton node data	90.27%

weight is 0.34 (P>0.05), and the P value of training years is 0.55 (P>0.05). As a result, there is no significant difference between the players in the experimental group and the control group. The two groups have good homogeneity, which can meet the experimental requirements and carry out this experiment.

The comparison results of the 4*10m reentry test before and after the experiment between the experimental group and the control group are revealed in Tables 3 to 6.

Before the experiment, SPSS software is used to input the results measured by the experimental and control groups. Then, the independent sample T-test is conducted. Table 3 reveals the final test results. From Table 3, the average score of the experimental group before the experiment is 13.9s, and its standard deviation is 1.0s. The average score of the control group is 14.3s, with a standard deviation of 1.1s. The final measured P value is 0.36, and P>0.05. Therefore, there is no significant difference between the two groups of basketball players in the experimental group and the control group in this test.

After 12 weeks of training with a total of 36 training sessions, all subjects in the experimental group are tested on the 4*10m reentry running index. According to the test results before the experiment and the test results after the experiment, a comparative analysis is carried out. Paired samples T-test is performed on the score data of the experimental group before and after the experiment. From Table 4, the average test score of the experimental group before the experiment is 13.9s, and the standard deviation is 1.0s. The average test score after the experiment is 13.3s, and its standard deviation is 0.8s. The average score after the experiment is 0.6s higher than the average score before the experiment. The measured P value is 0.01, and P<0.05. It is proved that the difference is significant.

After three months of regular foot movement training, the test scores of the control group before and after the experiment are substituted into SPSS software for the paired-sample T-test. The obtained results are shown in Table 5. From Table 5, the average score of the control group before the experiment is 14.3s, with a standard deviation of 1.1s. After 12 weeks of training, the average score improves to 14.2s with a standard deviation of 1.1s. It is 0.1s higher than the pre-experiment average. The measured P value is 0.18, and P>0.05. The difference is not significant.

Table 3. Comparison of test scores of 4*10m reentry run before the two groups of experiments

Test indicators	4*10m reentry run(s)
Experimental group $\bar{X} \pm S$	13.9 ± 1.0
Control group $\bar{X} \pm S$	14.3 ± 1.1
P value	0.36

Table 4. Comparison of the results of the experimental group before and after the experiment

Test indicators	4*10m reentry run(s)
$\bar{X} \pm S$ before experiment	13.9 ± 1.0
$\bar{X} \pm S$ after experiment	13.3 ± 0.8
P value	0.01

Table 5. Comparison of the Performance of the Control Group Before and After the Control Group Experiment

Test indicators	4*10m reentry run(s)
$\bar{X} \pm S$ before experiment	14.3 $\pm$ 1.1
$\bar{X} \pm S$ after experiment	14.2 $\pm$ 1.1
P value	0.18

Table 6. Comparison of 4*10m reentry test scores after the two groups of experiments

Test indicators	4*10m reentry run(s)
Experimental group $\bar{X} \pm S$	13.3 $\pm$ 0.8
Control group $\bar{X} \pm S$	14.2 $\pm$ 1.1
P value	0.04

After the experiment, the experimental and control groups are tested for a 4*10m reentry run. Then, the measured data is substituted into SPSS software, and the independent sample T-test after the experiment is carried out. The obtained data are presented in Table 6. After the experiment, the average performance of the experimental group in the 4*10m reentry run is 13.3s, and its standard deviation is 0.8s. The average performance of the 4*10m reentries run in the control group after the experiment is 14.2s, with a standard deviation of 1.1s. The measured P value is 0.04, and $P<0.05$. The difference is significant. It implies that after 12 weeks of training, the performance measured in the experimental group is significantly better than in the control group.

Analysis of Application Challenges

When conducting basketball footwork research, data collection is an important challenge. Firstly, it is necessary to consider how to obtain representative basketball footwork data, including the movements of athletes of different levels and styles. This may require collaboration with basketball clubs or schools to obtain video data from actual matches and training. In addition, to ensure data quality, it may be necessary to filter and clean the data to remove noise or inaccurate action data.

In terms of data annotation, it is necessary to carefully consider how to obtain accurate labeled data. Due to the potential involvement of complex movement patterns and sequences in basketball footwork, annotators need to have sufficient knowledge and experience in basketball. It may be necessary to recruit professional basketball coaches or athletes to participate in data annotation work to ensure the accuracy and professionalism of labeling.

For the proposed unsupervised transfer learning method, the research team needs to fully consider its applicability and stability on real basketball footwork data. A large amount of experimental verification and parameter tuning work may be required to ensure that the method can effectively migrate from the source domain to the target domain and maintain good generalization ability.

In terms of verifying the actual effectiveness of teaching applications, the research team can consider working closely with basketball coaches and students to design practical teaching scenarios and courses. By collaborating with actual educators and learners, feedback can be obtained in a timely manner and improvements can be made to ensure that the proposed method can truly improve the efficiency and effectiveness of basketball footwork teaching.

Overall, basketball footwork research may face multiple challenges and limitations in terms of data acquisition, annotation quality, validation of transfer learning methods, and validation of actual teaching effectiveness. In response to these challenges, the research team needs to adopt a systematic strategy and collaborate with professionals in relevant fields to ensure the smooth progress of the research and the effectiveness of the results.

Analysis of Practical Applications

Combining basketball training with video recognition technology can achieve the results of improving training effectiveness and individual skill levels.

- **Action analysis and feedback:** Use video action recognition technology to analyze players' actions, including shooting, dribbling, defense, etc. By monitoring and analyzing players' movements in real-time, real-time feedback and guidance can be provided to coaches and players, helping them improve their movement skills.
- **Personalized training plan:** Based on the results of video motion recognition technology, develop a personalized training plan tailored to each player's specific technical problems and improvement space. By quantitatively analyzing the actions of each player, targeted training plans can be tailored to improve training efficiency and results.
- **Game video analysis:** Using video action recognition technology to analyze game videos, identify the technical performance and shortcomings of teams and individuals in the game. This helps coaches and players to gain a deeper understanding of their performance in actual combat and identify areas for improvement.
- **Action demonstration and learning:** Utilize video action recognition technology to record and analyze the action performance of professional players, provide excellent action samples for players, and combine virtual reality and other technologies to help players imitate and learn excellent action skills.

Through the above methods, video motion recognition technology can play an important role in basketball training projects, helping players and coaches to train and improve their skills more scientifically, improving their technical level and overall competitive performance. In addition to basketball training, video motion recognition technology can also be applied to other sports such as yoga, gymnastics, running, etc. By real-time monitoring and analysis of movements, it helps athletes improve their posture and improve their athletic skills. In rehabilitation therapy, video motion recognition technology can be used to monitor the patient's motor recovery process, help rehabilitation doctors and physical therapists evaluate the patient's motor performance, and design more effective rehabilitation plans. In intelligent driving systems, video action recognition technology can be used to monitor the driver's status and behavior, assist the driver in driving decisions, and improve traffic safety. Overall, video motion recognition technology has broad application prospects in various fields such as sports training, medical rehabilitation, virtual reality, and can bring many conveniences and innovations to people's lives, work, and entertainment.

CONCLUSION

Basketball, as a collective, competitive, and recreational sport, is loved by most students (Ma & Li, 2023). In order to improve the effect of basketball teaching, we successfully explore the application of unsupervised transfer learning technology in the study of basketball footwork. We use CNN to extract features and predict basketball player's movements, and study the unsupervised human movement migration method, which provides a new idea for basketball footwork movement sequence data modeling. At the same time, we also discuss the principle of DL unsupervised transfer learning theory framework and successfully apply it to the teaching of basketball footwork. In the research, we carry out a large number of experiments on the unsupervised transfer learning method and achieve good performance on the real basketball footwork data set. We migrate the model under different data distribution and evaluate the effectiveness and stability of the method through cross validation and other methods. The results show that our method has important application value in solving basketball footwork data with inconsistent data distribution. The successful application of this study provides new ideas and methods for basketball footwork teaching and provides a feasible scheme for solving the problems existing in teaching and training. The successful application of unsupervised transfer learning technology not only improves the efficiency and effect of basketball footwork teaching, but it also provides a useful reference for using artificial intelligence technology to improve other physical education teaching.

Although the network proposed here performs well in image sequence recognition on RGB video datasets, there is still much room for improvement. Prospects and work include:

1. The model should be further optimized to improve the model's accuracy on the dataset. For example, introducing deeper and more complex neural network structures or adopting attention mechanisms to improve the accuracy of the model on the dataset. In addition to training and testing on the current dataset, the designed network model should also be used on other different datasets to verify its generalization ability and further improve its performance.
2. The datasets used here are all preprocessed, which differs from the real footstep transfer behavior recognition task. In practical applications, many videos containing complex backgrounds, interference, and low resolution need to be processed in real-time, so there is still a long way to go to truly efficiently and accurately identify human transfer without supervision. Therefore, we can actively collect real scene data containing complex backgrounds, interference, and low-resolution videos, and perform corresponding annotation work to construct more representative datasets.

Summatively, we must optimize the model for complex scenarios in practical applications to adapt to complex backgrounds, low resolution, and other conditions and improve the real-time processing ability of the model. In addition, research can be conducted on noise and interference processing methods for video data in complex scenes, such as background removal, image enhancement, etc., to improve the recognition performance of the model in real scenes.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

FUNDING STATEMENT

This work was not supported by any funds.

REFERENCES

Baye, Z., & Yusuf, K. (2023). The Role of Virtual Reality and Augmented Reality in Sports Psychology: Advancements, Applications, and Implications. [Journal of Sport Psychology]. *Revista de Psicología del Deporte*, *32*(3), 110–118. https://www.rpd-online.com/index.php/rpd/article/view/1388

Bu, X. (2023). Exploration of intelligent coaching systems: The application of Artificial intelligence in basketball training. *Saudi Journal of Humanities and Social Sciences*, *8*(09), 290–295. doi:10.36348/sjhss.2023.v08i09.007

Choi, T., Cho, K., & Sung, Y. (2023). Approaches that use Domain-Specific expertise: Behavioral-Cloning-Based advantage Actor-Critic in basketball games. [x]. *Mathematics*, *11*(5), 1110. doi:10.3390/math11051110

Duan, C., Hu, B., Liu, W., & Song, J. (2023). Motion capture for sporting events based on graph convolutional neural networks and single target pose estimation algorithms. *Applied Sciences (Basel, Switzerland)*, *13*(13), 7611. doi:10.3390/app13137611

Figueiredo, E., Yano, M. O., Da Silva, S., Moldovan, I. D., & Bud, M. A. (2023). Transfer learning to enhance the damage detection performance in bridges when using numerical models. *Journal of Bridge Engineering*, *28*(1), 04022134. doi:10.1061/(ASCE)BE.1943-5592.0001979

Gong, Y., & Srivastava, G. (2022). Multi-target trajectory tracking in multi-frame video images of basketball sports based on deep learning. *EAI Endorsed Transactions on Scalable Information Systems*, *10*(2), e9–e9. doi:10.4108/eetsis.v9i6.2591

Guo, X., Brown, E., Chan, P. P., Chan, R. H., & Cheung, R. T. (2023). Skill level classification in basketball Free-Throws using a single inertial sensor. *Applied Sciences (Basel, Switzerland)*, *13*(9), 5401. doi:10.3390/app13095401

Lamsiyah, S., Mahdaouy, A. E., Ouatik, S. E. A., & Espinasse, B. (2021). Unsupervised extractive multi-document summarization method based on transfer learning from BERT multi-task fine-tuning. *Journal of Information Science*, *49*(1), 164–182. doi:10.1177/0165551521990616

Lei, M., Fronteau, A., & Huang, J. (2023). Evaluation of the effect of virtual simulation teaching on learning behavior of college students. [IJET]. *International Journal of Emerging Technologies in Learning*, *18*(10), 204–218. doi:10.3991/ijet.v18i10.38707

Li, B., Wang, L., Jiang, Q., Li, W., & Huang, R. (2023). Sports risk prediction model based on automatic Encoder and Convolutional Neural Network. *Applied Sciences (Basel, Switzerland)*, *13*(13), 7839. doi:10.3390/app13137839

Li, H., Wang, J., & Wang, Y. (2023). Holistic transfer educational learning approach for higher education. *Computer Applications in Engineering Education*, *31*(3), 710–727. doi:10.1002/cae.22600

Liang, H. (2023). Improved EfficientDET algorithm for basketball players' upper limb movement trajectory recognition. *Applied Artificial Intelligence*, *37*(1), 2225906. doi:10.1080/08839514.2023.2225906

Liang, X. (2023). A video images-aware knowledge extraction method for intelligent healthcare management of basketball players. *Mathematical Biosciences and Engineering*, *20*(2), 1919–1937. doi:10.3934/mbe.2023088 PMID:36899515

Liu, Z., & Zhao, J. (2023). Data analysis of the development status of Basketball National Fitness based on FOG Computing. *Applied Artificial Intelligence*, *37*(1), 2221102. Advance online publication. doi:10.1080/08839514.2023.2221102

Ma, J., & Li, W. (2023). Efficient Image Segmentation of Cardiac Conditions after Basketball Using a Deep Neural Network. *Electronics (Basel)*, *12*(2), 466. doi:10.3390/electronics12020466

Naik, B. T., & Hashmi, M. F. (2023). LSTM-BEND: Predicting the Trajectories of Basketball. *IEEE Sensors Letters*, *7*(4), 1–4. doi:10.1109/LSENS.2023.3253863

Ren, F., Yang, C., & Nanehkaran, Y. A. (2023). MRI-based model for MCI conversion using deep zero-shot transfer learning. *The Journal of Supercomputing*, *79*(2), 1182–1200. doi:10.1007/s11227-022-04668-0

Shen, Q., Teso, S., Giunchiglia, F., & Xu, H. (2023). To Transfer or Not to Transfer and Why? Meta-Transfer learning for explainable and Controllable Cross-Individual Activity Recognition. *Electronics (Basel)*, *12*(10), 2275. doi:10.3390/electronics12102275

Su, A., Zhang, X., Zhang, C., Ding, D., Yang, Y. F., Wang, K., & She, Y. B. (2023). Deep transfer learning for predicting frontier orbital energies of organic materials using small data and its application to porphyrin photocatalysts. *Physical Chemistry Chemical Physics*, *25*(15), 10536–10549. doi:10.1039/D3CP00917C PMID:36987933

Wang, L., & Liu, F. (2023). Teaching knowledge sharing in virtual practice teaching. *International Journal of Emerging Technologies in Learning*, *18*(4), 170–185. doi:10.3991/ijet.v18i04.38239

Wang, T., & Shi, C. (2023). Basketball motion video target tracking algorithm based on improved gray neural network. *Neural Computing & Applications*, *35*(6), 4267–4282. doi:10.1007/s00521-022-07026-6

Wang, Y. (2023). Kinect Body Sensor Technology-Based Quantitative Assessment Method for Basketball Teaching. [IJDST]. *International Journal of Distributed Systems and Technologies*, *14*(2), 1–13. doi:10.4018/IJDST.317935

Zhao, K., Du, C., & Tan, G. (2023). Enhancing Basketball Game Outcome Prediction through Fused Graph Convolutional Networks and Random Forest Algorithm. *Entropy (Basel, Switzerland)*, *25*(5), 765. doi:10.3390/e25050765 PMID:37238520

Zhou, L., Zhang, C., & Wang, M. (2023). Emotion recognition algorithm of basketball players based on deep learning. *International Journal of Information and Communication Technology*, *22*(4), 377–390. doi:10.1504/IJICT.2023.131223