Received October 11, 2019, accepted October 19, 2019, date of publication October 25, 2019, date of current version November 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2949643

Implementation and Rehabilitation Application of Sports Medical Deep Learning Model Driven by Big Data

YANFENG SU
Institute of Physical Education, Xinxiang Medical University, Xinxiang 453003, China
e-mail : 081029@xxmu.edu.cn

ABSTRACT A large number of unlabeled and limited style data greatly reduces the reuse possibility of existing motion sequences. Effective classification and fragment splicing have become an important way of data reuse. Aiming at these two problems, this paper focuses on the great success of deep learning in the field of graphics and iconography. Based on the theory of Restricted Boltzmann Machine (RBM), a spatio-temporal feature extraction model for human skeleton medical motion sequences is established. The research results are mainly manifested in three aspects. (1) In this paper, stack factor decomposition spatiotemporal feature model and discriminate RBM are used to construct semi-supervised combination model. (2) The underlying model firstly uses the idea of weight decomposition to construct the three channel generative RBM model; and then it extracts the abstract temporal and spatial characteristics of the original motion sequence. Furthermore, it identifies the behavior style of the current input segment at the top using the discriminate RBM model. Finally conduct the stylistic statistics of the whole motion sequence in the voting space. (3) An unsupervised similar frame detection model is constructed by using 3D convolution RBM's perception of human adjacent joints' linkage. In this way, Candidate frames for constructing graph model nodes are obtained. Trajectory and style switching control is realized based on attitude similarity screening criteria. The simulation experiment verifies the superiority and reliability of the algorithm, and it is effectively applied in rehabilitation training.

INDEX TERMS Rehabilitation application, sports medical, deep learning, big data.

I. INTRODUCTION

As an important branch of multimedia technology and computer graphics, motion capture technology has been widely used in animation, games, medical treatment, virtual reality and other practical scenes [1], [2]. However, in the actual motion capture process, such as low precision of capture equipment, severe scene limitation and unavoidable environmental noise interference, which greatly reduces the availability of capture data [3]. With the continuous improvement of machine learning technology, extracting structural features from existing lossless data and reconstructing missing joint data have become one of the key technologies of data-driven character animation [4]. As a typical application of motion capture technology, the complex nonlinear structure between limb joint points and the strict temporal-spatial dependence makes it difficult for shallow machine learning to fully characterize the potential semantic structure of human motion data [5], [6]. Reference [7] summed up the recent progress in the research on the method of synthesizing human motion by using low dimensional features. It also pointed out that it is difficult for traditional algorithms to establish explicit correlation between low dimensional space and high-dimensional human pose [8]. In recent years, with the emergence of deep learning, a great breakthrough has been made in structural data modeling. However, human motion capture data contain complex kinematic articulation relations [9]. Due to its joint data, their rigorous spatial and temporal structures pose a great challenge to the effective application of the deep learning model [10].

Deep learning technology has been applied in many fields, and it also gradually develops from the early simple applications of static image processing and text information completion to high-dimensional applications, such as video and

The associate editor coordinating the review of this manuscript and approving it for publication was Yongtao Hao.

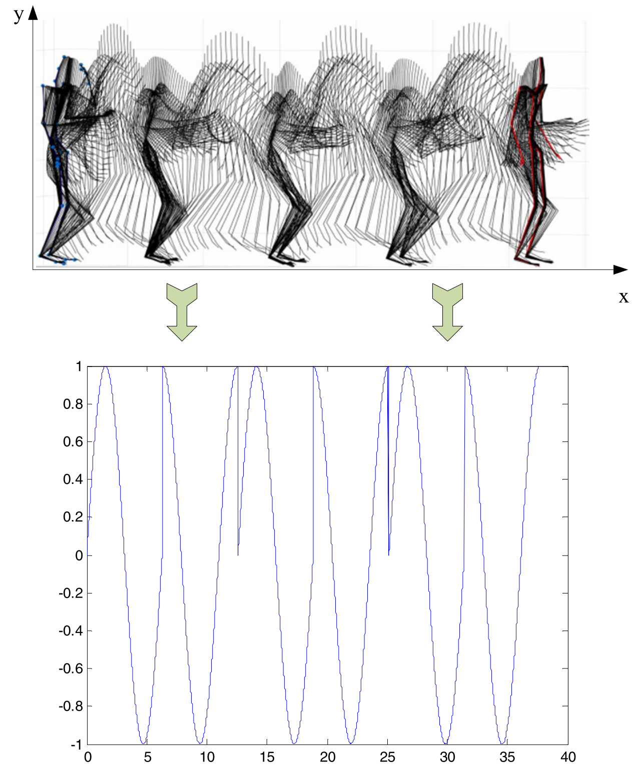

FIGURE 1. **The walking process capture in rehabilitation training.**

audio data processing [11]. With the support of big data, human intelligence can quickly extract the general rule and past data of low-level data by analyzing the past data [12]. On this basis, grasping the principles of pre-processing, and then creating new works in accordance with the principles are the most important problems [13]. Many methods try to accurately capture the movement during rehabilitation training, which helps patients to better recover and train their physical function [14]. The figure of Walking Process Capture in Rehabilitation Training is shown in figure 1. Collecting walking data of rehabilitation training and analysing the effectiveness of training is the important process during the sport medician. Deep learning is used to capture the training process, and accuracy of training detection is also improved. But none of these methods can accurately capture the movements in sports medicine [15], [16].

To improve the learning ability of the model in finite data sets, reference proposed a hierarchical FCRBM (HFCRBM) structure [17]. At the bottom of the HFCRBM model, the directional connection between the historical frame and the current frame in CRBM was removed [18]. The feature extracted from the continuous frame segment was used as the input of the high-level FCRBM. Style synthesis was achieved by linear interpolation in the middle hidden layer of the two models [19]. Similarly, the prototype CRBM and FCRBM are stacked directly [20]. Through unsupervised learning of the corresponding relationship between voice and human posture, the motion temporal features are learned in the bottom CRBM, and the corresponding data of attitude sequence are input in the upper FCRBM tag layer, which can make good use of the prosodic features of voice to drive posture generation [21].

It is found that the traditional RBM model based on Markov restriction can only extract local dynamic constraints [22]. In order to distinguish the spatial structure of human skeleton and the temporal characteristics of continuous motion accurately, reference decomposed the human trunk into five parts, using RBM to extract compression representation as the input of CRBM, and it achieved a smoother transition between different styles of motion [23], [24]. Similarly, reference used the ability of the encoder to reconstruct the original data, and established RBM training for each frame to get the spatial structure features, and connected the hidden layer to connect to get timing information [25]. Reference assumed that the whole motion sequence shared the same spatial transfer characteristics, by adding directional connections between adjacent input frames [26]. In addition, Hasinoff *et al.* [27] adopted the FCRBM weight decomposition idea, and constructed the conditional depth Boltzmann machine (conditional deep Boltzmann machine, CDBM) at the top level of CRBM to extract higher-order spatio-temporal interactions between human joints.. Considering the space-time characteristics, CRBM has more advantages than RBM in the growing motion sequence [28].

A large number of unlabeled and limited style data greatly reduces the reuse possibility of existing motion sequences. Effective classification and fragment splicing have become an important way of data reuse. Aiming at these two problems, this paper focuses on the great success of deep learning in the field of graphics and iconography. Based on the theory of Restricted Boltzmann Machine (RBM), a spatio-temporal feature extraction model for human skeleton medical motion sequences is established. The research contributions are mainly manifested in three aspects:

(1) In this paper, stack factor decomposition spatiotemporal feature model and discriminate RBM are used to construct semi-supervised combination model.

(2) The underlying models firstly use the idea of weight decomposition to construct the three channel generative RBM model; and then it extracts the abstract temporal and spatial characteristics of the original motion sequence. Furthermore, it identifies the behavior style of the current input segment at the top using the discriminate RBM model. Finally conduct the stylistic statistics of the whole motion sequence in the voting space.

(3) An unsupervised similar frame detection model is constructed by using 3D convolution RBM's perception of human adjacent joints' linkage. In this way, Candidate frames for constructing graph model nodes are obtained. Trajectory and style switching control is realized based on attitude similarity screening criteria.

The simulation experiment verifies the superiority and reliability of the algorithm, and it is effectively applied in rehabilitation training.

II. APPLICATION OF DEEP LEARNING IN MEDICAL SPORT TRAINING AND REHABILITAION

Deep neural network (DNN), as the foundation of deep learning theory, has reduced the problem of gradient diffusion and gradient explosion in the traditional back-propagation algorithm while introducing multi-storey computing units. According to the topological structure of DNN [29], deep modeling based on human motion capture data mainly includes 4 kinds of different learning structures and their motion generation applications [30].

A. RESTRICTED BOLTZMANN MACHINE

Restricted Boltzmann machine (RBM) is a stochastic generation neural network that can learn probability distribution through input data sets [31]. Restricted Boltzmann machine has been applied in dimension reduction, classification, collaborative filtering, feature learning and topic modeling. Depending on the task, restricted Boltzmann machine can be trained by supervised learning or unsupervised learning. RBM-based Framework of Training and Rehabilitation is shown in figure 2.

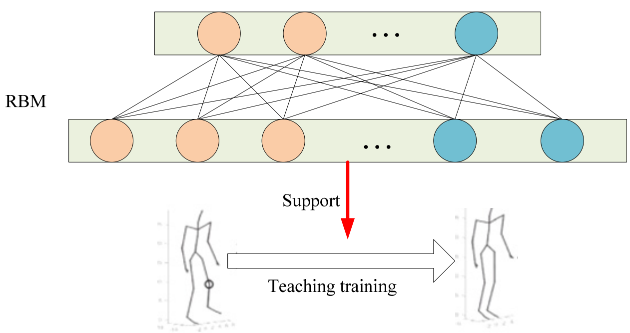

FIGURE 2. RBM-based framework of training and rehabilitation.

Deep neural networks consist of many restricted Boltzmann machine (RBM) stacks, and RBM is assumed to have no connection between the visible neurons and hidden neurons. The deep neural network uses hierarchical unsupervised greedy pre-training method to pre-train RBM hierarchically, and it takes the results as the initial value of the supervised learning training probability model, which greatly improves the learning performance. Unsupervised feature learning is to build a statistical model between the complex hierarchical structure of RBM and a large number of data sets. Through unsupervised pre-training, the network can acquire high-order Abstract features, and provide better initial weights. The weights are limited to the favorable range for global training. The local information between layers is used for layer-by-layer training, and the characteristics of training data are emphasized. The risk of over-fitting learning can be reduced, and the problem of excessive error accumulation and transmission in deep neural networks can be avoided.

Different from the traditional artificial neural network model, RBM is a kind of random neural network model with bipartite graph structure, symmetrical connection and no self-feedback. It is fully connected with the neurons between layers and has no connection within layers. In particular, unsupervised layer-by-layer greedy learning algorithm in RBM greatly improves its income by giving DNN appropriate initial parameters. At the same time, RBM is a generative model based on energy, which allows the network to extract the effective forward feature and reverse reconstruction of input data. Therefore, it has important reference significance and practical application value in learning joint linkage feature representation.

B. DEEP RECURRENT NEURAL NETWORK

Unlike RBM model, RNN is theoretically a directed graph model capable of modeling long time dependence. Its recursive structure on time axis is similar to TRBM, which allows adding feedback connection weights between hidden layers [32]. The difference is that the hidden layer information corresponding to the current frame is affected by the characteristics of the hidden layer, which is called Elman structure and Jordan RNN structure (JRNN) respectively. In practical applications, RNN learning often encounters problems such as gradient explosion or diffusion, long-term information forgetting, and so on. Therefore, a special RNN structure with forgetting gate, called long short term memory (LSTM), has emerged. Usually, in the process of motion sequence generation, besides the consideration of long-term dependence, it also needs to optimize and improve its spatial distribution learning.

As a classical time series modeling method in deep learning, RNN has a unique neuron called memory cell, which realizes the feedback of historical information through hidden layer autoregressive connection. The input of hidden layer includes not only the output of the input layer, but also the output of the hidden layer at the last time, which has been widely used in speech recognition and text generation. Aiming at human motion capture data, the use of RBM is empty. As shown in Figure 3, by using RBM structure to extract features of two different motion styles and interpolate them in feature space, a new motion sequence can be generated through decoding stage.

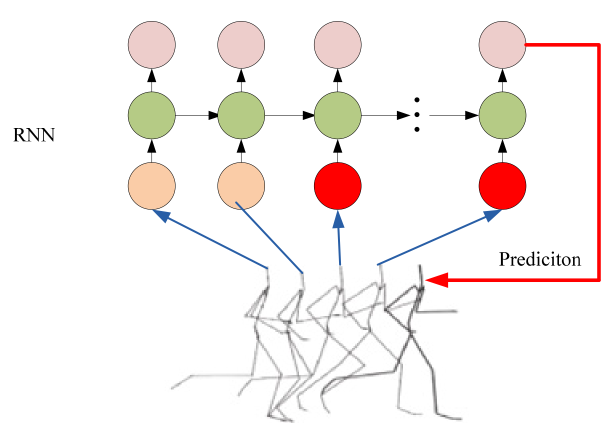

FIGURE 3. Sequential motion frame expansion based on RNN.

C. DEEP CONVOLUTION NEURAL NETWORK

Traditional neural networks have a fully connected structure, which can easily lead to "dimension disaster" and other serious problems [33]. Because of the local information perception, parameter sharing and translation rotation invariance of shared convolution kernels, CNN has achieved great success in the field of computer vision. In view of this, the capability of 3D convolution method in coordinate reconstruction of human joint point is fully embodied. It has a certain modeling ability for the overall structure and timing information of skeleton.

Because of the monotony of capturing data sets, it only stays in the synthesis of simple motion (walking, running, waving, and so on), and it can not achieve more complex motion generation, such as multiplayer scene interaction. User specified constraint satisfaction and so on. Holden et al. first demonstrated the ability of CNN self-encoders consisting of 3 layers convolution structure and de-convolution structure in motion generation. However, the drawback is that it can not guarantee some dynamic characteristics, such as foot touchdown information and joint length, and so on. Based on CNN, sports health care the complex frame diagram is shown in Figure 4.

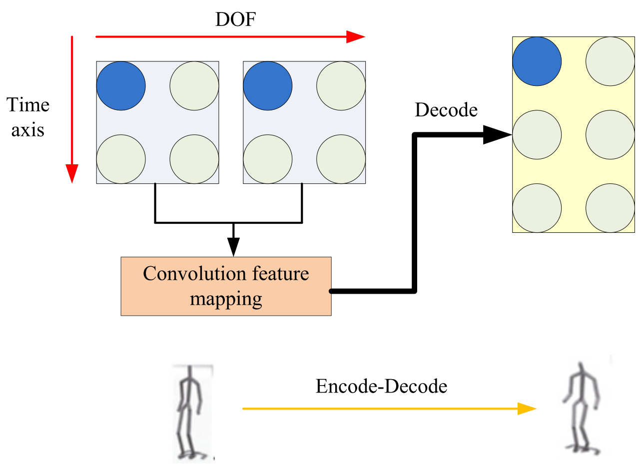

FIGURE 4. Sequential motion frame expansion based on CNN.

D. REINFORCEMENT LEARNING

In the process of learning, the environmental constraints are mapped to specific behaviors by optimizing the cumulative rewards obtained from the interaction between the target and the environment [34]. The famous robot AlphaGo is a milestone in the process of reinforcement learning. However, reinforcement learning has poor performance in modeling ability for high-dimensional human motion capture data. In addition, DRL has important research significance for real-time generation of role motion in complex terrain constrained scenes. It is also very helpful for us to interact with the environment in the process of medical rehabilitation and help patients recover as soon as possible. The framework based on DRL is shown in figure 5.

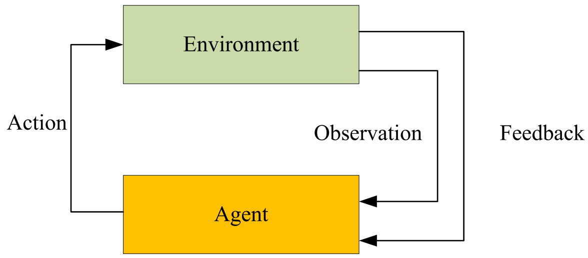

FIGURE 5. Sequential motion frame expansion based on DRL.

Deep enhancement learning aims at dealing with changeable user input in real time by training decision controller, and is suitable for generating more complex motion behaviors, such as changing target tracking, terrain adaptation, and so on. typical traditional reinforcement learning methods, such as controller decision making methods, are based on pre learning strategy table and the original discrete motion segments are entered according to control signals. However, the flexibility of this method is often limited by the size of the database, and its smoothness also depends on the selection of interpolation algorithm.

III. IMPLEMENTATION AND REHABILITATION APPLICATION OF SPORTS MEDICAL MODEL BASED ON IMPROVED RBM ALGORITHM

A. MEDICAL REHABILITATION FRAMEWORK BASED ON IMPROVED BOLTZMANN MACHINE

In view of the inaccurate detection of transition points and inefficient matching of track segments in existing motion graph construction methods, unsupervised feature learning and a motion segmentation method based on declination are proposed to construct motion maps. Firstly, aiming at the spatio-temporal structure of skeletal motion, a convolution constrained Boltzmann machine is used to extract spatiotemporal features automatically on time axis, and then the candidate transition points are detected. Subsequently, the whole motion is segmented based on the deflection angle. In this way, the obtained motion fragments can achieve higher retrieval and reuse efficiency in the stage of searching user's trajectory constraints. The experimental results show that compared with the traditional motion graph construction method, the key posture extracted by this method is more conducive to the use of fast quaternion interpolation to create natural motion transition sequence. At the same time, constraint matching based on skew angle also improves the reusability of original motion capture data, and it does not lack the flexibility of role control while taking into account the speed of motion generation.

Combining the features of deep learning method and feature extraction, this paper proposes an unsupervised depth learning model to extract spatio-temporal characteristics of motion capture sequences, and constructs motion maps based on the clips segmentation method based on skew angle, which

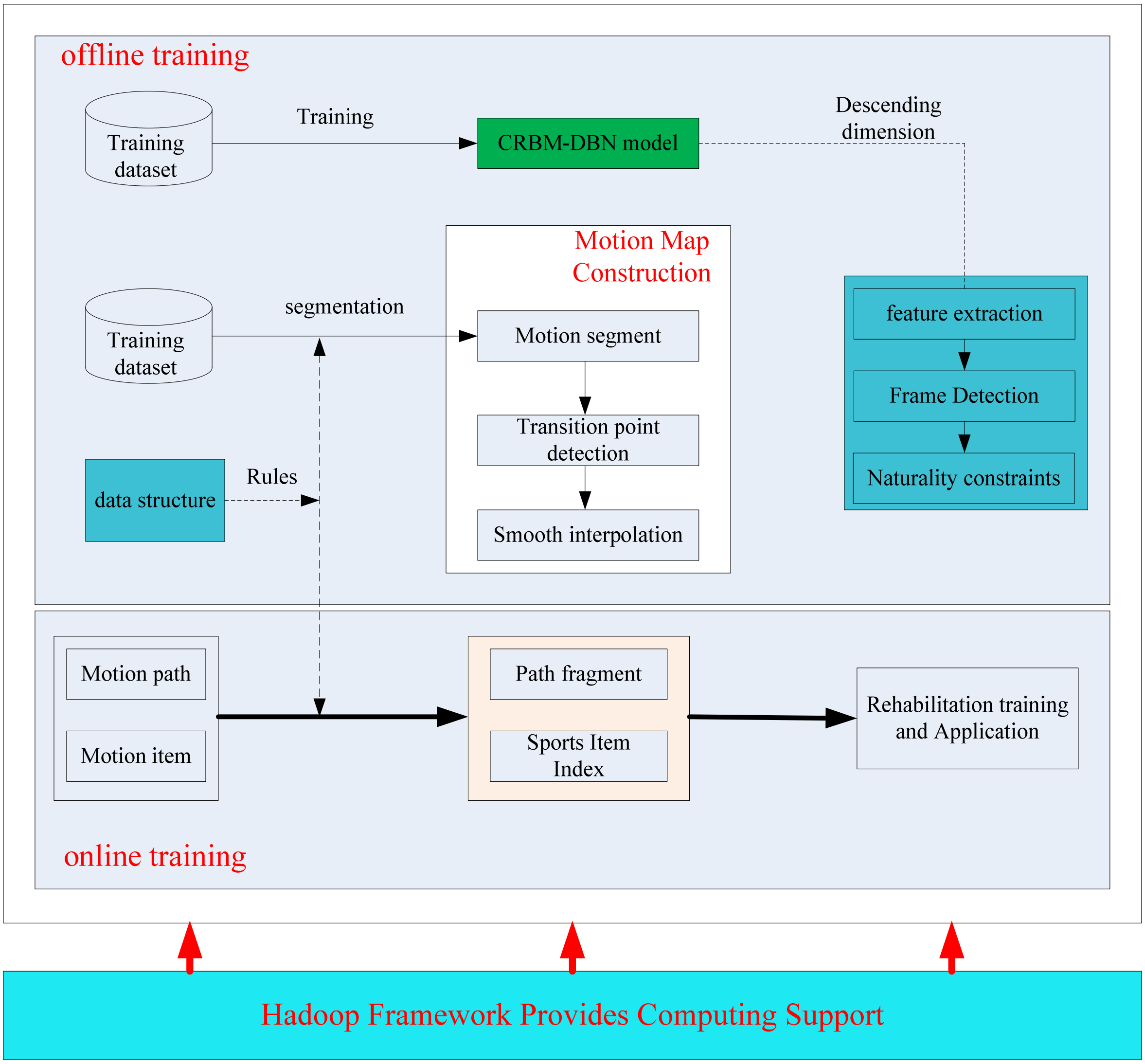

FIGURE 6. The whole architecture of CRBM model.

can achieve faster retrieval efficiency than minimum trajectory offset in the constraint matching process.

As shown in Figure 6, CRBM model is trained with a large number of sample data in the offline phase, which makes it possible to detect transition points in motion diagrams for user data sets. In the online phase, segment segmentation and style retrieval are performed for input trajectories, and the quadruple structure based on Scenes information can match constraints between online input and offline motion diagrams. The image processing of deep learning needs a lot of computation. In this paper, Hadoop cloud platform is used to support the deep learning framework. It can effectively improve the operation efficiency of the platform.

B. TRANSITION FRAME DETECTION MODEL BASED ON CONV-RBM

The common preprocessing method of DNN is to take the whole human body information of each frame as the input vector or separate the human body into five parts: left and right arms, left and right legs, and body. It relies on the full connection layer to extract the linkage of each part. This method lacks the consideration of the local correlation of skeleton and the joint information of limb articulation. To solve this problem, the human skeleton is reordered according to three adjacent joints, each of which represents the local perception domain of convolution operation, so the contribution of each joint to each movement style can be determined automatically. According to the convolution kernel parameter value, avoiding the sensitive operation of determining the weight of each joint is based on prior information in traditional methods.

In order to make full use of the advantages of DNN in automatic feature extraction, this paper proposes stacked 3D-Conv RBM and Deep Belief Network (DBN) for unsupervised learning based on the spatio-temporal structure of motion capture data. The model structure is shown

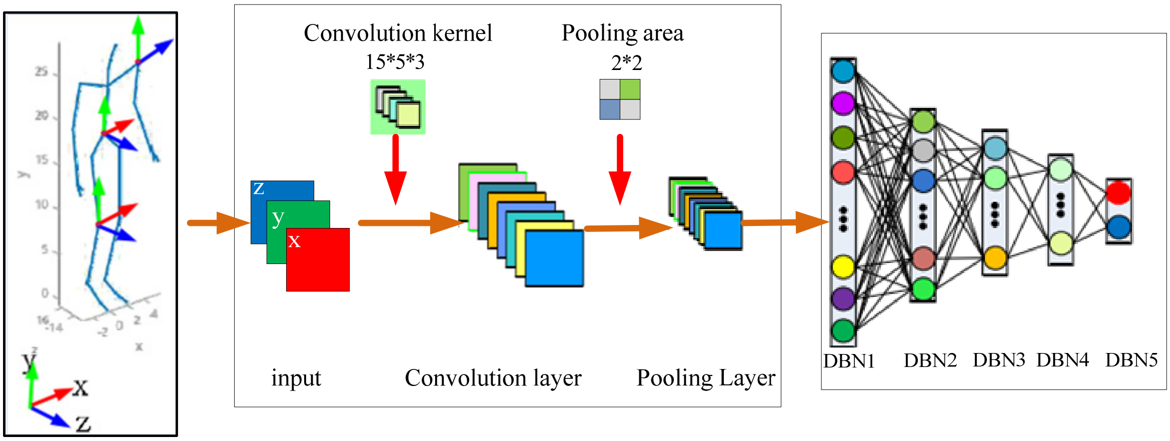

FIGURE 7. **The internal structure of transition frame detection model.**

in Figure 7. By mapping 96-dimensional point cloud data into low-dimensional feature space, candidates are screened according to preset threshold. Cross frame group.

C. THE PRINCIPLE OF CONV-RBM ALGORITHM

Conv-RBM is used to extract unsupervised local structure features of images, and then it is applied to speech recognition. These two applications fully demonstrate the advantages of convolution RBM in unsupervised learning for local structure and temporal features. In view of the inherent spatial structure of human skeleton in each frame of motion capture data and the temporal dependence between frame sequences, Conv-RBM is extended to extract spatial and temporal features of three-dimensional data, and candidate transition points are selected according to the Euclidean distance of frame segments in implicit space. Because 3D Conv-RBM for motion capture data expansion is still a contrast divergence algorithm for training RBM, it differs from the local information sharing field (weight) and bias in the same coordinate, so its energy function can be defined accordingly:

$$E(V,H;\theta) = -\sum_{c}\sum_{\alpha}\sum_{d}\frac{V_{\alpha}^{c}}{\delta}\otimes K_{c,d}\otimes C_{\alpha}^{d} -\sum_{c}\sum_{\alpha}\frac{(V_{\alpha}^{c}-b_c)^2}{2\delta_c^2}-\sum_{d}\sum_{\alpha}C_{\alpha}^{d}b_d \tag{1}$$

where V_{α}^{d} represents the perceptual region of the alpha block in the dimension C ∈ {x, y, z} of the input data; V_{α}^{c} represents the local temporal and spatial characteristics of the current block obtained by convolution operation, $K_{c,d}$ is the weights connecting the input data with the convolution output unit, and b_c and b_d are the neuron biases of the input layer and the convolution layer, respectively. Delta is the variance of training set. Because of the existence of normalization operation, its value is often set to 1.

In essence, each local sensing region in the 3D-ConvRBM model can form a miniature RBM structure with some neurons in the winding layer. According to Gibbs sampling, the signal value transferred from the bottom layer to the top level is expressed as:

$$C_{\alpha}^{d} = \sum_{C\in\{x,y,z\}}\frac{V_{\alpha}^{c}}{\delta}\otimes K_{c,d}+b_d \tag{2}$$

Therefore, when the binary implicit state is adopted, the energy contribution of the activation unit to the model is as follows:

$$C_{\alpha}^{d} = P(C_{\alpha}^{d}=1|V_{\alpha}^{c},\theta) = \frac{e^{C_{\alpha}^{d}}}{1+\sum_{d}e^{C_{\alpha}^{d}}} \tag{3}$$

In image processing, producing neural network not only eliminates the advantage of parameter reduction brought by weight sharing, but also endows the model with better anti-noise ability. This characteristic is also applicable to data collected by motion capture technology which is vulnerable to environmental interference. The most commonly used method in Conv-RBM is maximum probability pooling; however, it lacks more theoretical and experimental support for its effectiveness. Therefore, the maximum probability pooling method, which is widely used at present, is adopted in this paper:

$$P_{xy} = \max(C_{\alpha}^{d})_{2\times 2} \tag{4}$$

where $(\cdot)_{2\times 2}$ represents the size of the area selected by the pooling operation in the convolution layer C_{α}^{d}.

In the reverse reconstruction stage, the Gibbs sampling algorithm is still used. In order to overcome the dimension inconsistency caused by the pooling operation, we can make up the zero operation in the de-convolution operation. Unlike the forward inference process, the visible layer is a real value unit, and its value can be sampled from the specified Gauss distribution:

$$V_{\alpha}^{d} \sim N\left[\sum_{d}C_{\alpha}^{d}\otimes rot\left(K_{c,d},180^0\right)+b_c,\delta^2\right] \tag{5}$$

Formula (3) and (5) show the operation process of forward feature extraction and reverse data reconstruction using convolution RBM, respectively:

$$b_c = D_\alpha^c - (V_\alpha^c)^{(k)} \tag{6}$$
$$b_d = C_\alpha^d - (C_\alpha^d)^{(k)} \tag{7}$$
$$K_{c,d} = D_\alpha^c \times C_\alpha^d - (C_\alpha^d)^{(k)} \times (V_\alpha^c)^{(k)} \tag{8}$$

In the actual training process, RBM is often difficult to learn features from real data with small variance. Therefore, when connecting Conv-RBM with DBN model, the layer normalization method proposed in recent years is used to avoid the phenomenon of "covariate offset". The specific algorithm is as follows:

$$\mu = \frac{1}{H}\sum_{i=1}^{H} P, \quad F_1 = \sqrt{\frac{1}{H}\sum_{i=1}^{H}(P-\mu)^2} \tag{9}$$

μ is the mean value of all connection layer units connected by pooling layer, and F_1 is the connection hub of Conv-RBM and DBN structure.

Because Conv-RBM extracts features through multiple convolution cores, its dimension is generally much larger than the input dimension when its pooled neurons are fully connected. Effective data dimension reduction method is the core to support fast retrieval of similar frames.

D. FRAGMENT SEGMENTATION AND TRANSITION SEGMENT DECISION BASED ON QUATERNION DATA STRUCTURE

Motion graph is a kind of efficient method to explore data reuse and behavior control based on existing motion database. Reference pointed out that the construction of motion graph mainly includes three steps: motion segment segmentation, transition point detection and intermediate transition frame segment generation. Its basic criterion is to ensure that the graph structure still has good connectivity under the premise of satisfying user's arbitrary input constraints. From this point of view, we propose automatic segmentation of long motion sequences based on declination four Meta data structure, and can also be used for efficient retrieval of motion fragments. Candidate transition frames between segments are detected by deep learning feature extraction framework.

In essence, segmentation of original motion into segments is the necessary step to construct motion graphs. Obviously, more detailed partitioning can make the graph model more connected, but it will also increase the complexity of the model; thus, it increases the motion retrieval time. Therefore, an efficient motion graph must make a good trade-off between constraint matching time and graph connectivity.

Figure 8 depicts an example of two motion sequences separately divided into multiple motion segments, each of which is indexed by a quaternion structure. Specifically, the third motion segment of Motion 2 is indexed by a Clip, in which the "Frame" is the start-stop index of the segment in the frame sequence, the "Or" is the orientation letter of the start-stop frame of the segment, and the "10" represents the "Motion 2" in the current data set.

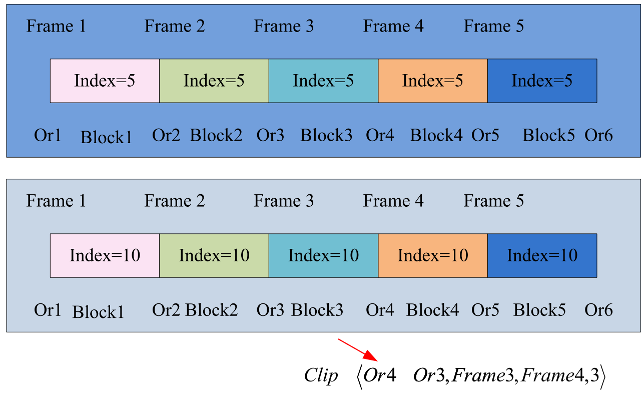

FIGURE 8. Medical motion sequence based on deviated quaternion Index.

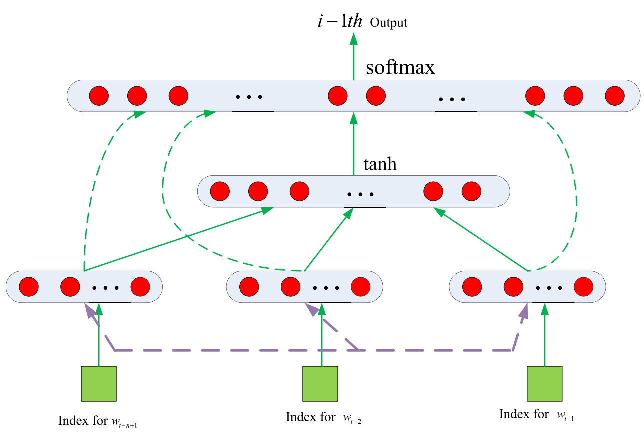

FIGURE 9. The fragment transition segment decision based on deep neural network.

A large number of candidate frames are saved in the motion segments indexed by Quad from algorithm 1. Although each frame pair has a small distance in the hidden space, it is not necessarily the most suitable transition point. Figure 9 shows the fragment transition segment decision based on deep neural network. However, because the transition direction of the left image is opposite to the motion direction, the stitched motion sequence will have a regression phenomenon. Therefore, it is reasonable to impose physical constraints to ensure the natural degree of the transition section. In the subject study, we stipulate that the transition of motion must occur, when at least one foot touches the ground, that is, the two feet are not allowed to transit to other motion sequences.

IV. EXPERIMENTS AND RESULTS

A. DATABASE DESCRIPTION

With the increasing application of digital information technology in the medical industry, clinical nursing information system has become an important part of hospital information construction. Mobile nursing information system is the extension of nurse workstation beside patient bed.

In recent years, the nursing information system based on the medical Internet of things has developed rapidly. On the basis of the existing hospital information management system, mobile handheld devices are used to complete the expansion of hospital information system data to bedside, and the instant delivery of terminal data to the system through the Internet of things technology and wireless local area network as the data interaction platform is also collected. It is accepted by hospital managers and nurses that nursing workflow should be optimized, link quality control should be strengthened, nursing work efficiency should be improved, and more convenient nursing services should be provided for patients. The Internet of Things (IOT) system for medical motion data acquisition is shown in figure 10.

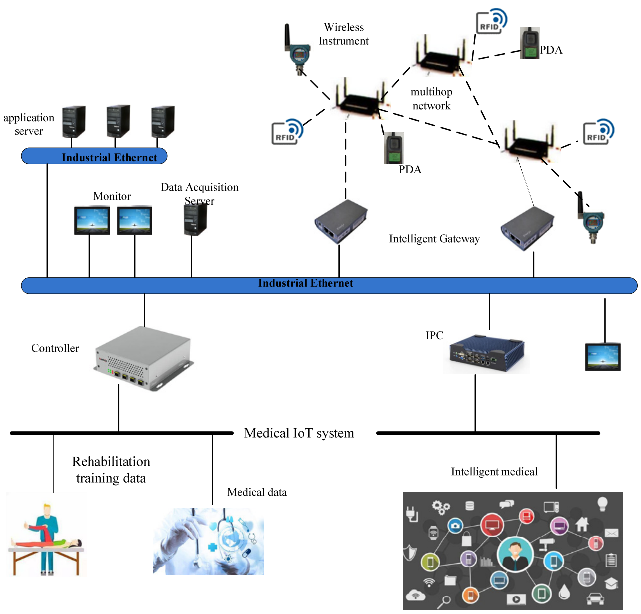

FIGURE 10. The designing diagram of an acquisition device based on the internet of things.

The performance of the proposed algorithm is validated. In the experiment, data records related to walking, running and jumping styles are extracted from the database. Specifically, the training data comes from the rehabilitation movement sequence of 1000 patients collected by the Internet of Things system, which contains 52801 frames, and the amount of data is about half of the training style of literature. Therefore, the amount of data used is enough for the model to extract the movement characteristics of various styles well. In addition, the data sets used to construct motion maps come from 13, 16 and 35, which include forward motion along Z axis, turning left and right to 45 degrees and 90 degrees.

The experimental encoding environment is Matlab 2015b, and the running environment is PC with Intel Core i3-3220 GHz processor and 8G running memory.

B. ACCURACY VERIFICATION OF MEDICAL MOTION MODEL DRIVEN BY BIG DATA

The quality of transition frame segments used to convert style directly affects the coherence of the whole motion sequence. Based on the idea that motion interpolation can produce smoother style conversion between frames with higher similarity, adjacent motion graph nodes representing similar frames and motion edges representing spherical linear interpolation are constructed respectively. In the experiment, 40 frames are selected as the window length of the linear synthesis algorithm, and the starting and ending frames of the mixed segment are the first and last frames of the motion segment to be stitched. The style transition of two kinds of motion synthesis methods based on RBM uses the method of introducing Gauss noise to guide the jump in style hidden space. In contrast, once two similar frames are detected, the quaternion interpolation algorithm can be directly used to generate intermediate frames based on the calculated transition length, which is similar to the key frame interpolation method in the open source engine framework Ogre. Taking the change of running style from walking to running as an example, the motion trajectories of two end effectors (left toe and left hand) are extracted as the basis for analyzing the quality of motion sequences generated by different methods.

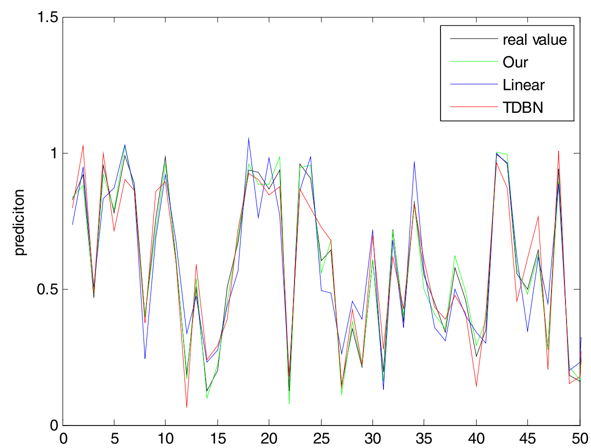

FIGURE 11. The designing diagram of an acquisition device based on the internet of things.

As shown in figure 11, the trajectories generated by various methods have their own characteristics, indicating that they contain different kinematic characteristics and visual effects. In fact, human motion is usually accompanied by slight body shaking, especially in the process of motion capture; the tension of patients directly affects the quality of data. For such reasons, the joint trajectory in the motion sequence based on the motion map constructed from the original motion segment will also have a certain degree of noise.

Therefore, the jitter amplitude and frequency of joint trajectory can be used as the quality evaluation index. Compared with the generative deep learning algorithm, the hybrid method based on windows and the direct quaternion interpolation method can generate smoother motion data. The main reason is that different deep learning models have different ability to model the dynamic characteristics of motion capture data. In order to achieve motion transition, the method of applying noise to realize hidden space jump makes this kind of algorithm have great limitations. Once it is necessary to model multiple motion styles, the noise level required for different styles of hidden space to jump each other can not be determined; furthermore, the motion quality produced by different Gauss noise can not be guaranteed, so it is easy to produce sliding steps.

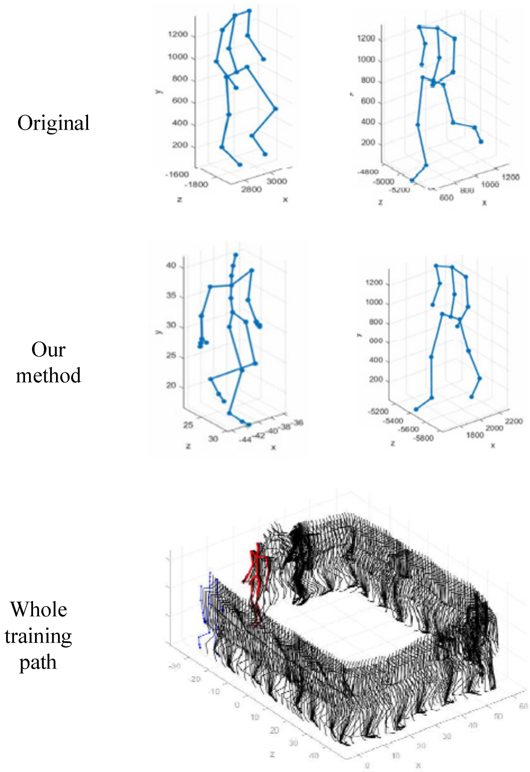

FIGURE 12. **The diagram of breast mammography with preprocessing.**

Figure 12 shows the partial frame visualization results of the motion data generated by the proposed model when the central distance vector is used as input. It can be found that the generated motion data can ensure the basic biological structure of the human body, without joint dislocation, skeletal deformation and other conditions. It shows that the optimized model structure based on Conv-RBM has good spatial feature extraction ability. Specifically, it is found from the comparison of visual effects that when Conv-RBM generates 400 frames of motion data, the center of gravity of the human body shifts and the upper and lower limbs are inclined. With the increase of time, the generated motion quickly converges to the mean posture, especially the changes of the hip and shoulder bones. Generally speaking, whether the spatial structure of the motion frames itself or the satisfaction degree of style constraints are in line with naked eye comfort.

C. SUPERIORITY VERIFICATION OF MEDICAL MOTION MODEL DRIVEN BY BIG DATA

The main purpose of the experiment is to explore the influence of coordinate and Euler angle input on the quality of motion generation, so it mainly includes the preprocessing of these two types of data. The first one is to extract the absolute position change of each joint relative to the root joint as the input of the model, using the central distance vector as the data processing method for the absolute position coordinates of the skeleton in the 3D space. The second is to use the joint rotation information as data input. Firstly, the Euler angle data is converted into an exponential mapping type to avoid the non-smoothness caused by the universal lock. In addition, the absolute position information of the root joint is replaced by the orientation and forward difference vectors of the human body in the current posture. In order to adapt to the method of model inference, it is necessary to normalize the processed data to ensure that its variance is 1. The comparison results between the proposed algorithm and traditional neural networks, RBM and RNN networks are shown in figure 13.

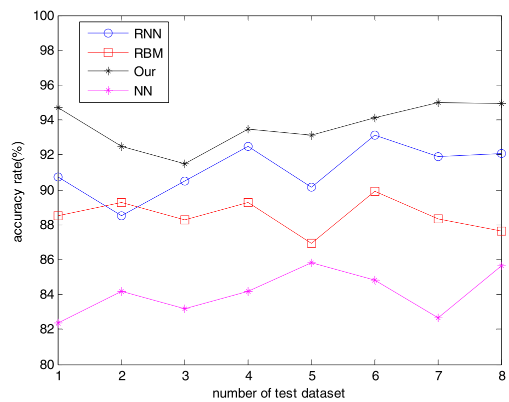

FIGURE 13. **The diagram of breast mammography with preprocessing.**

As shown in the figure 13, the proposed deep network model can achieve about 94% accuracy. Compared with RNN and RBM, they increased by 2% and 3%. Compared with the traditional neural network, our algorithm improves by nearly 10%. According to the quality of motion frame generated by two types of data, we can draw a preliminary conclusion. On the premise that the selection of model superparameters is close to optimal, the amount of spatio-temporal information transmission captured by different connection modes in graph model is also different, which directly affects the quality of data reconstruction. Therefore, further mining of more interactive information extraction model is parametric motion. In addition, the data based on coordinate centralization implicitly contains the body structure of the sports performers in the learning process, so they are better at switching between different skeletal structures, and it also have certain potential in the research of motion redirection. The motion frames generated based on rotation information input have higher data quality, and they are suitable

TABLE 1. **Comparison of experimental results of different NN models.**

Model	Our	RBM	RNN	NN
Time	321.311	333.921	351.321	251.012
Accuracy	0.945	0.925	0.916	0.832

for single person because they use the original skeletal data. The detailed compassion results are shown in Table 1. As is shown, our model has less time consumption and more accuracy compared with other deep neural network. Compared with traditional neural network, we consume more time but get more accuracy.

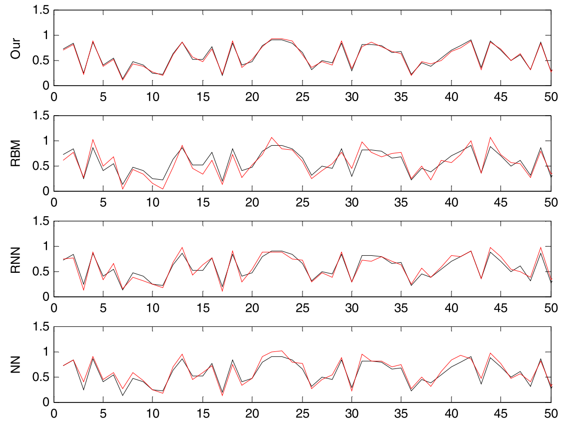

FIGURE 14. **The diagram of breast mammography with preprocessing.**

In order to better understand the smoothness of reconstructed motion on the time axis, figure 14 shows the center distance curves of the apex end effectors relative to the root joint of 400 frames of motion sequence generated by four models receiving two types of data input. The motion quality is judged by the magnitude and frequency of jitter. From the comparison graph, it can be clearly seen that the motion model for the central distance vector has less high-frequency fluctuation. In addition, the high-frequency jitter in the result graph of CRBM and RBM is also slightly lower than that of the motion sequence based on exponential mapping. RNN has the opposite result. It can be seen from figure 14 that the distance between the tips and the root joints of the motion sequence generated by the central distance vector gradually stabilizes, which indicates that the motion of the whole leg gradually stops, while the motion sequence generated by the exponential mapping still maintains a cyclic state. These two phenomena indicate that the quality of motion generation is closely related not only to the model structure, but also to the data processing. In designing the deep learning model for extracting temporal and spatial features, not only the expressive ability of the model, but also the inherent bio-kinematic characteristics of human body, such as skeletal priori, velocity and acceleration changes, should be considered.

In order to verify the convergence speed and training efficiency of the algorithm, this paper compares with two other RNN and RBM deep neural networks. As shown in the figure 15, the original model can have better and faster convergence. In addition, the convergence efficiency is much better than the other two network models. Experiments verify the efficiency of our algorithm.

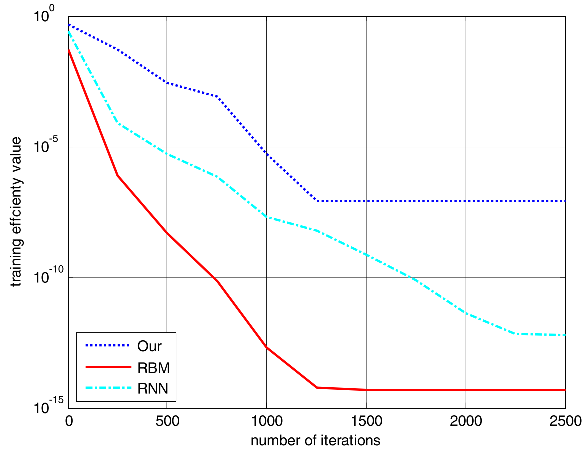

FIGURE 15. **The diagram of breast mammography with preprocessing.**

V. CONCLUSION

A large number of unlabeled and limited style data greatly reduces the reuse possibility of existing motion sequences. Effective classification and fragment splicing have become an important way of data reuse. Aiming at these two problems, this paper focuses on the great success of deep learning in the field of graphics and iconography. Based on the theory of Restricted Boltzmann Machine (RBM), a spatio-temporal feature extraction model for human skeleton medical motion sequences is established. The research results are mainly manifested in three aspects. (1) In this paper, stack factor decomposition spatiotemporal feature model and discriminate RBM are used to construct semi-supervised combination model. (2) The underlying model firstly uses the idea of weight decomposition to construct the three channel generative RBM model; and then it extracts the abstract temporal and spatial characteristics of the original motion sequence. Furthermore, it identifies the behavior style of the current input segment at the top using the discriminate RBM model. Finally conduct the stylistic statistics of the whole motion sequence in the voting space. (3) An unsupervised similar frame detection model is constructed by using 3D convolution RBM's perception of human adjacent joints' linkage. In this way, Candidate frames for constructing graph model nodes are obtained. Trajectory and style switching control is realized based on attitude similarity screening criteria. The simulation experiment verifies the superiority and reliability of the algorithm, and it is effectively applied in rehabilitation training. The research of deep learning on sporting training and rehabilitation training will be conducted in the future.

REFERENCES

[1] N. Sünderhauf, O. Brock, W. Scheirer, R. Hadsell, D. Fox, J. Leitner, B. Upcroft, P. Abbeel, W. Burgard, M. Milford, and P. Corke, "The limits and potentials of deep learning for robotics," *Int. J. Robot. Res.*, vol. 37, nos. 4–5, pp. 405–420, 2018.

[2] C. Zhang, H. Zhang, J. Qiao, D. Yuan, and M. Zhang, "Deep transfer learning for intelligent cellular traffic prediction based on cross-domain big data," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1389–1401, Jun. 2019.

[3] F. Kamali and E. M. S. Morovati, "Comparison of upper trapezius and infraspinatus myofascial trigger point therapy by dry needling in overhead athletes with unilateral shoulder impingement syndrome," *J. Sport Rehabil.*, vol. 28, no. 3, pp. 243–249, 2018.

[4] M. Zangger, D. Poethig, F. Meissner, M. von Wolff, and P. Stute, "Towards ICF implementation in menopause healthcare: A systematic review of ICF application in Switzerland," *Swiss Med. Weekly*, vol. 147, no. 5152, 2017, Art. no. w14574.

[5] A. Aboabat and H. Qannam, "Development of an Arabic inpatient satisfaction survey: Application in acute medical rehabilitation setting in Saudi Arabia," *BMC Health Services Res.*, vol. 17, no. 13, Sep. 2017, Art. no. 664.

[6] Y. Li, F. X. Wu, and A. Ngom, "A review on machine learning principles for multi-view biological data integration," *Briefings Bioinform.*, vol. 19, no. 2, pp. 325–340, 2018.

[7] H.-M. Gross, A. Scheidig, K. Debes, E. Einhorn, M. Eisenbach, S. Mueller, T. Schmiedel, T. Q. Trinh, C. Weinrich, T. Wengefeld, A. Bley, and C. Martin, "ROREAS: Robot coach for walking and orientation training in clinical post-stroke rehabilitation—Prototype implementation and evaluation in field trials," *Auton. Robots*, vol. 41, no. 3, pp. 679–698, 2017.

[8] T. Karras, T. Aila, S. Laine, A. Herva, and J. Lehtinen, "Audio-driven facial animation by joint end-to-end learning of pose and emotion," *ACM Trans. Graph.*, vol. 36, no. 4, Jul. 2017, Art. no. 94.

[9] A. Ziletti, D. Kumar, M. Scheffler, and L. M. Ghiringhelli, "Insightful classification of crystal structures using deep learning," *Nature Commun.*, vol. 9, no. 1, 2018, Art. no. 2775.

[10] R. Winter, F. Montanari, F. Noé, and D.-A. Clevert, "Learning continuous and data-driven molecular descriptors by translating equivalent chemical representations," *Chem. Sci.*, vol. 10, no. 6, pp. 1692–1701, Jul. 2019.

[11] A. Dutkiewicz, R. D. Müller, X. Wang, S. O'Callaghan, J. Cannon, and N. M. Wright, "Predicting sediment thickness on vanished ocean crust since 200 Ma," *Geochem. Geophys. Geosyst.*, vol. 18, no. 12, pp. 4586–4603, 2017.

[12] S. Liu, Z. Li, V. Srikumar, V. Pascucci, and P.-T. Bremer, "NLIZE: A perturbation-driven visual interrogation tool for analyzing and interpreting natural language inference models," *IEEE Trans. Vis. Comput. Graphics*, vol. 25, no. 1, pp. 651–660, Jan. 2019.

[13] S. R. Filbay, E. M. Roos, R. B. Frobell, F. Roemer, J. Ranstam, and L. S. Lohmander, "Delaying ACL reconstruction and treating with exercise therapy alone may alter prognostic factors for 5-year outcome: An exploratory analysis of the KANON trial," *Brit. J. Sports Med.*, vol. 55, no. 22, pp. 1622–1629, 2017.

[14] F. Hoekstra, M. A. G. van Offenbeek, R. Dekker, F. J. Hettinga, T. Hoekstra, L. H. V. van der Woude, and C. P. van der Schans, "Implementation fidelity trajectories of a health promotion program in multidisciplinary settings: Managing tensions in rehabilitation care," *Implement. Sci.*, vol. 12, no. 1, 2017, Art. no. 143.

[15] M.-E. Lamontagne, J. Gargaro, P. Marier-Deschêne, C. Truchon, M. T. Bayley, S. Marshall, C. Kagan, A. Brière, and B. Swaine, "A survey of perceived implementation gaps for a clinical practice guideline for the rehabilitation of adults with moderate to severe traumatic brain injury," *J. Head Trauma Rehabil.*, vol. 33, no. 5, pp. 306–316, 2018.

[16] M. Ahmed, O. Moselhi, and A. Bhowmick, "Two-tier data fusion method for bridge condition assessment," *Can. J. Civil Eng.*, vol. 45, no. 4, pp. 197–214, 2018.

[17] J. Neale, A. C. Smith, and N. C. Bishop, "Effects of exercise and sport in solid organ transplant recipients: A review," *Amer. J. Phys. Med. Rehabil.*, vol. 96, no. 4, pp. 273–288, 2017.

[18] X. Gao, S. Jin, C.-K. Wen, and G. Y. Li, "ComNet: Combination of deep learning and expert knowledge in OFDM receivers," *IEEE Commun. Lett.*, vol. 22, no. 12, pp. 2627–2630, Dec. 2018.

[19] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017.

[20] C. Feng, M. Cui, B.-M. Hodge, and J. Zhang, "A data-driven multi-model methodology with deep feature selection for short-term wind forecasting," *Appl. Energy*, vol. 190, no. 15, pp. 1245–1257, Mar. 2017.

[21] W. Ma, F. Cheng, and Y. Liu, "Deep-learning enabled on-demand design of chiral metamaterials," *ACS Nano*, vol. 12, no. 6, pp. 6326–6334, 2018.

[22] L. Hughes, B. Paton, B. Rosenblatt, C. Gissane, and S. D. Patterson, "Blood flow restriction training in clinical musculoskeletal rehabilitation: A systematic review and meta-analysis," *Brit. J. Sports Med.*, vol. 51, no. 13, pp. 1003–1011, 2017.

[23] H. F. Nweke, Y. W. Teh, M. A. Al-Garadi, and U. R. Alo, "Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges," *Expert Syst. Appl.*, vol. 105, pp. 233–261, Sep. 2018.

[24] S. E. P. Munce, I. D. Graham, N. M. Salbach, S. B. Jaglal, C. L. Richards, J. J. Eng, J. Desrosiers, M. MacKay-Lyons, S. Wood-Dauphinee, N. Korner-Bitensky, N. E. Mayo, R. W. Teasell, M. Zwarenstein, J. Mokry, S. Black, and M. T. Bayley, "Perspectives of health care professionals on the facilitators and barriers to the implementation of a stroke rehabilitation guidelines cluster randomized controlled trial," *BMC Health Services Res.*, vol. 17, Jun. 2017, Art. no. 440.

[25] E. King, A. Franklyn-Miller, C. Richter, E. O'Reilly, M. Doolan, K. A. Moran, S. C. Strike, and E. C. Falvey, "Clinical and biomechanical outcomes of rehabilitation targeting intersegmental control in athletic groin pain: Prospective cohort of 205 patients," *Brit. J. Sports Med.*, vol. 52, no. 16, pp. 1054–1062, 2018.

[26] T. M. Quan, T. Nguyen-Duc, and W.-K. Jeong, "Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1488–1497, Jun. 2018.

[27] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, "Deep bilateral learning for real-time image enhancement," *ACM Trans. Graph.*, vol. 36, no. 4, p. 118, 2017.

[28] G. Satat, M. Tancik, O. Gupta, B. Heshmat, and R. Raskar, "Object classification through scattering media with deep learning on time resolved measurement," *Opt. Express*, vol. 25, no. 15, pp. 17466–17479, 2017.

[29] S. Wang, M. Zhou, Z. Liu, D. Gu, Y. Zang, D. Dong, O. Gevaert, and J. Tian, "Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation," *Med. Image Anal.*, vol. 40, pp. 172–183, Aug. 2017.

[30] Y. Chen, Y. Wang, D. S. Kirschen, and B. Zhang, "Model-free renewable scenario generation using generative adversarial networks," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3265–3275, May 2018.

[31] H. Wen, J. Shi, W. Chen, and Z. Liu, "Transferring and generalizing deep-learning-based neural encoding models across subjects," *NeuroImage*, vol. 76, no. 3, pp. 152–163, Aug. 2018.

[32] S. Rasp and S. Lerch, "Neural networks for post-processing ensemble weather forecasts," *Monthly Weather Rev.*, vol. 146, no. 11, pp. 3885–3900, May 2018.

[33] J. D. Young, C. Cai, and X. Lu, "Unsupervised deep learning reveals prognostically relevant subtypes of glioblastoma," *BMC Bioinf.*, vol. 18, Oct. 2017, Art. no. 381.

[34] Y. Du, Z. Fu, and V. D. Calhoun, "Classification and prediction of brain disorders using functional connectivity: Promising but challenging," *Frontiers Neurosci.*, vol. 12, p. 525, Aug. 2018.

YANFENG SU was born in Han nationality, Henan, Xinxiang, in January 1977. He received the master's degree from Zhengzhou University, in 2008. Since 2002, he has been with Xinxiang Medical University for 17 years. His current research interest includes sports rehabilitation. He has published ten papers during his tenure as a Lecturer, including five Chinese cores in SCI. During his tenure, he instructed students to participate in the 2016 Henan University Students Athletics Association, the fifth place in the third place of 3000 meters.