Computer Aided Dance Art Action Identification and Classification Algorithm Based on Machine Learning

Xu Han[1] and Sifan Zhang[2]

[1]Department of Dance, Hanyang University, Seoul 04763, South Korea, hanxu313@hanyang.ac.kr
[2]School of Music, Shaanxi Normal University, Xi'an, Shaanxi 710062, China, zhangsifan@snun.edu.cn

Corresponding author: Xu Han, hanxu313@hanyang.ac.kr

Abstract. The purpose of this article is to propose an effective dance art movement identification and classification algorithm by combining ML (Machine Learning) and computer technology, and verify and test it with a large-scale dance data set. In this article, CNN (Convolutional Neural Network) combined with SVM (Support Vector Machine) algorithm is used to extract the features of dance actions, which are classified and recognized. In this article, a wealth of experiments are carried out, including the comparative analysis with other advanced algorithms, so as to comprehensively assess the performance of the proposed algorithm. The experimental results show that the accuracy, recall and F1 score of this algorithm are all above 92%, and the ROC curve also presents an ideal shape, which shows that this algorithm is superior and effective in the task of dance art movement identification and classification. The research results of this article can provide more possibilities for the research, teaching, creation and appreciation of dance art, and also promote the use of ML and computer vision technology in the fields of art and humanities.

Keywords: Machine Learning; Dance art; Motion identification; Action classification; Convolutional Neural Network; Support Vector Machine
DOI: https://doi.org/10.14733/cadaps.2024.S14.171-186

1 INTRODUCTION

In the field of dance art, dance action is the core element to express and convey dance emotion. Traditional dance action identification and classification mainly rely on manual observation and empirical judgment, which is time-consuming and easily influenced by subjective factors. With the rapid development of computer vision and deep learning technology, automatic recognition of human movements in RGB D images has become a hot research topic. Human motion recognition involves extracting features from videos or images, and then classifying and recognizing these features. Barkoky and Charkari [1] explored feature extraction methods based on complex networks and their applications in RGB-D human motion recognition. The feature extraction method based on complex

© 2024 U-turn Press LLC, http://www.cad-journal.net

networks has broad application prospects in RGB-D human motion recognition. By combining deep learning technology and complex network models, richer features can be extracted from images, improving recognition accuracy and robustness. Perform convolution operations on images through convolutional layers to extract image features. CNN has been widely used in RGB-D image processing. Processing sequence data through recurrent neural units is suitable for processing video or continuous frame image data.

A special RNN that can remember historical information and is suitable for processing image data with temporal relationships. Complex networks are a nonlinear network model with highly complex structures and dynamic characteristics. In RGB-D human motion recognition, complex networks can be used to extract richer features and improve recognition accuracy. ML can automatically learn rules and patterns from massive data, which provides a new idea for solving complex problems. Sports dance, as a unique artistic form, perfectly combines music rhythm, body movements, and spatial layout. In order to improve the expressive and visual effects of sports dance, visual rehearsal systems are widely used in daily training and competitions. With the development of technology, especially the advancement of three-dimensional mathematical modeling technology, new possibilities have been provided for the optimization of sports dance visual rehearsal systems. Chen et al. [2] explored how to apply three-dimensional mathematical modeling technology to a sports dance visual rehearsal system. In sports dance, the precise description of the dancer's posture and movements is the key to improving dance performance. 3D mathematical modeling techniques, such as human pose estimation and motion capture, can effectively obtain the full body posture and motion data of dancers. Through these data, it is possible to analyze whether the dancer's movements comply with norms, as well as their fluency and coordination in the dance. In sports dance, the relationship between dancers, partners, and stage space is an important factor affecting the dance effect. By utilizing three-dimensional mathematical modeling technology, a model of the dancer and stage space can be established, and precise planning and layout of the dance space can be achieved by simulating the dancer's movements and postures. Dance motivation is an important factor in determining the effectiveness of dance. By introducing mechanical models in physics, such as Newton's laws of motion and conservation of momentum, the dynamics and speed of dancers can be simulated and analyzed. Fan and Gao [3] proposed a scheme to enhance human activity recognition using wearable sensors through a hybrid feature selection method. This scheme combines rule-based and statistical methods for feature selection, which can more comprehensively and accurately describe human activities. The experimental results show that our method can achieve high accuracy recognition and has strong adaptability and robustness. This helps dancers understand and improve their dance motivation, thereby improving the quality of the dance. This technology can not only improve the efficiency of identification and classification, but also assess dance actions more objectively. Therefore, it is of great practical significance and application value to apply ML technology to computer-aided dance art movement identification and classification. In our digital world, technological advancements are constantly changing the way we live and work. Among them, the rapid development of computer vision and artificial intelligence technology has brought new possibilities to many fields, including dance recognition and evaluation. However, building an efficient and accurate automated dance recognition and evaluation tool still faces many challenges. Meanwhile, with the popularization of mobile devices and the advancement of network technology, mobile computing and communication have also become an indispensable part of this process. Faridee et al. [4] explored the challenges of building automated dance recognition and evaluation tools using Happyfeed and GetMobile, as well as the applications of mobile computing and communication. Dance is a highly expressive art form, with each dancer having their own unique style and expression. This makes it very difficult to establish a universal dance recognition and evaluation model. In dance performances, factors such as background, lighting, and clothing may interfere with the performance of computer vision systems.

Creativity, a mysterious and profound concept, has always been an important research object in disciplines such as human psychology, art, and sociology. In the field of contemporary music and dance, creativity is particularly prominent, not only playing a decisive role in the formation of artistic works, but also triggering new thinking in interdisciplinary fields such as human-computer

© 2024 U-turn Press LLC, http://www.cad-journal.net

interaction. Hsueh et al. [5] explored the non-linear processes and flowing roles of creativity in contemporary music and dance through deconstruction. Nonlinear processes, commonly applied in the fields of mathematics and physics, refer to phenomena that cannot be explained by a single causal relationship. In music and dance, nonlinear processes are manifested as the complexity and uncertainty of the creative process, which often endows the work with unique charm and depth. Taking improvisation dance as an example, the dancer's body creates coherent and expressive dance movements through interaction with music, space, and other dancers. This process is highly nonlinear, and dancers need to adapt flexibly and translate it into practical actions, while maintaining a global perception and response during this process. From the perspective of human-computer interaction, the deconstruction of creativity provides us with new ways to understand the role of humans in human-computer interaction. The role of a machine can flexibly transition between providing inspiration, stimulating human creativity, and assisting human performance. Computer vision is a science that studies how to make computers get information from images or videos, understand the content and make decisions. In the identification and classification of dance artistic movements, the theory and technology of computer vision can be used to extract the visual features of dance actions. Among them, image processing: involves image preprocessing, enhancement and transformation, which provides the basis for subsequent feature detection and identification. Feature extraction: by calculating the key points, edges, textures and other information in the image, the characteristics of dance actions are characterized. Target detection and tracking: used to determine the spatial position and time series relationship of dance actions. Head pitching is a common sport that involves rapid and complex sequences of movements. In order to understand and optimize the biomechanics of this movement, we need a system that can capture these rapid movements. For this purpose, Lapinski et al. [6] designed a wireless wearable inertial motion sensing system to capture rapid motion biomechanics in overhead pitching. The system consists of a series of wireless wearable sensors that can be fixed to the athlete's head and racket. Each sensor contains a three-axis accelerometer and a three-axis gyroscope, which can capture the dynamic motion of athletes. The data of the system is wirelessly transmitted to the receiver, and then transmitted to the computer for further processing through USB or Bluetooth. The system provides new tools for understanding and optimizing the biomechanics of rapid motion in overhead pitching. Future research directions may include improving the accuracy and response time of sensors, as well as increasing the wearability of systems to make them more comfortable and flexible. In addition, further research is needed to verify the effectiveness of our system in practical applications.

Dance artistic action is the basic unit of dance works, which contains rich emotions, stories and skills. The identification and classification of dance actions is the basis of in-depth analysis and understanding of dance. Among them, dance action identification is to accurately mark each movement in the dance by observing and analyzing the dance video. The classification of dance actions is to classify movements into different categories according to their attributes, styles and emotions. In order to realize automatic dance action identification and classification, it is needed to combine computer vision technology and ML algorithm to build an efficient and accurate identification and classification system. The research in this direction is of great significance to the inheritance, teaching, creation and automatic analysis of dance works. The goal of this study is to develop a computer-aided dance art movement identification and classification algorithm based on ML, so as to improve the identification accuracy of dance actions. Its innovations are as follows:

⊖ This article proposes to integrate the spatial and temporal multimodal features of dance actions, which can describe dance actions more comprehensively and accurately than traditional methods which only focus on a single modality.

⊜ This article innovatively combines the advantages of CNN and SVM. CNN is used to automatically extract advanced features, and SVM algorithm is used for efficient classification. This combination improves the overall performance of the algorithm.

⊛ This article designs an end-to-end training strategy, which can optimize the two steps of feature detection and classification at the same time, thus improving the accuracy of dance action identification.

© 2024 U-turn Press LLC, http://www.cad-journal.net

The article is structured as follows. Firstly, through the in-depth analysis and research of relevant literature, this study understands the research status and progress, and provides theoretical support for this study. Secondly, the feature detection and algorithm design of dance actions are carried out by using computer vision and ML technology. The specific technical route includes: collecting and preprocessing dance action data, extracting effective dance action features, designing dance action identification and classification algorithm based on ML, and conducting experiments to verify the performance of the algorithm. Finally, the results are deeply analyzed and discussed to assess the effectiveness and superiority of the proposed algorithm.

2 RELATED THEORETICAL BASIS

With the development of computer vision and pattern recognition technology, human motion recognition has become a hot research field. Human motion recognition refers to extracting human motion information from videos or images, and then identifying the meanings represented by these actions. In the field of human motion recognition, graphic based 3D human skeleton motion recognition method is an important technical means. Li and Leung [7] introduce the principle, advantages, and application scenarios of this method. The basic principle of a graphic based 3D human skeleton motion recognition method is to model the human skeleton, match the human model with the actual human body in images or videos, and identify the human body's movements. This method requires first obtaining the geometric shape and position information of the human body through techniques such as 3D scanning, and then using pattern recognition algorithms to process this information to identify human actions. By identifying and analyzing athletes' movements, it can help coaches and athletes better understand sports skills and improve sports performance. By identifying and analyzing the movements of the elderly and patients, medical staff can better monitor their health status and provide timely treatment and care. By identifying and analyzing player actions, a more realistic and immersive gaming experience can be developed. In the field of human motion recognition, the recognition of high dynamic dance movements is a challenging problem. High dynamic dance movements involve rapid body movements and complex posture changes, making traditional methods based on static images or single frame videos difficult to accurately recognize. Therefore, Luo and Ning [8] proposed a highly dynamic dance motion recognition method based on video visual analysis. Utilize deep learning models such as convolutional neural networks or recurrent neural networks to classify and recognize extracted features. By training the model, it can automatically learn the inherent features of dance movements, thereby improving recognition accuracy. Perform post-processing on the recognition results, including adjusting the action bounding box, correcting posture, etc., to obtain more accurate and refined recognition results. Experiments were conducted using publicly available high-dynamic dance action datasets to compare the performance of the proposed method with other mainstream methods. The experimental results show that the method proposed in this paper has advantages in recognition accuracy, stability, and real-time performance. Specifically, the proposed method can effectively identify high dynamic dance movements and has good robustness, which can play an important role in practical applications.

Through motion capture, dance movements can be transformed into digital information, which can be further used for analysis, editing, simulation, and reproduction. Mazian et al. [9] explored the theoretical framework for creating folk dance action templates using action capture. Motion capture technology is a technique that tracks and records human motion data. It has been widely used in fields such as film production, game development, motion analysis, and dance art. In the field of dance, motion capture can capture the subtle movements and expressions of dancers, providing feedback and helping them improve their skills and expression. During a dancer's performance, data is collected using motion capture devices to convert the dancer's movements into digital information. Process the collected data through post-processing software to generate dance action templates that can be edited and reproduced. Apply the generated dance action template to other dancers or virtual characters to reproduce the charm of folk dance. The theoretical framework for creating folk dance action templates using action capture can not only help us better record, analyze, and preserve the cultural heritage of folk dance, but also improve the quality and efficiency of dance education, provide

© 2024 U-turn Press LLC, http://www.cad-journal.net

real dance actions for virtual characters, and enrich the content of entertainment industries such as film, television, and games. Pan et al. [10] synthesized a group of dance actions using a keyed dance synthesis method based on a transformer controller. The experimental results show that this method can achieve high fidelity dance synthesis and has good flexibility and scalability. Compared with traditional dance synthesis methods, this method can achieve more precise control of the movements and postures of dancers, thereby improving the quality of dance synthesis. Transformer controller is a widely used controller in computer graphics, which can accurately control the position, rotation, scaling and other parameters of objects. In dance synthesis, the transformer controller can be used to accurately adjust the posture of dancers and render animations. By adjusting the parameters of the converter controller, flexible control of various movements and postures of dancers can be achieved. Keyed dance synthesis method is a keyframe based dance synthesis method that samples the actions of dancers to obtain a series of keyframes, and then interpolates to obtain intermediate frames. This method can achieve high fidelity simulation of dance movements and has good flexibility and scalability. The precise capture and visualization of human behavior in healthcare and virtual reality are of great significance. However, existing motion capture technologies often rely on expensive and complex devices, such as multi camera systems or depth sensors.

This limits its popularity and use in many applications. Therefore, developing a method for capturing and visualizing human behavior using a monocular camera is of great value. Su et al. [11] proposed a new method for capturing multi-layer human behavior using a monocular camera. This method combines computer vision, image processing, and computer graphics techniques to extract human motion information from images captured by a monocular camera. Firstly, we use a deep learning-based image segmentation algorithm to separate the human body part from the background part in the image. Then, we identify different body parts such as the head, hands, feet, etc. by analyzing the shape and movements of the human body. Next, we use a kinematic based model to convert the identified movements of body parts into data on human behavior. Finally, we use computer graphics methods to visualize the captured human behavior data into a 3D model. With the continuous development of computer technology and digital art, the application of computer three-dimensional assistance systems in the field of dance design is becoming increasingly widespread. This technology can simulate real dance movements, allowing designers to design and rehearse dance movements in a three-dimensional environment, thereby better understanding and mastering the details of dance movements. Tan and Yang [12] explored a dance action design method based on computer three-dimensional assistance systems. Using motion capture technology, capture the movements of real dancers into a computer and generate motion data for a 3D model. Designers can design dance movements based on this data, or import this data into a virtual environment for dance rehearsals. By using a computer-aided 3D system, designers can create 3D models and perform detailed design and rehearsal of dance movements. This technology can simulate real dance scenes, allowing designers to better grasp the details and rhythm of dance movements. Based on a computer-aided 3D system, designers can generate dance animations by adjusting the motion parameters of the 3D model. This technology can easily achieve repetitive, variable speed, playback and other operations of dance movements, enabling designers to better creatively design dance movements. By combining virtual reality technology, designers can simulate and demonstrate dance movements in a virtual environment, enabling the audience to better feel the atmosphere and emotions of the dance. Visual SLAM, as an important branch of it, plays an increasingly important role in intelligent transportation systems due to its ability to perceive and understand the environment. However, there is often a neglected issue in the visual SLAM process - the issue of re access. Tsintotas et al. [13] delved into this issue, analyzed its causes, and proposed corresponding solutions. Visual SLAM is a technology that simultaneously locates and navigates by constructing environmental maps. In the visual SLAM process, the system infers its own position and posture by analyzing image information obtained from the environment, and constructs a map containing its own position and posture information. However, when the system retrieves images from locations that have already been accessed, a so-called "re access problem" occurs. Re-access issues refer to the possibility that during the visual SLAM process, the system may mistakenly assume that it has accessed a new location, when in reality it is a previously accessed location. This

© 2024 U-turn Press LLC, http://www.cad-journal.net

error can lead to inaccurate maps and navigation errors in the system. The main reason for the re access problem is feature matching errors. When the system moves in the environment, it will see the same features from different perspectives and mistakenly match them.

Target recognition is one of the important research topics in the field of computer vision, which has broad application prospects in many practical applications, such as autonomous driving, intelligent monitoring, facial recognition, etc. In order to solve these problems, researchers have proposed many different methods, among which the neural morphological system is a highly concerned method. Yang et al. [14] introduced a high-speed target recognition method based on neural morphology systems. Neuromorphic systems are computational models that simulate the structure of biological neural networks, consisting of multiple neurons, each receiving input signals and generating output signals. Unlike traditional computer programs, neural morphological systems do not require explicit program control, but rather rely on interactions between neurons to achieve complex computational tasks. Due to its unique computational approach, neural morphological systems have high efficiency and flexibility in processing complex data such as images. To verify the feasibility and effectiveness of object recognition methods based on neural morphology systems, we conducted experiments on multiple datasets, including pedestrian detection, facial recognition, etc. The experimental results show that the target recognition method based on neural morphology systems has high accuracy and robustness when dealing with target recognition tasks in complex scenes. Compared with traditional computer vision methods, object recognition methods based on neural morphology systems have significant advantages in processing speed and effectiveness. With the rapid development of computer vision and deep learning technology, automatic recognition of dance movements has become a hot research field. Dance action recognition involves extracting features from videos or images, and then classifying and recognizing these features. Zhai [15] explored dance action recognition methods based on feature representation and attribute mining. Time characteristics describe the changes in dance movements over time, such as rhythm, speed, acceleration, etc. Time characteristics can be obtained by analyzing video or image sequences. Color features describe the color distribution of dance movements in an image, such as color histograms, local color features, etc. Color features can provide visual clues to dance movements and help identify different types of dance movements. This method extracts texture features of the image by calculating the local binary mode histogram, which is suitable for describing texture information such as clothing and background of dance movements. By calculating the directional gradient histogram of the image, the edge and texture information of the image is extracted, which is suitable for describing edge information such as posture and gestures of dance movements. By utilizing computer vision technology to extract the features of dance movements and utilizing parallel computing technology to improve the processing speed and efficiency of the system, this system can achieve automatic recognition and classification of sports dance movements. In order to verify the effectiveness and reliability of this system, we conducted a series of experiments. By classifying and recognizing features, each dance action is segmented from a video or image sequence. Learn and train a large number of dance movements, and establish a database containing multiple dance movements. Match the segmented dance actions with the actions in the action library to find the most similar action. The experimental results show that the automatic recognition system for sports dance movements based on computer vision and parallel computing has high accuracy and processing speed, which can meet the needs of practical applications. The specific experimental results are as follows: By processing and analyzing the test dataset, the average accuracy of this system has reached over 90%. In the experiment, this system can process a large amount of dance video or image data in a relatively short time, with an average processing speed of over 20 frames per second.

3 DATA PREPROCESSING AND FEATURE EXTRACTION

3.1 Data Collection and Collation

In the task of identifying and classifying dance artistic movements, data collection and arrangement is the first step. This step usually includes the following aspects:

© 2024 U-turn Press LLC, http://www.cad-journal.net

Data source: This article collects data from various public dance data sets, dance competition videos, online dance tutorials and other sources. These data should include dance actions with different styles, difficulties and perspectives.

Data annotation: For each dance video clip, professional dancers are required to annotate it, including the type of dance action, start and end time, etc.

Data collation: Collate the collected and marked data to form a standard data set suitable for ML training, including training set, verification set and test set.

3.2 Data Preprocessing Technology

Because the original data may have some problems such as noise and redundancy, it is needed to preprocess the data so that the ML model can learn the data characteristics better. Common technologies for data preprocessing include:

Data cleaning: Remove noise data, such as blurred, dim light or video clips with poor shooting angle.

Data normalization: Normalize the data features of dance actions to eliminate the differences in dimensions and values between features.

Data enhancement: By rotating, flipping and cutting the original data, the diversity of data is increased and the generalization ability of the model is improved.

3.3 Feature Detection Method

Feature detection is a key step in ML, which directly affects the performance of the model. In the identification and classification of dance art movements, the following features are usually extracted:

Morphological characteristics: Including the dancer's body posture, movement range, limb angle, etc., which can be obtained by computer vision technology.

Kinematic characteristics: Including the speed, acceleration and trajectory of the dancer, which can be obtained through the analysis of continuous frames.

Time series characteristics: Dance actions are a time series process, and it is needed to extract the time series relationship between movements, for example, by RNN. Effective feature detection methods can greatly improve the accuracy of dance art action identification and classification.

4 DESIGN OF DANCE ART ACTION CLASSIFICATION ALGORITHM BASED ON ML

In order to realize the classification algorithm of dance art movements, it is needed to collect a large quantity of dance action data at first. The collected data includes different kinds of dance actions, and each movement has a clear label or category. Moreover, in the data preparation stage, the original data needs to be processed in order to be input into the DL model. In order to realize the classification of dance artistic movements, it is needed to construct a DL model. This model can be CNN, RNN or LSTM (Long and Short Term Memory Network). In this study, we choose CNN as the basic model and improve it. First of all, the standard CNN is very effective in processing image data, but it may ignore the information in the time dimension when processing temporal and spatial data such as dance actions. Therefore, this article adds a spatio-temporal convolution layer to CNN to capture the characteristics of both spatial and temporal dimensions. Secondly, the dance action is not only related to the current action frame, but also related to the previous and subsequent action frames. In order to better capture this timing dependency, this article adds a bidirectional LSTM layer at the top of CNN. In this way, the model can use both past and future context information for action identification. Finally, convolution layers of different layers capture features of different scales. In order to make better use of these multi-scale features, this article adopts a feature pyramid network structure to fuse the features of different layers, thus providing more information for classification. By integrating the above improvement strategies, the CNN model in this article can achieve better performance in the task of dance art action classification. These improved strategies are aimed at

© 2024 U-turn Press LLC, http://www.cad-journal.net

enhancing the model's ability to extract features of dance actions and optimizing the classification decision, so as to achieve more accurate and reliable classification and identification of dance artistic movements. Assume that the equation of the full connection layer feature output x^{l} of the l layer is as follows:

$$x^{l}=f\left(w^{l}x^{l-1}+b^{l}\right) \tag{1}$$

Where w^{l} represents the weight parameter and b^{l} is the offset term. The functions to be learned are:

$$h_{w}\left(\vec{x}\right)=\frac{1}{\sum_{i=1}^{k}e^{\vec{w}_{i}\cdot\vec{x}+b_{i}}}\begin{bmatrix}\vec{w}_{1}\cdot\vec{x}+b_{1}\\ \vec{w}_{2}\cdot\vec{x}+b_{2}\\ \ldots\\ \vec{w}_{k}\cdot\vec{x}+b_{k}\end{bmatrix} \tag{2}$$

Among them, k is the quantity of categories to be classified, and b_{i}, and $\vec{w}_{i}$ represent the offset vector and weight vector corresponding to the i category. Sample $\vec{x}$ is the probability value of j class, and the equation is as follows:

$$P\left(y=j\middle|\vec{x}\right)=\frac{\vec{w}_{j}\cdot\vec{x}+b_{j}}{\sum_{i=1}^{k}e^{\vec{w}_{i}\cdot\vec{x}+b_{i}}}\qquad\sum_{j=1}^{k}P\left(y=j\middle|\vec{x}\right)=1 \tag{3}$$

After training and learning, $\vec{w}_{i}$ and b_{i} are obtained, and the target loss function can be expressed as:

$$J\ w,b\ =-\frac{1}{m}\sum_{j=1}^{m}\sum_{l=1}^{k}1\ y^{j}\ =l\ \log\frac{e^{\vec{w}_{l}\cdot\vec{x}+b_{l}}}{\sum_{i=1}^{k}e^{\vec{w}_{i}\cdot\vec{x}+b_{i}}} \tag{4}$$

Among them, m is the quantity of samples in the dance training set, k is the quantity of dance classification categories, and $1\{\cdot\}$ is the indicative function. Figure 1 shows the network structure of this model.

CNN can automatically learn the spatial features of dance actions, and through the stacking of convolution layer and pooling layer, it can gradually abstract advanced features. After constructing CNN model, we need to use the collected dance action data to train the model. During the training stage, the model will learn to extract features from the input data and classify dance actions according to these features. In order to optimize the performance of the model, this article will adopt some techniques and methods, such as data enhancement, increasing the diversity of data sets by flipping and rotating, and preventing over-fitting; Cross validation is used to assess the generalization ability of the model; And adjust the super parameters of the model, such as learning rate, batch size and iteration times, to find the optimal model configuration.

Moreover, the model introduces attention mechanism, which makes the model pay attention to the key parts of dance actions. For example, some dance actions may have specific gestures or steps that are more critical. Through the attention mechanism, the model can automatically learn which time and space to focus on. In this article, the data are normalized.

© 2024 U-turn Press LLC, http://www.cad-journal.net

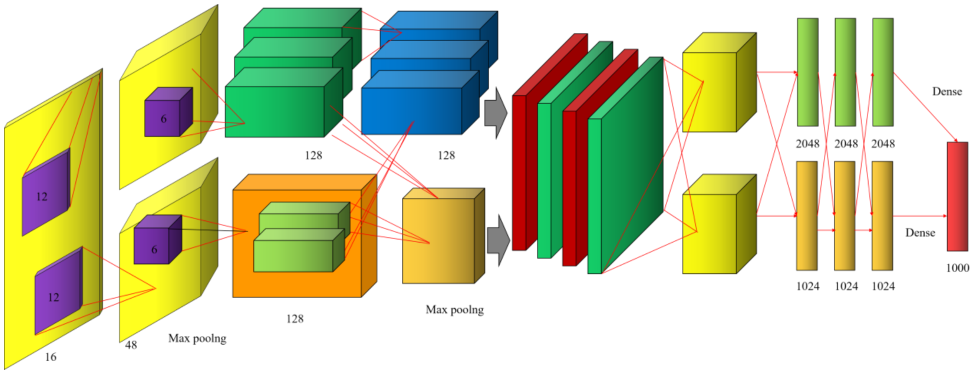

Figure 1: Network structure of this model.

The normalization equation is as follows:

$$x_i = \lambda_1 + (\lambda_2 - \lambda_1)\left(\frac{z_i - z_i^{\min}}{z_i^{\max} - z_i^{\min}}\right) \tag{5}$$

The output value is denormalized, and the denormalization equation is shown in the equation:

$$z_i = \left(\frac{x_i - \lambda_1}{\lambda_2 - \lambda_1}\right)(z_i^{\max} - z_i^{\min}) + z_i^{\min} \tag{6}$$

Where x_i is a normalized value; z_i is a denormalized value; λ_1 is the lower limit; λ_2 is the upper limit; $z_i^{\max}$ is the maximum value in the source data; $z_i^{\min}$ is the minimum value in the source data. After training and optimization, the trained DL model can be used to classify new dance actions.

5 DESIGN OF DANCE ART ACTION IDENTIFICATION ALGORITHM BASED ON ML

In the computer-aided dance art movement identification based on ML, this section proposes a specific dance art movement identification algorithm. This algorithm combines the advantages of feature detection and ML classifier to realize accurate and efficient dance action identification. First, we need to collect the data of dance actions. This can be collected by video recording, sensors and other equipment, and the digital representation of dance actions can be obtained. Then, these raw data are preprocessed, such as denoising and standardization, to prepare for subsequent feature detection and classification. Feature detection is a key step in dance action identification. For example, the dancer's body contour, joint angle change, motion trajectory and other features can be extracted. These features can describe the shape, dynamics and rhythm of dance actions. Next, we need to construct an ML classifier to classify and recognize the extracted features. In this algorithm, we can choose a suitable ML algorithm for training and classification, such as SVM, RF or DL model. This article adopts SVM algorithm. Specifically, for a given training sample set, SVM aims to find a hyperplane so that positive and negative samples can be separated best. This hyperplane has the greatest distance from the nearest points of positive and negative samples, which are called support vectors. For any symmetric function $K(x,x')$, it is an inner product operation in a characteristic space if and only if, for any $\Phi(x) \neq 0$ and $\int \Phi^2(x)\,dx < \infty$, the following equation holds:

$$\int\int K(X,X')\,\Phi(x)\,\Phi(x')\,dx\,dx' > 0 \tag{7}$$

Firstly, this article extracts meaningful features from dance action data. These features can be based on image, such as shape, texture, color, etc., and can also be based on motion, such as joint angle,

© 2024 U-turn Press LLC, http://www.cad-journal.net

motion trajectory, etc. For dance actions, we can consider extracting dynamic features such as the position, speed and acceleration of body parts, as well as static features such as the shape and angle of the dancer's limbs. Secondly, the extracted features are combined with the corresponding dance action tags to form a training set and a test set. Then, a classifier is trained by SVM algorithm using the training set. In this process, SVM will try to find an optimal hyperplane, so that the dance actions in the training set can be correctly classified according to their characteristics. In the training stage, the radial basis kernel function is selected and the related parameters, such as penalty coefficient C and kernel function parameters, are adjusted. Let $\alpha^* = \alpha_1^*,...,\alpha_l^{*\,T}$ be any solution to the problem of dance classification and identification, and choose a positive component of α^* to get the normal vector w^* and classification threshold b^* of the optimal classification hyperplane:

$$w^* = \sum_i \alpha_i^* y_i x_i - D\sum_i T_i * G_i \tag{8}$$

$$b^* = y_i - \sum_i y_i \alpha_i^* \; x_i \cdot x_j \; + D\sum_i x_j \cdot \; T_i * G_i \tag{9}$$

Thus, for any given unknown class sample x, its discriminant function is:

$$f \; x \; = \sum_i \alpha_i^* y_i \; x_i \cdot x \; + b^* \tag{10}$$

In order to show the optimization process of skeleton diagram topology in the training stage more intuitively, the skeleton topology diagram learned by a dance action sample during training is selected for visualization, as shown in Figure 2. From left to right, the three subgraphs respectively represent the skeleton topology diagram of the first layer, the fourth layer and the seventh layer in Figure 2.

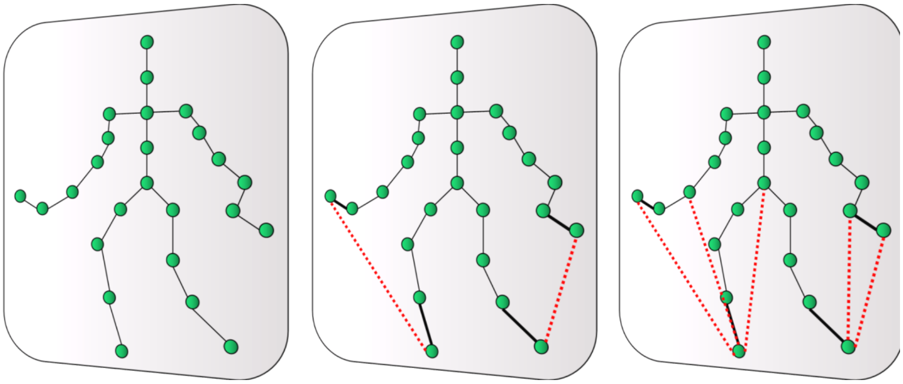

Figure 2: Skeleton topology diagram.

For the training of classifier, this section uses the marked dance action data as the training set, and optimizes the algorithm to train the parameters of classifier to minimize the classification error. Once the SVM classifier is trained, it can be used to identify new dance actions. Specifically, for a new dance action, the corresponding features are first extracted, and then these features are input into the trained SVM classifier for prediction. The SVM classifier will classify the dance action into a certain category according to the learned decision boundary.

6 SIMULATION EXPERIMENT AND RESULT ANALYSIS

In order to verify the performance of the designed dance art movement identification and classification algorithm, this section carries out simulation experiments. The experimental

© 2024 U-turn Press LLC, http://www.cad-journal.net

environment includes hardware equipment and software environment, such as computer processor, memory, operating system, etc. The data set adopts a large-scale dance action data set, which contains a variety of dance types and actions, and each action has a corresponding label. The data set is divided into training set and test set to ensure the accurate assessment of the algorithm performance. According to the experience and the characteristics of the data set, the key parameters in the algorithm are set. These parameters include learning rate, batch size, quantity of training rounds, etc. By adjusting these parameters, the performance of the algorithm can be optimized. The experimental design includes comparing the performance of different algorithms, including traditional ML-based algorithm and decision tree algorithm.

Figure 3 shows the precision result of the algorithm. Figure 4 shows the recall result of the algorithm. Figure 5 shows the F1 result of the algorithm.

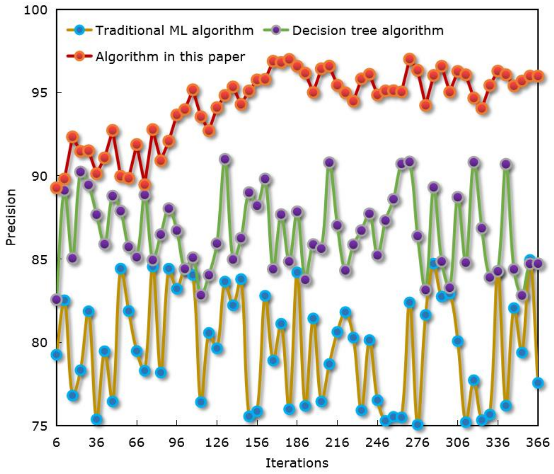

Figure 3: Precision result of the algorithm.

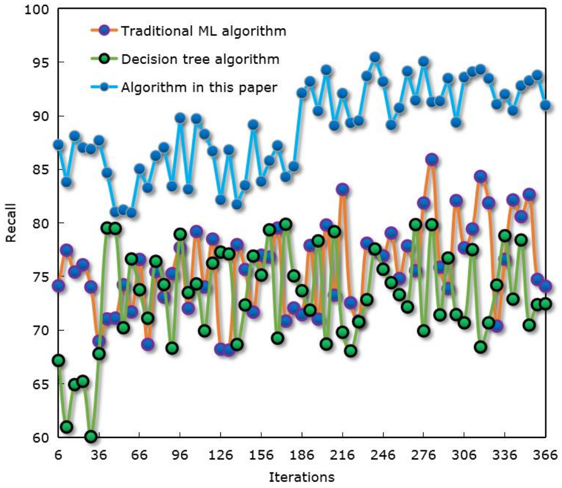

Figure 4: Recall result of algorithm.

The Precision, Recall and F1 scores of the algorithm in this article all show high results, basically reaching more than 92%. This means that the algorithm in this article has good performance in the task of dance art action identification and classification.

© 2024 U-turn Press LLC, http://www.cad-journal.net

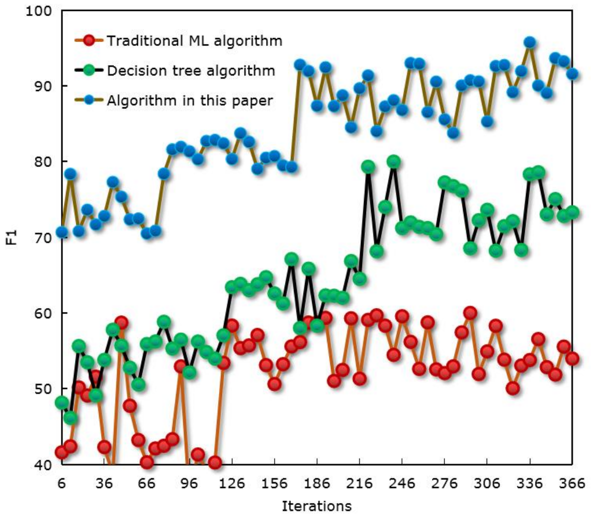

Figure 5: F1 result of the algorithm.

Among them, the high accuracy of the algorithm in this article shows that it can accurately identify dance actions with less misjudgment; The high recall rate of the algorithm shows that it can cover all the real dance actions well and classify them correctly. The high F1 score of the algorithm further verifies its superiority in the task of dance art movement identification and classification.

To sum up, through comparative analysis, the algorithm in this article shows superior performance in the task of dance art action identification and classification. This is due to the innovations in the algorithm, such as multi-modal feature fusion and the combination of DL and SVM, which effectively improve the accuracy and recall rate of the algorithm, and then improve the overall performance.

In order to measure the accuracy of the algorithm, RMSE (Root Mean Square Error) and Mean Absolute Error (MAE) are selected as assessment indexes. RMSE is sensitive to outliers, while MAE can better reflect the average error of predicted values. Through 10% cross-validation, the RMSE and MAE experimental results in Table 1 and Table 2 are obtained.

Training times	*RNN model*	*Decision tree model*	*This article model*
1	0.62	0.662	0.588
2	0.646	0.698	0.505
3	0.63	0.695	0.511
4	0.67	0.644	0.524
5	0.625	0.699	0.518
6	0.659	0.692	0.532
7	0.613	0.691	0.539
8	0.62	0.672	0.534
9	0.606	0.697	0.515
10	0.608	0.646	0.503
Average value	0.628	0.680	0.527

Table 1: RMSE results of the algorithm.

The algorithm in this article has achieved a low RMSE value in each fold, which indicates that the error between the prediction result of the algorithm and the real value is small. The low value of RMSE also shows that the algorithm has high accuracy for the identification and classification of dance actions.

© 2024 U-turn Press LLC, http://www.cad-journal.net

Training times	*RNN model*	*Decision tree model*	*This article model*
1	2.13	2.89	1.44
2	1.9	2.26	1.63
3	2.11	2.14	1.46
4	2.18	2.3	1.21
5	1.95	2.02	1.43
6	1.93	2.06	1.26
7	2.29	2.12	1.38
8	1.95	2.07	1.48
9	2.07	2.06	1.37
10	2.18	2.19	1.34
Average value	2.069	2.211	1.400

Table 2: MAE results of the algorithm.

Table 2 shows the MAE experimental results of the algorithm. The MAE value of the algorithm in this article is also relatively low on each fold. This means that the average prediction error of the algorithm for dance actions is small, which further confirms the accuracy of the algorithm in the task of dance art movement identification and classification. On the whole, the experimental results of 10% cross-validation show that the algorithm in this article shows high accuracy and stability in identifying and classifying dance actions. This is due to the effective feature detectionmethod and classification strategy adopted in the algorithm.

Figure 6 shows the ROC curve of the algorithm. ROC curve describes the relationship between TPR (True Positive Rate) and FPR (False Positive Rate). When drawing ROC curve, it is needed to choose the appropriate threshold to determine the classification of dance actions. In this article, the points of ROC curve are obtained by evaluating TPR and FPR under different thresholds on the verification set.

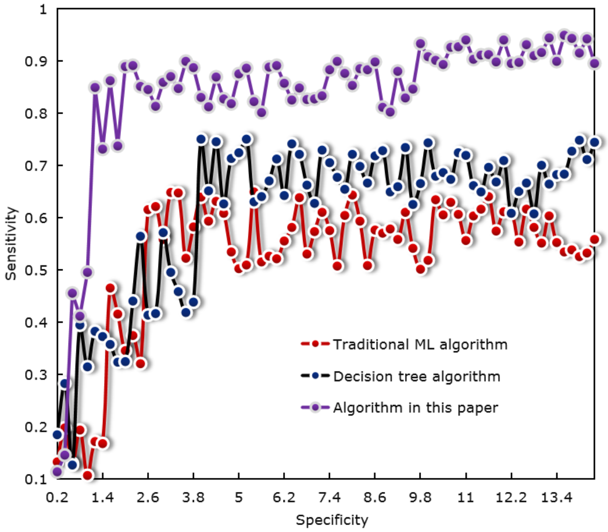

Figure 6: ROC curve of the algorithm.

By observing the ROC curve in Figure 6, we can see that the ROC curve of this model presents an ideal shape, close to the upper left corner. The dance art movement identification and classification algorithm proposed in this article has achieved ideal TPR and low FPR in the experiment. This shows

© 2024 U-turn Press LLC, http://www.cad-journal.net

that the model can accurately identify the real dance actions and reduce the false alarm rate, which further verifies the superiority and effectiveness of the model. This good performance can provide more possibilities for the research, teaching, creation and appreciation of dance art.

The following experiment uses a data set containing the information of dancers' joints. The data set obtains the coordinate data of the dancer's joints by motion capture technology, and carries out preprocessing, such as noise removal and standardization, to ensure the quality of the data. In the experiment, the coordinate information of the dancer's joint points is extracted and expressed as an action sequence. Each action sequence contains a series of joint position information, which is used to describe the dancer's artistic actions. The appropriate loss function and optimization algorithm are used to train the model, and the best performance is obtained by adjusting the hyperparameters. As shown in Figure 7, the dancer's joint action identification is indicated.

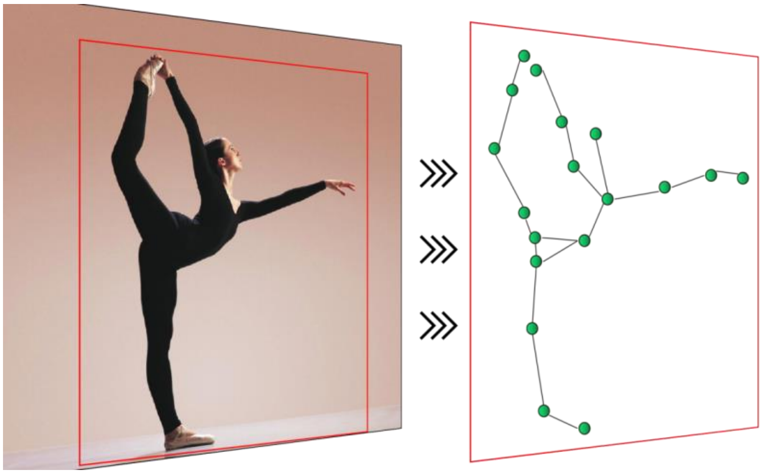

Figure 7: Motion identification of dancer's joint points.

By observing the motion identification of the dancer's joints in Figure 7, it can be concluded that the method proposed in this article can well recognize the dancer's artistic movements. Figure 7 shows the input of the model to the action sequence of the dancer's joint points, and outputs the corresponding dance action labels. By comparing the input action sequence with the output action tag, we can see that the model accurately identifies the dancer's actions. Through the experiment of joint motion identification, the effectiveness and superiority of the method proposed in this article in the field of dance art motion identification are further verified.

7 CONCLUSIONS

In this article, ML is applied to computer-aided dance art action identification and classification, and an action identification and classification algorithm based on CNN and SVM is proposed. Through the research of this article, ML technology has been successfully applied to the identification and classification of dance artistic movements, and a series of remarkable research results have been achieved. In this article, the algorithm is verified and tested on a large-scale dance data set, which is closer to the actual application scenario and proves the effectiveness of the algorithm in practical application. The proposed algorithm shows high accuracy, high recall and high F1 score in the experiment, which fully verifies its superiority and effectiveness. Moreover, through the analysis of ROC curve, the good performance of this algorithm in dance action classification task is further proved.

© 2024 U-turn Press LLC, http://www.cad-journal.net

These research results play an important role in promoting the field of dance art. First of all, the algorithm in this article can provide more accurate and objective tools for dance research and teaching, and help dancers to better analyze and understand dance actions. Secondly, through computer-aided dance action identification and classification, the algorithm in this article is helpful to improve the popularity and accessibility of dance art, so that more people can appreciate and experience the charm of dance. Finally, the research of this article also promotes the application of ML technology in the fields of art and humanities, and opens up new possibilities for interdisciplinary research. It is believed that in the future research, with the continuous progress and innovation of technology, the movement identification and classification of dance art will be more accurate, intelligent and expressive, which will bring more possibilities for the research, teaching, creation and appreciation of dance art.

Xu Han, https://orcid.org/0009-0003-3653-2906
Sifan Zhang, https://orcid.org/0009-0001-6932-8536

REFERENCES

[1] Barkoky, A.; Charkari, N.-M.: Complex network-based features extraction in RGB-D human action recognition, Journal of Visual Communication & Image Representation, 82(1), 2022, 103371. https://doi.org/10.1016/j.jvcir.2021.103371
[2] Chen, C.; Alotaibi, F.-S.; Omer, R.-E.: 3D mathematical modelling technology in visual rehearsal system of sports dance, Applied Mathematics and Nonlinear Sciences, 7(1), 2022, 113-122. https://doi.org/10.2478/amns.2021.2.00078
[3] Fan, C.; Gao, F.: Enhanced human activity recognition using wearable sensors via a hybrid feature selection method, Sensors, 21(19), 2021, 6434. https://doi.org/10.3390/s21196434
[4] Faridee, A.-Z.; Ramamurthy, S.-R.; Roy, N.: Happyfeet: Challenges in building an automated dance recognition and assessment tool, GetMobile: Mobile Computing and Communications, 22(3), 2019, 10-16. https://doi.org/10.1145/3308755.3308759
[5] Hsueh, S.; Alaoui, S.-F.; Mackay, W.-E.: Deconstructing creativity: non-linear processes and fluid roles in contemporary music and dance, Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 2019, 1-21. https://doi.org/10.1145/3359305
[6] Lapinski, M.; Brum, M.-C.; Moxley, S.-D.; Berkson, E.; Gill, T.-J.; Kepple, T.; Paradiso, J.-A.: A wide-range, wireless wearable inertial motion sensing system for capturing fast athletic biomechanics in overhead pitching, Sensors, 19(17), 2019, 3637. https://doi.org/10.3390/s19173637
[7] Li, M.; Leung, H.: Graph-based approach for 3D human skeletal action recognition, Pattern Recognition Letters, 87(2), 2017, 195-202. https://doi.org/10.1016/j.patrec.2016.07.021
[8] Luo, W.; Ning, B.: High-dynamic dance motion recognition method based on video visual analysis, Scientific Programming, 2022(10), 2022, 1-9. https://doi.org/10.1155/2022/6724892
[9] Mazian, A.-I.; Rizhan, W.; Rahim, N.; Jamal, A.-A.; Ismail, I.; Fadzli, S.-A.: A Theoretical framework for creating folk dance motion templates using motion capture, International Journal of Advanced Computer Science and Applications, 14(5), 2023, 445-451. https://doi.org/10.14569/IJACSA.2023.0140547
[10] Pan, J.; Wang, S.; Bai, J.; Dai, J.: Diverse dance synthesis via keyframes with transformer controllers, Computer Graphics Forum: Journal of the European Association for Computer Graphics, 40(7), 2021, 71-83. https://doi.org/10.1111/cgf.14402
[11] Su, Z.; Wan, W.; Yu, T.; Liu, L.; Fang, L.; Wang, W.; Liu, Y.: Mulaycap: Multi-layer human performance capture using a monocular video camera, IEEE Transactions on Visualization and Computer Graphics, 28(4), 2020, 1862-1879. https://doi.org/10.1109/TVCG.2020.3027763
[12] Tan, B.; Yang, F.: Dance movement design based on computer three-dimensional auxiliary system, Computer-Aided Design and Applications, 19(S3), 2021, 108-118. https://doi.org/10.14733/cadaps.2022.S3.108-118

© 2024 U-turn Press LLC, http://www.cad-journal.net

[13] Tsintotas, K.-A.; Bampis, L.; Gasteratos, A.: The revisiting problem in simultaneous localization and mapping: A survey on visual loop closure detection, IEEE Transactions on Intelligent Transportation Systems, 23(11), 2022, 19929-19953. https://doi.org/10.1109/TITS.2022.3175656
[14] Yang, Z.; Yang, L.; Bao, W.; Tao, L.; Zeng, Y.; Hu, D.; Shang, D.: High-speed object recognition based on a neuromorphic system, Electronics, 11(24), 2022, 4179. https://doi.org/10.3390/electronics11244179
[15] Zhai, X.: Dance movement recognition based on feature expression and attribute mining, Complexity, 2021(21), 2021, 1-12. https://doi.org/10.1155/2021/9935900

© 2024 U-turn Press LLC, http://www.cad-journal.net