ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Gender differences in infant and toddler dance rhythmic behavior: performance characteristics and computational-behavioral analysis

Nanhua Zeng[1,*], *Junfang Zheng*[1], *Ting Dong*[2] *and Haiyou Liu*[3]

[1] School of Preschool Education, Fuzhou Preschool Education College, Fuzhou, Jiangxi, 344000, China
[2] School of Arts and Sports, Fuzhou Preschool Education College, Fuzhou, Jiangxi, 344000, China
[3] Inspur Software Technology Co., Ltd, Nanchang, Jiangxi, 330000, China

SUMMARY: *In this paper, based on computational behavioral science, on the basis of the original OpenPose model, the backbone feature extraction network VGG19 was changed to ResNet18 network, the prediction network in the OpenPose model was optimized by depth-separable convolution, and a linear module was added to the nonlinear network to obtain a classification network for recognizing the basic dance rhythmic behaviors of infants and children. Then the data were collected and processed using qualisys miqus m3 three-dimensional motion capture system and Visual3D biomechanical analysis software, and the statistical analysis was completed by SAS JMP14.2. The results showed that the accuracy of the improved OpenPose network model in recognizing dance rhythmic behaviors of infants and toddlers reached more than 99%. the performance characteristics of dance rhythmic behaviors of infants and toddlers aged 2~3 years old did not differ significantly in basic rhythmic movements (P>0.1), but there were significant differences in body exploration and social interaction movements (P<0.001). Especially in the social interaction category dimension, male toddlers' completion of rotational imbalance and forward-leaning-backward movements was significantly higher than that of female toddlers. In conclusion, male and female infants and toddlers can synergistically control dance rhythmic behaviors through physiological structures, neuromotor control and behavioral patterns.*

KEYWORDS: *Improved OpenPose pose estimation algorithm; Artificial neural network; Dance rhythmic behavior; Infants and toddlers*

1 Introduction

Music is an art that is composed of melody, rhythm, beat, tune, etc. and emphasizes the expression of emotion. This strong emotionality also determines that although music can describe a certain atmosphere through melody, tune and so on, it can not be shown with clear and concrete images [1]. And dance is always accompanied by music, is symbiotic and coexist with music, and its movement can materialize the music, so that students can more intuitively feel the emotions expressed by the music, and enhance students' perception and understanding of music [2, 3]. The so-called dance rhythm is a kind of regular and rhythmic dance movement accomplished by human body movement along with music. At the same time, dance rhythm is also the basic element of dance as an art form. The difference from ordinary rhythm is that

*17779443909@163.com
https://doi.org/10.65102/is2026046 Published: 30 April, 2026

dance rhythm is a regular, rhythmic dance movement, rather than unconscious and regular blind rhythm [4].

Dance rhythm and music are interdependent in nature [5]. Dance is always displayed with the accompaniment of music, which shows beauty through sound, utilizing the combination of harmony, melody, rhythm and other elements to give people the enjoyment of beauty, and at the same time to cultivate people's ideological and moral sentiments [6-8]. And dance is an art that expresses the beauty of the human body, which relies on the changes of human body image, movement, line and posture to show a special beauty. This combination of music and dance forms a comprehensive art [9]. Secondly, dance rhythm has an important application value in music teaching. The use of dance rhythm in music teaching can change the students' previous passive learning mode, which is mainly based on technology and instilling knowledge [10-12]. The combination of music and dance rhythm can arouse students' musical instincts, activate their sensibility and imagination to music, so that they can better love music, aestheticize music and enjoy music [13].

Dance rhythm is a kind of instinct, which refers to the natural body movements of a person in the process of hearing music, and has an obvious promotion effect on students' creativity and imagination [14]. Effective integration of dance rhythm education in early education of infants and toddlers can allow infants and toddlers to feel the music, appreciate the dance, experience the emotions that the music and dance want to interpret, and integrate their own emotions, which can effectively improve the understanding of music and dance, stimulate the interest of infants and toddlers in dance and music, and promote the overall development of young children [15-17]. At the same time, infants and toddlers express themselves on the basis of their own cognition of the outside world, and if infants and toddlers lack the ability to understand things in the outside world, it will result in infants and toddlers not being able to express themselves accurately [18]. Dance rhythm can fully utilize the experience of the senses to help infants and toddlers deeply understand and master the basic knowledge, using different sounds, movements, auditory, visual and other ways to help infants and toddlers shape the cognition of music and dance from various perspectives, deepen the infants and toddlers' feelings towards music and dance, and promote the further enhancement of the infants and toddlers' comprehension ability [19-21]. However, at the present stage, the dance rhythm education enlightenment lacks professionalism, and in the dance rhythm teaching, it is mainly based on the basic knowledge of dance, without analyzing the characteristics of infants and young children or combining the interests of infants and young children, which lacks the value and role that the enlightenment should have, and has certain limitations.

In this paper, in order to analyze the performance characteristics in the dance rhythm behavior of infants and young children, a deep learning-based infant posture estimation model is constructed to estimate the posture of multiple infants and young children. In order to solve the shortcomings of the human posture estimation network OpenPose model with a large number of parameters, the VGG19 network in the original model is replaced by the ResNet18 network, which has a smaller number of parameters and higher accuracy; and then some of the convolution kernels in the prediction network are replaced by the depth-separable convolution to achieve the improvement of the feature extraction network and prediction network of the OpenPose model and to complete the model's lightweighting. Then the memory and generalization ability of the network is improved by adding a linear module to the traditional nonlinear network via artificial neural network. Then the kinematic and kinetic data of infants and toddlers in the process of dance rhythm were synchronously collected by the qualisys miqus m3 three-dimensional motion capture system and the bertecfully instrumented treadmills three-dimensional force measurement running platform, and then the data were processed by Visual3D biomechanical analysis software and Excel to realize the analysis of the dance rhythm

behavior of infants and toddlers. The computational-behavioral analysis of infants' dance rhythmic behavior was realized by combining the Visual3D biomechanical analysis software and Excel to process the data.

2 OpenPose-based behavioral recognition model for infant and toddler dance rhythms

2.1 Subjects of study

A total of 46 infants and toddlers were recruited through three channels: (1) infants and toddlers in the maternity and early childhood care departments of public hospitals, (2) members of community early childhood education centers, and (3) family members of volunteers of online parenting platforms, of whom 24 and 22 were male and female, respectively.

2.2 Deep learning based model for infant posture estimation

2.2.1 Human structure modeling

In the study of computer vision-based posture estimation for infants and toddlers, it is first necessary to select a suitable model to represent the action kinematics structure and body shape information. In this paper, we use the skeleton model, which can describe the structure of infants and young children using only the combination of points and line segments, which is simple and intuitive, with a relatively small amount of computation, and is suitable for the subsequent analysis of action posture.

The skeleton model, where the infant consists of skeletal joint point locations and connecting components, simplifies the infant skeleton, is flexible and intuitive in its representation, and is widely used in studies of 2D and 3D pose estimation. However, it has limitations in representing texture information and shape information.

For the planar model, the body parts are represented by rectangles that approximate the contours of the infant. The limbs and torso are represented as rectangles of different sizes. Further analysis was done through infant graphic silhouettes and contours. Stereoscopic models, which add depth and information about the infant's surface, can provide richer information for human-computer interaction, and the models are more complex.

2.2.2 OpenPose pose estimation algorithm

OpenPose is a deep neural network based algorithm for real-time multiplayer pose estimation, which employs the architecture of a two-branch multi-stage CNN, and Partial Affinity Fields (PAFs) are used to predict the position and orientation of the limbs.OpenPose can process inputs from live cameras, recorded videos, still images, and other kinds of inputs, and represent the results by skeletal keypoints that are connected into a human skeleton based on the body's physical structure connected to the human skeleton.

The OpenPose network employs a multi-stage inference structure, where each stage consists of two parallel convolutional network branches that predict the skeletal joint point confidence S^t and the partial affinity domain field L^t in the image as well as the input image features F extracted from the VGG-19, respectively. The S^t and L^t are computed as follows:

$$S^t = \rho^t\left(F, L^t, S^{t-1}\right), \forall\, t \geq 2 \tag{1}$$

$$L^t = \phi^t\left(F, L^{t-1}\right), \forall\, t \geq 2 \tag{2}$$

The ρ^t and ϕ^t are obtained from the CNN inference in the t stage, and together with the feature F are passed to the next stage for inference again to improve the prediction accuracy.

Confidence is the magnitude of the probability of predicting the location of the joints, and the affinity domain is the flow field that detects the vectorization of the pixels of the limb parts in the image region with directionality. The difference between the predicted and true values is measured by adding an L2 loss function at the end of the branching network stage. In practice, not all human joints can be recognized, so a spatially weighted loss function is used to measure the error, and the loss function at the two branches of stage t is:

$$f_s^t = \sum_{j=1}^{J}\sum_{p} W(p)\cdot\left\|S_j^t(p) - S_j^*(p)\right\|_2^2 \tag{3}$$

$$f_l^t = \sum_{c=1}^{C}\sum_{p} W(p)\cdot\left\|L_c^t(p) - L_c^*(p)\right\|_2^2 \tag{4}$$

$S_j^*(p)$ is the labeled true confidence, $L_c^*(p)$ is the labeled true affinity domain field, and $W(p)$ is the binary mask for when the pixel p is not detected, i.e. $W(p)=0$. is used to avoid undetected sites from adversely affecting the detected true predicted values. The global objective function is used to periodically supplement the gradient loss, i.e.:

$$f = \sum_{t=1}^{T}\left(f_l^t + f_s^t\right) \tag{5}$$

The j th joint of the i th person may generate multiple candidate joint positions at the same location in the inference stage, and the joint with maximum confidence is used as the peak of the joint. Finally, the partial affinity domain field is parsed, and the joints sharing the same body part are used to compose the full-body pose of multiple individuals using the Hungarian matching algorithm and the greedy algorithm.

2.3 Lightweight Infant Behavior Recognition Model Based on OpenPose Improvement

2.3.1 Feature Extraction Network

The original OpenPose model uses a VGG19 convolutional neural network for feature extraction. Convolutional neural network with too shallow depth is not effective for feature extraction, but for networks without residual structure such as VGG, with the increase of depth, it will affect the convergence speed of the network, instead of affecting the detection effect. Therefore, in this paper, the original VGG19 structure is changed to ResNet18 structure.

The ResNet18 network is divided into four residual layers for feature extraction from different scales. Each layer consists of two BasicBlock residual blocks, while the depth of the convolutional neural network is deepened, two shortcuts are created by using jump-level connections, and the ReLU activation function is added between the convolutional kernels to ensure the nonlinearity of the network, and the features of the main and convolutional paths are

continuously fused during the training process.

2.3.2 Feeder networks

In this paper, we adopt the idea of multiple small convolution kernels instead of large convolution kernels, and denote the size of the input feature map as $D_i \times D_i \times M$, the size of the convolution kernel as $D_c \times D_c \times M$, and the size of the output feature map as $D_o \times D_o \times N$. For the traditional convolution operation, each convolution kernel is convolved corresponding to each layer of the input feature map, then:

$$W_c = (D_c \times D_c \times M) \times N \tag{6}$$

Whereas for deep convolution, where multiple layers of convolution kernels are convolved separately on the input features to obtain a new feature map with the same number of input features, there is:

$$W_{dw} = (D_c \times D_c \times 1) \times M \tag{7}$$

Dotwise convolution, on the other hand, uses a certain amount of 1×1 convolution kernels on top of the previous step to fuse the intermediate features and end up with the input feature map, i.e:

$$W_{pw} = (1 \times 1 \times M) \times N \tag{8}$$

To summarize, the ratio of the total number of parameters of the conventional convolution operation and the depth-separable convolution is:

$$\mu = \frac{1}{N} + \frac{1}{D_c^2} \tag{9}$$

Since the conventional convolutional kernels being replaced are all of size 3×3 , the convolutional operations of the two-branch prediction structure will be reduced to $1/9$ of the original.

2.3.3 Algorithms for categorizing infant and toddler movements

Infant action form can be understood as the change of posture per unit of time, so the infant action classification algorithm in this paper mainly focuses on the 2D coordinate plane position change of the identified infant keypoints within a certain period of time.

The 18 keypoints of the human body output by OpenPose network are shown in Figure 1. If the position changes of all 18 key points of infants and young children recognized are all used as input data, many redundant inputs will be generated to affect the network running speed and will directly affect the classification effect. Therefore, in this paper, some of the keypoint data of the head and arm are discarded, so that the network focuses more on learning the position change information of the keypoints of the body parts.

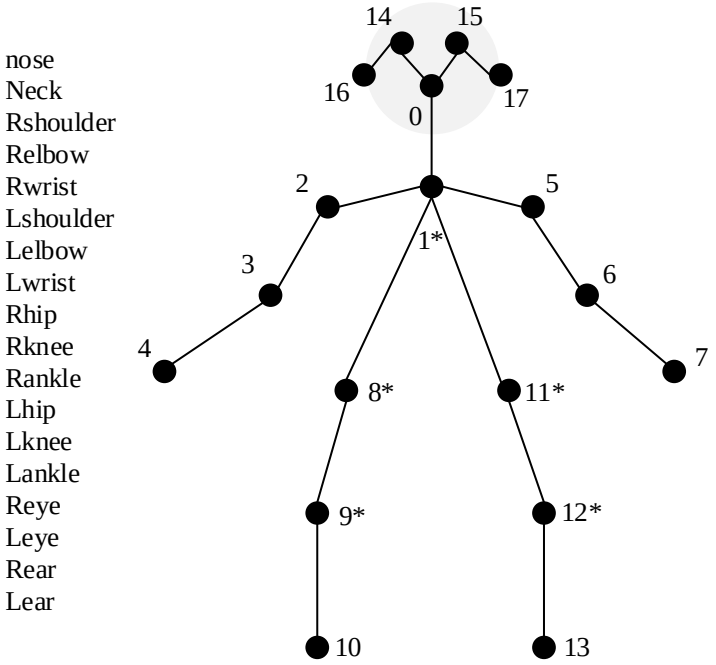


Figure 1: 18 Key Points of Human Body Output by OpenPose Network

2.3.4 Artificial neural networks

Artificial neural networks are also known as deep neural networks (DNN). A linear module is added to the nonlinear system of the traditional network, and the classification network trained by such an improvement possesses the memory capacity of the linear module and the generalization capacity of the nonlinear module with a certain trade-off. The final output is Softmax normalized exponential function to get the probability of 12 basic action states.

In this paper, three hidden layers are set up for nonlinear separation, and according to the general rule of the number of nodes in the hidden layer, the three hidden layers are set up as 30, 20 and 15 neurons respectively, in which there are independent weight values ω and bias values b on each neuron, and the transmission relationship between neurons is:

$$y_j^{h+1} = \sum_i \omega_i^h x_i^h + b_j^{h+1} \tag{10}$$

where: y_j^{h+1} denotes the output of the j neuron in the $h+1$th layer; x_i^h denotes the input of the i neuron in the hth layer; ω_i^h denotes the value of the i neuron weights in the hth layer; and b_j^{h+1} denotes the bias of the j neuron in the $h+1$th layer.

The linear module is able to learn the commonality of features and is learning specific features. The fusion information of three key points, namely, neck, right hip and left hip, can represent the posture change of the upper part of the human body; the fusion information of right knee, left knee, right hip and left hip can represent the posture change of the lower part of the human body. In this way, five types of fusion input parameters, namely, the horizontal coordinate of the center of the upper part U_x, the vertical coordinate of the center of the upper part U_y, the height of the center of the upper part U_h, the height of the center of the lower part L_h, and the composite factor CF, which contain the advancement information among the important key points and are more suitable to be used as the input parameters of the linear

module, are established in the following way:

$$U_x = \frac{\Delta_x[neck] + \Delta_x[Rhip] + \Delta_x[Lhip]}{3} \tag{11}$$

$$U_y = \frac{\Delta_y[neck] + \Delta_y[Rhip] + \Delta_y[Lhip]}{3} \tag{12}$$

$$U_h = \frac{\overline{[Rhip]_y^{t_2} + [Lhip]_y^{t_2}} - [neck]_y^{t_2}}{\overline{[Rhip]_y^{t_1} + [Lhip]_y^{t_1}} - [neck]_y^{t_1}} \tag{13}$$

$$L_h = \frac{\overline{[Rknee]_y^{t_2} + [Lknee]_y^{t_2}} - \overline{[Rhip]_y^{t_2} + [Lhip]_y^{t_2}}}{\overline{[Rknee]_y^{t_1} + [Lknee]_y^{t_1}} - \overline{[Rhip]_y^{t_1} + [Lhip]_y^{t_1}}} \tag{14}$$

$$CF = \frac{\overline{[Rhip]_y^{t_2} + [Lhip]_y^{t_2}} - [neck]_y^{t_2}}{\overline{[Rknee]_y^{t_2} + [Lknee]_y^{t_2}} - \overline{[Rhip]_y^{t_2} + [Lhip]_y^{t_2}}} \tag{15}$$

where $[key]$ denotes the name of the key point; Δ_x denotes the x-axis position change of the key point per unit of time; Δ_y denotes the y-axis position change of the key point per unit of time; $\overline{[key_1]_y^{t_1} + [key_2]_y^{t_1}}$ is the average of the y-axis coordinates of the key point key_1 and the key point key_2 at the t_1 moment; $[key]_y^{t_2}$ is the y-axis coordinate of the key point key at the t_2 moment.

The transmission relationship between the neurons of the linear module is:

$$y_j = \omega_i^T x_i + b_j \tag{16}$$

where ω is the weight value ith neuron weight value and b is the bias of the jth neuron.

2.3.5 Loss function

The multi-classification problem uses a cross-entropy loss function of:

$$Loss = -\sum_{i=1}^{K} y_i \log(p_i) \tag{17}$$

where K is the number of action classifications; y_i is the sign function, $y_i = 1$ if the category is i, otherwise take 0; and p_i is the output of the ith neural network.

2.4 Data Collection and Statistical Methods for Dance Rhythm Behavior of Infants and Toddlers

2.4.1 Dance Rhythm Behavior Collection and Processing Methods for Infants and Toddlers

In this study, a qualisys miqus m3 three-dimensional motion capture system (with 8 lenses) and a bertecfully instrumented treadmills three-dimensional force measuring platform (with 2 built-in force plates) were used to synchronously collect kinematic and kinetic data from the subjects during the dance rhythm, and the sampling frequencies of the instruments were set to 250Hz and 800Hz, respectively. Subjects were dressed in tight-fitting clothing without reflective markings. Subjects were dressed in non-reflective clothing and their test numbers and personal information were entered into the qualisys 3D motion capture system. Based on the bony markers on the infant's body surface, the tester applied the markers to the subject's lower limbs from the bottom to the top. Together with 4 fixed marker plates, there were a total of 36 reflective marker points. In order to minimize systematic errors and to ensure precise and fixed placement of the points, the reflective marking points of all subjects thereafter were completed by the same person. The indoor environment was examined by waving a t-bar through the use of the Qualisys 3D motion capture system to avoid external interference during the test.

The raw data collected were identified by the QTM data management software that comes with Qualisys, and the data were processed using Visual3D biomechanical analysis software and Excel. The three-dimensional coordinates of all the reflective markers were smoothed by Butterworth low-pass digital filtering, and the truncation frequency of the kinematic data was 12 Hz and that of the kinetic data was 45 Hz. The foot touching moment was regarded as the time when the vertical ground reaction force was greater than 5 N, and the foot leaving the ground moment was regarded as the time when the foot left the ground when the force was less than 5 N. The testers followed the “nodding law” and the “nodding law” to identify and recognize the data.

The testers tested the subjects according to 12 dance rhythmic movements: “nodding rhythm, bending knee, swinging arm sliding, turning wrist rotating, tiptoe undulating, twisting crotch swinging, clapping rhythm, stomping heavy beat, rotating imbalance, leaning forward and backward, imitating waving arm, and interacting with high-five”.

The specific test methods for the 12 movements are as follows:

Nodding rhythm: Demonstrate nodding up and down and observe whether the infant imitates with the rhythm.

Bending knees: Play music with strong beat and observe whether the knees bend and straighten naturally according to the beat.

Swinging and sliding: Guide infants and toddlers to imitate a bird flying or swinging their arms freely, and observe whether their arms can slide on the body test.

Wrist rotation: Demonstrate wrist rotation, such as the goodbye gesture, and observe whether they can make wrist rotation.

Tiptoe undulation: Observe if he/she tries to lift his/her heels to undulate the body when excited or imitating.

Hip wiggle: Demonstrate wiggling the buttocks from side to side and observe if the pelvis can wiggle from side to side with it.

Rhythmic clapping: Observe if the clapping is generally in time with the music, rather than random.

Stomping to the beat: Demonstrate stomping to the beat of the music. Observe whether the child can respond to the beat by stamping his/her foot forcefully.

Spinning off-balance: Observe if the person can spontaneously spin in a circle with arms

wide open, even if the body is swaying.

Leaning forward and backward: Sing a song with forward and backward imagery, such as “The Boating Song,” and observe if the person's torso swings forward and backward in response.

Imitation of arm waving: Perform exaggerated arm waving motions such as “Hercules” and observe the willingness to imitate and the similarity of the motions.

Interactive high-five: The tester holds the palm of the hand deep in the air and encourages the child to give a high-five with words or a look.

In this paper, the 12 types of movements are summarized into three categories according to their characteristics, namely, “basic gymnastics (nodding rhythm, knee bending, arm swinging and sliding, wrist turning), body exploration (tiptoe undulation, crotch twisting and swinging, hand clapping rhythm, foot stomping and heavy clapping), and social interaction (rotating imbalance, leaning forward and backward, imitation of waving the arms, and interaction of high-fiving)”. We analyzed infants' passive response to movements, active exploration, and interactive imitation to determine the differences in the performance of different genders of infants and toddlers.

Each action was tested 10 times, and 5520 valid data were obtained after screening. Subjects were scored according to their degree of completion of the movements, which were very consistent >80 points, qualified 60-80 points, and failed <60 points out of 100. In this paper, the frequency of behaviors of all infants and toddlers on the three types of actions has been quantified, and the data satisfy normal distribution.

2.4.2 Statistical methods

The data collected from the test were statistically analyzed by applying SAS JMP14.2 statistical analysis software, and the results of all indicators were presented in the form of mean and standard deviation. Differences between the indicators of male and female toddlers were tested using the independent samples t-test. One-way ANOVA was used to test and analyze the indicators, with $P < 0.05$* indicating a significant difference, $P < 0.01$** indicating a significant difference, and $P < 0.001$*** indicating a highly significant difference.

3 Analysis of Gender Differences in Performance Characteristics of Infants' and Toddlers' Dance Rhythm Behavior

3.1 Analysis of the results of the identification of infant and toddler dance rhythmic behavior

3.1.1 Model hyperparameter settings

This paper uses python language and tensorflow framework for network construction and algorithm implementation. In this paper, GPU is specified for computing through tf.device function in tensorflow, which greatly improves the training speed of the neural network through GPU. The memory is 16.0GB, batch_size represents the number of samples selected for each iteration 46, Learning_rate is the size of the learning rate 0.005, the samples to be trained in each epoch are all the training set 200, and m is the confidence threshold 0.85 in the loss function.

3.1.2 Network training and experimental results

In this paper, the loss curve and the accuracy curve during the training process are used to determine whether the model has reached a sufficient number of iterations. The loss curve of the improved OpenPose network model is shown in Figure 2. It can be seen that the improved OpenPose network in the beginning of the training phase of the loss value of the rapid decline in the Epoch reached 88 times, has been reduced to 0.05, when the Epoch reached 159 times, the loss value gradually stabilized, and the Loss value is only fluctuating within the range of 0 ~ 0.02, basically has reached a stable state.

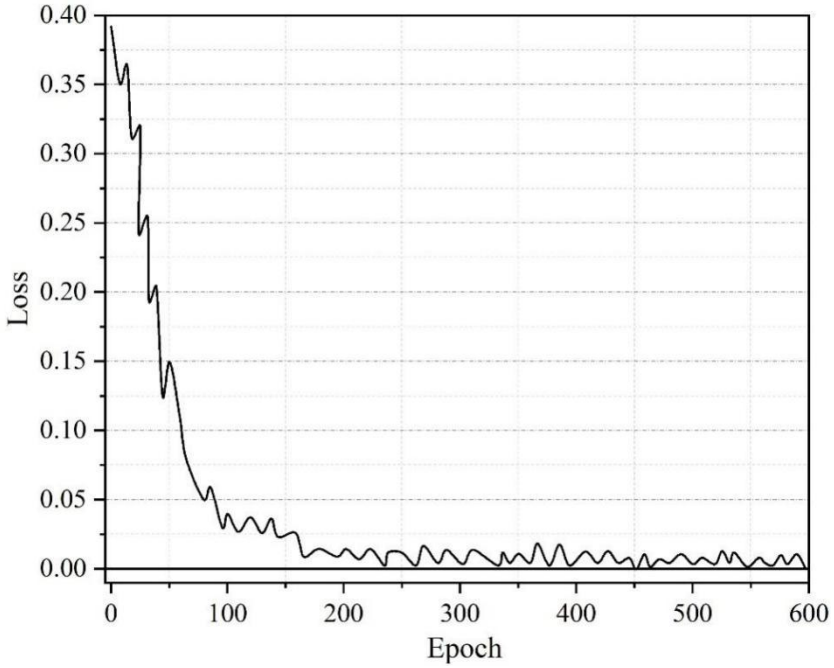


Figure 2: Improving the Loss Curve of OpenPose Network Model

The accuracy curve during the training process of the improved OpenPose network is shown in Figure 3. It can be seen that when the Epoch is 166, the accuracy rate has reached 90%, when the Epoch is 278, the accuracy rate reaches more than 95%, and basically reaches a smooth value, after which the accuracy rate is slowly improved, and after the Epoch is more than 350, the accuracy rate reaches 100%, and the network has already converged to achieve an excellent training effect. By comprehensively considering the changes in the loss curve and accuracy curve of the improved OpenPose network model during the training process, this paper finally determines that the Epoch is 350.

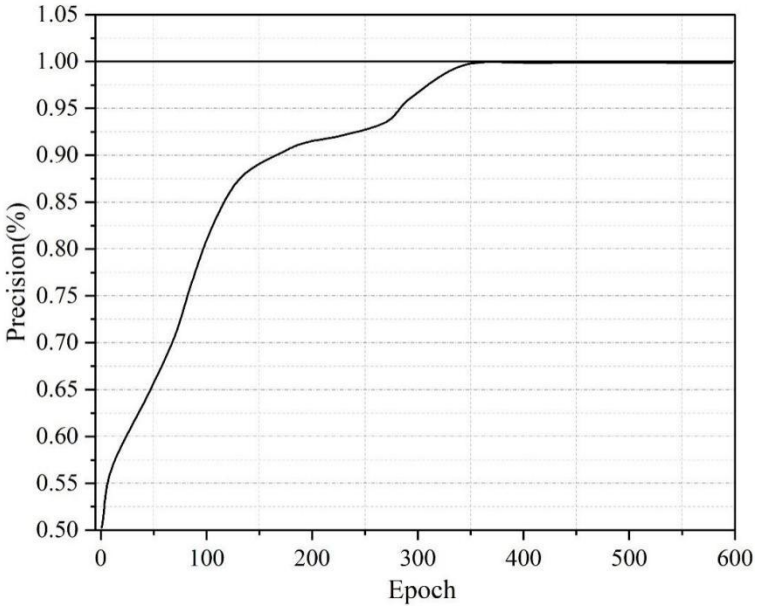


Figure 3: Improving the accuracy curve of OpenPose network training

In order to verify the recognition effect of the improved OpenPose network model proposed in this paper on 12 postures, this experiment tests the trained improved OpenPose network model through the test set, and the test results of each posture test sample in the test set are shown in Table 1. It can be seen that the improved OpenPose network model has the highest recognition accuracy for the basic rhythmic movements, and the recognition accuracy reaches more than 99% for all of them, especially for the head nodding rhythm and swinging arm sliding movements. The second is for social interaction actions, the number of recognition errors for such actions is between 3 and 7, and the recognition accuracy is above 98%. The relatively low recognition effect is for body exploration actions, which have 13~17 recognition error samples, but their overall recognition accuracy is still high, reaching more than 96%. Overall, the model in this paper has a recognition accuracy of more than 96% for the 12 types of movements in the process of infant dance rhythm, which achieves the expected effect.

Table 1: Test results for each posture test sample in the test set

Action category	Movement	Number of test samples	Correct identification number	Accuracy (%)
Basic moral principles	Nodding rhythm	552	552	100.00
	Flexion and extension of knee	552	551	99.82
	Swinging arm	552	552	100.00
	Rotary rotation	552	551	99.82
Body Exploration Category	Heel rise	552	539	97.64
	Swinging	552	537	97.28
	Clap rhythm	552	536	97.10
	Foot stamping	552	535	96.92
Social Interaction	Rotational imbalance	552	549	99.46
	Forward and backward tilt	552	545	98.73
	Imitative arm swing	552	548	99.28
	Clap	552	547	99.09

3.1.3 Comparison experiment of similar action recognition effect

In order to further verify the recognition effect of the improved OpenPose network proposed in this paper on similar actions, this experiment decides to train and test the improved OpenPose network using the ordinary pose dataset and the similar pose dataset respectively, and the recognition results of the improved OpenPose algorithm on different datasets are shown in Figure 4. It can be seen that as the number of training times increases, the accuracy of ordinary pose rises faster. The accuracy rate of the final validation set is at 100%, and the accuracy rate of the similar pose, although it rises relatively slower in the early stage and the final accuracy rate is slightly lower than that of the ordinary pose, still reaches 99.29%. It can be seen that the improved OpenPose network constructed in this paper achieves the expected effect for similar pose recognition of human body.

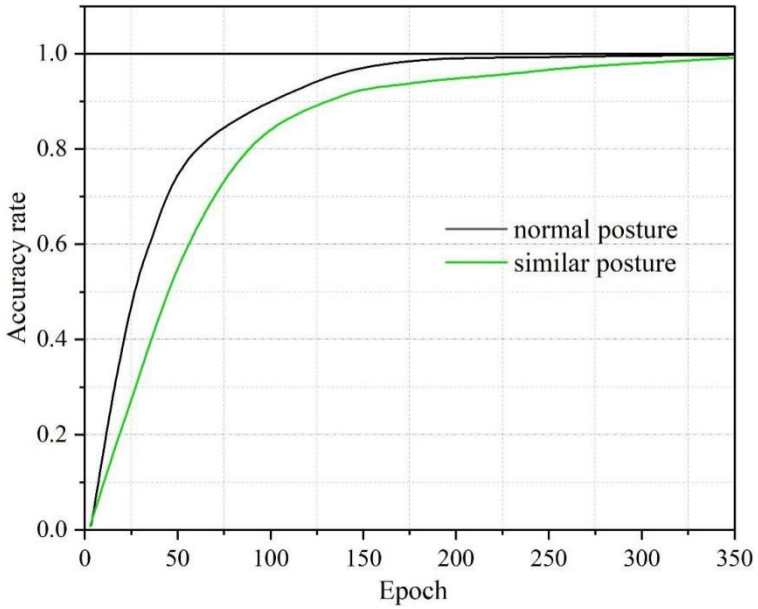


Figure 4: Improving OpenPose Algorithm for Different Data Sets

3.2 Characterization of infant and toddler dance rhythm performance

3.2.1 Differential Analysis of Overall Action Recognition

In this study, an ANOVA was conducted to analyze the gender-based factors that may influence the performance characteristics of infants' and toddlers' dance rhythms. The results of the gender-based analysis of the infant and toddler dance rhythm performance characteristics are shown in Table 2. The results showed that the total frequency of dance rhythmic performance characteristics of infants and toddlers showed a highly significant “gender” main effect ($F=12.953$, $p<0.01$). Therefore, the next step in this study will be to discuss the differences in the performance of basic movement skills of 2- to 3-year-old children at the gender level.

Table 2: Dance Rhythm Performance of Gender Based Infant and Toddler

Source of variation	Quadratic sum	Mean square	F	P
Revised model	20137.289[a]	1268.074	23.532	0.000
Nodal increment	2153622.038	2153622.031	38715.650	0.000
Sex	762.374	762.374	12.953	0.000**
Error	1123.495	57.302	—	—

3.2.2 Differential analysis of basic rhythmic movements

In order to understand the differences between 2~3 year old male and female infants in the completion of basic rhythmic movements, independent samples t-tests were conducted to analyze the total performance of the four movements, namely, “nodding rhythmic movement, bending knee and bouncing, swinging arm and sliding, and turning wrist and rotating”, by gender.

The results of the analysis of the differences in the basic rhythmic movements of infants and toddlers of different genders are shown in Table 3. It was found that there was no significant difference between male and female infants in the basic rhythmic movements (P>0.1). This is related to the fact that at the age of 2-3 years, male and female infants and toddlers are basically synchronized in the rate of perceptual and physiological development of basic rhythmic movements such as nodding rhythmic movements, bending knee bouncing and turning wrist rotating, and there is no significant innate difference.

Table 3: Analysis of the Differences of Basic Rhythm Movement in Children

Action category	Movement	Sex	Mean	SD	T	P
Basic moral principles	Nodding rhythm	Male	83.99	20.13	1.692	0.5830
		Female	82.25	15.11		
	Flexion and extension of knee	Male	87.52	15.76	1.305	0.7249
		Female	86.37	14.08		
	Swinging arm	Male	83.31	19.94	-0.610	0.8751
		Female	84.14	21.57		
	Rotary rotation	Male	80.73	26.59	-0.391	0.9934
		Female	81.35	18.05		

3.2.3 Differential analysis of movements in the body exploration category

The results of the differential analysis of the different genders of infants and toddlers on the body exploration type of movement are shown in Table 4. The results showed that there was no significant difference between infants and toddlers of different genders in the completion of tiptoe undulation movements (T=-3.6924, P=0.1309>0.1); however, there were significant differences between the completion of crotch-twisting and swinging, clapping rhythms, and stomping repetitions, with female toddlers' completion of crotch-twisting and swinging being significantly higher than that of male toddlers (T=-20.0977, P<0.001), while male toddlers had significantly better completion of hand clapping rhythm (T=14.9426, P<0.001) and foot stomping repetition (T=15.0218, P<0.001) movements than female toddlers (P<0.001). This may be related to the fact that toddlers of different genders have different control of body weight and strength under musical stimulation, in which the female toddlers have a more stable body weight, which is consistent with the conclusion of this paper that female toddlers have a higher degree of completion on the twisting and swinging movements than male toddlers.

Table 4: Analysis of the Differences in the Children's Exploration of the Body

Action category	Movement	Sex	Mean	SD	T	P
Body Exploration Category	Heel rise Swinging	Male	88.17	20.83	-3.6924	0.1309
		Female	89.53	21.67		
	Clap rhythm	Male	73.83	16.98	-20.0977	0.0000***
		Female	88.57	18.95		
	Heel rise Swinging	Male	89.15	21.94	14.9426	0.0000***
		Female	81.34	22.05		
	Clap rhythm	Male	89.19	22.83	15.0218	0.0000***
		Female	80.82	13.51		

3.2.4 Differential Analysis of Social Interaction Class Actions

The results of the analysis of the differences in the social interactions of infants and toddlers by gender are shown in Table 5. The results showed that there were significant differences in the completion levels of "rotating off-balance, leaning forward and backward, imitating arm waving and interacting with high-fives" among infants and toddlers aged 2~3 years old. The mean scores of "imitating arm swing" and "interactive high-five" were 14.30 and 12.04 points higher than those of boys, while the mean scores of "rotating off-balance" and "leaning forward and backward" were significantly higher than those of girls, with a difference of 11.48 and 14.20 points, respectively. Overall, through the introduction of games and situations, it can be observed that children of different genders will show different levels of participation, which may be caused by the difference between male and female children in terms of eye contact, intentionality and creativity of movements.

Table 5: Analysis of the Differences of Social Interaction Actions in Children

Action category	Movement	Sex	Mean	SD	T	P
Social Interaction Social Interaction	Rotational imbalance Forward and backward tilt	Male	88.17	20.83	18.529	0.0000***
		Female	73.87	25.68		
	Imitative arm swing Clap	Male	85.48	17.74	16.803	0.0000***
		Female	73.44	16.04		
	Rotational imbalance Forward and backward tilt	Male	79.3	25.04	-14.261	0.0000***
		Female	90.78	28.53		
	Imitative arm swing	Male	75.72	23.52	-18.114	0.0000***
		Female	89.92	21.36		

3.3 Kinematic analysis of dance rhythms in infants and toddlers of different genders

The joint angles involved in this study are all flexion-extension angles for the hip, knee, ankle and elbow joints, except for the shoulder angle, which is an abduction angle. In this section, we take the changes in the right and left knee joints during the dance rhythm and the upper limb trajectory during the right leg kick as an example to analyze their differences among infants and children of different genders.

3.3.1 Angular characterization of the right and left knee joints

The left and right knee angles of male and female toddlers at each characteristic moment of the flicking and kicking movements are shown in Table 6. The results show that, during the identification process of infants' dance rhythmic behavior, it was found that there was a

significant difference between male and female infants' left knee angles at the moment of right leg kicking out (P=0.0378<0.05), and there was a significant difference between genders' right knee angles at the moment of right leg dropping back (P=0.0432<0.05). In other performance characteristics, there were no significant differences in the dance rhythmic behavior of infants and children (P > 0.05). The reason for this result is related to the physiological structure, neuromotor control, and behavioral patterns of male and female infants. Specifically, female infants typically have a wider pelvis and greater anterior pelvic tilt, whereas this characteristic is not evident in male infants. The mean right knee angle at the moment of the right leg kick was 63.02°, indicating that female infants had a greater degree of knee hyperextension than male infants. This is related to the relatively high level of magnetic hormone in female infants, which may lead to greater laxity of joint ligaments throughout the body, making the stability and range of motion of the knee joint different from that of male infants in weight-bearing and locomotion.

Table 6: Comparison of angular characteristics of left and right knee joints

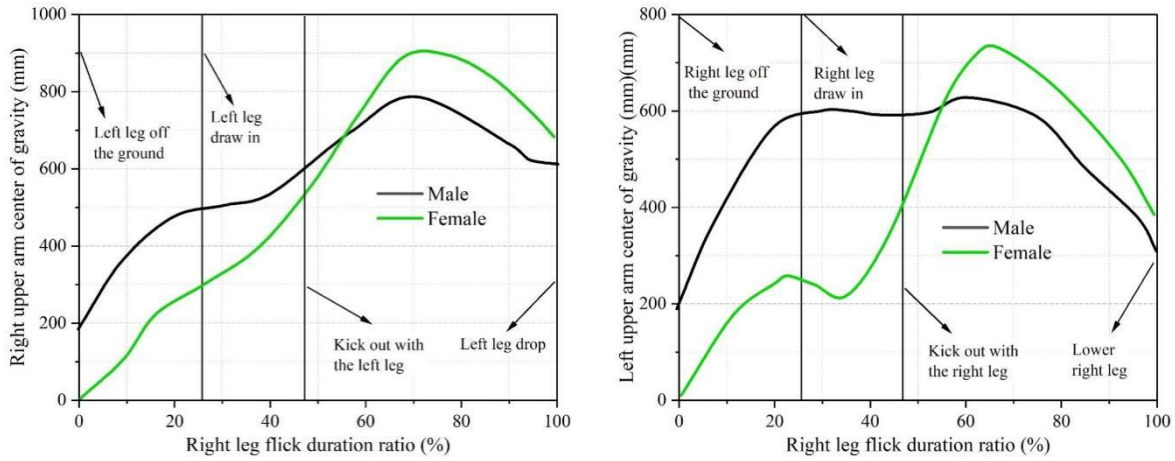
left and right knee	Characteristic time	Male		Female		P
		Mean(°)	SD	Mean(°)	SD	
Left knee joint angle	Left leg off the ground	48.11	10.82	51.12	7.69	0.7222
	Left leg draw in	27.04	6.25	19.51	11.11	0.7223
	Kick out with the left leg	88.78	7.06	89.65	8.38	0.9672
	Left leg drop	34.07	6.83	46.84	10.51	0.4609
	Right leg off the ground	60.32	10.94	59.02	11.39	0.8911
	Right leg draw in	68.38	9.18	69.81	9.86	0.927
	Kick out with the right leg	74.81	5.86	64.88	10.55	0.0378*
	Lower right leg	79.97	5.16	83.8	8.91	0.3028
Right knee joint angle	Left leg off the ground	46.09	19.18	52.06	6.94	0.2384
	Left leg draw in	52.66	7.25	52.48	9.73	0.5453
	Kick out with the left leg	56.87	14.75	49.47	4.63	0.0777
	Left leg drop	56.59	14.83	61.32	14.25	0.6899
	Right leg off the ground	29.35	17.47	42.6	26.71	0.4136
	Right leg draw in	28.34	18.14	22.99	19.26	0.7126
	Kick out with the right leg	49.18	18.88	63.02	20.67	0.5085
	Lower right leg	17.62	9.09	29.34	6.72	0.0432*

3.3.2 Trajectory analysis of the upper limb during the right leg stride

(1) Y-coordinate trajectory of the center of gravity of the upper limb during the right leg kicking process

The Y-coordinate trajectories of the center of gravity of the upper limb during the right leg kick are shown in Figure 5, where (a) and (b) represent the right upper arm and the left upper arm, respectively. It can be seen that in the right leg knee lifting stage, the kicking stage and the early stage of the leg closing stage, the Y-coordinate of the center of gravity of the left upper limb of both male and female children increased, and the increase in the magnitude and speed of increase (the degree of curve inclination) of female children was greater than that of male children; in the middle and late stage of the leg closing stage, the Y-coordinate of the center of gravity of the left upper limb of both male and female children decreased, and the magnitude of decline of the decline in the magnitude and speed of the decline in the decrease in the Y-coordinate of the center of gravity of the male child was slightly less than that of female child. In the right knee-lifting and kicking phases, the Y-coordinate of the center of gravity of the right

upper limb decreased for both boys and girls; in the leg-retracting phase, the Y-coordinate of the center of gravity of the right upper limb increased for both boys and girls. The magnitude of the curve change was similar for both sexes. Compared with male children, female children had greater forward swing distance and swing speed of the left upper limb in the first two movement phases and in the first part of the right leg retraction phase; and the backward swing distance and speed of the left upper limb were slightly greater than those of male children in the subsequent right leg recovery process.

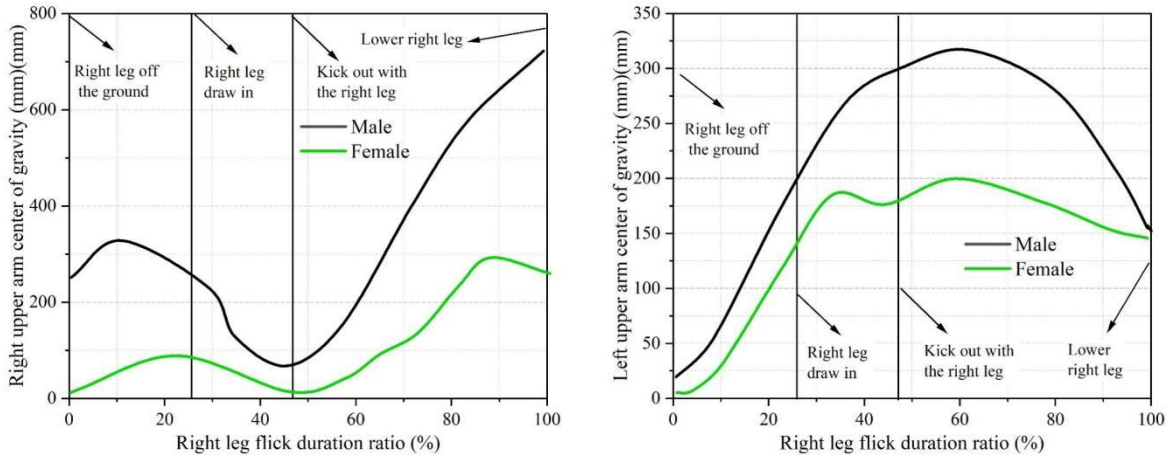


(a)Right upper arm (b)Left upper arm

Figure 5: Y-axis coordinate trajectory with shifting center of gravity

(2) X-coordinate trajectory of the center of gravity of the upper limb during the right leg kick.

The X-coordinate trajectories of the center of gravity of the upper limb during the right leg kick are shown in Fig. 6, where (a) and (b) represent the right upper arm and the left upper arm, respectively. In the right leg knee-lifting and kicking phases, the X-coordinate of the center of gravity of the left upper limb of both male and female children increased, while the increase in the magnitude and speed of increase (the degree of curve inclination) of female children was smaller than that of male children; in the leg-retracting phase, the X-coordinate of the center of gravity of the left upper limb of both males and females decreased, and the magnitude of decrease and the speed of decrease of male children were slightly larger than that of female children. In the right knee lifting and kicking phases, the X-coordinate of the center of gravity of the right upper limb of both sexes showed a small decrease, and the decrease in the center of gravity of the right upper limb of the female children was slightly smaller than that of the male children; in the leg closing phase, the X-coordinate of the center of gravity of the right upper limb of both sexes rebounded, and the magnitude and speed of rebound of the female children were smaller than that of the male children. Compared with male children, female children's right upper limb abduction distance and abduction speed were smaller in the leg closing phase.

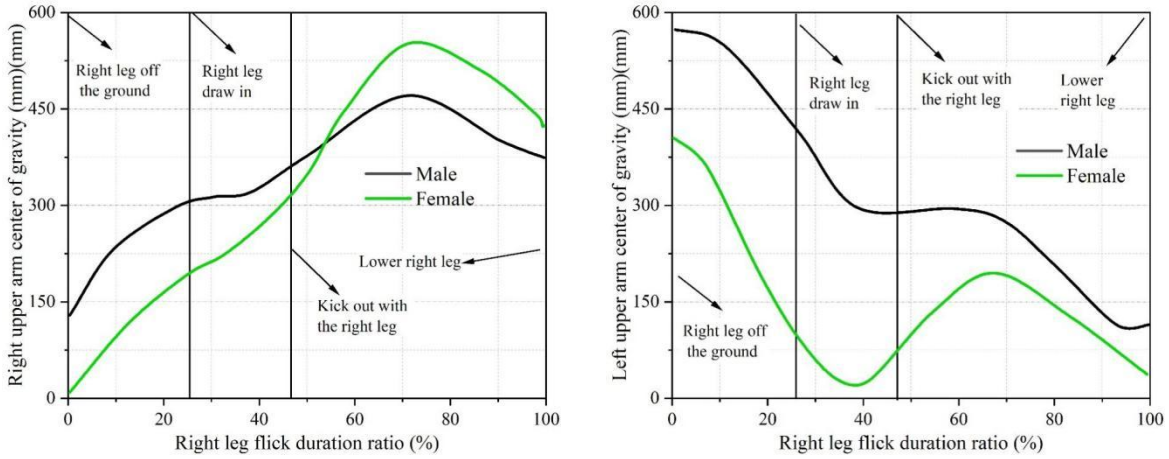


(a)Right upper arm (b)Left upper arm

Figure 6: X-axis coordinate trajectory with shifting center of gravity

(3) Z-coordinate trajectory of the center of gravity of the upper limb during the right leg kick.

The Z-coordinate trajectories of the center of gravity of the upper limb during the right leg kick are shown in Fig. 7, where (a) and (b) represent the right upper arm and the left upper arm, respectively. The Z-coordinate of the center of gravity of the right upper limb of both male and female children during the right leg kick showed a general upward trend, and the magnitude and speed of the upward movement were similar for both sexes. The Z-coordinate of the center of gravity of the left upper limb decreased during the right leg kick, and the magnitude and speed of the decrease were similar for both sexes.



(a)Right upper arm (b)Left upper arm

Figure 7: Z-axis coordinate trajectory with shifting center of gravity

4 Conclusion

In this paper, in order to realize the gender difference in the dancing behavior of infants and toddlers, we propose a lightweight infant behavior recognition model based on OpenPose improvement, and collect and process the data of infants and toddlers' dancing behavior through qualisys miqus m3 and Visual3D, to complete the extraction of infants and toddlers' dancing behavior performance and the computational behavioral analysis of infants and toddlers'

dancing behavior of different genders. The computational behavioral analysis of the dance rhythmic behavior of infants and toddlers of different genders was completed. The results show that:

The Epoch of the improved OpenPose network model gradually stabilizes after reaching 159 times, and its loss value fluctuates within the range of 0~0.02; when the Epoch reaches 350, the network has converged, the accuracy of the training effect reaches 100%, and the accuracy of the model in recognizing common movements and similar gestures is above 99%.

Male and female toddlers aged 2~3 years did not show any significant difference in basic rhythmic movements due to the basic synchronization of their perceptual and physiological development ($P>0.1$); female toddlers had a higher degree of completion than male toddlers in movements such as twisting and swinging and interactive high-five ($P<0.001$), and male toddlers had a significantly higher degree of completion than female toddlers in the rhythmic clapping, foot-stomping and heavy clapping, rotational imbalance and forward leaning and backward leaning; thus, it can be seen that different genders are more accurate in recognizing common movements and similar postures. This shows that infants and toddlers of different genders have significant differences in body exploration and social interaction movements. The main reasons for the differences in the kinematic analysis of dance movements between infants and toddlers of different genders are related to the innate tendency of skeletal structure and hormone-affected ligament characteristics, as well as the different movement patterns and behavioral choices of infants and toddlers.

Funding

This article is a project of Fuzhou Social Science Planning: Research on the Improvement of Observation and Evaluation Ability of Preschool Teachers in the Era of Digital Intelligence (project number: 25SK72).

About the Author

Nanhua Zeng was born in 1989 in Fuzhou City, Jiangxi Province, China. She obtained a master's degree from Jiangxi Normal University. Currently, She is working at the School of Preschool Education in Fuzhou Preschool Education College. Her main research direction is psychological development of preschool children.

Junfang Zheng was born in 1988 in Fuzhou City, Jiangxi Province, China. He obtained a master’s degree from Shaanxi Normal University in China. He is currently working at the School of Preschool Education, Fuzhou Preschool Education College. His main research directions are focusing on teacher education, preschool education and vocational education.

Ting Dong was born in 1990 in Fuzhou City, Jiangxi Province, China. She obtained her master's degree from Suan Sunandha Rajabhat University in Thailand. Currently, She is working at the School of Arts and Sports of Fuzhou Preschool Education College. Her main research direction is dance studies.

Haiyou Liu was born in 1988 in Fuzhou City, Jiangxi Province, China. He obtained a bachelor's degree from Jiangxi Normal University. Currently, He is working at the Inspur Software Technology Co., Ltd. His main research direction is computer technology and big data.

References

[1] Trevarthen, C., Delafield-Butt, J., & Schögler, B. (2016). Psychobiology of musical gesture: Innate rhythm, harmony and melody in movements of narration. In New perspectives on music and gesture (pp. 11-44). Routledge.

[2] Levitin, D. J., Grahn, J. A., & London, J. (2018). The psychology of music: Rhythm and movement. Annual review of psychology, 69(1), 51-75.

[3] Morley, I. (2012). A grand gesture: vocal and corporeal control in melody, rhythm, and emotion. Language and music as cognitive systems, 110-120.

[4] Wilson, J. M., & Henley, M. (2022). Experiencing rhythm in dance. Frontiers in Psychology, 13, 866805.

[5] Clarke, E. F. (2016). Rhythm/body/motion: Tricky's contradictory dance music. In Musical rhythm in the age of digital reproduction (pp. 105-120). Routledge.

[6] Moore, P. A. (2019). The Rhythm of the Dance. In Into the Illusive World: An Exploration of Animals' Perception (pp. 207-212). Cham: Springer International Publishing.

[7] Hamilton, A. (2020). Rhythm and Movement: The Conceptual Interdependence of Music, Dance and Poetry. Midwest Studies in Philosophy, 44(1), 161-182.

[8] Fitch, W. T. (2016). Dance, music, meter and groove: a forgotten partnership. Frontiers in human neuroscience, 10, 64.

[9] Popa, D. (2019). Dance, Rhythm, and Social Space. Studia Universitatis Babes-Bolyai-Philosophia, 64(2), 109-120.

[10] Zheng, W. (2020). An Application Based on the Body Rhythm Teaching Method in Order to Improve the Students' Musical Perception. Frontiers in Art Research, 2(6), 15-23.

[11] Kassing, G., Jay-Kirschenbaum, D., & Jay, D. M. (2021). Dance teaching methods and curriculum design: comprehensive K-12 dance education. Human Kinetics Publishers.

[12] Goldberg, T. L. (2018). Release, relax, ready: A rhythm-first, holistic approach to teaching tap dance. Dance Education in Practice, 4(3), 17-24.

[13] Xiu, Y., Liu, X., Yuan, Y., Zhao, H., & Zhang, C. (2023). A new method of dance rhythm training based on an immersive cave automatic virtual environment. IEEE Access, 11, 109422-109434.

[14] Huang, R., Hu, H., Wu, W., Sawada, K., Zhang, M., & Jiang, D. (2020, June). Dance revolution: Long-term dance generation with music via curriculum learning. In International conference on learning representations.

[15] Faber, R. (2017). Dance and early childhood cognition: The Isadora effect. Arts Education Policy Review, 118(3), 172-182.

[16] Kim, M., & Schachner, A. (2023). The origins of dance: Characterizing the development

of infants' earliest dance behavior. Developmental Psychology, 59(4), 691.

[17] Cirelli, L. K., & Kragness, H. E. (2025). The Development of Dance in Early Childhood. Current Directions in Psychological Science, 09637214251323490.

[18] Liya, A., & Katoningsih, S. (2022). The development of learning the arts of dance to the ability of early childhood gross motor development. Early Childhood Research Journal (ECRJ), 4(2), 1-8.

[19] Golding, A., Boes, C., & Nordin-Bates, S. M. (2016). Investigating learning through developmental dance movement as a kinaesthetic tool in the early years foundation stage. Research in Dance Education, 17(3), 235-267.

[20] Rossini, A. P., & Rossi, F. (2024). Corporeality and dance for babies: a systematic review study. Educação em Revista, 40, e40238.

[21] Kim, M. J. (2022). Meaningful Movements in Early Childhood: The Cognitive and Developmental Bases of Dance, Gesture, and Musical Actions. University of California, San Diego.