WILEY | Hindawi

Retraction

Retracted: A Hybrid Density Network-Based Dance Movement Generation Algorithm and Internet of Things for Young Children

Wireless Communications and Mobile Computing

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Fang, "A Hybrid Density Network-Based Dance Movement Generation Algorithm and Internet of Things for Young Children," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 2245284, 8 pages, 2022.

WILEY | Hindawi

Research Article

A Hybrid Density Network-Based Dance Movement Generation Algorithm and Internet of Things for Young Children

Yuting Fang [ORCID]

School of Computing, Yulin Normal University, Yulin 537000, China

Correspondence should be addressed to Yuting Fang; 19402309@masu.edu.cn

The music development matching model and the quantifiable planning model have an undesirable fit between the dance produced by the model and the music self in terms of music-driven dance development age (e.g., generated dance development is lacking, and long-distance dance arrangements are lacking in perfection and discernment). The traditional methodology cannot produce new dance moves or other associated concerns. To address these concerns, we are working on a dance age estimation based on technological developments and neural networks that will eliminate the need for voice and development planning. The first stage uses the prosody elements and sound beat highlights extracted from music as music highlights, while the second stage uses the directions of essential human body issues derived from dance recordings as movement highlights. The model's generating module acknowledges the vital planning of music and dance advancements to build a smooth dance posture in the next stage; the discriminator module acknowledges the autoencoder module agent has improved sound characteristics and the consistency of dance and music. In the third and final step, the model's transformed form changes the dance act succession into a good diversity of dance. Finally, a reasonable rendition of the dance that matches the music has been found (e.g., trial data is gathered from online dance recordings, and the exploratory outcomes are examined from five perspectives: poor work esteem, correlation of various baselines, assessment of grouping age influence, client examination, and genuine dance recording quality assessment). The proposed dance age model has a reasonable impact on converting into actual dance recordings, according to the results.

1. Introduction

Visual and auditory perceptions are intricately intertwined. Regardless of how long the article moves, visual changes will almost certainly result in the formation of audible sound. Currently, the majority of AI is still learning facts in a solitary mode. The transition from single-modular learning to multimodal learning has recently become the way to a greater understanding of machine perception, thanks to the rapid growth of artificial intelligence. A growing number of scientists are focusing on multimodular data learning, including cross-modular recovery, multimodular data joint decision-making and cross-modular age. The cross-modular age is aimed at orchestrating one or a few modules of information based on data from various modalities. Text to picture, text to video, sound to picture, and sound to video are all examples of cross-modular age, whether one-way or two-way [1].

With the recent turn of events and advocacy for profound learning, fake brain networks have been successfully used to the dancing age's developments. The most important advantage of involving profound learning for dancing age is that they may easily eliminate developed excerpts from raw data Deep brain networks can also come up with new dance movements. Using IoT in the sphere of education will be crucial for enhancing learners' aptitude and knowledge, particularly in children's education. Because of the individual and group use of technology in school, learning itself will alter. Technology has a significant impact on all learning-related activities. Since all students may access educational resources and information at any time and from any location, the IoT is a subcategory of internet technology in this context. In other words, the use of IoT technology intends to boost children's development while also piquing their interest in playing [2].

However, due to significant advancement, the dance age algorithm has a few flaws. Because of the start-to-finish model, for example, the produced dance may not be smooth when the casings are used. The produced dance's perceived impact is significantly worse; on the other hand, algorithmic maneuvers are frequently difficult to synchronize with music. Furthermore, people usually draw the human skeleton or perform liveliness handling in accordance with the directions of the central concerns of the human body for the perception of dancing, and there is room for further improvement in the representation impact.

2. Related Work

Body movement age, sound-driven picture age, and talking face video age are the three categories of cross-modular age from sound to video. A frequent cross-modular age assignment is to combine appropriate face footage using discourse or music. An early study on the age of talking faces was largely utilized to create a certain personality from a dataset due to erratic speech and sound. Kumar et al. [1] utilized a postponed LSTM [3] to build central issues that were synced to music and then another network built video outlines based on the central issues (it is the main network engineering that takes any text and turns it into a movie that compares voice and lip-sync to match photographs to the real world). Unlike previous distributed systems, they use only a completely teachable brain network and do not use typical PC illustration methods. The approach incorporates three primary modules to address these basic challenges: a Char2 Wav-based text-discourse network, a deferred LSTM for producing speech foci that are synchronized with sound, and a Pix2-Pix-based network for recording. As a result, Chung et al. [4] employed an EncoderDecoder CNN model, which used the combined implanting of face and sound to construct fabricated talking facial video outlines, to understand the relationship between the first sound and video [4]. (A still image and a sound voice fragment of the objective face are among the model data sources, as well as a lip-formed video of the objective face that is synchronized with the sound.) RNN and GAN [5] were used by Jalalifar et al. [6] to create a sequence of real countenances synchronized with the information sound by two networks. The LSTM network is one of them, and it is used to construct sound-based lip milestones. The other is dependent GAN, which uses a set of lip marks to build facial images. These two networks have the potential to collaborate and to create a distinctive talking face arrangement that is coordinated with the information sound track. Borra et al. [7] also presented a dynamic pixel misfortune period consistency approach [3]. Unlike the instantaneous sound with picture technique, this course strategy strives to avoid making fictitious linkages between the various media flags that are unrelated to the discourse topic. The authors presented a novel dynamic and adjustable pixel-level loss attention technique to eliminate pixel jitter problems by boosting the network's attention to audio-visual linked areas; they also demonstrated a new regression-based discriminator structure that considers sequence- and frame-level information to generate clearer images with well-synchronized face motion.

2.1. Dance Movement Recognition Technology. The movement of artists differs significantly from the movement of ordinary people on a daily basis. Several developments necessitate the artist's arms and legs being joined. It is critical to obtain the artist's whole body movement information while selecting the objective region for foundation acknowledgment. The correct understanding of their behavior and the view of artists' evolution can be divided into a few categories: size, shading, body shape, profundity, and so on, which are examples of static characteristics. The goal is to capture the artist's face. It can, for example, pass on a lot of information about the artist's progress through the artist's shadow [7]. This is a basic shape that you can draw. The artist's growth path recalls these characteristics in order to flesh out the concept of the artist's development path. The qualities of existence are most often presented in illustrated ways, such as the state of reality and focus points, such as artists' scenes, encompassing articles, and signals, to create conditions for exhibiting. Dance video picture recognition takes into account the characteristics of dance foundation and apparel when it comes to movement recognition. The changes should be kept in order to accurately document and fully reflect the artist's development information, static information assortment innovation, and artist's body development information. This collection provides a specific memory approach based on previous research [8]. These approaches are appropriate for the concept of artists' developments, can precisely distinguish and perceive artists' developments in dance photographs, work on the accuracy and competence of recognition, and have tremendous value applications. To depict the objective's direction, provide the typical direction, distance, bearing of movement, and normal speed. To begin, the mean shift technique is used to reconstruct plausible traffic classes for each property zone. More property combination algorithms are intended to intertwine the arrangement information by determining the relationship between the classes in various quality localities and then getting the bunching results of the information blend of multiple characteristic districts. Because information is incorporated at the gathering level, this is the case [9]. As a result, the issue of layered mix can be successfully avoided while consolidating at the spatial level. According to the substance of conduct, the examination of human conduct insight can be divided into three stages, from simple to complex: evolution of vision. There are two difficulties that are fundamentally related: one is in the identifiable proof cycle [10]. A considerable percentage of the information in the video activity overview is redundant or irrelevant to activity recognition. This increases the computational complexity and influences the accuracy of activity recognition. Furthermore, the primary method for selecting and comprehending ascribes is to collect portrayal and posture highlights, despite the fact that movement scenes and ensembles are various. Furthermore, given the concept of data signals, it is easy to miss the human body's static state [11].
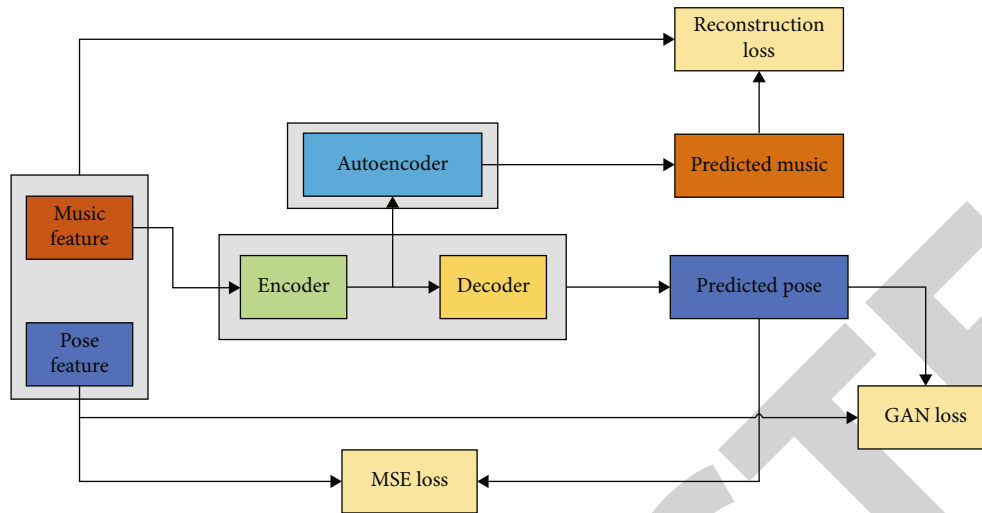
FIGURE 1: The dancing generative model's network structure is depicted in this diagram.
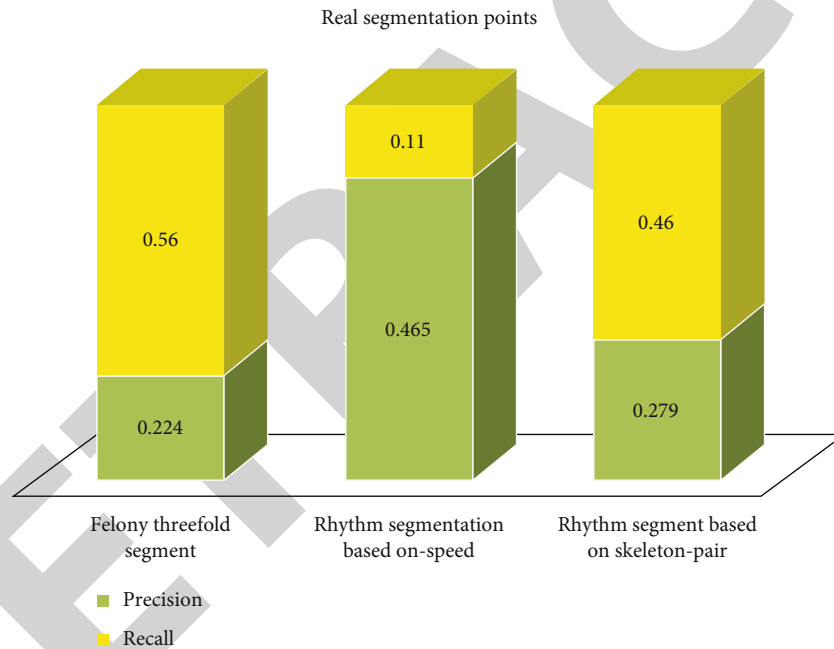

FIGURE 2: Real segmentation.

2.2. Dance Movement Generation. To move, a large range of contemporary AI-based development age methods have been used. The blend of dance advances has been studied using Secret Markov Models and their extensions. On movement catch information containing creative dance walk, expressive dance roll, disco, and complicated disco, Wang et al. created Hierarchical Hidden Markov Models with nonparametric yield appropriations (NPHHMM) [5]. Another strategy is based on dynamical frameworks that demonstrate the ability to capture the aspects of dance advances. Linear Dynamical Systems (LDS) were used by Li et al. to learn and produce dance innovations. They put their model through a 20-minute dance development with a skilled performer, who primarily performs disco. Their model learns movement sextons on its own, taking into account neighbor-ing development aspects. Every mind-boggling development grouping, according to the sextons, is made up of fundamental, repetitive examples [12]. A dance arrangement, for example, could include repetitive maneuvers like twisting, bouncing, kicking, and pussyfooting. The approach considers constant mix and provides numerous methods for developing innovations, such as key-outlining and commotion-driven age.

Fake brain networks have recently been effectively applied to the integration of dance innovations. For the game Dance Revolution, Donahue et al. focused on creating movements, which they handled as step graphs that encode the context and position of steps. They used LSTMs to create a new progression graph based on a rough sound recording. In any event, their technique is limited to the age of discrete
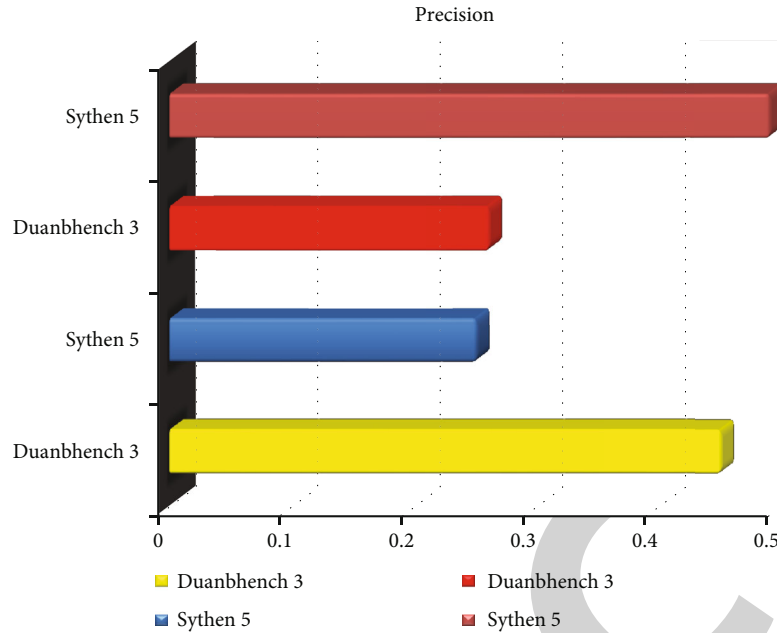
Figure 3: Precision.

Table 1: Segmentation points.

Method	Precision	Recall
Felony threefold segment	0.224	0.56
Rhythm segmentation based on speed	0.465	0.11
Rhythm segment based on skeleton-pair	0.279	0.46

Table 2: Wavelet function.

Wavelet function	Precision
Duanbhench 3	0.45
Scythe 5	0.25
Duanbhench 3	0.26
Scythe 5	0.49

progress pointer groupings rather than long-term improvements. To learn and produce movement, Crnkovic-Friis used recurrent neural networks (RNNs) with long short-term memory (LSTM) were used. They used 6 hours of contemporary dance data taken with Microsoft Kindest to train the model. This method provides no options for controlling the age and is devoid of music [13].

Controlling the movement generation based on machine learning and controlling the qualities of the advancements made by an AI model can be done in a number of ways: (1) train a distinct model for each recognition of a development quality, (2) use parametric quantifiable conveyances to capture the various types of development, and (3) expressly plan AI models to help with the task of controlling the age cycle [14]. In light of Taylor and Hinton's FCRBM design, we use the age controlling mechanism for GrooveNet. Nonetheless, our study differs from the neural network-based dance age techniques previously discussed in that we want

to create constant dancing advancements confined by a set sound track [15].

2.3. Music-Driven Dance Generation. While cross-modular grouping to arrangement planning is also considered in music-driven dance age, it is critical to emphasize the complexity of music-to-move planning. While similar hidden processes provide acoustic and movement cues in discourse, the relationships between music and evolution in dance are definitely more convoluted and chaotic. They are based on the entertainer's class and environment, as well as his or her ability and personal characteristics, and they exhibit a sophisticated pecking order of temporary designs, ranging from momentary synchronization of motions to the rhythm to long-term growth of dance designs [16].

Sound-driven moving symbols using HMM-based movement combinations were presented by Li et al. [17]. They want development to be physically remarked on into explicit demonstrations (or dance figures) coordinated with the beats as part of the preparation process. For the age, the sound is divided into sections based on beat location, and the movement examples are chosen based on Mel-Frequency Cepstral Coefficients recognition. Their methodology was also extended to include individual analysis of dance designs[18] Li et al. present three types of models: melodic measure models, interchangeable figures models, and figure change models, which address the many-to-one and one-to-numerous relationships between melodic instances and dancing figures, respectively. However, one of the major impediments to these techniques is the union methodology, which is based on the grouping of information melodic examples and so provides little opportunities for creating unique development ideas [19].
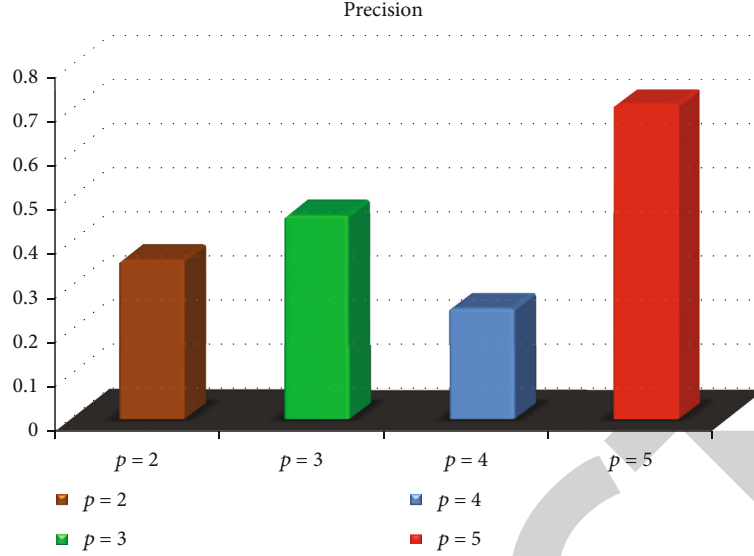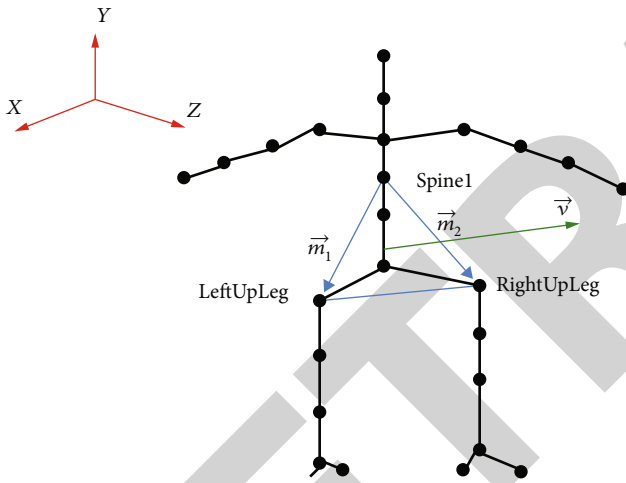
Precision

FIGURE 4: Precision values.

FIGURE 5: Feather Dress Dance.

3. Method

The proposed technique is described in detail in this section. Most notably, the network of long-and short-term memory (LSTM) is investigated, followed by discussions of its expansions, developments, and upgrades. In addition, the recipe has been updated. Furthermore, the foundation for the building of the dancing age model is deep learning. The strategy of prosody, which includes extraction, is then introduced. It is followed by a depiction of gathering information. Finally, the configuration of the generator is presented.

3.1. Long- and Short-Term Memory Network. The LSTM stands for long-and-temporary memory network, and it is a unique type of RNN. Its goal is to alleviate the long-term reliance on recurrent brain networks. Poutanen et al. [20] and Ginosar et al. [14] presented it. Many people worked on it after then refining and improving it. LSTM has pro-

duced excellent results. It is commonly used to solve a range of time series problems [21].

The first step in the LSTM procedure is to decide whether you want to use it; data from the cell should be discarded. The "neglected door," a sigmoid layer, constrains this decision. The ignoring entrance feeds input ht-1 and x1 to each component in the phone state Ct-1 The level of data held from the previous cell state Ct-1 to this phone is then returned as a number between 0 and 1. ft 1 denotes "keep this data," while ft 0 denotes "remove this data." The following is an updated version of the original recipe [22, 23]:

$$Lf(a, b) \frac{-ef}{Ct} \quad g(x, y, z) * I(a, b). \tag{1}$$

The following step is to determine what additional data the model will store in the cell state. There are two stages to this progression.

$$i, = \sigma(W \cdot [h_i - 1, x_t] + b_i),$$
$$\overline{\overline{C}} = \tan \ (W \, c[h_i - 1, x_t] + B_c. \tag{2}$$

After that, change from the previous cell state to the new cell state

$$C_t = f_t * C_{t-1} - 1 * C_{t^{\frown}}. \tag{3}$$

Finally, based to the information h_t and x_t, the final result is still up in the air after refreshing the cell express. The final result will be based on the current state of the cells, with some data sifted in.

$$O_t = \sigma \ (W0[h_i - 1, x_t] + B_0),$$
$$h_t = O_t' \tan \ (C_t). \tag{4}$$

3.2. Model Overall Design. In light of significant learning, Figure 1 depicts the overall arrangement of a dance age model. The red and blue boxes are for music and dancing signal quality, respectively, while the dim box is for the handling or network module. The light-orange box refers to the unpleasant work environment. Perform sound component extraction and activity including extraction on the dance information first, as shown in the figure. After that, enter the sound highlights into the dance generator to get the predicted dance posture, then perform MSE Loss with the real dance act; the autoencoder module is used to obtain the sound highlights [24]. The anticipated dance pose and the genuine dance pose are shipped off the discriminator for segregation, and the model is prepared for misfortune; the sound elements of the design are built, and the model is prepared for catastrophe; the anticipated dance posture and the genuine dance pose are sent to the discriminator for segregation; and the model is prepared for sound remaking deficit [10].

3.3. Research Methodology. Given the scarcity of human dance movement catch data sets, a dance movement information base based on movement catch information is built. The Notch movement catching framework, which has 12 inertial sensors and runs at 150 Hz to test recurrence, is used to record the BVH data. It has the potential to produce high-quality movement catch data. Figure 2 demonstrates the dance development's security system. Over 160 movement catch information is described in our analysis, and over 100 BVH-designed dance development part data sets are acquired [17].

4. Data Analysis

4.1. Segmentation Method. We compare the presentation of our division strategy to the speed restriction division approach described in [25]. The skeleton-pair distance is also used to guide the trial (as expressed in Section 3). This trial is based on a collection of 100 dance development parts created by BVH, each having a separate fundamental component development and a similar direction highlight. To evaluate the exhibition of all division methods, we use the accuracy rate and review rate as assessment pointers as show in Figure 3.

$$\text{Recall} = \frac{TP}{TP + EN},$$
$$\text{Precision} = \frac{TP}{TP + FN},$$
$$f(i) = \frac{2 * \text{precision}(i, j) * \text{recall}(i, J)}{\text{precision}(i, J) + \text{recall}(i, j)}, \quad (5)$$
$$TP, TP + FP, \text{ and } TP + FN.$$

The number of correct segmentation points, correct segmentation points, and actual segmentation points are represented by TP, TP + FP, and TP + FN, respectively,[18]; see Table 1.

From Table 2, our strategy achieves the maximum division precision possible. The musicality division has 94 percent accuracy in terms of speed, which is 19 percent greater than the speed limit division used in work [26–29]. Furthermore, in the context of kinematic highlights and beat, the impacts of division approaches are superior to those in the context of kinematic highlights. In terms of speed, accuracy, and review of skeleton-pair distance, the beat division is slightly greater. As a result, we chose the skeleton-pair distance as our activity division kinematics highlight [30].

The dimensionality of PCA should be chosen carefully. The division precision is also influenced by the component of the PCA vector. Tests are done at varied p of 2, 3, 4, and 5, independently, to confirm the optimal vector aspect. Furthermore, these p advantages are all met under state in which Ep is more than or equal to 85%, showing that the accuracy is relatively similar when $p = 3$ and $p = 4$. As a result, three were chosen in our experiment to ensure excellent distinguishing proof precision while reducing the aspect as shown in Figure 4 [31].

Rainbow and Feather Dress Dance were lab notated in part. Figure 5 shows four positions that correspond to the original movements [2, 32, 33].

5. Result and Discussion

We will use the examples of movement and to illustrate our point. The right leg is upstanding when moving, the upper level is "centre," and the flat heading is "back." As a result, the Laban image "centre, back, 1" is related to the left leg "centre, left forward, 2." The left arm is totally fixed, whereas the right has a greater degree of bend. As a result, the left lower arm is addressed by the Laban picture "centre, left back, 1," whereas the right lower arm is handled by the image "high, correct, 4." "Low, left forward" is the left foot, whereas "low, back" is the right foot. The head says, "High, right onward." Similarly, when moving, the person usually stands on two legs (left and right), with the Laban images "low, left forward, 1" and "low, right, 1," respectively. The left and right lower arms are labelled "centre, left forward, 2" and "centre, right, 2," respectively. The left and right feet are also "low, left forward" and "low, right forward" accordingly. As a result, the visuals and human movement produced are stable, according to the examination of the first movement and the Labanotation. As a result, the Labanotation generated is virtually correct [5].

6. Conclusion

The dancing age algorithm, which is based on substantial inclining, can contain music of any genre and style. It can also reveal the dance position and true individual to whom the song is appropriate. This article has completed the associated tasks by investigating the related models [20].

It came close with a massive number of domestic and unfamiliar connected records. Furthermore, based on deep learning, it comprehended the current situation and progress pattern of dancing age algorithms. There are various

challenges with combining the current well-known dance age algorithm with the traditional dance age methodology. The difficulty in developing smooth and effortless dance stances is the main issue. The next challenge is matching dancing movements with musical progressions. Our division approach has high precision in dividing movement groupings, according to tests, and the proposed technique performs substantially better on image precision in the programmed Labanotation age. Our work has a significant impact on the recording staff's productivity [34].

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References

[1] R. Kumar, J. Sotelo, K. Kumar, A. de Robison, and Y. Bengio, "Obama net: photo-realistic lip-sync from text," 2017, https://arxiv.org/abs/1801.01442.

[2] R. K. Jati, N. A. Suwastika, and R. Yasirandi, "Hopscotch game to support stimulus in children's gross motor skill using IoT," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, vol. 4, pp. 277–290, 2020.

[3] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Networks*, vol. 18, no. 5-6, pp. 602–610, 2005.

[4] J. S. Chung, A. J. Din, and A. Fisherman, "You said that?," 2017, https://arxiv.org/abs/1705.02966.

[5] I. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial nets," in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, pp. 2672–2680, Montreal Canada, 2014.

[6] S. A. Jalalifar, H. Hasani, and H. Aghajan, "Speech-driven facial reenactment using conditional generative adversarial net work," 2018, https://arxiv.org/abs/1803.07461.

[7] D. Borra, S. Fantozzi, and E. Magosso, "Interpretable and lightweight convolutional neural network for EEG decoding: application to movement execution and imagination," *Neural Networks*, vol. 129, pp. 55–74, 2020.

[8] L. Chen, S. Srivastava, Z. Duan, and C. Xu, "Deep cross-modal audio-visual generation," in *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, pp. 349–357, Mountain View, CA, USA, 2017.

[9] V. P. Brahmaiah, Y. P. Sai, and M. N. Giriprasad, "A new framework for recognizing normal and epileptic seizure from eye movement signals using genetic based convolutional neural network," *Traitement du Signal*, vol. 37, no. 3, pp. 493–501, 2020.

[10] M. S. Hossain and M. A. Yousuf, "Real time facial expression recognition for nonverbal communication," *International Arab Journal of Information Technology*, vol. 15, no. 2, pp. 278–288, 2018.

[11] A. C. Duarte, F. Roldan, M. Tubau et al., "Wav2Pix: speech-conditioned face generation using generative adversarial networks," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, 2019.

[12] P. Koch, M. Dreier, A. Larsen et al., "Regression of hand movements from sEMG data with recurrent neural networks," in *Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 3783–3787, Montreal, QC, Canada, 2020.

[13] M. B. Alazzam, F. Alassery, and A. Almulihi, "Development of a mobile application for interaction between patients and doctors in rural populations," *Mobile Information Systems*, vol. 2021, Article ID 5006151, 8 pages, 2021.

[14] S. Ginosar, A. Bar, G. Kohavi, C. Chan, A. Owens, and J. Malik, "Learning individual styles of conversational gesture," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3497–3506, Long Beach, CA, USA, 2019.

[15] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[16] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695–4708, 2012.

[17] Y. Li, D. K. Jha, A. Ray, and T. A. Wettergren, "Information fusion of passive sensors for detection of moving targets in dynamic environments," *IEEE Transactions on Cybernetics*, vol. 47, no. 1, pp. 93–104, 2017.

[18] J. Xia, Y. Feng, L. Liu, and D. Liu, "An information fusion model of innovation alliances based on the bayesian network," *Tsinghua Science and Technology*, vol. 23, no. 3, pp. 347–356, 2018.

[19] R. D. Renard, "Creating the other requires defining Thainess against which the other can exist: early-twentieth century definitions (redefining"otherness"from northern Thailand)," *Japanese Journal of Southeast Asian Studies*, vol. 44, no. 3, pp. 295–320, 2017.

[20] O. Poutanen, S. Ylirisku, and P. Hoppu, "Technology choreography: studying interactions in Microsoft's future visoins through dance," *Human Technology*, vol. 13, no. 1, pp. 10–31, 2017.

[21] J. H. Lee, H. S. Kim, and E. S. Shin, "A kinetic analysis of "Suyangsumurepdipimu" movement for Korean dance "Cheoyong-Mu"," *Korean Journal of Sports Science*, vol. 29, no. 2, pp. 1121–1132, 2020.

[22] O. Millard, "Dance as a social practice," *Idea Journal*, vol. 17, no. 2, pp. 335–349, 2020.

[23] H. Jing and S. Ravn, "Sharing the dance – on the reciprocity of movement in the case of elite sports dancers," *Phenomenology & the Cognitive Sciences*, vol. 17, no. 1, pp. 1–18, 2018.

[24] M. B. Alazzam, F. Alassery, and A. Almulihi, "A novel smart healthcare monitoring system using machine learning and the Internet of Things," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5078799, 7 pages, 2021.

[25] M. Bader Alazzam, F. Alassery, and A. Almulihi, "Identification of diabetic retinopathy through machine learning," *Mobile Information Systems*, vol. 2021, Article ID 1155116, 8 pages, 2021.

[26] F. Qi, W. Tianjiang, L. Fang, and L. HeFei, "Research on multicamera information fusion method for intelligent

perception," *Multimedia Tools & Applications*, vol. 77, no. 12, pp. 15003–15026, 2018.

[27] M. S. Chen, L. Huang, C. D. Wang, D. Huang, and J. H. Lai, "Relaxed multi-view clustering in latent embedding space," *Information Fusion*, vol. 68, no. 9, pp. 8–21, 2021.

[28] S. Bahrami, F. Dornaika, and A. Bosaghzadeh, "Joint auto-weighted graph fusion and scalable semi-supervised learning," *Information Fusion*, vol. 66, no. 1, pp. 213–228, 2021.

[29] W. Xu, L. Jing, J. Tan, and L. Dou, "A multimodel decision fusion method based on DCNN-IDST for fault diagnosis of rolling bearing," *Shock and Vibration*, vol. 2020, no. 1, Article ID 8856818, pp. 1–12, 2020.

[30] H. L. Li and Y. F. Ma, "Target recognition method based on multi feature fusion and hybrid kernel SVM," *Shenyang GongyeDaxueXuebao/Journal of Shenyang University of Technology*, vol. 40, no. 4, pp. 441–446, 2018.

[31] Y. Cai, J. Liu, Y. Guo, S. Hu, and S. Lang, "Video anomaly detection with multi-scale feature and temporal information fusion," *Neurocomputing*, vol. 423, no. 5, pp. 264–273, 2021.

[32] A. Kaklauskas and R. Gudauskas, "Intelligent decision-support systems and the internet of things for the smart built environment," in *Start-Up Creation*, pp. 413–449, Woodhead Publishing, 2016.

[33] J. Yu and X. Bai, "Analysis of classroom learning behaviors based on internet of things and image processing," *Traitement du Signal*, vol. 38, no. 3, pp. 845–851, 2021.

[34] X. Liu, Y. Li, and Q. Wang, "Multi-view hierarchical bidirectional recurrent neural network for depth video sequence based action recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 10, article 1850033, 2018.