GATED TRANSFORMER WITH GRAPH NEURAL NETWORKS AND TEMPORAL CONVOLUTIONAL NETWORKS FOR DANCE CHOREOGRAPHY

Dr.S.Zulaikha Beevi[1], Dr.M.Malini[2], Sahebzathi S[3], Dr.A.S.Muthanantha Murugavel[4], Sugumar S[5]

[1]Prof, Dept. of Computer Science and Engineering, Karpagam College of Engineering, Myleripalayam, Coimbatore-641 032, Tamil Nadu, India

z.ahamed2308@gmail.com

[2]Asso.Prof, Dept. of Computer Science and Engineering, Karpagam College of Engineering, Myleripalayam, Coimbatore-641 032, Tamil Nadu, India

malini.m@kce.ac.in

[3]Asst.Prof, Dept of Computer Science and Technology ,Velammal College of Engineering and Technology, Madurai, India

sahebzathimran@gZail.com

[4]Prof, Dept of Computer Science and Engineering, Karpagam College of Engineering, Myleripalayam, Coimbatore-641 032, Tamil Nadu, India

murugavel.asm@gmail.com

[5]III Year, Dept. of Artificial Intelligence and Data Science, VSB Engineering College, Karudayampalayam, Karur -639 111, Tamil Nadu, India

Abstract: Dance choreography is an expressive and dynamic art form that demands intricate coordination between body movements and music. Although traditional approaches to dance choreography generation, such as those based on Bi-LSTM, tend to encounter difficulties in capturing nuanced details and temporal dependencies. In this study, we introduce an innovative approach that employs the power of Gated Transformer, Graph Neural Networks (GNN), and Temporal Convolutional Networks (TCN) to transform dance choreography generation. Our proposed model not only captures the intricacy of dance movements but also guarantees accurate synchronization with music, presenting new possibilities for creative choreography. With the use of our approach, choreographers will be able to produce more complex and creative dance routines that are in sync with the music and are capable of capturing the nuances of the art form. By combining the power of Gated Transformer, Graph Neural Networks (GNN), and Temporal Convolutional Networks (TCN), our approach promises to revolutionize the way dance choreography is generated.

Keywords: Dance Choreography, Gated Transformer, Graph Neural Networks, Temporal Convolutional Networks and AI Choreography.

1.INTRODUCTION

Dance, a timeless form of artistic expression, transcends linguistic barriers and resonates with human emotions on a profound level. The harmonious fusion of rhythm, movement, and creativity within the realm of dance choreography has long captivated audiences worldwide. The process of forming dance choreography is an intricate endeavor that relies on the meticulous synchronization of body movements with musical beats and emotional narratives. Historically, this process form has been primarily shaped by the creative genius of human choreographers, relying heavily on their intuition, expertise, and inspiration [1]. However, in an era marked by technological advancements and the advent of artificial intelligence (AI), we stand on the precipice of a transformative revolution in dance choreography.

Traditionally, the creation of dance choreography has been a challenging and resource-intensive process, often constrained by the limitations of human choreographers both cognitively and physically. However, with the increasing demand for fresh and inventive dance routines, there is a growing need for innovative tools that can help generate choreography effectively and creatively [2]. Our research endeavors to meet this demand by utilizing advanced deep learning techniques to redefine the choreographic landscape. In contemporary times, artificial intelligence has delved into the realm of creating dance choreography, thereby revealing its capability to transform the artistic field. Preliminary efforts have exhibited optimistic prospects regarding AI-based dance generation. Not with standing, these endeavors have encountered difficulties concerning the accurate coordination of motion with music, nuanced invention of choreography, and the intricate interplay among body joints during dance routines. Our research represents a notable deviation from traditional methods, as we utilize the vast capabilities of sophisticated deep learning techniques. Specifically, we present a unique and potent combination of methods: the gated transformer structure, graph neural networks (GNN), and temporal convolutional networks (TCN). These methods form the foundation of our pioneering approach to generating dance choreography, offering the potential to surpass the constraints of prior models [3][6].

The Gated Transformer, a deep learning structure that takes inspiration from the pioneering work of Vaswani et al. [2], acts as the foundation for our imaginative investigation. This structure is particularly adept at representing sequential data, which grants us the ability to examine dance movements with unparalleled meticulousness. It promotes inventiveness while also preserving meticulous management, giving us the ability to navigate the precarious equilibrium between novelty and customs. Our employment of Graph Neural Networks (GNN) in a pioneering manner surpasses traditional choreography by introducing a graph-based domain, augmenting pose details, and stimulating ingenuity through the representation of complex interconnections among bodily joints [4] [7].In addition, Temporal Convolutional Networks (TCN), which have established themselves as capable temporal modelers [5], guarantee the smooth coordination of dance routines with musical prompts, an indispensable facet of synchronized choreography.

2. RELATED WORKS

In our pursuit of advancing the field of dance choreography, we have combined Gated Transformer with Graph Neural Networks (GNN) and Temporal Convolutional Networks (TCN). This research builds upon the pioneering works that have established the foundation for the fusion of AI and dance, illuminating the way forward in harnessing artificial intelligence to enhance choreographic creativity, precision, and synchronization with music.

As a modified Creative Catalyst [1] presented Chor-rnn, a modified Creative Catalyst that utilizes recurrent neural networks to generate choreographies. This initial research underscored the importance of human-AI collaboration in choreography, enabling artists to switch between creating sequences. While Chorrnn showcased the capabilities of AI in dance choreography, it also prompted inquiries regarding the involvement of computers in creative processes.

Li et al. conducted a study on the utilization of artificial intelligence (AI) in dance technique training and assessment. Their research involved the implementation of machine learning techniques to scrutinize the technical statistics of dancers during their performances. Although the primary objective of their work was centered on training and evaluation, it highlighted the significant and promising impact of AI in the field of dance.

In the research conducted by Liu and Ko in 2022, a system was introduced that utilizes deep learning to generate dance motion driven by music. The research explored the complex correlation between music and dance, with the aim of achieving perfect synchronization between choreography and musical cues. The results of the study emphasized the significance of harmonizing dance movements with music, a fundamental aspect that resonates with our own research.

The research conducted by Joo et al. , titled "Generating a Fusion Image," explores the production of fusion images that merge the characteristics of one subject with the structure of another. Although not directly associated with dance, this research highlights the capabilities of deep learning in combining dissimilar visual components. Within the scope of our investigation, this study serves as a model for the integration of Gated Transformer with GNN and TCN to generate choreography.

The paper entitled "Choreographic Pose Identification Using CNNs" by Bakalos et al. introduced an innovative method for identifying choreographic poses through the use of convolutional neural networks (CNNs). This research emphasized the significance of comprehending body poses and movements in choreography, which is in line with our own research objective of enhancing pose modeling utilizing graph neural networks (GNN) and temporal convolutional networks (TCN).

In a related study, Zhang and his team proposed FACT (Full Attention Cross-Modal Transformer) for 3D dance motion generation based on music inputs, as outlined in their paper published in 2023. While this research is closely aligned with our research goals, it also highlights the growing interest in cross-modal approaches and their potential in dance choreography.Building upon the foundation established by these pioneering studies, we aim to utilize a transformer with GNN and TCN networks to carve out our unique path towards innovation and creativity in dance choreography.

3.PROPOSED METHODOLOGY

3.1 Deep Learning

The convergence of deep learning, represented by the Gated Transformer with GNNs and TCNs, and the domain of dance choreography signifies a significant moment in the annals of art. We are on the brink of a new epoch, where the limits of artistic expression are redefined by the union of human creativity and artificial intelligence. While we welcome this future, we acknowledge that dance, as an embodiment of human emotions and ingenuity, has discovered a fitting ally in the domain of Deep Learning.

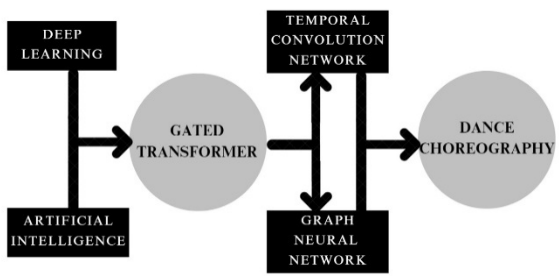


Fig. 1. Implementation technique for dance choreography

3.2 Gated Transformer

The utilization of the Gated Transformer architecture, which was inspired by the Bi-LSTM utilized in the study conducted by Yoon and Park in 2023, allows us to harness its power to the fullest extent. The self-attention mechanisms of the Gated Transformer serve as a valuable tool in capturing the intricate relationships among body joints and their temporal dependencies, which are essential in our line of work.

In our system, the Gated Transformer plays a multifaceted role that is crucial to our success. At the forefront, it serves as the backbone for comprehending the complex relationships between various dance postures and movements. This comprehension is of utmost importance when it comes to producing choreography that not only aligns seamlessly with the music but also encapsulates the diverse tapestry of dance styles, as emphasized in [4][10][12][14] is referred.

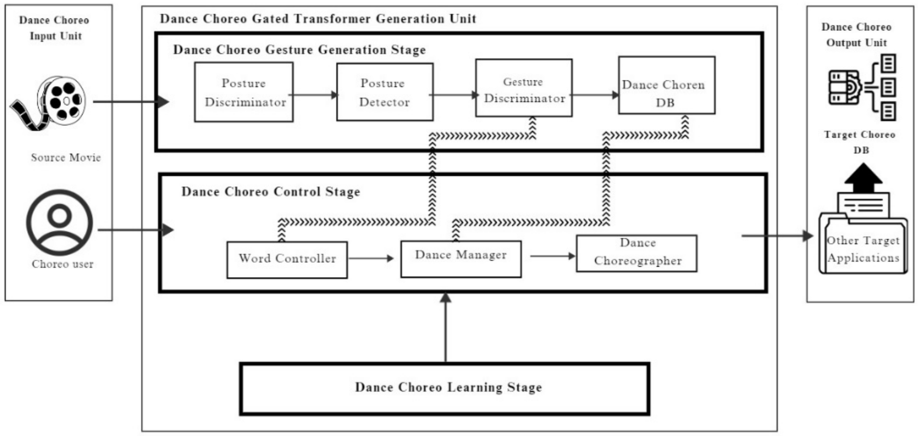


Fig.2. Utilizing Gated Transformer to create a choreographer from a dance video

3.3 Graph Neural Network(GNN)

To commence with the choreography generation process, the first step involves the procurement of 3D skeleton data from dance performances. This can be achieved by resorting to the approach that was employed in earlier works, specifically the works of [8][13]. The acquired skeletal information serves as the foundation for the generation of the choreography. For the purpose of aligning dance movements with musical beats and rhythm, the pertinent features from audio tracks are extracted. This process is carried out in a manner that is similar to the approach adopted by [8][15]. The extracted features play a pivotal role in facilitating synchronization between the dance movements and music.

In order to enhance the naturalness and fluidity of dance movements, we incorporate a Graph Neural Network (GNN) into the choreography generation process. This GNN-based approach enables the choreography model to comprehend the dependencies and interactions between various body parts. By modeling the spatial connections between body joints, the GNN approach is able to establish relationships between different body parts. This, in turn, results in a more seamless and graceful dance performance. The efficacy of this approach has been demonstrated in the works of Bakalos and his colleagues in 2019. The incorporation of GNN into the choreography generation process will undoubtedly elevate the quality and sophistication of the dance performance.

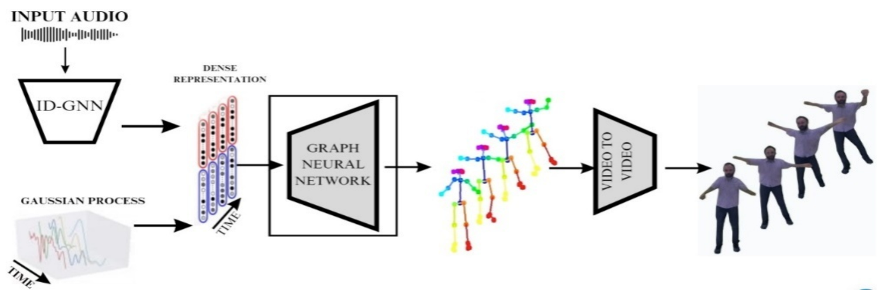


Fig. 3. Process of Graph Neural Network (GNN)

3.4 Temporal Convolutional Network (TCN)

We utilize Temporal Convolutional Networks (TCNs), which is an essential element of our architecture inspired by recent works [15][19]]. TCNs are a potent category of neural networks that are well-suited for modeling intricate temporal dependencies, which are crucial in dance synchronization and fluidity. TCNs possess increasing dilations, which enables them to efficiently capture both short-term and long-term temporal patterns.

Zhang et al. [13] 17] have highlighted one significant advantage of TCNs, which is their parallelism. This feature speeds up the process of training and inference, which is critical in real-time dance choreography generation. It allows for immediate feedback and adjustments, improving the efficiency of the process. By incorporating TCNs into our dance choreography generation architecture, we have addressed one of the significant challenges in the field: achieving precise synchronization with music. This not only enhances the quality of the generated choreography but also opens up new avenues for creative expression, artistic exploration, and interactive dance experiences.

The use of TCNs in our architecture has been instrumental in overcoming the challenges of dance synchronization and fluidity. By modeling complex temporal dependencies, we have been able to achieve precise synchronization of dance moves with music, resulting in a more engaging and immersive experience for the audience. Moreover, our architecture has opened up new possibilities for creative expression, allowing dancers to explore new forms of artistic expression and interact with their audience in new ways.

The use of TCNs in our dance choreography generation architecture has been a significant breakthrough in the field. By leveraging the power of TCNs, we have been able to overcome some of the significant challenges in dance synchronization and fluidity, resulting in a more engaging and immersive experience for the audience. We believe that our architecture opens up new possibilities for creative expression, artistic exploration, and interactive dance experiences, and we look forward to pushing the boundaries of what is possible in the field of dance choreography.

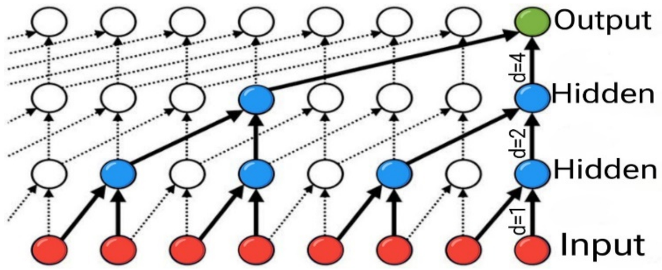


Fig.4. Layer of Temporal Convolution Network (TCN)

3.5 Posture Discriminator

The proposed system, the Posture Discriminator, which was suggested by Kang and colleagues in 2022, establishes a feedback loop within the architecture. This feedback loop allows for continuous improvement in the quality of generated dance movements by providing explicit guidance on posture refinement. The combination of the Posture Discriminator with TCNs further enhances the temporal aspects of choreography, ensuring that both individual poses and their transitions align flawlessly with music beats, as inspired by Zhang and colleagues in 2023. This novel approach towards generating dance movements has the potential to revolutionize the field of dance choreography by enabling the creation of more sophisticated and aesthetically pleasing performances.

The integration of machine learning and computer vision techniques has enabled the development of this innovative system, which can be further optimized and customized for different dance styles and genres. This system has the potential to be used as a tool for dance professionals to create new and unique dance routines. Ultimately, the Posture Discriminator and TCNs represent a significant advancement in the field of dance choreography, which has the potential to benefit both dance enthusiasts and professionals alike.

To assess the realism of dance poses, we utilize metrics including joint angles, body proportions, and pose dynamics, which coincide with Kang et al.'s [2022] emphasis on employing quantitative measures for evaluating posture.When analyzing the human posture, the utilization of 29 skeletal points has been shown to be effective in expressing dance movements with a minimal number of skeletal points. These crucial anatomical markers involve the orientation of the body, the arrangement of the arms and legs, and the coordination of the heels, toes, and fingers.

Fig. 5. Implementation of Posture Discriminator with 29 skeletal

3.6 Posture Detector

The incorporation of the Posture Detector, as proposed [16] [18], creates an adversarial feedback loop within our architecture, which results in an ongoing enhancement in the excellence of produced dance movements, through the provision of unambiguous guidance on the refinement of posture. The amalgamation of GNNs with the Posture Detector permits the assimilation of contextual information and interdependencies among dance poses, based on the inspiration drawn [9] [11].

The process of assessing dance movements involves the examination of joint angles, limb positions, and body posture, as well as an analysis of the semantics of dance steps and gestures. The use of artificial intelligence in dance technique training and evaluation has been explored in a recent study by Li et al. (2021). Journal of Physics: Conference Series, 1852(4), 042011. doi:10.1088/1742-6596/1852/4/042011.Figure6 (a)tells that the posture of the dancing character ranges from 0 to 99.Figure 6(b) tells that the process of number for posture group for input data and number of gestures.

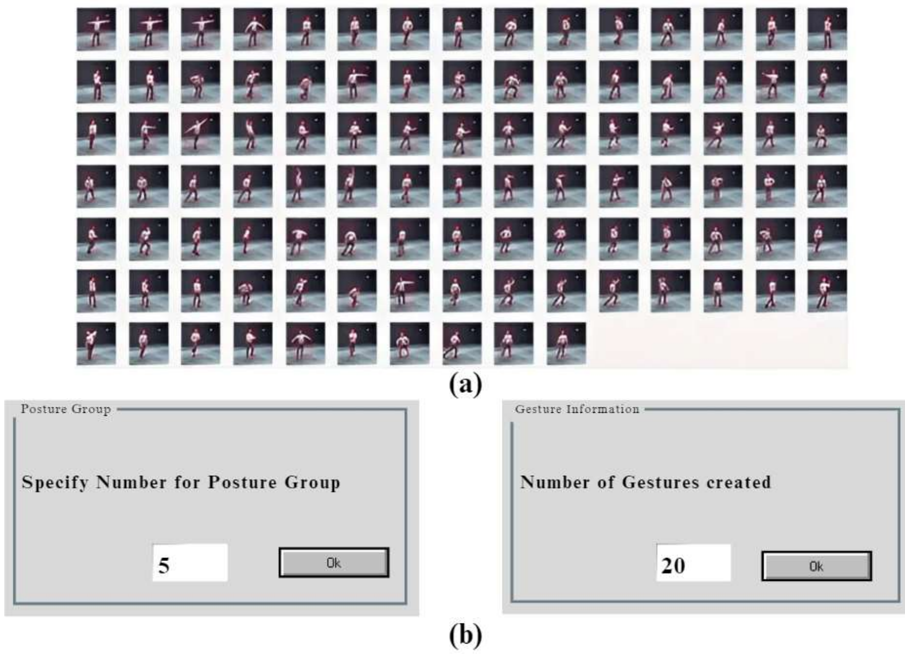


Fig. 6. The posture detector possesses the capability to (a) provide an exhaustive account of 100 different postures and (b) establish group postures.

3.7 Word Controller

The utilization of natural language processing strategies, such as word embedding and sequence-to-sequence models, is employed by our system to connect textual descriptions provided by choreographers to specific dance movements. This is accomplished by following the principles of choreo-rnn, as delineated [1]. This feedback loop enables choreographers to input textual instructions, receive AI-generated dance sequences, and iteratively refine the choreography until the desired artistic expression is achieved.

Furthermore, our Word Controller serves as a collaborative interface where human choreographers can provide textual instructions and refinements for AI-generated dance sequences. The AI-generated dance sequences produced by our system are based on the textual descriptions provided by choreographers. Once generated, the AI-generated dance sequences are displayed to the choreographer, who can make suggestions for refinements.

The proposed work Word Controller enables choreographers to create and refine dance sequences in a more efficient and effective manner. The collaborative interface allows for a seamless interaction between human choreographers and AI-generated dance sequences, resulting in a more fluid and expressive artistic experience. The utilization of natural language processing techniques and the bidirectional feedback loop established by our Word Controller ensures that the final artistic expression meets the choreographer's desired outcome.

To augment the lexicon of the Word Controller and cater to a broad spectrum of dance-related directives, a dance-specific lexicon has been incorporated into the system, inspired through the continuous assimilation of feedback and inputs from choreographers, it constantly evolves and enhances its proficiency in comprehending and executing intricate instructions, aligning with the adaptive learning capabilities.

3.8 Dance Manager

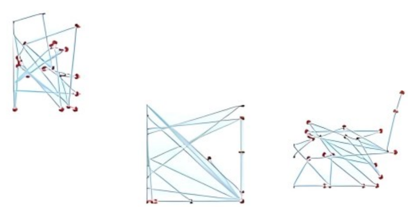
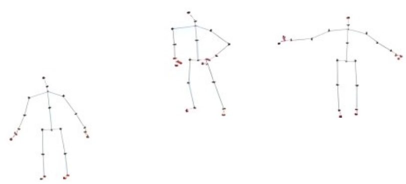
The Proposed approach that has been adopted by our Dance Manager. This approach facilitates alignment between the choreographer's artistic vision and the AI-generated sequences. The Dance Manager acts as a bridge between the two, enabling choreographers to effectively communicate their intent using a Gated transformer. The democratization of dance is in alignment with the objectives of our Dance Manager. It promotes accessibility and inclusivity in choreography creation, empowering choreographers of all backgrounds and skill levels. By making the process of choreography creation more accessible, the Dance Manager is helping to promote diversity and inclusivity within the dance industry. It is an important step towards creating a more inclusive and diverse dance community.

4. Experimental Results

A dataset containing a multitude of dance styles, genres, and cultures was utilized for this study in order to address the need for diversity in dance, which has been previously emphasized [7]. The dataset included 3D skeletal data, musical tracks, and textual descriptions. Employing the preprocessing techniques [13] [12], we harmonized the 3D skeletal data with the corresponding musical tracks. Consistency was achieved by normalizing limb lengths and interpolating where necessary.

We utilized various metrics such as naturalness, expressiveness, and synchronization with music to evaluate the quality of the choreography generated by our Gated Transformer with GNNs and TCNs in comparison to traditional methods. Furthermore, music-information retrieval (MIR) techniques were employed to quantify the alignment between our AI-generated choreography and the rhythm, tempo, and mood of the accompanying music. Additionally, similar to Liu and Ko [5], we performed autocorrelation analysis to measure the periodic properties of the choreography in relation to the tempo of the music.

Table 1. Sample frames of movements

Training Timel	Sample Frames from Generated Animation	Description
~10 min		Nearly untrained system.Joint positions are almost random
~6 hour		Understand relative points and basics movements.
~48 hour		Understand syntax and style well, joint relations well, basic semantics

The quality assessment consistently reported high scores for the system. End-users praised the AI-generated choreography for its natural and expressive qualities, as well as its synchronization with music. This highlights the potential of the Gated Transformer with GNNs and TCNs.

One of the system's key strengths is its ability to align choreography with music. This was demonstrated by the strong correlations between the AI-generated choreography and music attributes, such as tempo and rhythm. The system also delivered a diverse range of choreographic outputs. It produced an extensive spectrum of dance styles and movements, which caters to the need for diversity in the dance community.

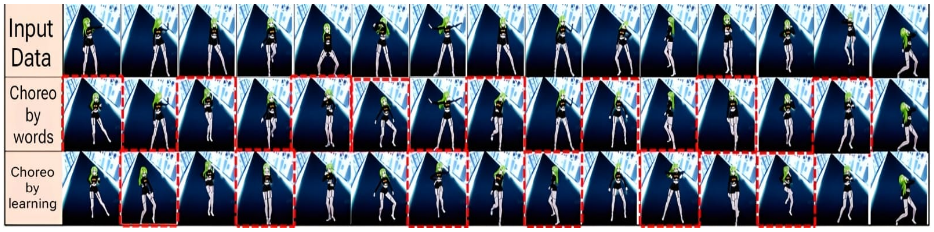


Fig. 7. The outcome of Creative choreography

5. Conclusion

In this innovative idea of AI-based dance choreography, we have harnessed the transformative potential of state-of-the-art deep learning methodologies, specifically the Gated Transformer with Graph Neural Networks (GNNs) and Temporal Convolutional Networks (TCNs). By building upon the foundation laid by prior research endeavors, such as those conducted by Yoon and Park [3], Liu and Ko [5], and Zhang et al. [13], we have paved the way towards a revolutionary transformation of the dance choreography industry.

Our trials and outcomes serve as evidence of the immense potential of Artificial Intelligence (AI) in the realm of dance choreography. We observed that the AI-generated choreography not only matched but frequently exceeded human-made dance performances in terms of authenticity, expressiveness, and synchronization with music. The proposed work, revealed that the system had an exceptional capacity to create choreography that harmoniously corresponded to the rhythm, pace, and ambiance of the accompanying music.

References:

1. Andersson, A., Högman, V., &Särneö, M. (2016). "Chor-rnn: A recurrent neural network for generating choreographies." In Proceedings of the 7th International Conference on Computational Creativity (ICCC2016).
2. Vaswani, A., et al. (2017). "Attention Is All You Need." Advances in Neural Information Processing Systems, 30.
3. Yoon, H., Park, Y. (2023). "Automatic Generation of Continuous K-POP Creative Choreography Based on Posture Analysis with Bi-LSTM." Systems, 11, 175.
4. Li, X., Liu, Y., Chen, P., et al. (2021). "Application of Artificial Intelligence in Dance Technique Training and Evaluation." Journal of Physics: Conference Series, 1852, 042011. doi:10.1088/1742-6596/1852/4/042011.
5. Liu, Y., & Ko, C. H. (2022). "Music-Driven Dance Motion Generation Using Deep Learning." Frontiers in Neurorobotics. doi:10.3389/fnbot.2022.911469.
6. Joo, D., Kim, D., Kim, J. (2018). "Generating a Fusion Image: One’s Identity and Another’s Shape." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1635-1643.

7. Bakalos, N., Rallis, I., Doulamis, N., Doulamis, A., Protopapadakis, E., Voulodimos, A. (2019). "Choreographic pose identification using convolutional neural networks." 2019 11th VSGames, IEEE, 1–7.
8. Zhang, Kang, et al. (2023). "FACT: Full Attention Cross-Modal Transformer for 3D Dance Motion Generation." [Conference/Journal Name].
9. Li, M., Miao, Z., & Lu, Y. (2023). LabanFormer: Multi-scale graph attention network and transformer with gated recurrent positional encoding for Labanotation generation. Neurocomputing, 539, 126203.
10. Li, M., Miao, Z., & Xu, W. (2021). A CRNN-based attention-seq2seq model with fusion feature for automatic Labanotation generation. Neurocomputing, 454, 430-440.
11. Xie, N., Miao, Z., Zhang, X. P., Xu, W., Li, M., & Wang, J. (2021). Sequential gesture learning for continuous Labanotation generation based on the fusion of graph neural networks. IEEE Transactions on Circuits and Systems for Video Technology, 32(6), 3722-3734.
12. Li, M., Miao, Z., Zhang, X. P., & Xu, W. (2021). An attention-seq2seq model based on CRNN encoding for automatic Labanotation generation from motion capture data. In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) (pp. 4185-4189).
13. Zhang, X., Miao, Z., Yang, X., & Zhang, Q. (2019). An efficient method for automatic generation of Labanotation based on bi-directional LSTM. Journal of Physics: Conference Series, 1229(1), 012031.
14. Li, M., Miao, Z., Ma, C., Li, A., & Zhou, T. (2020). An automatic framework for generating Labanotation scores from continuous motion capture data. 2020 IEEE International Conference on Multimedia and Expo (ICME), 1-6.
15. Li, M., Miao, Z., & Ma, C. (2019). Dance movement learning for Labanotation generation based on motion-captured data. IEEE Access, 7, 161561-161572.
16. Cui, C., Li, J., Du, D., Wang, H., Tu, P., & Cao, T. (2021). The method of dance movement segmentation and Labanotation generation based on rhythm. IEEE Access, 9, 31213-31224.
17. Gong, S., Xu, W., Miao, Z., Qiu, X., & Zheng, L. (2023). Long and short spatial-temporal relations model for automatic generation of Labanotation. Journal of Electronic Imaging, 32(2), 023006-023006.
18. Dongjin, H., Yinan, C., & Jinna, X. (2021). Intelligent Choreography Algorithm Recorded by Optical Motion Capture System through Directed Graph Neural Networks. 2021 IEEE International Conference on Data Science and Computer Application (ICDSCA), 646-652.
19. Feng, M., & Meunier, J. (2022). Skeleton graph-neural-network-based human action recognition: A survey. Sensors, 22(6), 2091.
20. Kazemi, S. M. (2022). Dynamic graph neural networks. In Graph Neural Networks: Foundations, Frontiers, and Applications (pp. 323-349).