© February 2025 | IJIRT | Volume 11 Issue 9 | ISSN: 2349-6002

AI-Driven Dance Generation and Music Composition System

Anshu Sakhare[1], Lekha Pulavarthy[2], Riya Kamble[3], Shruti Parade[4], Prof. Amit Narote[5]
[1,2,3,4] Student, Department of Information Technology, Xavier Institute of Engineering.
[5] Professor and Guide, Department of Information Technology, Xavier Institute of Engineering.

***Abstract*—The field of AI-generated music and dance is experiencing rapid growth, employing deep learning and generative models to produce synchronized sound and movement. This research investigates recent innovations in AI-based music continuation and choreography creation, with an emphasis on sequence modelling and motion prediction methods. We investigate cutting-edge approaches for extending musical pieces using LSTM-based models and creating dance movements in response to input music through machine learning algorithms. Primary areas of investigation include data pre-processing techniques, music feature extraction, and motion synchronization strategies. A comparative evaluation of neural networks for music and dance generation assesses model efficiency, synchronization precision, and artistic consistency. Additionally, we examine multimodal fusion techniques that improve the integration of auditory and visual signals, ensuring smooth and natural dance choreography. This study aims to produce dynamic dance movements from given music samples and create extended compositions from brief musical segments, utilizing deep learning techniques in sequence modelling.**

***Index Terms*—AI-generated choreography, Music composition, Deep learning, AIST++, Long Short-Term Memory (LSTM), Virtual Reality (VR), Motion Capture.**

I. INTRODUCTION

Music and dance have long been regarded as highly creative fields requiring extensive training and human creativity. However, recent AI advancements have opened up possibilities for automating and customizing this artistic endeavor. This research presents an AI-driven system that allows users to input a brief musical sample, which the model then expands using LSTM - based music generation, or provide a musical piece to create synchronized dance movements. The system utilizes deep learning frameworks like TensorFlow and Keras for effective sequence modelling, alongside Librosa for extracting audio features from Dutschl and AIST++ for generating motion. By connecting sound and movement through AI, this platform simplifies the processes of music composition and dance creation, making artistic expression more accessible and interactive for a diverse user base.

II SYSTEM ARCHITECTURE

The AI-powered Music and Dance Creation System offers users two main interaction options: submitting a music sample to create matching dance routines or providing a brief musical excerpt for AI-based music expansion and composition. This system utilizes advanced machine learning techniques to examine, produce, and enhance both audio and choreographic outputs. For choreography creation, the system initially examines the input music to detect key rhythmic and tempo elements. A Convolutional Neural Network (CNN) extracts musical characteristics, while a Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) forecasts coordinated dance movements based on these features. These movements are subsequently applied to a digital dancer model, ensuring the choreography naturally aligns with the music's flow. In terms of music composition, the system employs LSTM or Transformer-based models to generate melodies and harmonies that build upon the user-provided musical snippet while preserving coherence in structure and style. The generated audio then undergoes post-processing methods, such as pitch adjustment and harmony synchronization, to improve its overall quality. By incorporating cutting-edge deep learning models, this system delivers a smooth and interactive platform for users to explore AI-assisted music composition and dance routine generation.

III. LITERATURE SURVEY

DanceNet3D: Music-Based Dance Generation concentrates on creating key poses from musical beats, forecasting motion curves between these poses, and utilizing an encoder-decoder structure. The system incorporates adversarial training to enhance the model, while MoTrans employs Kinematic Chain Networks and Learned Local Attention. Training is conducted using the PhantomDance dataset, which offers synchronized dance information. This method produces smooth and elegant dance performances, achieving superior music synchronization compared to earlier approaches. The motion curve representation improves performance fluidity, and the two-stage framework yields a high beat consistency score. Nevertheless, the model's deterministic nature restricts motion variety, and imprecise 2D keypoint detection impacts 3D reconstruction quality.

Learn to Dance with AIST Music-Conditioned 3D Dance employs a cross-modal transformer-based neural network to understand audio-motion relationships for creating 3D motion. The AIST dataset is developed to facilitate conditional 3D motion generation, and user studies assess the model's realism and compare it to baseline techniques. This approach generates lifelike 3D motion sequences from music, effectively learns audio-motion correspondence, and prevents freezing or drifting during motion creation. It also produces diverse dance movements for various music inputs and utilizes a large, multi-genre 3D dance dataset. However, the absence of 3D ground-truth data in the AIST dataset affects validation, and the model experiences freezing after multiple iterations in practice.

EDGE: Editable Dance Generation utilizes a transformer-based diffusion model to produce realistic, extended dance sequences based on music input. This model offers joint-wise and temporal editing capabilities, and auxiliary losses are employed to enhance physical realism. EDGE generates realistic, physically-plausible dance motions and provides powerful editing features for dance creation, including joint-wise conditioning and in-betweening flexibility. A novel metric for physical plausibility is introduced to evaluate the generated motions. However, automated dance quality metrics are complex and subjective, and the current FID metrics for dance evaluation are unreliable. Additionally, the AIST dataset is insufficient for comprehensive evaluation due to its limited size.

The AI-Based Dance Generation Applications Review examines diverse AI systems, including Chor-RNN, GrooveNet, and AI Choreographer, which employ multimodal learning techniques to scrutinize and produce dance movements using audio and visual inputs. These systems facilitate the creation of intricate dance sequences and real-time dance generation driven by music. Artificial intelligence enhances choreographic innovation and accessibility, while multimodal learning techniques improve the comprehension of dance patterns. Virtual reality tools provide interactive dance experiences, and AI models simplify the dance creation process for users. Nevertheless, the training of dance models remains unstable and intricate, and the limited availability of training data impacts the maturity of dance generation systems. Furthermore, these models face challenges in generalizing beyond their training datasets.

The Music-Driven Dance Generation methodology employs an LSTM-SA Seq2Seq network to produce dance movements from musical sequences. It introduces a trio of evaluation criteria for model assessment and aims to generate natural and diverse dance sequences. This model effectively maps music to dance movements and utilizes LSTM for stable training and performance, incorporating an attention mechanism to enhance sequence representation. However, LSTM's compression of the entire input sequence into a fixed representation restricts flexibility. Moreover, the attention model lacks parallelization, which increases training duration, and there is a scarcity of research on evaluating dance sequences.

IV. SYSTEM MODELS

The platform generates a coordinated dance routine that corresponds to the musical input's rhythm, beats, and tempo. It employs signal processing methods to detect key musical components, including tempo (BPM), rhythmic patterns, beat locations, and instrument sections with volume ranges [5][4]. A pretrained deep learning model connects these identified musical features to appropriate dance moves, drawing from a comprehensive dance move database. This model ensures the choreography is in sync with the input music's pace, genre, and atmosphere. To produce smooth and realistic dance movements, a pose estimation algorithm is utilized to generate a sequence of body joint positions. The dance routine is divided into segments, incorporating both existing and AI-generated movements.

Users can also use the system to extend short musical inputs into longer compositions through AI-driven music continuation. The platform identifies and analyses musical characteristics by recognizing patterns, learning from KenScores to detect melodic, harmonic, and rhythmic elements. Long Short-Term Memory Networks (LSTMs) are used to predict and create new musical sequences that naturally extend the input. The AI ensures the generated music maintains the input snippet's style, tempo, and structure while introducing variations for organic progression [3]. Furthermore, the produced music expands dynamically based on pattern complexity and inferred structure, maintaining consistency and delivering a coherent, engaging result.

V. ALGORITHM

A. Convolutional Neural Network for Dance Routine generation

The system employs a Convolutional Neural Network (CNN) to analyse musical rhythms and beat patterns, subsequently creating synchronized dance sequences. Dance Sequence Generation Process:

1. Information Gathering – The system learns from a database of dance movements labelled with their associated musical beats and tempos.
2. Feature Extraction – The CNN identifies elements like tempo, rhythmic intensity, and beat changes from the input music.
3. Genre Identification – The system matches extracted features to predefined dance categories (such as hip-hop, ballet, freestyle).
4. Pose Estimation – The model generates dance moves by predicting key body positions at various beats.
5. Routine Construction – The final dance sequence is assembled, ensuring smooth transitions between movements. The system ensures real-time dance routine creation, adjusting movements based on the complexity and speed of the given music.

B. Musical Composition

The system uses Deep Learning methods, specifically Long Short-Term Memory (LSTM) networks, to produce extended musical pieces based on a brief musical input. The platform examines the provided segment, recognizes its harmonic, rhythmic, and melodic patterns, and creates a coherent continuation of the music. AI Music Creation Steps:

1. Input Analysis – The user provides a short musical segment as input, which acts as the foundation for generating an extended composition.
2. Feature Recognition – The model extracts key musical characteristics such as tempo, chord progression, note sequences, and rhythm from the input segment.
3. Music Synthesis – An LSTM-based neural network predicts and generates the subsequent sequence of musical notes while maintaining consistency with the original input.
4. Audio Production – The generated music is synthesized using MIDI or WAV generation techniques, ensuring the output is a fully playable and musically coherent composition.

VI. CONCLUSION

This research introduced an artificial intelligence-based system that automatically generates dance choreography from musical input and composes music based on provided music bit. The proposed system aims to employ deep learning methods to examine musical characteristics and produce synchronized dance movements while ensuring musical compositions are both coherent and innovative. This investigation establishes a basis for creating an intelligent system capable of comprehending musical components and converting them into expressive dance sequences. Subsequent research will concentrate on developing and training deep learning models, enhancing performance, and improving user experience through customizing choreography styles and music generation.

REFERENCES

[1] S. Lee, J. Park, and S. Yoon, "Music2Dance: Learning Music-Conditioned 3D Dance

Generation with Score-Based Diffusion," IEEE Transactions on Visualization and Computer Graphics, 2023.
[2] M. Luo, W. Zheng, and H. Liu, "Danceformer: Music-Driven Dance Generation with Transformer-Based Motion Prediction," arXiv preprint arXiv:2301.09812, 2023.
[3] T. Wang, J. Zhang, and Y. Jiang, "Transformer-GAN for Symbolic Music Generation," IEEE Access, 2022.
[4] J. Li, C. Xu, and T. Shiratori, "DanceNet3D: Music-Conditioned 3D Dance Generation with Parametric Motion," arXiv, 2020.
[5] C. Jacob et al., "Dance Generation with AI: A Comprehensive Review," Frontiers in Robotics and AI, 2021.
[6] C. Xu, T. Shiratori, and J. Li, "Learning to Generate Dance Movements from Music: A Deep Neural Network Approach," ACM Transactions on Multimedia Computing, Communications, and Applications, 2020.
[7] S. Zhang, Y. Liu, and L. Xing, "DeepDance: Music-Driven Dance Motion Generation Using Deep Learning," IEEE Transactions on Circuits and Systems for Video Technology, 2021.
[8] J. Zhang, M. S. Guralnick, and R. A. K. Ranjan, "MIDI-based Music and Dance Generation with RNNs," Proceedings of the 2021 International Conference on Machine Learning, 2021.
[9] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," International Conference on Learning Representations, 2014.
[10] H. Kim, C. Zhang, and P. L. Bartlett, "Unsupervised Music Generation with Variational Autoencoders," Journal of Machine Learning Research, 2018.
[11] S. Liao, M. Li, and Z. Zheng, "AI-Assisted Music and Dance Generation for Virtual Performances," IEEE Transactions on Artificial Intelligence, 2022.
[12] A. Patel, D. Huang, and S. Song, "Music-to-Movement: Predicting Dance Poses from Music with a Cross-modal Generative Model," IEEE Transactions on Neural Networks and Learning Systems, 2023.
[13] S. Xu, C. He, and W. Zhang, "Deep Music-to-Dance Generation with Neural Networks," Journal of Artificial Intelligence Research, 2022.
[14] X. Zhang, J. Liu, and Y. Sun, "Using Generative Adversarial Networks for Dance Pose Generation from Music," ACM SIGGRAPH Conference on Computer Graphics and Interactive Techniques, 2022.
[15] T. Zeng, M. Liu, and Y. Wang, "Dance-to-Music: A Survey on Dance Generation from Music Data," IEEE Transactions on Multimedia, 2022.
[16] X. Wei, H. Tian, and F. Liu, "Music-Driven Dance Generation with Temporal Convolutional Networks," IEEE International Conference on Acoustics, Speech, and Signal Processing, 2022.
[17] J. Chen, R. Wang, and J. Lu, "Cross-Modal Music and Dance Generation Using Transformers," arXiv preprint arXiv:2206.09051, 2022.
[18] M. Lee, Y. Shin, and S. Kim, "Music-Driven Dance Generation Using Reinforcement Learning," International Conference on Machine Learning, 2021.
[19] L. Chang, X. Zhang, and Z. Liu, "Real-Time Music-Driven Dance Animation Generation Using Deep Learning," IEEE Transactions on Visualization and Computer Graphics, 2021.
[20] L. Ma, X. Li, and T. Jiang, "Dance Generation Using Music Style and Mood Analysis," ACM Computing Surveys, 2021.