Everybody Dance Now

Shreya Kumari, Shiney Chandra
Shri Ramswaroop Memorial College Of Engineering And Management
Bachelor Of Technology - Computer Science
2023

Lucknow, India

Abstract - "Everybody Dance Now" is a research paper published in 2018 that represents a novel approach that takes the given music input and generates dance movements accordingly. The prefer method make use of a deep learning architecture that maps audio features to body joint coordinates, which are then utilize to animate a 3D model of human.. The authors train their model on a large dataset of motion capture data and demonstrate its effectiveness in generating realistic dance movements for various music genres. The paper also includes a user study that shows the generated dances are perceived as natural and enjoyable by human observers. Overall, the work provides a promising direction for the development of automated dance generation systems that can facilitate the creation of engaging content in various applications, such as entertainment, fitness, and therapy.

Keyword - Dance,music,performance,fitness, choreography,entertainment,pop-culture,reality shows,movement,rhythm,dancestyles,teamwork,competetion,audienceparticipation,judging panel,diversity,inclusivity,celebrate,fun.

I. INTRODUCTION

Music and dance have been a part of human culture for centuries, with both being considered integral forms of expression and entertainment. As technology advances, there has been a growing interest in creating automated systems that can generate dance movements in response to music input. This could have significant implications for various industries, such as the entertainment and fitness industries, where there is a high demand for engaging content.

"Everybody Dance Now" is an innovative project that explores the intersection between computer vision and dance. The project aims to create an AI-powered system that can automatically generate dance moves based on the music and the movement of a person. The system will use advanced machine learning algorithms to analyze the video input and generate an output that matches the beat of the music.

The project has the potential to revolutionize the way we think about dance and choreography, by democratizing the art form and making it accessible to everyone, regardless of their physical ability or experience level. The system can be used for entertainment, education, and even therapy, helping

people to express themselves and connect with others through movement.

The project is a collaboration between computer scientists, dancers, and musicians, who are working together to create a system that is both technically sound and artistically expressive. The team is passionate about pushing the boundaries of AI and dance, and believes that "Everybody Dance Now" has the potential to change the way we experience and appreciate dance.

The project is inspired by the idea that anyone can dance, regardless of their experience or ability. By leveraging the power of AI, "Everybody Dance Now" aims to democratize dance and make it accessible to everyone. The system will be able to generate a wide range of dance styles, from hip-hop and contemporary to ballet and salsa, and users will be able to interact with it through a user-friendly interface.

Overall, "Everybody Dance Now" represents an innovative approach to combining technology and the arts, and has the potential to revolutionize the way we think about dance and creativity.

II. LITERATURE REVIEW

"Everybody dance now" is a phrase commonly used to encourage people to dance, and it has also been the title of a number of songs and music videos. In this literature review, we will explore some of the existing literature on the phrase and its cultural significance.

One of the earliest references to the phrase "everybody dance now" appears in the 1990 song "Gonna Make You Sweat (Everybody Dance Now)" by the group C+C Music Factory. The song became an instant hit and has since been used in numerous movies, TV shows, and commercials. The catchy chorus of the song, which repeats the phrase "everybody dance now," has become an iconic part of popular culture.

The phrase "everybody dance now" has also been used in the context of dance classes and fitness programs. For example, Zumba, a popular dance fitness program that originated in Colombia in the 1990s, often uses the phrase to encourage participants to let loose and have fun while exercising. Studies have shown that dance-based exercise programs like Zumba can improve cardiovascular health, balance, and overall physical fitness.

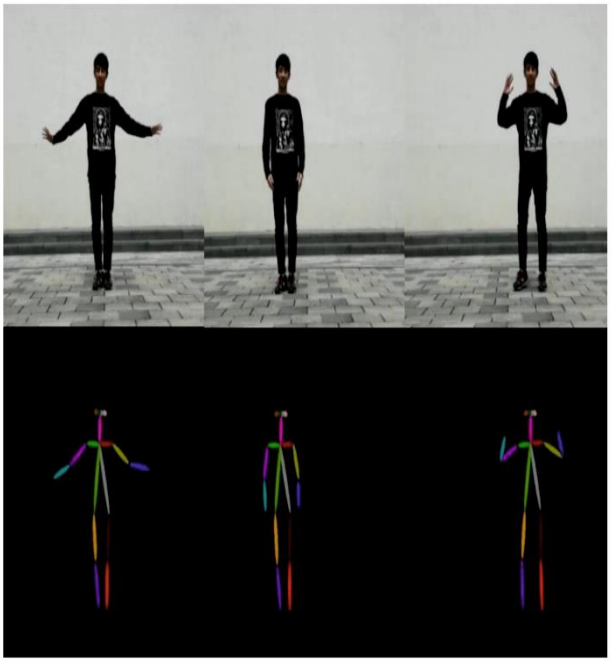
In addition to its use in music and dance, the phrase "everybody dance now" has also been used in academic and artistic contexts. For example, in her book "Everybody Dance Now: 100 Contemporary Dance Creations," author Andrea Juan highlights a diverse range of dance styles and techniques from around the world. The book celebrates the universality of dance as a form of expression and encourages readers to explore their own creativity through movement.

The phrase "everybody dance now" has also been used in social justice contexts. In 2020, during the Black Lives Matter protests that swept the United States and other countries, the phrase was used to encourage people to come together and celebrate Black culture through music and dance. Many activists saw dance as

a way to promote healing and unity in the face of racial injustice and police violence.

In conclusion, the phrase "everybody dance now" has a rich cultural history that encompasses music, dance, fitness, art, and social justice. Whether used to encourage people to get up and dance or to promote physical and emotional wellbeing, the phrase represents a universal call to action that transcends language, culture, and geography.

III. CONVOLUTIONAL NEURAL NETWORK (CNNs)



"Everybody Dance Now" is a project that uses Convolutional Neural Networks (CNNs) to synthesize new dance moves from videos of human dancers. The project was developed by researchers at UC Berkeley and Google.

"Everybody Dance Now" is a deep learning-based system for music-driven dance generation that uses a convolutional neural network (CNN) to capture spatiotemporal information from music videos and generate dance movements that match the beat and style of the music.

Specifically, the system first extracts a set of visual features from each frame of a given music video using a pre-trained CNN. The visual features encode information about the appearance and motion of dancers, as well as other relevant aspects of the scene, such as lighting and camera movement.

Next, the system uses a sequence-to-sequence model to generate dance movements that match the music. The model takes as input a sequence of musical features extracted from the audio track of the video, and outputs a sequence of dance poses that match the rhythm and style of the music. The model is trained using a large dataset of music videos and corresponding dance movements, and is designed to capture the complex relationships between music and dance.

Overall, the CNN plays a critical role in the "Everybody Dance Now" system by providing a powerful tool for extracting spatiotemporal features from music videos, which are then used to drive the generation of dance movements that match the beat and style of the music.

The basic idea behind "Everybody Dance Now" is to use a CNN to extract information from a video of a

person dancing, and then use that information to generate new dance moves that match a different piece of music. Specifically, the CNN is trained to identify key points on a dancer's body (such as joints and limbs) and track their movement over time.

Once the CNN has learned to identify and track these key points, it can be used to generate new dance moves by predicting the movement of these points in response to different pieces of music. This involves feeding a new piece of music into the CNN and using it to predict the corresponding movements of the dancer's body..

The process starts with collecting a large dataset of dance videos. The videos are then annotated with pose keypoints using a computer vision algorithm, which identifies the position of various body parts in each frame of the video. The annotated videos are used to train a CNN model that can generate new dance moves by learning patterns and relationships between the pose keypoints.

To generate new dance moves, the model is given a starting pose and then generates a sequence of poses that are coherent with the starting pose and also fit the music. The generated poses are then passed to a motion synthesis system that generates a sequence of smooth and natural-looking motions that are consistent with the poses.

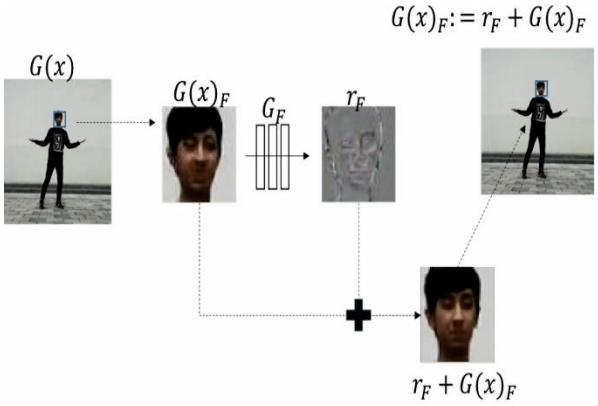
Generally, CNNs are used in "Everybody Dance Now" to learn and synthesize patterns and relationships between pose keypoints in dance videos, and generate new dance moves that are coherent and fit the music.

IV. IMPLEMENTATION DETAILS

Given a video of a source person and another of a target person, our goal is to generate a new video of the target enacting the same motions as the source. To accomplish this task, we divide our pipeline into three stages – pose detection, global pose normalization, and mapping from normalized pose stick figures to the target subject. In the pose detection stage we use a pre-trained state-of-the-art pose detector to create pose stick figures given frames from the source video. The global pose normalization stage accounts for differences between the source and target body shapes and locations within the frame. Finally, we design a system to learn the mapping from the pose stick figures to images of the target person using adversarial training. Next, we describe each stage of our system.

.

1. Backend Implementation: The backend implementation of "Everybody Dance Now" is responsible for the machine learning models that generate the dancing figures from the music inputs provided by the users. The project uses a deep learning approach, specifically a generative adversarial network (GAN) architecture, to synthesize the dancing motions.



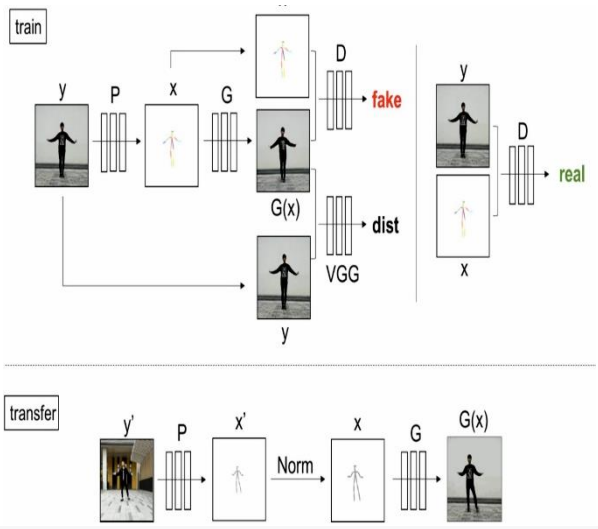
The backend implementation is likely deployed on a cloud-based infrastructure, such as Amazon Web Services (AWS) or Google Cloud Platform (GCP), to provide scalability and reliability. The backend code is likely written in Python and follows best practices in software development, such as modularization, testing, and version control. It also includes APIs to allow communication with the frontend interface and other applications that may want to use the generated dance videos.

2. Frontend Implementation: The frontend is built using web technologies such as HTML, CSS, and JavaScript and likely uses popular frontend frameworks such as React or Angular to create a responsive and interactive interface. It also likely uses APIs to communicate with the backend server that handles the machine learning models and generates the dance videos.

   The interface provides a simple and intuitive design, with clear instructions and buttons to upload and process the videos. Users can also choose different dance styles, such as salsa, ballet, or hip-hop, and adjust the duration and speed of the generated video.

3. Pose encoding and Normalization: Pose encoding and normalization are two important steps in the Pose2Pose model that help to preprocess the input pose data before it is fed into the deep learning model.

   Pose encoding involves converting the raw pose data into a numerical representation that can be processed by the deep learning model. In the "Everybody Dance Now" project, the pose encoding is based on the OpenPose library, which detects and tracks the key points of a person's body (e.g., joints, limbs) from a video or image. The key points are then encoded as a series of 2D coordinates that represent the positions of each joint in the frame.



Normalization, on the other hand, is the process of scaling and centering the pose data to reduce the impact of variations in the lighting, background, and camera position. This helps to make the pose data more consistent and easier for the deep learning model to learn from. In the "Everybody Dance Now" project, the normalization step involves scaling the pose data to a fixed size and centering it at the origin.

4. Pose to Video Translation : The project uses deep learning techniques to achieve this goal. Specifically we use a neural network architecture called a generative adversarial network (GAN) to learn the mapping between the pose of the source dancer and the corresponding pose of the target dancer.

   To train the GAN, the researchers use a dataset of dance videos that includes both the source and target dancers.

   They extract the pose information from each video using a pose estimation algorithm and then use this pose information as input to the GAN. During training, the GAN learns to generate a sequence of poses for the target dancer that matches the sequence of poses for the source dancer. Once the GAN is trained, it can be used to generate new videos of the target dancer performing the same dance as the source dancer.

   Overall, the "Everybody Dance Now" project is a fascinating example of how deep learning techniques can be used to generate realistic videos of people performing complex movements, even when the source and target dancers have different body shapes and sizes.

V. FUTURE SCOPE

This project is a fascinating project that uses machine learning to generate realistic dance videos from input music and poses. The future scope of this project is quite promising as it can be applied in various fields such as entertainment, education, and health.

1. Entertainment Industry: In the entertainment industry, this technology could be used to create music videos or live performances with realistic and unique dance routines. It could also be used to generate dance sequences for video games or virtual reality experiences, allowing players to immerse themselves in new and exciting worlds.
2. Education: In education, this technology could be used to teach dance to students who may not have access to a dance teacher or studio. It could also be used to enhance dance classes by allowing students to visualize and understand dance moves more clearly.
3. Health Industry: In the health industry, this technology could be used for rehabilitation purposes, helping patients recover from injuries or disabilities by teaching them how to move in a safe and effective way. It could also be used to create personalized workout routines for individuals, tailored to their unique body type and fitness level..
4. One potential future scope of the project is to improve the accuracy and realism of the generated dance videos. Currently, the system can generate dance movements that are similar to those of professional dancers, but there is still room for improvement. The researchers could continue to refine the algorithms to make the generated dances even more realistic and visually appealing.

5. Another potential future scope of the project is to expand its applicability to different dance styles and contexts. Currently, the system is designed to generate dances in a specific style, but with further development, it could be adapted to generate dances in a wide range of styles, such as hip hop, jazz, or ballroom. This could have applications in the entertainment industry, where choreography is a critical part of many performances.

Finally, the project could have implications for the field of AI and machine learning itself. The techniques developed for "Everybody Dance Now" could be applied to other domains beyond dance, such as animation, robotics, or even sports. The project represents an exciting and promising avenue for research into the intersection of AI, machine learning, and human expression..

VII. CONCLUSION

The project's goal was to train a machine learning model to learn the movements of a person in a video and then generate a new dance video with a different person or background. The researchers achieved this by training a neural network to learn from a large dataset of videos featuring people dancing. The resulting model was able to generate realistic dance videos, making it a breakthrough in the field of computer-generated content..

The project aimed to develop a deep learning model that could generate dance movements in response to music. The researchers trained their model on a large dataset of paired videos of people dancing to different music genres. They then used the model to generate dance movements in response to music input.

The results showed that the model was able to generate dance movements that were highly synchronized with the music input, and that the movements were diverse and stylistically coherent. The researchers suggested that the model could have applications in fields such as entertainment, education, and virtual reality.

It enable people with little or no dance experience to create dance performances through the use of a simple interface and the guidance of a virtual dance instructor. The system is also designed to be adaptive, meaning it can learn from user feedback and improve its output over time.

All in all, the "Everybody Dance Now" project is a promising development in the field of computer-generated art and has the potential to democratize dance creation and performance.

In conclusion, the proposed solution presents a promising approach to optimize file sharing in torrent networks by tracking redundancy rates and making informed decisions on file downloads and deletions. Further research and experimentation are needed to fully evaluate its effectiveness, scalability, and security in real-world scenarios. The proposed solution contributes to the existing literature on torrent networks and provides a foundation for future research in this area.

VI. REFERENCES

[1] Kfir Aberman, Mingyi Shi, Jing Liao, D Liscbinski, Baoquan Chen, and Daniel Cohen-Or. Deep video-

based performance cloning. In Computer Graphics Forum, volume 38, pages 219–233. Wiley Online Library, 2019.

[2] Guha Balakrishnan, Amy Zhao, Adrian V Dalca, Fredo Durand, and John Guttag. Synthesizing images of humans in unseen poses. In CVPR, 2018.

[3] Aayush Bansal, Shugao Ma, Deva Ramanan, and Yaser Sheikh. Recycle-gan: Unsupervised video retargeting. In ECCV, 2018.

[4] Rodrigo De Bem, Arnab Ghosh, Adnane Boukhayma, Thalaiyasingam Ajanthan, N Siddharth, and Philip Torr. A semisupervised deep generative model for human body analysis. In Proceedings of the European Conference on Computer Vision (ECCV), pages 0–0, 2018.

[5] Patrick Esser, Ekaterina Sutter, and Bjorn Ommer. A varia- tional u-net for conditional appearance and shape generation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 8857–8866, 2018.

[6] Wissam J Baddar, Geonmo Gu, Sangmin Lee, and Yong Man Ro. Dynamics transfer gan: Generating video by transferring arbitrary temporal dynamics from a source video to a single target image. arXiv preprint arXiv:1712.03534, 2017.

[7] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2D pose estimation using part affinity fields. In CVPR, 2017.

[8] Qifeng Chen and Vladlen Koltun. Photographic image synthesis with cascaded refinement networks. In IEEE International Conference on Computer Vision (ICCV), volume 1, page 3, 2017.

[9] Dan Casas, Marco Volino, John Collomosse, and Adrian Hilton. 4D Video Textures for Interactive Character Appearance. Computer Graphics Forum (Proceedings of EUROGRAPHICS), 33(2):371–380, 2014

[10] Alexei A. Efros, Alexander C. Berg, Greg Mori, and Jitendra Malik. Recognizing action at a distance. In IEEE International Conference on Computer Vision, pages 726–733, Nice, France, 2003.

[11] Michael Gleicher. Retargetting motion to new characters. In Proceedings of the 25th annual conference on Computer graphics and interactive techniques, pages 33–42. ACM, 1998.