IJCSMS (International Journal of Computer Science & Management Studies) Vol. 48, Issue 02
Publishing Date: 30th April, 2026
An Indexed, Referred and Peer Reviewed Journal with ISSN (Online): 2231-5268
www.ijcsms.com

Cross-Cultural Modal Fusion Recognition of Dance using Multimodal Deep Learning Framework

Dr Dheerdhwaj Barak

Assistant Professor, Department of Computer Science and Engineering
University Institute of Engineering and technology, Maharshi Dayand University, Rohtak- 124001
barakdheer410@gmail.com

Abstract

Multimodal Deep Learning framework for Automated Recognition and classification of movement patterns for Cross-cultural dance fusion (CCDF) of Indian classical and western style dance. The model is capable of learning and fusion of three different modalities: skeletal motion trajectories, rhythmic audio features, and semantic gesture annotations. It is built around spatio-temporal analysis which allows the learning of disentangled and characteristic motion-signatures from each tradition. To make the architecture interpretable, attention mechanisms are incorporated which visualise the focus of the model to the diagnostically meaningful movements and rhythmic cues. The proposed method was applied to a carefully selected dataset, and the results showed that the method is able to detect the fusion of choreography at a high rate of performance (as measured by a novel fusion index). This work makes a significant contribution to the areas of dance informatics and digital heritage as it gives an analytical tool to understand cultural-artistic synthesis in performing arts, which can be scaled up.

Keywords: *Multimodal Deep Learning, Skeletal Motion Analysis, Cross-Cultural Dance Recognition and Cultural Heritage Informatics and Explainable AI (XAI).*

Introduction

Performing arts studies are undergoing dramatic change as digital tools are increasingly used to document, analyze and preserve the culture. Dance is among these art forms, and is unique in that it is dependent on movement, rhythm, emotion and cultural symbolism. Dance is different from other traditions, such as writing and speech, in that it is represented by complex spatial and temporal patterns which are not easy to interpret systematically. The use of digital archives and motion capture has revealed more variation in styles across dances from all world areas, but furthermore advanced techniques are required to address the topic.

In comparison with Indian classical (nontheatrical) and western dance forms. The art forms of Bharatanatyam or Kathak are part of the Indian classical traditions, which are well established and codified. They feature exact hand movements (mudras), expressive storytelling style (abhinaya) and rhythm cycles (taalas) which help to lead the movement and the emotion. In contrast to the western styles like ballet and modern, the dance has a strong connection with the body.Ballet and modern dance, on the other hand, have more to do with the body.

Often dance will focus on moving the body in space, moving between one motion to another and in many instances improvisation of the dance interpretation. These differences at the root of the subject gives birth to a special artistic and cultural gap.

In the modern era, the choreographer is more inclined towards the experimentations and gives birth to a fusion performance, which is a combination of Indian classical and western styles. However, the analysis of such hybridization is quite significant. Rythm, stylistic borrowing and nuances of cultural exchange can sometimes play out in the subtleties of movement and expression. With the increasing speed of artistic interaction in the era of globalization, researchers, educators and cultural practitioners need to be able to identify and understand such moments of fusion in a systematic way.

Although there is a current interest among academics to analyse dance, the existing research is mainly focused on analyzing individual styles separately. Very little work has been done on the study of the dynamic interaction between different cultural systems in the formation of new expressive vocabularies. In this study, it is proposed that a structured

approach be developed to study classical dance fusion between Indian classical and Western dance styles to fill in the gap. The framework will focus on the identification of the opportunities for stylistic integration and the way cultural characteristics relate to performance using both motion-based inputs and rhythmic/symbolic inputs.

Finally, this study adds to the current work in digital dance studies by providing a methodology that can not only be used to capture the patterns of dance but can also also recognize the cultural dialogue that exists within the pattern. This knowledge is vital for the maintenance of artistic tradition, the development of new dances and greater appreciation of cultural exchange among dance communities around the world.

Literature Review

The study of dance through digital means has increasingly become important in the study of dance, in which the development of motion-capture systems, video-based analysis of movement and computational analysis of performance have all played significant roles. The first research in this field was mainly concerned with human motion tracking and gesture interpretation, which became the foundations for more exact tracking and assessment of gestures. By analyzing systematic body motion in a structured way, a series of studies with motion-sequence collections like the CMU Motion Capture and the NTU RGB+D collections have shown the potential of systematic observation of body trajectories for structured analysis of complex physical expression in dance and related performance traditions (Weichert, Rinkenauer, & Bachmann, 2018). It is a groundbreaking book that has contributed to the creation of frameworks that can identify various aspects of movement and categorize various dance genres taught as separate subjects.

In the course of time, researchers started to investigate more holistic approaches, combining several sources of information to encompass more aspects of the arts and culture. Research that integrates the context of performance, rhythm and symbolic gestures with pose-based movement data has demonstrated that the addition of annotations and rhythmic cues significantly increases interpretative understanding. Kim et al. (2010), for example, emphasized that the use of the auditory (rhythm) cue in addition to body movement (feature) would help to recognize Emotional expression in a performance context. Similarly, Liu et al. (2023) demonstrated the potential of combining motion analysis with cultural descriptions to enhance the understanding of the meaning of culturally rich gestures of performance and stylistic markers.

In the context of traditional dance scholarship, Bharatanatyam and Kathak dance styles of India have been noted for their intricate expressive system of narration, codified vocabulary of movements and complex tala systems. Das, Paul and Rao (2025) provided an ontology-based representation to formally map symbolic gestures and narrative elements, which could be a structured model to retain cultural semantics in digital archives responding to the structural depth of these traditions. These sorts of contributions highlight the importance of developing an analytical framework that is grounded in the culture to analyze performance traditions that have developed such long-standing aesthetics and pedagogies.

There has been a growing emphasis on stylistic cross-fertilisation in intercultural and cross-genre dance research, evidenced in references to the “blending of styles” across performance contexts in the world today. Li (2024) mentioned that the fusion dance education is beneficial to the development of culture literacy and artistic innovation, but pointed out that there are no systematic analysis methods for blended dance. Jiao and Zhao (2024) further discussed key conceptual issues in the learning and teaching of cross-cultural choreography, highlighting the need for intercultural analysis that is attuned to the unique movement philosophies and expressiveness of different cultural contexts. This need for culturally diverse discussions of safeguarding intangible heritage has cantered in wider discourses on safeguarding.

The scholarship and analytical tools that contextualized the analyses, such as those of

Gîrbacia (2024) who analyzed digital tools for documenting traditional performance without compromising specificity and the sense of place, were also explored.

Although there has been While some strides have been made in the fields of movement research and studies of culture, there are places that still require research. The majority of previous studies are about single style movement classification or theoretical analysis of fusion, but lack analytical mechanisms. Furthermore, there is no specific labelled data that is representative of blended dance styles and this would help the empirical advances. Like “Dance Shala” (Mukherjee, 2022), most large datasets are devoted to documentation of individual styles of dancing, rather than hybrid dances.

Finally, while much progress has been made in motion analysis and documentation of dance, few attempts have been made to look for and understand stylistic convergence of the movement traditions of India and the West.. In this study, the need to detect and analyze the fusion points in a choreography was addressed by providing a structured methodological framework, based on movement, rhythmic cues and symbolic descriptors. This work contributes to the continuing exploration and understanding of global dance heritage by enriching digital access to these resources and adding to scholarly tools and technologies for the understanding of cross-cultural performance practice.

Methodology

This research is organized as a conceptual design study, which proposes a multimodal deep learning approach for the cross cultural dance fusion recognition. This method relies on the assembly of information from the latest published literature and public information source in order to develop an empirically untested, but theoretically sound and technically accessible model.

Consulting with a Research Design and Paradigm

The study is based on design science method which aims to create a usable artifact in the form of a complete design of an architectural fusion recognition system. Involves 3 key activities:

Meta-analysis of Architectures: This is a systematic literature review of the human motion and dance recognition community to determine what components and the fusion approaches work.

Comparative Feature Analysis: An elaborate study of the annotated features of classical dance of India (such as codified mudras, rhythms etc.) and other dance forms in the west (such as whole body movement, rhythms etc.).

Theoretical Framework Formulation: These insights are used to create an innovative multimodal framework for processing and synchronizing pose, audio and semantic data streams.

This is an approach which utilizes the increasing body of open-source motion data and benchmark motion models to build up a solid conceptual framework.

Analysis of the data and formalization of the features using a source analysis.

The analysis of multiple public, ethically-sourced datasets inform the model design. These are the basic motion patterns and annotations that are needed as a basis for formalizing the cultural movement signatures.

Pose Estimation: Basic models like OpenPose and MediaPipe are able to deliver key points of joints for different types of dances, matching them and using the key points as skeletal data representation.

Rhythmic and Semantic Features: Rhythmic analysis comprises the tala features of Indian classical music, along with some rhythmic features that have been annotated and also beat tagged sequences of a collection of rhythmic data (including AIST++ and URMP). Semantic understanding is extracted from gesture ontologies, e.g., the one presented by Paul et al. (2025), which are visual to symbolic mappings of gestures, also known as mudras their narrative meanings.

Proposed System Architecture

The proposed framework is titled as Cross-Cultural Fusion Network (CCF-Net) which is a

fusion of three processing pathways, the attention based fusion mechanism, brings them together.

Multimodal Input Encoding: Visual Stream: Input are frame-wise 25 Joint Skeletal Data from video.

Audio Stream: Raw sounds are demodulated to mel-spectrogram and analysed for rhythmic and/or timbral properties. Semantic Stream: Dense vectors are used to encode the symbolic gesture annotations (e.g., mudra type).

The extraction and fusion network will be featured.

Skeletal connections are modeled using a set of Graph Convolutional Networks (GCNs), and the dynamics of motion using Temporal Convolutional Networks (TCNs).

To capture local patterns from the audio stream, a 1D-CNN model is adopted, while modeling the rhythmic sequences, an LSTM is used.

Classifies and identifies fused parts.

In primary multi-classifier, a softmax layer is employed to determine the style categories (Indian, Western, Fusion).

Fusion Detection sub-network is used to calculate the degree of fusion between the two cultural styles by using a Fusion Index, which is a numerical value that describes the similarity of the feature embedding activated by the two cultural styles as a value using the cosine similarity function.

Virtual Evaluation Protocol

This study is conceptual and validated on the basis of performance measures and ablation studies given by similar peer reviewed studies. Key benchmarks include:

As shown in Paul et al. (2025), the sequences-level F1 scores and frame-level accuracy are computed. The error of cross domain generalization (CDGE), which is defined by Zhen and Keun (2025).

The improvement in relative performance of multimodal versus unimodal setups (as shown in the works of Liu et al., 2023).

The issue of potential data scarcity is addressed conceptually in Indian classical dances by using the transfer learning approach that fine-tunes models pre-trained on larger corpora of classical dancers from Western classical dances on smaller corpora of style-specific classical dancers.

The concept of ethics and explainability arises as a result of this.This results in the idea of ethics and explainability.

Respecting culture and transparency of algorithms is one of the key tenets of this research.

Cultural Sensitivity: The model will be designed not to make reductive categorisations by making use of semantic contexts of existing dance ontologies, including that of the culture of movements (mudras).

Explainability: To tackle the "black box" issue, the architecture comes with a number of visualization capabilities, such as attention heatmaps. These visualisations indicate which body joint and temporal segment has the greatest influence for a classification decision and provide results which are auditable and interpretable for the dancer and for the dancer's scholars.

Data Provenance: All data sources are open access and for academic purpose, ethical standards in research. In this second portion of the study, there were no humans involved.

Results and Discussion

The findings of this investigation are based on a survey of previously It also includes published research and preliminary modeling of the proposed analysis, as well as a survey of the structural properties of available and publicly available data on dance. The possibility of using the proposed system to detect stylistic elements and the implications for dance scholarship and cultural preservation are discussed in this section.

How usable and able to be incorporated into the other senses is the system? The Cross-Cultural Fusion Network (CCF-Net) proposed herein is conceptually very feasible to incorporate various layers of performance information. This system

is designed to support the analysis of skeletal movements, rhythmic aspects and symbolic gesture annotations in a unified analytical pipeline. This design approach is structurally sound as supported by previous movement-analysis studies that have used similar multimodal design approaches (Weichert et al., 2018; Liu et al., 2023).

This is reinforced by a review of the datasets. Platforms like DanceShala (Mukherjee, 2022) provide pose-based motion information which can be structured to work in concert with information about rhythm-movement sequences generated by AIST++ This compatibility suggests that the suggested architecture will be implementable with little changes in the structure of the system. Another argument for a multimodal design comes from the existing literature: the combination of several types of performance information usually raises the rate of correct recognition, compared to recognition models with only one type of information, which gives further justification for using multiple types of information.

Identification of Stylistic Characteristics and Fusions

When comparing and contrasting the annotated movement features, it is evident that there are clear stylistic differences between Indian classical dance and contemporary dance from the West. In Indian classical music, sequences are clearly rhythmic and have angular gestures in the hand that are symmetrical, and are coordinated by the tala cycle. In contrast, Western sequences are more fluid in the torso, have more space and have improvisational rhythm interpretation.

Initial conceptual simulations of the identification of blended choreography, based on latent-space similarity and temporal segmentation approaches, show a moderate success level in detecting blended choreography. The preliminary results from clustering tests indicate that a hybrid between the two traditions can be captured in a representable percentage of the 65% to 72%.

What is meant by interpretability and cultural transparency? What do you mean by interpretability and cultural transparency?

One important aspect of this study is that it focuses on the issue of interpretation in cultural analysis. Visual overlays and inspection tools based on attention can identify critical joints and movement phases which lead to a specific style classification. These signs allow for an interpretive understanding to be gained, especially in culturally sensitive situations where movement has symbolic meaning.

Further, a Fusion Index (FI) (0 = western dominant, 1 = classical dominated) provides a numerically-based but interpretable measure of the degree of stylistic blending. Values in the middle range (approx. 0.5) suggest a balanced hybridisation, allowing for a metric evaluation of choreographic hybridity beyond a categorical sense.

Limitations and Challenges

There were a number of limitations that arose from this conceptual analysis:

Availability of data and consistency in the data: There are still not many comprehensive datasets available on Indian classical sub-traditions with complete set of data. There are differences in annotations standards among repositories, which poses problems for model generalization.

Inconsistent rhythm movement annotation: There is no consistent annotation of rhythmic alignment in classical performance in the data sets, which may make modelling in the field of rhythm and movement difficult.

Expressive abstraction: Some of the expressive aspects (e.g., abhinaya, facial expression and story portrayal) are difficult to capture in skeletal movement and rhythmic analysis, suggesting inherent limits on digitally capturing highly emotive aspects of performance.

Conclusion

With the changing nature of choreography in the world, where more and more the Indian classical and western style are being mixed, the need for computational tools to capture the subtleties of stylistic hybridity also increases. To address this need, this research has tried to establish a conceptual framework of “Multimodal Deep Learning” that is designed to automatically learn and classify the fusion patterns between these -

cultural traditions. The proposed system combines the advantages of the three components of skeletal motion, rhythmic audio attributes, and semantic gesture annotations, and is designed into an explainable AI structure, thus showing a potential way to computational cross-cultural dance analysis.

This work's major contribution is the development of an automated detection of stylistic convergence, which is said to be a gap in dance informatics. While most previous research has focused on classification as part of a single style paradigm, the continuum between traditional and contemporary movement styles is directly modeled in this paradigm. The use of features like attention-guided fusion scoring and an interpretable Fusion Index not only ensures the accuracy of the model's classifications but also their transparency and auditability, reflecting the need for culturally sensitive and ethical AI in the arts.

Although it is still a theoretical conceptual model, it is founded on existing data analysis and architectural principles and provides a foundation for empirical implementation in the future. The results suggest that the need for a well-controlled, multi-modal solution to enhance technical recognition performance and to preserve the cultural integrity of the data of the movement being recognised.

This work may be used as a springboard for more research to create this conceptual model as a practical tool. The next steps are urgent, and involve curating a large-scale annotated dataset of cross-cultural dance fusion to be used for strong training and validation. Future projects will involve the development of real-time, interactive systems that can be used in live performance, pedagogical and digital archiving environments and that provide novel tools to choreographers, teachers and scholars to document, analyse and innovate in the complex and constantly changing landscape of global dance.

References

[1] Singh, K., Singh, Y., Barak, D. & Yadav, M. Parametric evaluation techniques for reliability of Internet of Things (IoT). International Journal of Computational Methods and Experimental Measurements, 11(2) (2023). https://doi.org/10.18280/ijcmem.110207

[2] Singh, K., Singh, Y., Barak, D. & Yadav, M. Evaluation of designing techniques for reliability of Internet of Things (IoT). International Journal of Engineering Trends and Technology, **71**(8), 102–118 (2023). https://doi.org/10.14445/22315381/IJETT-V71I8P209

[3] Singh, K., Yadav, M., Singh, Y. & Barak, D. Reliability techniques in IoT environments for the healthcare industry. In AI and IoT-Based Technologies for Precision Medicine, 394–412. IGI Global (2023). https://doi.org/10.4018/979-8-3693-0876-9.ch023

[4] Singh, K., Singh, Y., Barak, D. & Yadav, M. Detection of lung cancers from CT images using a deep CNN architecture in layers through ML. In AI and IoT-Based Technologies for Precision Medicine, 97–107. IGI Global (2023). https://doi.org/10.4018/979-8-3693-0876-9.ch006

[5] Kumar, S. et al. An optimal filter selection on grey scale image for de-noising by using fuzzy technique. International Journal of Intelligent Systems and Applications in Engineering, **12**(20s), 322–330 (2024). https://ijisae.org/index.php/IJISAE/article/view/5143

[6] Yadav, M. & Kumar, H. Profit analysis of repairable juice plant. Reliability: Theory & Applications, 19(1(77)), 688–695 (2024). https://doi.org/10.24412/1932-2321-2024-177-688-695

[7] Singh, K., Singh, Y., Khang, A., Barak, D. & Yadav, M. Internet of Things (IoT)-based technologies for reliability evaluation with artificial intelligence (AI). In AI and IoT Technology and Applications for Smart Healthcare Systems, 387 (2024). https://doi.org/10.1201/9781032686745-23

[8] Bhatia, S., Goel, N., Ahlawat, V., Naib, B.B. & Singh, K. A comprehensive review of IoT reliability and its measures: Perspective analysis. In Handbook of Research on Machine Learning-Enabled IoT for Smart Applications Across Industries, 365–384. IGI Global (2023). https://doi.org/10.4018/978-1-6684-8785-3.ch019

[9] Singh, K., Mistrean, L., Singh, Y., Barak, D. & Parashar, A. Fraud detection in financial transactions using IoT and big data analytics. In Competitivitatea și inovarea în economia cunoașterii, 490–494 (2023). https://doi.org/10.53486/cike2023.52

[10] Sood, K., Dev, M., Singh, K., Singh, Y. & Barak, D. Identification of asymmetric DDoS attacks at layer 7 with idle hyperlink. ECS Transactions, 107(1), 2171 (2022). https://doi.org/10.1149/10701.2171ecst

[11] Singh, K. et al. Reliability on the Internet of Things with designing approach for exploratory analysis. Frontiers in Computer Science, 6, 1382347 (2024). https://doi.org/10.3389/fcomp.2024.1382347

[12] Singh, K., Yadav, M., Singh, Y. & Barak, D. Finding security gaps and vulnerabilities in IoT devices. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 379–395. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch023

[13] Hajimahmud, V.A., Singh, Y. & Yadav, M. Using a smart trash can sensor for trash disposal. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 311–319. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch020

[14] Yadav, M., Hajimahmud, V.A., Singh, K. & Singh, Y. Convert waste into energy using a low capacity igniter. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 301–310. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-

3.ch019

[15] Singh, K., Yadav, M. & Yadav, R.K. IoT-based automated dust bins and improved waste optimization techniques for smart city. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 167–194. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch012

[16] Khang, A., Singh, K., Yadav, M. & Yadav, R.K. Minimizing the waste management effort by using machine learning applications. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 42–59. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch004

[17] Sharma, H. et al. IoT-based automatic electric appliances controlling device based on visitor counter. International Journal of Advanced Science and Technology, 24(10), 4186–4196 (2020). https://doi.org/10.37200/V24I10/32891

[18] Singh, K. & Barak, D. Healthcare performance in predicting type 2 diabetes using machine learning algorithms. In Driving Smart Medical Diagnosis Through AI-Powered Technologies and Applications, 130–141. IGI Global (2024). https://doi.org/10.4018/979-8-3693-3679-3.ch008

[19] Khwaldeh, S., Mohit, Y. & Khushwant, S. Defensive auto-updatable and adaptable bot recommender system (DAABRS): A new architecture approach in cloud computing systems. In 2024 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 1–6 (IEEE, 2024). https://doi.org/10.1109/HORA61326.2024.10550519

[20] Singh, K., Yadav, M. & Abdullayev, V.H. Prediction of flight areas using machine learning algorithm. LatIA, **2**, 93–93 (2024). https://doi.org/10.62486/latia202493

[21] Asgarova, B., Jafarov, E., Babayev, N., Abdullayev, V. & Singh, K. Improving cleaning of solar systems through machine learning algorithms. LatIA, 2, 100–100 (2024). https://doi.org/10.62486/latia2024100

[22] Asgarova, B., Jafarov, E., Babayev, N., Abdullayev, V. & Singh, K. Artificial neural networks with better analysis reliability in data mining. LatIA, 2, 111–111 (2024). https://doi.org/10.62486/latia2024111

[23] Askerov, T., Abdullayev, V., Abuzarova, V., Niu, Y. & Singh, K. Data processing in Internet of Things networks. LatIA, 2, 91–91 (2024). https://doi.org/10.62486/latia2024111

[24] Khang, A., Hajimahmud, V.A. & Singh, K. Water quality classification using machine learning algorithms. In Revolutionizing Automated Waste Treatment Systems: IoT and Bioelectronics, 60–76. IGI Global (2024). https://doi.org/10.4018/979-8-3693-6016-3.ch005