AI-Powered Bharatanatyam Mudra Identification and Description System: Preserving Heritage through Technology

Geethika K. P. K. Nambiar
Department of Computer Science and Applications
Amrita School of Computing
Amrita Vishwa Vidyapeetham Amritapuri
Kollam, India
am.sc.p2mca23016@am.students.amrita.edu

Nandana M. Kumar
Department of Computer Science and Applications
Amrita School of Computing
Amrita Vishwa Vidyapeetham Amritapuri
Kollam, India
am.sc.p2mca23024@am.students.amrita.edu

Veena G.
Department of Computer Science and Applications
Amrita School of Computing
Amrita Vishwa Vidyapeetham Amritapuri
Kollam, India
veenag@am.amrita.edu

M. G. Thushara
Department of Computer Science and Applications, Amrita School of Computing
Amrita Vishwa Vidyapeetham Amritapuri
Kollam, India
thusharamg@am.amrita.edu

***Abstract*—Bharatanatyam is a traditional Indian dance form. It involves hand gestures, known as mudras, and facial ex- pressions to represent emotions and convey stories. In this paper describes a new system for classifying and recognizing Bharatanatyam mudras of single-hand and double-hand styles. Advanced computer vision techniques have been used, including the You Only Look Once (YOLO) framework for object detection, MediaPipe for accurate hand keypoint tracking, and MobileNet for efficient mudra classification. High accuracy has been achieved by the system in detecting mudras across various poses and conditions. The curated data set on which the system has been trained contains various lighting scenarios and camera angles. Each identified mudra is linked to its cultural and symbolic meaning, with this information stored in an Excel sheet to enhance the educational value system. The Excel sheet has been intended to serve as a reference for students, researchers, and enthusiasts interested in the symbolic aspects of the dance. The proposed system is expected to contribute significantly to the learning, research, and preservation of Bharatanatyam performance. Its potential application can also be seen in cultural heritage preservation and educational contexts. The educational value of the system has been enriched, and students and researchers have been provided with a detailed understanding of the mudras. The finding shows that the system can successfully identify the mudra and provide a description, helping Bharatanatyam to be preserved and studied.**

Keywords—Bharatanatyam, Mudra Recognition, Hand Gesture Classification, Deep Learning, MediaPipe, MobileNetV2, Dance Pose Recognition, Real-Time Gesture Recognition.

I. Introduction

Bharatanatyam is one of the oldest classical dance styles in India known for its complex narrative and elaborate hand gestures called mudras [1]. A significant aspect of the dance is represented by mudra to communicate feelings, stories, and cultural importance. Hence, these have been termed as an essential method of communication of the narrative depth and meaning behind each performance.

As time passes, a risk has been posed to traditional knowledge, including the nuances of Bharatanatyam mudras. The preservation of this knowledge can be made easier by an AI system by providing a digital record, with rapid advancement of computer vision and machine learning techniques that allow learning to be improved [2]. Many advantages have been offered to students through self-evaluation tools, cultural preservation with digital archives, and support for the development of interactive resources to increase participation in the art [3].

The literature survey has indicated that although the recognition of mudras in Bharatanatyam has improved, the intricacies remain when it comes to dealing with occlusion, small objects, intricate gestures, and many conditions for a scene. Inspired by various studies [4], [5], [6], [7], [8], and [9] which have demonstrated the effectiveness of YOLO, it fails if complex crowded scenes are involved [10]. Even keypoint detection techniques and hybrid models, such as the Convolutional Neural Network-Recurrent Neural Net- work (CNN-RNN) and HOG-CNN, suffer from computational overheads [11]. Transformer and meta-learning techniques lack scalability and real-time applicability [12]. This study overcomes these gaps with the use of fine-tuned YOLO for accurate hand detection [13], MediaPipe for detailed skeletal keypoints [14], and MobileNetV2 for real-time classification [15]. The diverse dataset and the incorporation of cultural descriptions further address previous limitations while proposing contextual storylines enhance the cultural preservation and usability of the system.

These complex and ancient Indian classical Bharatanatyam dances significantly employ hand mudras as sources to express meaningful stories [16]. However, this poses threats to preservation and understanding for different interpretations that lack systemic documentation and the appropriate tools to apply to analyze their meaning [17]. The existing methods for studying mudras are inexact for recognition under various conditions, including different lighting and camera angles, making it difficult to develop a complete system for their classification and contextual understanding [18]. To overcome this, a robust system for the accurate identification and categorization of Bharatanatyam mudras has been developed [19]. The created data set consists of an appropriate capture of mudras in varying lighting conditions as well as angles of the camera, thereby considering real-world scenarios. Moreover, it is built from advanced computer vision techniques for detecting hands and extracting skeletal keypoints, plus a fine-tuned model for the classification process is used [20]. The system is very precise in identifying mudras, and the description given with the help of this system is also quite elaborate. In this regard, with these technologies, the system also contributes to documenting and analyzing Bharatanatyam mudras in such a way that it could encourage preservation and create greater knowledge about the same cultural heritage [21].

A dataset comprising 52 mudras, including both Asamyuktha and Samyuktha types, was created under various lighting conditions and camera angles. Accurate real-time mudra classification has been achieved using a fine-tuned YOLO model for hand detection, skeleton extraction through MediaPipe, and classification with the MobileNetV2 model. The limitations of previous approaches, such as challenges in handling complex backgrounds and initiating gestures, have been addressed through the proposal, a scalable yet effective solution. In addition, integration of contextual storylines has been introduced, enhancing its usability and relevance for cultural documentation.

II. RELATED WORK

The notable progress made in recent times in object detection and classification, particularly in real-time applications, has been attributed to deep learning techniques. Challenging opportunities for gesture and pose recognition are presented by classical Indian dances such as Bharatanatyam, which typically rely on mudras (hand gestures), facial expressions, and body posture. Therefore, the progression of the literature review has explored the use of CNN, deep learning models, and the YOLO framework, specifically focusing on Bharatanatyam mudras.

A. Object Detection

Redmon et al. [10] introduce a real- time object detection framework known as YOLO, unlike other traditional methods like RCNN this method unifies object localization and classifies this into a single CNN. The entire detection pipeline is optimized using a single neural network of 24 convolutional layers and 2 fully connected layers. Hussain et al. [13] conducted research exploring different types of YOLO versions to analyze architectural advances and performance improvements between YOLO versions. From their research, they found that YOLOv5 has a balance of speed and accuracy. YOLOV8 builds the foundation of YOLOv5 with more features and better performance. YOLOv10 is the best choice for advanced applications, it targets lower power consumption and faster inference. Diwan et al. [11] proves why the YOLO model is better for object detection in real time than those faster models such as RCNN. The authors used most of the YOLO models for comparison. They provide a comparison of YOLO with other object detection methods, especially those that are region-based models like RCNN that provide higher accuracy, and they proved that YOLO is best for real-time processing.

B. Key Point Detection

Zhang et al. [14] provides a real-time hand gesture detection and tracking pipeline that is optimized for resource-constrained devices. The system makes use of a multistage machine learning architecture comprising the BlazePalm model for palm detection, thereby reducing computational complexity but enhancing performance. It is robust against occlusions and varied hand poses with a low latency palm detection approach of 95.7% precision. It applies to augmented reality, game, and human-computer interaction applications. Yaseen et al. [20] are exploring the possibility of CNNs and RNNs to improve gesture recognition performance. They argue that inter-specialization of different dance forms is feasible. A dynamic hand gesture recognition model using MediaPipe, Inception-v3, and Long Short-Term Memory (LSTM) architecture is proposed. Aloysius et al. [12] improve transformer- based models for the task of recognition of sign language with relative positional encoding to overcome the difficulties associated with the spatial representation of hand gestures. This enhancement leads to increased recognition, interpretability, and performance and has possible applications for real-time sign language recognition and gesture-based communication systems. Luangaphirom et al. [22] have created a web application using MediaPipe and OpenCV that allows beginners to do basic body weight exercises with the correct posture. The system captures real-time video from a web camera, determines key points of the body, and calculates the angles for the exercise. The system provides real-time feedback on posture and counts the number of repetitions, helping through a visual and auditory signal. This allows people to exercise safely and effectively at home without a personal trainer.

C. Classification

Sadhana et al. [3] apply deep and meta-learning to the recognition of Bharatanatyam mudras, considering the scarcity of data on classical dance forms. It is an adaptable approach towards the efficient recognition of new classes with minimal training data, so it is applicable to sparse datasets. Varsha et al. [16] developed a hybrid approach combining the spatial feature extraction using HOG and a fine- tuned CNN for high-level features to recognize Bharatanatyam mudras. A combined feature vector was fed into a support vector machines (SVM) classifier, achieving superior accuracy and F1-scores compared to standalone models especially under varying lighting and occlusion. Raj et al. [19] discussed variability issues in shape, orientation, and the background in recognition of the mudra. Their contribution focused on data augmentation improving classification accuracy. Rich datasets were necessary for better generalization. Another contribution focused on an ensemble learning approach combining multiple classifiers for enhancing hand gesture recognition. Kumar et al. [18] [23] utilized deep learning with attention mechanisms for Bharatanatyam pose recognition. According to Anil et al. [18] wavelet multi-head progressive attention enhanced the extraction of spatial and temporal features against pose variations. Based on [23] the multi-frame attention model helped improve recognition by paying more attention to the spatial and sequential aspects between frames. Both methods applied attention in capturing fine-grained details and contextual cues and thus provided robust solutions to Indian classical dance analysis. Sneha et al. [1] classify Bharatanatyam mudras using deep learning models built with the enhanced CNN approach, where the variability of hand shape, orientations, and complexity of the backgrounds are tackled. The data set utilized is diverse with data augmentation to improve robust- ness and accuracy of the model. Their results conclude that it is the fine detail capture for better generalization in gesture recognition, toward the preservation of Indian classical dance heritage. Malavath et al. [21] present the utilization of deep learning for classifying hand mudras in Indian classical dance videos. They use pre-trained CNNs fine-tuned to this task. They address problems such as variability in hand position, lighting, and motion blur. The model's accuracy was improved by using data augmentation and transfer learning, and deep learning can be used to automate the study of traditional dance without diminishing its value. Santhosh et al. [2] explores various machine learning techniques, such as decision trees, (SVM), and neural networks to improve the accuracy level of classification performance. The methodologies for combining these classifiers and enabling effective combinations are described by the authors and provide insights into how aggregation techniques can reduce the vulnerabilities of the weakest model. In addition, feature selection was covered in the paper along with the importance of data augmentation in the development of the ensemble model for enhancing its performance. It is indicated that recognition rates have improved significantly, suggesting that this is probably viable path for further research in the classification of dance gestures. Sadhana et al. [3] focuses on recognizing Bharatanatyam mudras using a combination of deep learning and meta- learning techniques. Given the constraints in the availability of data regarding classical dance forms, the use of meta- learning algorithms in this study allows the model to adapt quickly to new mudra classes even with less training data. The approach is adaptable and efficient, which makes it relevant for dynamic underrepresented Bharatanatyam datasets. The study underlines the potential of deep learning with meta-learning in overcoming the scarcity of data for gesture recognition, thus being a valuable contribution to the analysis of Indian classical dance. Zhu et al. [15] uses an improved version of Mo- bileNetv2 known as I-MobileNetV2 which has many features like reverse fusion mechanism, RELU activation function. The improved network reduces the computational load while preserving a lightweight profile. Naaz et al. [17] discuss how the challenges in real-time video-based dance action detection, such as motion blur and occlusion, can be addressed using a Temporal Neural Network (TCN) in capturing space-time features. It opens the avenues to real-time gesture recognition in the study, which can now go to deep learning techniques for Bharatanatyam mudra classification. A CNN-based study [24] evaluates SVM performance in classifying Bharatanatyam mudras, highlighting customized feature extraction for dance gestures improved classification accuracy. The work lays a foundation for the application of machine learning in classical dance and recommends optimization of feature extraction. Srimani et al. [25] apply meta-learning to adapt the system in just a few trials on a new mudra class with limited data, as classical dance faces data scarcity. They also investigate Samyuktha hand gesture recognition using skeleton matching and gradient orientation techniques, enhancing the adaptability and efficiency of the system. Malavath et al. [26] used CNN and Naive Bayes to categorize Kathakali Asamyuktha Hasta Mudras. Among the features of extraction techniques were edge detection, Haar wavelets, and HOG. With an accuracy of 88%, the CNN model beat Naive Bayes, proving the usefulness of deep learning for mudra classification. Human anomaly detection and dance recognition rely on CNNs, LSTMs, and attention mechanisms but are plagued with issues such as occlusions and noise [27], and [28]. Wavelet-based feature extraction is more robust but not maximally exploited. Alternating Wavelet Channel and Spatial Attention Mechanisms are novel approaches suggested to enhance recognition accuracy in both the fields [27], and [28]. By adjusting important hyper- parameters through ablation tests, Jisha Raj et al. [29] created a six-layer CNN for Bharatanatyam mudra categorization. Their model's accuracy with 15,396 single-hand and 13,035 double- hand mudra pictures was 95.5% and 97.8%, respectively. This demonstrates how systematic tweaking works in deep learning-based mudra identification.

It appears that the literature progresses on the reconstitution of Bharatanatyam mudra recognition, but with limitations. In crowded scenes, YOLO does not perform well on small objects. Key point detection techniques, which include palm-based and CNN-RNN models, fail in complex situations. Transformer-based models offer better recognition capabilities, but have no hand and body pose detection. Meta-learning- based approaches are less scalable for real-time applications. Hybrid approaches such as HOG-CNN are more accurate, but computationally intensive. The ensemble methods add complexity to the approach. The Attention mechanism and I-MobileNetV2 are precision-intensive.

This paper goes one step further than other works by strongly dealing with complex backgrounds due to YOLO- based hand detection for proper localization of hands in a cluttered scene. The 21 keypoints of MediaPipe can produce the minute information required to detect complex mudras. MobileNetV2 provides Scalability and real- time applicability without the inefficiencies shown by earlier models. This is a data set that has all different conditions: lighting, angles, Samyuktha and Asamyuktha mudras. Also provide descriptive modules to add culture and educational background. An integration of YOLO, MediaPipe and MobileNetV2 creates a far stronger and efficient pipeline compared to traditional pipelines with the proposed generation of context-based storylines using integrated mudras, facial expression, and song description that has enhanced system usability and cultural relevance of the system.

III. Proposed System

The system has been designed for the detection of Bharatanatyam mudras by integrating hand detection with gesture classification, allowing single-hand as well as double- hand gestures to be recognized. Two primary components have been utilized: MediaPipe for real-time tracking of hands and keypoints and YOLO for hand detection. A Mudra model has been created using MobileNetV2, which classifies mudras as identified hand gestures. In addition, an Excel spreadsheet has been used to maintain and refer to extensive descriptions of each mudra. The process requires the input of images or video frames. After the processing step, the identified mudra is provided along with a description. The overall system architecture consists of four primary components: (1) Object detection using YOLO, (2) Key point identification using MediaPipe, (3) Mudra classification using MobileNetV2, and (4) Mudra description retrieval from the Excel sheet. The architecture has been demonstrated in Fig. 1, showing how input images are processed, mudras are identified, and the corresponding descriptions are retrieved.

A. Experimental Work

The dataset was prepared in such a way that images of Samyuktha mudras (two-hand), which include 24 mudras, and Asamyuktha mudras (single-hand), which include 28 mudras, were collected under varying conditions. The images were gathered from Bharatanatyam performances, instructional videos, and custom captures, incorporating variations in lighting (natural, artificial, low light) and camera angles (frontal, angled, top-down). An Excel sheet containing the names of mudras along with their meanings as shown in Table 1. The meanings were pulled from this sheet after categorization, thus providing the mudra name along with its possible meanings. This structured approach was adopted to ensure a precise and sensible dataset for training and evaluation. The dataset that was created by ourself representing different mudras has been posted on Kaggle. https://www.kaggle.com/datasets/geethikanambiar/mudra- dataset.

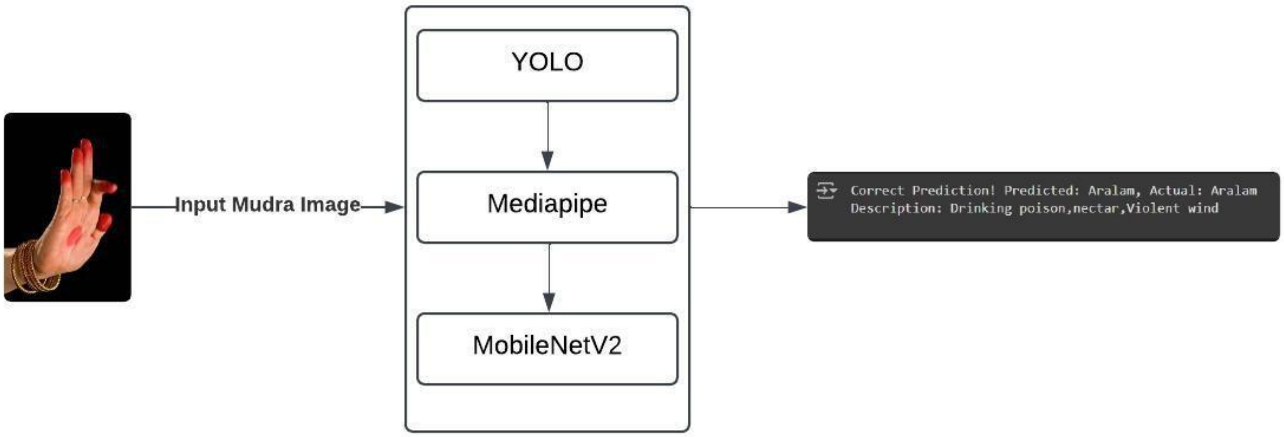

Fig. 1. Architecture Diagram

TABLE I. DESCRIPTIONS OF BHARATANATYAM MUDRAS

Mudra	Description
Alapadmam	Viraha (yearning for the beloved), Mukura (mirror), Tataka (pond or lake), Udrathakopa (great anger), Shakata (cart), fresh ghee, sweets, ball, dancing, palace, cluster of flowers
Aralam	Drinking poison, nectar, Violent wind
Arthachandran	It denotes the eighth phase of the waning fortnight of the moon. (Half Moon), A hand seizing the throat, A Spear, Consecrating and bathing an image, A dining plate, source or beginning, Waist, contemplation, Meditation, Mudra for numbers
Arthapataka	Leaves, A board or a slab for writing, Bank of the river, To Indicate "Both", A knife, A banner, A tower, An animal horn
Bramaram	Bee, parrot, wing, cuckoo, heron
Chaduram	Musk, Gold, Copper, Iron, Dampness, to indicate lesser quality, grief, Aesthetic pleasure, eyes, difference in caste, proof, sweetness, slower gait, breaking into pieces, face, oil and ghee
Chandrakala	Moon, face, index of measure, objects of similar shape, The Ganga River, A club (weapon)
Hamsapaksham	Number six, Construction of a bridge, engraving with nails, to cover
Hamsasyam	An auspicious occasion or festival, tying thread, Ascertaining the imparted instructions, Horripilation (Romancha), Pearl, Light a lamp, A touchstone (stone meant to test gold), Flowers like Jasmine, to draw picture, Sting, Impress
Kadakhamukham	Plucking or picking flowers, holding a necklace or a garland, Pulling the bow string, offering betel leaves, to show preparing a paste of sandals or musk, to mix, to smell, to speak, to see
Kangulam	To represent Lakuca fruit, Bells worn by children, Bell, Chakora bird, Betel-nut tree, The bosoms of young maiden, White Lilly flower, Caataka bird (Skylark), Coconut
Kapitham	Goddess Lakshmi and Saraswati, milking cows, holding cymbals, holding flowers at the time of making love, Grasping the end of the robes, offering incense or light, Collyrium (applying Kajal)
Kartharimukaham	A Scissors, Separation of a couple, Opposition, Looting, to show two different things, Corner of an eye, Death, Lightning, Sleeping, Falling and Weeping, Creeper
Mayooram	A Creeper, A Bird, Vomiting, Separating the hair locks, Applying Tilak on the forehead, dispersing water of the river, Something Famous, Discussing the Shastras
Mrigasheersham	A Deer's head, Lord Krishna (when held with both hands), Women's cheek, A Wheel, Fears, Quarrel, Costume or dress, Tripundraka made on the forehead (Tilak on Lord Shiva's forehead), A lute, Massage on the feet, Holding Up
Mukulam	Stepping and calling the beloved
Mushti	Grasping objects, Combative position of the wrestlers, Steady fastness of a person
Padmakosham	Fruit, Woodapple, Bosoms of women, Circular movement, Ball, A vessel, to eat, Bud/ Can be used to show a garland, Mango, Flower strewn around, Cluster of flowers/ bouquet, Offering flowers for worship, A bell, An idol, Ant Hill

B. Model Components

Three models are employed by the system: YOLO for hand detection, MediaPipe for hand key point detection, and MobileNetV2 for Mudra classification.

1) YOLO

The model was used solely for the detection of hand movements in images while isolating the regions of interest for the classification of mudras in Bharatanatyam. Objects were detected in a single pass, making the approach highly suitable for fast and accurate hand detection. Effective performance was observed even on complex backgrounds. A pretrained YOLO model was specifically fine-tuned for hand detection. Bounding boxes were output around the detected hands, which were extracted for further processing. The detected hand regions cropped were cropped and transferred to the next stage for the extraction of features required for mudra classification. To ensure a robust and efficient hand isolation process, the YOLO technique was adopted for hand detection itself [10].

2) MediaPipe

The MediaPipe Hand Keypoint Detection was employed to obtain highly detailed skeletal information from the regions detected as hands using YOLO. This step was crucial in capturing the fine subtleties of hand poses required to classify specific mudras in Bharatanatyam. The required 21 hand landmarks, covering key joints and fingertips, were detected by MediaPipe. Its lightweight architecture allowed it to be well suited for handling real-time applications. The 3D coordinates (x, y, z) for all keypoints were provided, offering a comprehensive interpretation of the hand pose. The output of the previous step, produced by YOLO, was processed through the MediaPipe module. MediaPipe detected 21 key points for each hand, including the wrist, knuckles, and finger- tips. After the extraction of keypoints, the feature vectors were used as input to the model where mudras were classification of mudras were performed. Skeletal features were captured efficiently and accurately, allowing the system to detect minute variations that distinguish different Bharatanatyam mudras. This stage significantly contributed to the general accuracy and reliability of this classification pipeline [14].

3) MobileNetV2

The base model used to classify the Bharatanatyam mudras was developed based on the key points of the hands extracted using the MediaPipe model. A Mudra classification model, created using MobileNetV2, was optimized to minimize computational requirements, making it suitable for real-time execution. Although the model is lightweight, it achieves high accuracy in distinguishing intricate mudras. As input, 21 normalized hand key points extracted from MediaPipe were provided to the model. A fine-tuned MobileNetV2 was utilized for the mudra dataset. The input was classified into predefined mudra categories such as Anjali, Pataka, or Tripataka. To map the classified mudra label to its corresponding description, a curated Excel sheet was used. As a result, the name of the mudra and its meaning were provided to the user [15].

C. Retrieving Mudra Descriptions from Excel Data

After a mudra is identified, a corresponding description is retrieved from an Excel sheet. This sheet contains predefined descriptions of each mudra, detailing their cultural significance, meaning, and execution techniques. The entries for all 52 mudras are included in the Excel file, comprising 24 Samyuktha mudras (two-hand mudras) and 28 Asamyuktha mudras (single-hand mudras), each entry providing the mudra name and its description. Once a mudra is classified, the relevant description is searched for in the Excel sheet and displayed to the user. The retrieved description is presented alongside the recognized mudra, offering users an educational explanation of the gesture [1].

IV. Experiment and Result

The dataset consists of various Bharatanatyam mudras. These images were pre-processed to ensure that the hands were clearly visible. The images were passed through the YOLO object detection module, which efficiently captured the hands in complex backgrounds. After that, MediaPipe was used to identify keypoints, as shown in Fig. 2. The images were first passed through the YOLO object detection module. The module efficiently captured a hand within the image. The capability of YOLO to detect objects even in cluttered backgrounds ensured that the correct location of the hand produces the mudra. The output of this stage, as shown in Fig. 2, demonstrates how well YOLO performs in isolating regions of interest in an image.

Fig. 2. Object Detection Using YOLO

The detected region of the hand was then passed to MediaPipe's keypoint detection module, which was customized for hand landmark detection. MediaPipe identified 21 key points that represent the skeletal structure of the hand, as depicted in Fig. 3, providing a detailed spatial output of the pose of the hand, including the location of the wrist, knuckles, and fingertips. These details were crucial in distinguishing the Aralam mudra from similar hand postures.

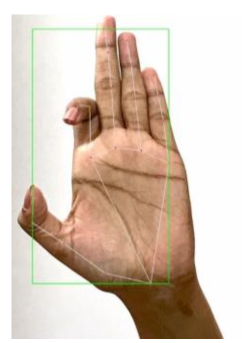

Fig. 3. Key Points Detection by using MediaPipe

Following the detection of keypoints, the extracted key-points were passed to the MobileNetV2 classifier. The classification in this stage of the pipeline assigned the mudra in the image, with a high degree of confidence, as Aralam. The system also used the filtered database stored in a Microsoft Excel file to retrieve a description of the Aralam mudra. The final output is shown in Fig. 4.

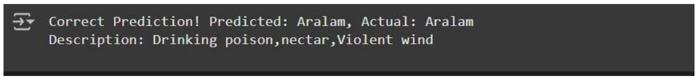

Fig. 4. Output

The fine-tuned model converged more effectively compared to the base model. The validation accuracy of the fine-tuned model exceeded that of the base model by the eighth epoch, reaching 83%, thus generalizing well with minimal overfitting. Fig. 5 visualizes the accuracy of the model.

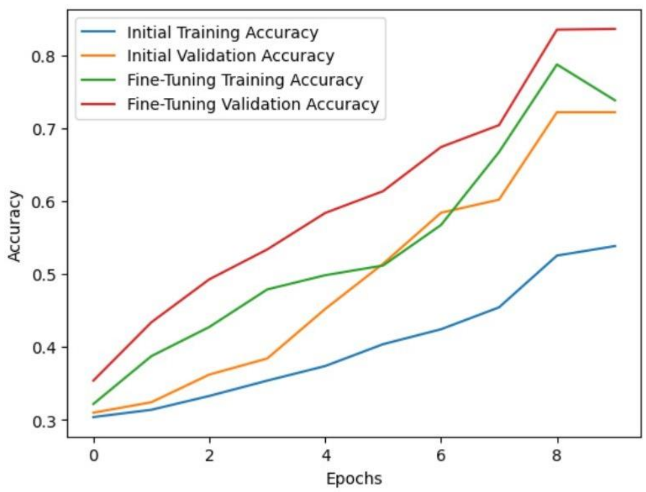

Fig. 5. Accuracy vs Epochs Graph.(X-axis: Number of epochs (count); Y-axis: Accuracy (%)

V. Conclusion

A system has been designed using a multistage methodology that includes advanced computer vision and deep learning techniques. A large dataset of 52 mudras, consisting of Asamyuktha and Samyuktha varieties, has been produced by collecting images from performances, training video captures, and user-triggered captures under varied illuminations and viewpoints. The detection phase has been carried out using YOLO to identify the region of interest, particularly the hand Segments in complex backgrounds. These detected regions have been further processed using MediaPipe, which extracts 21 key points that represent the skeletal features of the hand with commendable accuracy. A fine-tuned MobileNetV2 model optimized for lightweight execution in real-time has been used for classification. Each recognized mudra has been linked to its corresponding cultural and symbolic descriptions from an Excel sheet, allowing users to learn about its significance. The system has not only facilitated the accurate recognition of Bharatanatyam mudras but has also emphasized the cultural and educational impact of modernizing traditional arts through technology.

The system has proven to be a valuable resource for performers, educators, and researchers in Indian classical dance. With the incorporation of AI, the intangible heritage has been preserved in innovative ways. In some way, this program results in an integrated documentation toolset, analysis, and interpretation system for cultural practices. By incorporating additional elements such as facial expressions, songs, and narratives, an AI-driven platform can be created to capture the essence of Bharatanatyam and ensure its relevance for future generations while celebrating its deep societal and cultural significance.

For future work, the Bharatanatyam video is planned to be converted into its storyline. For this, Mudra's identification along with its description, as well as facial expression recognition and its description, will be utilized. The significance of Bharatanatyam songs is acknowledged, and songs in different languages such as Sanskrit, Telugu, Tamil, and Malayalam will be translated into English with their descriptions extracted. By combining mudra descriptions, facial expressions, and translated songs, the information will be provided to the Large Language Model (LLM) model which will generate the Bharatanatyam storyline in PDF format. This will help ordinary people who have no prior knowledge of dance but are interested in it to easily understand the meaning of the video using the system.

References

[1] S. Haridas et al., "Detection and classification of indian classical bharathanatyam mudras using enhanced deep learning technique," in 2022 International Conference on Innovations in Science and Technology for Sustainable Development (ICISTSD), pp. 18-23, IEEE, 2022.

[2] R. Santhosh and R. KK, "Classification of asmyukta mudras in indian classical dance using handcrafted and pre-trained features with machine learning and deep learning methods," Available at SSRN 4818826, 2024.

[3] P. Sadhana et al., "Bharatanatyam mudra recognition using deep learning and meta-learning techniques," in 2024 1st International Conference on Communications and Computer Science (InCCCS), pp. 1-6, IEEE, 2024.

[4] M. Thalapilly et al., "Enhanced aerial vehicle detection using a customized yolov8 model for real-world applications," in 2024 9th International Conference on Communication and Electronics Systems (ICCES), pp. 1751-1757, IEEE, 2024.

[5] P. Krishnan et al., "Assistive eye: A comparative analysis of yolo object detection models on edge devices," in Proceedings of the 2024 Sixteenth International Conference on Contemporary Computing, pp. 104-108, 2024.

[6] S. K. Manoharan et al., "Elevating yolov5 performance: Hyperparameter optimization for numeric, environmental, and healthcare applications," in 2024 3rd International Conference for Advancement in Technology (ICONAT), pp. 1-6, IEEE, 2024.

[7] R. A. Krishna et al., "An efficient model for camera mounted helmet and number plate detection on custom dataset," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1-6, IEEE, 2024.

[8] B. Nagavasista et al., "A robust hand gesture recognition model for assistive typing systems using yolov8," in 2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA), pp. 553-560, IEEE, 2024.

[9] P. Nandal et al., "Enhancing real time object detection for autonomous driving using yolo-nas algorithm with cleo optimizer," International Journal of Information Technology, vol. 17, no. 3, pp. 1321-1328, 2025.

[10] J. Redmon et al., "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788, 2016.

[11] T. Diwan et al., "Object detection using yolo: Challenges, architectural successors, datasets and applications," Multimedia Tools and Applications, vol. 82, no. 6, pp. 9243-9275, 2023.

[12] N. Aloysius et al., "Incorporating relative position information in transformer-based sign language recognition and translation," IEEE Access, vol. 9, pp. 145929-145942, 2021.

[13] M. Hussain, "Yolov5, yolov8 and yolov10: The go-to detectors for real-time vision," arXiv preprint arXiv:2407.02988, 2024.

[14] F. Zhang et al., "Mediapipe hands: On-device real-time hand tracking," arXiv preprint arXiv:2006.10214, 2020.

[15] Q. Zhu et al., "A study on expression recognition based on improved mobilenetv2 network," Scientific Reports, vol. 14, no. 1, pp. 8121, 2024.

[16] K. Varsha and M. L. Pai, "Bharatanatyam hand mudra classification using SVM classifier with HOG feature extraction," in Innovations in Computer Science and Engineering: Proceedings of 7th ICICSE, pp. 175-183, 2020.

[17] S. Naaz and K. S. Kumar, "Integrated deep learning classification of mudras of bharatanatyam: A case of hand gesture recognition," The Scientific Temper, vol. 14, no. 4, pp. 1374-1380, 2023.

[18] D. A. Kumar et al., "Deep bharatanatyam pose recognition: a wavelet multi head progressive attention," Pattern Analysis and Applications, vol. 27, no. 2, p. 53, 2024.

[19] V. A. Raj and G. Malu, "Engesto: An ensemble learning approach for classification of hand gestures," IEEE Access, 2024.

[20] Yaseen et al., "Next-gen dynamic hand gesture recognition: Mediapipe, inception-v3 and LSTM based enhanced deep learning model," Electronics, vol. 13, no. 16, p. 3233, 2024.

[21] P. Malavath and N. Devarakonda, "Natya shastra: Deep learning for automatic classification of hand mudra in Indian classical dance videos," Revue d'Intelligence Artificielle, vol. 37, no. 3, 2023.

[22] T. Luangaphirom et al., "Real-time weight training counting and correction using MediaPipe," Advances in Computational Intelligence, vol. 4, no. 2, p. 3, 2024.

[23] A. Kumar et al., "Multi frame multi-head attention learning on deep features for recognizing Indian classical dance poses," Journal of Visual Communication and Image Representation, vol. 99, p. 104091, 2024.

[24] M. Yan and Z. He, "Dance action recognition model using deep learning network in streaming media environment," Journal of Environmental and Public Health, vol. 2022, no. 1, p. 8955326, 2022.

[25] P. Srimani and S. Kavitha, "Recognizing samyuktha hand gestures of Bharatanatyam using skeleton matching and gradient orientation," International Journal of Current Research, vol. 5, no. 6, pp. 1457-1462, 2013.

[26] P. Malavath and N. Devarakonda, "Classification of Kathakali asamyuktha hasta mudras using Naive Bayes classifier and convolutional neural networks," in Applied Computing for Software and Smart Systems: Proceedings of ACSS 2022, pp. 131-144, Springer, 2023.

[27] S. Nou et al., "Human pose feature enhancement for human anomaly detection and tracking," International Journal of Information Technology, pp. 1-10, 2024.

[28] P. Kishore et al., "Alternating wavelet channel and spatial attention mechanism for online video-based Indian classical dance recognition," International Journal of Information Technology, pp. 1-19, 2024.

[29] R. Jisha Raj et al., "Systematic approach to tuning a deep CNN classifying Bharatanatyam mudras," in Proceedings of the Satellite Workshops of ICVGIP 2021, pp. 3-23, Springer, 2022.