International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Aggregation Signature of Multi Scale Features from Super Resolution Images for Bharatanatyam Mudra Classification for Augmented Reality Based Learning

Saba Naaz[1], Dr. K. B. ShivaKumar[2], Dr. Parameshachari B. D.[3]

Submitted: 01/11/2022 **Accepted:** 03/02/2023

Abstract: Hand gesture is an important non verbal communication mechanism of Indian classical dances especially Bharatanatyam. The hand gestures in Bharatanatyam are called as mudras and there are total 52 mudras with 28 single hand mudras and 24 double hand mudras. Many computer aid mudras classification systems were designed to infer the non verbal theme communicated via mudras. But unlike other hand gesture recognition system, accurate classification of mudra is challenging due to high structural similarity between mudras. This work proposesdeep learning multi scale feature guided aggregated signature for accurate classification of mudras. The deep learning multi scalefeatures are extracted from multi scale images after super resolution and thus self similarities between mudras can be easily differentiated. In addition the features are scale and orientation invariant. Aggregation signature is constructed based on multi scale super resolution features to reduce the classification time. The proposed solution is able to provide an average accuracy of 96% which is atleast 2% higher compared to existing works. Finally the proof of concept of application of proposed mudra classification system in augmented reality based learning system is presented.

Keywords: *communication, orientation, resolution, Aggregation, classification*

1. Introduction

Indian classical dances are performed with rigid code and conventions and Bharatanatyam is one such classical dance practiced mostly in southern parts of India. The dance is based on coordinated movement of foot, hand, face and body [1]. Use of hand gesture to convey non verbal communication is a salient feature of Indian classical music and it is more profound in Bharatanatyam. Hand gestures or mudras in Bharatanatyam convey various outer event, inner feelings like joy, fear etc. Mudras performed in one hand are called Asamyukta and mudras performed in two hands are called Samyukta [2]. Computer aided mudra detection systems are being recently used for self study and interpreting the information communicated by the performer. It is very challenging to interpret 27 Asamyukta and 24 Samyukta mudras as there are high structural differences in the mudras and considering the pose variations, error in accurate interpretation of mudra is high. Conventional hand gesture recognitions methods based on hand crafted features like shape descriptors, moments, SIFT features etc cannot discriminate between structurally similar mudras. Conventional hand crafted features combined with traditional machine learning algorithms have higher false positives [4]. Recently many deep learning based hand gesture recognition systems are being proposed. These deep learning systems avoid the need for hand crafted features. Features learnt at convolutional layers are reduced in the pooling and fully connected layers. The features are further at the classification layer to the mudras. But the problem in these deep learning models is that, it becomes overfit and has higher false positives for mudras recognition. To solve this problem of overfit, this work proposes a integration of deep learning model with multi scale super resolution SIFT.

The proposed solution constructs multiple super resolved image and extract characteristics SIFT features. The high relevant SIFT points are selected. Deep learning features extracted around the SIFT points are aggregated to create a deep aggregation signature. A novel hamming based distance operator is proposed over the deep aggregation signature to match mudras. Due to use of multi scale based SIFT, the discriminating ability of features in presence of structural similarity is enhanced in the proposed solution. In addition hamming based aggregation signature matching reduced the mudra recognition time. Following are the novel contributions of this paper work

(i) A novel multi scale super resolved features integrated with deep learning for feature extraction.

(ii) A novel aggregation signature based on multi scale

[1]*Visvesvaraya Technological University. Belagavi, Karnataka,*[2]*HOD, Dept of CSE, Don Bosco Institute Technology, Bengaluru*

[3]*Professor, Dept of E&CE, NITTE Meenakshi Institute of Technology, Bengaluru.*

feature guided deep learning feature for accurate classification of mudras

(iii) A fuzzy membership function to match mudra based on aggregation signature.

The rest of the paper is organized as follows: In Section II, a detailed survey on existing methods in two categories: hand gesture recognition and mudra recognition. The problems in mudra recognition are found and research gaps are identified. In Section III, the proposed aggregation signature of multi scale features from super resolved image for mudra recognition is detailed. In Section IV, the results of proposed solution for different classes of Mudras are presented and compared with existing works. In Section V, the conclusion and scope for future research are presented.

2. Related Work

Saha et al [5] addressed the problem of mudra recognition using fuzzy membership functions. Texture segmentation is applied to segment the hand regions. From the segmented hand regions, eight spatial points are localized and distance of these points from center of hand region is calculated as feature vectors. Fuzzy membership function between the feature vectors and the mudra is constructed. But in presence of structurally similar mudras, the features have lower discriminating ability as it is based only on eight spatial locations.

Parameshwaran et al [6] improved the accuracy of CNN for mudra classification using transfer learning. The accuracy of CNN can be improved by training with large dataset. To increase the volume of the dataset, authors applied data augmentation. VGG16 architecture was used as the CNN model. The method was able to achieve about 98% accuracy but the mudras classes were limited and did not consider structural similarity scenarios. Anami et al [7] classified mudras using a three stage methodology. Canny edge detection is applied to extract hand contour. From the hand contour, cell features are extracted. Rule based classification is applied on the contour features to recognize mudras. The false positive is higher under structurally similar mudras. Kumar et al [8] classified mudras by using the histogram of gradient(HoG) features. From the segmented hand region, HoG features are extracted and classified to mudras using SVM classifier. The HoG features are not scale and rotation invariant. Also HoG features doesn’t have discriminative ability against structurally similar mudras. Kopuklu et al [9]classified hand gestures using 3D CNN. Feature learning ability is increased by doubling the connected layer and increasing the training volume using data augmentation. The suitability of model for recognizing structurally similar mudras is untested. Sahoo et al [10] recognized hand gestures by using multi CNN score level fusion technique. AlexNet and VGG-16 results are fused with weighted fusion technique to get the final score. Involving multiple CNN models increase the confidence of gesture recognition but the proposed solution was not tested for structurally similar gestures. Patil et al [11] classified hand gestures using Hough transform features. Spatio Hough transform features are extracted from segmented hand gesture images and classified using ANN. Authors extracted features at multi scale and fused it for higher gesture recognition accuracy. The concept of extracting features at multiple scales can be extended for mudra recognition. Fang et al [12] classified hand gestures using geometric feature encoded in Fisher vector notation. Authors extracted geometric features of distances, angles and curvatures from hand gesture images. These features are encoded using Fisher vector. Support vector machine classifier is trained to classify the fisher vector to hand gestures. Geometric features considered in this work were not scale or rotation invariant. Gadekallu et al [13] improved the performance of CNN for hand gesture recognition by fine tuning the hyper parameters of CNN. Authors applied Harris Hawks optimization algorithm for fine tuning the hyper parameters. Hyper parameter fine tuning yielded significant performance improved and suitability of it mudra classification need to be experimented. Lim et al [14] segmented the hand gesture from image and extracted feature covariance matrix from it. The features are then matched to gestures using Eigen distance matching. Though the scheme is light weight, it is not scale or translation invariant. Jain et al [15] experimented with use of deep convolutional neural networks for classification of Indian classical dance postures. Authors adapted the Resnet model with different layers so that posture classification accuracy improves. Dance postures across Indian classical dances did have large structural differences so the method performed well. But it may not suit for the case of postures with higher structural similarity. Kumar et al [16]classified Indian classical dances using multi feature fusion approach. From the segmented dance postures, features of ernike moments, Hu moments, shape features, LBP features and Haar are extracted. Adaboost multi classier is trained to classify the dance types. The method is not suitable for fine grained classification at level of postures. Kishore et al [17] experimented with optimization to CNN to increase accuracy in classification of mudras. Pooling mechanism in CNNs were adapted with stochastic pooling techniques to optimize feature selection. But the pooling techniques removed the necessary features to solve the structural similarity problems. Mozarkar et al [18] used wavelet features for mudra classification. Wavelet features were extracted from hand gestured images using Quaternion Fourier transform. These features are then classified to mudras using KNN classifier. The features were not sufficient to detect structurally similar mudras. Kishore et al [19] recognized hand gestures using shape and texture feature extracted from the segmented hand gestures. Active contour model is applied to segment the hand gesture regions. The shape and texture features extracted from the hand gesture regions were encoded to fisher vectors. Artificial neural network is trained with the fisher vector to classify hand gestures. The method is not scale or rotation invariant. High structural similarity is the important problem not addressed in existing mudra recognition systems. Though existing methods attempted multiple techniques like data augmentation, transfer learning, hyper parameter fine tuning, multi scale features they could not solve the structural similarity problem in mudra recognition. Among multiple techniques, multi scale features looks promising to solve the problem of structural similarity among mudras and this work adopts multi scale features approach.

3. Aggregation Signature Of Multi Scale Features

The architecture of the proposed solution is given in Figure 1. The mudra image is first processed using multi scale Gaussian pyramid to locate interest points which are scale and orientation invariant. From the interest points, equal size patches are extracted. From the patches, deep learning features are extracted using Quaternion discrete cosine transform (QDCT) integrated convolutional neural network model. From the deep learning features, aggregation signature is constructed and used for matching the mudras. Matching of mudras is done with fuzzy membership functions constructed over the aggregation signature. The important stages in proposed solution (i) multi scale interest point localization, (ii) QDCT based deep feature extraction (iii) Aggregation signature generation and (iv) mudra classification are detailed in below subsections.

A. *Multi scale interest point localization*

Multi scale Harris corners are extracted as interest points from the input mudra image. For the input image $I(x, y)$, Gaussian pyramid $P_l(x, y)$ is repeatedly formed by super sampling at rate $s = 2$ and pyramid smoothing width of $m_p = 1$. Interest points are localized in each pyramid level images. Gaussian pyramid is computed by applying

Gaussian kernel $G(x, y, \sigma)$. It is represented as

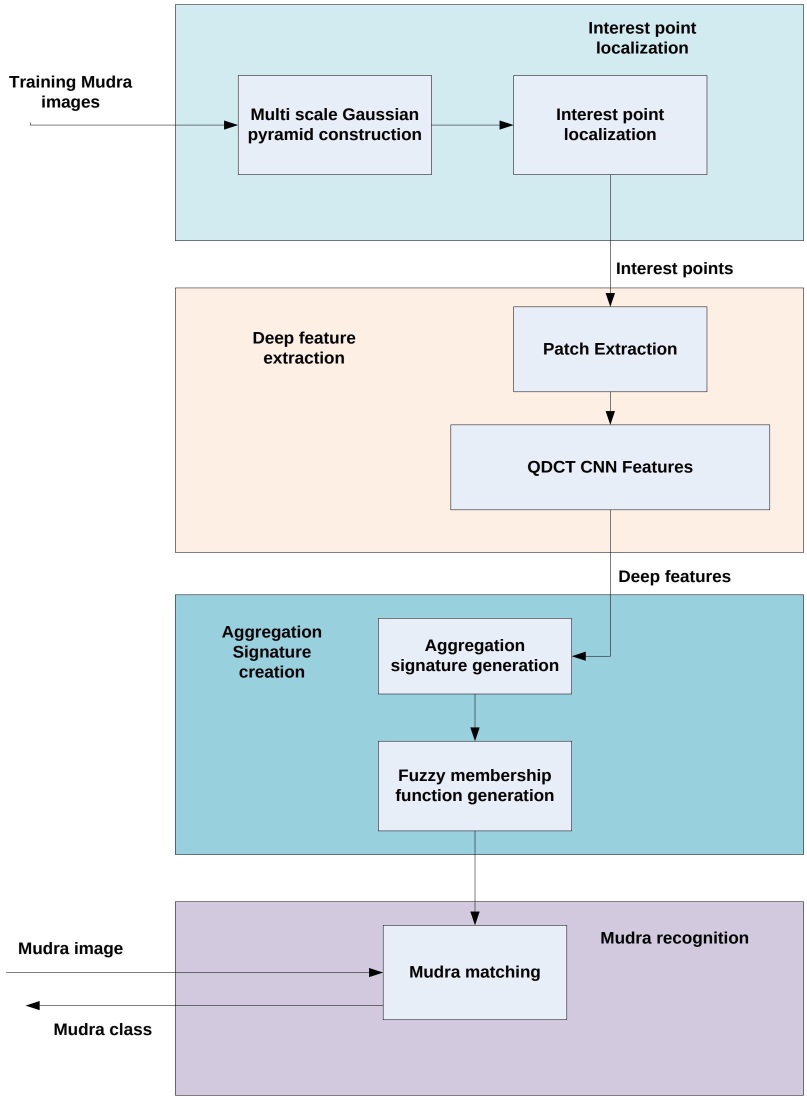

Fig. 1 Proposed mudra recognition with aggregation signature

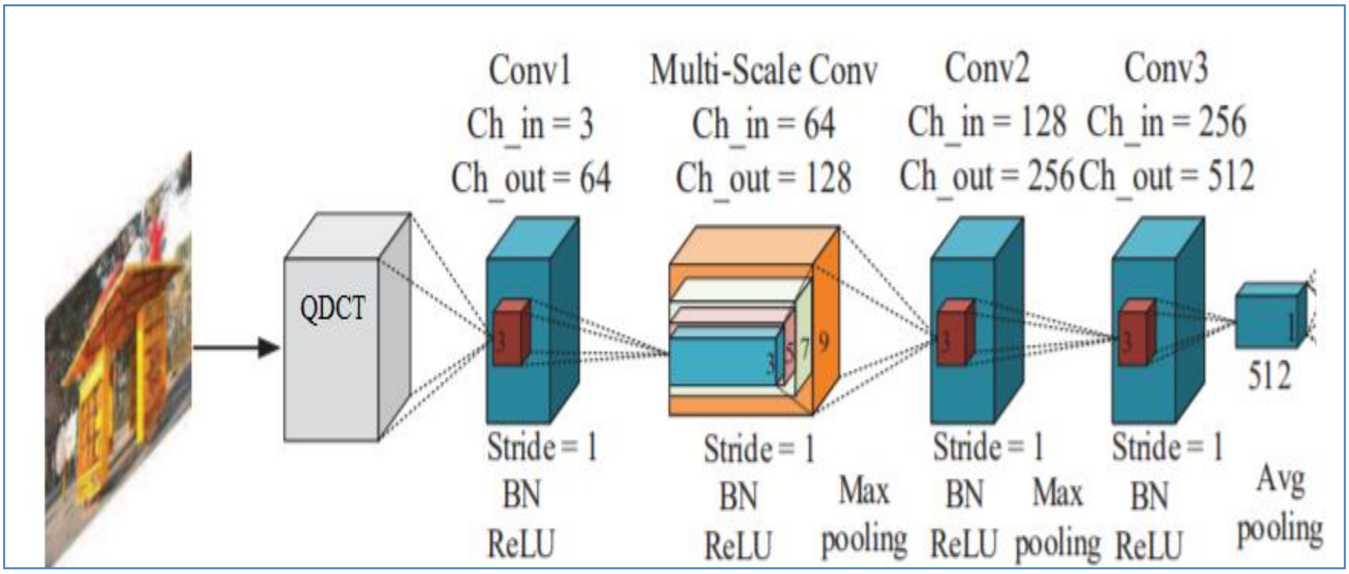

Fig. 2 QDCT CNN for feature extraction

Table 1 CNN configuration

Layer	Input	Filter	Stride	Padding	Out
MaxPool	32×32×128	2×2	2	0	16×16×128
Conv2	16×16×128	3×3	1	1	16×16×256
MaxPool	16×16×256	2×2	2	0	8×8×256
Conv3	8×8×256	3×3	1	1	8×8×512
AvgPool	8×8×512	8×8	1	0	1×1×512
FC	512×1	-	-	-	2×1

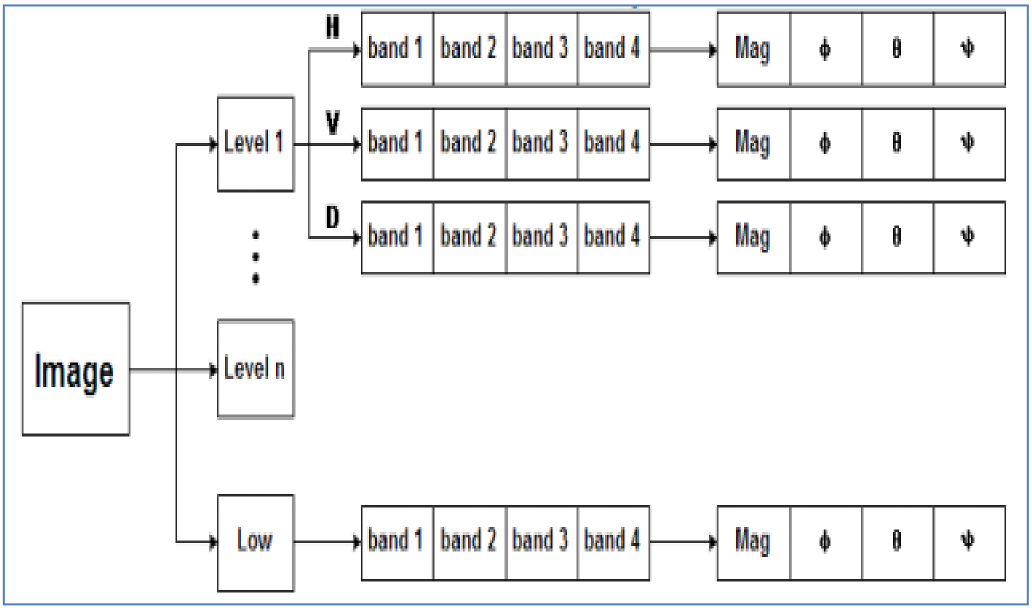

Fig. 3 QDCT Coefficients of Image

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)^2}{2\sigma}}$$

The kernel is applied on the image to calculated super resolution image as

$$L(x, y, \sigma) = G(x, y, \sigma) \times I(x, y)$$

The interest point is found by calculating a corner strength function and thresholding it. The corner strength function is calculated as

$$f_s = \frac{\det H_l(x, y)}{tr\ H_l(x, y)}$$

Where $H_l(x, y)$ is the Harris matrix at level 1 which is computed as smoothed outer product of gradients

$$H_l(x, y) = \nabla_{\sigma_d} P_l(x, y) \nabla_{\sigma_d} P_l(x, y)^T \times g_{\sigma_i}(x, y)$$

σ_i is the integration scale, σ_d is the derivative scale. A point (x, y) is selected as interest point when f_s is local maximum in a 3×3 neighborhood and above a threshold value of 10. Once local-maxima have been detected, their position is refined to sub-pixel accuracy by fitting a 2D quadratic to the corner strength function in the local 3×3 neighborhood and finding its maximum.

B. *QDCT based deep feature extraction*

A image patches of size $m \times n$is extracted with interest point as center. QDCT is applied on each of the image patch to get low frequency and high frequency components. QDCT on the image patch with (x, y) as center is calculated as

$$QDCT(x, y) = A_n^q f(x, y) + \sum_{s=1}^{n} [D_{s,1}^q f(x, y) + D_{s,2}^q f(x, y) + D_{s,3}^q f(x, y)]$$

Where $A_n^q f(x, y)$ is the low frequency band and $D_{s,1}^q f(x, y)$ is the high frequency band of the image.After QDCT is applied on the image a low frequency part, n groups of high frequency parts are obtained.

The frequency of the coefficients is given in Figure 3. To reduce the dimension of the coefficients, average fusion is done for low frequency sub bands. High frequency sub bands are fused using a fusion rule based on maximum value of energy of coefficients. The average fusion rule for fusing the low frequency bands is given as average of the coefficients pair wise between the Low frequency coefficients of two patch images. The fusion rule for fusing the high frequency sub bands is given as selecting the maximum value of coefficient between the pair wise high frequency sub bands.

The QDCT coefficients are given as input to a frequency domain convolutional neural network (Figure 2). The coefficients pass through a sequence of ReLU and max pooling layer and a final average pooling layer to provide an output of 1× 512 dimension feature vector. The CNN configuration used for feature extraction is given in Table 1.

C. *Aggregation signature generation*

An aggregation signature is constructed from the feature vectors belonging to same image patch as below:

- A unit random vector of dimension d (d<512) is generated $\{r_0, r_1, \dots r_d\}$. Each element is sampled from a Gaussian function with mean 0 and variance 1. The d vector is put together into a matrix D of dimension 512 × d. This is generated on time at time of collecting the video as input for tracking.
- A inner product between the feature vectors v and the matrix D is done to get vector $u = D^T v$
- For every vector u, following transformation function tf is applied produce the transformed feature vector $\bar{u}$.
- $tf(u) = \begin{cases} 1\ r.u \geq 0 \\ 0, r.u < 0 \end{cases}$
- $\bar{u} = \{\mathrm{tf_{r1}(u), tf_{r2}(u), \dots . tf_{rd}(u)}\}$
- The feature vectors belonging to same image patch is now represented as bit stream of length d called as aggregation signature of the target image patch.

The benefits of converting the features of same patch to binary bit stream of aggregation signature have two benefits of: compressed form and reduced time complexity for matching the aggregation signature.

D. *Mudra recognition*

Aggregation signature is calculated for each of the mudra classes. From the aggregation signatures belong to each mudra class, a fuzzy membership function is constructed for each class. A dataset of aggregation signature vs mudra class is first constructed. Fuzzy C Means clustering is done on the dataset with P as the number of mudra classes. The cluster center after the fuzzy C means clustering is defined as

$$\mathrm{D} = \{\ \mathrm{D_{e,q}}\ , \mathrm{e} = 1,2 \dots \mathrm{P\ and\ q} = 1,2,3\}$$

Where $\mathrm{D_{e,q}}$ is the qthcoordinating of the eth cluster.

The closeness of the qth feature of the rth data f r,q with qth coordinate of eth cluster is defined using Gaussian function as

$$\mathrm{G}(\mathrm{f_{r,q}}, \mathrm{D_{e,q}}, \sigma_{\mathrm{e,q}}) = \mathrm{e}^{\frac{(\mathrm{f_{r,q}} - \mathrm{D_{e,q}})^2}{\sigma_{\mathrm{e,p}}^2}}$$

Where

$$\sigma_{\mathrm{e,q}} = \frac{1}{\mathrm{N_e}} \sum_{\mathrm{r=1}}^{\mathrm{N_e}} (\mathrm{f_{r,q}} - \mathrm{D_{e,q}})^2$$

The closeness of features of rth data to the eth cluster is given as

$$\Psi_{r,e} = \prod_{q=1}^{P} G\left(f_{r,q}, D_{e,q}, \sigma_{e,q}\right)$$

The output label for eth cluster is found from the linear regression of input features $f_{r,q}$ as

$$\Phi_{r,e} = W_{e,0} + \sum_{q=1}^{P} W_{e,q,f_{r,q}}$$

Where W is the regression coefficient of the eth cluster. Since each of the rth data has membership value to all P clusters, final label of that particular link is given by weighting the label of the link with its membership value as

$$\overline{N}(r) = \sum_{e=1}^{P} \Psi_{r,e}\Phi_{r,e}$$

The value of $\overline{N}(r)$ calculated above may have an error with respect to $N(r)$ from training. The total error is calculated as

$$E = \sum_{r=1}^{N} ||\overline{N}(r) - N(r)||^2$$

The Gaussian parameters $D_{e,q}, \sigma_{e,q}$ and the regression coefficients $W_{e,p}$ are tuned to reduce the error defined above using gradient decent method.

$$D_{e,q}(t+1) = D_{e,q}(t) + \eta_C \frac{\partial E}{\partial D_{e,q}}$$

$$\sigma_{e,q}(t+1) = \sigma_{e,q}(t) + \eta_\sigma \frac{\partial E}{\partial \sigma_{e,q}}$$

$$W_{e,q}(t+1) = W_{e,q}(t) + \eta_W \frac{\partial E}{\partial W_{e,q}}$$

Where t is the iteration number and $\eta_C, \eta_\sigma, \eta_W$ are the learning parameters. The iteration is stopped when error threshold is reached. From training the Fuzzy Gaussian membership functions are obtained for each mudra class in terms of its aggregation signature.

When a mudra image is given, it aggregation signature is calculated. The aggregation signature is passed to fuzzy membership function of each mudra class and the label of the highest value fuzzy membership function is given as output mudra label for the input image.

4. Results

The performance of the proposed solution is tested against Indian classical mudras dataset [20]. The dataset has 10 classes of mudras. The images for rest of 18 classes of mudras were downloaded from websites and Chalearn dataset [21]. The performance of the proposed solution is measured in terms of standard metrics: accuracy, precision, recall and F1-score. The performance of proposed two stage segmentation method is measured in terms of: spatial accuracy index, structural similarity index and Hausdorff distance. The

Table 2 Mudra class wise accuracy

Mudra	Proposed	Jain et al [15]	Anami et al [7]	Anami et al [4]
Pataka 	0.982	0.936	0.835	0.765
Tripataka 	0.971	0.931	0.831	0.761

Ardha-pataka 	0.973	0.922	0.829	0.759
Kartari-Mukha 	0.982	0.941	0.839	0.764
Mayura 	0.951	0.911	0.827	0.752
Ardha-chandra 	0.952	0.901	0.801	0.741
Arala 	0.945	0.891	0.812	0.749
Suk-Tun-daka 	0.951	0.902	0.813	0.747
Mushti 	0.972	0.923	0.821	0.751
Sikhara	0.961	0.927	0.824	0.753

				
Kapitha 	0.952	0.941	0.825	0.753
Kataka-Mukha 	0.971	0.930	0.827	0.753
Suci 	0.961	0.925	0.829	0.757
Chanora-kala 	0.963	0.929	0.830	0.758
Padma-kosha 	0.951	0.931	0.831	0.759
Sarpa-sisha 	0.956	0.912	0.827	0.758
Mrga-sisha 	0.958	0.922	0.820	0.757

Simha-mukhaa 	0.951	0.924	0.815	0.758
Khan-gu-lah 	0.962	0.920	0.813	0.761
Alapadma 	0.971	0.932	0.814	0.761
Chaturo 	0.943	0.903	0.814	0.743
Bhramara 	0.949	0.904	0.821	0.745
Hamasaa 	0.943	0.905	0.816	0.734

performance is measured for each class of mudras. The performance of proposed solution is compared against vertical-horizontal-intersections feature based approach proposed by Anami etl [8], eigen value based matching proposed by Anami et al [5] and Deep convolutional neural network proposed by Jain et al [16].

The performance averages for all the 28 mudra classes are measured and given in Table 1.

Measures	Proposed	Jain et al [15]	Anami et al [7]	Anami et al [4]
Precision	0.963	0.937	0.835	0.787
Recall	0.949	0.936	0.832	0.764

Accuracy	0.961	0.935	0.832	0.763
F1-score	0.958	0.931	0.832	0.759

.

The accuracy in proposed solution is 2.6% higher compared to Jain et al [15], 12.9% higher compared to Anami et al [7] and 19.8% higher compared to Anami et al [4]. The accuracy has improved in proposed solution due to use of interest point localization in multiple scales and extraction of deep learning features around the multi scale interest points. The interest points were localized around the Harris corners and this clearly discriminated between the structurally similar mudras.

The classification accuracy is measured for each of the mudra classes and the result is given in Table 2. From the table, it can be seen that proposed solution has higher discriminating ability (more than 90% accuracy) in recognizing structurally similar mudras like Pataka, TriPataka & Arala , Sarpa-sisha &Chaturo, Mrga-sisha &Simha-mukhaa. The proposed solution performed consistently higher compared to existing works.

The accuracy of mudra recognition is tested by varying the Gaussian pyramid levels used for interest point localization and the result is given in Figure 4.

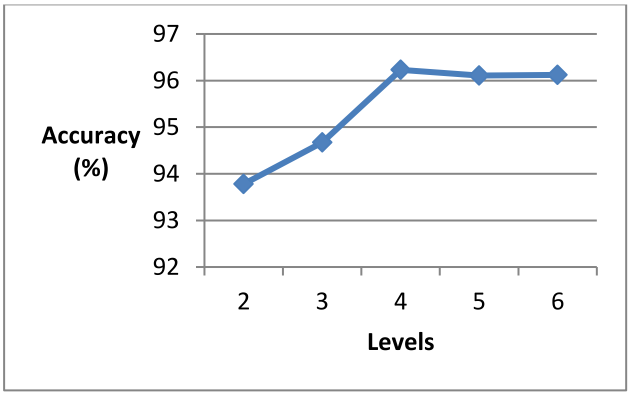

Fig. 4 Accuracy vs levels

As the levels increases, the accuracy increased but for mudra images, after 4 th level, there is not much significant increase in accuracy.

5. Augmented Reality Based Learning

Many choreographers have focused upon implementation of computer technology to improve artistic skills. With lack of trained professionals, learning becomes challenging and Augmented reality (AR) based learning is a solution. Learning through AR helps the students with external amplification, internal rewards, challenge, and increased self-confidence. AR is a technique which makes use of computer vision techniques to collaborate computer generated virtual objects with real-time environment in order to increase or to enhance what can be visualized by the human user. AR has been increasingly used in dance education.

The proposed mudra classification system is integrated with Microsoft Kinect V2 for a proof of concept AR based Mudra learning system.

The architecture of the proof of concept system is given in Figure 5

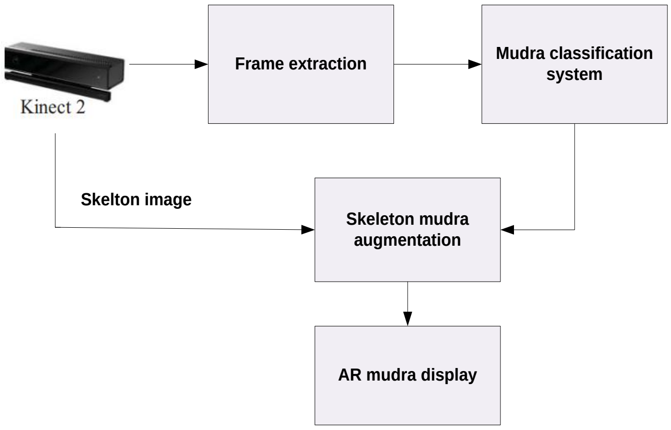

Fig. 5 Proof of concept mudra learning system

Video frames are extracted are extracted from Kinect 2 video camera. The frames images are passed to Mudra classification system to recognize mudras. Once the mudra is recognized, the skeletal template image of mudra is replaced in the hand region of the skeletal image provided by Kinect 2 to generate a augmented dance skeleton with mudra and mudra label is also displayed. By this way leaner can mudras effectively. Some of the mudra augmented skeleton AR scenes generated by the proof of concept system is given in Figure 6

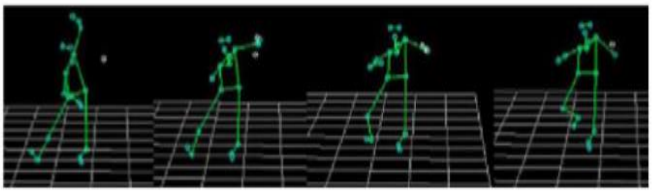

Fig. 6 Mudra augmented skeletal images

6. Conclusion

An aggregation signature constructed based on deep features extracted from multi scale interest points was proposed in this work for mudra recognition.

Interest points were found using Harris corner based strength function constructed based on multi super scale Gaussian transformed images. Deep learning features around the interest points are aggregated to a signature. Fuzzy membership function is constructed to map the aggregation signature to mudra classes. The proposed solution was able to achieve an average accuracy of 96.1% which is atleast 2.6% higher compared to existing works. Extending the solution for double hand mudras is in scope of future work.

References

[1] Mallik, A., Chaudhury, S., and Ghosh, H. Nrityakosha: Preserving the Intangible Heritage of Indian Classical Dance. Journal on Computing and Cultural Heritage (JOCCH) 4, 3 (2011), 11

[2] Tanwi Mallick, Partha Pratim Das, Arun Kumar Majumdar:Bharatanatyam Dance Transcription using Multimedia Ontology and Machine Learning. CoRR abs/2004.11994 (2020)

[3] Sangeeta Jadav, and Sasikumar Mukundan, "A Computational Model for Bharata Natyam Choreography," (IJCSIS) International Journal of Computer Science and Information Security, Vol. 8, No. 7, October, 2010.

[4] Anami, B.S., & Bhandage, V.A. (2018). A Comparative Study of Suitability of Certain Features in Classification of Bharatanatyam Mudra Images Using Artificial Neural Network. *Neural Processing Letters, 50*, 741-769.

[5] Saha, S., Ghosh, L., Konar, A. and Janarthanan, R. (2013) 'Fuzzy L membership function based hand gesture recognition for Bharatanatyam dance', Proceedings of the 5th International conference on Computational Intelligence and Communication Networks (CICN), pp.331–335, IEEE

[6] A. P. Parameshwaran, H. P. Desai, R. Sunderraman and M. Weeks, "Transfer Learning for Classifying Single Hand Gestures on Comprehensive Bharatanatyam Mudra Dataset," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019, pp. 508-510

[7] B.S. Anami and V.A. Bhandage, "A vertical-horizontal-intersections feature based method for identification of bharatanatyam double hand mudra images", *Multimedia Tools and Applications*, vol. 77, no. 23, pp. 31021-31040, 2018

[8] K.V.V. Kumar and P.V.V. Kishore, "Indian Classical Dance Mudra Classification Using HOG Features and SVM Classifier", *International Journal of Electrical & Computer Engineering (2088–8708)*, vol. 7, no. 5, 2017

[9] Kopuklu, Okan & Gunduz, Ahmet & Kose, Neslihan & Rigoll, Gerhard. (2019). Real-time Hand Gesture Detection and Classification Using Convolutional Neural Networks. 1-8. 10.1109/FG.2019.8756576.

[10] Sahoo JP, Prakash AJ, Pławiak P, Samantray S. Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network. *Sensors*. 2022; 22(3):706.

[11] Patil, A.R.; Subbaraman, S. A spatiotemporal approach for vision-based hand gesture recognition using Hough transform and neural network. Signal, Image Video Process. 2019, 13, 413–421

[12] Fang, L.; Liang, N.; Kang, W.; Wang, Z.; Feng, D.D. Real-time hand posture recognition using hand geometric features and fisher vector. Signal Process. Image Commun. 2020, 82, 115729

[13] Gadekallu, Thippa & Srivastava, Gautam & Liyanage, Madhusanka & Meenakshisundaram, Iyapparaja & Chowdhary, Chiranji & Koppu, Srinivas & Reddy, Praveen. (2022). Hand gesture recognition based on a Harris Hawks optimized Convolution Neural Network. Computers & Electrical Engineering. 100. 107836. 10.1016/j.compeleceng.2022.107836.

[14] Lim, Kian & Tan, Alan & Tan, Shing. (2016). A Feature Covariance Matrix with Serial Particle Filter for Isolated Sign Language Recognition. Expert Systems with Applications. 54. 10.1016/j.eswa.2016.01.047.

[15] Jain N, Bansal V, Virmani D, Gupta V, Salas-Morera L, Garcia-Hernandez L. An Enhanced Deep Convolutional Neural Network for Classifying Indian Classical Dance Forms. *Applied Sciences*. 2021; 11(14):6253.

[16] Kumar, K.V.V.; Kishore, P.V.V.; Kumar, D.A. Indian Classical Dance Classification with Adaboost Multiclass Classifier on Multifeature Fusion. Math. Probl. Eng. 2017, 2017, 1–18.

[17] Kishore, P.V.V.; Kumar, K.V.V.; Kumar, E.K.; Sastry, A.S.C.S.; Kiran, M.T.; Kumar, D.A.; Prasad, M.V.D. Indian Classical Dance Action Identification and Classification with Convolutional Neural Networks. Adv. Multimedia 2018, 2018, 1–10

[18] S. Mozarkar and C. S. Warnekar, "Recognizing bharatnatyam mudra using principles of gesture recognition gesture recognition," International Journal of Computer Science and Network, vol. 2, no. 2, pp. 46–52, 2013

[19] Kishore, P.V.V., Kishore, S.R.C., Prasad, M.V.D.: Conglomeration of hand shapes and texture information for recognizing gestures of Indian sign language using feed forward neural networks. Int. J. Eng. Technol. (IJET), ISSN: 0975-4024 (2013).

[20] https://www.kaggle.com/datasets/ishanishah8/indian-classical-mudras-classification

[21] https://www.kaggle.com/competitions/GestureChallenge/overview/data-description-2