ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

A Deep Learning Model to Detect Kathak Facial Expression

ApekshaKhopkar[1], Dr.Ashish Adholiya[2]

[1]Research Scholar, Pacific University, Udaipur, India

[2] Assistant Professor MMS, Pacific Institute of Management, Udaipur, India

Email-[1]apeksha.gaje@gmail.com, [2] asia_1983@rediffmail.com

ABSTRACT:

Emotion" is referred to as "rasa," while "state of mind" is referred to as "bhava." Kathak is a kind of traditional Indian dance that is renowned for the expressive movements and facial emotions that are included in the performance. There hasn't been a lot of studies done on the facial expressions that Kathak dancers use. Deep learning is a strong technique that may be used for the task of analyzing and classifying the facial expressions of Kathak dancers. In order to find an answer to this research issue, a dataset consisting of the facial expressions shown by Kathak dancers will be compiled and then subjected to deep learning algorithms for analysis. The findings of this research will help us get a better knowledge of the facial expressions used by Kathak dancers as well as the feelings that are communicated via this style of dance. Kathak face expressions are nine types namely- *Adbhut*(wonder), *Hasya*(joy), *Shringar*(affection), *Bheebatsya*(loathing), *Bhayanak*(terror), *Roudra*(anger), *Veer*(pride), *Karuna*(sadness) and *Shant*(peace). These expressions are mapped to Kathak dancers' expressions for better classification through face detection and extraction. Image preprocessing and classification are performed to achieve maximum accuracy. Using Convolutional Neural Network(CNN). The scope of the paper is to identify the facial expressions of Kathak dancers to have a better understanding of the art in the teaching and learning paradigm.

Keywords: Deep Learning, Facial Expressions Recognition (FER), Convolutional Neural Network CNN, Kathak.

INTRODUCTION:

Every single human person on the globe communicates with one another via movement on the most fundamental level. It conveys the meaning of what cannot be spoken in any other way (Takalkar& Xu, 2017). Indian dances are intimately connected to a variety of emotions and are sometimes described as having the appearance of a perfect balance between body movement and classical music. People in India are able to easily connect with Indian Classical Dances since they are a reflection of India's traditional culture and the individuals in India find it simple to do so. These dances communicate a variety of sentiments and moods via a mix of physical movements, postures, and facial gestures. *Bharatnatyam, Odissi, Kathak, Kuchipudi, Manipuri, Kathakali,* and *Mohiniattam*are the seven most important forms of classical dance in India (Grover, 2015). On a fundamental level, there are six primary expression categories, including *"joy," "sadness," "terror," "anger,"* and *"loathing."*

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

Fig. 1. Nine Facial Expressions of Kathak Dancer

- *Shant*(peace): The nose and mouth of this expression are in their relaxed shape. Eyes are closed. Eyebrows here are very close to the eye line with an equivalent distance. This shows the peaceful posture of the face. (figure.1:- row:1, 1st image)
- *Raudra*(angry): Here, eyes are pointed downwards. Nose is elongated and eyebrows are in the shape of pointed inverted parabola. The mouth length depicted is very less. This expression shows the anger emotion. (figure.1:- row:2,2nd image)
- *Veer*(pride): Mouth and eyebrows in this expression are moved almost straight. It is mostly connected with a king's face and his act of pride. The eyes are made to a little elongated shape with uplifted eyebrows. Nose is moved little sideways. (fig.1:- row:1,3rd image)
- *Bhibatsya*(loathing): In this expression, eyebrows are shrunk toward the nose.Mouth here is in the shape of an inverted parabola. The eyes are reduced to 3/4th of the original size. Nose is compressed in length. Something really awful is depicted in the expression. (fig.1:-row:2,1st image)
- *Adhbut*(wonder): In this expression, eyes are almost circular and completely open. Nose is little projected outward and the mouth part is a little shrunk. The expression shows unexpected happiness here. (fig.1:- row:2,2nd image)
- *Bhayanak*(terror): Mouth is shaped like an inverted parabola. Both eyes appear to-ward the center of the nose. It shows the facial expression of having fear of some- thing. (fig.1:- row:2,3rd image)
- *Hasya*(joy): It is very similar to sringara but the difference is that eyes are opened to it's normal size. The eyebrows are a little raised from their original position. Nose is in its original shape. It is used to show funny instances. (fig.1:- row:3,1st image)

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

- *Karun* (sadness): In this expression, eyes are too hard to be found out in size. Mouth is shaped like an inverted parabola but the length is too small. This expression shows feeling sympathy for someone. (fig.1:- row:3, 2nd image)
- *Shringar*(affection): A very pleasant smile is expressed on the face. The eyebrowsare perfectly parallel to the eyes and the original size of mouth is retained here. The eyes are almost half of its actual size slightly to the corner. Nose is in its original shape and usually depicts a face of love. (fig.1:- row:3, 3rd image)

It is possible to use deep learning to automatically extract facial expressions from video footage of Kathak dancers. These facial expressions may then be statistically examined to get a better understanding of the spectrum of emotions that can be represented via this dance style. This study may provide light on the significance of facial expressions in Kathak, as well as how different feelingsare communicated via the various components of the dance. Convolutional neural networks(CNN) are one of the important categories in neural networks for image processing, classification, segmentation and also for other autocorrelation data. So this is the one main reason to choose this method for Kathak face expression detection.

CNN algorithms provide good accuracy when there is a large amount of data avail- able for training. The computation cost is very high and the algorithm requires good GUI , otherwise, a lot of time is invested in training. There exists no previous dataset for Kathak facial expressions to the best knowledge. The works that were previously done are concentrated on dance forms like Kuchipudi, Bharatanatyam, Odissi, etc. Creating a dataset for all expressions from scratch was the main challenge. Guidance of Kathak exponents was taken for the collection of data. More than 1150 images of the novel dataset for Kathak face expressions in which every nine expressions has around 100-105 images. The diversity of the dataset was taken into consideration. Factors considered are different positions by different people, gender, background, some pictures with makeup and some without makeup, etc., in creating the dataset.

The intention of the work is due to the following reasons:

- The findings of this research will help us get a better knowledge of the facial expressions used by Kathak dancers as well as the feelings that are communicated via this style of dance.
- However, these facial expressions cannot be easily understood by a common human being. this study has the potential to aid in understanding the facial expressions that Kathak dancers use to portray their emotions as well as the feelings that they are trying to convey.

This work is instantiated due to the following motives:

1. Most research has been done on hand gestures and body movements in Indian classical dance except for Kathak dancer's facial expressions recognition.

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

2. No dataset for Kathak Facial Expressions. This work proposes a FER dataset of Kathak dancers. The implication of this model will be helpful for online teaching and learning of Indian classical dance, especially facial expression.

LITERATURE REVIEW:

2.1 Facial Expressions in Indian Classical Dance – Kathak

ANatya performance uses dance as a theatre medium and uses role-playing through dance to tell intricate tales. Both Nritya and Natya are interested in communicating concepts, themes, moods, and emotions (Massey and Reginald, 1999). Figure 2.1 in (Massey and Reginald, 1999) depicts many kinds of Natya.

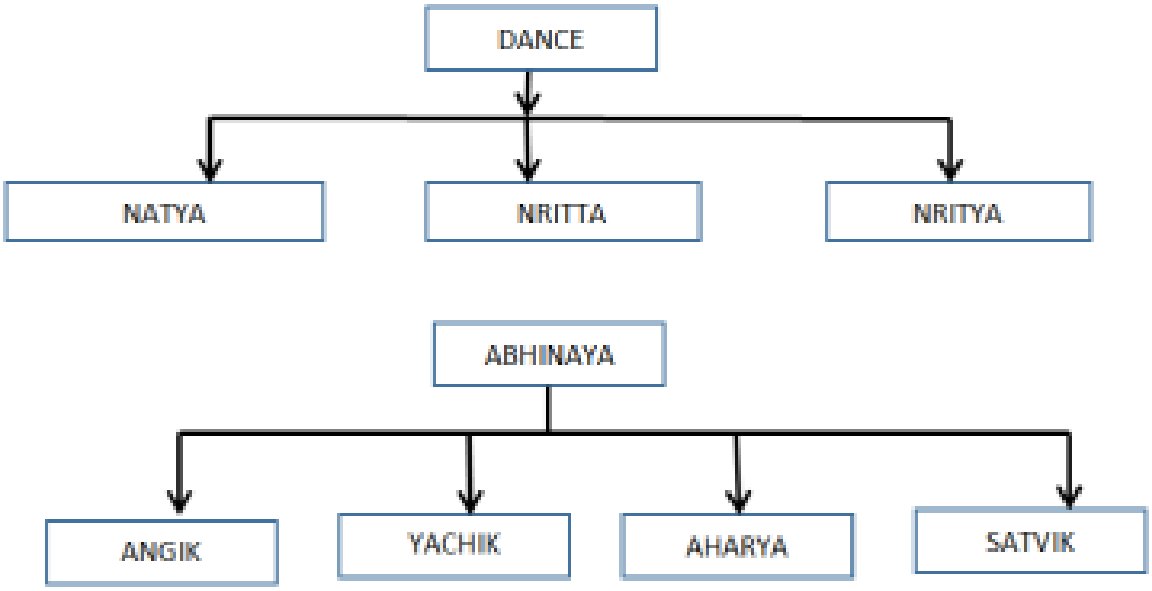
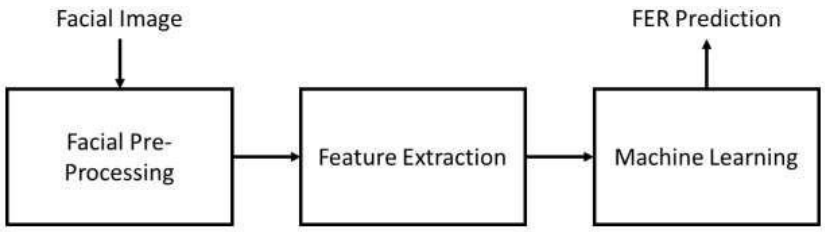


Fig. 2.1. Categories of dance according to Natya Shastra

2.2 Deep Learning Based Facial Expressions Recognition (FER) approach

Facial Emotion Recognition (FER) is a booming research field where industry advancements like machine-to-machine communication and automatic translation systems are taking place. In order to categorize facial emotions, computer vision and deep learning are being used in the non-verbal field of facial expression recognition. (Revanth Krishna). Understanding how a typical system is created and the steps necessary tomove from image to expression is necessary in order to create an effective FER system. As illustrated in Fig. 2.3



ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

Fig. 2.2. Facial Emotion Recognition

In this paper, while curating the dataset we have considered professional dancers of Indian Classical Dance (ICD) in an unrestricted environment, a dataset of the Navarasa was created for a deep learning-based module that uses Convolutional Neural Networks (CNN) to recognize emotions. There was no work done in the area of facial expression analysis in relation to Navras through the recognition of expressions specifically for the classical dance Kathak.

The two most well-known methods for classifying any type of data are machine learning [Pandey. S. et al., 2018] and deep learning [SudarshanaTamuly et al., 2019]. Deep learning is regarded as one of the most accurate techniques for categorizing image data into different labels. Convolution Neural Network (CNN) plays a significant role in this situation in extracting the key features and labeling them appropriately. Four steps make up CNN [SudarshanaTamuly et al., 2019]: convolution, pooling, flattening, and full connection. Here, a kernel of x*y size is convolved with an image of M*N size, hence the dot product of the original image and the kernel must be found. The first layer to extract features from the supplied images is convolution. For various tasks, such as identification, sharpening, and edge detection, many filters are offered. The following stage is pooling, which is used to lower the parameter when the size of the image is too huge. Even while pooling reduces the image's size, it keeps the crucial details. The matrix is then flattened into vector form and input into a neural network with fully connected layers. Adding many convolutional layers and pooling results in the extraction of anticipated features and a very accurate classification of the image.

2.3 Literature Survey

Table 2.1. Literature survey

Author	Paper title	Work done	Performance measure
Srimani P.K et al.,2021	Analysis Of Facial Expressions With Respect To Navarasas In Bharathanatym Styles Using Image Processing	This author utilizes image processing techniques to study facial expression analysis in relation to Bharatanatyam, a traditional dance form from southern India. Several aspects of the face are taken into account.	It has been shown that most of the time, a face's alteration in kurtosis corresponds to a change in navarasa expression. The observed modifications amount to 85%. But when We evaluated the facial expressions without cosmetics, the alterations were determined to be 95% equivalent.
Mo-Hantyand Sahay, (2018)	"Rasabodha : Understanding Indian classical dance by	In an effort to explain the relevance of "Navarasas" in relation to ICD, this study proposes a dataset of various	There are many opportunities to strengthen the offered technique and manage the 725 difficulties associated with real world

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

Author	Paper title	Work done	Performance measure
	recognizing emotions using deep learning"	emotions ("Navarasas") displayed in ICD that includes "RGB" pictures and additional depth information. The "Microsoft Kinect sensor" is used to acquire this data. It is advised to use CNN's "deep learning frame-work" to understand the semantic meaning associated with "ICD" films by recognising "Navarasas"supported by the artist.	footage of "ICD" with the help of the emotions portrayed by the dancer. Future research will concentrate on con- trolling obstacles and variants for identifying "Navarasas" in realworld dancing films.
SudarshanaTamuly et al.,2018	Deep Learning Model for Image Classification	This work attempts to design a way to automatically classify images as pleasant, unpleasant, and neutral using deep learning and tools related with it.	A model with a VGG-16 architecture was suggested, and when it was applied to the IAPS dataset with 1037 pictures, the training accuracy was 61% and the testing accuracy was 54%. The IAPS, NUSEF, and MIT-300datasets were used once more, and the obtained training accuracy was 71% and the achieved testing accuracy was 65%. Finally, the training accuracy was 88% and the testing accuracy was 80% after applying to 20 pictures from the IAPS dataset and its accompanying heat and focus maps of 220 images.
P. V. V. Kishore et. al., 2018	Indian classical dance action identification using adaboost multiclass classifier on multifeature fusion	This research uses un-constrained video sequences from Indian traditional dance styles to solve a challenging class issue. Discrete wavelet transform and local binary pattern (LBP) characteristics are used to create a novel segmentation model. Following video frames' local changes in human form are used to construct a 2Dpoint cloud.	Due to the training and testing datasets utilised in this experiment, there is extremely little false matching. According to the findings, the dataset's dancing videos have an average score of 0.99.
Basavaraj S.	A Comparative	In this instance, preprocessing	This method was able to

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

Author	Paper title	Work done	Performance measure
et al., 2019	Study of Suitability of Certain Features in Classification of Bharatanatyam Mudra Images Using Artificial Neural Network	images using a smart edge detector is the first step in obtaining the mudra's contour. The second phase involved extracting humoments, eigenvalues, and intersection features. The third stage involved classifyingmudras using an artificial neural network (ANN).	categorise 28 distinct types of the asamyuktha (single hand) mudra employed in Bharatanatyam with 93.64% accuracy when run across 50 epochs.
Anuja P et al., 2019	Transfer Learning for Classifying Single Hand Gestures on Comprehensive Bharatanatyam MudraDataset	Using CNN, the Bharatanatyam mudras were categorised, and two models—transfer learning and double transfer learning— were used for testing.	The dataset had 27 classes of 2D hand mudras, and after training and testing, the dataset's accuracy scores were 94.56% and 98.25%, respectively.

PROPOSED ARCHITECTURE:

3.1 The image capturing and curating of a dataset:

To curate, the dataset of each expression guidance from profound Kathak artists of Mumbai was taken. Images collected had different angles for each expression. These images were saved in .jpg format on the drive storage. A wide collection of 1000-plus images where captured processed, and cleaned for each of the nine expressions.

3.2 Pre-processing and noise elimination:

1. Prewitt filters are applied to remove noise and also to get a clearer image.
2. Gray scaling and resizing of the images are used to increase the efficiency of the CNN classifier.
3. The unfilled pixels or scattered pixels are replaced by using the nearest neighbor method which uses the pixel values of the nearest pixel.
4. Fuzzy Algorithm rulesets are considered.

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

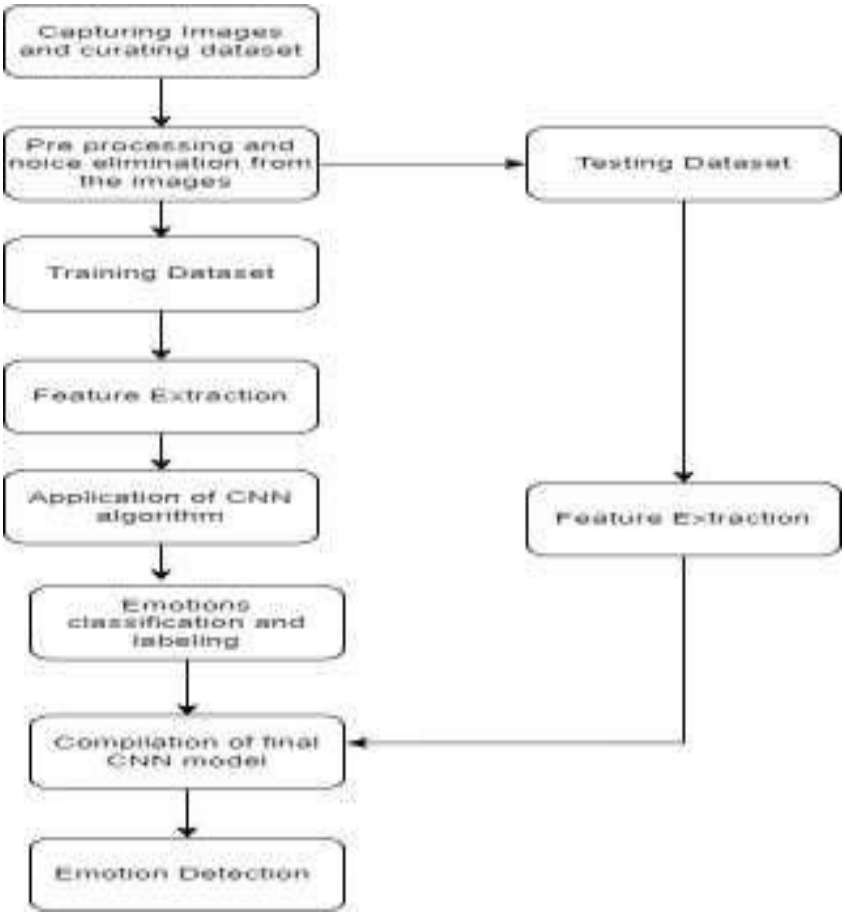


Fig. 3. Proposed flow of the architecture

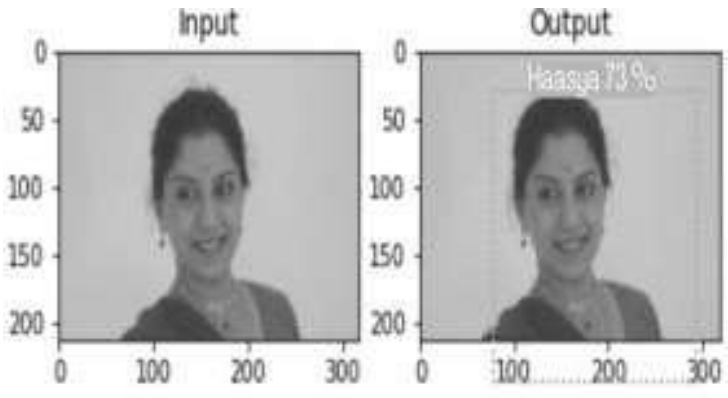


Fig. 3.1. Pre-proposed flow of the image

3.3 Training and validation:

1052 plus images, in order to have better accuracy the chosen ratio of training to testing dataset is 70:30. 70% of the training data that is 735 images are equally distributed among 9 expressions. Each of the 9 expressions now contains 81 training images. 30% of the testing data that is 315 images are equally distributed among 9 expressions. Each of the 9expressions

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

now contains 35 testing images. This split is adopted because the data is redundant and inconsistent.

3.4 CNN classifier:

The Classifier CNN has a special architecture to detect complex features in data. CNNs are involved with many layers starting with an input layer followed by hidden layers usually consisting of convolutional layers, ReLU layers, pooling layers and an Output layer. Using a stack of convolution layers and max pooling for the exact detection of images, a model was made and a dense layer was applied for the detection of expressions using relu as an activation function.

3.5 Performance Evaluation and Measures:

- Training accuracy: Accuracy obtained during the training of the dataset.
- Validation accuracy: Accuracy obtained during the testing of the dataset.
- Training loss: Training loss is the error in the training set of data.
- Validation loss: Validation loss is the error after running the validation set of data through the trained network.

RESULT ANALYSIS AND DISCUSSION:

4.1 Dataset

The algorithm allows only jpg format for the image dataset and so the images are all in jpg format. Also, there are 9 classes of expressions dataset which is distributed in 70% and 30% ratio for training and testing respectively.

4.2 Tools or Techniques used:

Packages used for general plotting of graphs and reading of datasets are matplotlib, numpy, seaborn, pandas, os, and sys. The packages used for cnn training are Keras, TensorFlow, and image data generator.

4.3 Pre-processing Techniques:

Filters like the Prewitt filter are used for removing noise in the data. Relu (Rectified linear unit) and SoftMax are the activation functions used for quick convergence and optimization. Relu relies only on positive values of the outputs whereas Softmax con- siders negative values as actual outputs. Image is passed through the first stack of 2 convolution layers of the very small receptive size of 3 x 3, followed by ReLU activations, each with two layers containing 64 filters. The convolution stride and the padding are fixed at 1 pixel. This configuration preserves the spatial resolution, and the size of the output activation map is the same as the input image dimensions. The activation maps are then passed through spatial max

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

pooling over a 2 x 2-pixel window, with a stride of 2 pixels. This halves the size of the activations. Thus the sizeof the activations at the end of the first stack is 112 x 112 x 64.

The activations then flow through a similar second stack, but with 128 filters as against 64 in the first one. Consequently, the size after the second stack becomes 56 x 56 x 128. This is followed by the third stack with three convolutional layers and a max pool layer. The number of filters applied here is 256, making the output size of the stack 28 x 28 x 256. This is followed by two stacks of three convolutional layers, with each containing 512 filters. The output at the end of both these stacks will be 7 x 7 x 512.

The stacks of convolutional layers are followed by three fully connected layers with a flattening layer in between. The first two has 4,096 neurons each, and the last fully connected layer serves as the output layer and has 1,000 neurons corresponding to the 1,000 possible classes for the ImageNet dataset. The output layer is followed by the Softmax activation layer used for categorical classification.

4.4 Parameter tuning

For nine classes, the CNN model was compiled using sparse categorical cross entropy which is for compiling the methods that have more than two classes.

4.5 Prediction for validation dataset

The dataset is classified accordingly into 9 expressions and was split into training and testing datasets. CNN predefined model has been tested for accuracy using a confusion matrix. The evaluation was done for a few testing samples to identify the expression.

Mean training accuracy is around 88% and mean validation accuracy is around 66%

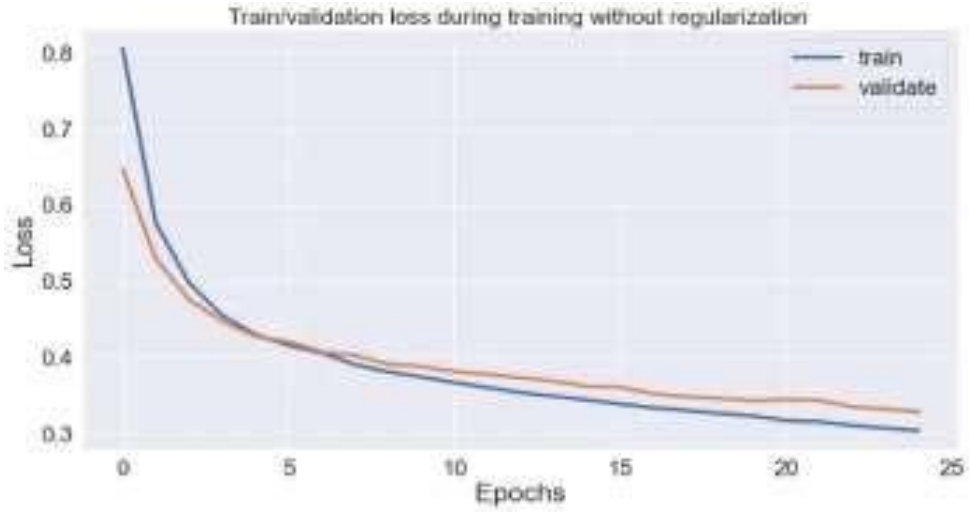


Fig. 4.1. Training/Validation loss without regularization

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

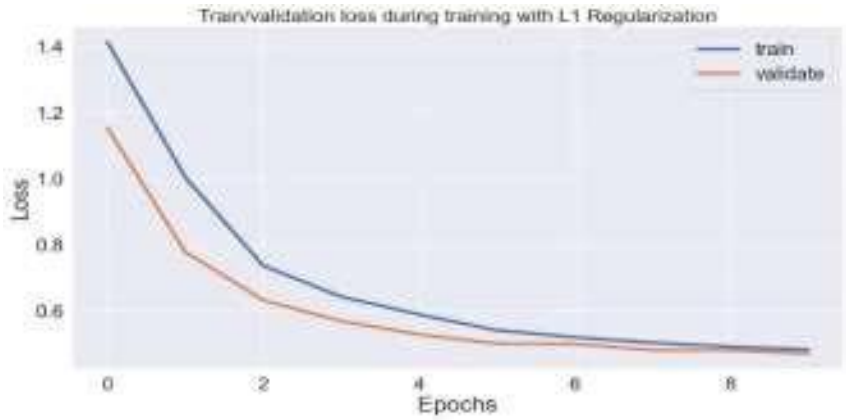


Fig. 4.2. Training/ Validation loss during L1 regularization

4.6 Expression detection:

The training of CNN is done using Keras and the result is mentioned in figure 9. After training the class labels were collected for 9 facial expressions and at last, the emotion for a particular expression was detected.

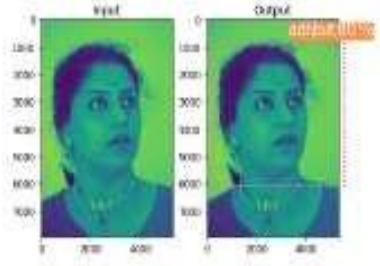


Fig. 4.6.1. Testing Adhbut expression

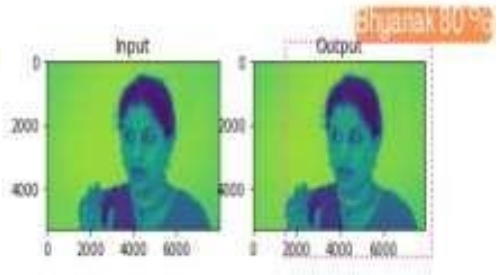


Fig. 4.6.2. Testing Bhayanak expression

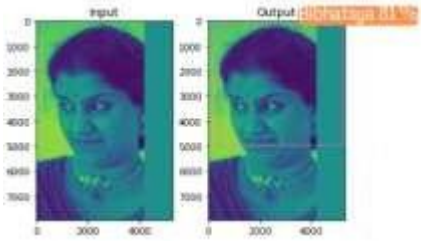


Fig. 4.6.3 Testing Bibhatsya expression

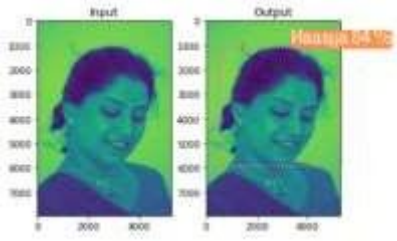


Fig. 4.6.4 Testing Hassya expression

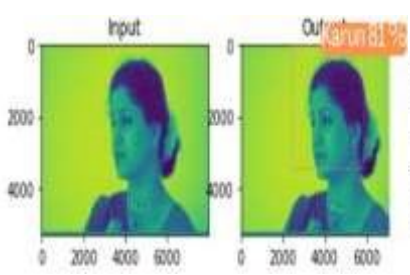


Fig. 4.3.5 Testing Karun expression

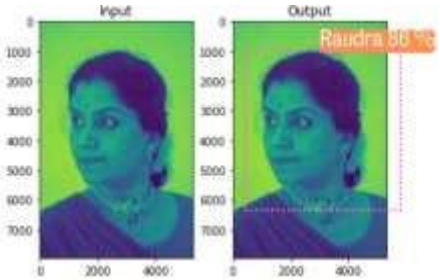


Fig. 4.3.6 Testing Rudra expression

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

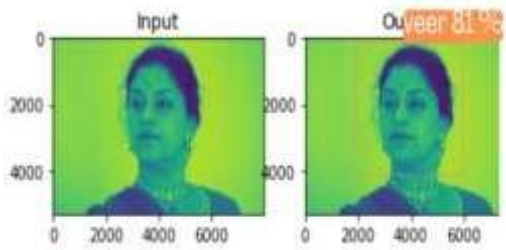


Fig. 4.3.7 Testing Veer expression

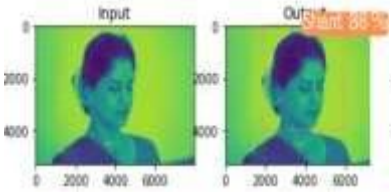


Fig. 4.3.8 Testing Shant expression

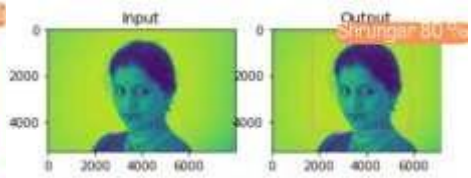


Fig. 4.3.9 Testing Shrungar expression

CONCLUSION AND FUTURE WORK:

In this paper, the dataset is created for nine expressions of Kathak with the help of the classical dancers. The dataset consists of 1052 images with 70% training images and 30% testing images. The features were extracted from the designed Kathak Facial Expressions Dataset using max pooling layers and reached an accuracy of 88%. This is the first attempt of this kind to create a dataset for facial expressions and classify the expressions using deep learning techniques. However, the process of turning the right parameters could only make the model successful. The drawback is that Kathak expressions have very fine differences between the two expressions sometimes can be over- lapping. Which makes the model not predict well at times? Hence, this issue can be resolved only if more samples of the datasets of classical dance facial expressions areavailable. This research will surely contribute to the teaching and learning fraternity of Indian classical dance.

REFERENCES:

1. Takalkar, M. A., & Xu, M. (2017). Image Based Facial Micro-Expression Recognition Using Deep Learning on Small Datasets. *DICTA 2017 - 2017 International Conference on Digital Image Computing: Techniques and Applications*, *2017-December*. https://doi.org/10.1109/DICTA.2017.8227443

2. Grover, B. (2015). NEW EXPERIMENT IN KATHAK DANCE: ANALYTICAL STUDY. *International Journal of Research -GRANTHAALAYAH*, *3*(1SE). https://doi.org/10.29121/granthaalayah.v3.i1se.2015.3446

3. Grover, B. (2021). ILLUSTRATION OF NATURE IN THE BANDS OF KATHAK RAIGAD GHARANA. *ShodhKosh: Journal of Visual and Performing Arts*, *2*(2). https://doi.org/10.29121/shodhkosh.v2.i2.2021.31

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, S Iss 2, 2022

4. Massey, Reginald. India's Kathak Dance, Past Present, Future. Abhinav Publications,1999.

5. SudarshanaTamuly, C. Jyotsna, Amudha J, "Deep Learning Model for Image Classification", International Conference On Computational Vision and Bio Inspired Computing (ICCVBIC 2019).

6. SudarshanaTamuly, C. Jyotsna, Amudha J, "Effective spam image classification using CNN and transfer learning", 3rd International Conference on Computational Vision and Bio In- spired Computing, ICCVBIC 2019

7. Mohanty, A., & Sahay, R. R. (2018). Rasabodha: Understanding Indian classical dance by recognizing emotions using deep learning. Pattern Recognition, 79, 97–113.

8. Kishore K. Reddy and Mubarak Shah, "Recognizing 50 human action categories of web videos", published in springer, pp. 971–98, November 2012.

9. Kishore, K. V. V. Kumar, E. Kiran Kumar, A S. C. S. Sastry, M. Teja Kiran, D. Anil Kumar and M. V. D. Prasad, "Indian Classical Dance Action Identification and Classification with Convolutional Neural Networks", article in Hindawi:5141402, 2018

10. Kishore, K. V. V. Kumar, E. Kiran Kumar, D. Anil Kumar," Indian classical dance action identification using adaboost multiclass classifier on multifeature fusion" , In Proceedings of the IEEE Conference:17632790 , Jan 2018.

11. Basavaraj S. Anami ,Venkatesh A, Bhandage , "A Comparative Study of Suitability of Certain Features in Classification of Bharatanatyam Mudra Images Using Artificial Neural Net- work", Published in Springer (Volume 50, pp. 741–769, August 2019).

12. Anuja P. Parameshwaran, Heta P. Desai, RajshekharSunderraman Michael Weeks," Transfer Learning for Classifying Single Hand Gestures on Comprehensive Bharatanatyam Mu- dra Dataset", In proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)(pp. 0-0, 2019).

13. C Selvi, Y Anvitha, C H Asritha and P B Sayannah "Kathakali face expression detection using deep learning techniques", Journal of Physics: Conference Series 2062 (2021) 012018 IOP Publishing doi:10.1088/1742-6596/2062/1/012018

14. Khopkar, A., &Adholiya, A. (2021). Facial expression recognition using CNN with Keras. *BiosciBiotechnol Res Commun*.

15. Khopkar, A., &Adholiya, A. (2022). An Exhaustive Survey of Deep Learning Models for Analysis and Classification of Facial Expressions of Kathak Dancers Asian Journal of Organic & Medicinal Chemistry.