Jyothi Engineering College
NAAC Accredited College with NBA Accredited Programmes*
Approved by AICTE & affiliated to APJ Abdul Kalam Technological University
A CENTRE OF EXCELLENCE IN SCIENCE & TECHNOLOGY BY THE CATHOLIC ARCHDIOCESE OF TRICHUR
JYOTHI HILLS, VETTIKATTIRI P.O, CHERUTHURUTHY, THRISSUR. PIN-679531 PH : +91- 4884-259000, 274423 FAX : 04884-274777
NBA accredited B.Tech Programmes in Computer Science & Engineering, Electronics & Communication Engineering, Electrical & Electronics Engineering and Mechanical Engineering valid for the academic years 2016-2022. NBA accredited B.Tech Programme in Civil Engineering valid for the academic years 2019-2022.

Bhavas Classification using Deep Learning

June 10, 2021

Group No: 24

Group Members :

Nova Dileep	**JEC17CS078**
Shinoz Mohammed P P	**JEC17CS094**
Sreelakshmi C (Team Leader)	**JEC17CS098**

Name of mentor:

Ms. Reshma K V

Assistant Professor

Vision of the Department

- Creating eminent and ethical leaders in the domain of Computational Sciences through quality professional education with a focus on holistic learning and excellence.

Mission of the Department

- To create technically competent and ethically conscious graduates in the field of Computer Science and Engineering by encouraging holistic learning and excellence.
- To prepare students for careers in Industry, Academia and the Government.
- To instill Entrepreneurial Orientation and research motivation among the students of the department.
- To emerge as a leader in education in the region by encouraging teaching, learning, industry and societal connect.

Contents

Introduction

- Facial expressions are important factors in human communication that helps us to understand feelings and intentions of others
- It is one of the major significant features of recognition of human emotion
- Use the facial recognition techniques to identify different ‘Bhavas’ in the dance form


Sringara-Love
Hasya-Comedy
Karuna-Pity
Raudra-Furious
Veeryam-Heroic
Bhayanakam-Fearful
Beebhatsa-Disgust
Adbhutha-Wonder
Santham-Peace

- Sringara : Happy
- Hasya : Comic
- Karuna : Compassion
- Roudra : Anger
- Veera : Valor
- Bhayankara : Fear
- Bheebhatsya : Disgust
- Adbhutha : Wonder
- Shanta : Tranquility

- Bhavas classification is used to classify the ‘navarasam’ of classical dance
- Easily understands the meaning of each bhavas

Problem Statement

- Non trained dancers can't understand meaning of navarasams(bhavas) in classical dances
- Can't enjoy the dance performances in it's full extent
- Existing system classifies only human emotions and not bhavas
- This system identifies and classifies bhavas and therefore is a better option

Existing system

1. Convolutional Neural Network(CNN) is used to identify 6 elementary emotions
    - i.Happy
    - ii.Sad
    - iii.Fear
    - iv.Anger
    - v.Surprise
    - vi.Disgust

2.Classification is done through simple image processing

3.Different types of layers are present

i.Input layer

ii.Convolution layer

iii.Activation layer

iv.Pooling layer

a.Mean pooling

b.Max pooling

v. Fully connected layer

vi.Softmax layer

4. Accuracy of 91.6 % has been achieved by the test result

Advantages & Disadvantages

1. Advantages
   - i. High accuracy
   - ii. Automatic extraction of features
2. Disadvantages
   - i. Classifies input images to atmost 5 facial expressions
   - ii. Less accuracy for disgust and surprise expressions

Literature Survey

Sl.No	Objective	Dataset	Methodology	Conclusion	Contribution to our project	Disadvantage
1	1.Convolutional Neural Network (CNN) 2. Identify 6 basic emotions 3.Implemented in MATLAB	JAFFE (Japanese Female Facial Expression)	1.Take the input image 2.Image is pass through CNN 3.Feature extraction and Classification is performed 4.Accuracy	1. complete view of facial emotion recognition system. 2.An accuracy of 91.6 % has been achieved by the test result.	1.Gives a complete view of facial emotion recognition system 2.CNN algorithm can be used for the project	1. Classifies input images to atmost 7 facial expressions 2. less accuracy for disgust and surprised expressions

2.	1.Detect people's emotions in natural conditions 2.outperform traditional approaches 3.Human-level performance 4..Convolutional Neural Network (CNN)	FER-2013	1.Take the input image 2.Image is passed through a)Baseline model b)Five layer model c)ResNet50 model d)ED NET 3.Feature extraction and Classification is performed 4.Accuracy calculation	1.Accuracies of some of the models achieved on the FER2013 test data a)Human_level - 65±5% b) Yichuan Tan-71.2% c) Shervin Minaee-70.02% d) Baseline -61% e) Five-layer model -66.2% f) ResNet50 - 73% g) ED NET - 73.9% h) State_of_the_art - 76.8%	1. CNN algorithm can be used for the project	1. Only 7 facial expression 2. More time is required

3.	1.To improve the accuracy in classifying emotions 2. To classify the facial expressions to angry, happy, fear, neutral, sad and surprise using CNN architecture	Combination of CK+, RaFD, MUG, KDEF	1.Take the input image 2.Fine tuning CNN architecture by VGG 3.Fine tuning CNN architecture by first model 4.Accuracy calculation	1.Accuracies obtained on different datasets: a)CK+ - 99.33% b)MUG – 87.65% c)RaFD – 93.33% 2.Highest accuracy is obtained for CK+	1. CNN algorithm can be used for the project	Considered images which has face in one position only

4.	1. To avoid the explicit feature extraction process 2. To Classify facial expressions to worried, angry, disgust, surprise, anxious, happy, sad, neutral emotions	CLDC	1.Image is normalized 2.Features are extracted and dimension is reduced using convolution layers and max pooling 3.RPN is used for region proposal 4.Softmax and regressor for classification and predicting bounding box coordinates	1.mean Average Precision obtained on different networks: a)VGG_CNN_M_1024 – 0.8200 b)ZF – 0.8203 c)VGG16 – 0.8312 2. Not very good identification for surprise and neutral	1.Normalization technique can be used for preprocessing	Recognition rate for surprise and neutral is less

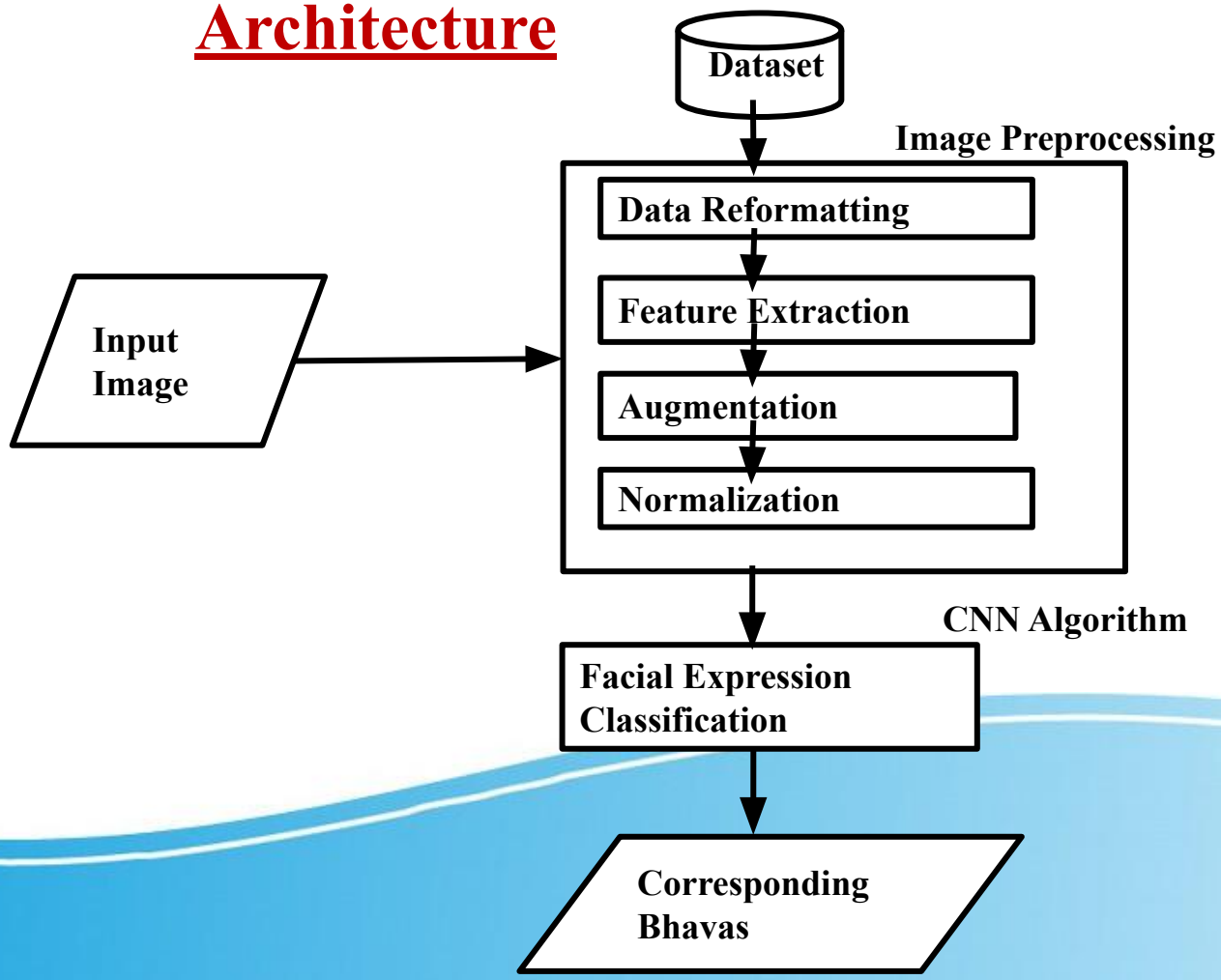
5.	1. To Classify facial expressions into 6 emotions. 2.Classifications with and without pre-processing of images. 3. Angry, sad, happy, disgust, surprise, fear and neutral.	FER-2013	1.Image pre-processing 2.Feature extraction and Classification using CNN. 6 Convolutional layers used. 3.Training set 4.Accuracy	1.Obtained a test accuracy of 61.7% on FER2013 in a seven-classes classification task.	1.CNN algorithm can be used in our project	Misclassified images came from the emotions of fear and sad with 43.95% and 49.77% accuracy

6.	1. Classify facial images into 6 emotions using CNN architecture. 2. using dataset FER2013 for training.	FER-2013	1.Face detection and pre-processing 2.Feature extraction using trainable convolution kernels. 3.Max pooling by pooling layer 4.Classification by softmax layer into the 6 basic classification	1. An accuracy of 65.1 % has been achieved by the test result by CNN. 2. CNN architecture and CNN parameter training explained.	1.CNN architecture can be used in our project	Distortion of image is high

SRS

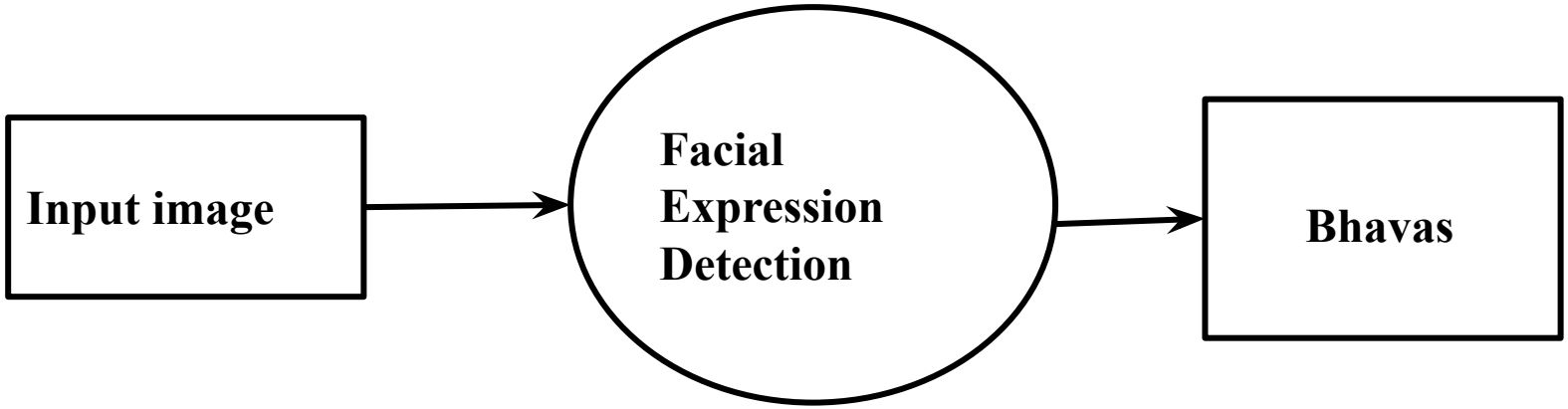
Functional Requirement	Non Functional Requirement
The system should classify an image into one of 4 bhavas and remaining bhavas in 'others' class	The system should be implemented in Google Colaboratory
The system should include an face expression detection algorithm	

Architecture

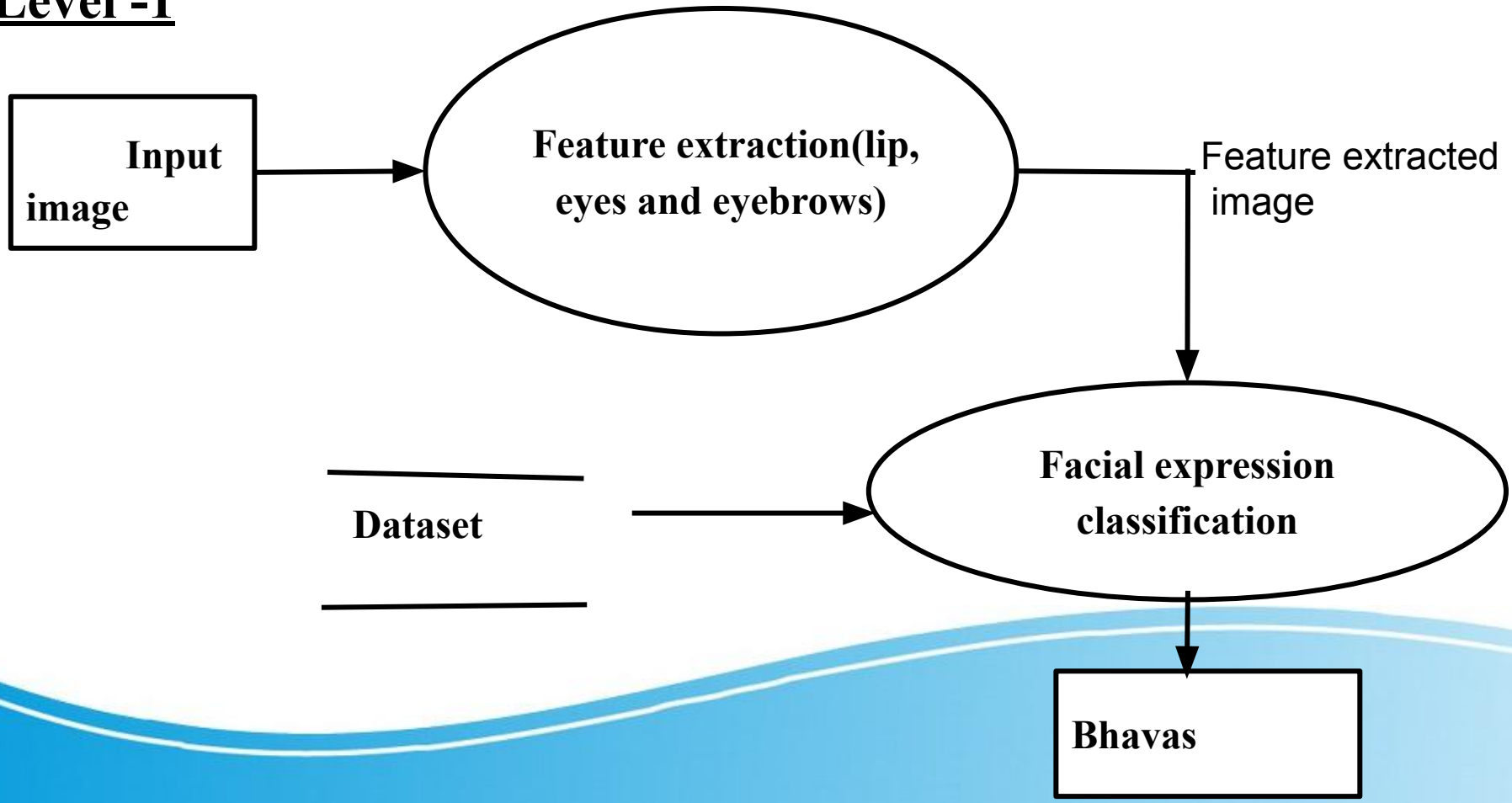


DFD

Level -0



Level -1


Input image
Feature extraction(lip, eyes and eyebrows)
Feature extracted image
Dataset
Facial expression classification
Bhavas

Proposed System

1. Dataset Creation -
   a. A dataset created with 5 subfolders that includes 4 bhavas - 'Bhayanakam', 'Bheebatsam', 'Karunam' & 'Veeram'
   b. Also another folder named 'Others' that contains the images of the remaining 5 bhavas
   c. There are a total of 4664 augmented images in the dataset:
      i. Bhayanakam - 720
      ii. Bheebatsam - 620
      iii. Karunam - 580
      iv. Veeram - 778
      v. Others - 1966

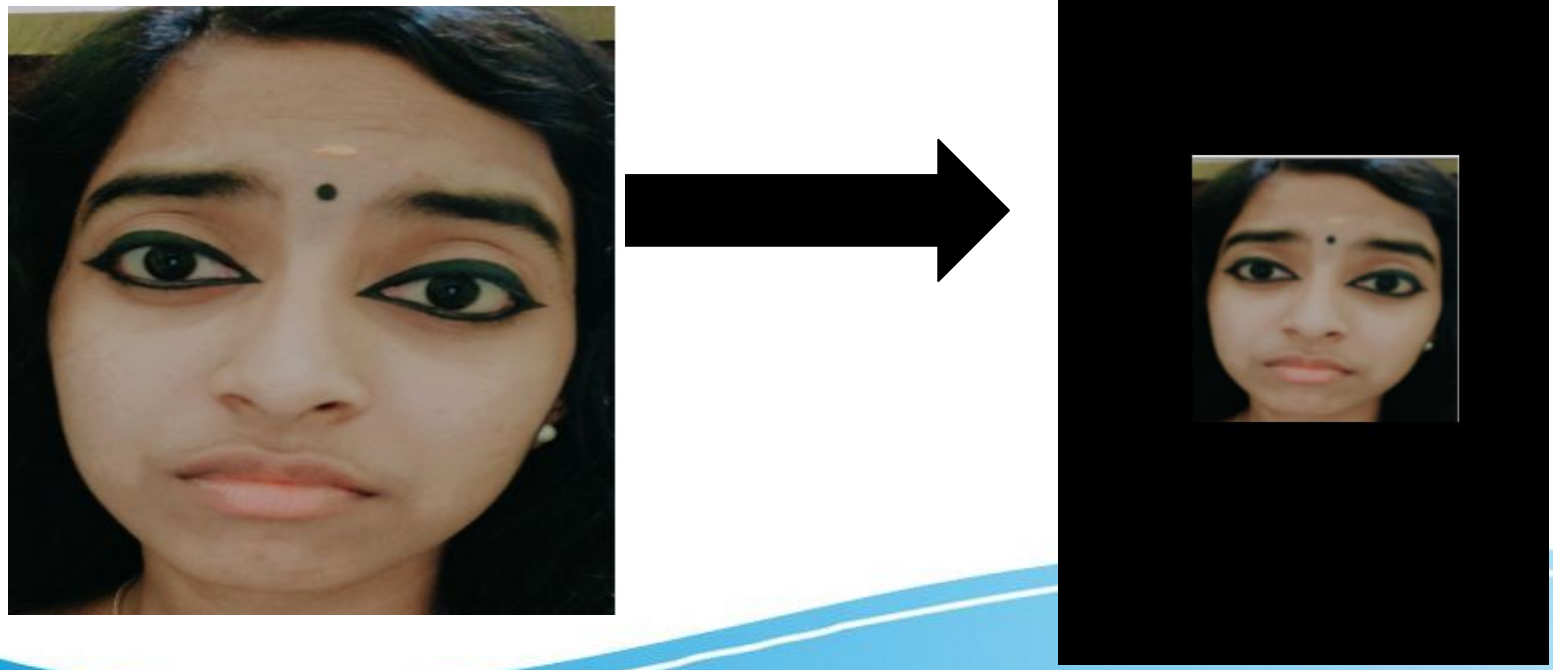
2. Image preprocessing-
   a. Transforms the raw sourced data into a format that enables successful model training
   b. Steps involved in preprocessing modules are -
      i. Image reformatting
      ii. Feature Extraction
      iii. Augmentation
      iv. Normalization
   c. In image reformatting, all images in the dataset are rescaled in to a fixed size of 150x150
   d. The extraction of facial features requires translating the input data into a set of features
   e. Features included in this project are - eyes, lips and eyebrows

f. The training masks are generated by detecting and localizing facial landmarks on a face image
g. In augmentation, images are flipped, rotated, sheared, zoomed which helps to increase number of images
h. Normalization involves normalizing data per pixel

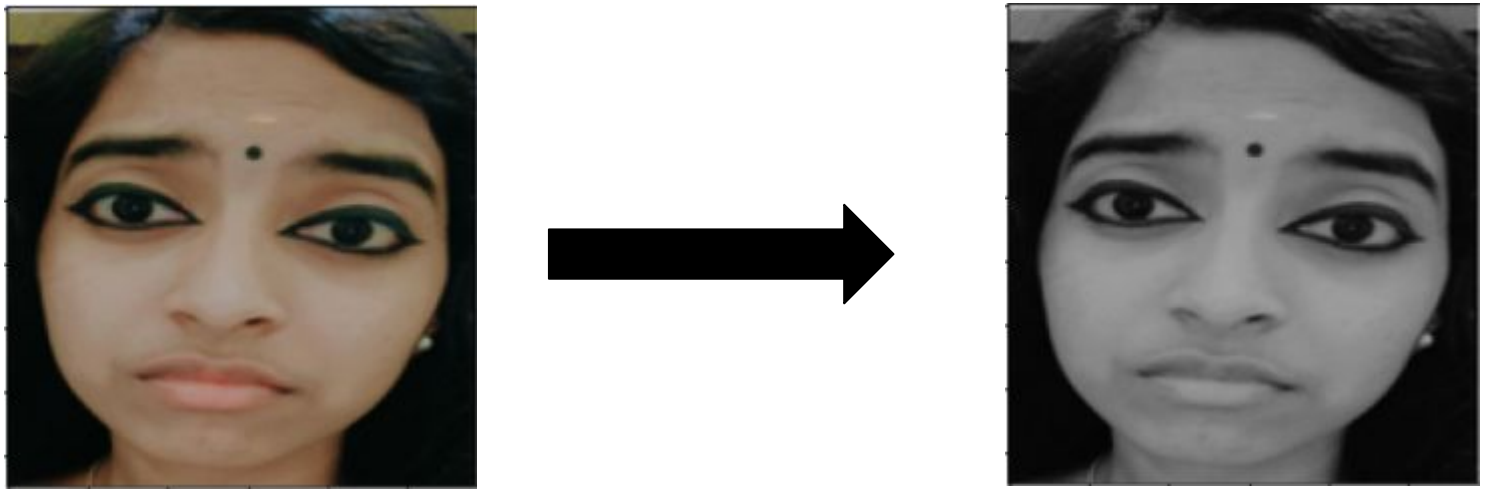
3. Classification
   a. Pre trained model EfficientNetB5 is used
   b. 5 dense layers are used
   c. Softmax layer is used for classification
   d. Classify image to bhavas

Implementation

1 .Resizing images



2.Converting RGB image to grayscale



3.Landmark detection

4.Segmentation

5.Augmentation



6. Model creation

a. Pre-trained ‘EfficientNetB5’ model is used
b. On top of it, dense layers with ‘Relu’ activation functions are used
c. Total of 582 layers are used
d. 576 layers of EfficientNetB5
e. 5 dense layers
f. 1 softmax layer

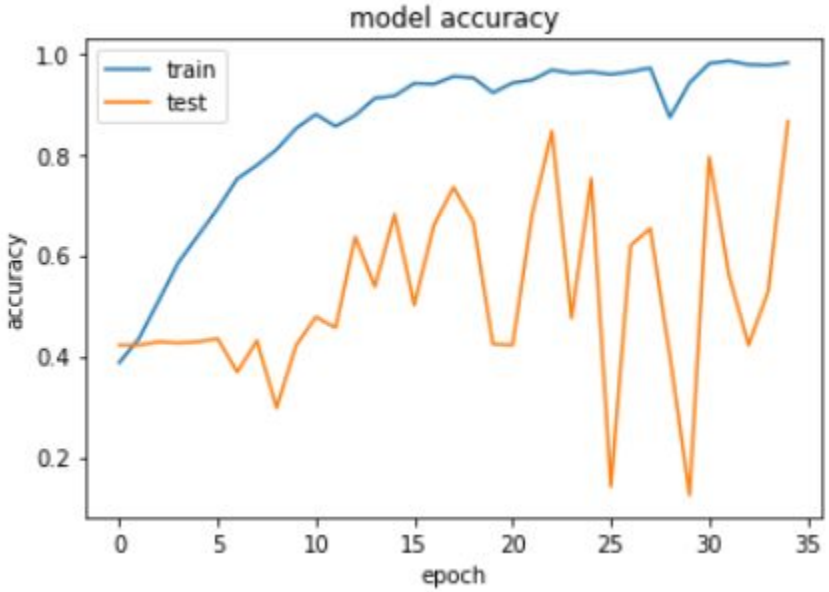
7. Training and Testing

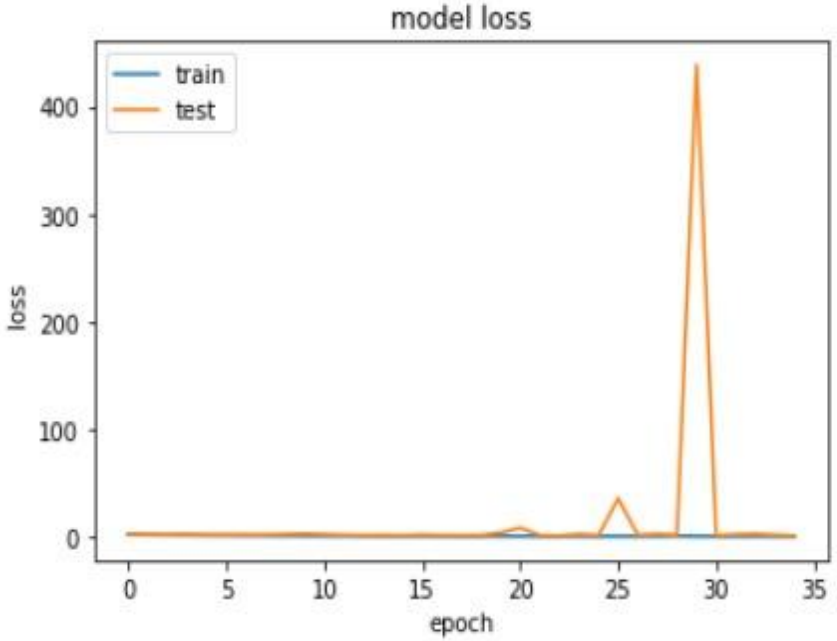
```
132/132 [==============================] - 73s 555ms/step - loss: 0.5308 - accuracy: 0.8378 - val_loss: 1.4865 - val_accuracy: 0.4004
Epoch 30/35
132/132 [==============================] - 73s 554ms/step - loss: 0.2386 - accuracy: 0.9253 - val_loss: 438.6175 - val_accuracy: 0.1242
Epoch 31/35
132/132 [==============================] - 73s 554ms/step - loss: 0.0858 - accuracy: 0.9751 - val_loss: 0.7479 - val_accuracy: 0.7944
Epoch 32/35
132/132 [==============================] - 74s 561ms/step - loss: 0.0486 - accuracy: 0.9851 - val_loss: 1.8499 - val_accuracy: 0.5610
Epoch 33/35
132/132 [==============================] - 73s 553ms/step - loss: 0.0857 - accuracy: 0.9778 - val_loss: 2.2539 - val_accuracy: 0.4218
Epoch 34/35
132/132 [==============================] - 73s 554ms/step - loss: 0.0604 - accuracy: 0.9790 - val_loss: 1.3795 - val_accuracy: 0.5289
Epoch 35/35
132/132 [==============================] - 73s 554ms/step - loss: 0.0520 - accuracy: 0.9849 - val_loss: 0.4799 - val_accuracy: 0.8651
```

```
model.evaluate(X_test,Y_test)
```

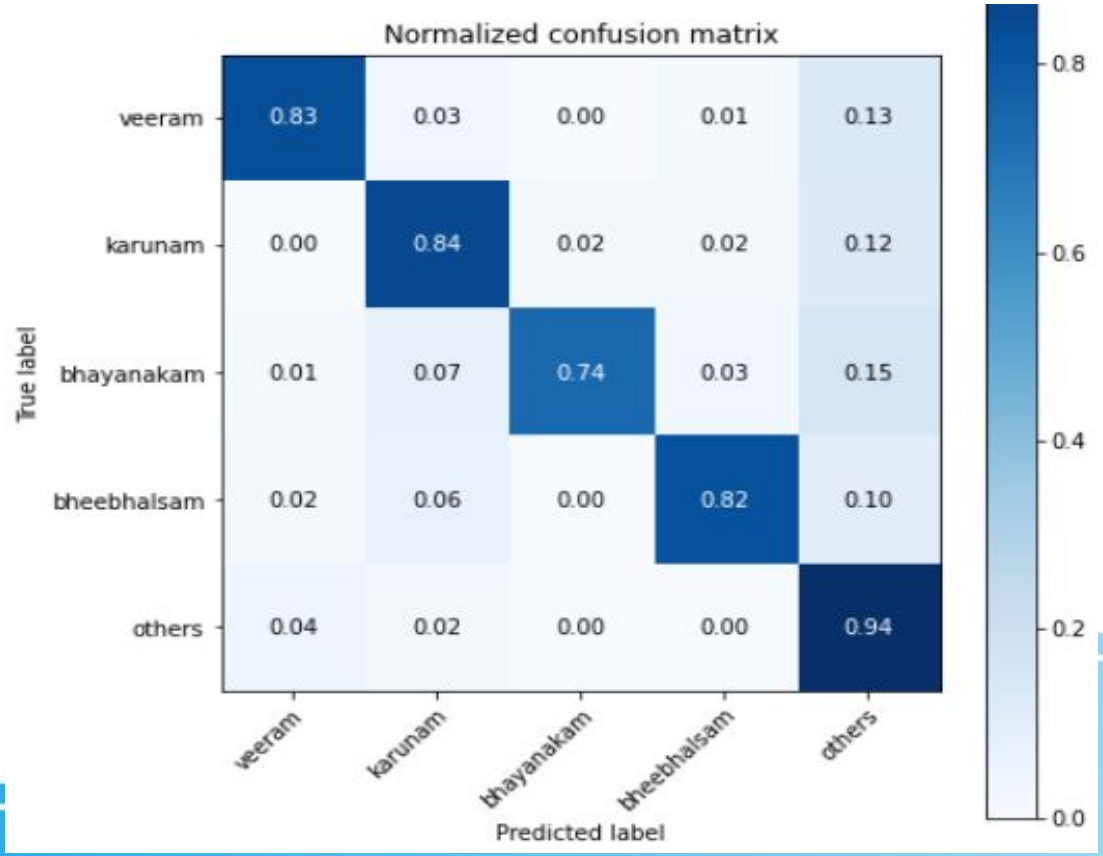
```
15/15 [==============================] - 46s 3s/step - loss: 0.4799 - accuracy: 0.8651
[0.47993364930152893, 0.8650963306427002]
```

- Accuracy and Loss





8. Confusion Matrix



9. Model Prediction

```
pred=model.predict(img)
```

```
print(d[np.argmax(pred[0])])
```

```
karunam
```

Conclusion

- This project helps to identify each bhavas with the help of modern technology
- Classifies additional human facial expressions compared to existing systems
- Preliminary tests achieved 86.5% bhavas recognition accuracy

Future Scope

- To predict the remaining 5 bhavas - ‘Sringara’, ‘Hasya’, ‘Raudra’, ‘Adbhuta’ and ‘Shanta’
- To predict bhavas of images from classical dance related videos

Publication

Paper Title:Review of Facial Expression Detection using Deep Learning.

Link:https://www.ijert.org/research/review-of-facial-expression-detection-using-deep-learning-IJERTV10IS040305.pdf

Reference

1. S. Harshitha, N. Sangeetha, A. P. Shirly and C. D. Abraham, "Human facial expression recognition using deep learning technique," 2019 2nd International Conference on Signal Processing and Communication (ICSPC), Coimbatore, India, 2019, pp. 339-342, doi: 10.1109/ICSPC46172.2019.8976876.

2. Vamshi. N, R. and Raja S, B., “Facial Expression Recognition using Deep Learning”, arXiv e-prints, 2020.

3. Y. Lv, Z. Feng and C. Xu, "Facial expression recognition via deep learning," 2014 International Conference on Smart Computing, Hong Kong, 2014, pp. 303-308, doi: 10.1109/SMARTCOMP.2014.7043872

4. Jiaxing Li, Dexiang Zhang, Jingjing Zhang, Jun Zhang, Teng Li, Yi Xia, Qing Yan, Lina Xun,Facial Expression Recognition with Faster R-CNN,Procedia Computer Science,https://doi.org/10.1016/j.procs.2017.03.069.
5. .S. Singh and F. Nasoz, "Facial Expression Recognition with Convolutional Neural Networks," 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2020, pp. 0324-0328, doi: 10.1109/CCWC47524.2020.9031283.
6. .L. Xu, M. Fei, W. Zhou and A. Yang, "Face Expression Recognition Based on Convolutional Neural Network*," 2018 Australian & New Zealand Control Conference (ANZCC), Melbourne, VIC, 2018, pp. 115-118, doi: 10.1109/ANZCC.2018.8606597.

Thank You