Facial Expression Recognition Based Rating System

[1]T . Prathyusha, [2]Dr . K . Suresh Babu

[1]Student, [2]Associate Professor

[1]M.TECH. Software Engineering

[1]School Of Information Technology, JNTU Hyderabad

***Abstract* :** Nowadays, most of the restaurants are being automated and the presence of staff is minimal. Due to the less staff, the management was not able to collect the experiences from all the customers who visit the restaurant. This paper presents a rating system based on FER with pre-trained convolutional neural network (CNN) models. For interactive human and computer interface (HCI) the computer must understand the facial expressions of humans. HCI reduces the gap between computers and humans. Computers can interact more appropriately with humans by judging their expressions. There are various techniques for facial expression recognition which focus on getting good results of human expressions and then the food is supposed to be rated. Compared to the text-based rating systems, the information contained in it would be less. The rating is simple and playful. In addition to this, the management can estimate the age group experiences on the restaurant's concepts. Here we explain three expressions (Happy, neutral, and sad) in the rating system.

***Keywords*-** Deep Learning, Machine Learning, Facial expression recognition, HCI,CNN, automated rating system, automated restaurants.

I. INTRODUCTION

Collection of proper feedback from all the customers who visit the restaurant is a difficult task in unmanned restaurants. It is even more difficult for the restaurant management to estimate how the concept and the food are experienced by the customers. Few rating systems, such as Google, Zomato, Swiggy and TripSavvy, only partially solve this problem, as they only cover a part of the customer's opinions. The customers who feel extremely happy and the customers who are completely unsatisfied usually take initiative in rating the restaurants on independent rating platforms.

The above problem can be solved by motivating all the customers to give proper feedback on the restaurant concepts. The rating system based on facial expression detection can be something interesting and interactive for the customers compared to the text-based rating system. This rating system uses FER detection using pre-trained convolutional neural network(CNN) models. The live detection and capturing of a customer's facial expression results in the rating of the food. The restaurant management can collect a wider range of opinions about the experiences of the customers with the restaurant concept.

II. BLOCK DIAGRAM

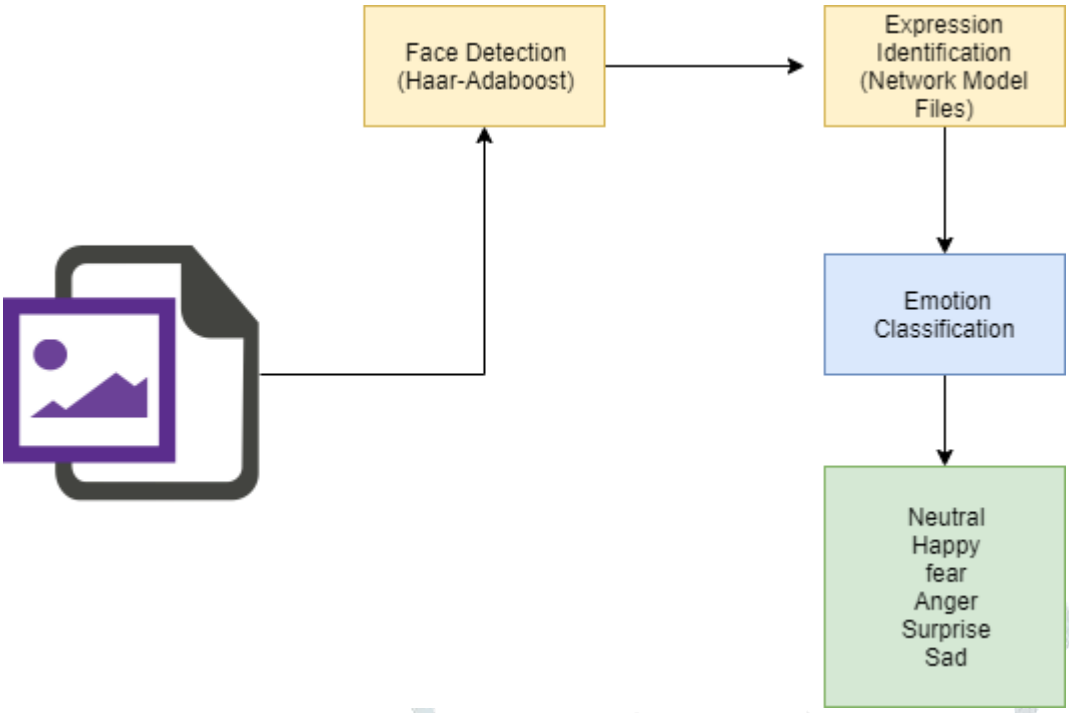

Fig.1. System Architecture.

The above diagram shows the architecture view of the designed system. It shows the flow among various elements throughout the system in an abstract view.

III. PROCEDURE

3.1Collecting Dataset

The two training models used were haarcascade_frontalface_default.xml file and fer2013_mini_XCEPTION.hdf5 file. The OpenCV is a huge open-source library for machine learning, image processing and computer vision. Haar-Cascade are classifiers that are used to detect features by superimposing predefined patterns over face segments and are used as xml files. OpenCV provides a training method or pretrained models that can be read using cv::CascadeClassifier::load method. Fer2013 dataset contains 48x48 pixel grayscale pictures of faces. It categorize facial expressions into seven categories. This set consists of 28,709 examples.

3.2 Face Detection

Detection of the face is an important step for image classification since only the principal component of the face such as nose, eyes, mouth are needed for classification. Face detection algorithms are classified into feature, knowledge, template, and appearance-based methods. In our detection models, the system uses the viola jones object detection algorithm for face localization which comes under feature-based classification. Viola-Jones object detection algorithm uses Haar feature-based cascade classifiers. The Haar cascade classifier is an extremely important element of face detection. The presence of the features in any of the input images is determined by the Haar features.

Sample Code:

```
image_path=”../FER/images/capture.jpg”
detection_model_path=’../FER/trained_models/detection_models/haarcascade_frontalface_default.xm’
face_detection=load_detection_model(detection_model_path)
rgb_image=cv2.imread(image_path)
gray_image=cv2.cvtColor(rgb_image,cv2.COLOR_BGR2GRAY)
gray_image=np.squeeze(gray_image)
```

```
gray_image=gray_image.astype('uint8')
faces=detect_faces(face_detection,gray_image)
 if faces==():
print("[WARNING] Recapture the image")
return
```

3.3Facial Expression Recognition Classification

The deep learning features were trained, the final step of FER(Facial Expression Recognition) is to classify the given face into one of the basic emotion categories. In traditional methods, the feature extraction step and the feature classification step are independent, deep networks can perform FER in an end-to-end way. The loss layer is added at the end of the network to regulate the back-propagation error, then, the prediction probability of each sample can be directly generated as output by the network. In a convolutional neural network, the softmax loss is the most commonly used function that minimizes the cross-entropy between the estimated class probabilities and the ground truth distribution.

Sample Code:

```
emotion_model_path='../FER/trained_models/emotion_models/fer2013_mini_XCEPTION.hdf5'
For face_coordinates in faces:
x1,x2,y1,y2=apply_offsets(face_coordinates,emotion_offsets)
gray_face=gray_image[y1:y2,x1:x2]
Try:
gray_face=cv2.resize(gray_face,(emotion_target_size))
 except:
Continue
gray_face=preprocess_input(gray_face,True)
gray_face=np.expand_dims(gray_face,0)
gray_face=np.expand_dims(gray_face,-1)
emotion_label_arg=np.argmax(emotion_classifier.predict(gray_face))
emotion_text=emotion_labels[emotion_label_arg]
print(emotion_text)
```

3.4Training And Generation

The convolutional neural network has been extensively used in diverse computer vision applications, including FER. The captured image is given as an input to the input layer. In the convolution layer, we build the core block of the input data, this layer has a set of kernels means learnable filters using the small blocks of the data, and learn features of the data. The activation layer will apply the element-wise RELU activation function to the output of the convolution layer. The pooling layer extracts the features from the single size of the data splitting to multilayer. The computation cost is reduced.

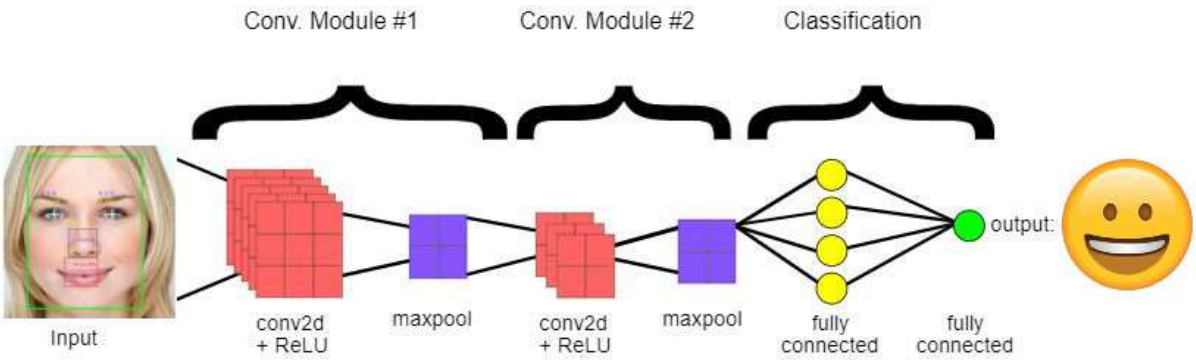

Fig. 2. Convolutional Neural Network.

IV. RESULTS

4.1 Main Home Page

Fig. 3. Home page.

This is the main home page where a user selects facial expression recognition and submit the feedback.

4.2 Select the food item

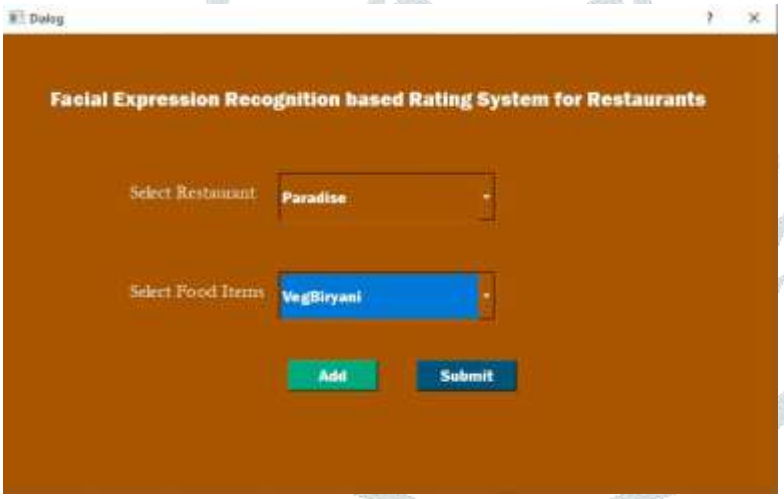

Fig. 4. Select the food item.

The user can select the restaurants and the food items. Two or more food items from the same restaurant can also be selected at once. To save the food items click on the submit button.

4.3 Camera Detection

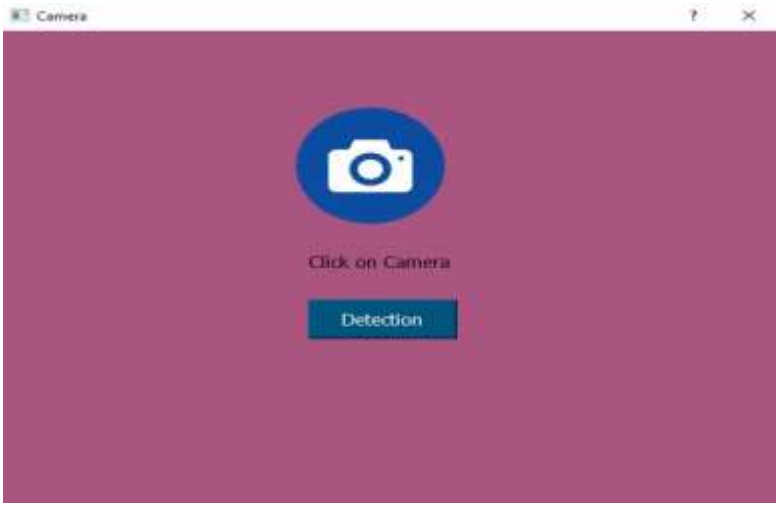

Fig. 5. Camera Detection.

The user can click on the camera button and capture the facial expression. The detection button is used to classify the captured image and helps in submitting the feedback for those food items. If the expression of the user doesn't match with classified models in the system, user need to retake the picture.

4.4 Submit the rating

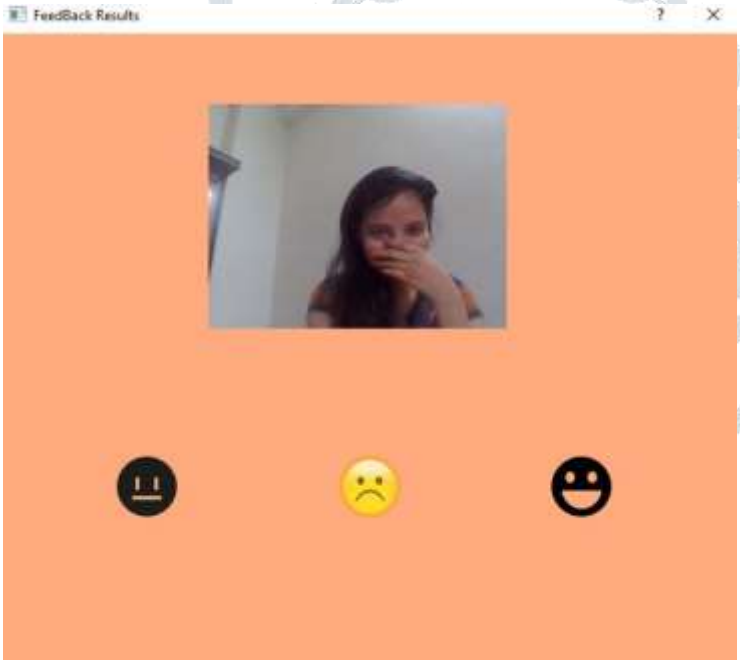

Fig. 6. Submit the rating.

The user rating is submitted to the restaurant admin and the rating is also displayed on the screen for the user to verify the feedback. If the user wants to change the given feedback they can retake the picture and submit the feedback.

V. CONCLUSION

In this paper, the human facial expressions are captured from a live webcam in rating the restaurant's food items. Three expressions (Happy, Sad, and Neutral) were captured as of now in this application. This application can also be used in office meetings where we can know all the expressions of the members who attended the session and among which how many members were happy, sad, and neutral.

VI. REFERENCES

[1] https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html.

[2]https://www.restaurantindia.in/article/why-customer-feedback-is-important-for-your-restaurant-s-growth.13558

[3] https://www.cs.auckland.ac.nz/~rklette/CCV-CIMAT/pdfs/B27-Haar_VJ_AB.pdf

[4]https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53?gi=e91115e517d

[5]https://towardsdatascience.com/the-intuition-behind-facial-detection-the-viola-jones-algorithm-29d9106b6999

[6] https://numpy.org/doc/stable/user/absolute_beginners.html

[7]https://analyticsindiamag.com/gradient-descent-everything-you-need-to-know-with-implementation-in-python/

[8] https://pypi.org/project/visualizer/

[9] https://www.w3schools.com/python/python_mysql_getstarted.asap