A Deep Learning Facial Expression Recognition Based Scoring System For Restaurants

P.SWAPNA [1], P ANUPAMA [2]

[1] PG Scholars, Department of CSE, Sree Dattha Institute Of Engineering and Science, Sheriguda,Hyderabad.
[2] Assistant Professor, Department of CSE, Sree Dattha Institute Of Engineering and Science, Sheriguda,Hyderabad.

ABSTRACT

Recently, the popularity of automated and unmanned restaurants has increased. Due to the absence of staff, there is no direct perception of the customers' impressions in order to find out what their experiences with the restaurant concept are like. For this purpose, this paper presents a rating system based on facial expression recognition with pre-trained convolutional neural network (CNN) models. It is composed of an Android mobile application, a web server, and a pre-trained AIserver. Both the food and the environment are supposed to be rated. Currently, three expressions (satisfied, neutral and disappointed) are provided by the scoring system.

INTRODUCTION

F ACIAL expression is one of the most powerful, natural and universal signals for human beings to convey their emotional states and intentions [1], [2]. Numerous studies have been conducted on automatic facial expression analysis because of its practical importance in sociable robotics, medical treatment, driver fatigue surveillance, and many other human-computer interaction systems. In the field of computer vision and machine learning, various facial expression recognition (FER) systems have been explored to encode expression information from facial representations. As early as the twentieth century, Ekman and Friesen [3] defined six basic emotions based on cross-culture study [4], which indicated that humans perceive certain basic emotions in the same way regardless of culture. These prototypical facial expressions are anger, disgust, fear, happiness, sadness, and surprise. Contempt was subsequently added as one of the basic emotions [5]. Recently, advanced research on neuroscience and psychology argued that the model of six basic emotions are culture-specific and not universal [6].

Although the affect model based on basic emotions is limited in the ability to represent the complexity and subtlety of our daily affective displays [7], [8], [9], and other emotion

description models, such as the Facial Action Coding System (FACS) [10] and the continuous model using affect dimensions [11], are considered to represent a wider range of emotions, the categorical model that describes emotions in terms of discrete basic emotions is still the most popular perspective for FER, due to its pioneering investigations along with the direct and intuitive definition of facial expressions. And in this survey, we will limit our discussion on FER based on the categorical model.

FER systems can be divided into two main categories according to the feature representations: static image FER and dynamic sequence FER. In static-based methods [12], [13], [14], the feature representation is encoded with only spatial information from the current single image, whereas dynamic-based methods [15], [16], [17] consider the temporal relation among contiguous frames in the input facial expression sequence. Based on these two visionbased methods, other modalities, such as audio and physiological channels, have also been used in multimodal systems [18] to assist the recognition of expression.

The majority of the traditional methods have used handcrafted features or shallow learning (e.g., local binary patterns (LBP) [12], LBP on three orthogonal planes (LBP-TOP) [15], non-negative matrix factorization (NMF) [19] and sparse

learning [20]) for FER. However, since 2013, emotion recognition competitions such as FER2013 [21] and Emotion Recognition in the Wild (EmotiW) [22], [23], [24] have collected relatively sufficient training data from challenging real-world scenarios, which implicitly promote the transition of FER from lab-controlled to in-the-wild settings. In the meanwhile, due to the dramatically increased chip processing abilities (e.g., GPU units) and well-designed network architecture, studies in various fields have begun to transfer to deep learning methods, which have achieved the state-of-the-art recognition accuracy and exceeded previous results by a large margin (e.g., [25], [26], [27], [28]). Likewise, given with more effective training data of facial expression, deep learning techniques have increasingly been implemented to handle the challenging factors for emotion recognition in the wild. Figure 1 illustrates this evolution on FER in the aspect of algorithms and datasets.

Exhaustive surveys on automatic expression analysis have been published in recent years [7], [8], [29], [30]. These surveys have established a set of standard algorithmic pipelines for FER. However, they focus on traditional methods, and deep learning has rarely been reviewed. Very recently, FER based on deep learning has been surveyed in [31], which is a brief review without introductions on FER datasets and technical details on deep FER. Therefore, in this paper, we

make a systematic research on deep learning for FER tasks based on both static images and videos (image sequences). We aim to give a newcomer to this filed an overview of the systematic framework and prime skills for deep

Despite the powerful feature learning ability of deep learning, problems remain when applied to FER. First, deep neural networks require a large amount of training data to avoid overfitting. However, the existing facial expression databases are not sufficient to train the well-known neural network with deep architecture that achieved the most promising results in object recognition tasks. Additionally, high inter-subject variations exist due to different personal attributes, such as age, gender, ethnic backgrounds and level of expressiveness [32]. In addition to subject identity bias, variations in pose, illumination and occlusions are common in unconstrained facial expression scenarios. These factors are nonlinearly coupled with facial expressions and therefore strengthen the requirement of deep networks to address the large intra-class variability and to learn effective expression-specific representations.

EXISTING SYSTEM

As there is no staff available in unmanned restaurants, it is difficult for the restaurant management to estimate how the concept and the food is experienced by the customers. Existing rating systems, such as Google and TripAdvisor, only partially solve this problem, as they only cover a part of the customer's opinions. These rating systems are only used by a subset of the customers who rate the restaurant on independent rating platforms on their own initiative. This applies mainly to customers who experience their visit as very positive or negative.

PROPOSED SYSTEM

In order to solve the above problem, all customers must be motivated to give a rating. This paper introduces an approach for a restaurant rating system that asks every customer for a rating after their visit to increase the number of ratings as much as possible. This system can be used unmanned restaurants; the scoring system is based on facial expression detection using pretrained convolutional neural network (CNN) models. It allows the customer to rate the food by taking or capturing a picture of his face that reflects the corresponding feelings. Compared to text-based rating system, there is much less information and no individual experience reports collected. However, this simple fast and playful rating system should give a wider range of opinions about the experiences of the customers with the restaurant concept.

IMPLEMENTATION

MODULES:

Face Detection:

Face detection or localization is an important step for image classification since only the principal component of face such as nose, eyes, mouth are needed for classification. Face detection algorithms can be broadly classified into feature, knowledge, template and appearance

based methods. Our proposed system uses Viola Jones object detection algorithm for face localization which comes under feature based classification. Viola Jones object detection algorithm uses Haar featurebased cascade classifiers. The Haar Cascade classifier is extremely important element of the face detection. The presence of the features in any of the input image is

determined by the Haar features.

Facial Expression Recognition classification:

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

Convolutional neural network (CNN):

CNN has been extensively used in diverse computer vision applications, including FER. At the beginning of the 21st century, several studies in the FER literature found that the CNN is robust to face location changes and scale variations and behaves better than the multilayer perceptron (MLP) in the case of previously unseen face pose variations, employed the CNN to address the problems of subject independence as well as translation, rotation, and scale invariance in the recognition of facial expressions.
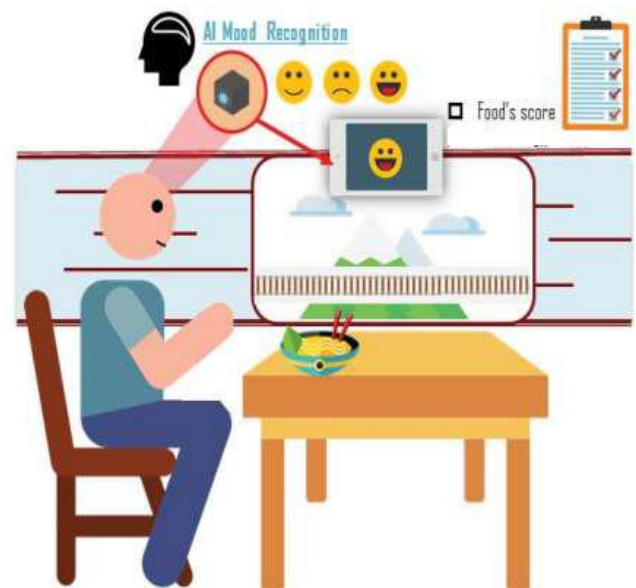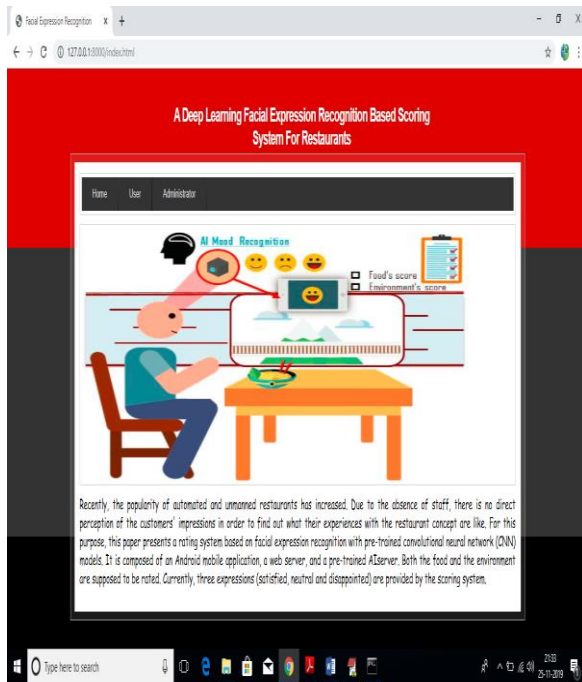
System Architecture:-



Fig. System Architecture

Results

To run this project install MYSQL and then create database by copying content from 'DB.txt' file and paste in MYSQL. Install python and then install DJANGO web server and deploy code on DJANGO. After deployment start server and run the code from browser.

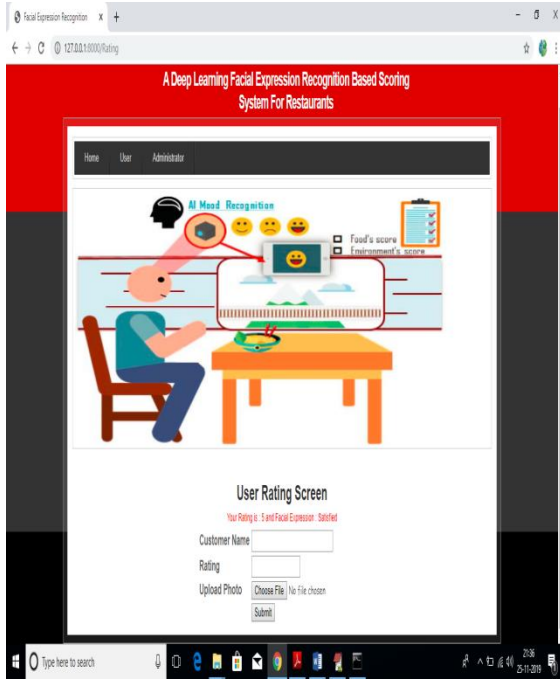In above screen click on 'User' link to get below screen where user can upload photo and give ratings
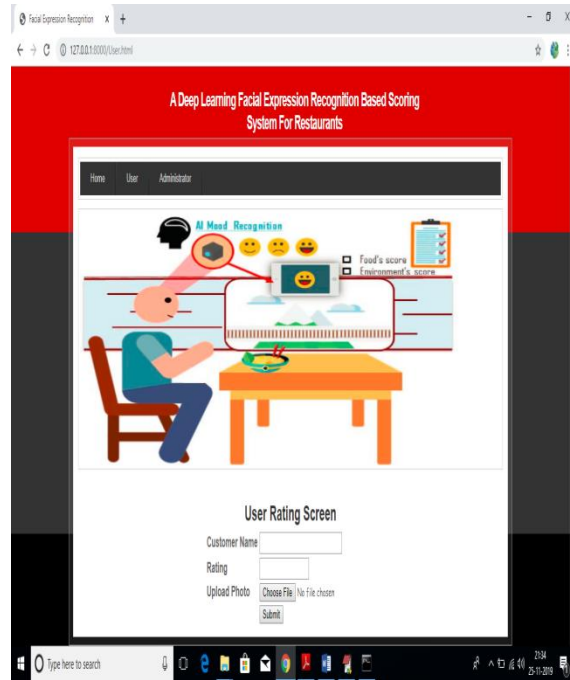
User will fill above from and upload photo
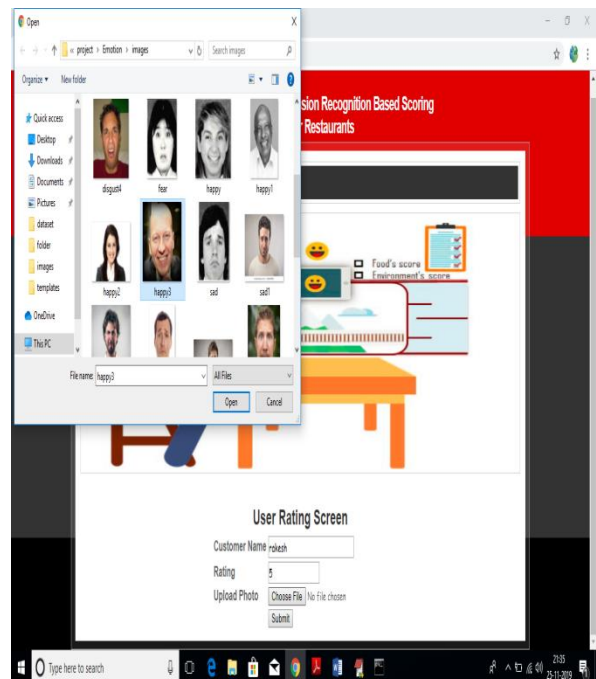
In above screen I filled form and uploading one happy image and then click on 'Open' button and then click 'Submit' button to send data to
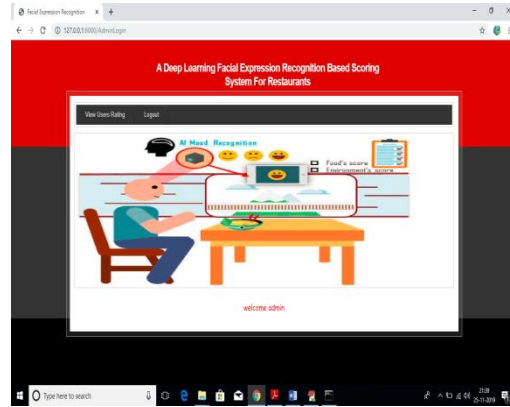
webserver. After processing above data will get below results.

In above screen we can see output message as given rating and from photo extracted facial expression is satisfied. Now go to 'Administrator' link and login as admin by giving username as 'admin' and password as 'admin'. See below screen.
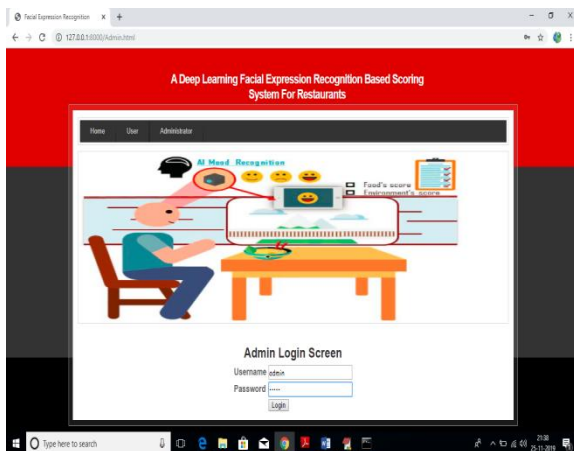
After login will get below screen

In above screen click admin can click on 'View Users Rating' link to get all customers feedback. See below screen
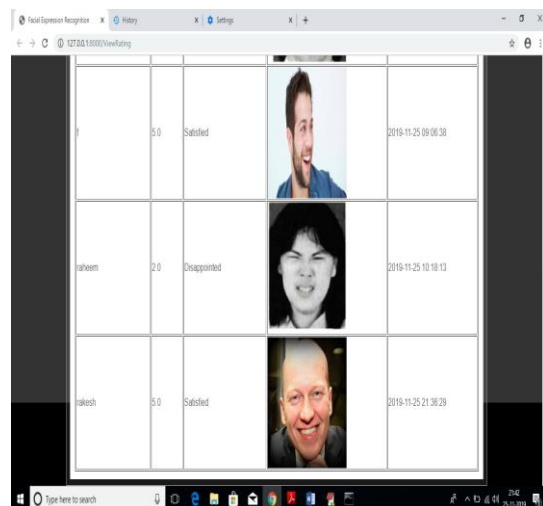
From above screens admin can see photos and their facial expressio

CONCLUSION

Data bias and inconsistent annotations are very common among different facial expression datasets due to different collecting conditions and the subjectiveness of annotating. Researchers commonly evaluate their algorithms within a specific dataset and can achieve satisfactory performance. However, early cross- database experiments have indicated that discrepancies between databases exist due to the different collection environments and construction indicators [12]; hence, algorithms evaluated via intra-database protocols lack generalizability on unseen test data, and the performance in cross-dataset settings is greatly deteriorated. Deep domain adaption and knowledge distillation are alternatives to address this bias [226], [251]. Furthermore, because of the inconsistent expression annotations, FER performance cannot keep improving when enlarging the training data by directly merging multiple datasets [167]. Another common problem in facial expression is class imbalance, which is a result of the practicalities of data acquisition: eliciting and annotating a smile is easy.

REFERENCES

- [1] C. Darwin and P. Prodger, The expression of the emotions in man and animals. Oxford University Press, USA, 1998.
- [2] Y.-I. Tian, T. Kanade, and J. F. Cohn, "Recognizing action units for facial expression analysis," IEEE Transactions on pattern analysis and machine intelligence, vol. 23, no. 2, pp. 97–115, 2001.
- [3] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion." Journal of personality and social psychology, vol. 17, no. 2, pp. 124–129, 1971.
- [4] P. Ekman, "Strong evidence for universals in facial expressions: a reply to russell's mistaken critique," Psychological bulletin, vol. 115, no. 2, pp. 268–287, 1994. [5] D. Matsumoto, "More evidence for the universality of acontempt expression," Motivation and Emotion, vol. 16, no. 4, pp. 363–368, 1992. [6] R. E. Jack, O. G.Garrod, H. Yu, R. Caldara, and P. G. Schyns, "Facial expressions of emotion are not culturally universal,"
- Proceedings of the National Academy of Sciences, vol. 109, no. 19, pp. 7241–7244, 2012.