International Journal of Computing and Artificial Intelligence

E-ISSN: 2707-658X
P-ISSN: 2707-6571
IJCAI 2020; 1(2): 05-10
Received: 04-05-2020
Accepted: 07-06-2020

Bana Leelavathi
Dept. of Computer Science, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

Corresponding Author:
Bana Leelavathi
Dept. of Computer Science, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

Facial expression recognition based scoring system for restaurants by using deep learning concepts

Bana Leelavathi

DOI: https://doi.org/10.33545/27076571.2020.v1.i2a.10

Abstract
Now-a-days in advance countries automated unmanned restaurants are more popular as this restaurants will not have any human power to take customer feedbacks about food quality and service and to automate this process author has introduce concept called 'Deep Learning Facial Expression Recognition Based Scoring System For Restaurants' where customers will be asked to give rating to food and upload his photo and based on user facial expression application will inform whether customer was satisfied or not. To extract facial expressions from photo we are using CNN (Convolution Neural Networks) machine learning algorithm. This algorithm can predict 3 different expression from photo such as satisfied, neutral or disappointed.

Keywords: facial expression recognition, scoring system, deep learning concepts

1. Introduction
Facial expression is one of the most remarkable, normal and widespread signs for people to pass on their passionate states and expectations [1], [2]. Various examinations have been directed on programmed outward appearance investigation in view of its down to earth significance in agreeable mechanical technology, clinical treatment, driver weakness reconnaissance, and numerous other human-PC association frameworks. In the field of PC vision and AI, different outward appearance acknowledgment (FER) frameworks have been investigated to encode demeanor data from facial portrayals. As right on time as the twentieth century, Ekman and Friesen [3] characterized six essential feelings dependent on cross-culture study [4], which demonstrated that people see certain fundamental feelings similarly paying little mind to culture.
These prototypical outward appearances are outrage, appall, dread, satisfaction, pity, and shock. Scorn was hence included as one of the essential feelings [5]. As of late, propelled explore on neuroscience and brain science contended that the model of six fundamental feelings are culture-explicit and not widespread [6]. In spite of the fact that the influence model dependent on fundamental feelings is restricted in the capacity to speak to the intricacy and nuance of our day by day full of feeling shows [7], [8], [9], and other feeling portrayal models, for example, the Facial Action Coding System (FACS) [10] and the ceaseless model utilizing influence measurements [11], are considered to speak to a more extensive scope of feelings, the clear cut model that depicts feelings as far as discrete essential feelings is as yet the most mainstream viewpoint for FER, because of its spearheading examinations alongside the immediate and instinctive meaning of outward appearances. Furthermore, in this overview, we will restrain our conversation on FER dependent on the all out model. FER frameworks can be partitioned into two primary classifications as indicated by the element portrayals: static picture FER and dynamic succession FER. In static-based strategies [12], [13], [14], the component portrayal is encoded with just spatial data from the present single picture, while dynamic-based techniques [15], [16], [17] consider the worldly connection among touching edges in the information outward appearance arrangement. In view of these two vision based strategies, different modalities, for example, sound and physiological channels, have additionally been utilized in multimodal frameworks [18] to help the acknowledgment of articulation. Most of the conventional techniques have utilized high quality highlights or shallow learning (e.g., nearby double examples (LBP) [12], LBP on three symmetrical planes (LBP-TOP) [15], non-negative grid factorization (NMF) [19] and inadequate learning [20]) for FER. Be that as it may, since 2013, feeling acknowledgment rivalries, for example, FER 2013 [21] and

Emotion Recognition in the Wild (EmotiW) [22], [23], [24] have gathered generally adequate preparing information from testing genuine situations, which verifiably advance the change of FER from lab-controlled to in-the-wild settings. In the in the interim, because of the significantly expanded chip preparing capacities (e.g., GPU units) and all around planned system design, concentrates in different fields have started to move to profound learning techniques, which have accomplished the best in class acknowledgment precision and surpassed past outcomes by an enormous edge (e.g., [25], [26], [27], [28]). In like manner, given with progressively viable preparing information of outward appearance, profound learning strategies have progressively been executed to deal with the difficult components for feeling acknowledgment in nature. Figure 1 outlines this advancement on FER in the part of calculations and datasets.

2. Literature Survey

2.1 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.

Ekman and Friesen [14] built up the Facial Action Coding System (FACS) for portraying outward appearances by activity units (AUs). Of 44 FACS AUs that they characterized, 30 AUs are anatomically identified with the withdrawals of explicit facial muscles: 12 are for upper face, and 18 are for lower face. AUs can happen either separately or in blend. At the point when AUs happen in blend they might be added substance, in which the mix doesn't change the presence of the constituent AUs, or no additive, in which the presence of the constituents changes. Despite the fact that the quantity of nuclear activity units is moderately little, in excess of 7,000 diverse AU mixes have been watched [30]. FACS gives the elucidating power important to portray the subtleties of outward appearance. Usually happening AUs and a portion of the added substance and no additive AU blends are appeared in Tables 1 and 2. For instance of a no additive impact, AU 4 shows up contrastingly relying upon whether it happens alone or in mix with AU 1 (as in AU 1 ‡ 4). At the point when AU 4 happens alone, the foreheads are drawn together and brought down. In AU 1 ‡ 4, the foreheads are drawn together however are raised because of the activity of AU 1. AU 1 ‡ 2 is another case of no additive mixes. At the point when AU 2 happens alone, it raises the external forehead, yet in addition regularly pulls up the internal temple which brings about a fundamentally the same as appearance to AU 1 ‡ 2. These impacts of the no additive AU mixes increment the challenges of AU acknowledgment

2.2 P. Ekman and W. V. Friesen, "Constants across societies in the face and feeling." Journal of character and social brain research, vol. 17, no. 2, pp. 124–129, 1971

Explored the subject of whether any outward appearances of feeling are widespread. Ongoing investigations indicating that individuals from educated societies connected a similar feeling ideas with a similar facial practices couldn't exhibit that probably some outward appearances of feeling are all inclusive; the way of life contrasted had all been uncovered with a portion of a similar broad communications introductions of outward appearance, and these may have shown the individuals in each culture to perceive the remarkable outward appearances of different societies. To show that individuals from a preliterate culture who had negligible introduction to proficient societies would connect a similar feeling ideas with indistinguishable facial practices from do individuals from Western and Eastern educated societies, information were accumulated in New Guinea by recounting to 342 Ss a story, demonstrating them a lot of 3 faces, and requesting that they select the face which indicated the feeling proper to the story. Ss were individuals from the Fore phonetic social gathering, which up until 12 yr. prior was a confined, Neolithic, material culture. Results give proof on the side of the theory. (30 ref.) (PsycINFO Database Record (c) 2016 APA, all rights saved)

2.3 R. E. Jack, O. G. Garrod, H. Yu, R. Caldara, and P. G. Schyns, "Outward appearances of feeling are not socially all inclusive," Proceedings of the National Academy of Sciences, vol. 109, no. 19, pp. 7241–7244, 2012.

Since Darwin's original works, the all-inclusiveness of outward appearances of feeling has stayed one of the longest standing discussions in the organic and sociologies. Quickly expressed, the comprehensiveness speculation guarantees that all people impart six fundamental inward enthusiastic states (glad, shock, dread, disturb, outrage, and tragic) utilizing a similar facial developments by uprightness of their organic and transformative starting points [Susskind JM, *et al*. (2008) Nat Neurosci 11:843–850]. Here, we disprove this expected all-inclusiveness. Utilizing a one of a kind PC designs stage that consolidates generative syntaxes [Chomsky N (1965) MIT Press, Cambridge, MA] with visual discernment, we got to the imagination of 30 Western and Eastern culture people and remade their psychological portrayals of the six fundamental outward appearances of feeling. Multifaceted examinations of the psychological portrayals challenge all-inclusiveness on two separate checks. To begin with, while Westerners speak to every one of the six fundamental feelings with an unmistakable arrangement of facial developments normal to the gathering, Easterners don't. Second, Easterners speak to passionate power with particular unique eye movement. By discrediting the long-standing all-inclusiveness speculation, our information feature the ground-breaking impact of culture on molding essential practices once considered naturally designed. Therefore, our information open an extraordinary nature–sustain banter across expansive fields from transformative brain science and social neuroscience to long range interpersonal communication by means of advanced symbols.

3. Proposed work

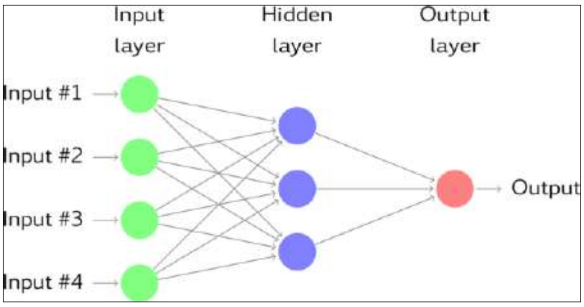
In order to solve the above problem, all customers must be motivated to give a rating. This paper introduces an approach for a restaurant rating system that asks every customer for a rating after their visit to increase the number of ratings as much as possible. This system can be used unmanned restaurants; the scoring system is based on facial expression detection using pre- trained 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

Advantages

- Security
- Improving memory utilization through parallelization

3.1 CNN working procedure

- To demonstrate how to build a convolutional neural network based image classifier, we shall build a 6 layer neural network that will identify and separate one image from other. This network that we shall build is a very small network that we can run on a CPU as well. Traditional neural networks that are very good at doing image classification have many more parameters and take a lot of time if trained on normal CPU. However, our objective is to show how to build a real-world convolutional neural network using TENSORFLOW.
- Neural Networks are essentially mathematical models to solve an optimization problem. They are made of neurons, the basic computation unit of neural networks. A neuron takes an input (say x), do some computation on it (say: multiply it with a variable w and adds another variable b) to produce a value (say; z= wx + b). This value is passed to a non-linear function called activation function (f) to produce the final output (activation) of a neuron. There are many kinds of activation functions. One of the popular activation function is Sigmoid. The neuron which uses sigmoid function as an activation function will be called sigmoid neuron. Depending on the activation functions, neurons are named and there are many kinds of them like RELU, TanH.
- If you stack neurons in a single line, it's called a layer; which is the next building block of neural networks. See below image with layers



- To predict image class multiple layers operate on each other to get best match layer and this process continues till no more improvement left.

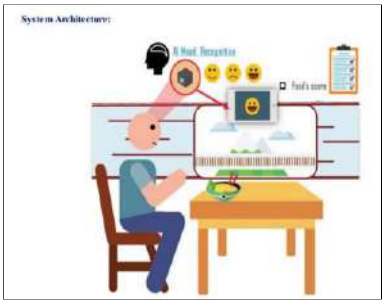


Fig 3.1: System Architecture

4. Results and Discussions

Fig 4.1: In above screen click on 'User' link to get below screen where user can upload photo and give ratings

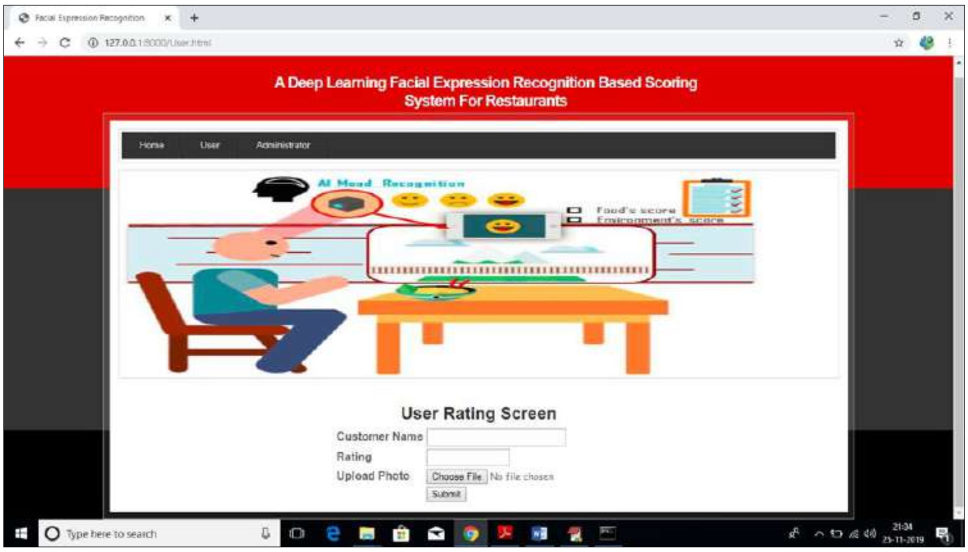


Fig 4.2: User will fill above from and upload photo

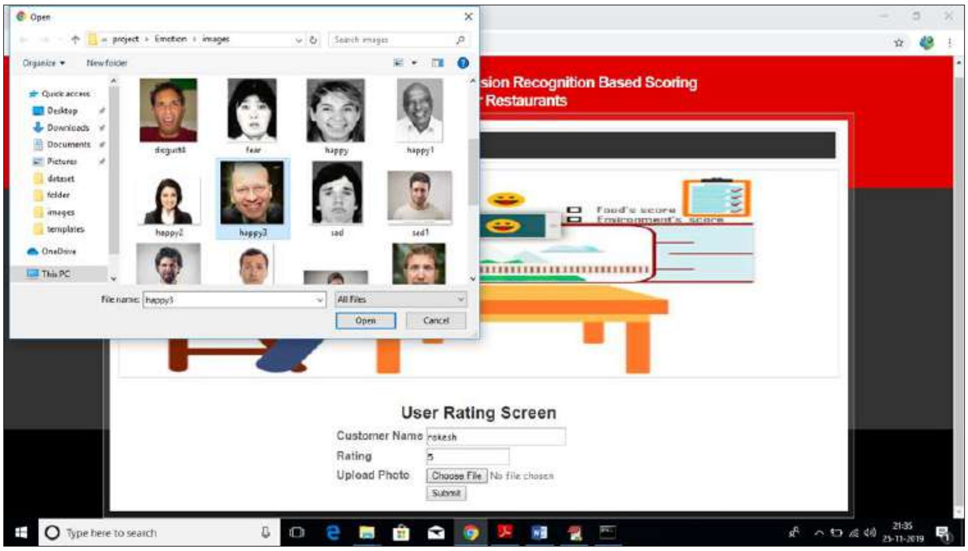


Fig 4.3: In above screen I filled form and uploading one happy image and then click on ‘Open’ button and then click ‘Submit’ button to send data to webserver. After processing above data will get below results.

Fig 4.4: 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.

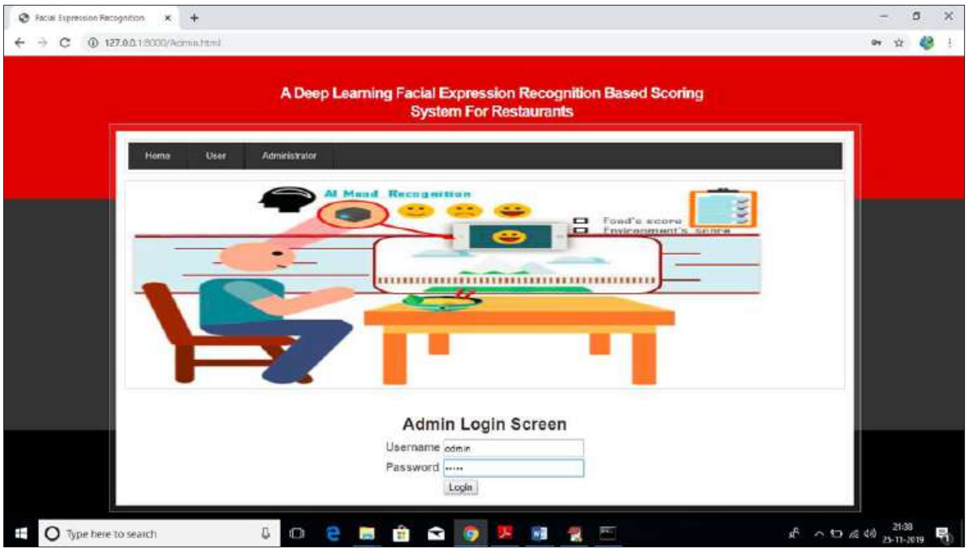
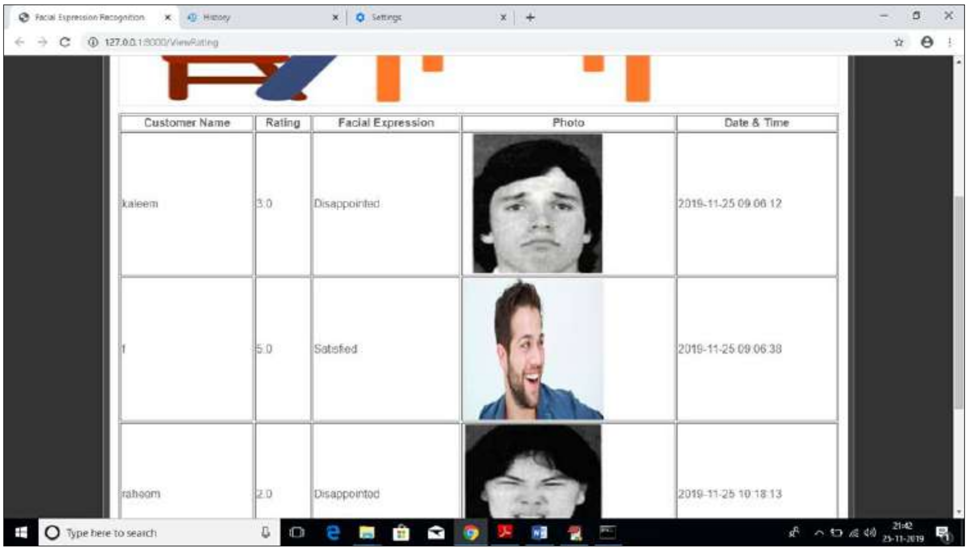


Fig 4.5: After login will get below screen

Fig 4.6: In above screen click admin can click on 'View Users Rating' link to get all customers feedback. See below screen



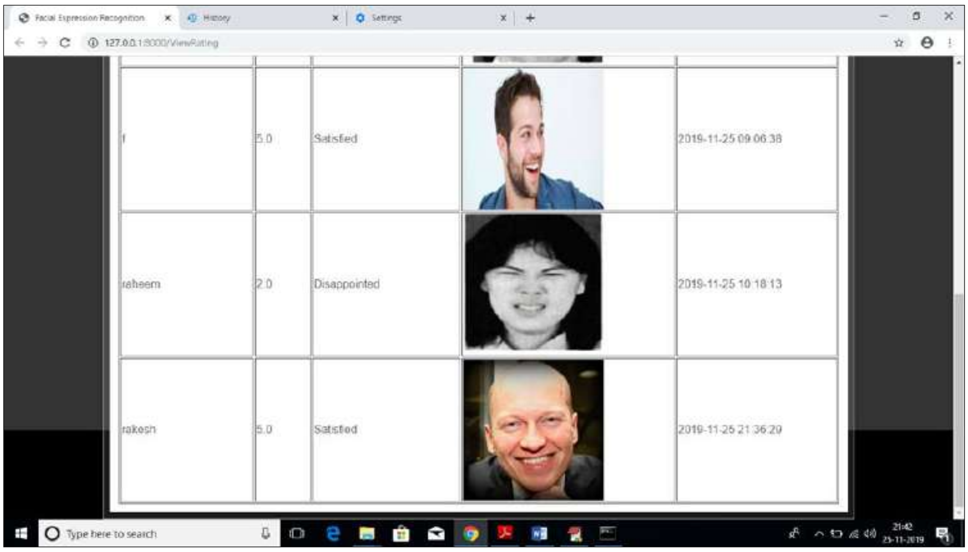


Fig 4.7: From above screens admin can see photos and their facial expressions

5. 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.

6. References

1. Darwin C, Prodger P. The expression of the emotions in man and animals. Oxford University Press, USA, 1998.
2. Tian YI, Kanade T, Cohn JF. Recognizing action units for facial expression analysis, IEEE Transactions on pattern analysis and machine intelligence. 2001; 23(2):97-115.
3. Ekman P, Friesen WV. Constants across cultures in the face and emotion.” Journal of personality and social psychology. 1971; 17(2):124-129.
4. Ekman P. Strong evidence for universals in facial expressions: a reply to russell’s mistaken critique, Psychological bulletin. 1994; 115(2)268–287.
5. Matsumoto D. More evidence for the universality of acontempt expression,” Motivation and Emotion, vol. 1992; 16(4):363-368.
6. Jack RE. Garrod OG, Yu H, Caldara R, Schyns PG. Facial expressions of emotion are not culturally universal, Proceedings of the National Academy of Sciences. 2012; 109(19):7241–7244,