Multi Facial Expression Recognition (MFER) for Identifying Customer Satisfaction on Products using Deep CNN and Haar Cascade Classifier

DNVSLS Indira, I. Sumalatha, Babu Rao Markapudi
1Associate Professor, 2Professor and Director of Evaluation, 3Professor and Head
1Dept. of Information Technology, 2Dept. of CSE
1Dept. of CSE, Jawaharlal Nehru Technological University Kakinada (JNTUK), Kakinada.
3Gudlavalleru Engineering College, Gudlavalleru – 521356

1,2,3E-mail: indiragamini@gmail.com, lsumalatha@jntucek.ac.in, baburaompd@gmail.com

Abstract. Face Expression is one of the most normal, remarkable and a general sign for individuals to convey on their enthusiastic states and it is not restricted to national borders, linguistics and gender. This article presents the modeling of a framework that plans to foresee the fulfillment of a customer through his facial feelings. The cutting edge innovation of Facial Expression Recognition framework is the consumer satisfaction estimation. MFER, a Novel procedure is proposed in this paper for identifying consumer satisfaction levels. This sound methodology of client satisfaction estimation is an alternative option of the ordinary method of gathering clients' reaction. This model must anticipate client's behavior in the dynamic cycle. To expect consumer trustworthiness, we have characterized mathematical highlights of the face by utilizing Deep CNN and Haar Cascade Classifier. The kinds of consumer fulfillment are classified as satisfied, not-satisfied and neutral. Our framework shows a decent exhibition, testing it on the FER2013 dataset. Our MFER –Multi Facial Expression Recognition procedure identifies multiple objects in the same image which consists of same and different expressions.

Keywords: MFER, SVM, Feature Extraction, Deep CNN, Haar Cascade Classifier, Recommendation Systems

1. Introduction

Facial expression investigation includes estimation and acknowledgment of articulation. Computerization of Facial expression prediction comprises of three ideas initial one Face Detection (FD), second one Facial Feature Extraction (FFE) lastly the Facial Expression Recognition (FER)[2]. Computerized Facial Recognition Systems have the ability to distinguish people from their facial highlights. Facial feelings are significant elements in human correspondence that assist us with understanding the goals of others. As a rule, individuals derive the passionate conditions of others, for example some expressions like anger, joy and sadness, utilizing facial expressions and vocal tone[4][5]. In this way, it is regular that exploration of facial feeling has been picking up part of consideration over past a long time with applications in the perceptual and psychological sciences, but also in affective computing and computer animations.

Facial Expression Recognition (FER) [13] is the way toward identifying human feelings from outward appearances. Individuals can perceive the articles by observing them with less exertion. Given a picture,
making the machines to perceive and comprehend the item is a difficult assignment. These image recognition application models can be used to detect the expression or emotion in the face. One of the applications of identifying the emotion of a person is opinion mining to measure the customer satisfaction about a product[6]. These customer satisfactions can be used for further recommendation on products, movies and for any business intelligence applications when users give their opinions or reviews as a facial expression image file. This article presents literature survey in the next section. In the subsequent sections it covers the importance of Haar cascade classifier in the FER in phase 3 and then proposed MFER procedure and finally results and references.

2. Literature Survey

Face is one of the main techniques for human correspondence. It accepts a central occupation in each and every social affiliation. Facial enunciations are non-verbal insinuations to emotions. To start with, face securing is a preparing stage to naturally find the face area in the information pictures. The subsequent stage is to extricate and speak to facial changes brought about by outward appearances. At long last, the arrangement task permits to derive the outward appearances. The extraction strategies are utilized to portray face appearances and changes on a face [1].

The creators in [2][3] characterized facial segments from which they extricated HOG include descriptors. The goal of this work is to perceive consequently the six fundamental feelings just as the nonpartisan state by applying appearance-highlight strategies like LBP and its variations to new explicit face-areas. Different authors discussed and worked on different methods. The following table gives brief information about the algorithms mentioned in the reference articles.

| Reference Nos. | Year | Methods Discussed                                      |
|----------------|------|--------------------------------------------------------|
| [1]            | 2014 | LTP, MTP, CLBP.                                        |
| [2]            | 2015 | HOG, PCA, LDA and Wavelet                              |
| [3]            | 2015 | MDA and SVM                                            |
| [4]            | 2016 | EEG signal                                             |
| [5]            | 2016 | Event Related Potentials (ERP’s)                       |
| [6]            | 2016 | Harris Algorithm                                       |
| [7]            | 2017 | CNN                                                    |
| [8]            | 2017 | FERS                                                   |
| [9]            | 2018 | LBP and SIFT into CNN                                  |
| [10]           | 2018 | Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) |
| [11]           | 2019 | Deep CNN                                               |
| [12]           | 2019 | DLP-CNN                                                |
| [13]           | 2019 | SVM                                                    |
2.1. Existing Linear Support Vector Machine for Recognizing Face

All most all authors till now they used SVM for recognizing face in the given image. SVM or Support Vector Machine is a straight model for arrangement and relapse issues. It can take care of direct and non-straight issues and function admirably for some handy issues[13]. The possibility of SVM is straightforward. SVM identifies the ROI (regions of interest) in the face, which consists of eyes (left, right), eyebrows(left, right), mouth, nose and distance between eyebrows These features are fed to SVM to recognize the face in the image. Our proposed method identifies multiple faces in the same image with different expressions using haar_cascade classifier.

3. PROPOSED METHODOLOGY

When arranging a modified system, three things are thought of; they are face recognition, facial features extraction, and characterization of articulations or expressions. At first, face recognition is a phase to normally discover the face district in the given pictures. The accompanying stage is to discover facial changes achieved by articulations. Finally, the arrangement determines the outward appearances that are the expressions. Here are three significant strides for building this model. One is face recognition which identifies the person who is giving the response and the later one is feature extraction by identifying geometric features and the final one is emotion recognition which indicates the type of response from that person.

![Flow Diagram](image)

The proposed framework actualizes a Deep CNN model and a classifier for recognizing expressions by building system. The reason for MFER is used to identify the customer satisfaction on a purchased product when he/she gives opinion as an image file like an expression for further recommendations.

3.1. Face Recognition using Haar_Cascade Classifier

Haar_Cascade Classifier is used to detect objects in an image in a set of machine learning algorithms. It was proposed by Paul Viola and Michel for recognizing faces in the given image. Haar features are different from primary features of images those are color, text and shape. Generally these Haar features are Edge, Line and Rectangle like in convolution networks. It is just like kernel in CNN. Every feature
obtained from this is a single value. This can be achieved by subtracting sum of pixels in white area by sum of pixels in black area in one edge.

**Figure 2: Haar Features**

This algorithm passes into four steps. 1. Selection of Haar Features 2. Integral Image Creation 3. Training of input images dataset with ADAboost algorithm. 4. Finally Applying Cascading Classifier for detecting face image and non-face image.

**Figure 3: Calculating output one pixel value using four rectangular edge feature in the input image**

### 3.2. Proposed MFER Steps in Python

1. Collect the data from Kaggle, the data set is FER2013.
2. Load the data set.
3. Perform data visualization on the data set.
4. Perform data preprocessing on the data set to remove the missing values and outliers.
5. Split the data set into Train and Validation set.
6. Define the model
   A. Sequential => linear stack of NN layers -> feed forward cnn
   B. layers -> for almost any nn
   C. CNN layers -> convolution2D , maxpooling2D
7. Perform the model summary using Model.Summary().
8. Data Augmentation is done by taking a batch and applying some series of random transformation to increase generalizability of model.
9. Neural networks learn more complex features from input image.
   A. input layer => feeding input
3.3. The System Architecture of MFER

It consists of the following steps:
1. Data Preparation
2. Feature Extraction using Deep CNN
3. Visualization
4. Training and Testing of Data
5. Identifying Expressions
6. Result with the Polarity of the Person like: Satisfactory, Not Satisfactory and Neutral

![Figure 4: System Architecture for Facial Emotion Recognition](image)

3.3.1 Data Preparation

Initially, the data set to train the emotions is resourced from FER2013 which contains seven categories of emotions such as happy, fear, sad, anger, disgust, surprise and neutral. The required data is only any of the three categories which define the ratings of satisfactory, not satisfactory and neutral.
3.3.2 Feature Extraction using Deep CNN

The CNN utilizes various algorithms which calculate the measures of features of face. This is called feature extraction. Each dimension of eye, nose, brows, mouth, jaw line and so on. Different fuzzy logic techniques are used to find the distances like Euclidian distance.[11]

The customized pictures that are being perceived are spared in a folder in a similar catalog. These are the enlisted understudies pictures gathered prior to getting to the criticism framework. The advantage of this methodology is a lot more noteworthy authentic productivity as it accomplishes best in class face acknowledgment execution. The method relies upon learning an Euclidean embedding per picture using a significant convolution organization. The system is set up with the ultimate objective that the squared L2 eliminates in the embedding space clearly identify with go up against comparability. FaceNet[10] utilizes a profound convolution organization. Using FaceNet algorithm we extracted geometric features by highlighting ROI on faces. This method calculate 8 different features in the image which has eyes, nose, mouth, eyebrows etc.

3.3.3. Visualization

Data portrayal is the control of endeavouring to understand statistics by placing it in a visual setting so models, examples and connections that may for no situation be recognized can be revealed. Python offers different exceptional graphing libraries that come squeezed with heaps of different features. Despite if you have to make savvy, live or significantly changed plots python has a surprising library for us.
3.3.4. Training and Test Data

Almost all Data mining applications use supervised data classification. For this the entire data is classified into training and test data sets. Here in this article the data set that is used is FER2013. It consists of nearly 3500 records. We used 3000 records as training data and 500 records for test data. Apart from that we collected some images manually from different websites that expresses opinion on products and own data prepared by taking photos of our friends by giving opinion on product by facial expression. This model works beautifully on test data which gives about 89% accuracy for some expressions.

3.3.5. Identifying Expressions

Outward appearance recognition as a rule brings about classes as per either the Facial Actions Coding System (FACS) or the six fundamental outward appearances. The information layer gets six feelings and the yield layer speaks to 6 potential facial appearances.

3.3.6. Result with The Polarity of the Person

The result of the model is a box on the face of a person with emotion on the top left of the image. These six emotions are classified into three customer satisfaction levels. Happy and Surprise are classified as Satisfactory, Sad, Angry and Disgust are Completely Not Satisfactory and Neutral are Neutral class.

4. Results and Discussion:

![Image](shutterstock.com • 1235498808)

Figure 7: Output from the proposed MFER on test data. Which shows the multiple expressions on each face in the given input image.
Facial Emotion Recognition is crucial in many ways providing details about the emotional state. Many researches and studies about Emotion Recognition are used for recognizing the emotions are performed by learning about six basic emotions. Emotions in real life are very complex, so the future scope for this work is to build larger databases and create a powerful model that recognizes all basic and secondary emotions. This paper uses Deep CNN on Kaggle FER2013 dataset which contains 3500 images for emotion recognition. Accuracy rate of about 78% has been achieved. The article identifies expressions of multiple faces in a single image. This type of identification is not possible for the system which are in use.
5. Conclusion

As facial expression recognition technology improves in accuracy, the range of its application will grow, both in the innovation improves in precision, the scope of its application will develop, both in mechanical turn of events and past. The model is intended to take input or audit from a client devouring an item. The survey about the item helps in improving the item qualities. A Facial Expression Recognition framework is intended to take certifiable reaction from the purchaser or client who uses an item. The real reaction is taken as outward appearance, which indicates what is running in the psyche of a client. The report made so ever from the reaction of the client improves the profitability in a business. This framework at times clarifies that what sorts of individuals are utilizing the item and how they are expecting the conveyance or how the administration could be stretched out to arrive at clients desires.

This work can be applied in various divisions or units such as education, business and software. In software field it works between a manager and his subordinates while analyzing their skills. Proposed MFER outputs are compared with existing SVM classifier output. Our model gives best performance in initial stage of Face Recognition (FE) then the next steps automatically improves accuracy for finding customer satisfaction on products when the test data is sample of consumer interest on goods.

6. Future Scope

Machine learning models are biased in accordance to their training data. In my specific application our trained model that adopts Haar_cascade classifier along with Deep CNN is used to fine the emotions. This work can be integrated with speech recognition project and can be extended to find emotions or expressions from video.

References

[1] F. Ahmed, H. Bari and E. Hossain, "Person-independent facial expression recognition based on compound local binary pattern (clbp)”, Int. Arab J. Inf. Technol., vol. 11, no. 2, pp. 195-203, 2014.
[2] P. Carcagni, M. Coco, M. Leo and C. Distante, “Facial expression recognition and histograms of oriented gradients: a comprehensive study”, SpringerPlus, vol. 4, no. 1, pp. 1, 2015.
[3] S. Kumar and M. K. Bhuyan, “Neutral expression modeling in feature domain for facial expression recognition”, IEEE Recent Advances in Intelligent Computational Systems (RAICS), 10-12 Dec. 2015.
[4] A. M. Bhatti, M. Majid, S. M. Anwar and B. Khan, “Human emotion recognition and analysis in response to audio music using brain signals”, Computers in Human Behavior 65:267-275, Dec 2016.
[5] P. Rajhans, S. Jessen, M. Missana and T. Grossmann, “Putting the face in context: Body expressions impact facial emotion processing in human infants”, in Developmental Cognitive Neuroscience, Vol 19, 115-121, 2016.
[6] Vikrant Chaugule, Abhishek D, Aadheeshwar Vijayakumar, Pravin Bhaskar Ramteke, and Shashidhar G. Koolagudi, “Product Review Based on Optimized Facial Expression Detection”, 978-1-5090-3251-8, IEEE, 2016.
[7] B Hasani and M H. Mahoor, "Spatio-Temporal Facial Expression Recognition Using Convolutional Neural Networks and Conditional Random Fields[C]", IEEE International Conference on Automatic Face & Gesture Recognition, pp. 790-795, 2017.
[8] U. Ayvaz, H. Gürüler and M. O. Devrim, "USE OF FACIAL EMOTION RECOGNITION IN E-LEARNING SYSTEMS", Information Technologies and Learning Tools, vol. 60, no. 4, pp. 95, sept. 2017.
[9] B. Yang, J. Cao, R. Ni and Y. Zhang, "Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images", IEEE Access, vol. 6, pp. 4630-4640, 2018.

[10] S R Livingstone and FA Russo, "The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic multimodal set of facial and vocal expressions in North American English", PLoS ONE, vol. 13, no. 5, pp. e0196391, 2018.

[11] N. Veeranjaneyulu, Jyostna Devi Bodapati, “Facial Emotion Recognition Using Deep Cnn Based Features ISSN: 2278-3075, vol. 8, no. 7, May 2019.

[12] W. Deng and S. Li, "Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition", IEEE Transactions on Image Processing, vol. 28, no. 1, pp. 356-370, 2019.

[13] Moulay Smail Bouzakraoui, Abdelalim Sadiq, Abdessamad Youssfi Alaoui, “Appreciation of Customer Satisfaction Through Analysis Facial Expressions and Emotions Recognition”, 978-1-7281-1232-9, IEEE, 2019.