IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-2, Feb 2019 E-ISSN: 2347-2693

Recognizing Human Emotional State from a Machine

Saniya Zahoor

Srinagar, J&K, India

DOI: https://doi.org/10.26438/ijcse/v7i2.674678 | Available online at: www.ijcseonline.org

Accepted: 17/Feb/2019, Published: 28/Feb/2019

Abstract— Emotion recognition is an approach to detect and analyze different human emotions from their facial expressions. We, humans, have versatile and consistent abilities through variable among individuals in detecting emotions, but the high-speed computer/ machines lag humans in this capability. Here, an approach is developed which will make the machines detect human emotions and learn to behave with them by analyzing their emotions. In this paper, computer vision techniques are used for feature extraction and machine learning tools and Artificial Intelligence tools for training purposes. This paper proposes a neural network based solution combined with image processing in classifying the universal emotions – Happiness, Sadness, Anger, Disgust, Surprise, and Fear.

Keywords—Emotions, Neural Networks, Artificial Intelligence, Data Mining, Image Processing.

I. Introduction

Recognizing human facial expression and emotion by a machine is challenging task [1][2]. There has been recent trend towards improving the interaction between humans and computers. There is a need for the computer to interact naturally with the users in order achieve human-computer intelligent interaction. Humans interact with each other through speech and body gestures to display of emotions. Emotions are displayed by visual, vocal and other physiological means. Various studies reveal that emotional skills are part of 'intelligence'. One of the most important ways for humans to display emotions is through facial expressions, so recognizing the emotional state of the human from the human face could prove to be an individual tool. This will help us to achieve more effective human-computer interaction.

A facial expression is a gesture executed with the facial muscles, which convey emotional state of a human and sends a message about a person's internal feeling. Although there is wide range of verbal languages, facial expression has a great role in interactions. Expressions and emotions go hand in hand i.e., special combinations of face muscular actions reflect a particular emotion.

In communication with others, humans can recognize emotions of another human with a considerable level of accuracy, but with varying consistency, while machines especially computers are lagging in this respect, and hence the development of Human-Computer Interface (HCI) has not advanced to that level [3]. In this paper, an approach is proposed that would bridge the interaction gap between the machines and humans by recognizing the emotions from their facial expressions with the help of Image Processing and Artificial neural network tools.

The organization of the paper is: Section II presents a brief literature survey; Section III gives implementation; Results are shown in Section IV, and Section V gives the conclusion.

II. Related Work

There has been an extensive study of human facial expressions from early 1970's. Recent studies have found evidences to support universality in facial expressions [4] and these 'universal facial expressions' represents happiness, sadness, anger, fear, surprise, and disgust. Study has analysed expressions in many cultures and found many commonalities in the expression and recognition of emotions on the face. However, there are differences as well. [5] developed a coding system for facial expressions where movements of the face are described as a set of action units and each action unit has some related muscular basis. Recent studies has been inspired to use the image and video processing to automatically track facial features to use them for categorization of different expressions. The computer programmers often use Paul Ekman's Facial Action Coding System (FACS) as a guide. Emotion recognition is used for a variety of reasons. Affectiva also makes a Q-sensor that gauges the emotions of autistic children [6]. Emotient utilized artificial intelligence to predict attitudes and actions based on facial expressions.

Facial Action Coding System is a system to taxonomize human facial movements by their appearance on the face, based on a system originally. Movements of individual facial muscles are encoded by FACS from slight different instant changes in facial appearance. It is common standard to systematically categorize the physical expression of emotions, and it has proven useful to psychologists and to animators. Due to subjectivity and time-consuming issues, FACS has been established as a computed automated system that detects faces in videos, extracts the geographical features of the faces, and then produces temporal profiles of each facial movement [7,8,].

III. Implementation

For facial feature extraction and emotion detection, image datasets are usually used and KDEF (Karolinska Directed Emotional Faces), a set of totally 4900 pictures of human facial expressions of emotion. It is the set of 70 individuals, each displaying 7 different emotional expressions, each expression being photographed (twice) from 5 different angles.
For the implementation, KDEF dataset is used. Out of 4900 images, 980 images are used i.e., 7 emotions per subject. For this dataset, the input images to the system should follow the following constraints:

- The subject needs to be seated at a distance of approximately three meters from the camera.
- The image needs to be of the size 256*256 pixels.
- The picture should contain only the face of the subject ideally part above the neck and shoulders.

The core of the work is to detect facial features which include:

- Center of Mouth
- Mouth Width
- Mouth height
- Center of the right eye
- The height of right eye
- Center of the left eye
- The height of left eye
- The distance of left eyebrow from the center of the mouth
- The distance of right eyebrow from the center of the mouth

These nine features were to be detected or calculated which in return meant for us to calculate 13 points in an image as:

- 1 point for the center of the mouth
- 2 points for Mouth width
- 2 points for Mouth height
- 1 point for Center of the right eye
- 2 points for Height of right eye
- 1 point for Center of the left eye
- 2 points for Height of left eye
- 1 point for Distance of left eyebrow from the center of the mouth
- 1 point for Distance of right eyebrow from the center of the mouth

Figure 1 shows the feature points in an image that are extracted.

Figure 1. Extracted Feature Points

Histogram of Oriented Gradients (HOG) is a feature extraction method used for object detection. It counts occurrences of gradient orientation in localized portions of an image. It detected HOG features in the input image. Its primary use isto detect the subject, which is the input image and plot orientation bins. These orientation bins detect each and every corner in the image, which will be used further for the feature extraction. To extract HOG features in MATLAB followings commands were used:

```
[featureVector, hogVisualization] =
extractHOGFeatures(Image);
```

The above command returns HOG features extracted around specified point locations of the image. The example image is shown in figure 2.

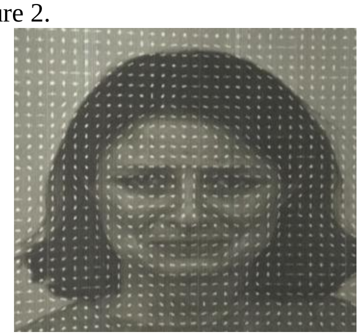

Figure 2 – HOG Implementation

As seen in figure 2, almost all the curvatures in the image are detected. But only portion around eyebrows, eyes, and mouth are needed, a narrow down of output from HOG feature extraction method has to be done.

As already discussed, Haar-like features are digital image features to detect eyebrows, left eye, right eye and mouth. In MATLAB, Haar-like features can be applied by using CascadeObjectDetector which comes under the package vision.

Haar cascade is applied and detected five rectangles for two eyebrows, two eyes, and mouth as shown in figure 3.

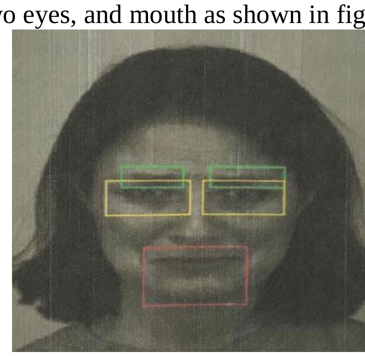

Figure 3 – Haar Cascade Implementation

After detecting these five features there is a need to pinpoint the exact coordinates for maximum accuracy. For that, corner detection algorithms are used. The rectangular regions detected above there would be intensity variations. As for example, the iris of the eye should be darker than its neighboring pixels. So, to detect these intensity changes, corner detection algorithm is used. Few of the corner detection algorithms in MATLAB are detectFASTFeatures, detectMinEigenFeatures,detectMSERFeatures, etc.

As already discussed, I have used KDEF image dataset and filtered out 980 images for this work. The next step isto build training data for the neural network. A program is developed to calculate feature points. These feature points include Mouth width, Mouth weight, Eye height and distance from the center of mouth to the eyebrow.

Among these 980 images, few images were not compatible with the program and were ruled out. It isleft with 913 images. So, a matrix of four rows with each row depicting each feature for 913 images is created. Finally, it had a 4*980 matrix dataset. Few details of the database are:

Table 1. Image Samples

Emotion	Number of Subjects
Afraid	128
Angry	130
Disgust	134
Normal	121
Sad	131
Happy	132
Surprise	137

Comparison of the emotions based on the attributes selected is given in Table 2. In order to efficiently classify these emotions, the training algorithm and tool used would have to create classifiers equal to the number of emotions i.e. seven. However, to solve mapping from 4*913 matrixes to seven classifiers will be highly inaccurate. To compensate for the higher accuracy, it is decided to detect only three basic emotions – Sad, Happy and Surprise.

Table 2.Image Attributes

Emotion	Mouth Width	Mouth Height	Eye Height	Eyebrow Point
Afraid	39.64	13.80	10.35	69.68
angry	44.06	14.87	19.27	68.73
Disgust	55.32	18.23	12.89	70.23
Normal	56.21	16.22	13.56	72.67
Sad	54.21	16.28	13.77	72.11
Happy	28.32	12.56	15.83	63.42
Surprise	40.80	16.50	20.86	75.09

The reasons for choosing these three emotions include the following:

- It included both positive and negative emotions – happy and surprised emotions are considered positive and sad emotion is considered as negative. Here, positive and negative are in the sense of brain activity; High for positive emotions and less for negative emotions.
- These three emotions are universal emotions. By universal, means these emotions are across each and every culture.
- These three emotions are also very common. Hence, this work can be applied in any field.
- The higher accuracy set as a platform for future work, and if all the emotions are detected with the same accuracy, it would be a huge breakthrough.
- Marketing research, advertisements, amusement parks, customer feedback, etc are few fields where this work can be applied because not all emotions are present in these scenarios.
- These three emotions can help us detect emotions other than basic emotion like depressed, joy, etc.

- In the medical field, it can help the patients with autism discern emotion.

Hence, the final data is reduced to 4*400 matrix containing data of sad, happy and surprise emotions only. For training purpose, it is required to provide input data and target data to the training tool. The input data is already calculated. For training data, a target values are set for each emotion. Following target valued is selected:

Table 3. Emotion and target values

Emotion	Target Value
Happy	0
Surprise	30
Sad	60

Finally, a matrix is created with dimensions 1*400 containing 0 (132 occurrences) parallel to happy emotion, 30 (131 occurrences) parallel to surprise emotion to surprise emotion and 60 (137 occurrences) parallel to sad emotion. After creating the input data and target data, it proceeded to neutral networking training in MATLAB.
In MATLAB, a tool to create a network and train it for the data is used. It is trained for different algorithms and then choose the most accurate one. Following algorithms were used for training the data: Lavenberg-Marquardt (trainlm in MATLAB), Bayesian Regularization (trainbr in MATLAB), Gradient Descent (traingd in MATLAB) and Stochastic Gradient Descent (triangda in MATLAB).

IV. RESULTS

Few properties of emotions obtained by training are shown:

A. Accuracy vs number of features selected

I have selected only one feature (mouth width) at first and trained the network. It reached an accuracy of about 30%. The network is trained again again but now with two features (mouth width and height) to reach up to the accuracy of about of about 55%. The third feature (eyebrow point) didn't affect the accuracy much and the accuracy of approximately 60% is achieved.

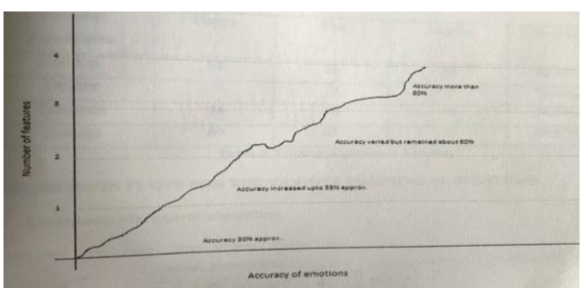

Figure 4. Accuracy with different number of features

B. Accuracy with different algorithms

The accuracy didn't change much when different algorithms were used. However, for the Lavenberg-Marquardt algorithm, the accuracy is considerably high. After training, we gave a hundred samples for each emotion to the network and got following results:

Table 4. The accuracy of Training algorithms

Training algorithms	Happy	Sad	Surprise	Accuracy
Trainlm	88	86	85	86.34%
Trainbr	77	78	79	78.00%
Traingda	74	70	79	74.34%
Traingd	69	67	73	69.67%

C. Regression plots for the various training algorithms:

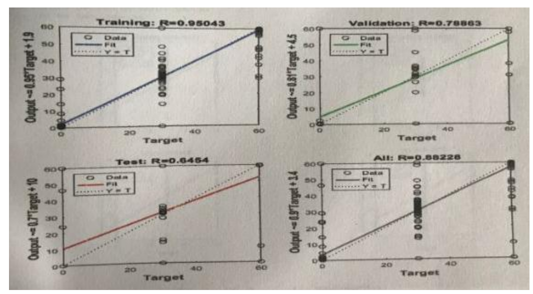

Figure 5. Regression Plot (Lavenberg-Marquardt)

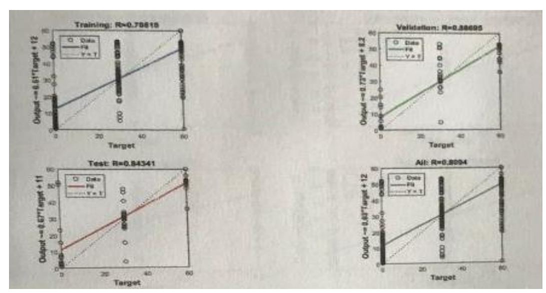
Figure 6. Regression Plot (Bayesian Regularization)

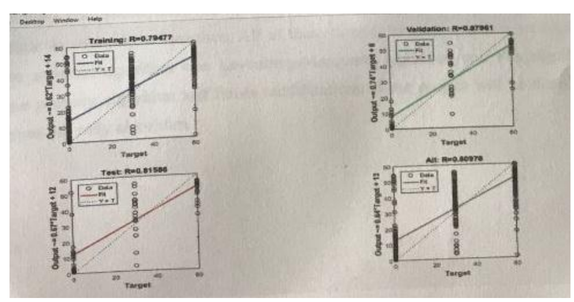
Figure 7. Regression Plot (Stochastic Gradient Descent)

D. Training algorithm selected:

We trained the neural network by four algorithms and all of them have a separate command button in the system interface. All of them provide correct output most of the time but the most accurate algorithm is Levenberg-Marquardt algorithm.

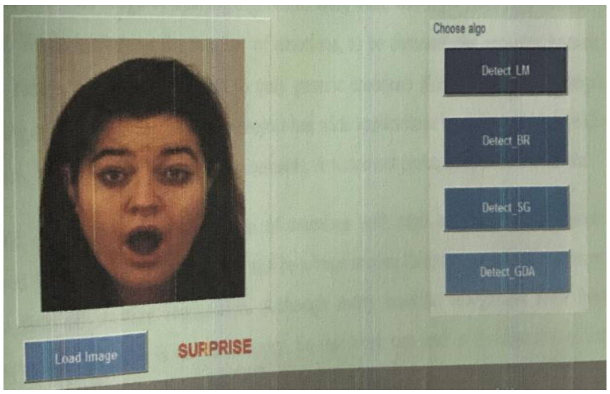

Figure 8 – System Interface

V. CONCLUSIONS

This paper detects the emotions with maximum accuracy. It is able to get the accuracy of more than 85%. A wide range of training algorithms isused, which provides high flexibility to the system. While increasing the number of emotions, to be detected, the accuracy became a vital factor. Hence, the work is limited to only generic emotions like sad, smile and surprise. By detecting only generic emotions, the work has wide applications in various fields such as customer feedback, marketing research, advertisements, amusement parks, student feedback, etc. The implementation is such that it can be upgraded to real-time detection, which will widen its scope to these online applications as well.

REFERENCES

[1] Yacoob, Y. and Davis, L.S., 1996. Recognizing human facial expressions from long image sequences using optical flow. IEEE Transactions on pattern analysis and machine intelligence, 18(6), pp.636-642.

[2] Kanade, T., Cohn, J.F., and Tian, Y., 2000. Comprehensive database for facial expression analysis. In Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on (pp. 46-53). IEEE.

[3] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W. and Taylor, J.G., 2001. Emotion recognition in human-computer interaction. IEEE Signal processing magazine, 18(1), pp.32-80.

[4] Ekman, P. and Oster, H., 1979. Facial expressions of emotion. Annual review of psychology, 30(1), pp.527-554.

[5] Ekman, P. and Rosenberg, E.L. eds., 1997. What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford University Press, USA.

[6] Sayette, M.A., Cohn, J.F., Wertz, J.M., Perrott, M.A. and Parrott, D.J., 2001. A psychometric evaluation of the facial action coding system for assessing spontaneous expression. Journal of Nonverbal Behavior, 25(3), pp.167-185.

[7] Ekman, P., Hager, J.C. and Friesen, W.V., 1981. The symmetry of emotional and deliberate facial actions. Psychophysiology, 18(2), pp.101-106.

[8] Gupta A, Sharma E, Sachan N, Tiwari N. Door Lock System through Face Recognition Using MATLAB. International Journal of Scientific Research in Computer Science and Engineering. 2013 Jun;1(3).

Authors Profile

Ms. Saniya Zahoor pursed Bachelor of Engineering from University of kashmir in 2011 and Master of Enginering from Pune University in year 2015. She has published 19 research papers in reputed jourals, international conferences and book chapters in IEEE, Springer, Elsievier and ACM. Her main research work focuses on Internet of Things, Personalization, Big Data Analytics, Data Mining.