

Facial Emotion Recognition for Smart Applications

Fathimath Hafisa Hariz¹, K Nithin Upadhyaya², Sumayya³, T F Mohammad Danish⁴,

Bharatesh B⁵, Shashank M Gowda⁶

Department of Electronics and Communication Engineering

Yenepoya Institute of Technology, Moodbidri

Karnataka, India

Abstract - Facial expressions play a significant role in nonverbal communication. One of the important fields of study for human-computer interaction (HCI) is detecting facial expressions using emotion detection systems. Facial expressions are detected by analyzing the discrepancy in various features of human faces such as color, posture, expression, orientation etc. To detect the expression of a human face it is required to detect the different facial features such as the movements of eye, nose, lips, etc. and then classify them by comparing with trained data using a suitable classifier for expression recognition. The proposed Facial Emotion Recognition (FER) system uses Viola Jones algorithm for face detection, a pre-trained Convolutional Neural Network (CNN) called Alexnet for feature extraction and the support vector machine (SVM), a machine learning algorithm for classification of emotions. To demonstrate its application the proposed FER system is then applied for smart applications such as in smart cars to control the speed of the automobiles based on driver's emotions, regulating the drivers emotions using mood lighting and displaying the detected emotions on an LCD display.

Keywords - HCI; pre-trained CNN; SVM; Alexnet; facial expressions

I. INTRODUCTION

Human emotions play an important role in the interpersonal relationship. Emotions are reflected from speech, hand and gestures of the body and through facial expressions. The human brain tends to recognize the emotions of a person more often by analyzing his face. Hence extracting and understanding of emotions has a high importance in the field of human machine interaction.

Facial expression recognition (FER) systems uses computer based algorithms for the instantaneous detection of facial expressions. For the computer to recognize and classify the emotions accordingly, its accuracy rate needs to be high. To achieve higher this, a Convolutional Neural Network (CNN) model is used. CNN is a type of Neural Network method. Neural Network is a combination a number of neurons which take in input and provide an output by applying an activation function. The activation function is a function which is applied to the input in order to achieve an output. In CNN model of Neural Network the activation function is Convolution. A CNN model works well with larger database as it trains the model with specific features of each image in the database which helps it to recognize and classify the images into different

categories as per the requirement. The time consumed by this model is significantly less while compared to other models.

The main objective is to design a suitable facial emotion recognition system to identify various human emotional states by analyzing facial expressions which is used to improve Human-Computer Interaction (HCI). This is achieved by analyzing hundreds of images of different emotional states of a human being. The analysis is carried out by a computer system which uses different technologies to recognize and study the different emotional states. This system can be applied to smart cars for improving automatic driver assistance systems.

II. FACIAL EMOTION RECOGNITION

The proposed FER system uses Extended Cohn-Kanade database (CK+) and Japanese Female Facial Expression (JAFFE) database [6] for training the model. Facial emotion recognition involves three major steps i.e., face detection, feature extraction and expression classification.

A. Face Detection

The face is detected using CascadeObjectDetector which is an inbuilt Matlab function. This function is based on Viola-Jones algorithm and is used to detect human faces [5]. The detected face is then cropped to eliminate other non-expressive features from the image.

B. Feature extraction

The features required for emotion classification are extracted using a pre-trained Convolutional Neural Network model named Alexnet. It consists of eight layers, out of which five are convolutional layers and three are fully connected layers. Alexnet requires input images of size 227x227x3. Alexnet is trained on the Image-Net Large-Scale Visual Recognition Challenge (ILSVRC) which consists of 1.2 million images in the Image-Net database. Alexnet is trained to classify these images into 1000 different classes. This paper focuses on classifying the images into seven different classes based on the seven different expressions as suggested by Dr.Paul Ekman [1]. The features are extracted from the fully-connected fc7 layer of Alexnet.

C. Expression classification

The extracted features are passed through a classifier to find its corresponding labels. The classifier used here is a multiclass Support Vector Machine [10]. A Support Vector Machine (SVM) is a supervised machine learning model that uses classification algorithms for two-group classification problems. The multi-class SVM is implemented for a set of data with M classes, where M binary classifiers can be trained that can distinguish each class against all other classes and then select the class that classifies the test sample with the greatest margin.

III. IMPLEMENTATION

The proposed system performs two main tasks i.e., training the classifier model and testing an unseen image using the trained classifier model. This system has been developed using Matlab .

A. Training the classifier model

The customized database contains 1195 images consisting of both JAFFE as well as CK+ datasets. JAFFE database contains 213 images of 7 facial expressions (6 basic facial expressions and 1 neutral). In this experiment 80% of the database is used as the training data and the rest 20% as validation data. Fig. 1 show the steps involved in training the classifier.

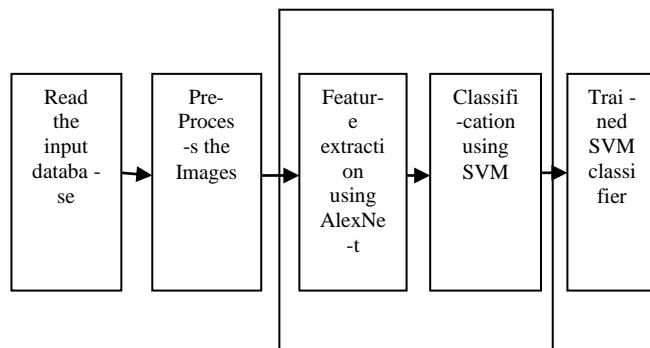


Fig. 1. Training the classifier model

Since AlexNet only accepts the image with input size of 227x227x3, the channel is replicated three times to convert single channel gray scale images of JAFFE and CK+ to three channels and also the input images are resized to 227x227. In this model, the features are extracted from the fully connected layer fc7 of Alexnet. The feature vector along with the training labels is fed into the SVM classifier for classification and validation.

B. Testing an unseen image

Fig. 2 shows the steps involved in testing an unseen image. The test image is captured using a webcam and is read. The face is detected using cascade object detector (Viola Jones). The bounding box enclosing the detected face is cropped and resized to 227x227.

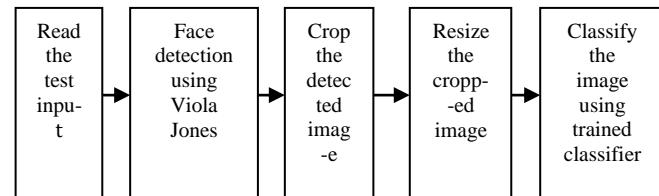


Fig. 2. Testing an unseen image

The test image label is then predicted using the trained classifier model. The application of the proposed FER system is tested using the setup shown in Fig. 3.

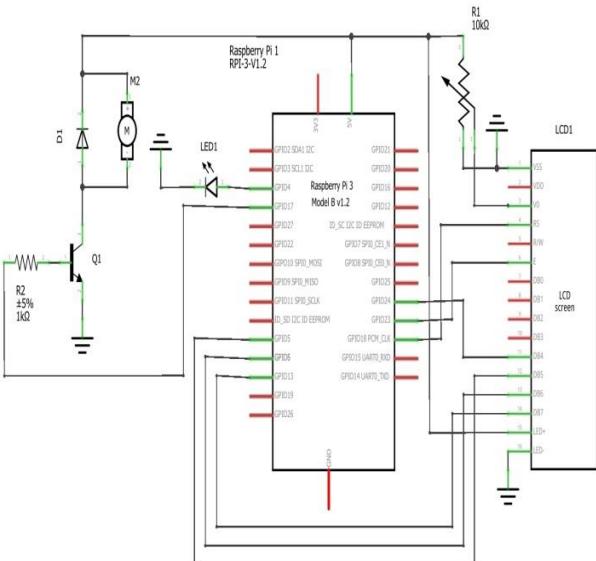


Fig. 3. Circuit diagram of FER application setup

The setup consists of an Arduino UNO microcontroller. The mood lighting effect is demonstrated using a blue LED. A brushed DC motor is used in this setup to demonstrate the speed control of the automobile. The detected emotions are displayed on an LCD display.

IV. TESTING AND RESULT

The classifier model is trained using JAFFE and CK+ database. A snapshot of the predicted label of four (out of the 20% of the database) of the validation data along with their images is shown in Fig. 4.

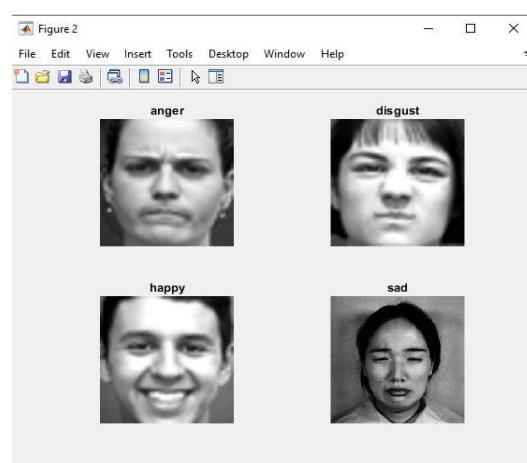


Fig. 4. Predicted labels of validation data

A confusion matrix that shows the accuracy of the predicted labels of the validation data is shown in Fig. 5. This shows that the proposed system has an accuracy of

Figure 3

		anger	disgust	fear	happy	neutral	sad	surprise
True Class	anger	48	1				1	1
	disgust		66					
fear			42					
happy				78	4			
neutral					31			
sad						49		
surprise	2							91

Fig. 5. Confusion matrix of True labels vs predicted labels of validation data

97.83% over the validation data that consists of the JAFFE and CK+ database images.

The proposed system is tested in laboratory conditions. The system showed lower accuracy over the test images. The bar chart in Fig. 6 shows the accuracy of the seven emotions. It is seen that the system shows accuracy rate of 50%, 40%, 20%, 40%, 20% and 50% in detecting the expressions anger, disgust, happy, neutral, and sad and surprise respectively, whereas the system was unable to detect fear expression.

The advantages of the FER system proposed are that, it has an accuracy of 97.83% in lab-controlled conditions. This system is resistant to variant head poses. By tracking driver's emotions, this system can help prevent inattentiveness, road rages and other potential safety issues on road. The proposed system can also help improve various other human computer interactive systems.

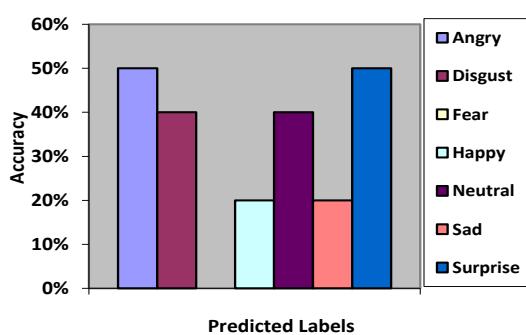


Fig. 6. Bar chart showing accuracy of predicted labels of test images

Facial emotion recognition has wide range of applications in the field of Human computer interactions. Augment reality.

virtual reality, Affective computing and Advanced driver assistance systems (ADSSS). FER can also be applied to different areas of website customization, gaming, healthcare and education. It has greater potential in the field of surveillance and military.

V. CONCLUSION AND FUTURE WORK

A Facial Emotion Recognition system to detect and classify facial emotions is developed. The classifier model consisted of a pre-trained neural network for feature extraction and SVM for emotion classification. The classifier model is trained on 1195 images of JAFFE and CK+ database. The face is detected using Viola Jones algorithm and the emotions are classified using the trained classifier model. The proposed model has an accuracy rate of 97.83% over validation data. The practical application of the detected emotions in affective computing are tested using a setup to control various applications that help in improving driver assistance systems, such as speed control and mood lighting.

The proposed model does not work well in low light conditions, hence this model can be improved to function properly in low light conditions. The proposed system also shows low accuracy rate for the detection of fear and neutral expression. This can be improved by training the images on a larger database. The accuracy can be increased by creating a customized database to train the model. Advanced hardware implementation of this model needs to be done.

REFERENCES

- [1] Ekman, P. & Keltner, D. (1997). "Universal facial expressions of emotion: An old controversy and new findings". Nonverbal communication where nature meets culture (pp. 27-46). Mahwah, NJ: Lawrence Erlbaum Associates.
- [2] Michael Revina, W.R. Sam Emmanuel, "A Survey on Human Face Expression Recognition Techniques" Journal of King Saud University – Computer and Information Sciences, 2018.
- [3] Mrs. Jyothi S Nayak, Preeti G, Manisha Vatsa, Manisha Reddy Kadiri, Samiksha S "Facial Expression Recognition: A Literature Survey" International Journal of Computer Trends and Technology (IJCTT) Volume-48 Number-1 2017.
- [4] Young Hoon Joo, "Emotion Detection Algorithm Using Frontal Face Image" ICCAS2005 June2-5, KINTEX, Gyeonggi-Do, Korea.
- [5] Jing Huang, Yunyi Shang and Hai Chen "Improved Viola Jones face detection algorithm based on Holo Lens" EURASIP Journal on Image and Video Processing, 2019.
- [6] Tanner Gilligan, Baris Akis "Emotion AI, Real-Time Emotion Detection using CNN" Stanford University, 2016. [Courtesy https://web.stanford.edu/class/cs231a/prev_projects_2016/emotion-ai-real.pdf]
- [7] David Fernando Tobón Orozco, Christopher Lee, Yevgeny Arabadzhi, Dr. Vijay Kumar Gupta "Transfer learning for Facial Expression Recognition", semantics scholar, 2018.
- [8] Zheng-Wu Yuana, Jun Zhang "Feature Extraction and Image Retrieval Based on AlexNet", Eighth International Conference on Digital Image Processing (ICDIP), 2016.
- [9] Yichuan Tang "Deep Learning using Linear Support Vector Machines", University of Toronto 2013. [Courtesy <https://arxiv.org/abs/1306.0239>]
- [10] Maram.G Alaslani and Lamiaa A. Elrefaei "Convolutional Neural Network based Feature Extraction for Iris Recognition "International Journal of Computer Science & Information Technology (IJCSIT) Vol. 10, No. 2, April 2018.