



Machine Learning Based Method to Design a Facial Emotion Detection and Chat Bot System

Tanisha Majumder¹, Kajal², Vijay B³, Santhosh S R⁴, Vinay Kumar M⁵

^{1,2,3,4}UG Student School of Computing and Information Technology.

⁵School of Computing and Information Technology

REVA Institute of Technology and Management

Bangalore, India.

Email: tanishamajumder80@gmail.com¹, kajalsandal006@gmail.com², appuv098@gmail.com³, santhosh7259951@gmail.com⁴, vinaymkv@gmail.com⁵

*Corresponding author's E-mail: tanishamajumder80@gmail.com

Article History	Abstract
Received: 06 June 2023 Revised: 05 Sept 2023 Accepted: 30 Nov 2023 CC License CC-BY-NC-SA 4.0	One of the active areas of research trends is recognizing emotions in images. This project aims to identify facial emotions. The research concept in Emotion Recognition is included in the flow of our emotion recognition. These involve image acquisition, image pre-processing, face detection, feature extraction, and classification, with the machine being applied after the emotion have been classified. Our framework relies on already-existing still images. This project aims to improve automated facial emotion recognition and build interaction between the system and the user (bot). Keywords: Face detection, Face recognition, OpenCV, Bot

1. Introduction

Human-computer interaction and human-mediated communication have become an important part of human life, even though we somehow lack the basic features to identify and respond to facial expressions of perspectives, feelings, and consciousness, that we underestimate in human thoughts and communication. So, in real-life applications, a very challenging task is the detection of human emotion. Facial expression recognition systems require overcoming the human face having multiple variabilities such as expression by displaying the particular emoji. In our framework we are taking the pre-existing still images as well as live images then the model will scan and detect the particular expression such as happy, sad, angry, surprise, and neutral. We are taking few expressions to recognize the emotions and establishing a communication between the system and the user. Face detection is a stage that helps to identify the face area in input images or sequences. The next step is to retrieve essential information from the machine's facial expressions once the face has been located. Facial expression recognition recognizes and identifies the feeling of the human for example. Happy, Angry, Sad, Neutral, and Surprise. The main aim of this model is to identify facial emotion expression and to engage the user in an interactive session which make them feel better. Emotion analysis can be beneficial in a lot of circumstances. Human-computer interaction is one such area. Computers can make better decisions to assist users with benefit from emotion recognition. We are particularly interested in assessing research efforts that analyze emotions based on text and speech. The paper is organized as Section II deals with the background and related work. Section III deals with methods such as Convolutional Neural Network Section IV, various algorithms are used for training and detecting. Finally, Section VI concludes our work that our method scales an accuracy of 63.02%.

Background and Related Work

The problem is well-known and many have already studied this problem. R. S. Deshmukh [1] In this paper, For emotion classification, a machine learning algorithm is used. They developed an API that categorizes emotion information given to the system, and the Viola-Jones algorithm is used to implement face feature detection for better results. According to the system, it classifies the emotion based on the individuals, and then with the facial expression, the songs were played.

A. A. Varghese [6] In their paper A framework for the classification of emotional states was generated, which was supported by still depictions of the face. The process comprises training an active appearance model (AAM) on face images from a public access repository to describe forms and texture variation, which are relevant for expression recognition. It can help with the development of initial face search algorithms as well as the extraction of important data from each shape and texture. The ANN predicts the user's emotion and can enhance its results as a response. It utilizes facial expressions, speech, and multi-mode data to determine emotion recognition. R. Pathar [8] In this paper The FER2013 data set was used in the facial expression recognition competition at the ICML 2013 workshop. The dataset comprises of 35,887 very much organized grayscale pictures with a resolution of 48x48 pixels. They built Shallow and Deep Convolutional Neural to obtain expression recognition results. Our network takes the image as input and assesses how it makes us feel. The seven emotions are anger, happiness, fear, sorrow, disgust, surprise, and neutral. By comparing accuracy at various depths, a 58 percent precision was attained. K. Lekdioui [9] FER classifies static images and CNNs without the need for preprocessing feature extraction tasks. The illustrations of this paper are based on colors like green square for face detection and yellow for eyes, nose, and mouth. They used horizontal mirroring and uniformly cropped image data for training data enlargement, and the accuracy was 61%. We have to select either speech or text option from the interface. Once the speech is selected, the emotion starts detecting, we propose a bot to establish a communication between the user and the system. According to the emotion detected such as sad where the bot interacts and asks the user reason behind the emotion. To cheer up the user, a song is played by the bot. Similarly, when the text is selected emotion starts detecting then the bot asks the user to chat and interact with it.

Dataset

FER2013 in Kaggle was utilized for this study, and it was initially provided as part of the ICML 2013 workshop's facial expression recognition challenge [9]. The dataset consists of about 24,282 well-structured grayscale pictures with a resolution of 48x48 pixels. The entire dataset is structured in an unlabelled manner. There are five different emotions represented in the dataset: happy, sad, angry, surprise, and neutral. The FER2013 dataset is accessible in Kaggle as a single file. For training/testing, I transformed the images into a PNG dataset and provided it as the dataset.

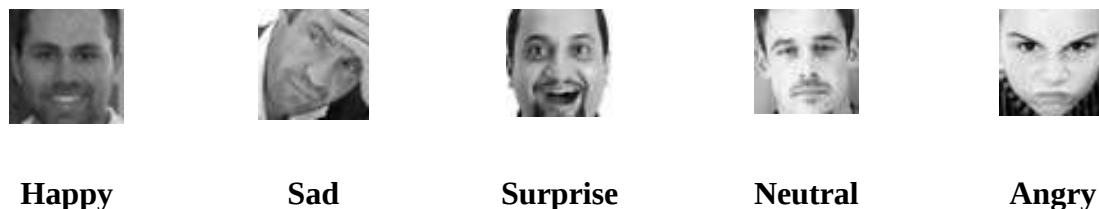


Figure 1: each emotion group is represented by a single example

2. Materials And Methods

It will capture live images then using open cv it retrieves the frames which are then retrieved. Using a detector detects the emotions then the system responds to the emotions which are detected like happy, sad, angry, etc. Then the input and output parts are for the communication of the bot. So, when the emotion is detected, the bot takes the emotion as input and then it uses the speech-to-text method through the server or system to respond to the user using predefined dialogues like hey user, why are you sad? It's kind of interacting with the user through the model. Then by using GTTS (Google Text to Speech) it interacts with the user and delivers the type of emotion as an output.

Dataset

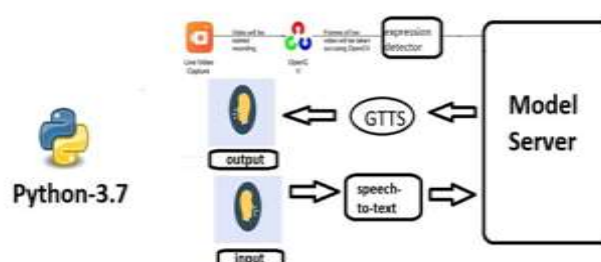


Fig 2: Architecture

Face Algorithms

The following algorithms are available:

- A. *Haarcascade_frontalface_default.xml*
- B. *Convolutional Neural Network*

A. Face Detection using Haar Cascade

A haar cascade is a sequence of "square-shaped" capacities printed from either a set of wavelets or a foundation. It's also aimed at "Haar Wavelets," which use a hair rippling approach driven by Paul Viola and Michael Jones' paper to group pixels on an image into squares using To divide pixels on an image into squares, "rapid object recognition victimization with such an extended cascade of simple features" is used. [11]. It's a computer-assisted learning process in which a cascade of positive and negative representations is generated.

This method is used to detect faces in each frame of the webcam feed. The face-containing region of the frame is resized to 48x48 pixels and fed into the CNN. The network outputs a list of softmax scores for the five classes of emotions. On the screen, the emotions with the best performance are shown.

B. Training the expressions

The data of five expressions from the dataset is used. The training set was fed into the CNN architecture, which was then trained using Python parameters. For obtaining the prediction effects of five articulations, the architecture is divided into an input layer, three layers of convolution and pooling, and fully connected layers. A 2*2 convolution kernel is used. in the first layer of convolution. Each layer's pooling is done using maximum pooling. The parameters of the training result are stored for subsequent use after a sufficient number of training iterations of the above architecture.

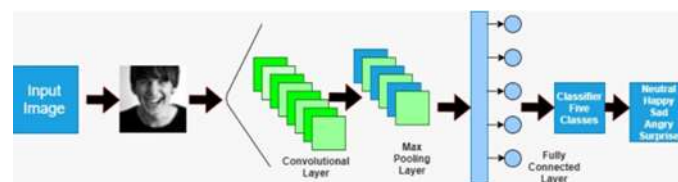


Fig 3: CNN Architecture

3. Results and Discussion

The above experiments results show that the methods of detecting emotions accurately.

In other words, the effectiveness of each way is the same if a static image is used to detect faces.

A. Result of implementation:

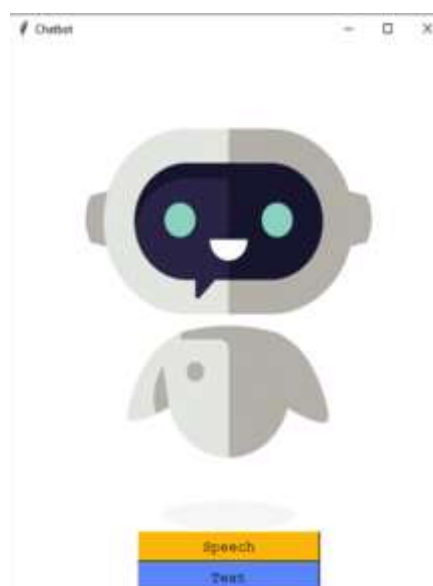


Fig 4: Interface



Fig 5: By selecting the speech option, the user can capture real-time face detection via webcam.

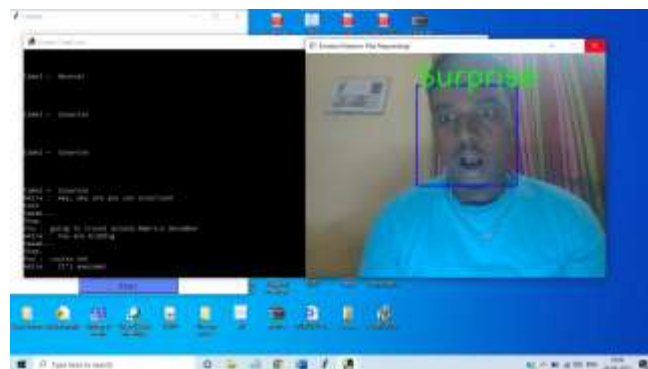


Fig 6: By selecting the speech option, the user can capture real-time face detection via webcam.

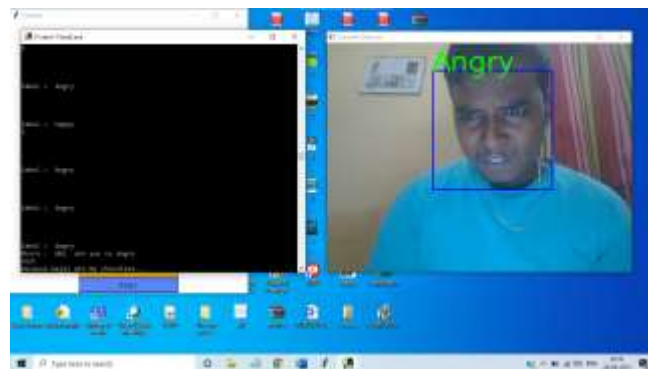


Fig 7: By selecting the text option, user can capture real-time face detection via webcam.

A. Accuracy

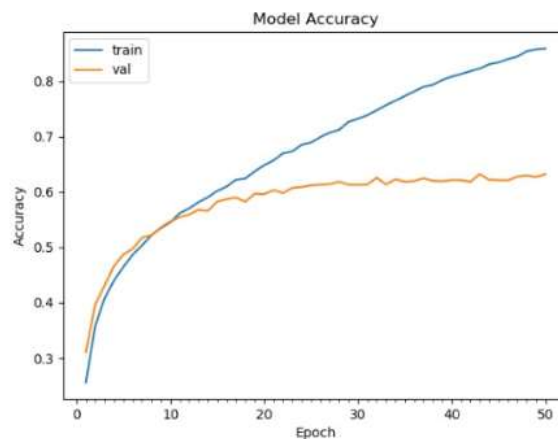
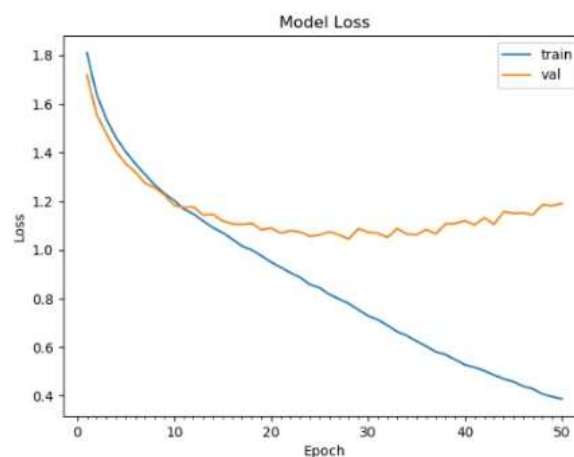


Fig 8: Accuracy of both train and test



results

Fig 9: Loss of data during training and testing

The test precision reached 63.2% in 50 epochs.

4. Conclusion

Models must be able to understand and respond to human speech as well as perform signs of feelings and mental states. Besides, facial expressions are important. Specifically, the expression refers to a wide range of emotions and mental conditions in the brain that is not part of the basic feelings set. This is a challenging task because automating the analysis of facial expressions in images is difficult. The aim of this work, facial expressions, is to recognize emotions in human facial expressions from pre-existing images. The proposed system is to use a bot to interact with the user to make them feel better.

Acknowledgement

We would like to express our special thanks to our Guide, Professor Vinay Kumar M, REVA University, for continuous support and guidance. We would also like to extend our sincere thanks to Dr. Sunil Kumar Manvi, Director, REVA University, for providing us with all the support.

References:

- [1] R. S. Deshmukh, V. Jagtap and S. Paygude, "Facial emotion recognition system through machine learning approach," in *proc of 2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, pp. 272-277, 2017.
- [2] M. Swartz and A. Crooks, "Comparison of Emoji Use in Names, Profiles, and Tweets," in *proc of 2020 IEEE 14th International Conference on Semantic Computing (ICSC)*, San Diego, CA, USA, pp. 375-380, 2020.

- [3] S. Cappello, S. Svetlichnaya, P. Garrigues, T. Mensink and C. G. M. Snoek, "New Modality: Emoji Challenges in Prediction, Anticipation, and Retrieval," in *proc of IEEE Transactions on Multimedia*, vol. 21, no. 2, pp. 402-415, Feb. 2019.
- [4] A. Illendula, K. V. Manohar and M. R. Yedulla, "Which Emoji Talks Best for My Picture?" in *proc of 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, Santiago, pp. 514-519, 2018.
- [5] M. Healy, R. Donovan, P. Walsh, and H. Zheng, "A Machine Learning Emotion Detection Platform to Support Affective Well Being," in *proc of 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Madrid, Spain, pp. 2694-2700, 2018.
- [6] A. A. Varghese, J. P. Cherian and J. J. Kizhakkethottam, "Overview on emotion recognition system," in *proc of 2018 International Conference on Soft-Computing and Networks Security (ICSNS)*, Coimbatore, pp. 1-5, 2018.
- [7] M. Xiaoxi, L. Weisi, H. Dongyan, D. Minghui and H. Li, "Facial emotion recognition," in *proc of 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP)*, Singapore, pp. 77-81, 2019.
- [8] R. Pathar, A. Adivarekar, A. Mishra and A. Deshmukh, "Human Emotion Recognition using Convolutional Neural Network in Real-Time," in *proc of 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT)*, CHENNAI, India, pp. 1-7, 2019.
- [9] S. Singh and F. Nasoz, "Facial Expression Recognition with Convolutional Neural Networks," 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, pp. 0324-0328, 2020.
- [10] A. V. Nefian and M. H. Hayes, "Hidden Markov models for face recognition," *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98* (Cat. No.98CH36181), Seattle, WA, USA, 1998.
- [11] K. Lekdioui, Y. Ruichek, R. Messoussi, Y. Chaabi, and R. Touahni, "Facial expression recognition using face-regions," in *proc of 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, Fez, pp. 1-6, 2019.
- [12] A. Raheel, M. Majid, and S. M. Anwar, "Facial Expression Recognition based on Electroencephalography," in *proc of 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, Sukkur, Pakistan, pp. 1-5, 2019.
- [13] Paul Viola, Matthew Jones Conference paper- IEEE Computer Society Conference on Computer Vision and Pattern Recognition. "Rapid object detection using a boosted cascade of simple features."