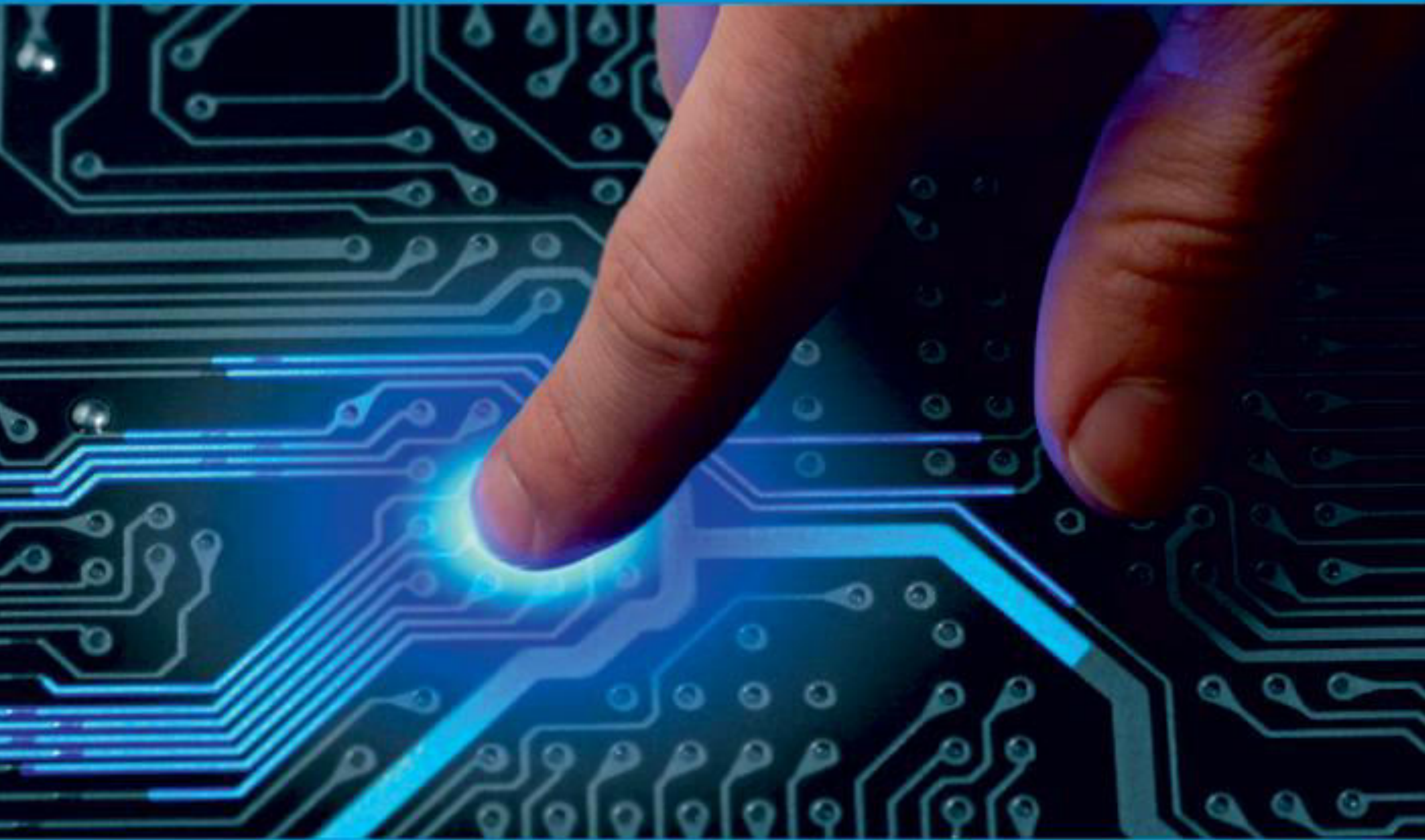
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 11, November 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 11, November 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1111030 |

Survey Paper On: "From Faces to Emojis: A Deep Learning Approach to Convert Facial Expressions into Emoticons"

Prof. Balaji Chaugule, Bhakti Sambal, Janakee Shelke, Aachal Chimankar, Unnati Shinde

Dept. of Information Technology, Zeal College of Engineering and Research, Narhe, Pune, India

ABSTRACT*:* *Emojis* are an inevitable records nsmg throughout the remaining years, from marketing, virtual verbal exchange in particular, to Recovery of statistics associated with sentiment evaluation and Viewpoint mining. Emoji allow people to specific emotions and their identities greater "authentically" via way of means of growing semantic exceptional of visible messages. Emojis also are utilized in remarks forms. The remark shape fee having emojis is more than different strategies for remarks. The emotions represented via way of means of the textual content or its severity are modified via way of means of emojis. Indeed, via means of simulating facial gestures, emojis may be utilized in Informal Text Communication to specific emotion including sarcasm, irony or non-textual humor. Emoji lets in customers to make a selection out from extensive lists, is one manner to show nonverbal signs. Emotional popularity the use of facial features via emoji in actual time is explored on these studies' thesis. Moreover, in addition develops the standards of facial features evaluation and actual-time belief of facial emotion popularity. The created software carries six human expressions that encompasses feelings which might be happy, angry, sad, surprise, neutral, fearful, disgusted. The actual expressions which might be being expressed are those expressions transmitted via means of human beings. Because in there potential to higher talk emotional responses and the manner they sell touch among people, the investigations of such speech are important. The output of the task suggests the emoji with the respective face emotion

KEYWORDS*:* CNN, Gestures, image to Emotion, Emoji, HAAR

I.INTRODUCTION

Emojify using deep learning is a project that aims to recognize and classify facial expressions in real-time and associate them with appropriate emoji representations. It utilizes computer vision techniques and deep learning algorithms to analyze facial images captured from a webcam or video stream and predicts the corresponding emotion. The project provides an interactive and engaging user experience by overlaying the detected emotion with a relevant emoji on the user's face. The underlying deep learning model is built using a convolutional neural network (CNN) architecture. CNN is trained on a large dataset of facial images labeled with different emotion categories such as anger, disgust, fear, happiness, neutrality, sadness, and surprise. During training, the model learns to extract meaningful features from the input images and classify them into the appropriate emotion categories. The Emojify project also incorporates various components to enhance the user experience. It includes a graphical user interface (GUI) that displays the captured video feed along with the detected emotion and corresponding emoji overlay on the user's face in real-time. The GUI allows users to interact with the application, capture snapshots, generate reports, and visualize emotion statistics. Additionally, the project provides functionalities to generate reports and analyze the distribution of detected emotions. Users can capture snapshots of their emotions, which are saved along with relevant information such as the username, date, and time. The project also generates reports containing this information for further analysis or documentation purposes. Furthermore, users can visualize the distribution of detected emotions through pie charts or bar plots, enabling them to gain insights into their emotional expressions over time. Overall, the Emojify using deep learning project combines computer vision, deep learning, and interactive user interface components to provide a fun and insightful way of recognizing and representing emotions through emojis. This project allows users to interact with the application, capture snapshots, generate reports, and visualize emotion statistics. Additionally, the project provides functionalities to generate reports and analyze the distribution of detected emotions. Users can capture snapshots of their emotions, which are saved along with relevant information such as the username, date, and time. The project also generates reports containing this information for further analysis or documentation purposes. Furthermore, users can visualize the distribution of detected emotions through pie charts or bar plots, enabling them to gain insights into their emotional expressions over time. It showcases the potential of deep learning models in real-time emotion recognition applications and

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 11, November 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1111030 |

offers a practical application for understanding and analyzing human expressions.

II. LITERATURE SURVEY

[l].Deep Leaming Models for Facial Expression Recognition: Atul Sajjanhar, ZhaoQi Wu, Quan Wen, 2018, IEEE. In this paper, they train and test a CNN model for facial expression recognition and evaluate the performance of VGG which are pre-trained for object recognition, and compare these with VGG-Face which is pre-trained for face recognition. But the limitation of their project was the model was able to depict few expressions in total and wasn't flexible.

[2].Facial Expression Recognition via Deep Leaming: Abir Fathallah, Lotti Abdi, Ali Douik, March 2018, IEEE. In this paper an effective approach system based on ConvNets for facial expression recognition. We proposed a new architecture in which the input of the system is an image; then, they use CNN to predict the facial expression label which should be one of these labels[9]: anger,happiness, sadness, disgust, surprise and neutral. But their model is not able to capture the screenshots along with time, date, path of screenshot

[3].Facial Expressions detection using convolutional neural network and SoftMax function on captured images: Ashish Deopa, Abhishek Sinha, Aditya Prakash, Rupesh Sinha, 2019, IEEE. In their study they used Convolutional neural network for the image recognition using FER2013 dataset from Kaggle that consist of 35,887 faces and can capture 7 emotions.But the limitation of their project is this project does not include and provide any social media operations to the user for a better and advance use

[4].EMOJIFY-CREATE YOUR OWN EMOJIS WITH DEEP LEARNING: Sagar Chilivery, Sandeep Pukale, Yashraj Sonawane, February 2022, IRJETS. This paper is primarily based totally on a gadget which implements Convolutional Neural Network and Fer2013 Dataset to hit upon feelings from facial expressions and change them to personalized emojis. But the limitation of the project is that the model is not able to capture the screenshot and also show emotion statistics.

[5].EMOJI CLASSIFICATION USING CNN: Pradeep Bharati, Omkar Singh, April 2023, IJCRT. In their study they first train the machine learning model with TensorFlow to detect face expressions like happy, sad, fearful and so one. Then map those expressions to emoji's in real time. But the limitation is that their paper does not provide any additional features that may enhance the user experience such as 'Like' and 'share.

III. PROPOSED SYSTEM

The idea of the proposed system is to detect the face after which the image can be processed using HAAR cascade for facial feature extraction. SVM Classifier is then used to categorize the emotions into its seven distinct types. Using HAAR of OpenCV package, the corresponding emojis of the emotions can get superimposed over the subjects' faces. In any camera module of any leading social networking apps, the use of APis can reduce the processing time for face detection for which they have their in-built face detection algorithm which can detect the face smoothly and followed by which the emoticons can be implemented over the faces as filters. Once the facial expression is recognized, the corresponding emoji is displayed on the screen, providing a visual representation of the detected emotion. It includes a feature to capture a screenshot of the application and generate a report containing information such as the username, date, time, and the path to the captured screenshot This model also generates statistics about the detected emotions. We also provide visualizations such as bar plots and pie charts to represent the frequency and distribution of different emotions detected during the Emojify session

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 11, November 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1111030 |

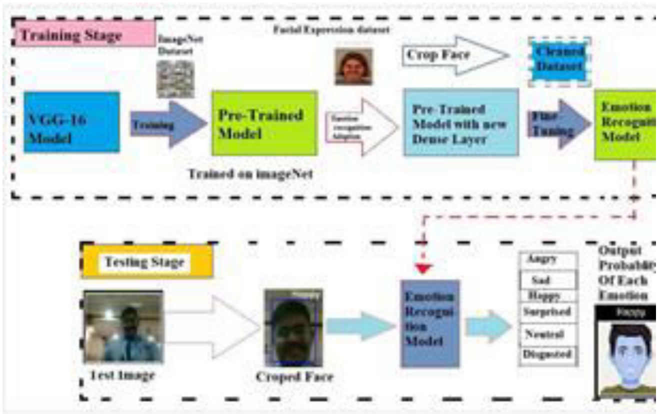


Fig 1. Architecture Diagram[2]

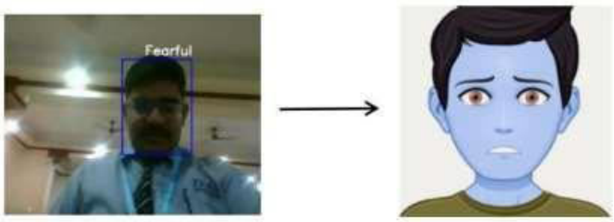


Fig 2.Transformation of Facial Expression into Corresponding Emoji

IV. METHODOLOGY

A. Convolutional Neural Network:

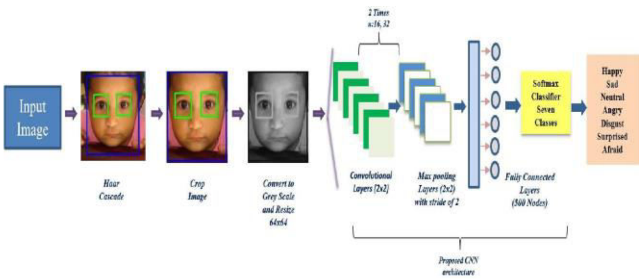


Fig 3. Architectural Diagram of Facial Generation using CNN

A typical architecture of a convolutional neural network contains an input layer, some convolutional layers, some fully-connected layers, and an output layer. It has 6 layers without considering input and output. The architecture of the Convolution Neural Network used in the project is shown in the above figure[3].

1. Input Layer: The input layer has predetermined, fixed dimensions, so the image must be pre-processed before it can be fed into the layer. Normalized gray scale images of size 48
X 48 pixels from Fer2013 dataset are used for training, validation and testing. For testing proposed laptop webcam images are also used, in which the face is detected and cropped using OpenCV Haar Cascade Classifier and normalized.
2. Convolution and Pooling (ConvPool) Layers: Convolution and pooling is done based on batch processing. Each batch has N images and CNN filter weights are updated on those batches. Each convolution layer takes image batch input of four dimensions N x Color-Channel x width x height. Feature maps or filters for convolution are also four dimensional (Number of feature maps in, number of feature maps out, filter width, filter height). In each convolution layer, four dimensional convolution is calculated between image batch and feature maps. After convolution only the parameter that changes is image width and height.
New image width = old image width - filter width +1 New image height= old image height - filter height+ 1

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 11, November 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1111030 |

After each convolution layer downsampling / subsampling is done for dimensionality reduction. This process is called Pooling. Max pooling and Average Pooling are two famous pooling methods. In this project max pooling is done after convolution. Pool size of (2 x 2) is taken, which splits the image into a grid of blocks each of size (2 x 2) and takes a maximum of 4 pixels. After pooling only height and width are affected. Two convolution layers and a pooling layer are used in the architecture. The first convolution layer size of the input image batch is (N x I x 48 x 48). Here, the size of the image batch is N, number of color channels is I and both image height and width are 48 pixels. Convolution with a feature map of (1 x 20 x 5 x 5) results in an image batch of size (N x 20 x 44 x 44). After convolution pooling is done with a pool size of (2 x 2), which results in an image batch of size (N x 20 x 22 x 22). This is followed by a second convolution layer with a feature map of 20x20x5x5, which results in an image batch of size (N x 20 x 18 x 18). This is followed by a pooling layer with pool size (2 x 2), which results in an image batch of size (N x 20 x 9 x 9).

3. Fully Connected Layer: It takes a large number of input features and transforms features through layers connected with trainable weights. Two hidden layers of size 500 and 300 units are used in fully-connected layers. The weights of these layers are trained by forward propagation of training data then backward propagation of its errors. Back propagation starts from evaluating the difference between prediction and true value, and back calculates the weight adjustment needed to every layer before. We can control the training speed and the complexity of the architecture by tuning the hyper-parameters, such as learning rate and network density. Hyper-parameters for this layer include learning rate, momentum, regularization parameter, and decay. The output from the second pooling layer is of size Nx20x9x9 and input of the first hidden layer of fully-connected layer is of size Nx500. So, the output of the pooling layer is flattened to Nx1620 size and fed to the first hidden layer. Output from the first hidden layer is fed to the second hidden layer. Second hidden layer is of size Nx300 and its output is fed to the output layer of size equal to the number of facial expression classes.

4. Output Layer: Output from the second hidden layer is connected to the output layer having seven distinct classes. Using Softmax activation function, output is obtained using the probabilities for each of the seven classes. The class with the highest probability is the predicted class.

B. Haar Cascade:

The image that is supplied by the API is then provided to the HAAR cascade in which some dataset has been given for training the data. For the development of a working model, we will use the Fer2013 dataset. HAAR-Like features have high accuracy to detect faces from different angles

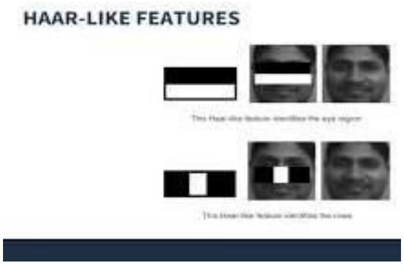


Fig 4. HAAR-Like feature for face detection

V. CONCLUSION

As Today's technology humans are loving the fashion of speaking with non-verbal cues like emoticons so we have no idea why we no longer deliver our personal emojis. With improvements in laptop imaginative prescient and deep learning, we are now capable of hitting upon human feelings from pies. In this deep learning task, we are able to classify human facial expressions to clear out and map corresponding emojis or avatars. The end result we're anticipating is the usage of emojify in the chatting world. We need humans to talk with their personal customisable emoji. The task will apprehend one's present emotion and convert that emotion's emoji in order that the consumer receives emoji of their face and use it in chatting

REFERENCES

[1]. Deep Leaming Models for Facial Expression Recognition Atul Sajjanhar, ZhaoQi Wu, Quan Wen, in 2018, IEEE

[2]. Facial Expression Recognition via Deep Leaming: Abir Fathallah, Lotfi Abdi, Ali Douik, in March 2018, IEEE

[3].Facial Expressions detection using convolutional neural network and SoftMax function on captured images

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 || A Monthly Peer Reviewed & Referred Journal |

|| Volume 11, Issue 11, November 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1111030 |

presented by Ashish Deopa, Abhishek Sinha, Aditya Prakash, Rupesh Sinha in 2019, IEEE.

[4].EMOJIFY-CREATE YOUR OWN EMOJIS WITH DEEP LEARNING Sagar Chilivery, Sandeep Pukale, Yashraj Sonawane in February 2022, IRJETS

[5].EMOJI CLASSIFICATION USING CNN; presented by Pradeep Bharati, Omkar Singh in April 2023, IJCRT

[6].D. H. Kim, W. J. Baddar, **J.** Jg, etY. M. Ro,« Multi-Objective Based Spatio-Temporal Feature Representation Learning Robust to Expression Intensity Variations for Facial Expression Recognition», IEEE Trans. Affect. Comput, 2019

[7].C. Marechal et al., « Survey on Al-Based Multimodal Methods for Emotion Detection », in High-Performance Modelling and Simulation for Big Data Applications: Selected Results of the COST Action IC1406 cHiPSet, J. Kolodziej et H.Gonzalez-Velez, Ed. Cham: Springer International Publishing, 2019, p. 307-324

[8].Facial emotion recognition using deep learning: review and insights; presented by Wafa Mellouka, Wahida Handouzi, August 2020, ScienceDirect.

[9].C. Xu, "Analysis of Emoji Generation Based on DCGAN Model," 2021 2nd International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), 2021,pp. 130-134, doi: 10.1109/ICHCI54629.2021.00034.

[10].Emotion Detection through Facial Expression using Deep Learning presented by Manish Kumar, Swati Srivastava in February 2022, IEEE.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING