Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
2326-9865

Emoji Generation Using Facial Emotion Classifier

Ramya. G. Franklin[1] , Dr. A.C. Santha Sheela [2]*,

[1] Assistant Professor, [2] Associate Professor, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

[1] mikella.prabu@gmail.com, [2] shantha.jasmine@gmail.com

Article Info
Page Number: 142-148
Publication Issue:
Vol 71 No. 3s2 (2022)

Article History
Article Received: 28 April 2022
Revised: 15 May 2022
Accepted: 20 June 2022
Publication: 21 July 2022

Abstract. Face expressions are an intrinsic aspect of nonverbal communication and play a significant role in Human Computer Interaction. The formation of facial emoticons is a human-computer interaction device. Emoji generation in real time based on a person's facial expression has always been difficult. Human social conversations depend on facial expressions. Since the world is becoming more technologically sophisticated day in day out, there are more interactive encounters, such as text messages, than physical ones. Emoticons promote virtual social interaction by reducing the amount of words exchanged. This paper describes an Emoji generation methodology based on Facial Expression Recognition (FER) and Convolutional Neural Networks (CNN) coupled with Machine Learning and Deep Learning. This CNN-based model can be put to work to evaluate feelings as users watch movie trailers or video lessons, as well as to help people with autism regulate their emotions.

Keywords: Emoticons, Human-Computer Interaction, Facial Expression Recognition, Machine Learning, Deep learning.

1. Introduction

Emojis and avatars are visual representations of nonverbal signals. These signals have been ingrained in online chatting, product reviews, brand emotion, and a variety of other activities. It also encourages future data science exploration into emoji-driven storytelling. Human emotional contact relies heavily on facial gestures. The studied features are generalized by a facial expression classifier to distinguish various expressions from unknown faces. It is now possible to identify human emotions from images more accurately thanks to advances in computer vision and deep learning. We chose CNN because it has the ability to detect essential features without the need for human intervention, which finds significant traits without the need for human intervention.

The aim of this deep learning project is to effectively identify human facial expressions so that corresponding emoji or avatars can be filtered and mapped. The application can be used in large corporations or businesses to receive real-time consumer reviews. The application's findings can be used in further research and development.

2. Literature Survey

Ko et al proposed a brief review of the approaches followed in FER. The two main streams in which these approaches fall include conventional FER approaches and deep learning-based FER approaches. The former approach comprises of three steps namely, facial part detection, feature extraction and classification. The latter approach enables end-to-end learning from the input images directly. To keep track of the temporal features, a hybrid approach was proposed in which a Convolutional Neural Network (CNN) that represents spatial features is combined with Long Short Term Memory (LSTM) to represent temporal features of consecutive frames. LSTM possesses a chainlike structure that is designed to solve the dependency in long term using short term memory. LSTM supports both fixed and variable length input and output. Moreover, they support straightforward end-to-end fine tuning when they are integrated with other models such as CNN.

Ying et al used distance vectors as facial features for FER. The distance vectors are selected based on FAU, a component of Facial Action Coding System (FACS). To determine the AUs that were involved in facial expressions, statistical analysis was done. Based on that analysis, fourteen facial points were termed to be significant for FER. Then, it aimed to measure the intensity of facial expressions using 3D distance vectors that were computed from these facial points. All these salient facial points reside around eyes, brows and mouth region only. It revealed the fact that it is not necessary to rely on all the facial feature points for estimating the facial expression intensity. Their studies concluded that the mean and standard deviation of distance measures cannot identify which action units must be considered for sad expression. Many reviews are presented in literature by many researchers with respect to ecommerce applications in different domain [16][17][18]. This analysis will surely enable the researchers with the idea of deep learning technique in different applications [19][20][21][22][23]. Different issues also discussed in machine learning applications[24][25].

3. Proposed System

Human emotion expression is one of the key subjects in facial recognition outside of laboratory trials,

which can produce both technical and everyday applications. This project creates a deep learning model that can classify a given picture of a human facial expression into one of the seven essential human emotions. The classified emotion would then be mapped to an emoji or avatar.

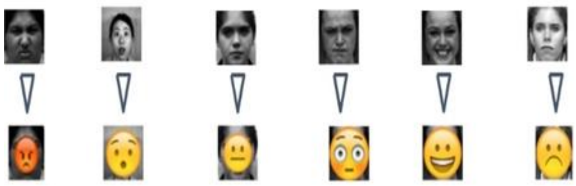

Fig. 1: The emoji are correctly imposed on their corresponding faces. The input will be raw image of the expression, and output will be shown as above

The major goal of this project is to improve on the baseline accuracy of a model that can identify seven basic emotions: joyful, sad, surprise, anger, disgust, neutral, and fear. Furthermore, our project aims to analyze the outcomes of our model in terms of accuracy for each class. The model is expected to perform wild emotion recognition in the future, with more complex state variance than lab condition photos.

Throughout this paper, we use CNN to develop a robust model for generating emoticons. For the goals of our project, the essential problem is to focus on the classification of 7 basic emotions shown at below:

Fig. 2 Classification of 7 basic emotions

We construct a deep neural classifier to identify facial expression and generate emoticons for this. We use the Haar cascade classifier for face detection before using the deep neural classifier, and then the neural classifier for expression classification. Facial expression recognition employs the deep CNN process, as well as the Haar Cascade face classifier and computer vision techniques. To remove higher-level functions, CNN convolve with the original images or function maps using a convolutional kernel.

4. System Architecture

The system architecture is as follows: our model recognizes faces and classifies facial expressions using an open-source computer vision library. It is a library of programming functions primarily geared at real-time computer vision. For each classified expression there will be a generated emoticon as the output in the application.

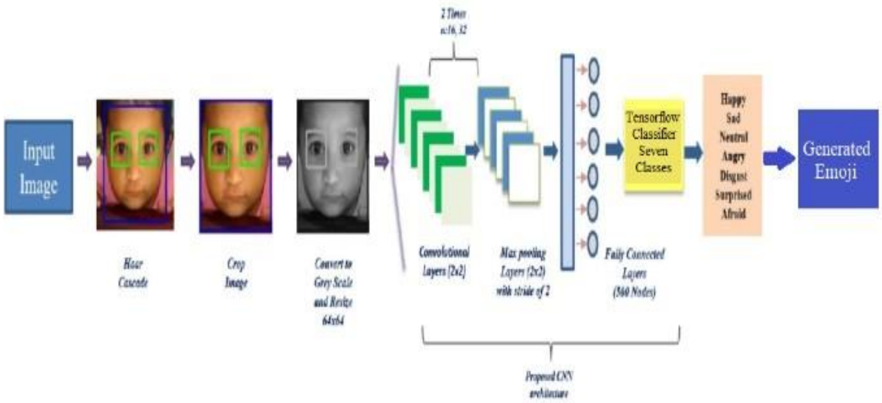

Fig. 3 System Architecture

The implementation process is divided into four modules which simplify the understanding of our problem statement. These are as follows:

Data Pre processing

Let's talk about the datasets until we get into the preprocessing operations. The datasets has a significant impact on model efficiency.

FER2013 is the name of the datasets we used in our article, and it is available on Kaggle. It contains 35,685 grayscale images of faces with a resolution of 48x48 pixels. Happy, neutral, sad, anger, disappointment, disgust, and fear are the emotions displayed in the facial expressions that are classified as happy, neutral, sad, angry, surprise, disgust, and fear.

One of the preprocessing procedures that aids in visualizing the total amount of photos is data visualization. Data augmentation is another level of preparation that entails rotating and flipping each image in our datasets. More photographs are added to our datasets as a result of this. It is being achieved in order to broaden the range of data that can be used to train models. After that, we normalize the data before saving and targeting it.

4.2 Training the CNN

Before we undergo training, we must first build the CNN, which consists of several layers Conv2D, MaxPooling2D, Flatten, Dropout, and Dense are just a few examples. Keras library's sequential API helps us to incrementally build a new layer for our model. The Conv2D is the initial layer, with 200 filters and a filter size or kernel size of 3x3. The 'Relu' activation function is employed in the initial layer. This Relu function, which originally stood for Rectified Linear Unit, outputs the user input if it is positive, and zero otherwise. The maxpooling2d has been utilized in the second layer, with a pool capacity of 2x2. The succeeding layer is another Conv2D layer with 100 filters of the size (3x3) and the 'Relu' activation feature. Following the Conv2D layer is a MaxPooling2D layer with a pool size of 2x2.

Deep learning neural networks are trained using the gradient descent optimization algorithm. As part of the optimization procedure, the error for the current state of the model must be determined repeatedly. This entails choosing an error function, also known as a loss function, to estimate the model's loss and modifying the weights to minimize the loss on the next evaluation.

Categorical cross-entropy, often known as Softmax Loss, is a loss function used in multi-class classification applications. These are jobs in which an example can only be classified into one of several options, and the model must choose one. The categorical cross-entropy loss function sums

$$Loss = - \sum_{i=1}^{\substack{Output \\ size}} y_i \cdot \log \hat{y}_i$$

the following values to compute an example's loss:
where yi corresponds to the goal value,

$\hat{y}_i$ output size is the number of scalar values in the model output, and is the i-th scalar value in the model output. This loss is a great way to tell how different two discrete probability distributions are from one another. We have seven categories: ‘Angry,' ‘Disgusted,' ‘Fearful,' ‘Happy,' ‘Sad,' ‘Neutral,' and ‘Surprised,' so, for stochastic gradient descent, we utilize Categorical cross entropy as the loss function and Adam as the optimizer.

Finally, the CNN model is trained for 100 epochs with seven classes, on denoting the class of images with different facial expression which will be then converted into emoticons.

Detecting face and Generating Emoticon

In this Phase the trained data model has been loaded. Face detection must be implemented prior to emoticon generation. For detecting facial features, we will use the Haar Feature-based Cascade Classifier. Thousands of photos are used to train this OpenCV Cascade Classifier to detect the frontal face.

We label seven probabilities, '0-6' for different face expressions, and using the RGB, you may also change the color of the boundary rectangle. value blue which shows the the detected face. In the final phase, we utilise the OpenCV package to run an endless loop with our web camera, where we use the Cascade Classifier to detect the face. The use of a webcam is indicated by the code. The label will be picked and displayed above the detected faces in white characters based on which likelihood is higher.

5. **Result and Discussion**

The Python language was used to run the model in an Anaconda environment. It makes use of open-source libraries such as pandas and NumPy, as well as classification techniques such as Tensorflow and Keras.
The experiment was carried out on an Intel i5-7200U processor running at 2.50GHz with 8GB of RAM

```
y: 0.6295
Epoch 100/100
448/448 [==============================] - 142s 318ms/step - loss: 0.1400 - accuracy: 0.9534 - val_loss: 1.5770 - val_accurac
y: 0.6295
```

For 100 epochs, the model was conditioned. However, unless over-fitting happens, we will train for further epochs to improve precision. We will see that our model has 95.3 percent accuracy after the 100th epoch. This indicates that it has been well trained and is not over-fitted.

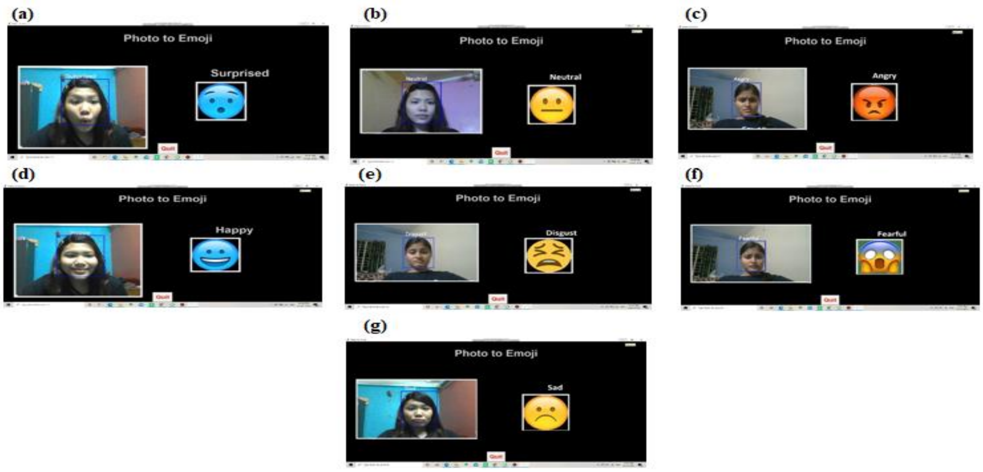

Fig. 4(a)-(f) Superised, Neutral, Angry, Happy, Disgust, Fearful and Sad faces of Emoji Generation

The model was able to generate emoji from the expression detected on the faces. If the expression is happy it will generate happy face emoticon, if it is angry it will give angry face emoticon and similarly to other expression also.

6. Conclusion

In this research, we offer a method for creating emoticons that uses Convolutional Neural Networks and Deep Learning algorithms to recognize face expressions. This paper demonstrates an improved FER model with 95 percent accuracy, and the produced outcome correlates to the person's facial expression. It demonstrates how the model can work in real-time by capturing facial expressions through the webcam and also by using images as input data.

References

[1] Alshamsi, Humaid, Veton Kepuska, and Hongying Meng. "Real time automated facial expression recognition app development on smart phones." In 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), pp. 384-392. IEEE, 2017.
[2] Ekman .P & Keltner, D Universal facial expressions of emotion: An old controversy and new findings. In U. C. Segerstråle & P. Molnár (Eds.), Nonverbal communication: Where nature meets culture (pp. 27–46). Lawrence Erlbaum Associates, Inc. 1997.
[3] Fathallah, Abir, Lotfi Abdi, and Ali Douik. "Facial expression recognition via deep learning." In 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), pp. 745-750. IEEE, 2017.
[4] Goodfellow, Ian J., Yaroslav Bulatov, Julian Ibarz, Sacha Arnoud, and Vinay Shet. "Multi-digit number recognition from street view imagery using deep convolutional neural networks." arXiv preprint arXiv:1312.6082, 2013.
[5] McDuff, D., Mahmoud, A., Mavadati, M., Amr, M., Turcot, J., & Kaliouby, R. E. (2016, May). AFFDEX SDK: a cross-platform realtime multi-face expression recognition toolkit. In Proceedings of the CHI conference extended abstracts on human factors in computing systems (pp. 3723-3726). ACM, 2016.
[6] J. Chen, Y. Lv, R. Xu, and C. Xu, "Automatic social signal analysis: Facial expression recognition using difference convolution neural network," Journal of Parallel and Distributed Computing, vol. 131, pp. 97-102, 2019.
[7] Barsoum, Emad, et al, Training deep networks for facial expression recognition with crowd-sourced label distribution, ACM International Conference on Multimodal Interaction ACM, pp. 279—283, 2016.
[8] Martinez, Brais, et al, Automatic analysis of facial actions: A survey, IEEE Transactions on Affective Computing, 2017.
[9] J. M. B. Fugate, A. J. O’Hare, and W. S. Emmanuel, "Emotion words: Facing change," Journal of Experimental Social Psychology, vol. 79, pp. 264-274, 2018.
[10] Clawson, Kathy & Delicato, Louise & Bowerman, Chris. Human Centric Facial Expression Recognition. 10.14236/ewic/HCI2018.44, 2018.

[11] T. Cao and M. Li, "Facial Expression Recognition Algorithm Based on the Combination of CNN and K-Means," presented at the Proceedings of the 2019 11th International Conference on Machine Learning and Computing, Zhuhai, China, 2019.

[12] H.-D. Nguyen, S. Yeom, G.-S. Lee, H.-J. Yang, I. Na, and S. H. Kim, "Facial Emotion Recognition Using an Ensemble of MultiLevel Convolutional Neural Networks," International Journal of Pattern Recognition and Artificial Intelligence, 2018.

[13] N. Christou and N. Kanojiya, "Human Facial Expression Recognition with Convolutional Neural Networks," Singapore, pp. 539-545: Springer Singapore, 2019.

[14] Ying, ACC, Ujir, H & Hipiny, I 2017, '3D facial expression intensity measurement analysis', Proceedings of the 6th International Conference of Computing & Informatics, pp. 43-48.

[15] Ko BC 2018, 'A Brief Review of Facial Emotion Recognition Based on Visual Information', Sensors (Basel), vol.18, no.2, pii. E401.

[16] Kanyadara Saakshara, Kandula Pranathi, R.M. Gomathi, A. Sivasangari, P. Ajitha, T. Anandhi, "Speaker Recognition System using Gaussian Mixture Model", 2020 International Conference on Communication and Signal Processing (ICCSP), pp.1041-1044, July 28 - 30, 2020.

[17] R. M. Gomathi, P. Ajitha, G. H. S. Krishna and I. H. Pranay, "Restaurant Recommendation System for User Preference and Services Based on Rating and Amenities," 2019 International Conference on Computational Intelligence in Data Science (ICCIDS), 2019, pp. 1-6, doi: 10.1109/ICCIDS.2019.8862048.

[18] Subhashini R , Milani V, "IMPLEMENTING GEOGRAPHICAL INFORMATION SYSTEM TO PROVIDE EVIDENT SUPPORRT FOR CRIME ANALYSIS", Procedia Computer Science, 2015, 48(C), pp. 537–540

[19]Harish P, Subhashini R, Priya K, "Intruder detection by extracting semantic content from surveillance videos", Proceeding of the IEEE International Conference on Green Computing, Communication and Electrical Engineering, ICGCCEE 2014, 2014, 6922469.

[20] Sivasangari, A., Krishna Reddy, B.J., Kiran, A., Ajitha, P.(2020), " Diagnosis of liver disease using machine learning models", ISMAC 2020, 2020, pp. 627–630, 9243375

[21] Sivasangari, A., Nivetha, S., Pavithra,, Ajitha, P., Gomathi, R.M. (2020)," Indian Traffic Sign Board Recognition and Driver Alert System Using CNN", 4th International Conference on Computer, Communication and Signal Processing, ICCCSP 2020, 2020, 9315260

[22] Ajitha, P., Lavanya Chowdary, J., Joshika, K., Sivasangari, A., Gomathi, R.M., "Third Vision for Women Using Deep Learning Techniques", 4th International Conference on Computer, Communication and Signal Processing, ICCCSP 2020, 2020, 9315196

[23] Ajitha, P.Sivasangari, A.Gomathi, R.M.Indira, K."Prediction of customer plan using churn analysis for telecom industry",Recent Advances in Computer Science and Communications,Volume 13, Issue 5, 2020, Pages 926-929.

[24] Gowri, S. and Divya, G., 2015, February. Automation of garden tools monitored using mobile application. In International Confernce on Innovation Information in Computing Technologies (pp. 1-6). IEEE.

[25]Gowri, S., and J. Jabez. "Novel Methodology of Data Management in Ad Hoc Network Formulated Using Nanosensors for Detection of Industrial Pollutants." In International Conference on Computational Intelligence, Communications, and Business Analytics, pp. 206-216. Springer, Singapore, 2017.