Proceedings of the Fifth International Conference on Trends in Electronics and Informatics (ICOEI).
IEEE Xplore Part Number:CFP21J32-ART; ISBN:978-1-6654-1571-2

Emotion Recognition Based Emoji Retrieval Using Deep Learning

Swati Srivastava
Department of Computer Engineering & Application
GLA University
Mathura, India

Prateek Gupta
Department of Computer Engineering & Application
GLA University
Mathura, India

Pravendra Kumar
Department of Computer Engineering & Application
GLA University
Mathura, India

2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) | 978-1-6654-1571-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICOEI51242.2021.9452832

***Abstract*— Facial expression is a verbal act of speech that is expressed on the face in terms of our emotions. Emotional recognition to identify facial expression plays crucial role in various applications like psychology, linguistics etc. Playing an important role in the fields of artificial intelligence and robotics, the automatic recognition of facial expression is therefore a need for generation. Emotion recognition plays an important role in the area of human machine communication. Emotional recognition is usually done in four stages which include pre-processing, facial recognition, feature extraction, and classification. In this paper, we have used deep learning to identify the seven main human emotions: anger, disgust, fear, happiness, sadness, surprise and neutrality.**

Keywords— Facial emotion recognition, Convolutional neural networks, Deep-learning feature, Support vector machines, Expression specific features

I. INTRODUCTION

Emojis or avatars are ways to indicate non-verbally the feelings through the face. These cues have become an important part of online chatting, product review, brand emotion, and many more. Dedicated towards emoji-driven storytelling, it also leads to rapidly increasing data science research. With progress in computer vision and deep learning, it is easy to map and detect human emotions from images. In this work, we distinguished the shape of a person's face to filter and map out matching emojis or avatars. Emotional recognition allows actual-time research, as well as speculation of sensible emotional states from a person's facial expression, tone of voice, posture and body signal, and text on social media. Emotional pattern recognition based on explicit and implicit traits extracted by wearable devices and others can be determined by computer model. At present, sensitivity and statistical recognition are important in the detection and diagnosis of patients who may have an emotional disorder. In these days with the continued development of artificial intelligence and machine learning applications facial emotion recognition has become more popular. The emotion recognition plays a crucial role in interaction technology. In interaction technology the verbal section only plays a one third of communication and the non-verbal section plays a two third of communication. A facial emotion recognition (FER) method is used for detecting facial expressions. Facial expression plays a major role in expressing one's feelings and it expresses innermost attitude and his or her mental situation or human perspective. This paper aims to identify basic human emotions. The facial emotions such as happy, sad, angry, fear, surprised, neutral emotions are considered as basic emotions. Here it is used to identify and detect faces in real time. These images are captured by using anchor boxes for accuracy. Emotional analysis attracted the attention of investigators. Many past works in the field of artificial intelligence focuses on recognize sentiment rather than digging into the reason why emotions are unknown or perceived inappropriately. The interplay of emotions responsible for the access of senses. In this paper, we focused to put the difference between emotional identified and emotional mines with native language written from news. The group of emotions, which is given as non-understandable, is largely due to prejudice. There are three types of indications and two deep learning neural-network architecture are expressed as reaction connected from emotion recognition. Test results define in spontaneous statements, reactions are easily provided as smile. Statements like to describe the feelings of care-spreading anger. In targeted stories, it is easy to see the emotions of the text as love and cause the spread of fear. These findings may provide insight into applications related to interactive communication, such as public opinion, social networking, and human-to-computer communication.

The main methods of emotion recognition include the following emotional indicators: moral character; body pattern, meaning emotional signal; and psychological and emotional measures, such as textual information and multidimensional emotional details. Faces could also be among the foremost crucial functions for interaction of reactions. From 1970s it started as, countenance identification which is the maximum readable field in pure reactions, mostly within the USA and Japan. Facial emotions are considered to be visible tellers of interior emotional parts, which make emotional differences available. Emotional recognition is based on abstract observations and emotional magnitude, with the power of the CNS and PNS. Recognition of single emotion and perception of multiple emotions depends on facial expressions, speech, posture, body language, and text, which may work separately or simultaneously. Compiling the emotional data of many types, emotional recognition remains a challenge, and more research is needed on how they relate to human impact.

The paper is organized as: Section 2 includes related work. Proposed work is explained in Section 3. Section 4 represents the results. Finally, the work is concluded in Section 5.

II. Related Work

While people are able to see faces with no effort, facial recognition can be a competing problem for computer stance recognition. Face recognition programs plan to find a person's face, that is three-dimensional and convert the presence of light and face, supported by its dual image. Many face recognition models finds facial characteristics by finding symbols, or characteristics, by the face images. That is, a procedure can evaluate a related position, size, and shape of the nose, ears, lips, cheeks and eyes. These characteristics are therefore want to look for remaining images with parallel characteristics. Convolutional Neural Networks (CNNs) is one of the most widely used computer vison systems. Convolutional Neural Networks gain excellent exactitude with larger details. André Teixeira Lopes and Edilson de Aguiar suggested a conjunction of the Convolutional Neural Network and some image technological process to inform facial expressions [1]. Identifying suitable breaking points may be an important task under consideration for computing mistakes within the field of gesture to identity public facial expression. Asit Barman proposed Facial Expression identifies with the help of margin and size characteristics [2]. In [3], Facial Emotion Recognition supported Visual Information. They categorized the facial emotion recognition system into two parts; with either the features are handcrafted or generated through the output of a deep neural network. In [4], the authors proposed FERAtt: Facial expression recognition attentively Net. They used dual-branch network, first they separate the face structure from other parts in image including hairs than detect the expression from the faces created with their approach. In various fields, scientists are using Facial Expression Recognition attentively Net because it's CNN based method and deep Convolutional Neural Networks (CNN) has recently shown excellent performance during a major photo segmentation project. Current facial emotion recognition systems may suffer from many weaknesses, two of them may be: facial image may collapse to some conversions within the images while using various approaches. If the classifier isn't powerfully built, it degrades the performance. So to seek out the solutions of those problems [5] proposed intellectual facial recognition is based on constant wavelet entropy in which they use a single hidden network to feed the neural network because it is a detector. In [6], authors have proposed countenance recognition system with a variant of evolutionary firefly algorithm for feature optimization. It is used for feature selection as it outperforms various optimizations. The characteristics selection for facial emotion recognition using memetic algorithm to beat the matter of reluctancy and irrelevant data after various steps of other algorithms [7]. Nowadays, hybrid features are required for reading learning and theorization. Recently, in [8], the researchers proposed hybrid feature representation for the popularity of facial expressions. This model basically uses the support vector machine (SVM) technique which combined the features of the image by the mixture of SIFT and deep learning models to classify the facial expressions. As deep neural network has high capability of information retrieval [11], it is widely used in retrieval based applications. In [9], human movement in real life has been identified using RF classifier. The methods based on autoencoders, deep Boltzman machine, sum product network, Recurrent NN have been summarized in [10]. In the proposed work, deep learning is used to identify the emotions of human faces and then map the emotions to the emojis.

III. Proposed Work

In this paper, we proposed a deep learning model to classify facial expressions from the images. Then we mapped the classified emotion to a corresponding emoji or an avatar. We built a convolution neural network to acknowledge facial emotions. We trained our model on the FER2013 dataset. Then mapping of those emotions with the corresponding emojis or avatars have been performed. Using OpenCV's haar cascade xml, we obtained the bounding box of the faces within the webcam. Then we feed these boxes to the trained model for classification. In the below will build a convolution neural network structure and train the model on FER2013 dataset for Emotion recognition from images. The FER2013 dataset (facial expression recognition) consists of 48*48pixel grayscale face images. The images are centered and filled with an equal amount of space. This dataset consists of facial emotions of following descriptions; angry, disgust, fear, happy, sad, surprise, natural. The training set consists of 25000+ images and the public test set consists of 3500 images.

A. Proposed Framework

(a) Data preprocessing

We used image processing techniques such as reshape, resizing, converting to greyscale, and standardization. NumPy and pandas library are used for implementation. With the help of power vectorization array and pandas data frames, images have been processed whenever required the images with the help of NumPy arrays and generate the output by pandas. Preprocessing results in an improved image that suppresses undesired distortions or enhances some image features relevant for further processing and analysis task. First we have reduced the noise of images with the help of auto-encoding algorithm[12] which help in reduction of unwanted data. It also helps to reduce dimensionality of data. Then we have converted the images into grayscale. We have used image segmentation through non-contextual thresholding and contextual segmentation[13] so that it can convert part of an image into multiple parts and separating foreground from background so that emotion on the face in an image can be easily detected[14][15]. With geometric transformation, positions of pixels in an image are modified. Figure 1 represents the proposed framework.

(b) Make model using Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN)[16] is a deep learning algorithm that we have used in this work which can capture image input, give value (readable and discriminatory bias) to other elements within the photo, and possibly may be distinguish one from the other. The previous processing wanted for CNN is as small as the equivalent of other class algorithms. CNN can read the features without extracting the features manually. Keras with TensorFlow has been used that works as a background for building Neural Networks. Further, we have various layers for cameras to be integrated; Pooling layer, Convolution layer, Batch normalization, Activation Layer, Dropout Layer, Flatten Layer, Dense layer.

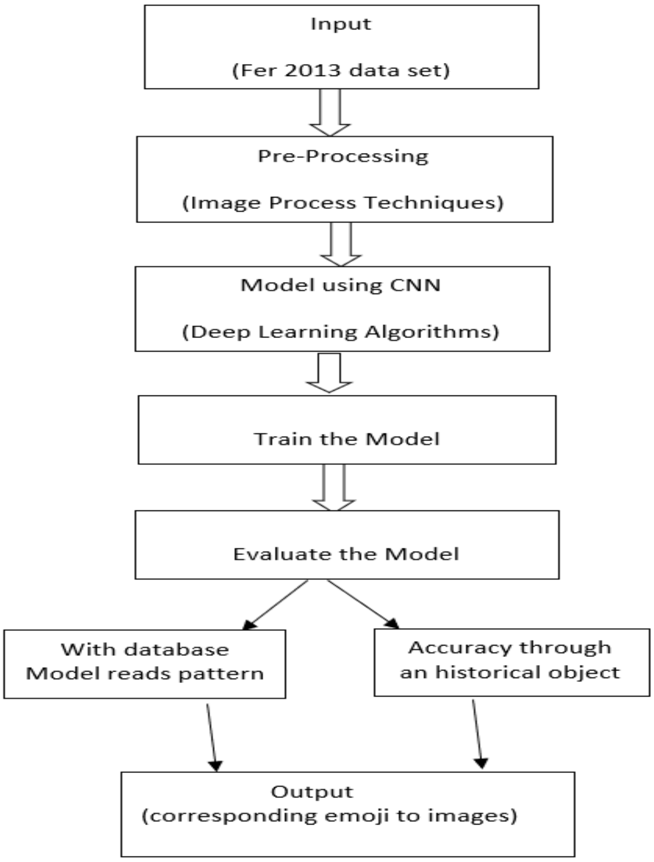

Fig.1 : Proposed Framework

(c)Training the model

This is time to train the model. Coaching is a learning process. Here we describe hyperparameters such as quantity of epochs, batch size, reading of rate etc. We have used callback function to write down model presentation during the model is working.

(d) Evaluating the model

With the help of the database, it's time to take a look at how well a model reads patterns. Let's take a look at how accuracy changes with epochs using a historical object. Given are the discharge callbacks which are recorded through trained network. Write down training metrics of all epoch, the history object is backed by calls to the fit() function for training the representation, Metrics are written in glossary in the history of the come back item, plot loss vs number of epochs.

IV. Results

After going through various step of processes we obtained the result of this proposed model which will detect the facial expressions and map those emotions on the faces with the corresponding emojis or avatars.. Figure 2 shows some sample emojis which will be used for emotion recognition. Figure 3 shows some of the outputs that we got as a result of the proposed system. Table1 and 2 shows the training and testing accuracy respectively.

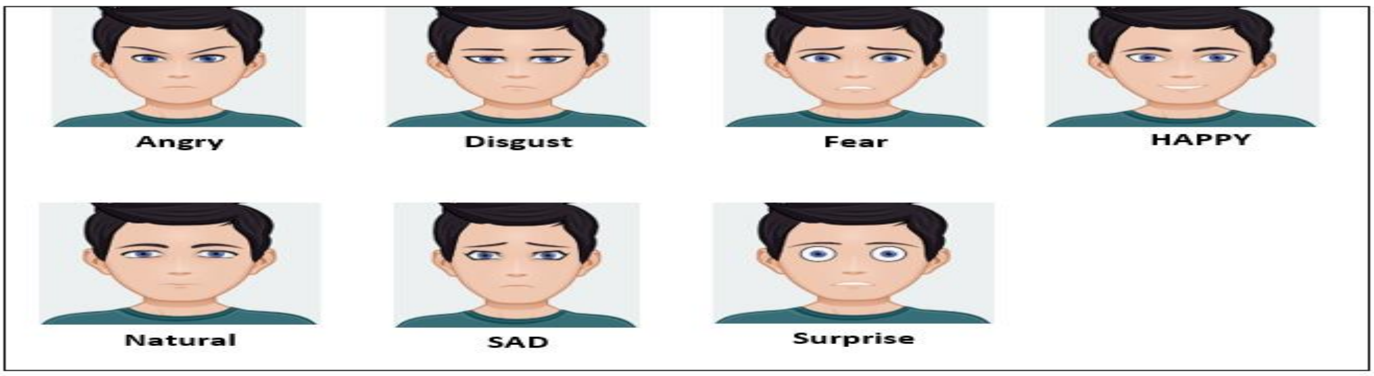

Fig. 2. Sample emojis

Table1: Training Accuracy

EMOTION	TRAINING IMAGES	CORRECT RESULTS	INCORRECT RESULTS	ACCURACY (%)
ANGRY	1000	897	103	89.7
DISGUST	1000	868	132	86.8
FEAR	1000	912	088	91.2
HAPPY	1000	924	076	92.4
NATURAL	1000	851	149	85.1
SAD	1000	872	128	87.2
SURPRISE	1000	902	098	90.2

Table2: Testing Accuracy

EMOTION	TESTING IMAGES	CORRECT RESULTS	INCORRECT RESULTS	ACCURACY (%)
ANGRY	500	452	48	90.4
DISGUST	500	437	63	87.4
FEAR	500	454	46	90.8
HAPPY	500	447	53	89.4
NATURAL	500	421	79	84.2
SAD	500	449	51	89.8
SURPRISE	500	426	74	85.2

Fig. 3. Results of the proposed Model

V. CONCLUSION

In this research, we developed an improved emotion recognition model for facial expression images. We have proposed a convolutional neural network based deep learning model which is tested on FER2013 dataset. During these testing we finds results more accurate and performance is also high. Future researchers can improve the efficacy of the system by changing the preprocessing steps. Also in classification module, researchers can go for complex domain computing instead of real domain for better results.

REFERENCES

[1] Lopes, A. T., de Aguiar, E., De Souza, A. F., & Oliveira-Santos, T. (2017). Facial expression recognition with

convolutional neural net-works: coping with few data and the training sample order. Pattern Recognition, 61, 610-628.

[2] Barman, A., & Dutta, P. (2017). Facial expression recognition us-ing distance and shape signature features. Pattern Recognition Let-ters.

[3] Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. sensors, 18(2), 401.

[4] Marrero Fernandez, P. D., Guerrero Pena, F. A., Ren, T., & Cunha, A. (2019). Feratt: Facial expression recognition with attention net. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops.

[5] Wang, S. H., Phillips, P., Dong, Z. C., & Zhang, Y. D. (2018). Intel-ligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm. Neurocomputing, 272, 668-676.

[6] Zhang, L., Mistry, K., Neoh, S. C., & Lim, C. P. (2016). Intelligent facial emotion recognition using moth-firefly optimiza-tion. Knowledge-Based Systems, 111, 248-267.

[7] Ghosh, M., Kundu, T., Ghosh, D., & Sarkar, R. (2019). Feature se-lection for facial emotion recognition using late hill-climbing based memetic algorithm. Multimedia Tools and Applications, 78(18), 25753-25779.

[8] Wang, F., Lv, J., Ying, G., Chen, S., & Zhang, C. (2019). Facial ex-pression recognition from image based on hybrid features under-standing. Journal of Visual Communication and Image Representation, 59, 84-88.

[9] Ranganathan, G. (2020). Real Life Human Movement Realization in Multimodal Group Communication Using Depth Map Information and Machine Learning. Journal of Innovative Image Processing (JIIP), 2(02), 93-101.

[10] Smys, S., Chen, J. I. Z., & Shakya, S. (2020). Survey on Neural Network Architectures with Deep Learning. Journal of Soft Computing Paradigm (JSCP), 2(03), 186-194.

[11] Suma, V. (2020). A Novel Information retrieval system for distributed cloud using Hybrid Deep Fuzzy Hashing Algorithm. JITDW, 2(03), 151-160.

[12] Pan, Y., He, F., & Yu, H. (2020). Learning social representations with deep autoencoder for recommender system. World Wide Web, 23(4), 2259-2279.

[13] Bhargavi, K., & Jyothi, S. (2014). A survey on threshold based segmentation technique in image processing. International Journal of Innovative Research and Development, 3(12), 234-239.

[14] Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. Information Fusion, 59, 103-126.

[15] Saxena, A., Khanna, A., & Gupta, D. (2020). Emotion recognition and detection methods: A comprehensive survey. Journal of Artificial Intelligence and Systems, 2(1), 53-79.

[16] Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017, August). Understanding of a convolutional neural network. In 2017 International Conference on Engineering and Technology (ICET) (pp. 1-6). Ieee.