A Review of Various Deep Learning Methods Used in Automatic Facial Expression Recognition System

Rajan Saini, Shainy Bakshi
Scholar (M.Tech), Assistant Professor
Department of Computer Science Engineering
Yamuna Institute of Engineering and Technology, Yamunanagar, Haryana.
rajansaini848@gmail.com, shainybakshi15@gmail.com

Abstract- The most effective and simplest wordless emotion interaction is facial expression. The possible advantages of individual FER termed as facial emotion recognition has stimulated the interest of academic researchers. During FER, the major aim is to correlate distinct facial features with distinct emotional reactions. Facial Expression Recognition (FER) is used in a variety of experimental fields, including the treatment of mental health conditions and the identification of social relationships. FER techniques have been built to directly support usage situations rather than scientific conditions, thanks to growing modern technology in equipment and sensor systems. There are different types of human facial emotions such as anger, fear, joy, sadness, etc. The different categories of emotions are explained in this paper. The face recognition system with basic architecture is also presented in the work. The results of various existing techniques of the facial emotion recognition system are also compared in this paper.

Keywords – Emotion, Facial Expression Recognition, Deep Learning Methods

I. INTRODUCTION

Emotion is a broad concept which comprises a huge range of significant mental factors. Whenever it comes to a single individual, item, or event, certain feelings are quite particular. Pain, happiness [1], and sadness, for example, are fairly broad terms. Many feelings, including a momentary blush of shame or a flash with wrath, are fleeting yet hardly noticeable [2]. The different categories and types of emotions are shown in table 1.

Basic emotions are those that have been formed to ensure an individual's survival. Whenever confronted with a scenario involving a basic emotion, the person doesn't analyze; the response is instantaneous and therefore does not necessitate the use of thought patterns. Basic emotions are the most natural feel. The facial expressions of basic emotions are shown in fig.1.

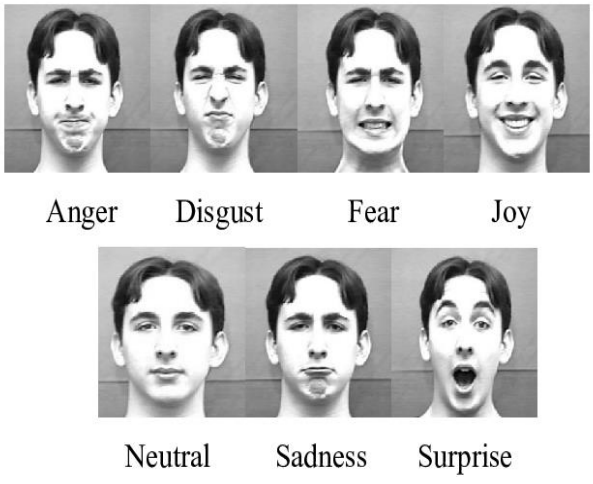

Fig. 1 Basic types of emotions [3]

Table 1: Different categories and types of emotions

Sr. No.	Category of Emotions	Different Types
1.	Basic	• Anger • Fear • Joy • Disgust • Neutral • Sadness • Surprise
2.	Complex	• Pride • Shame • Anxiety • Jealousy
3.	Ambiguous	• Hope • Compassion

1.1 Complex Emotions: The experience of these complex emotions seems to have a serious influence on someone's personality as well as self-esteem, including satisfaction in positive feedback that boosts self-esteem. Regret, on the other hand, encourages people to repair harm because this is an unpleasant sensation, and this is a method for rebuilding somebody's morals even though it has an adverse influence on someone's self-esteem.

1.2 Ambiguous Emotions: Although ambiguous feelings are much more complicated than most other kinds of stimuli, their growth takes longer. Furthermore, the impact of all these feelings may vary based mostly on a person's improvement procedure while experiencing things.

In paper organization as face emotion recognition and process are explained in sect 2. The deep learning based face recognition system is explained in section 3. The existing methods are surveyed in section 4. The comparison

of existing methods based on the emotions accuracies are depicted in section 5. The graphical representations of results are shown on section 6. The overall conclusion and future scope is explained in section 7.

II. FACE EMOTION RECOGNITION

Several actions or postures of the muscular underneath the surface of the skin constitute the facial features. Such motions, as per some set of contentious beliefs, communicate an unpleasant psychological condition for spectators. Nonverbal communication is the kind of facial movements. Individuals utilize facial movements to transmit a multitude of interpretations in a variety of situations. The implications extend from simple, maybe inherent social and economic ideas like "shocked" to sophisticated, society notions like "thoughtlessly." Face emotions are used in a wide variety of circumstances, from reactions to environmental stimuli to specific linguistic formulations across idiomatic phrases [4]. The technique of identifying emotional responses using facial movements is known as facial emotion recognition. The nervous system perceives feelings instinctively, and technology that really can recognize emotions has recently been created. The technologies are constantly improving, and it will soon be able to recognize feelings equally accurately as human minds. Through recognizing what another face signifies and apply this information to new data, Intelligence can recognize sentiments. Emotional AI, or emotion machine intelligence, seems to be a system that can detect, imitate, understand, and react to various face gestures as well as feelings [5]. The basic architecture of the face recognition system is shown in fig. 2.

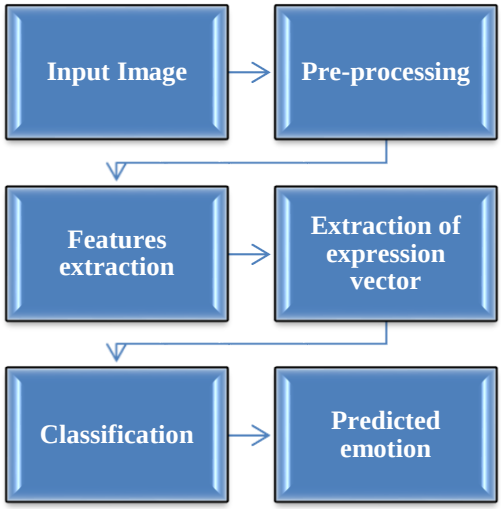

Fig. 2 Process of Recognition of Face Emotions

III. DEEP LEARNING-BASED FACE RECOGNITION SYSTEM

Despite the widespread effectiveness of classical facial emotion recognition systems based mostly on extraction from handmade components, scientists have drawn to deep learning over the last years because of its great appropriate classification capability [6]. The architecture of the facial emotion recognition system is presented in fig.2. There are multiple steps used for emotion detection. The initial step is pre-processing step in which the image is converted into a grayscale image. The essential features are extracted for the training of the classifier. The following steps are used in the facial emotion recognition system.

3.1 Pre-processing: The initial stage of the facial emotion recognition system is the pre-processing stage. In this stage, the image is enhanced by using various filters and all types of noises are removed from the image. The contrast of the image is also evaluated in this step. In the area of image processing, pre-processing is a crucial stage in image representation. Developed by the united must be used to identify critical characteristics before facial expressions can be extracted.

3.2 Feature Extraction: Feature extraction is the second stage of the emotion recognition system. The essential features are extracted from the images in this stage. The extracted features are then used for the training purpose of the model.

3.3 Classification and Detection: At the classification stage, several emotions are classified from the images. There are different types of facial emotions such as sadness, happiness, anger, etc. The actual facial emotion is recognized after the classification stage [6].

IV. LITERATURE SURVEY

Proposed deep learning (DL) based system for facial expression recognition by **Li, B., et al., [7]**. ResNet-50 was the pre-trained model employed in the suggested architecture. The key features were extracted using the convolution layer, and the features were reduced using the pooling layer. The efficiency of the proposed model was improved by the activation function. The overfitting problem was reduced by the cross-validation method. The authors were demonstrated that the proposed approach outperforms existing mainstream face facial expression recognition models in face emotion identification using experimental simulation of the provided collected data. According to **Gupta et al., [8]** emotions and feelings, are made up of a range of sub-emotions that are difficult to categorize. ResNet, an attentive component, and CNN, which provides spatial visibility to the system, were integrated with the suggested deep-learning system. The proposed framework was more effective at detecting real-world facial reactions. Mostly on the FER sample, the developed approach had produced satisfying performance and demonstrated effectiveness. By reaching better accuracy of 85.76 per cent throughout the training process and 64.40 per cent throughout the test phase, the developed scheme beat established CNN-based models. Also on the FER database, the prediction design performed much better and might be employed in application scenarios. With the use of

frame and video level features, **Viegas, C., et al. [9]** suggested an emotion identification system. The suggested study uses two-stage categorization to analyze a seven-class predictor for emotion recognition categorization. The MER termed as Multimodal Emotion Recognition dataset was used for the evaluation of the system. The unimodal and multimodal attributes are compared. The multimodal had achieved 50 per cent of accuracy and the seven-class predictor had achieved 49% of accuracy. **Kim, S., et al. [10]** created a DNN (deep neural network) based facial emotion identification system. The authors discovered that the assumption of facial emotion had no explanation. Based on the notion that facial expression was a blend of facial muscles, the scientists identified a link between Emotion marks and Facial Action Coding Units in the CK+ Dataset. The authors had introduced a framework that takes facial action units to describe the prediction performance of a CNN-based neural model. The convolutional neural network classifier was constructed on the CK+ Dataset and uses derived features to identify the feeling. The CNN neural model's derived characteristics and expression groups were used to classify the numerous facial action coding units in the interpretation framework. **Chen et al. [11]** proposed a structure for addressing the problem of facial emotion detection. In our study, we used two quite different visual and aural modes. A new component predictor called Histogram of Oriented Gradients from Three Orthogonal Planes (HOG-TOP) was presented to remove complex details through action recognition and to characterize visual illumination fluctuations. To identify differences specification, a potentially innovative architectural function resulting from warp modification of face images was proposed. The experiment was conducted upon the enhanced Cohn-Kanade (CK+) dataset and AFEW (Acted Facial Expression in Wild 4.0) dataset to demonstrate that the proposed methodology was more effective than other state-of-the-art strategies in interacting through multimedia facial expression recognition issues in an experimental system and even in the general. For the detection of face emotions, **Jain, N., et al. [12]** used a hybrid methodology. Convolutional neural network (CNN) and recurrent neural network (RNN) were used in the methodology. The public available datasets were used for the evaluation of the proposed model. The authors found that the combination of two neural networks can enhance the efficiency and identification results. On the JAFFE dataset, 94.91 per cent of accuracy was achieved and with the MMI dataset, 92.07% of accuracy was attained. The various existing techniques of face emotion recognition with classification and compared techniques are depicted in Table 2. Table 3 is the advantages and problem statement with the dataset of existing techniques.

Table 2: Several Existing Techniques of Face Emotion Recognition System

Author Name	**Year**	**Proposed Methods**	**Classification Techniques**	**Compared Techniques**
Li, B., et al. [7]	2021	Face recognition framework based on deep learning	ResNet-50	HWT CSO BBO
Gupta, et al. [8]	2020	Self-attention deep network	ResNet	CNN CNN+ Residual connection
Viegas, C., et al. [9]	2020	Two-stage multi feature-based face emotion recognition system	Seven class classifier	Unimodal and multimodal
Kim, S., et al. [10]	2019	Convolutional neural network-based framework	CNN	-
Chen et al. [11]	2019	Fusion of multiple features based approach	Prior Fusion Net (PFN)	CNN DeRF
Jain, N., et al. [12]	2018	A hybrid methodology based on deep neural network	CNN	CNN CNN-RNN CNN- RNN + ReLU

Table 3: Different Existing Techniques of Dataset used, advantages and Issues

Reference	Advantages	Issues/Gaps	Dataset Used
[7]	Provided efficient results in the multi-class classification of facial emotions	Poor optimization results	Collection of self-capture images
[8]	Reduces the vanishing gradients issues,	Classify only basic emotions	FER dataset
[9]	Extracted multiple features	Less accurate recognition results	eNTERFACEO5 dataset
[10]	Provided efficient results for classification	Hard to handle multi-modal features	MER dataset
[11]	Provide efficient results in multiple classifications	Provided poor results in a similar type of expression detection	CK+
[12]	More number of layers provided better classification results	Inefficient results for the detection of complex emotions	MMI JAFFE

V. EXISTING METHOD BASED RESULT EVALUATIONS

In this section, the existing results are evaluating in facial emotional recognition system. The various existing techniques based on deep learning methods comparison are shown in table 4. The comparison of existing methods based on accuracy rate (%).

Table 4: Comparison Analysis

Author's and Year	Classification/Feature Extraction Methods	Happy	Fear	Anger	Disgust	Neutral	Sadness
Li, B., et al. 2021[7]	Convolution operations	95.90%	95.70%	95.50%	94.50%	96.40%	94.40%
Gupta et al., 2020[8]	ResNet	64%	64%	64%	64%	64%	64%
Viegas, C., et al., 2020 [9]	Frame level features extraction	49%	49%	49%	49%	49%	49%
Kim, S., et al., 2020 [10]	CNN model	75%	-	96%	-	-	-
Chen et al., 2018 [11]	CNN model	98%	98%	98%	98%	98%	98%
Jain, N., et al., 2018 [12]	CNN-RNN	93.68%	64.36%	94.40%	64.36%	72.12%	62.36%

The graphical representations of results of various existing techniques are shown in fig. 3 and fig. 4.

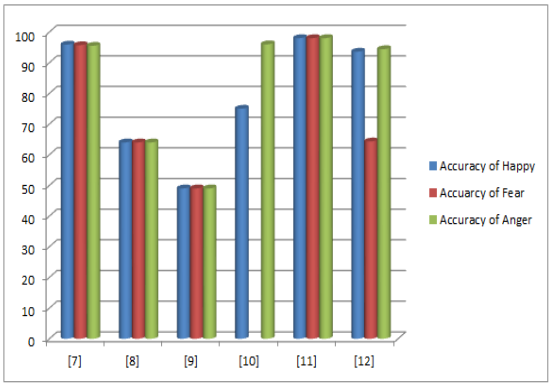

Fig. 3 Graphical representation of existing techniques based on the accuracy of different emotions

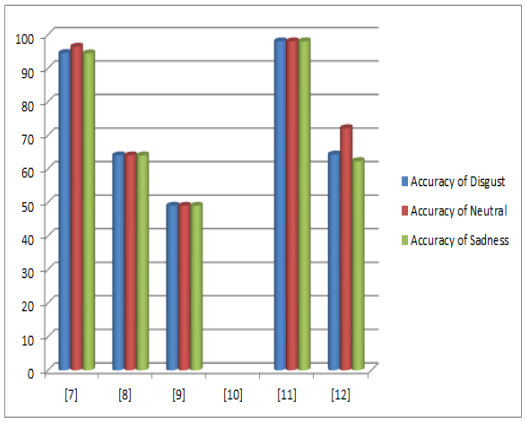

Fig. 4 Graphical representation of existing techniques based on the accuracy of different emotions

VI. CONCLUSION AND FUTURE SCOPE

In this conclusion, various existing techniques of the facial emotion recognition system are surveyed in this paper. Deep learning has a great impact on recognition systems; the basic architecture of a deep learning-based face emotion recognition system with different phases such as Facial image pre-processing, Feature extraction, classification and recognition. The accuracy rate results of different emotions are presented in the given table 4. The accuracy results based on different emotions of existing techniques are presented graphically. FERs are amongst the most essential means of delivering feedback over an individual's psychological response, and they are often constrained by just knowing the basic emotions including neutrality. It clashes with what each feels in everyday life that has more complicated feelings. In the future, complex emotions will be surveyed and compared with different deep learning methods.

VII. REFERENCES

[1] Emotion - The physical expression of emotion. (2021). Retrieved 8 July 2021, from https://www.britannica.com/science/emotion/The-physical-expression-of-emotion.

[2] Types of emotions - Kindling a flame. (2021). Retrieved 8 July 2021, from https://sites.google.com/site/kindlingaflame/2-socio-affective-development/types-of-emotions.

[3] Kwong, J. C. T., Garcia, F. C. C., Abu, P. A. R., & Reyes, R. S. (2018, October). Emotion recognition via facial expression: Utilization of numerous feature descriptors in different machine learning algorithms. In TENCON 2018-2018 IEEE Region 10 Conference (pp. 2045-2049). IEEE.

[4] Elliott, E. A., & Jacobs, A. M. (2013). Facial expressions, emotions, and sign languages. Frontiers in psychology, 4, 115.

[5] Mellouk, W., & Handouzi, W. (2020). Facial emotion recognition using deep learning: review and insights. Procedia Computer Science, 175, 689-694.

[6] Jaiswal, A., Raju, A. K., & Deb, S. (2020, June). Facial emotion detection using deep learning. In 2020 International Conference for Emerging Technology (INCET) (pp. 1-5). IEEE.

[7] Li, B., & Lima, D. (2021). Facial expression recognition via ResNet-50. International Journal of Cognitive Computing in Engineering, 2, 57-64.

[8] Gupta, A., Arunachalam, S., & Balakrishnan, R. (2020). Deep self-attention network for facial emotion recognition. Procedia Computer Science, 171, 1527-1534.

[9] Viegas, C. (2020, November). Two Stage Emotion Recognition using Frame-level and Video-level Features. In 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020) (pp. 912-915). IEEE.

[10] Kim, S., & Kim, H. (2019, February). Deep explanation model for facial expression recognition through facial action coding unit. In 2019 IEEE International Conference on Big Data and Smart Computing (BigComp) (pp. 1-4). IEEE.

[11] Chen, Y., Wang, J., Chen, S., Shi, Z., & Cai, J. (2019, December). Facial motion prior networks for facial expression recognition. In 2019 IEEE Visual Communications and Image Processing (VCIP) (pp. 1-4). IEEE.

[12] Jain, N., Kumar, S., Kumar, A., Shamsolmoali, P., & Zareapoor, M. (2018). Hybrid deep neural networks for face emotion recognition. Pattern Recognition Letters, 115, 101-106.