IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A HIGH PERFORMANCE AND COMPACT MODEL FOR FACIAL EXPRESSION RECOGNITION USING DEEP LEARNING

[1]Smt.R.Lavanya, [2]P.Jahnavi, [3]V.Pavana Chandrika, [4]R.Dileep, [5]P.Sai Rahul

[1]Assistant Professor, [2]UG Scholar, [3]UG Scholar, [4]UG Scholar, [5]UG Scholar

Dept. of Electronics & Communication Engineering,
NBKR Institute of Science & Technology, Andhra Pradesh, Vidyanagar, India.

Abstract: Facial emotion recognition (FER) is an important topic in the fields of computer vision and artificial intelligence owing to its significant academic and commercial potential. Although FER can be conducted using multiple sensors, this review focuses on studies that exclusively use facial images, because visual expressions are one of the main information channels in interpersonal communication. This project provides a brief review of researches in the field of FER conducted over the past decades.

First, conventional FER approaches are described along with a summary of the representative categories of FER systems and their main algorithms. FER approaches using neural networks enabling "end-to-end" learning are then presented. This work also focuses on an up-to-date hybrid learning approach combining a neural network (NN) for the spatial features of an individual frame and long short-term memory (LSTM) for temporal features of consecutive frames. In the later part of this paper, a brief review of publicly available evaluation metrics is given, and a comparison with benchmark results, which are a standard for a quantitative comparison of FER researches, is described. This review can serve as a brief guidebook to newcomers in the field of FER, providing basic knowledge and a general understanding of the latest state-of-the-art studies, as well as to experienced researchers looking for productive directions for future work.

Index Terms **- –** Facial Emotion Recognition, Conventional FER, Convolutional Neural Networks, Deep Learning- based FER, Principal component Analysis (PCA).

I. INTRODUCION

Facial emotions are important factors in human communication that help us understand the intentions of others. general, people infer the emotional states of other people, such as joy, sadness, and anger, using facial expressions and vocal tone. According to different surveys, verbal components convey one-third of human communication, and nonverbal components convey two-thirds. Among several non-verbal components, by carrying emotional meaning, facial expressions are one of the main information channels in interpersonal communication. Therefore, it is natural that research of facial emotion has been gaining lot of attention over the past decades with applications not only in the perceptual and cognitive sciences, but also in affective computing and computer animations.

Interest in automatic facial emotion recognition (FER) (Expanded form of the acronym FER is different in every paper, such as facial emotion recognition and facial expression recognition. In this paper, the term FER refers to facial emotion recognition as this study deals with the general aspects of recognition of facial emotion expression.) has also been increasing recently with the rapid development of artificial intelligent techniques, including in human-computer interaction (HCI), virtual reality (VR), augment reality (AR), advanced driver assistant systems (ADASs), and entertainment. This paper first divides researches on

automatic FER into two groups according to whether the features are handcrafted or generated through the output of a deep neural network.

Deep-learning-based FER approaches highly reduce the dependence on face-physics-based models and other pre-processing techniques by enabling "end-to-end" learning to occur in the pipeline directly from the input images. Among the several deep-learning models available, the convolutional neural network (CNN), a particular type of deep learning, is the most popular network model.

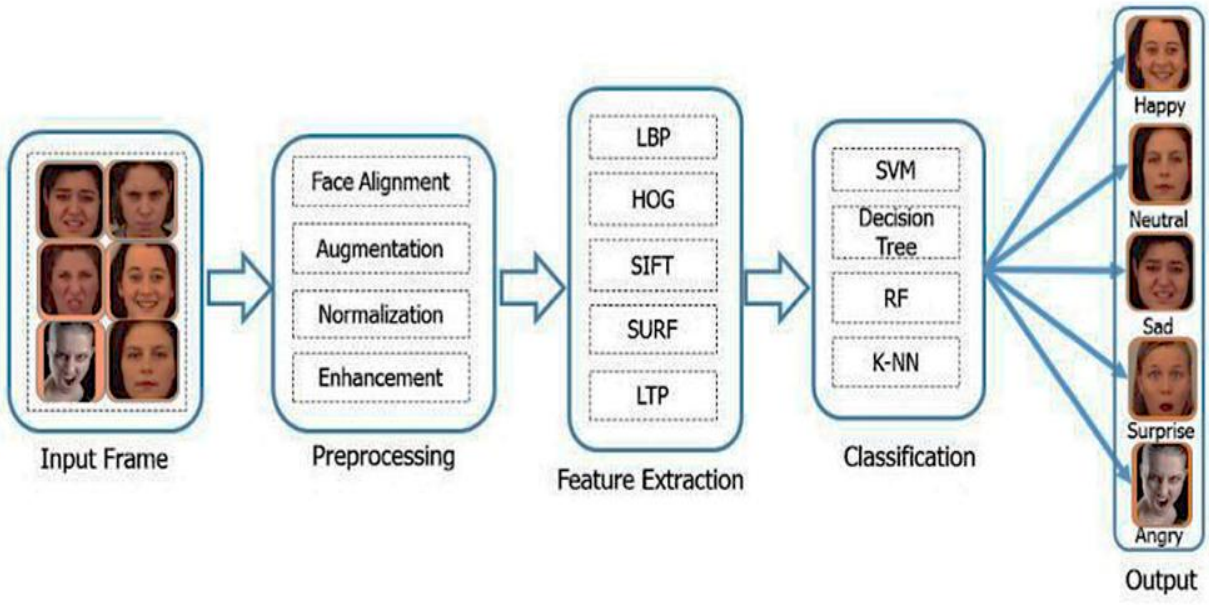

Fig.1. Architecture of Face Emotion Recognition

II. LITERATURE SURVEY

Souza, A. F. De and Oliveira-santos, Classified and featured extraction CNN algorithm are used as a classifier to identify the person's facial expression. CK+, JAFFE and BU3DFE databases is used in the experiment; CNN algorithm classification rate of 82.10% With JAFFE database. On the other hand, it gives accuracy rate of 96.76% With the CK+ database and accuracy rate of 82% With BU3DFE database.

Sönmez, E. B., and Kelimeler, on the other hand, used the Extracting temporal information technique for feature extraction. To identify a person's face expression, a sparse representation technique is utilized as a classifier for pattern classification. The experiment uses the ADFES-BIV database, This features 12 North European actors (5 male, 7 female) and 10 Mediterranean actors (5 male, 5 female) showing the six fundamental emotions as well as three complicated emotions of scorn, pride, and humiliation, as well as neutral expressions. Sparse representation algorithm classification rate in low intensity: 66.9%, middle intensity: 79.6%, and high intensity: 80.3 percent.

Mohammad, M., et al. suggested a deep learning-based system for human emotion recognition. For feature extraction, the suggested system uses Gabor filters, followed by a deep convolutional neural network. The proposed characteristics improved the speed and accuracy of training the neural network, according to the results of the experiments. As a result of using the Gabor filter, the system's learning time was reduced and its accuracy increased.

Sun and Yu used various methods for feature extraction Gabor, and LBP. For pattern classification purposes, the Linear, RBF and polynomial kernel SVM are used as a classifier to identify the person's facial expression. An extended Cohn-Kanade (CK+), database was used in the experiment. Contained 593 sequences from 123 subjects, and the 6-basic emotions beside neutral. By comparing RBF SVM, it gives the classification rate of 97.42% will the, polynomial SVM gives the accuracy rate of 96.45% and linear SVM gives the accuracy rate of 97.10% for classification.

Piparsaniyan et al. used Gabor filter for feature extraction technique. For pattern classification Bayesian algorithm are used as a classifier to identify the person's facial expression. JAFFE database is used in the experiment; JAFFE Contains 213 images of female facial expressions expressed by 10 subjects. Bayesian algorithm classification rate of 96.73%.

Souza and Oliveira-santos et al classified and featured extraction CNN algorithm are used as a classifier to identify the person's facial expression. CK+, JAFFE and BU3DFE databases is used in the experiment; CNN algorithm classification rate of 82.10% With JAFFE database. On the other hand, it gives accuracy rate of 96.76% With the CK+ database and accuracy rate of 82% With BU3DFE database.

III. METHODOLOGY

A complete facial emotions recognition system can be split into three stages; image acquisition, feature extraction, and pattern classification. The main methodology of this work is illustrated in Fig 2. Firstly, brief information about the dataset in this study is presented. Next, the concept transfer learning are explained for extracting the feature from the facial image. To enhance the performance of facial emotion recognition system. Finally, details about the classification algorithm ((VGG16) transfer learning + Classic Neural Networks).

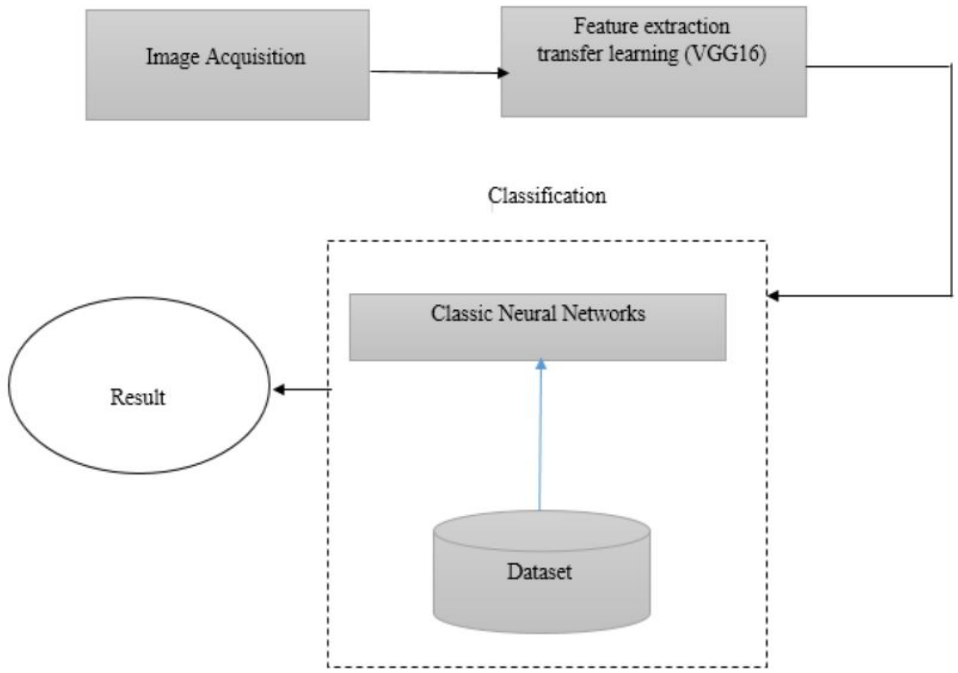

Fig. 2. Block Diagram of Facial emotion recognition System

A. Dataset

Facial Expression Recognition (FER-2013) was used as the database, There are 35,887 photos in this collection. It's broken down into 28,709 training samples, 3,589 validation samples, and 3,589 test samples, with basic expression labels provided for each photos with a resolution of 48 × 48 pixels. It was made by compiling the results of each emotion's Google picture search, as well as synonyms for the emotions.

Fig. 3. shows the sample of (FER-2013) facial emotion dataset

B. Feature extraction

Used transfer learning (VGG16) principles to use the pretrained model and train on FER-2013 dataset images. Of all the configurations, VGG16 was identified to be the best performing model on the ImageNet dataset. Fig 4 shows the actual architecture of VGG16 model . Facial feature extraction is the process of extracting face component features like eyes, nose, mouth, etc from human face image. Facial feature extraction is very much important for the initialization of processing techniques like face tracking, facial expression recognition or face recognition.

Feature extraction is the most vital stage in pattern recognition and data mining. In this stage, the meaningful feature subset is extracted from original data by applying certain rules. For reliable recognition, it is desirable to extract appropriate features space, since all the extracted features may not contribute to the classification positively. In this paper, some feature extraction methods and algorithms were studied, compared and means of improving feature selection through dimension reduction was explained.

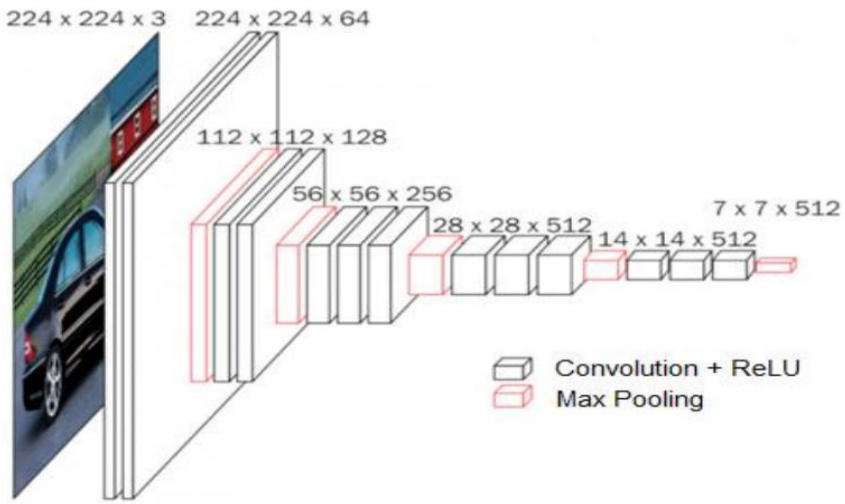

Fig. 4. shows the actual architecture of VGG16 model.

C. Classification

Classic Neural Networks are used to classify the data. The architecture classification system is depicted in Fig 5.

Classic Neural Networks It's also known as Fully Connected Neural Networks, and it's distinguished by its multilayer perceptron, which connects the neurons to a continuous layer. Fran Rosenblatt, an American Psychologist, created it in 1958. It entails the transformation of the model into basic binary data inputs. The following are the three functions incorporated in this model**:**

Linear function It shows a single line that multiplies its inputs with a constant multiplier, as the name implies.

Non-Linear function: is one that does not follow a straight line. It's further broken down into three sections:

a. Softmax OR Sigmoid Curve: The softmax and sigmoid functions are both activation functions used in neural networks. The primary difference between them is that the softmax function works for multiclass classification problems and the sigmoid function is a better option for binary class problems.
b. Tangent hyperbolic (tanh): The Tanh (Hyperbolic Tangent) activation function is one of the fundamental activation functions for deep learning.
c. Rectified Linear Unit (ReLU): The Rectified linear activation function or ReLU for short is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero.

Fig. 5. shows the architecture classification system.

IV. SOFTWARE REQUIREMENTS

MATLAB (matrix laboratory) is a multi-paradigm numerical computing environment and fourth generation programming language. A proprietary programming language developed by math works. MATLAB allows matrix manipulations plotting of functions and data, implementation of algorithms creation of user interferences, and interfacing with programs written in other languages including c, c++, java, FOTRAN and python. Although MATLAB is intended primarily for numerical computing an optional toolbox uses the MiPads symbolic engine, allowing access to symbolic computing, capabilities. An additional package, Simulink, adds graphical multi domain simulation and model-based design for dynamic and embedded systems.

MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. This allows you to solve many technical computing problems, especially those with matrix and vector formulations, in a fraction of the time it would take to write a program in a scalar non interactive language such as C or Fortran. The name MATLAB stands for matrix laboratory.

MATLAB was originally written to provide easy access to matrix software developed by the LINPACK and EISPACK projects. Today, MATLAB engines incorporate the LAPACK and BLAS libraries, embedding the state of the art in software for matrix computation. MATLAB has evolved over a period of years with input from many users. MATLAB features a family of add-on application-specific solutions called toolboxes. Very important to most users of MATLAB, toolboxes allow you to Learn and apply specialized technology. Toolboxes are comprehensive collections of MATLAB functions (M-files) that extend the MATLAB environment to solve classes of problems. Areas in which toolboxes are available include signal processing, control systems, neural networks, fuzzy logic, wavelets, simulation, and many others.

V. RESULTS

A set of baselines are presented to better understand the performance of the experiments compared to other publications. The problem to be solved is detection of faces in an image. A human can do this easily, but a computer needs precise instructions and constraints. To make the task more manageable, Viola–Jones requires full view frontal upright faces. Thus in order to be detected, the entire face must point towards the camera and should not be tilted to either side. While it seems these constraints could diminish the algorithm's utility somewhat, because the detection step is most often followed by a recognition step, in practice these limits on pose are quite acceptable.

The relevance of this work is in the performance evaluation of the PCA (Principal component Analysis) formulation in compressing digital images from the measurement of the degree of compression and the degree of information loss that the PCA introduces into the compressed images in discarding some principal components.

- Extraction of major features, this enhances the speed of operation.
- Recognition based on Spectral features.
- Neural network enhances the classification process..
- Error recognition due to illumination changes and motion blur will be reduced due to face feature extraction and matching.
- Increased error rate of mismatching due to out-bound blur.
- Time delay on face detection is high due to extraction of more number of features
- Personalization is crucial in the above scenario since there is a large inter-subject variability in the way gestures are executed.

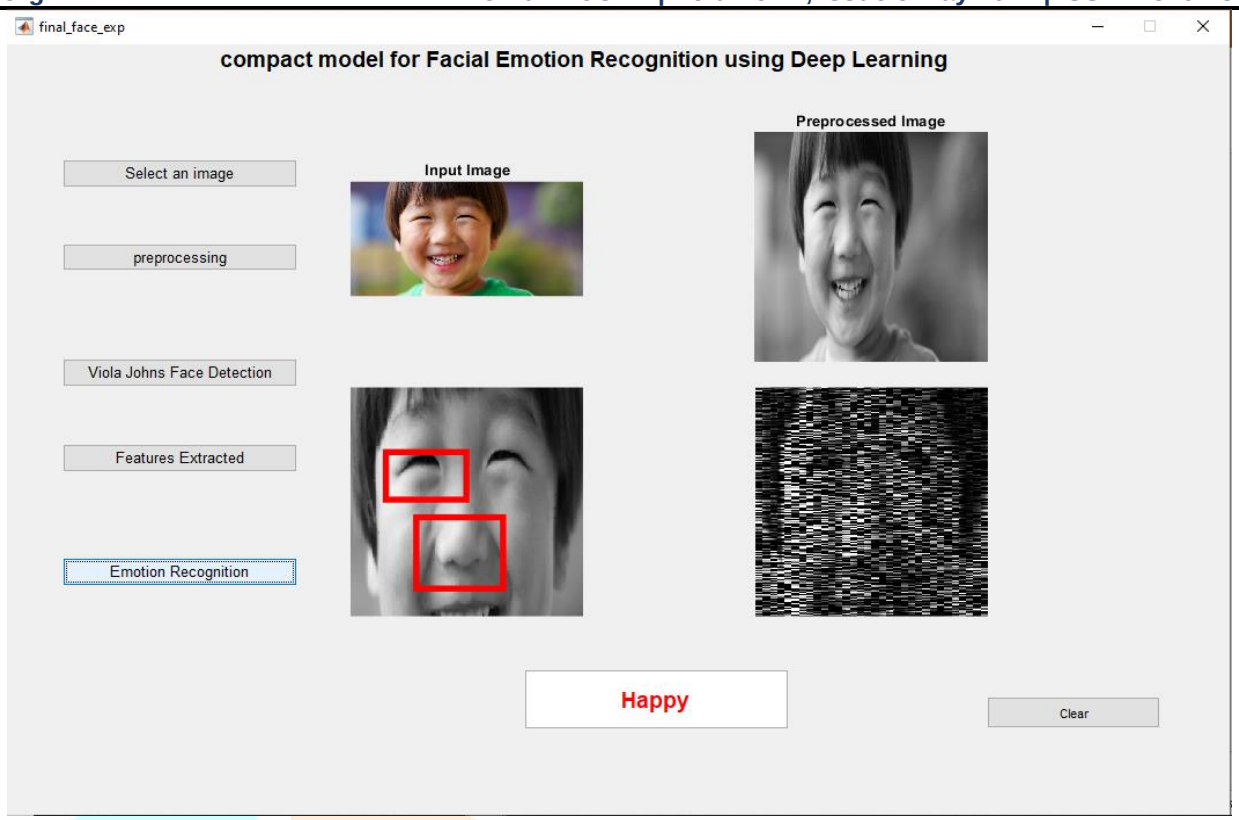

Fig. 6. Compact model for Facial Emotion Recognition

VI.CONCLUSION

In general, Facial emotion is one of the most currently used traits for diagnose mental and neurological diseases such as autism as well as predict the behavior of criminals. This paper has presented an approach of facial emotion recognition system, which was tested using the FER-2013 database of facial emotion images. Where VGG16 for extracting the feature from normalized facial emotion image. This technique is most promising in terms of performance and accuracy. Classic Neural Networks is applied as a classifier. Some important human emotions are happiness, sadness, joy, guilty, mightiness, cry, laughing, boring, fear, aggressiveness and jealousy and so on This research paper gives idea about fundamental concept how could A.I engineers able to engineer and implement human-like emotions in Humanoid with the help of "facial expression" and "speech/sound" recognition combination A.I engineering as emotion input to process and identify by Humanoid. In the field of image processing it is very interesting to recognize the human gesture for general life applications. For example, observing the gesture of a driver when he/she is driving and alerting him/her when in sleepy mood will be quite useful. Human gestures can be identified by observing the different movements of eyes, mouth, nose and hands. In this paper we are focusing on the human face for recognizing expression. Many techniques are available to recognize face. In this paper, face is detected using the Viola and Jones technique.

VII. REFERENCES

[a]. S. W. Lee, V. Mail, J. Jang, J. Shin, and J. Paik, "Noise-adaptive spatiotemporal filter for real-time noise removal in low light level images," IEEE Trans. Consumer Electron., vol. 51, no. 2, pp. 648-653, May. 2005.

[b]. E. Bennett and L. McMillan, "Video enhancement using per-pixel virtual exposures," ACM Trans. Graphics, vol. 24, no. 3, pp. 845-852, Jul. 2005.

[c]. H. Malm, M. Oskarsson, E. Warrant, P. Clarberg, J. Hasselgren, and C. Lejdfors, "Adaptive enhancement and noise reduction in very low lightlevel video," In Proc. IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, pp. 1-8, Oct. 2007.

[d]. Q. Xu, H. Jiang, R. Scopigno, and M. Sbert, "A new approach for very dark video denoising and enhancement," In Proc. IEEE International Conference on Image Processing, Hong Kong, China, pp. 1185-1188, Sept. 2010.

[e]. 2022 IEEE 2nd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA) Sabratha, Libya 23-25 May 2022.

[f]. Byoung chul ko., Department of Computer Engineering, Keimyung University, Daegu 42601, Korea; niceko@kmu.ac.kr; Tel.: +82-10-3559-4564, Sensors 2018, 18, 401; doi:10.3390/s18020401.

[g]. X. Dong, G. Wang, Y. Pang, W. Li, J. Wen, W. Meng, and Y. Lu, "Fast efficient algorithm for enhancement of low lighting video," In Proc. IEEE International Conference on Multimedia and Expo, Barcelona, Spain, pp. 1-6, Jul. 2011.

[h]. X. Zhang, P. Shen, L. Luo, L. Zhang, and J. Song, "Enhancement and noise reduction of very low light level images," In Proc. International Conference on Pattern Recognition, Tsukuba, Japan, pp. 2034-2037, Nov. 2012.

[i]. M. Kim, D. Park, D. K. Han, and H. Ko, "A novel framework for extremely low-light video enhancement," in Proc. IEEE International Conference on Consumer Electronics, Las Vegas, USA, pp. 95-96, Jan. 2014.

[j]. A. Loza, D. Bull, and A. Achim, "Automatic contrast enhancement of low-light images based on local statistics of wavelet coefficients," In Proc. IEEE International Conference on Image Processing, Hong Kong, China, pp. 3553-3556, Sept. 2010.

[k]. F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," Computer Graphics Forum, vol. 22, no. 3, pp. 419-426, Sept. 2003.

[l]. Tian Y., Kanade T., Cohn J. Handbook of Face Recognition. Springer; New York, NY, USA: 2005. Facial expression analysis; pp. 247–275.