International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

An Optimal MSER Descriptor Based Facial Expression Recognition System using Artificial Intelligence Method

Danish Meiraj, Ashish Oberoi

***Abstract:** In this Artificial intelligence based Facial emotion recognition system (AI_FERS) model, emotions of facial expressions through performing some predefined steps such as face acquisition, pre-processing of images, face detection, feature extraction & classification have recognized. In the pre-processing of the image phase include the approaches used for face detection is: Knowledge-based, Feature-based, Template-based, and Appearance-based approach. Binary image computation, Skin-color segmentation and morphological filtering, which includes the dilation of Binary images and Gray Images are being extensively applied. For features extraction from images MSER (Maximally Stable External Regions) technique is used. At the final step categorize of emotion into six parts: surprise, fear, disgust, anger, happiness, and sadness come as an outcome using ANN (Artificial Neural Network) technique. The efficiency of the system is examined based on performance parameters such as FAR, FRR, accuracy and execution time. The average accuracy of the AI_FERS model examined is about 98.23 %.*

***Keywords:** Facial expression, feature descriptor, deep neural network.*

I. INTRODUCTION

Nowadays, the application of emotional recognition in the "Human-Machine/Computer Intercommunication systems" is expanding and becoming more and more important. Emotion recognition is the procedure of recognizing human emotions dependent on Vocal Sounds, Hand Signals, and Facial Gestures. According to the "Psychological Hypothesis", Humans show their fundamental emotions states in the six classes;

Anger Joy Surprise Disgust Sadness Fear

When Human beings' express signals by change in the facial muscles motion and tone of the speech.

Revised Manuscript Received on January 30, 2020.
* Correspondence Author
Danish Meiraj, Department of Computer Science & Engineering, RIMT University, Gobindgarh, Punjab, India. Email ID: er.danishmeiraj@gmail.com
Dr. Ashish Oberoi*, Department of Computer Science & Engineering, RIMT University, Gobindgarh, Punjab, India. Email ID: a_oberoi01@yahoo.co.in

Humans easily recognize objects like emotional signals by processing the available statistics obtained by the Ears and Eyes. An attempt has also been made to design a facial expression recognition system that recognizes six emotions called Anger, Joy, Surprise, Disgust, Sadness, and fear 1, 2, 3. Inspired by these thoughts, **De Silva et al.** have carried out research using 18 people that have been used to identify emotion by utilizing both set Visual and Auditory information individually obtained from a Video and an Audio dataset. From the experiment, it has been carried out that few emotions like Sadness and Fear were clearly distinguished by means of audio. Whereas, the few emotions Anger and Happiness were clearly recognized through video 4.
Facial expression of emotion is one of the most significant tools for interaction, humans utilize universal gestures to express their emotions responses and intentions 5. Achieving high recognition rates using a single artificial Neural network (ANN) as a feature classification schemes is computationally complex and challenging. The newest machine learning framework designed for emotions identification using MSER as a feature extraction approach, to distinguish facial emotion from the sample images 6, 7. A number of techniques have been applied to design an emotional recognition system, but, most of the work was focused on hand engineering features. Due to the variation in picture size and availability of different dataset, deep neural network is the best and appropriate technique for image processing. ANN has the capability to handle simple spatial images 8.

II. RELATED WORK

The accuracy achieved in average up to 97.10 %; this accuracy has been improved by using ANN for classification purpose 9. **Lopes et al. (2018)**, have proposed a mechanism to categories the facial expressions of the old aged and presented the differences among the facial expressions of old age and another group of age along with the schemes to solve these differences. To extract the faces, Viola-Jones and Haar Features have been utilized and to extract the feature of faces used Gabor Filter mechanism. In this work classify the emotion by using Multiclass Support Vector Machine, produces accuracy in the detection of neutral state, happiness, and sadness correspondingly 90.32%, 84.61%, and 66.6% in case of old age group 3. **Ko, B. (2018)**, have focused on research using facial pictures solely due to visual signals is a unique main method of social exchange information. Have offered a short overview of FER-related research carried out over the previous decades.

In the first place, Standard FER methods are explained included with a summary of the representative classifications of FER schemes and its primary algorithms.
Deep-learning is contingent on Facial Emotion Recognition Systems are utilized for Hierarchical Deep Neural Networks (HDNN) which enables "end-to-end" (E2E) learning 10. **Chen et al. (2018)**, have proposed Deep Sparse Auto Encoder is exploited for studying face characteristics that determines the thinly scattered of invisible parts for high-level learning constructions. In the meantime, Softmax Regression (SR) is applied to arrange expression functionality to resolve the issue of DNN that suffered from challenges of learning-capability and computational-complexity 11. **Mehta et al. (2018)**, have provides a short explanation of the different emotion identification methods and methods. This task includes a brief overview of the databases regarded as information sets for algorithms that detect feelings through face expressions. Afterward, Mixed Reality Technology (MRT) is used in “HoloLens” devices for noticing emotion recognition. Also provides an overview of its sensors, their emotional recognition. Finally, this work concluded by comparing the “HoloLens” emotion identification outcomes with a frequent webcam 12. **Zhentao et al. (2017)**, have addressed a facial expression emotion recognition dependent on human-robot interaction (FEER-HRI) scheme that is a four-layer structure is intended. Furthermore, a technique of recognition of facial emotions based on 2D-Gabor, standardized extraction of LBP features and various classifiers of ELM is provided, which is applied to the FFER-HRI scheme 13.

III. PROPOSED AI_FERS MODEL

In this section, the process of work carried out to design efficient facial expression recognition is provided in detail. In the AI_FERS model, facial emotion recognition system using MSER descriptor along with the ANN approach has been provided. ANN is used as a classifier to train the facial emotion recognition system.

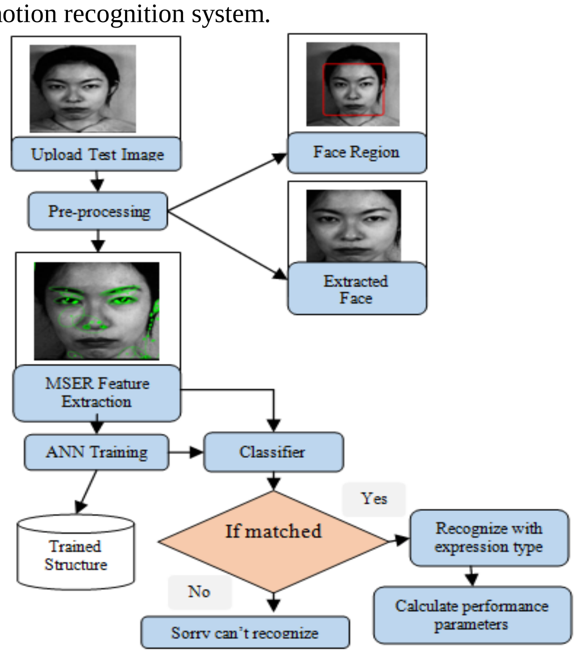

Fig. 1. AI_FERS (Artificial intelligence based Facial emotion recognition system) Model

In ANN, the first part is training and second is classification, and ANN is trained by appropriate feature sets of facial expression images, so, in classification section, test data can be easily classified with the help of trained structure. At the last of facial emotion recognition system, the performance metrics of AI_FERS model will be calculated with the comparison of AI_FERS model with the existing work. Fig. 1 The structural outline represents the entire framework of the proposed “Artificial intelligence based Facial emotion recognition system (AI_FERS) model” in Fig. 1.

A. MSER Feature Extraction Approach

Kimmel, (2011) have presented The MSER is an area of interest detector and shape descriptor because its area may be much larger than another point of interest method (such as Harris or the function from accelerated segmentation test (FAST) 14. The main aim of developing MSER detector was to address the parallax correspondence in broad baseline stereo systems. This method includes sorting pixels into a group of regions, which is based on a binary intensity threshold 1. An area with similar pixel values within the threshold range. The pixel value, which lies in the considerable range of pattern, is taken as maximally stable component. The image obtained after applying feature extraction technique is displayed in Fig. 2.

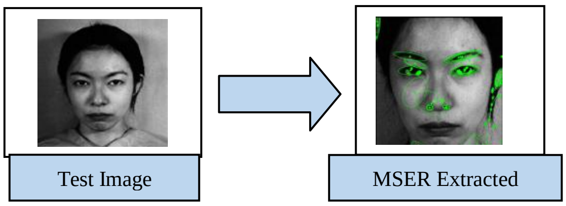

Fig. 2: Feature Extracted Image

B. Artificial Neural network Approach

ANN is utilized to train the designed model based on the feature extracted images. The ANN is a 3-Tiers structure containing input, hidden and output tiers are utilized to train the system based on the images 15. The designed structure is shown in Fig. 3.

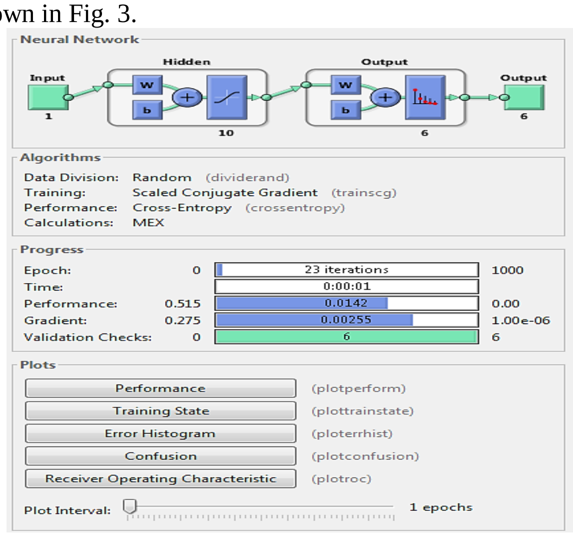

Fig. 3. ANN structure

Retrieval Number: B7522129219/2020©BEIESP
DOI: 10.35940/ijitee.B7522.019320
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

As shown in Fig. 3, for a single input variable, ANN provides six different outputs. Also, in the hidden layer there are 10 numbers of neurons that are passed to obtain the desired output.

IV. RESULT AND DISCUSSIONS

The capability of the AI_FERS system is determined on the basis of the following parameters.

1. False acceptance ratio (FAR)

$$\mathrm{FAR} = \frac{(\text{Total feature} - \text{Falsely accepted feature})}{\text{Total feature}}$$

2. False rejection rate (FRR)

$$\mathrm{FRR} = \frac{(\text{Total feature} - \text{Falsely rejected feature})}{\text{Total feature}}$$

3. Recognition rate

The efficiency of AI_FERS model except error is known as the recognition rate.

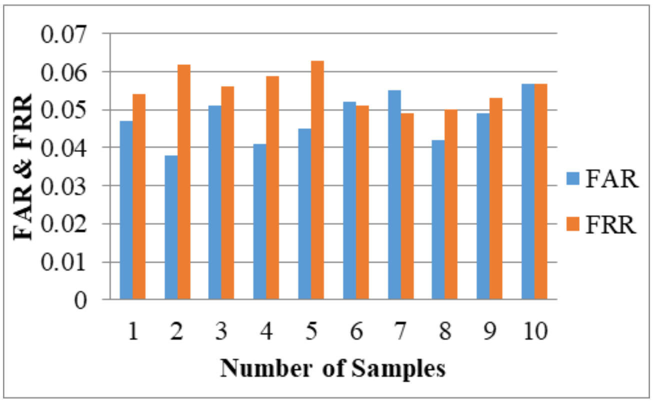

Fig. 4. FAR and FRR

FAR and FRR parameters measured for the AI_FERS model are shown in Fig. 4. The blue bar and the red bar represent the values for FAR and FRR respectively. The experiment has been performed for 10 different samples as denoted by the x-axis of the graph. The average value of FAR and FRR for the AI_FERS model measured is 0.0477 and 0.0554 respectively.

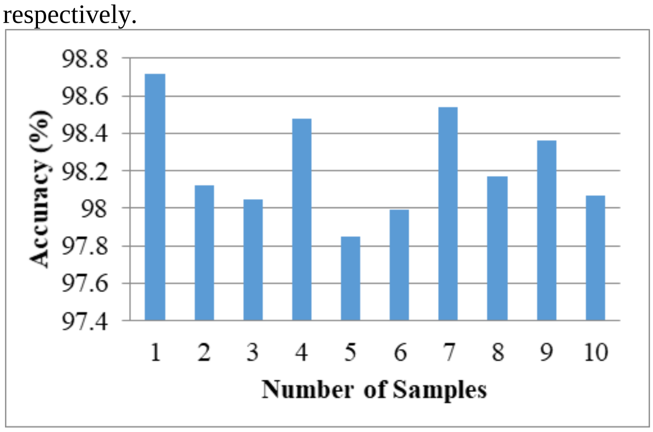

Fig. 5. Accuracy

The detection accuracy measured for the AI_FERS model is shown in Fig. 5. The horizontal-axis and the Vertical-Axis represents the percentage of measurement accuracy and number of samples respectively. The average accuracy measured for the designed facial emotion recognition system is 98.23 %.

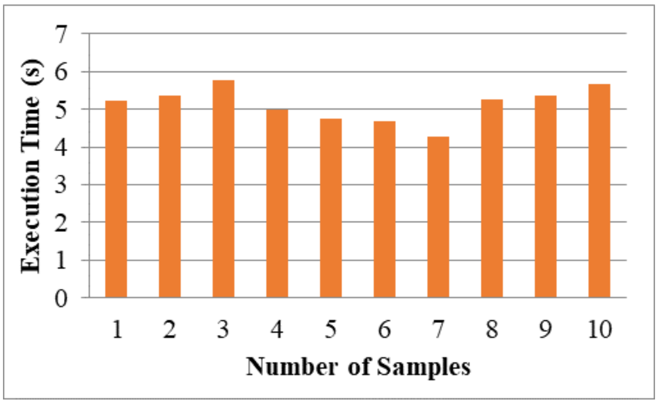

Fig. 6. Execution Time

The execution time measured for the AI_FERS model for different 10 numbers of samples is shown in Fig. 6. The average execution time analyzed for the entire work is 5.138s. The comparison of detecting accuracy against the existing work performed by Tsai and Chang is shown in Fig. 6.

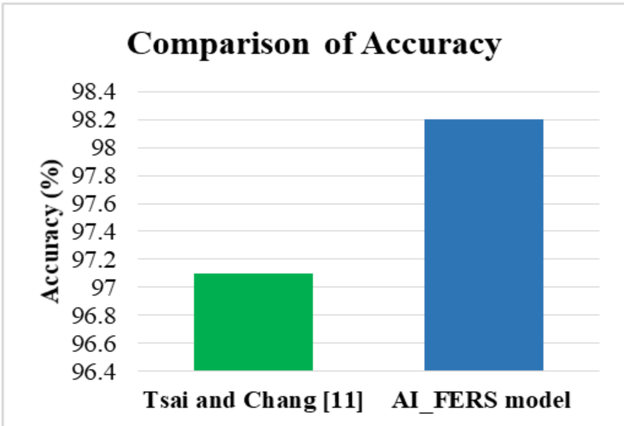

Fig. 7. Comparison of Accuracy

The comparison of the accuracy of the AI_FERS model with existing work is shown in Fig. 7. From the figure it is clear that the average detection accuracy obtains for the proposed and the existing work are 98.23 and 97.10 respectively. Thus, there is an improvement of 1.16 % has been attained compared to the existing work.

V. CONCLUSION

Facial expression recognition system has been designed and implemented by utilizing deep neural networks. The emotions are sorted into six classes, including; Anger, Joy, Surprise, Disgust, Sadness, and fear from face images datasets. To carry out the particular operation on the image data set firstly makes the data suitable to use for this purpose Pre-processing has been done. From the experiment results performed in MATLAB simulator and the parameters such as FAR, FRR, accuracy and execution time are measured for six different facial emotions. The average executed values for FAR, FRR and accuracy observed are 0.0477, 0.0554 and 98.23 % respectively with the average execution time of 5.138s.
Also, to show effectiveness of the AI_FERS model with respect to the existing work,

it is evident that the proposed model performs better with enhanced accuracy of 1.16 % compared to the existing work.

REFERENCES

1. K. Prasada Rao, Chandra Sekhara Rao & N Hemanth Chowdary. (2019). An integrated approach to emotion recognition and gender classification. Journal of Visual Communication and Image Representation, 60, 339–345. [CrossRef]
2. Busso et al. (2004). Analysis of emotion recognition using facial expressions, speech and multimodal information. Proceedings of the 6th International Conference on Multimodal Interfaces - ICMI '04. [CrossRef]
3. Lopes, N., Silva, A., Khanal, S. R., Reis, A., Barroso, J., Filipe, V., & Sampaio, J. (2018, June). Facial emotion recognition in the elderly using an SVM classifier. In 2018 2nd International Conference on Technology and Innovation in Sports, Health and Wellbeing (TISHW) (pp. 1-5). IEEE. [CrossRef]
4. De Silva, L. C., Miyasato, T., & Nakatsu, R. (1997, September). Facial emotion recognition using multi-modal information. In Proceedings of ICICS, 1997 International Conference on Information, Communications and Signal Processing. Theme: Trends in Information Systems Engineering and Wireless Multimedia Communications (Cat. (Vol.1, pp.397-401). IEEE. [CrossRef]
5. Jain, D. K., Shamsolmoali, P., & Sehdev, P. (2019). Extended deep neural network for facial emotion recognition. Pattern Recognition Letters, 120, 69–74. [CrossRef]
6. Dagar, D., Hudait, A., Tripathy, H. K., & Das, M. N. (2016, May). Automatic emotion detection model from facial expression. In 2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT) (pp. 77-85). IEEE. [CrossRef]
7. Tarnowski, P., Kołodziej, M., Majkowski, A., & Rak, R. J. (2017). Emotion recognition using facial expressions. Procedia Computer Science, 108, 1175-1184. [CrossRef]
8. Ratliff, M. S., & Patterson, E. (2008, March). Emotion recognition using facial expressions with active appearance models. In Proc. of HRI. [PDF] [GoogleScholar]
9. Tsai, H. H., & Chang, Y. C. (2018). Facial expression recognition using acombination of multiple facial features and support vector machine. Soft Computing, 22(13), 4389-4405. [PDF] [CrossRef]
10. Ko, B. (2018). A brief review of facial emotion recognition based on visual information. sensors, 18(2), 401. [CrossRef]
11. Chen, L., Zhou, M., Su, W., Wu, M., She, J., & Hirota, K. (2018). Softmax regression based deep sparse autoencoder network for facial emotion recognition in human-robot interaction. Information Sciences, 428, 49-61. [CrossRef]
12. Mehta, D., Siddiqui, M., & Javaid, A. (2018). Facial emotion recognition: A survey and real-world user experiences in mixed reality. Sensors, 18(2), 416. [CrossRef]
13. Zhentao et al. (2017). A facial expression emotion recognition based human-robot interaction system. IEEE/CAA Journal of Automatica Sinica, 4(4), 668-676. [CrossRef]
14. Ron Kimmel, Cuiping Zhang, Bronstein, A. M., & Bronstein, M. M. (2011). Are MSER Features Really Interesting? IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(11), 2316–2320. [CrossRef]
15. Mollahosseini, A., Chan, D., & Mahoor, M. H. (2016). Going deeper in facial expression recognition using deep neural networks. 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). [CrossRef]

AUTHORS PROFILE

Danish Meiraj, is a student pursuing Masters of Technology in Computer Science and Engineering from RIMT University, Punjab, India. His areas of interest include Image Processing, Artificial Intelligence, Machine Learning, and Networking.

Dr. Ashish Oberoi, is working as a Professor (HOD) in Department of Computer Science and Engineering from RIMT University, Punjab, India. He was awarded his Ph.D. His areas of interest include Image Processing, Internet of Things, Artificial Intelligence and Machine Learning He has published 28 papers in various International/ National Journals.

Retrieval Number: B7522129219/2020©BEIESP
DOI: 10.35940/ijitee.B7522.019320
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)