International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 06 | Jun 2022 www.irjet.net p-ISSN: 2395-0072

A Real-Time Emotion Recognition from Facial Expression Using Conventional Neural Network Architecture

Thaneshwari Sahu[1], Shagufta Farzana[2]

[1]PG(M.Tech) Student, Dept. Of Computer Science & Engineering, Dr. C. V. Raman Institute Of Technology kota Bilaspur, Chhattisgarh, India
[2] Assistant Professor Shagufta Farzana, Dept. Of Computer Science & Engineering, Dr. C. V. Raman Institute Of Technology kota Bilaspur, Chhattisgarh, India

---

Abstract - *Emotion is a important issue in various fields such as brain science neuroscience and wellbeing. Expression investigation could be important for result of mind and mental issues. Recently, deep learning has developed more in the area of picture classification. In this survey, we present a Convolutional Neural Network (CNN) for facial expression recognition. We utilized face emotion data-set and real-time video as an input image for facial emotion recognition system and utilizes Haar classifier for face detection. In this paper, we accomplished accuracy of 94.61% for classification of 7 unique emotions through facial expressions.*

Key Words: **Convolutional Neural Network; Deep Learning; Emotion Recognition; Facial Expressions, Real Time Detection.**

1. INTRODUCTION

Facial Expression Recognition framework has come in design in late time as the connection between the human and PC machine has expanded immensely as well as has a broadrange of utilization of it in practical existence [1]. Video-based FER has been begun specialization, and the features of the expressions on the essence of people is finished through this framework. Video-based a large number of researchers has determined on FER frameworks, as this architectureattend with a critical issue of fitting the hole found in the visual elements in the video and the feelings seen on face in recordings [2]. Broad scope of use of video-based FER should be visible in businesses like mechanical technology, clinical area, driver security and so on [3]. The essential facial expressions that should be visible arranged are sad, surprised, disgust, angry, happy, fearand neutral [4].

The face emotion recognition architecture should be noticeable in two ways, and it incline to be ordered in prospect of the description of the elements [5]. The firstsegment depends on the static pictures, and the face emotions are perceived from this spatial data existent in the static pictures. It is by all accounts a simple undertaking as contrasted with the second segment of FER system. The next one sort is the video-based FER in which the face emotions is received from the video clasps or continuous recordings. This sort of FER is examining as the common work to do as in this the work is in the middle of between the two adjoining video outlines and the worldly in the middle of between these two frames is completed. So, this work allotted the spatial and transient information present in the video arrangement and handling it instead of the standard static pictures. The video-based FER works in three phases, the first is video pre-management, sending with the extrication of visual-elements, and finally, the recognition of emotions is completed [6]. Trimming of pictures of the face is finished in the pre-processing video stage, and the video outline facial pictures are utilized for the extraction of visual highlights. Eventually, the segmentingaccount is putting for the recognition of emotions.

A segment of the profound learning methods like CNN have been in need for gaining the highlights from the video for face emotion recognition [7]. CNN provides higher accuracy to the arrangement assignment for the monstrous video data-sets. The data-sets are having noise as well as are accessible in less amount. The information like spatial and fleeting is cooperative for the showing of the recordings having differentemotions[8][9].

As of late the work in video-based-FER has been done in light of the spatial and transient highlights and 3D-convolutions. The situation found in 3D-convolution is that there is trouble found in scaling it, though the spatial and worldly can perform on just brief recordings [10-14]. So, the abuse of the extensive element has been a fundamental key for the FER from the video[15].

Here in our paper, we have introduced a methodology utilizing CNN architecture that is prepared start to finish for the collection and extrication of the elements present in the static pictures in the middle of between the fleeting surges of video. Here in the center of the descriptors, we have carried out a spatial-worldly conglomeration layer named as Emotional-VLAN that is working for the clustering in the static pictures we have and the recognition of the activities in the recordings. The varioussegments of the CNN model assist in the conglomeration of highlights and the FER with entrusting are helped out through the penultimate layer on the CNN-model.

2. Literature Survey

Table -1: Comparisions of various techniques and method used in existing system

S.No.	Ref. Year	Approach	Dataset	Method Used	Algorithm Used	Advantage	Accuracy
1	[16] 2021	an effective technique of classification and feature selection for Face Emotion Detection from sequence of face pictures.	CK+	minimum features selected by chisquare	decision-making board, a random forest, radial bias, MLP, SVM, KNN and MLP.	support a consistency-assessment of various controlled algorithms to identify face emotion based on limited chi-square characteristics.	94.23 %
2	[17] 2021	Effective method for Face Emotion dataset and classify number of facial images	CK+	Reliever-F approach	Decision Tree, SVM, KNN, Random Forestry, radial basis functions and MLP	Concentrating on the usage of a minimal amount of assets, it concentrates at identification the exactness of six classifiers based on the Reliever-F techniques for function attributes	94.93%
3	[18] 2020	To achieve effective modality fusion.	SAVEE, eNTERFACE'0 5, and AFEW	LBP-TOP and spectrogram	Fuzzy-Fusion based two-stage neural-network (TSFFCNN)	The function in which every modal data provide to expression investigation is strongly im-balanced is supported well by TSFFCNN.	SAVEE 99.79%,eNTERFACE' 05 90.82%, AFEW 50.28%
4	[19] 2020	The eight basic emotions of human facial to be investigated	CK+	Co-relation gains ratio and data gain	using Naïve Bayes, decision tree, Naïve Bayes, decision tree, KNN and MLP algorithms	Provided FER by comparison method based on three methods of characteristics choose: co-relation, gains ratio, and information gain to examine the max outstanding features of facial pictures	-
5	[20] 2020	To improve the impromptu recognition of face micro-expression by world wise extrication method by hand	CASMEII, SMIC and SAMM.	The unique methodology of the facial motion is expansion to intake number - rank pooling to input critical pictures.	CNN utilizes for the recognition of microexpressions in critical images requested	Comprehensive as well as effective but plane technique analysis as well as classification of ME	CASMEI, 78.5%, SMIC 67.3%, SAMM 66.67%

6	[21] 2020	To permit a particular function to take input from the action units' features (AUs) (eyes, nose, eyebrows, eyes, nose, mouth).	CK+ OuluCASIA MMI, AffectNet	-VGG16 structure - Standardization cluster (BN) layers Bi-directional Short-Memory (BiLSTM)	- CNN as well as RNN	Experiment result indicate that the SAANet is improved rapidly in comparison to other trained techniques.	Dataset CK+ is 99.54%, Oulu CASIA MM is 88.33%, AffectNetis 87.06%
7	[22] 2020	Select the difficult research of the facial	CK+ Oulu CASIA and MMI	-3D-InceptionResNet network architecture	A strain and Spatial-Temporal Focus Module (STCAM)	The result find and suggest that the method is more or less achieved than state of the art techniques	CK+ 99.08%, CASI 89.16%,
8	[23] 2020	To achieve the growth of normal emotion captured from different region of the facial.	MMI AFEW BP4D	(UGMM), (MIK), (IMK)	Dynamic kernelbased facial expression representation	For any emotion detection method based on measurement time and a dynamic kernel is an essential selection	BP4D 74.5%
9	[24] 2019	To perform better dynamic features, learning -Enlarge the capacities between practice from easy confused features	-	enhanced deep the pooling-covariance and residual network-units	-a newly DNN for face emotion - a new loss features is proposed to enhanced the gap between samples	The neural-network construct will preserve the loss of gradients	86.47%
10	[25] 2019	To improve the accuracy.	(CK+, MMI, AFEW)	Dropout layer, Island layer, Softmax layer Layer, Stem Layer, 3D LayoutResNets, GRU layer	- CNN video Face Emotion Recognition technique	It avoid the consequences of facial emotion features and adapt and gain max important of identification as well as generalization	CK+ 94.39%, MMI 80.43%, AFEW 82.36%

3.Problem Identification

With reference to literature survey, the author has proposed a mechanism for facial expression recognition using SVM and pre-processing activity. But the pre-processing activity just crops the image into pre-defined size and does nothing more than that in the pre-processing activity. The SVM alone is used for classification which is much slower than Neural-Network Approach.

4.Praposed Methodology

Our system proposed facial-expression recognition classifier system which can take picture data-set via two methods to recognized emotion-

i) Upload facial image from system

ii) Real-time video system

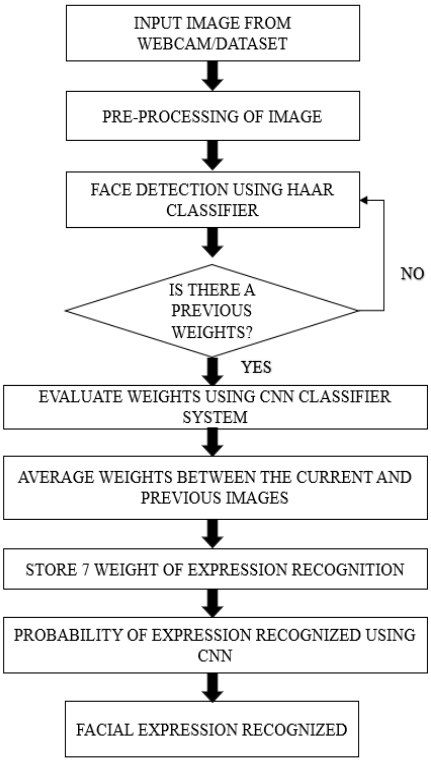

Fig -1: Proposed system architecture for real-time image/ dataset image for facial expression recognition

(a) Image Processing

The face border, maxessentially, was identified utilizing the Haar-Cascade library from the pictures. Then, these identified rectangular-facial emotions were cut and recorded to anidentical size. Additionally, the pixel-values in the pictures were shifted over totally to gray pictures size of 64x64 to be put in neural-networks.

(b) Image Dataset

There are two-methods to take input pictures ae as follows:

(i) Facial expression recognition (FER) data-set

(ii) Real-time image captured via webcam

(a) (b)

Fig -2: Input Image (a) FER data-set (b) Real time Image

(c) Face Detection

For face recognition, we utilized a webcam with the Open CV face catch program and the borders from the face preparing set haarcascade_frontalface_alt.xml. The recognition was programed in Python to accomplish a discovery pace of around 10 pictures each second. This progression was trailed by grayscale change and resizing for each picture to right away and immediately input the caught pictures into the trained framework.

Face detection is the essential advance for any FER framework. For face location, Haar overflows were utilized (Viola and Jones, 2001). The Haar overflows, otherwise called the Viola Jones finders, are classifiers which identify an article in a picture or video for which they have been prepared. They are prepared over a bunch of positive and negative facial pictures. Haar overflows have ended up being a productive method for object discovery in pictures and give high accuracy.

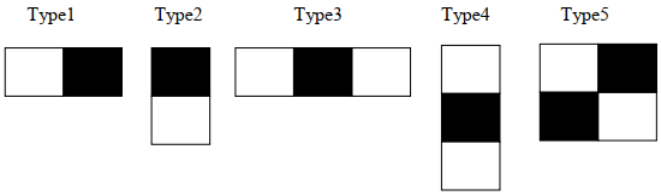

Fig -3: Haar characteristics in Viola-Jones algorithm

(d) Emotion Classification

The framework orders the image into one of the seven different expressions - Happiness, Sadness, Anger, Surprise, Disgust, Fear, and Neutral by using CNN.

The emotions arrangement step comprises of the following stages-

(i) Splitting of data

Splits the data into 3 categories

- Training (use for generation of model)
- Private test (use for generation of model)
- Public test (use for evaluating the model)

(ii) Training and Generation of model

The neural network framework consists of the following layers-

- Conventional layer - In the convolution layer, a haphazardly launched learnable channel is slid, or convolved over the input.
- Max pooling - The pooling layer is utilized to decrease the spatial size of the info layer to bring down the size of information and the calculation cost.

- Fully connected layer - In the completely associated layer, every neuron from the past layer is associated with the result neurons by which input image is to be classified.

- Activation function - Activation function is utilized to reduce the overfitting. In the CNN architecture, the ReLu activation function has been used.

$$f(x) = \max(0, x) \qquad .1$$

Equation 1: Equation of ReLu Activation Function.

- Softmax - The softmax function takes a vector of N-real numbers as well as normalizes that vector into a range of values between (0, 1).

- Batch Normalization - The batch-normalizer speeds up the preparation procedure and implements a transformation that regulate the mean-activation close to 0 as well as the activation standard-deviation close to 1.

(iii) Using model to evaluate FER for real-time video input image/ dataset

The model creates during the training process comprise of pretrained weights and values, which can be utilized for implementation of a new FER problem.

As the model created already contains weights, FER becomes faster for real-time images. The CNN framework is shown in fig. 4

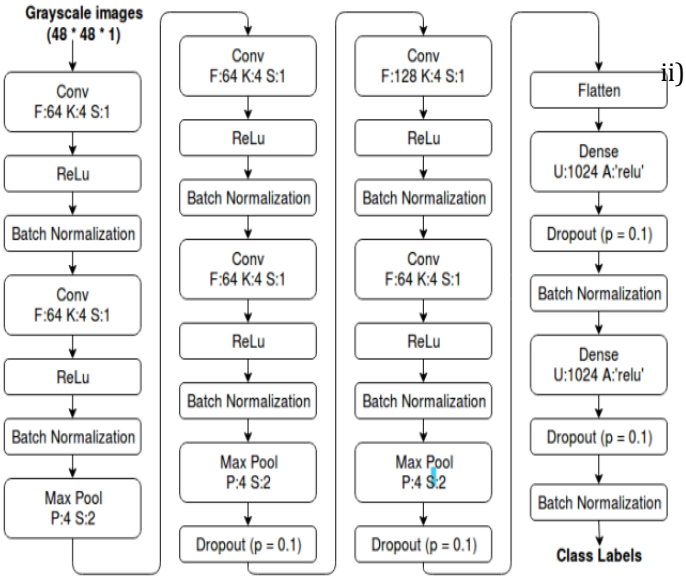

Fig -4: CNN framework

5. Result and Discussion

We have analyzed image of FER dataset and also real time video input captured via webcam.

i) Prediction for FER dataset

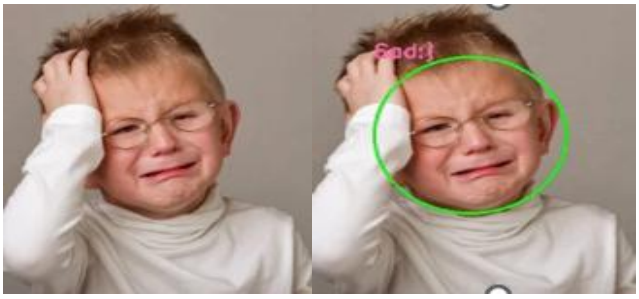

Fig -5: Shows the of prediction of FER of the input image

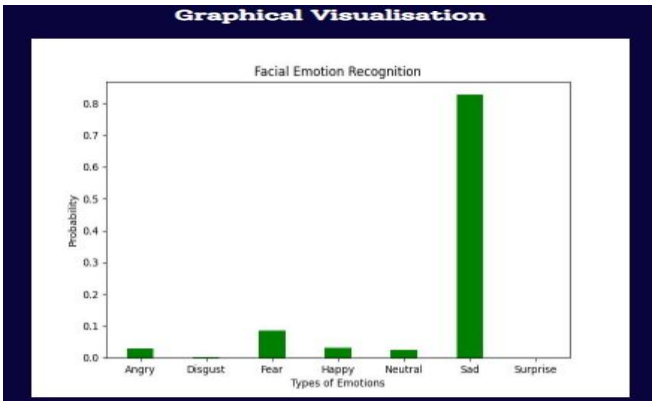

Fig -6: Graphical visualization of predicted sad FER

ii) Prediction for real-time input image captured from webcam.

Fig -7: Prediction of real-time facial expression image

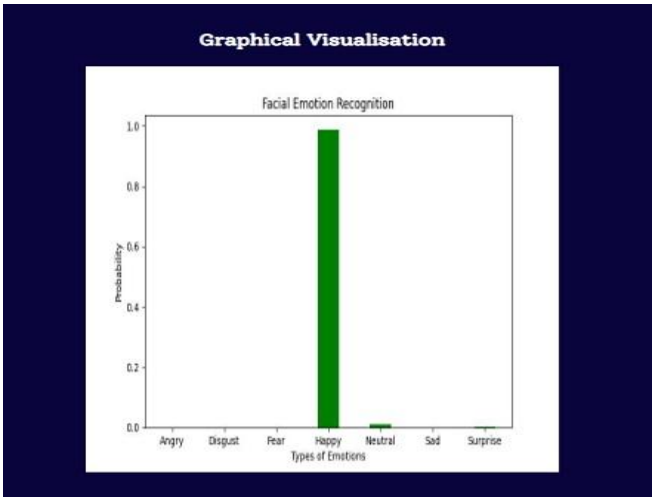

Fig -8:.Graphical representation of real-time of FER image

This section presents the values obtained after performing experiments. We have utilized CNN, and Haar Classifier algorithm for facial expression recognition. Table II. Shows the accuracy obtained from 2 different frameworks and compared with ours.

Table -2: Accuracy obtained from different frameworks

S. NO	Author	Year	Technique Used	Data set Used	Accuracy
1	Mahmood M. R. et. al	2021	MLP, SVM	CK+	94.23 %
2	Proposed System	2022	Voila Jones, Artificial Neural Network and Principal Component Analysis	Facial expression image and real-time video image	94.61 %

6. Conclusion

We have proposed an efficient and accurate facial expression system using 7 different emotions of facial expression image as a database and also our system can investigate emotion from the real-time video image. Proposed method is the combination of Haar classifier for face recognitionand convolutional neural network for arrangement of facial datasets. The performance is compared with the existing method and is more efficient than all. The proposed algorithm achieves approximately 94.61% accuracy gain.

In this project, most of the work is done over the facial expression database as well as real-time video input image. The databases are self-sufficient for training and testing of framework. But under some circumstances the classifier fails and hence some improvement while training. So, the future scope might be gathering more data-set for analysis.

References

[1] S. Wang, S. Member, B. Pan, H. Chen, Q. Ji, "Thermal Augmented Expression Recognition," pp. 1–12, 2018.

[2] X. Pan, S. Zhang, W. Guo, X. Zhao, Y. Chen, and H Zhang, "Video-Based Facial Expression Recognition using Deep Temporal – Spatial Networks" vol. 4602, 2019, doi: 10.1080/02564602.2019.1645620.

[3] T. Ben, A. Radhouane, M. Hammami, "Facial-expression recognition based on a low-dimensional temporal feature space," 2017.

[4] L. Greche, M. Akil, R. Kachouri, N. Es, "ORIGINAL RESEARCH PAPER A new pipeline for the recognition of universal expressions of multiple faces in a video sequence," J.Real-Time Image Process.,no. 0123456789, 2019, doi:10.1007/s11554-019-00896-5.

[5] M.M. Hassan, M. Alrubaian, G. Fortino, S. Member, "Facial Expression Recognition Utilizing Local Direction - based Robust Features and Deep Belief Network," vol. 3536, no. c, 2017, doi: 10.1109/ ACCESS.2017.2676238.

[6] S. Lee, S.H. Lee, K.N.K. Plataniotis, Y.M. Ro, "Experimental investigation of facial expressions associated with visual discomfort: Feasibility study towards an objective measurement of visual discomfort based on facial expression," no. c, 2016, doi: 10.1109/JDT.2016.2616419.

[7] E.S. Networks, K. Zhang, Y. Huang, Y. Du, S. Member, "Facial Expression Recognition Based on Deep," vol. 14, no. 8, 2017, doi: 10.1109/ TIP.2017.2689999.

[8] Alshamsi, H., Kepuska, V., & Meng, H., "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.

[9] Fathallah, A., Abdi, L., &Douik, A., "Facial expression recognition via deep learning", In 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA) (pp. 745-750), IEEE.

[10] Kumar, G. R., Kumar, R. K., & Sanyal, G., "Facial emotion analysis using deep convolution neural network", In 2017 International Conference on Signal Processing and Communication (ICSPC) (pp. 369-374), IEEE.

[11] Amin, D., Chase, P., & Sinha, K.," Touchy Feely: An Emotion Recognition Challenge", Palo alto: Stanford.

[12] Shan, K., Guo, J., You, W., Lu, D., &Bie, R., "Automatic facial expression recognition based on a deep convolutional neural-network structure", In 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA) (pp. 123-128), IEEE.

[13] Kulkarni, K. R., &Bagal, S. B. (2015, December), "Facial expression recognition", In 2015 Annual IEEE India Conference (INDICON) (pp. 1-5). IEEE.

[14] Minaee, S., &Abdolrashidi, A. (2019, "Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network", arXiv preprint arXiv:1902.01019.

[15] Patil M.N., Brijesh Iyer, Rajeev Arya, "Performance Evaluation of PCA and ICA Algorithm for Facial Expression Recognition Application", In: Pant M., Deep K., Bansal J., Nagar A., Das K. (eds) Proceedings of Fifth International Conference on Soft Computing for Problem Solving. Advances in Intelligent Systems and Computing, vol 436. Springer, Singapore.

[16] M. R. Mahmood, M. B. Abdulrazzaq, S. Zeebaree, A. K. Ibrahim, R. R. Zebari, and H. I. Dino, "Classification techniques' performance evaluation for facial expression recognition," Indonesian Journalof Electrical Engineering and Computer Science, vol. 21, pp. 176-1184, 2021.

[17] M. B. Abdulrazaq, M. R. Mahmood, S. R. Zeebaree, M. H. Abdulwahab, R. R. Zebari, and A. B. Sallow, "An Analytical Appraisal for Supervised Classifiers' Performance on Facial Expression Recognition Based on Relief-F Feature Selection," in Journal of Physics: Conference Series, 2021, p. 012055.

[18] M. Wu, W. Su, L. Chen, W. Pedrycz, and K. Hirota, "Two-stage Fuzzy Fusion basedConvolution Neural Network for Dynamic Emotion Recognition," IEEE Transactions on Affective Computing, 2020.

[19] H. I. Dino and M. B. Abdulrazzaq, "A Comparison of Four Classification Algorithms for Facial Expression Recognition," Polytechnic Journal, vol. 10, pp. 74-80, 2020.

[20] T. T. Q. Le, T.-K. Tran, and M. Rege, "Dynamic image for micro-expression recognition on region-based framework," in 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI), 2020, pp. 75-81.

[21] Liu, X. Ouyang, S. Xu, P. Zhou, K. He, and S. Wen, "SAANet: Siamese action-units attention network for improving dynamic facial expression recognition," Neurocomputing, vol. 413, pp. 145-157, 2020.

[22] W. Chen, D. Zhang, M. Li, and D.-J. Lee, "STCAM: Spatial-Temporal and Channel Attention Module for Dynamic Facial Expression Recognition," IEEE Transactions on Affective Computing, 2020.

[23] N. Perveen, D. Roy, and K. M. Chalavadi, "Facial Expression Recognition in Videos Using Dynamic Kernels," IEEE Transactions on Image Processing, vol. 29, pp. 8316-8325, 2020.

[24] G. Wen, T. Chang, H. Li, and L. Jiang, "Dynamic Objectives Learning for Facial Expression Recognition," IEEE Transactions on Multimedia, 2020.

[25] L. Deng, Q. Wang, and D. Yuan, "Dynamic Facial Expression Recognition Based on Deep Learning," in 2019 14th International Conference on Computer Science & Education (ICCSE), 2019, pp. 32-37.

BIOGRAPHIES

Thaneshwari Sahu, M.Tech in Computer Science & Engineering at Dr. C. V. Raman Institute Of Technology kota Bilaspur, Chhattisgarh, India

Shagufta Farzana Assistant Professor in Computer Science & Engineering at Dr. C. V. Raman Institute Of Technology kota Bilaspur, Chhattisgarh, India