The International Journal of Engineering and Science (IJES)
|| Volume || 11 || Issue || 12 || December || Pages || PP 25-31 || 2022 ||
ISSN (e): 2319-1813 ISSN (p): 20-24-1805

Comparative Analysis of Feature extraction and image enhancement combinations for facial expression recognition

Sushil kumar
[1]*Computer Science and Engineering, Maharaja Surajmal Institute of Technology, New Delhi, India*
Corresponding Author: Sushil Kumar

--ABSTRACT--
The aim of this work is to analyse a classifier for different image enhancement and feature reduction techniques for emotion classification. Therefore, Convolutional neural network and FER2013 dataset is considered for classifying facial expressions. The images from FER2013 were first enhanced using combination of brightness, sharpness, and contrast followed by features extraction using Principal reduction technique and independent component analysis. The features were then fed to CNN. The performance of classifier is analysed with different performances metrics. The analysis revealed that performance of CNN with different combination of image enhancement and feature reduction has impact in terms of accuracy and precision.

Date of Submission: 10-12-2022 Date of Acceptance: 25-12-2022

I. INTRODUCTION

Human face exhibits numerous expressions and these expressions helps in different ways to fetch useful information in uncertainty inducing conditions. Facial expression recognition (FER) system help extract correct information in such situations. Therefore, the practical necessity of FER has gained importance. The application of FER is not only limited to detecting expressions but contribute to other areas like mood analysis, fatigue detection, neuro marketing, medical fields, and behaviour analysis etc. Despite practical necessity and application, the study of FER is still facing various challenges. The major factor which is posing challenge is feature engineering which is further dependent on clarity of image. Another factor is datasets and techniques for data acquisition. The acquisition of label dataset does not guarantee the proper definition and justification of feature. This is due to that; the expressionsare dependent on individual. An image of similar person may have variations due to lighting, background, and head pose [1]. In the past researchers have identified emotion using geometric, textual, and statistical information [2]. The accuracy of FER is affected by other factors also like intensity of emotion, variation in face due to age, facial hair, presence of mask, spectacles etc. The accuracy is highly dependent on image resolution, size, contrast, brightness, and amount of noise in image etc. The objective of this work is as follows:

- The focus of this work is to apply image preprocessing for image enhancement.
- Extracting features of enhanced images and designing machine learning model for testing on datasets.

The results of models are compared considering different combination of image preprocessing and feature extraction techniques.

The remaining paper is organized as follows. Section 2 is literature survey to discuss related work in area of expression recognition and classification. The tools and techniques used are discussed in section 3. The results of proposed FER as given in section 4 and final section 5 presents the conclusion and future scope of the work.

II. LITERATURE REVIEW

Human beings can easily predict the facial expression by just observing the various movement of facial muscles and different formation formed due to this movement. Ekman and Friesen [3] developed the Facial Action Coding System (FACS) for describing facial expressions by action units (AUs). It states 32 atomic facial muscle actions, called as Action Units (AU). However, Hui Fang [4] presented a dynamic expression recognition system without using the action units. The important feature is frame with peak expression and is extracted from the video recording. The proposed method is effective with certain uncontrolled conditions.Ekman [5] classified facial expressions into one of six universal emotions i.e., happiness, sadness, fear, disgust, surprise, and anger.Abboud et al. [6] proposed an active appearance model for facial expressions recognition, model is further used to create expressions synthetically using two approaches.

PCA is used to merge the shape and grey level appearance whereas linear regression allows the construction of expression models. Results shows that the artificial expressions created synthetically by the proposed technique are closer to the real face expressions. Mlakar et al. [7] suggested histogram of oriented gradient and difference features vectors for recognizing facial emotions. The developed system reduces the size of feature vector and efficiency of support vector machine classifier is also increased. Edwin et al. [8] suggested temporal pattern of oriented edge magnitude features for FER. The results showed increased accuracy in comparison to state of art methods and while keeping low execution time. Shui-Hua Wang et al. [9] used stationary wavelet entropy as feature extractor and single hidden layer feedforward neural network for classification and achieved accuracy of 96.80%. Local pattern in facial regions and relation in these patterns are important to capture for expression recognition. Graph signal processing techniques is suggested [10] to represent the facial regions. Spectral graph wavelet transform is further used for creating feature descriptors. The experiment is conducted on CK+ and JAFFEE datasets to confirm the effectiveness. Khan et al. [11] proposed a memory and time efficient model for expression recognition in low resolution images which extract features in a pyramidal fashion. The model is reliable facial expression recognition system. P. Geetha et al. [12] has used combination of discrete wavelet transform and fuzzy to enhance images and classified expression using modified eye and mouthmap algorithm. Barman and Dutta [13] suggested new features set and expression recognition is done using distance and shape signatures. The features are further enhanced by augmenting statis-tical features like range, moment, entropy etc. The combinations of three features are test-ed on a Multilayer Perceptron to detect the corresponding facial expressions. Deep convolutional neural network is utilized by Li et al. [14]. Two CNNs are used for feature extraction and learning, later two features are concatenated and given as input to fully connected layers resulting into a new model. The proposed model has enhanced the accuracy of facial expression identification.

III. AUTOMATIC FACIAL EXPRESSION RECOGNITION

A facial expression recognition system has many stages which involves pre-processing, face detection or finding Region of interest, feature extraction and finally expression classification as shown in Fig. 1.

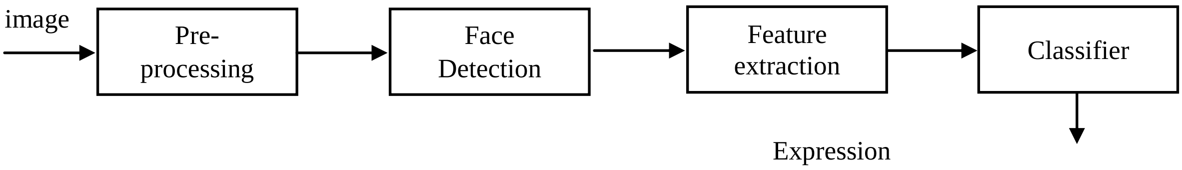

Figure 1: Facial Expression recognition system

Image preprocessing basically eliminates the effects which affects the system accuracy like noise, background etc. Preprocessing is also done to enhance image, alignment of face, pose correction etc. for effective feature extraction. The next phase is detecting region of interest (ROI) and in case of FER it is face detection, viola & jones is most popular and commonly used for detecting face. The location of eyes, lip, midpoint of lower lip, nose etc. are unique locations on face and represent the facial landmarks. The effective detection of facial landmarks results in efficient FER. Further, the ROI of pixel data of facial landmarks is converted to different representation to reduce the input shape. The input shape is sometimes exceptionally large and may require to be reduced so that only dominant and enough features are considered for classification. Numerous feature extraction techniques are reported in literature like, Principal Component Analysis, Linear Discriminant Analysis [15], Local Binary Patterns, and Local Gradient Code [16]. Finally, classifier classify the features into expressions. Commonly used classification methods are Support Vector Machine, Nearest Neighbor [17] and Convolution neural network etc.

Image Preprocessing

The considered image preprocessing operations on an input image are brightness, contrast, and sharpness to enhance image for effective feature extraction.

- Brightness of image refer to its darkness and lightness. To alter brightness a constant value is added to each pixel of image.
- Contrast is the relation between the darkest and brightest part of an image.
- Sharpness is defined as contrast between the edges of an object in an image.

A factor of 1.5 is under consideration for brightness, contrast and sharpness in the proposed model FER as shown in fig. 2.

Feature extraction

Feature extraction is locating and depicting of true features of interest from an input image for further processing. The feature extraction methods are categorized as texture feature-based method, edge-based method, global and local feature-based method geometric feature-based method and patch-based method.

Principal component analysis (PCA)

Principal component analysis is a feature reduction method where large dataset is reduced in term of number of variables or dimensions, by transforming large set of features into smaller one containing most of the information. PCA is computed as follows:

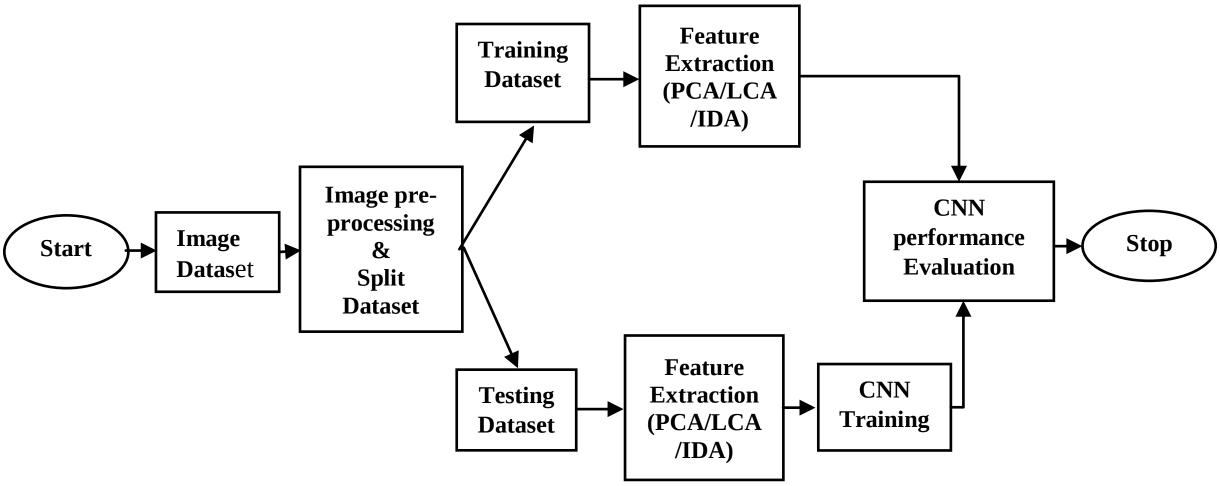

Figure 2: Flow chart of proposed methodology

1. **Standardize the data:** For the better performance all the features are required to be on the same scale. Standardization is done by using the calculation:

$$f_s^i = \frac{f^i + \mu_f}{\sigma_f} \qquad (1)$$

Where, μ_f is the mean of feature column and σ_f is the sample variance.
The result of this is all values having mean equal to 0 and standard deviation 1.

2. **Computing covariance matrix:** covariance is the measure of variation between two features. A positive variance represents the simultaneous increase and decrease of the feature. Whereas negative value indicates feature variations on opposite direction. For feature vectorsf_j and $f_k$the covariance σ_{jk} is computed as

$$\sigma_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}(f_j^i - \mu_j)(f_k^i - \mu_k) \qquad (2)$$

The covariance matrix contains the covariance values among the features and has shape $l \times m$.

3. **The Eigen vector and Eigen values:** The Eigen vectors and Eigenvalues and computed for thecovariance matrix to determine the principal components of the dataset.

4. **Computing the principal components:** To calculate Principal component (PC) Eigen vectors and Eigen values are ordered in decreasing order, where the Eigen Vectors with highest Eigen values forms the first PC and is the most significant.

5. **Dimension reduction of the dataset:** In the last step of PCA arrange original data with final PC's representing maximum most dominant information of any dataset.

Independent Component analysis (ICA)

Independent component analysis is a statistic technique used to uncover hidden independent factor in the dataset. ICA is considered as alternative to PCA to divide multivariate signals to sub signals independent to each other and non-Gaussian in nature

Convolutional Neural Network

Convolution neural network (CNN) consist of three layers namely convolution, pooling and fully connected layers as shown in fig.3. The Convolution layer of CNN uses a special method called convolution. Convolution is an operation applied on two function to produce third one as output. With image as input to convolution layer, it generates activation maps summarizing the features present in the input. The convolution operation is performed between the input image and kernel of Size N×N as shown in fig.4.

To improve the efficiency of FER and to reduce the computational burden the size of image must be reduced. The pooling layer contributes towards this for reducing the spatial size of image. The max polling is commonly used layer. The output of this layer is finally passes to fully connected layer. The CNN uses backpropagation algorithm with rectified Linear Unit (ReLU) as activation function.

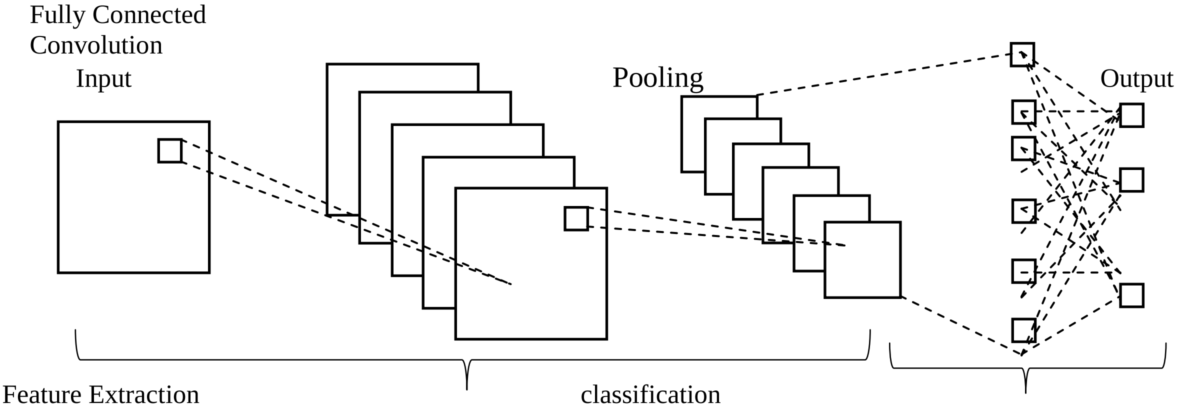

Figure 3: General Structure of Convolution Neural Network

Input Image

0	0	0	0	0	0	0
0	1	0	0	0	1	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	1	0	0	0	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

Feature Detector:

0	0	1
1	0	0
0	1	1

×

Feature Map:

0	1	0	0	0
0	1	1	1	0
1	0	0	2	1
1	4	2	1	0
0	0	1	2	1

Figure 4: Convolution Operation

IV. RESULT AND DISCUSSION

In this work images from dataset FER2013 are given as input to CNN after enhancing images with combination of brightness, contrast, and sharpness by factor of 1.5 and individually followed by extracting features using PCA and ICA as per table 1 and the proposed model of FER shown in fig. 2.

Table 1: Combination of image enhancement and feature reduction for CCN

Combinations of Image enhancement's	Brightness (B)	Contrast (C)	Sharpness (S)	B+C	B+S	C+S	B+C+S
Feature Reduction Techniques	Principal component analysis/Independent component analysis						

The considered CNN model consists of 6 convolutional, 3 max pooling and 2 fully connected layers. Fully connected is further connected to a softmax classifier. All convolutional layers have 3×3 kernels and max pooling layer has 2×2 pool size. To control overfitting dropout layers are also added.Finally, the proposed CNN is classifying seven emotions. The parameters of model areset using the keras library to set hyper-parameters of the model. The CNN model is simulated as per the parameters in table 2.

Table 2: CNN training Parameter Values

Parameter	Value
Learning Rate	0.001
Decay rate (First moment)	0.0001
Decay rate (Second moment)	0.999
Epochs	100
Batch Size	64
Training Dataset	70
Validation Dataset	30
Loss	categorical_crossentropy

The evaluation of CNN model is visualized using confusion matrix. Confusion matrix represent predicted and actual classes in row and column format. The correctly classified classes of positive and negative are denoted as true positive (TP), true negative (TN), and false positive (FP) and false negative (FN). Similarly false positive and negative for misclassified classes. The confusion matrix is also used to generate other evaluation metrics for the evaluation of classification models.

Accuracy

Accuracy is defined as fraction of prediction correctly made by the model and iscalculated as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

The Accuracy of CNN on different feature reduction and combination of image enhancements is summarised in table 2.

Table 3: Accuracy Comparison of CNN

	Model	Feature Reduction	Image Enhancements					
			B	C	S	BC	BS	CS
Accuracy	CNN	PCA	63%	62%	**64%**	60%	63%	63%
		ICA	68%	64%	62%	67%	**69%**	64%

Precision/ Positive Predictive Value (PPV)

Precision can be defined exactness of classifier and is defined as positive patterns detected correctly out of the total predicted as positive. Precision is calculated as:

$$PPV = \frac{TP}{TP+FP} \quad (4)$$

Recall/Sensitivity or the True Positive Rate (TPR)

Completeness of classifier is called recall or sensitivity. It is the ratio of positivesidentified correctly. And is defined as

$$TPR = \frac{TP}{TP+FN} \quad (5)$$

F1 Score/F Score/F Measure

It is defined as harmonic mean between PPV and TPR.

$$F1\ Score = 2 * \frac{PPV * TPR}{PPV + TPR} \quad (6)$$

Table 4:Quantitative analysis of Convolution Neural Network

	CNN PCA								
	B			C			S		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Angry	0.43	0.04	0.08	0.55	0.04	0.08	0.64	0.15	0.25
Disgust	0.67	0.01	0.02	1	0.01	0.01	0	0	0
Fear	0.67	0.4	0.5	0.49	0.49	0.49	0.67	0.3	0.41
Happy	0.73	0.97	0.84	0.8	0.99	0.88	0.78	0.97	0.86
Sad	0.54	0.63	0.58	0.49	0.63	0.55	0.46	0.68	0.55
Surprise	0.76	0.87	0.81	0.74	0.87	0.8	0.75	0.88	0.81
Neutral	0.48	0.69	0.57	0.5	0.55	0.52	0.55	0.69	0.61

Table 5: Quantitative analysis of Convolution Neural Network

	CNN ICA								
	B			C			S		
	Precision	**Recall**	**F1**	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**
Angry	0.53	0.12	0.2	0.41	0.03	0.05	0.57	0.18	0.27
Disgust	0.87	0.12	0.22	1	0.01	0.02	1	0.03	0.05
Fear	0.58	0.59	0.58	0.66	0.45	0.53	0.73	0.33	0.45
Happy	0.84	0.95	0.89	0.8	0.96	0.87	0.74	0.97	0.84
Sad	0.5	0.71	0.59	0.49	0.7	0.58	0.43	0.56	0.48
Surprise	0.82	0.95	0.88	0.82	0.89	0.86	0.76	0.92	0.83
Neutral	0.65	0.64	0.64	0.5	0.69	0.58	0.53	0.64	0.58

The 70% Images of Dataset are used for training and 30% for testing. The Results of CNN classifier is shown in table 3. The obtained results shows that the accuracy of combination sharpness, PCA with sharpness and CNN has highest accuracy that is of 64%. Whereas enhancing brightness and sharpness with ICA for CNN classifiershas accuracy 69%. Although the accuracy of other combinations is also comparable. From table 4 through 7 it isalso revealed that classifier is detecting happy, surprise and disgust with comparable accuracies. Other metricspredicted that emotion which lowest accuracy are angry, disgust and neutral for some combinations. Overall, the results are satisfactory and comparable.

Table 6: Quantitative analysis of Convolution Neural Network

	CNN PCA								
	BC			BS			CS		
	Precision	**Recall**	**F1**	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**
Angry	0.43	0.04	0.07	0.35	0.04	0.08	0.45	0.18	0.25
Disgust	1	0.01	0.01	0	0	0	0	0	0
Fear	0.58	0.22	0.31	0.74	0.41	0.53	0.58	0.34	0.43
Happy	0.81	0.99	0.89	0.8	0.95	0.87	0.83	0.96	0.89
Sad	0.49	0.64	0.55	0.54	0.49	0.51	0.5	0.7	0.58
Surprise	0.71	0.9	0.79	0.83	0.9	0.86	0.67	0.88	0.76
Neutral	0.41	0.65	0.5	0.42	0.82	0.56	0.51	0.58	0.54

Table 7: Quantitative analysis of Convolution Neural Network

	CNN ICA								
	BC			BS			CS		
	Precision	**Recall**	**F1**	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**
Angry	0.58	0.2	0.29	0.5	0.36	0.42	0.41	0.03	0.05
Disgust	1	0.07	0.13	0.93	0.16	0.27	1	0.01	0.02
Fear	0.62	0.58	0.6	0.59	0.51	0.55	0.66	0.45	0.53
Happy	0.76	0.98	0.86	0.76	0.97	0.85	0.8	0.96	0.87
Sad	0.59	0.56	0.57	0.61	0.57	0.59	0.49	0.7	0.58
Surprise	0.85	0.88	0.86	0.79	0.92	0.85	0.82	0.89	0.86
Neutral	0.56	0.71	0.63	0.71	0.7	0.7	0.5	0.69	0.58

V. CONCLUSION

In this work the focus is to design CNN for facial emotion classification considering different combination of image enhancement with PCA and ICA as feature extraction techniques. The performance of classifier is evaluated considering accuracy and other metric derived from confusion matrix. It is revealed from results that CNN has comparable precisions for different emotions and has also comparable accuracy. In future, the accuracy can be enhanced by using combination of other pre-processing techniques andfeature extraction techniques. Classifier will also be tested on different balanced dataset for better precision of FER.

REFERENCES

[1]. Andre Teixeira Lopes, Edilsonde Aguiar, Alberto F. DeSouza, Thiago Oliveira-Santos, “Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order,” Pattern Recognition, vol. 61, pp. 610–628, 2017.

[2]. A. Yadav and D. K. Vishwakarma, "A Weighted Text Representation framework for Sentiment Analysis of Medical Drug Reviews," IEEE Sixth International Conference on Multimedia Big Data (BigMM), pp. 326-332, 2020.

[3]. P. Ekman, W. Friesen, and J. C. Hager, *"Facial Action Coding System. A Human Face,"* 2002.

[4]. Hui Fang, Neil Mac Parthaláin, Andrew J.Aubrey, Gary K.L. Tam, RitaBorgo, Paul L. Rosin, Philip W. Grant, David Marshall, Min Chen, *"Facial expression recognition in dynamic sequences: An integrated approach,"* Pattern Recognition, vol. 47, pp. 1271–1281, 2014.

[5]. P. Ekman, "*Facial Expression and Emotion*," Am. Psychologist,vol. 48, pp. 384-392, 1993.

[6]. BouchraAbboud, Franck Davoine, Mo Dang, *"Facial expression recognition and synthesis based on an appearance model,"* Signal Processing: Image Communication, vol. 19, pp. 723–740, 2004.

[7]. UrošMlakar, IztokFister, Janez Brest, BožidaPotocnik, "Multi-Objective Differential Evolution for feature selection in Facial Expression Recognition systems," Expert Systems with Applications, vol. 89, pp. 129-137, 2017.

[8]. Edwin Alberto Silva Cruz, Claudio Rosito Jung, Carlos Humberto Esparza Franco, Facial Expression Recognition using temporal POEM features, Pattern Recognition Letters (2017), doi: 10.1016/j.patrec.2017.08.008.

[9]. S. H. Wang, P. Phillips, Z. C. Dong, and Y. D. Zhang, "Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm," Neurocomputing, vol. 272, pp. 668–676, 2018, doi: 10.1016/j.neucom.2017.08.015.

[10]. M. Li, H. Xu, X. Huang, Z. Song, X. Liu, and X. Li, "Facial Expression Recognition with Identity and Emotion Joint Learning," IEEE Trans. Affect. Comput., vol. 12, no. 2, pp. 544–550, 2021, doi: 10.1109/TAFFC.2018.2880201.

[11]. R. A. Khan, A. Meyer, H. Konik, and S. Bouakaz, "Framework for reliable, real-time facial expression recognition for low resolution images," Pattern Recognit. Lett., vol. 34, no. 10, pp. 1159–1168, 2013, doi: 10.1016/j.patrec.2013.03.022.

[12]. A. Joseph and P. Geetha, "Facial emotion detection using modified eyemap–mouthmap algorithm on an enhanced image and classification with tensorflow," Vis. Comput., vol. 36, no. 3, pp. 529–539, 2020, doi: 10.1007/s00371-019-01628-3.

[13]. Asit Barman, Paramartha Dutta, "Facial Expression Recognition using Distance and Shape Signature features," Pattern Recognition Letters, 2017, doi: 10.1016/j.patrec.2017.06.018.

[14]. M. Li, H. Xu, X. Huang, Z. Song, X. Liu and X. Li, "Facial Expression Recognition with Identity and Emotion Joint Learning,"IEEE Transactions on Affective Computing, doi: 10.1109/TAFFC.2018. 2880201.

[15]. Lyons MJ, Budynek J, Akamatsu S., "Automatic classification of single facial images". IEEE Transactions on Pattern Analysis and Machine Intelligence 1999; 21:1357-62.

[16]. Bartlett MS, Movellan JR, Sejnowski TJ., "Face recognition by independent component analysis. Neural Networks," IEEE Transaction on 2002; 13:1450-64.

[17]. Belhumeur PN, Hespanha JP, Kriegman D. "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. Pattern Analysis and Machine Intelligence," IEEE Transactions on 1997; 19:711-20.

Sushil kumar. "Comparative Analysis of Feature extraction and image enhancement combinations for facial expression recognition." *The International Journal of Engineering and Science (IJES)*, 11(12), (2022): pp. 25-31.