SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V10I7P116

Volume 10 Issue 7, 181-187, July 2023

Original Article

Facial Expression Recognition for Low Resolution Images using Local and Global Features with SVM Classifier

Shubhangi Patil-Kashid[1], Y. M. Patil[2], Vijaya R. Pawar[3] , A. S. Kashid[4]

[1]*Department of Electronics and Telecommunication Engineering, N. K. Orchid College of Engineering and Technology, Solapur, Maharashtra, India.*
[2]*Department of Electronics and Telecommunication Engineering, KIT College of Engineering, Kolhapur, Maharashtra, India.*
[3]*Department of Electronics and Telecommunication Engineering, Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra, India.*
[4]*Department of Mechanical Engineering, N. K. Orchid College of Engineering and Technology, Solapur, Maharashtra, India.*

[1]*Corresponding Author : kashidshubhangi@gmail.com*

Received: 21 April 2023 Revised: 22 June 2023 Accepted: 07 July 2023 Published: 31 July 2023

Abstract - *Human Machine interaction similar to human-human interaction is possible if machines can recognize expressions on human faces during communication. Facial expression recognition is easy for humans but a difficult task for machines. Facial expression recognition (FER) by machines is expected in the next generation of computers. Image acquisition, preprocessing, feature extraction and classification are the steps in FER. The accuracy of recognition depends on all these steps. Happy, sad, angry, fearful, surprised, disgusted and neutral are the seven expressions considered for classification. Local Binary Pattern (LBP), Histogram of Orientations (HOG), Shift Invariant Feature Transform (SIFT) and Optical Flow (OPF) are the four different feature extraction methods that are used. This paper compares the different feature extraction methods, and multiclass SVM is used as a classifier. The experiment is performed on three different datasets. The novelty of the work is the low-resolution dataset and combination of local and global features. Local feature extractors LBP and HOG are more efficient than global feature extractors SIFT and OPF. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are used to reduce the dimensionality and improve the speed of execution. The confusion matrix is used to compute the accuracy of recognition, specificity and sensitivity. Recognition accuracy is suitable for images captured in controlled laboratory scenarios, but still, more work is required for wild images on both feature extraction and classifier.*

Keywords - *Facial expression recognition, Feature extraction, HOG, LBP, SIFT.*

1. Introduction

Human beings communicate with machines using input and output devices such as keyboards, mouse, monitor, printers etc. Human communication includes spoken words, facial expressions, gestures, body movements, etc. Communication has two types verbal and nonverbal communication. In human communication, 55 % occurs via body language, 7 % by spoken words and 38 % by voice tone [1]. Verbal communication includes spoken words, sounds and text messages. Nonverbal communication takes place via facial expressions, gestures and body movements[2]. Humans can easily understand facial expressions from their childhood. This task is difficult in computer vision as every person's face size, shape, and colour differs. It depends on gender, also. The expressions have a wide range. Distance between face and camera, resolution, intensity of expressions, and illumination are also not constant. Occlusion, head pose variation, and similarity in expressions are obstacles in Facial Expression Recognition Systems (FERS)[3, 4].

The applications of FERS are in searching for lie detection, online gaming, student engagement detection in online schools, and mental diagnosis. Much work is carried out in FERS, but still, it is attracting researchers as real-time FERS has very low accuracy. Man-machine interaction can become possible if machines recognize facial expressions.

This paper compares the performance of feature extraction methods Local Binary Pattern (LBP), Histogram of Gradient (HOG), Shift Invariant Feature Transform (SIFT), Optical Flow (OPF) and their combinations for face expression recognition using SVM classifier. Broadly there are 30 expressions, and with slight varying these 30 expressions, humans can produce nearly 50000 facial expressions [1]. Ekman defines six facial expressions in [5] as happy, sad, disgusted, angry, surprised and fearful. The non-expressional face is considered neutral. The intensity of expressions is not always the same. It varies from micro expressions to extensive expressions.

Expressions appear on the face due to changes in the muscle beneath the skin. These changes are captured using different features and classified using a multiclass support vector machine (SVM) classifier. The four basic steps in FERS are image acquisition, preprocessing, feature extraction and classification.

Every step plays a vital role in improving the recognition accuracy of FERS. The research paper is divided into five sections. The survey of different feature extraction and classification methods is elaborated in this paper's second section, i.e., the literature survey. The complete block diagram of FERS is explained in the third section. Results obtained by applying different feature extraction methods are discussed in the fourth section. The conclusion of the research carried out is given in the fifth section.

2. Literature Survey

Different researchers apply different algorithms for preprocessing, feature extraction and classification. This paper is related to the performance comparison of feature extraction methods. So literature survey of 6 recent papers with different feature extraction methods is elaborated here. Zahra Abbasvandi Et. al [6] extracted features using local binary pattern (LBP) and classified them using SVM with radial basis function. 97% recognition accuracy is achieved with ten-fold cross-validation on the CK+ dataset.

Michel Revina Et. al [7] computed features using Local Directional Number (LDN) and Dominant Gradient Local Ternary Pattern (DGLTP) and classified using SVM. He experimented on JAFEE and C.K. datasets and obtained 88 % recognition accuracy. Ying He [8] Et. al; proposed a new feature High-Order Singular Value Decomposition (HOSVD), combined geometric and texture features and performed human-independent FERS. A. A. Pise et al. discussed methods of FERS using different architectures of neural network and concluded that deep learning neural network gives the most accurate recognition results [9, 10].

Safa Rajaa et al. proposed FERS using HOG descriptors and SVM classifier and validated results using True Success Rate. He concluded that FERS using HOG and SVM is more accurate than the existing state-of-the-art methods [11]. Tingxuan Zhang worked on four expressions, positive, negative, focused, and surprise, required for online classroom teaching. He used a deep-learning neural network for data augmentation and preprocessing. To increase performance speed and reduce parameters, he applied Global Average Pooling and depth-wise separable convolution [12]. To enhance the expression recognition accuracy of FER I.

Michael Revina proposed Whale- Grasshopper Optimization algorithm-based Multi-Support Vector Neural Network. Scale-Invariant Feature Transform (SIFT) and the Scatter Local Directional Pattern (SLDP) were used to extract the features from face images and were classified using the proposed algorithms. The proposed SLDP descriptor is modelled using the LDP, the scattering transform gives robust features, and the proposed classifier increases recognition accuracy [13]. I. Michael Revina et al. combined two phases of Weber Local Descriptor (WLD) based on a similar group and classified using Firefly excited Radial Basis Function Neural Network (F-RBFNN) in which hidden neurons are developed [14].

Tanoy Debnath et al. proposed a new model of FERS in which features are extracted by Local Binary Pattern region based Oriented FAST and rotated BRIEF (ORB) and Convolutional Neural network (CNN) were used to develop the classification model. The proposed four-layer convolution neural network model contains two fully connected layers [15]. Ben Niu et al. extracted rotated BRIEF and LBP features from the expressive part of the face to avoid the concentration of the features and increase computational speed. The features are invariant to scale and rotation. SVM is used for classification [16, 17].

Sunil MP and Hariprasad . S.A. proposed a modified deep learning method by concatenating Xception and ResNet50 architectures to take advantage of both methods. ResNet50 can capture more profound, complex patterns, and Xception efficiently extracts the features [18]. Junjie Wu et al. proposed micro-expression recognition using an optical flow filtering mechanism [19].

S. L. Happy et al. proposed Histogram equalization to reduce the illumination variations by enhancing the contrast of the face image [20]. Dahmane M. proposed a face alignment technique using Scale Invariant Feature Transform, which calculates the reference image first. Then, all images in the database are aligned through related reference images [21]. The review shows that most JAFEE dataset is used for most FERS. As JAFEE contains posed images captured in controlled scenario and training and testing contains images from the same subject, FERS becomes the subject.

Also, the resolution of images is high.Practically it is impossible to get training and testing images from the same subject, and resolution varies if the camera changes. The accuracy of FERS for JAFEE and CK+ datasets is excellent, but more work is still required on preprocessing, feature extraction and classifier for authentic world images.

3. Block Diagram of FERS and Methodology

Figure 1 shows a block diagram of FERS. The experiment is performed using three datasets.

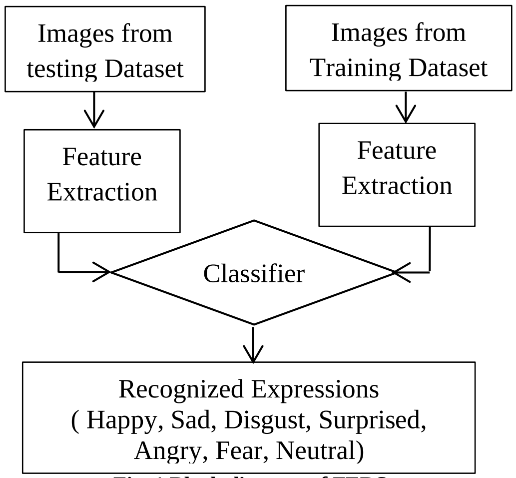


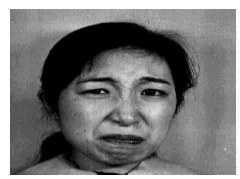
Fig. 1 Block diagram of FERS

Japanese Female Facial Expression (JAFFE) dataset [22] contains 213 posed images from 10 female subjects captured in a controlled scenario at the University of Chicago, Japan [23]. The dataset has preprocessed grayscale images with the resolution 256 X 256 stored in TIFF format. There are two to four images of each person showing all seven expressions. Researchers mostly use this dataset for FERS.

CK+ contains 593 images captured from 123 subjects aged 18 to 30 years. These images are stored with 8-bit precision and 640X 490 pixels in grayscale. A dataset of 700 images named DB2 is created by selecting available face images from the internet.

The subjects include females, males and children from the age group 4 years to 55 years. The image size is 48 X 48, and the format is JPG. This dataset is wild, as images are not captured in a controlled scenario. They are collected without considering illumination, background, or camera distance. Table 1 shows sample images from the JAFFE dataset.

Table 1. Sample images from the JAFEE dataset

Expression	Happy	Sad	Disgust	Neutral
Image				

The experiment is carried out in Matlab20018, an on-core i5 processor laptop. Table 2 shows the number of images used for training and testing purposes. 70 % of the images are used for training and 30% for testing. As all the images in JAFFE are preprocessed, direct features are computed using different algorithms. Viola Jones algorithm of object detection [24] is used to find the face area from the image. It is a fast and efficient algorithm. A local facial region, such as eye and mouth regions, contains more information for classification, so local descriptors such as LBP and HOG are used for feature extraction. The face region is divided into small spatial regions called cells. Local features LBP and HOG are computed by dividing the face area into 8X8 cell sizes. Global features SIFT and OPF are computed from the entire face area. The accuracy of expression recognition depends on the features computed.

4. Result and Discussion

70 % of the images are used for training and 30% for testing. Table 2 shows the images used for training and testing from all three datasets. The experiment is carried out in Matlab20018a on a core i5 processor laptop. As all the images in JAFFE are preprocessed, direct features are computed using different algorithms. Viola Jones algorithm of object detection [22] is used to find the face area from the image. It is a fast and efficient algorithm. A local facial region, such as eye and mouth regions, contains more information for classification, so local descriptors such as LBP and HOG are used for feature extraction. The accuracy of expression recognition depends on the features computed. The face region is divided into small spatial regions called cells. Local features LBP and HOG are computed by dividing face area into 8X8, 16X16, 32X32, and 64X64 cell sizes.

To calculate LBP 3X3 matrix is used. The matrix's centre value is a threshold and is compared with other neighbour values to find the binary code of each pixel. If the pixel value is less than the threshold, it is replaced by 0; if it is greater than the threshold, it is replaced by 1. The histogram bin is constructed to know the frequency of each binary pattern.

The 3X3 matrix uses 8 pixels, so the histogram bin has 28 values. The feature vector of uniform LBP obtained by dividing the face region into 8X8 cell size is reduced to length 59. HOG is also a local feature extraction method in which the magnitude at each cell is computed using

Pythagoras Theorem, and the orientation is calculated using a tangential angle. The histogram of gradient and orientation computed on each cell are combined to form HOG representation.

HOG features are invariant to illumination and shadowing.The contrast is normalized using the energy of the spatial block region, referred to as normalized HOG features. If the change in neighbouring pixel magnitude is more, it indicates the change in intensity value.

Table 3 shows the recognition accuracy obtained and execution time taken using LBP and HOG features and SVM classifier. HOG features give the highest recognition accuracy and lowest execution time. Also, the highest recognition accuracy is obtained with JAFEE dataset as it contains posed images in a controlled scenario. 16X16 cell size gives more accuracy than any other cell size. For the DB2 dataset, recognition accuracy is less and execution time is more as compared to JAFEE and CK+ datasets.

It shows that only HOG or LBP features are insufficient for low-resolution, non-posed, subject-independent FERS. Global features SIFT and OPF are computed from the entire face area. SIFT is a technique to detect stable salient points in the image. As the name indicates, SIFT features are invariant to rotation and scale. So SIFT features do not change with illumination, scale and object position. Combining HOG features with Scale Invariant Feature Transform (SIFT) gives matching scale-invariant vital points.

Table 2. Number of images used for training and testing purposes from different datasets

Dataset	**JAFEE**			**CK+**			**DB2 (IMAGES FROM NET)**		
Expression	Test	Train	Test	Test	Test	Total	Test	Train	Total
Angry	10	21	31	26	59	85	30	70	100
Disgusted	10	21	31	26	59	85	30	70	100
Fear	10	21	31	26	59	85	30	70	100
Happy	9	21	30	25	59	84	30	70	100
Neutral	9	21	30	25	59	84	30	70	100
Sad	9	21	30	26	59	85	30	70	100
Surprised	9	21	30	26	59	85	30	70	100
Total Images	66	147	213	180	411	593	30	70	100

Table 3. Recognition accuracy obtained and execution time taken by FERS for local features

Features	**Database**	**Cell size**	**Accuracy (%)**	**Execution time (sec)**	**Sensitivity**	**Specificity**
HOG	JAFEE	8	95.52	5.61	0.96	0.99
HOG	JAFEE	16	97.18	4.97	0.96	0.99
HOG	JAFEE	32	95.52	4.78	0.96	0.99
HOG	JAFEE	64	98.29	5.87	0.96	0.99
HOG	CK+	8	95.15	5.57	0.96	0.99
HOG	CK+	16	96.32	5.31	0.96	0.99
HOG	CK+	32	95.52	4.97	0.96	0.99
HOG	CK+	64	94.82	5.06	0.96	0.99
HOG	DB2	8	50.95	8.15	0.51	0.92
HOG	DB2	16	54.29	7.71	0.49	0.91
HOG	DB2	32	51.16	7.38	0.44	0.91
HOG	DB2	64	48.17	7.53	0.44	0.91
LBP	JAFEE	8	95.52	6.29	0.96	0.99
LBP	JAFEE	16	97.01	5.47	0.96	0.99
LBP	JAFEE	32	94.82	4.93	0.96	0.99
LBP	JAFEE	64	94.16	5.31	0.96	0.99
LBP	CK+	8	93.28	5.93	0.95	0.98
LBP	CK+	16	94.61	5.38	0.95	0.98
LBP	CK+	32	93.53	5.84	0.95	0.98
LBP	CK+	64	92.87	6.21	0.95	0.98
LBP	DB2	8	41.90	7.86	0.43	0.90
LBP	DB2	16	48.57	7.33	0.43	0.90
LBP	DB2	32	43.33	8.44	0.43	0.90
LBP	DB2	64	41.29	8.73	0.43	0.90

Which helps capture edge and gradient change due to expression. OPF features are essential at edges and corners to track the motion of points from frame to frame. SIFT and OPF are the global features computed from the entire face area. The features thus computed are classified using seven class SVM with radial basis function. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are used to reduce the dimensionality and improve the speed of execution without affecting recognition accuracy. Table 4 shows the recognition accuracy, sensitivity, and specificity of FERS in which global and a combination of local and global features are used. The cell size selected is 16X16, and the dataset used is JAFEE. The recognition accuracy is computed using the formula given in Equation 1.

$$Accuracy = \frac{\sum number\ of\ correctly\ classified\ expressions}{Total\ number\ of\ expressions\ of\ that\ class} \quad (1)$$

We get the number of correctly classified expressions from the confusion matrix's upper left to lower correct diagonal. The classifier's performance is better tested by sensitivity and specificity than only recognition accuracy. To find the misleading results confusion matrix is also computed for every combination. 7 X 7 confusion matrix is used to evaluate the performance of seven expressions. The confusion matrix gives a complete analysis by comparing classified and labelled target values. Predicted values are plotted on the x-axis, and true values are plotted on the y-axis. It helps to understand errors in classification Sensitivity indicates a true positive rate.

Sensitivity is the ratio of correctly classified optimistic targets to all positive targets in the dataset. Specificity indicates a true negative rate. Specificity is the ratio of negatively identified targets to all negative targets in the dataset. Sensitivity and specificity are computed from each confusion matrix by using the formula given in Equation 2 and Equation 3.

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (2)$$

$$Specificity = \frac{True\ Negative}{True\ Negative + False\ Positive} \quad (3)$$

An actual positive result means the result is accurate and correctly identified as valid.

True negative indicates that the result is negative and correctly identified as negative. False negative class indicates the number of actual results but is mistakenly considered wrong.

Figure 2 shows the sample confusion matrix used to calculate accuracy, sensitivity and specificity. From Figure 3, one image from surprise is wrongly identified as fear and sad is identified as disgust. In both surprise and fear, eye size increases, sometimes leading to wrong identification. Accuracy, Sensitivity and Specificity are calculated for each combination of feature extraction methods, and the result is shown in Tables 3 and 4.

Table 4. The recognition accuracy of proposed FRES with SVM classifier

Features Used	Accuracy (%)	Execution Time (sec)	Sensitivity	Specificity
SIFT	13.43%	9.35	0.19	0.58
OPF	25.37%	10.29	0.27	0.64
LBP + SIFT	98.06%	5.67	0.96	0.99
LBP + OPF	96.57%	5.83	0.96	0.99
HOG + SIFT	98.40%	5.29	0.96	0.99
HOG + OPF	97.07%	4.81	0.96	0.99
LBP + SIFT + PCA	98.06%	5.43	0.96	0.99
LBP + SIFT + LDA	97.04%	5.72	0.96	0.99
LBP + OPF+ PCA	96.57%	5.39	0.96	0.99
LBP + OPF + LDA	89.55%	5.31	0.94	0.97
HOG + SIFT + PCA	97.57	4.91	0.96	0.99
HOG +SIFT +LDA	95.16	5.37	0.96	0.99
HOG + OPF + PCA	96.57	4.21	0.96	0.99
HOG + OPF + LDA	96.84	4.36	0.96	0.99

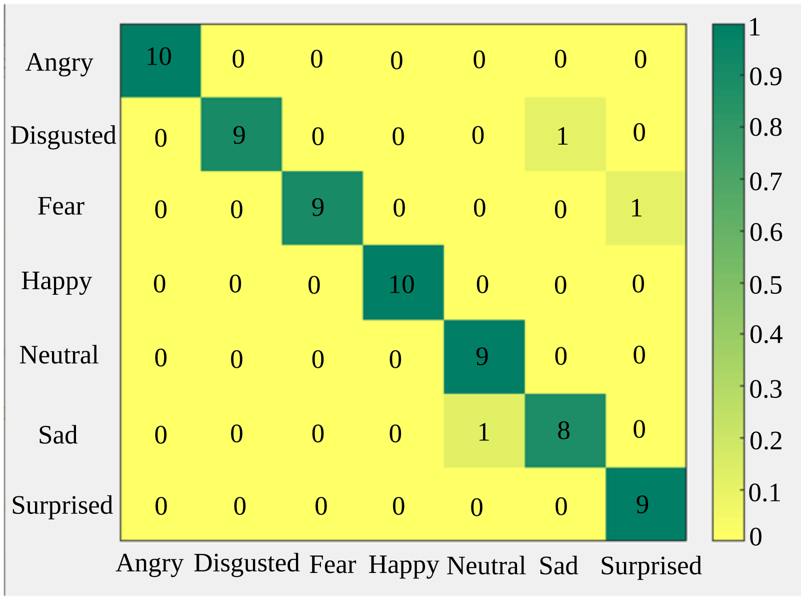


Fig. 2 Sample confusion matrix

5. Conclusion

The accuracy of recognition is computed by applying local and global features separately. Local and global features are also used to check the performance. The highest 98.40 % recognition accuracy is obtained with HOG features combined with SIFT. LBP features also give good recognition accuracy. Compared with LBP and HOG, the performance of HOG is best for recognition accuracy. SIFT and OPF features alone give the slightest expression recognition accuracy. It indicates that local features are more efficient than global features for expression recognition. PCA and LDA increase execution time as only selected features are used for classification without affecting recognition accuracy significantly. More execution time is required for low-resolution non-posed images, and recognition accuracy is less. Still, more work is required to design a robust algorithm for low-resolution wild images to initiate human-machine communication.

References

[1] Mehrabian, *A Communication without Words*, *Psychology Today*, vol. 9, no. 2, pp. 52-55, 1968.

[2] Linkai Li et al., "Integrated Access Control System of Face Recognition and Non-Contact Temperature Measurement Based on Arduino," *International Journal of Computer and Organization Trends*, vol. 12, no. 2, pp. 1-5, 2022. [CrossRef] [Publisher Link]

[3] Shubhangi Patil, and Y. M. Patil, "A Review of Human Facial Expression Recognition Systems Applied for Wild Dataset in Last Decade," *AIP Conference Proceedings*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Pooja G. Nair, and Sneha .R, "A Review: Facial Recognition using Machine Learning," *International Journal of Recent Engineering Science*, vol. 7, no. 3, pp. 85-89, 2020. [Google Scholar] [Publisher Link]

[5] Paul Ekman, and Wallace V. Friesen, *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*, Prentice Hall, Englewood Cliffs, 2003. [Google Scholar] [Publisher Link]

[6] Zahra Abbasvandi et al., "Recognition of Facial Expressions by Extracting Local Binary Pattern Features," *Proceedings of International Conference on New Research Achievements in Electrical and Computer Engineering*, 2016. [Publisher Link]

[7] I. Michael Revina, and W.R. Sam Emmanuel, "Face Expression Recognition using LDN and Dominant Gradient Local Ternary Pattern Descriptors," *Journal of King Saud University – Computer and Information Sciences*, vol. 33, pp. 392-398, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Ying He, and Xiaoju He, "Facial Expression Recognition Based on Multi-Feature Fusion and HOSVD," *Information Technology, Networking, Electronic and Automation Control Conference*, pp. 638-643, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Anil Audumbar Pise et al., "Methods for Facial Expression Recognition with Applications in Challenging Situations," *Hindawi Computational Intelligence and Neuroscience*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Joyassree Sen et al., "Face Recognition using Deep Convolutional Network and One-Shot Learning," *SSRG International Journal of Computer Science and Engineering*, vol. 7, no. 4, pp. 23-29, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Safa Rajaa et al., "Facial Expression Recognition System Based on SVM and HOG Techniques," *International Journal of Image Processing*, vol. 15, no. 2, pp. 14-21, 2021. [Publisher Link]

[12] Tingxuan Zhang, "Face Expression Recognition Based on Deep Learning," *Journal of Physics: Conference Series*, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[13] I. Michael Revina, and W.R. Sam Emmanuel, "Face Expression Recognition with the Optimization based Multi-SVNN Classifier and the Modified LDP Features," *Journal of Visual Communication and Image Representation*, vol. 62, pp. 43–55, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] I.Michael Revina, and W.R.Sam Emmanuel, "Face Expression Recognition using Weber Local Descriptor and F-RBFNN," *Proceedings of the Second International Conference on Intelligent Computing and Control Systems*, *IEEE Xplore Compliant Part Number*, pp. 196-199, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[15] Tanoy Debnath et al., "Four-Layer ConvNet to Facial Emotion Recognition with Minimal Epochs and the Signifcance of Data Diversity," *Scientific Reports*, vol. 12, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Ben Niu, Zhenxing Gao, and Bingbing Guo, "Facial Expression Recognition with LBP and ORB Features," *Hindawi Computational Intelligence and Neuroscience*, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Anju, and Anu Rani, "Face Recognition using PCA Gobor Filter SVM Techniques," *International Journal of Computer & Organization Trends*, vol. 6, no. 3, pp. 20-23, 2016. [Publisher Link]

[18] M. P. Sunil, and S. A. Hariprasad, "Facial Emotion Recognition using a Modified Deep Convolutional Neural Network Based on the Concatenation of XCEPTION and RESNET50 V2," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 6, pp. 94-105, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Junjie Wu et al., "Optical Flow Filtering-Based Micro-Expression Recognition Method," *Electronics*, vol. 9, no. 12, p. 2056, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] S. L. Happy, and S. Routray, "Automatic Facial Expression Recognition using Features of Salient Facial Patches," *IEEE Transactions on Affective Computing*, vol. 6, no. 1, pp. 1–12, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[21] Mohamed Dahmane, and Jean Meunier, "Prototype-Based Modeling for Facial Expression Analysis," *IEEE Transactions on Multimedia*, vol. 16, no. 6, pp. 1574–1584, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[22] Shubhangi Patil, and Y. M. Patil, "Face Expression Recognition using SVM and KNN Classifier with HOG Features," *Smart Innovation, Systems and Technologies, Springer, Singapore*, vol. 303, pp. 416-424, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] The Japanese Female Facial Expression (JAFFE) Database, 2020. [Online]. Availabele: http://www.kasrl. org/jaffe.html

[24] P. Viola, and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai, HI, USA, pp. 511–518, 2001. [CrossRef] [Google Scholar] [Publisher Link]