An Intelligent System Design for Emotion Recognition and Rectification Using Machine Learning

Mr. Umakant Dinkar Butkar, Manisha J Waghmare
Sir Visvesvaraya Institute of Technology, Nashik, India.
Matoshri College of Engineering & Research Centre

Abstract:
Nonverbal cues that are sent through facial expressions are crucial in interpersonal interactions. Automatic facial expression recognition can be a crucial component of typical human-machine interfaces. A facial expression can be viewed from the perspective of automatic recognition as consisting of changes to the facial components' spatial relationships, changes in the pigmentation of the face, or both. The CVIP tools help us achieve our goals. We used a train data set of six facial expressions from three different people. For the train data set, we used a total of 90 border mask samples, and for the test data set, we used 30 border mask samples.
Keywords: CVIP tools, RST- Invariant, KNN, Human-Machine Interfaces.
DOI: 10.24297/j.cims.2023.2.3

1. Introduction

A facial expression can be viewed from the standpoint of automated recognition as being made up of changes in the pigmentation of the face or deformations of facial components and their spatial relationships[1].Face expressions are the variations in facial appearance that occur in response to an individual's internal emotional states, social cues, or intentions[2][3]. The most effective, immediate, natural, nonverbal method for humans to convey emotions and intentions is through facial expression[4][5]. Instead of using words, facial expressions are a faster way to convey feelings. As robots and people begin to share more and more jobs, the need for effective lines of communication between them grows more and more important. Human machine interaction (HMI) systems are systems that create these communication channels.Technology advancements enable the creation of more beneficial HMI systems that don't rely on common objects like keyboards, mice, and displays but instead receive commands directly from users' voices and mimics. Such systems don't need any artificial equipment and merely use human-to-human communication channels to imitate human-human interaction[6]. Before machines take over more aspects of our lives, human-machine contact needs to be improved to more closely resemble human-to-human interaction. Facial expression recognition (FER) will be one of the finest approaches for enhancing HMI systems because changing facial expressions are a powerful way to convey emotions[7][8].Face detection, facial feature point extraction, and facial expression categorization make up the three primary components of an autonomous facial expression identification system[9][10]. The system first acquires the input image and applies various image processing methods to it in order to identify the facial region. Face tracking in movies is referred to as face localization in static images. Once the face has been identified in the picture or video frame, it can be assessed to see if any facial activity is occurring. A feature is a subject of curiosity or knowledge[11]. Geometric characteristics and appearance features are the two categories of features that are typically employed to represent facial emotion.While appearance features depict the change in texture of the face when a particular motion is done, geometric features assess the displacements of specific portions of the face, such as the corners of the mouth or the brows[12]. Geometric feature measurement is frequently related to face area analysis, namely the identification and tracking of critical spots in the face region. Occlusions, the presence of facial hair or glasses, and obstructions could occur during the task of decomposing the face. Furthermore, it can be challenging to define the feature set because characteristics should be descriptive and sometimes uncorrelated.The classification job is the final component of the Facial Expressions Recognition system that is based on machine learning theory. The classifier takes as input a set of features that were extracted from the face region in the previous step. In order to explain the face expression, the set of features is made. Training for classification requires supervision, hence labeled data should be included in the training set[13]. Once trained, the classifier can identify incoming photos by labeling them with a certain class. The most common facial expressions are categorized using both the Facial Action Coding System's Action Units and the six universal emotions of happiness, sorrow, anger, surprise, disgust, and fear.K-Nearest Neighbors, artificial neural networks, support vector machines, hidden markov models, expert systems with rule-based classifiers, bayesian networks, and boosting approaches are only a few examples of the numerous machine learning techniques available for classification tasks. Selecting a solid feature set, using an effective machine learning technique, and using a distinct database for training are the three key challenges in classification tasks. Features that are distinctive and particular to a certain expression should make up the feature set[14]. The kind of feature set that is used to determine the machine learning technique is common. The training set database should be sufficiently large and contain a range of data. The pair of eyes, the nostrils, and the mouth region are represented as region of interests in facial expression identification.A region of interest is a broad area that includes the point that we want to detect. Only some facial regions are employed for discrimination in facial expression recognition systems[15]. The primary features of any facial expression recognition system are the regions of the eyes, brows, mouth, and nose. By removing landmarks or other components from a photograph of the subject's face, some facial recognition algorithms are able to recognize different parts of the face[16].These characteristics are then used to look for photographs with the same characteristics. The bulk of facial expression recognition techniques that have been described so far concentrate on identifying six main kinds of emotion, including happiness, sadness, fear, anger, contempt, and grief[17].

2. Related work

A fully programmed FERS was shown by Giorgana and Ploeger (2012) to be able to recognize the emotions of grief, joy, and surprise.They employed Gabor filters and Principal Component Analysis to extract characteristics from facial photographs.To aggregate the output of neural network classifiers, utilize MD-Adaboost.The face expression image sequences are used to extract the geometric features.The Elastic Bunch Graph Multi-resolution approach serves as the foundation for the landmark initialization and tracking.For the purpose of locating and separating the face portion from the images, they used the integral projection approach.The fisherface approach was used to extract the features. The neural network was employed as a classifier to distinguish between the facial expressions.The proposed system had a recognition rate of 86.85%.The most effective SVM classifier combination for facial expression recognition is found using the Multi Objective Genetic Algorithm. A capable method for recognizing face expression in image sequences was developed by Sarawagi and Arya (2013). They placed a strong emphasis on face feature extraction and color normalization. They utilized a local binary pattern to find features. A new automatic landmark identification method that makes advantage of local binary patterns is presented to find facial landmarks. The accuracy for the Indian database is 94.7%. The characteristics were retrieved by Hablani et al. (2013) using the Local Binary Pattern. They have chosen several significant face characteristics, such as the smaller portions of the eyes, nose, and lips.To categorize the expression, they employed template matching to compare the retrieved facial feature templates. Compared to methods that make use of the entire face image, their suggested strategy is superior. To outperform more traditional expressive systems, the suggested solution includes human identification. The results demonstrate that self-training with unlabelled data can greatly enhance the performances attained with a limited number of labeled training instances. In order to maintain the system current and create a stronger facial recognition system, self-trained individuals are used. An technique for face expression was developed by Zhang et al. (2013) by identifying connections between visual feature and Local Binary Pattern (LBP).Therefore, face expression recognition tends to be more understandable without reducing recognition outcomes. Das (2014) stated that it is difficult to discern facial expressions in real time. They put out a technique for real-time, automatic facial expression recognition that was based on geometrical features. The suggested solution performs better for the JAFFE dataset and operates in real-time on video data. Using ANN and KNN, Hai et al. (2015) suggested a model for categorizing facial expressions. To extract facial traits, they employed Independent Component Analysis. For the JAFEE database's classification of seven fundamental face expressions, they used the ANN KNN model. 92.38% of classes were correctly classified. De et al. (2015) created a model of the eigenface technique to recognize facial expressions in people.Eigen space, which described the overall ariations among face images, is the ideal feature space to represent variation in the eigenfaces. They suggested using the Hue Saturation Value (HSV) color model to identify faces in images.This greater local knowledge of faces on significant fiducial sites is captured. The proposed technique has a 96.19% recognition rate. For the purpose of recognizing facial expressions, Yu and Liu (2015) incorporated geometric features and appearance descriptors from the image. Covariance descriptors with various textural properties are generated to reflect facial appearance. A face emotion identification system using attributes extrapolated from facial motions was proposed by Zhang and Tjondronegoro (2011).

3. Methodology

Support Vector Machine, Artificial Neural Network, and K Nearest Neighbors are the methods we employ (KNN). In pre-processing, we start with the image and apply these methods before extracting the features. Some preset positions are used as facial features in the features extraction approach. Next, they were categorised based on their feature choices. Figure 1.shown that facial expression recognition system overview below:

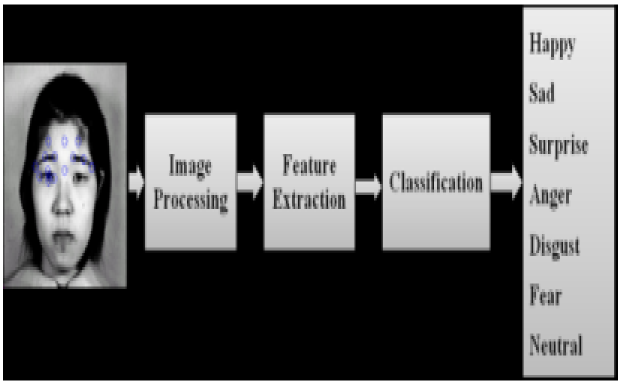


Figure 1: Facial Expression Recognition System Overview

3.1. Image Acquisition

Static images are those that are utilized to identify face expressions. We utilize a Panasonic camera (Model DMC-LS5) with a 5mm lens length to capture pictures of people's facial expressions. The photographs are 24 bit color JPEG files with a 4320 by 3240 pixel resolution. Images of each subject's six most common expressions were taken at a four-foot distance from the camera.

3.2. Image Pre-processing

A crucial stage in the face expression identification challenge is the image preparation step. Pre-processing aims to capture photos with a normalized intensity, uniform size and shape, and representation of only a single face conveying a specific emotion. Additionally, the pre-processing step should lessen the impacts of lighting and illumination. The translation, scale, and rotation of the head in a photograph can have an impact on how an expression is represented.

3.3. Feature Extraction

The most crucial stage in creating a reliable facial expression detection system is feature extraction. For further expression categorization, the split representation is used. Typically, feature extraction reduces the dimensionality of the information space. Important data should be maintained during the reduction process with strong segregation and security.

3.4. Feature Selection

The goal of feature selection is to pick from a wider pool of candidate features a subset of characteristics that are absolutely necessary to carry out the classification operation. The practice of choosing a smaller collection of features increases the classifier's effectiveness and speeds up execution.

3.5. Classification

Finalizing Facial Expressions Based on the collected features, recognition systems are designed to identify facial expression. A classification method uses an algorithm to identify a given expression as one of a set of possible expressions. For classification, we employ the K-Nearest Neighbor classifier.

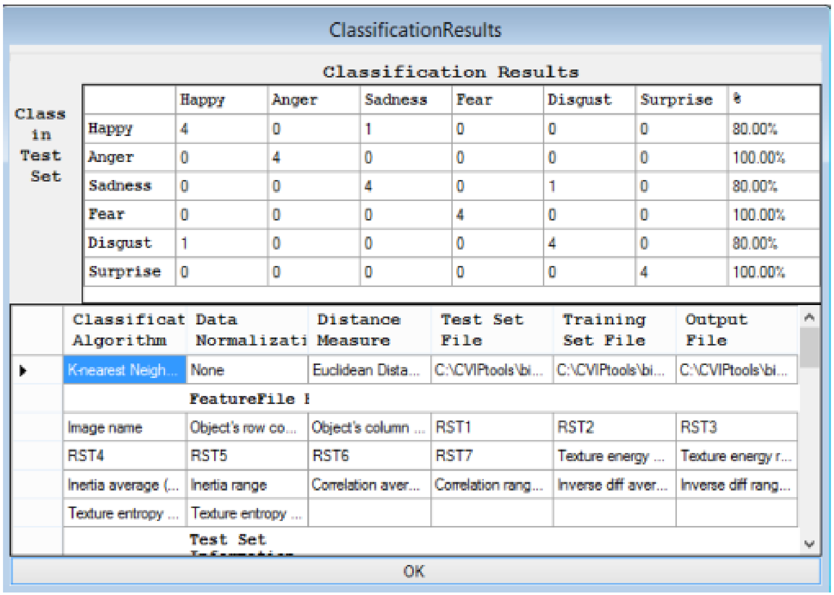
3.6. Physically Data Set:

Here we have the different data sets. Figure 2 shown that actual data set.

Figure 2: Actual data set

4. Results And Discussion

The CVIP tools help us achieve our goals. Figure 3 shown that Confusion matrix for Individual.



	Happy	Anger	Sadness	Fear	Disgust	Surprise	%
Happy	4	0	1	0	0	0	80.00%
Anger	0	4	0	0	0	0	100.00%
Sadness	0	0	4	0	1	0	80.00%
Fear	0	0	0	4	0	0	100.00%
Disgust	1	0	0	0	4	0	80.00%
Surprise	0	0	0	0	0	4	100.00%

Classificat Algorithm	Data Normalizati	Distance Measure	Test Set File	Training Set File	Output File
K-nearest Neigh...	None	Euclidean Dista...	C:\CVIPtools\bi...	C:\CVIPtools\bi...	C:\CVIPtools\bi...
	FeatureFile I				
Image name	Object's row co...	Object's column ...	RST1	RST2	RST3
RST4	RST5	RST6	RST7	Texture energy ...	Texture energy r...
Inertia average (...	Inertia range	Correlation aver...	Correlation rang...	Inverse diff aver...	Inverse diff rang...
Texture entropy ...	Texture entropy ...				
	Test Set				

Figure 3: Confusion matrix for Individual

A confusion table comprising six facial expressions of person A is shown in the table above. The train data set is used. For categorization, we have a test data set with a ratio of 30, 70. There are two different textures. Following that, pattern classification is carried out. There is no standardization of the data. Euclidean distance is the unit of measurement. Accuracy for happiness is 100%, for anger 80%, for sadness 80%, for fear 100%, for disgust 80%, and for surprise 100%. 90% of the time is accurate overall. Figure 4 shows that classification of person a's facial expressions below.

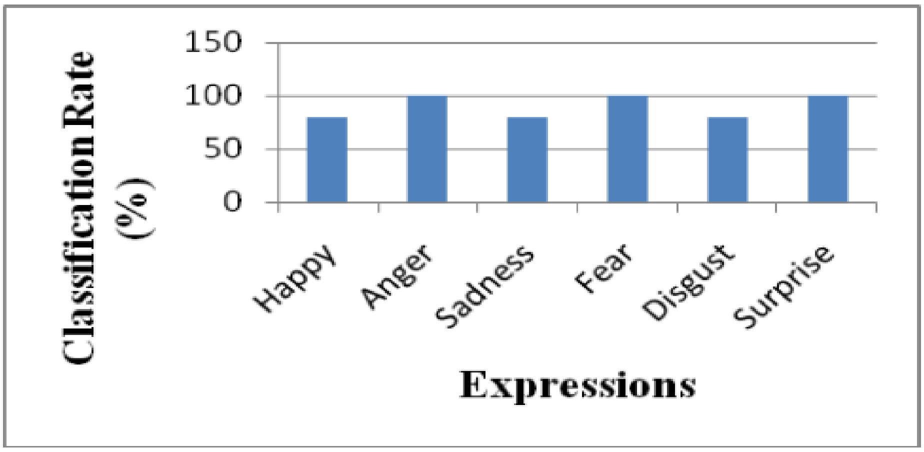
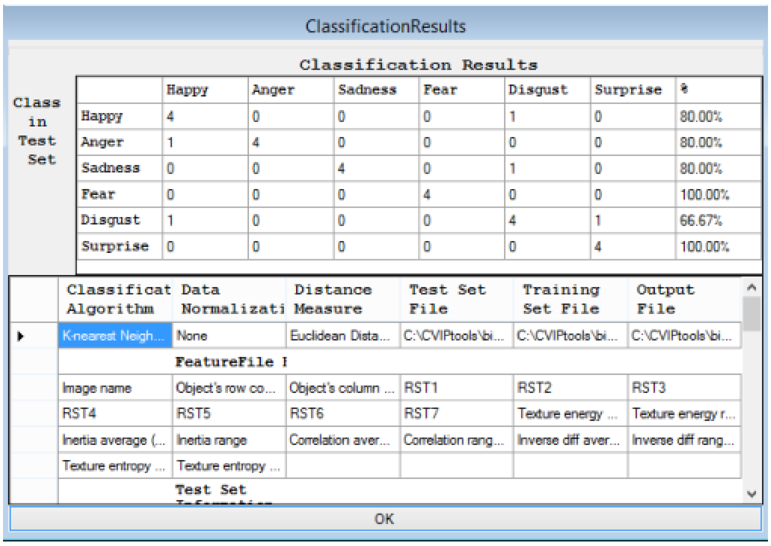


Figure 4: Classification of Person A's facial expressions

Six facial expressions of participant A were classified in the graph up top. Accuracy for happiness is 100%, for anger 80%, for sadness 80%, for fear 100%, for disgust 80%, and for surprise 100%. Figure 5 shown that Confusion for Person B below:



	Happy	Anger	Sadness	Fear	Disgust	Surprise	%
Happy	4	0	0	0	1	0	80.00%
Anger	1	4	0	0	0	0	80.00%
Sadness	0	0	4	0	1	0	80.00%
Fear	0	0	0	4	0	0	100.00%
Disgust	1	0	0	0	4	1	66.67%
Surprise	0	0	0	0	0	4	100.00%

Classificat Algorithm	Data Normalizati	Distance Measure	Test Set File	Training Set File	Output File
K-nearest Neigh...	None	Euclidean Dista...	C:\CVIPtools\bi...	C:\CVIPtools\bi...	C:\CVIPtools\bi...
	FeatureFile I				
Image name	Object's row co...	Object's column ...	RST1	RST2	RST3
RST4	RST5	RST6	RST7	Texture energy ...	Texture energy r...
Inertia average (...	Inertia range	Correlation aver...	Correlation rang...	Inverse diff aver...	Inverse diff rang...
Texture entropy ...	Texture entropy ...				
	Test Set				

Figure 5: Confusion for Person B

A confusion table of person B's six facial expressions is shown in the table above. The train data set is used. For categorization, we have a test data set with a ratio of 30, 70. There are two different textures. Following that, pattern classification is carried out. There is no standardization of the data. Euclidean distance is the unit of measurement. The 80% accuracy rate for happiness, 80% for anger, 80% for sadness, 100% for fear, 66.67% for contempt, and 100% for surprise. 84% of the total is accurate. Figure 6 shown that classification of person b's facial expressions below:

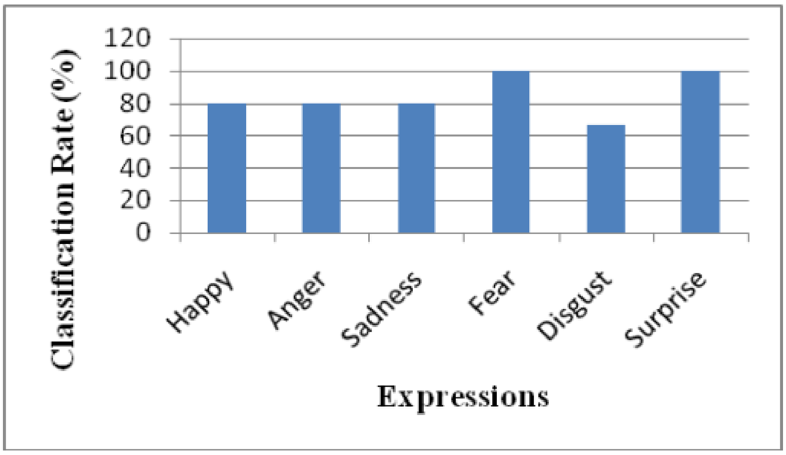
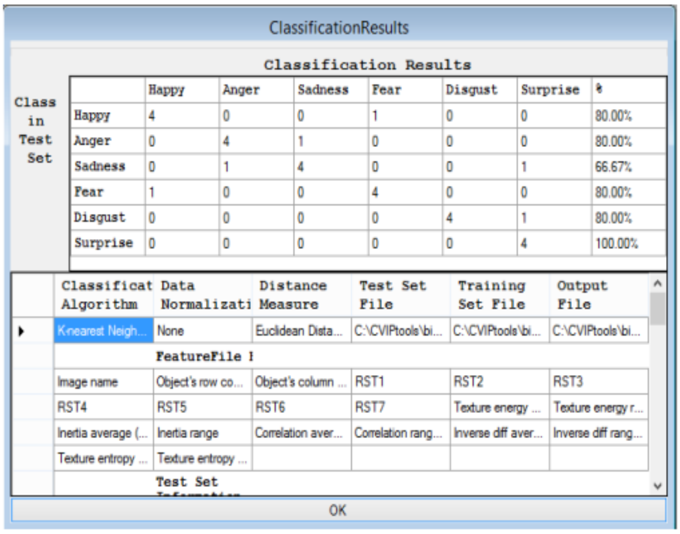


Figure 6: Classification of Person B's Facial Expressions

The 80% accuracy rate for happiness, 80% for anger, 80% for sadness, 100% for fear, 66.67% for contempt, and 100% for surprise. Figure 7 shown that confusion for person c.



	Happy	Anger	Sadness	Fear	Disgust	Surprise	%
Happy	4	0	0	1	0	0	80.00%
Anger	0	4	1	0	0	0	80.00%
Sadness	0	1	4	0	0	1	66.67%
Fear	1	0	0	4	0	0	80.00%
Disgust	0	0	0	0	4	1	80.00%
Surprise	0	0	0	0	0	4	100.00%

Classificat Algorithm	Data Normalizati	Distance Measure	Test Set File	Training Set File	Output File
K-nearest Neigh...	None	Euclidean Dista...	C:\CVIPtools\bi...	C:\CVIPtools\bi...	C:\CVIPtools\bi...
	FeatureFile I				
Image name	Object's row co...	Object's column ...	RST1	RST2	RST3
RST4	RST5	RST6	RST7	Texture energy ...	Texture energy r...
Inertia average (...	Inertia range	Correlation aver...	Correlation rang...	Inverse diff aver...	Inverse diff rang...
Texture entropy ...	Texture entropy ...				
	Test Set				

Figure 7: Confusion matrix for Person C

A list of person C's six possible facial expressions is shown in the table above. The train data set is used. For categorization, we have a test data set with a ratio of 30, 70. There are two different textures. Following that, pattern classification is carried out. There is no standardization of the data. Euclidean distance is the unit of measurement. Happiness has an accuracy of 80%, anger 80%, sadness 66.67%, fear 80%, disgust 80%, and surprise 100%. 81% is the overall accuracy rate. Figure 8 shown that classification of person c's facial expressions.

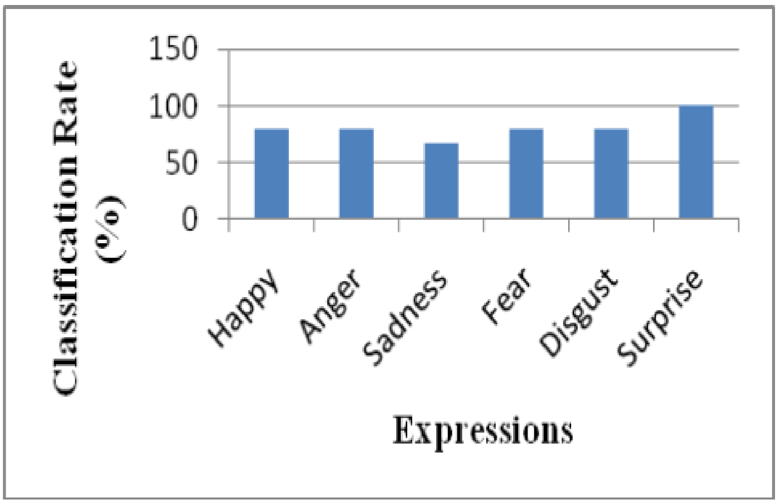


Figure 8: Classification of Person C's Facial Expressions

Six facial expressions of individual C were classified in the graph up top. Happiness has an accuracy of 80%, anger 80%, sadness 66.67%, fear 80%, disgust 80%, and surprise 100%.

5. Conclusion

Six facial expressions were classified differently in each of the Confusion tables. In the confusion table 4.1, where k is the k-Nearest Neighbor method, the maximum accuracy was 90% at k=2. All of the confusion tables demonstrate that the Surprise expression yields superior results, with a 100% accuracy rate, than other expressions. RST-Invariant features and texture features are used for improved results.

Conflict of Interest

There is no conflict of interest.

Data availability statement:

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding Statement:

This study did not receive any funding in any form.

References

1. Giorgana, G., and Ploeger, P. G., 2012, Facial expression recognition for domestic service robots, In Robo Cup 2011: Robot Soccer World Cup XV, pp. 353-364.
2. Wang, Z. and Xiao, N., 2013, Using MD-Adaboost to Enhance Classifier of Facial Expression Recognition, Journal of Computational Information Systems, 9(3), 923- 932.
3. Ghimire, D., and Lee, J., 2013, Geometric feature-based facial expression recognition in image sequences using multi-class Adaboost and support vector machines, Sensors, 13(6), 7714-7734.
4. Abidin, Z., and Harjoko, A., 2012, A neural network based facial expression recognition using fisherface, International Journal of Computer Applications, 59(3), 30-34.
5. Zavaschi, T. H., Britto, A. S., Oliveira, L. E., and Koerich, A. L., 2013, Fusion of feature sets and classifiers for facial expression recognition, Expert Systems with Applications, 40(2), 646-655.

6. Sarawagi, V., and Arya, K. V., 2013, Automatic facial expression recognition for image sequences, In Contemporary Computing (IC3), Sixth International Conference on, 278-282.
7. Hablani, R., Chaudhari, N., and Tanwani, S., 2013, Recognition of Facial Expressions using Local Binary Patterns of Important Facial Parts, International Journal of Image Processing (IJIP), 7(2), 163- 170.
8. Lingareddy, M. and D, Haritha., 2013, An Optimal Face Recognition Tool, International Journal of Research in Engineering and Technology, 2(3), 351-357.
9. Zhang, H., Luo, S., and Yoshie, O., 2013, Facial expression recognition by analyzing features of conceptual regions, In Computer and Information Science, 529-534.
10. Das, D., 2014, Human' s Facial Parts Extraction to Recognize Facial Expression, International journal on information theory, 3(3), 65-72.
11. Hai, T. S., Thai, L. H. and Thuy, N. T., 2015, Facial Expression Classification Using Artificial Neural Network and K-Nearest Neighbor, International Journal of Information Technology and Computer Science (IJITCS), 7(3), 27- 32.
12. De, A., Saha, A., and Pal, M. C., 2015, A Human Facial Expression Recognition Model Based on Eigen Face Approach, Procedia Computer Science, 45, 282-289.
13. Chao, W. L., Ding, J. J., and Liu, J. Z., 2015, Facial expression recognition based on improved local binary pattern and class-regularized locality preserving projection, Signal Processing, 4(7), 1-39.
14. Yu, H., and Liu, H., 2015), Combining appearance and geometric features for facial expression recognition, In Sixth International Conference on Graphic and Image Processing , 944308-944308.
15. Zhang, L., and Tjondronegoro, D., 2011,. Facial expression recognition using facial movement features, IEEE Transactions on Affective Computing, 2(4), 219-229.
16. Sriram A, Sekhar Reddy G, Anand Babu GL, Bachanna P, Gurpreet SC, Moyal V, Shubhangi DC, Sahu AK, Bhonsle D, Madana Mohana R, Srihari K, Chamato FA. A Smart Solution for Cancer Patient Monitoring Based on Internet of Medical Things Using Machine Learning Approach. Evid Based Complement Alternat Med. 2022 Jun 24;2022:2056807. doi: 10.1155/2022/2056807. PMID: 35783507; PMCID: PMC9249483.
17. Maheshwari, V., Mahmood, M. R., Sravanthi, S., Arivazhagan, N., ParimalaGandhi, A., Srihari, K., & Sundramurthy, V. P. (2021). Nanotechnology-Based Sensitive Biosensors for COVID-19 Prediction Using Fuzzy Logic Control. Journal of Nanomaterials, 2021