



UNSUPERVISED FACIAL EXPRESSION DETECTION USING GENETIC ALGORITHM

R. Dembani, W. Zheng*, M. Sun*, Nooruddin**

Abstract: Interpersonal communication can be done by understanding the clues of facial expressions. As its importance increase in behavior and clinical studies, so automatic detection of facial expressions is an open research area for the last few decades. Efforts of expression detection by a human being are easy and effective but the machine needs some more understanding. This paper proposes a face expression clustering using a genetic algorithm. Image get convert into binary format for finding the related cluster selection in different phases of genetic algorithm. Proposed work has utilized a modified teacher learning-based optimization algorithm where the population gets updated in each phase to get the best representative features. A real dataset of facial expression was used in this work. A comparison of the proposed model was done with existing models on different evaluation parameters. It was obtained that the proposed work has improved precision, recall, the accuracy of facial expression identification without any training.

Key words: *facial expressions, facial sentiment detection, digital image processing, color format, genetic algorithm*

Received: October 21, 2019

DOI: 10.14311/NNW.2020.30.005

Revised and accepted: February 29, 2020

1. Introduction

The facial motion and expressions play a major part in conveying the message while interacting with others, and so we can say that the facial features make an important contribution in Interactions between people. Knowledge of facial expression with greater precision is still a challenging job owing to the complexity and variation of facial expressions even after a lot of prior effort.

Computer vision is a method of extracting spatial data from a 2d picture. Recognition of facial expression uses its feature to identify if a person lies or detects emotions. Processing the entire picture might be computationally costly so that only a viable portion of the picture is considered. The primary area of concern is that those with mouth, nose, eyes, and eyebrows are the extracts in the face picture. Communication between people can either be non-verbal or verbal. It can be done

*Rahool Dembani – Corresponding author; Wang Zheng; Meijun Sun; Nooruddin; College of Intelligence and Computing, Tianjin University, Tianjin, China, E-mail: rahuldembani@tju.edu.cn, wzheng@tju.edu.cn, sunmeijun@tju.edu.cn nooruddin@ibacc.edu.pk

not only by Words as well as by facial expressions, feelings and head motion. For the purpose of evaluation, the seven main basic emotions such as angry, fear, disgust, surprise, neutral, happy, sadness are considered for the purpose of assessment [2].

The Facial Action Coding System (FACS) is a famous system of facial coding, used by computer scientists and leading psychologists like [1,3]. In order to classify emotions, the system utilizes action units describing motions of certain muscle groups and facial muscles. Action units detail specifics of the facial motion, such as the raising of the internal or pulling or puckering lips or outer brow or dilating cheeks, as well as optional intensity data for such movements. Facial Activity Coding System appears discrete and perceivable facial motions and manipulations in understanding with the interest of emotions, visual facial feature, and digital image processing analysis can enable predictors to be trained for effective facial expression.

Facial expression (FE) identification system is fast becoming a well-known feature for many distinct reasons in ‘applications’ and websites [4, 5]. In addition, the individuality of the facial feature is very efficient in the biometric credentials that automatically identifies a individual person from a digital image or a video image. Facial expression is not only used to express our feelings, but also to provide significant talkative indications during social interaction, Like our attention rate.

Most sentiment-based image classification methods have made many complicated calculations, such as the number of features gathered to identify the right class [6,7]. This improves the work’s execution time. Predicting sentiment class for the unknown picture leads to a false alarm in the case of the supervised learning model. So following were the objective of this dissertation:

- Reduce Execution time for identifying facial sentiment.
- Accuracy of Classification of images into right sentiment should be increased.
- Develop an algorithm which not need any prior training or do unsupervised classification.
- Reduce the calculation complexity of the sentiment classification.

2. Related work

So far, the recognition of distinct facial expression is commonly based on six fundamental emotions or its subsets: happiness, fear, sadness, anger, surprise, disgust. Qi C. et al. [8] have proposed a paper on Recognition of facial expression based on binary pattern and cognition. In texture feature extraction local binary pattern has an advantage. Facial images have been taken from volunteers and made smooth as noise is being removed. The confusion matrix has been employed. The basic emotion classification model and the circumflex emotion model are discussed and contrasted on the basis of dimension space theory. And the accuracy of recognition of the circumflex emotion model is higher. [15] This based on facial expression recognition and evaluated the efficiency of the well-known Local Binary Patterns (LBPs) feature extractor under different facial expressions.

Uddin M.Z., et.al. [9] Have a paper on recognition of facial expression using the silent feature and convolution neural network. They have considered a waste amount of data set. The proposed method based on feature extraction has tolerance against illumination variation. Besides, the features are adjusted for testing and training using CNN-based deep learning. Compared to other traditional approaches on various databases, the proposed method showed its superiority over other techniques. Ding Y., et.al [10] have proposed Face Expression Recognition Paper based on LBP and expansion of Taylor's picture sequence. In this paper peak frame detection from the image, the sequence is performed. They have divide image into Nodes and apply the double LBP on each pixel. Then build a histogram for each block and then concatenate this to a vector. Ghimire D. et al. [14] extracted region-specific appearance aspects by way of separating the whole face area into domain-specific close by regions. Important local areas are identified by incremental search strategy, resulting in feature sizes being reduced and recognition accuracy has been an enhancement.

In [11] This paper deals with the problems of how to customize the test sample generic model without information on the test label. Weighted Center Regression Adaptive Feature Mapping (W-CR-AFM) It is primarily suggested that the feature distribution of test samples should be transformed into the feature distribution of trained samples. According to minimizing the error between each test sample feature and the most appropriate category center, W-CR-AFM can take the features of test samples to the centers of expression categories around the decision boundary; therefore, their expected labels can be fixed

3. Proposed methodology

The entire work is explained in this section as shown in Fig. 2. This work has categorized data, acquired from the picture of the face with the tag of emotion information. This work is explained in modules for initially is the population generation side and then distinctive cluster set with feature values was captured in the image. While teacher Learning Based Optimization (TLBO) was updated or modified as per the genetic algorithm in the second module population set.

3.1 Pre-Processing

The data set is resized into the fixed aspect of the facial region in this step input image. As some portion of the image like hair, background increase noise for expression identification. In this step feature from the image are extracted by converting input gray image into binary format. Here 1 represents the edge portion of the image and 0 represents the non-edge region. To identify the edges in the image, use the Canny algorithm to convert it into a gray format. That converting a gray picture into a binary picture, So an analysis of each picture element is performed for this.

Steps of pre-processing were shown in Fig. 1 where Fig. 1(a) represent input image, (b) is binary image of input. Finally Fig. 1(c) is face region of input image.

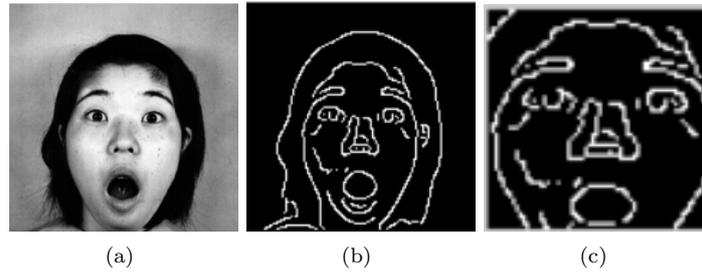


Fig. 1 Pre-Processing steps of input dataset.

3.2 Generate population

In this step, different chromosome set was generated using Eq. (1). Each chromosome is a set of image feature where as per the number of expression numbers of images were selected. So each expression feature set act as the chromosome while the collection of all set is termed as population. This can be assumed as let $Cc = [F_1, F_2, \dots, F_m]$ as the chromosome set where m is a number of expressions in a set, while $P = [Cc_1, Cc_2, \dots, Cc_n]$ n is number of chromosomes.

$$P \leftarrow \text{Random}(n, m) \tag{1}$$

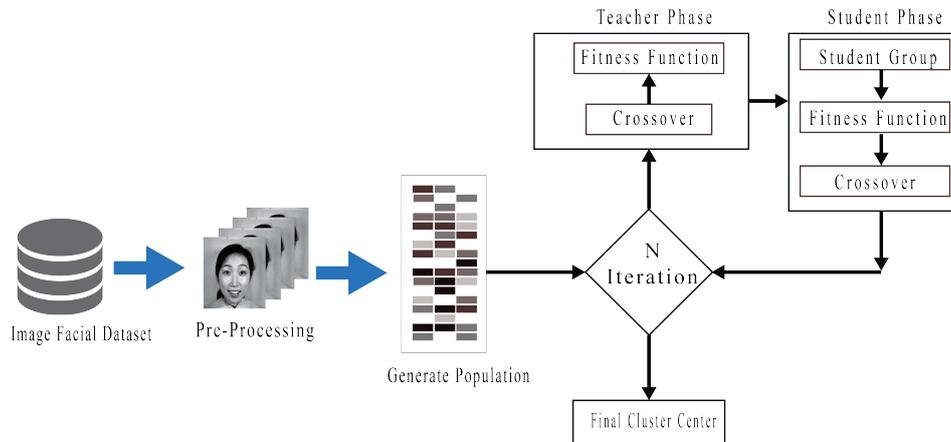


Fig. 2 The proposed work block diagram.

3.3 Fitness function

The set of chromosomes present in the population were evaluated on the fitness function values. In this work, In this function the center set is transferred in order to get a good chromosome from the bunch of available set fitness values of each probable cluster. So the value of fitness returns. The Euclidean distance

formula has been used to calculate the fitness value. This can be understood as it is appropriate to measure the fitness value of the cluster set $Cc = [C1, C2]$. The distance from each non-cluster center image has been evaluated. So the distance value of each cluster center form the elements of the cluster is sum up to get the fitness value of the chromosome. This can be understood as

$$D_{m,t} = |Cc_f - I_f|. \quad (2)$$

In above Eq. (2) t is number of images in the dataset, while f is feature value.

Algorithm 1 Fitness function algorithm.

Input: I, P // I: Image dataset feature, P: Population

Output: F // F: Fitness value

```

while  $n \neq 0$  do
  /* t: number of images in dataset */
  while  $t \neq 0$  do
    while  $m \neq 0$  do
      D[m] ← Dist( $P[n, m], I[t]$ )
    end while
    S[t] ← Minimum( $D$ )
  end while
  F[n] ← S
end while

```

3.4 Teacher phase

Here it is the best $C_{teacher,i}$ solution from P which acts as a teacher and selection based entirely on the fitness value. After sorting, the highest possible solution will operate as the teacher for other feasible solutions. Presently chosen teacher will educate other conceivable solutions $C_{student,i}$ by means of replacing expression feature value existing in teacher solution. By this, solution that operates as a student will learn how to behave as a teacher from the best solution. The random position value of the crossover operation is replicated from the teacher chromosome and substituted by the non-teacher chromosome by Eq. (3) [12]. This improves the population quality.

$$C_{new,i} \leftarrow C_{teacher,i} \quad i \in \{1, 2, \dots, m\}, \quad (3)$$

where $C_{new,i}$ is $C_{student,i}$ updated value. Accept $C_{teacher,i}$ value.

3.5 Student phase

Population in this phase processes into a group of two or more chromosome. The fitness value of each group chromosome was evaluated and fittest chromosome perform crossover operation as done in the teacher phase. The population gets an update in this phase if offspring fitness value is better as compared to the previous chromosome. Each new chromosome was checked for better fitness than

before, if fitness increases then the population includes fresh chromosomes and the older one is removed. If the fitness value does not improve, vice versa.

3.6 Final solution

TLBO algorithm is completed if two consecutive iterations give the same cluster center set of chromosome. So obtained the final cluster center classify all other images of the dataset. Each cluster represents one facial expression.

Algorithm 2 Proposed algorithm.

Input: t : number of images in dataset, m : number of expression,

B : BestSolutionactasTeacher, I : Pre Processing(Image Dataset)

Output: C // Cluster

```

while  $t \neq 0$  do
  while  $m \neq 0$  do
     $D[m] \leftarrow \text{Dist}(B, I[t])$ 
  end while
   $m \leftarrow \text{Minimum}(D)$ 
   $C[m] \leftarrow I[t]$ 
end while

```

4. Experiment and result

Using the MATLAB software, all calculations and utility measures were performed. The experiments were performed out on an Intel Core i5 2.27 GHz computer equipped with 8 GB RAM and running on Windows 7 Professional.

4.1 Dataset

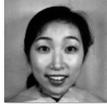
Consider well-known outward-looking information collected from JAFFE Dataset at that stage for an image. In the real image, the JAFFE database contains 213 image records of different outward appearances where basic six are common expressions and one is outrageous [13]. Ten young Japanese ladies' posture was gathering here. Tab. I shows a sample of image dataset with various classes.

4.2 Evaluation parameter

There are many parameters for evaluating outcomes, such as F-score, recall, accuracy, etc. For getting the superior results obtain values can be put mention parameter formula

$$\text{Precision Value} = \frac{\text{True_Positive}}{\text{False_Positive} + \text{True_Positive}}$$

$$\text{Recall Value} = \frac{\text{True_Positive}}{\text{False_Negative} + \text{True_Positive}}$$

Expression Class	Angry	Disgust	Fear	Happy	Neutral	Sadness	Surprise
Image Set							
Image Set 2							

Tab. I Image dataset example.

$$F_Score = \frac{2 \times \text{precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

$$\text{Accuracy Value} = \frac{\text{Correct_Classification}}{\text{Incorrect_Classification} + \text{Correct_Classification}}$$

4.3 Time execution

We understand that time execution is very essential in any procedure and it must be less. Classifying user data on the server is the algorithm's time period. It is an exceptionally significant parameter. In terms of seconds the execution time is measured.

4.4 Results

Results of the proposed graph-based clustering of bots and real user were compared with the existing method developed.

Fig. 3 and Tab. II shows that Precision value compared to the previous algorithms, It can be observed that the method of selection of centroid in the proposed work is as effective as in the previous one. Here iteration increases the accuracy

Expression Class	Khorsheed [15]	Wu [11]	Proposed Work
Angry	0.4556	0.7273	0.9000
Disgust	0.3499	0.5000	0.7500
Fear	0.5600	0.6223	0.8571
Happy	0.7111	0.4000	0.9000
Neutral	0.5565	0.8571	0.8571
Sadness	0.4999	0.5000	0.9000
Surprise	0.6000	1.0000	1.0000

Tab. II Precision value comparison of previous and proposed work.

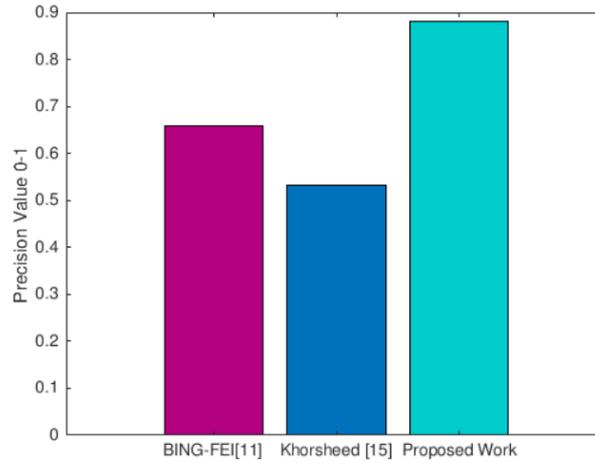


Fig. 3 Precision Average value comparison of previous and proposed work

value in both works, but choosing the distinct set of clustering characteristics makes the proposed work highly accurate.

Fig. 4 and Tab. III shows that Recall value compared to previous algorithms, Compared to the previous one, it can be observed that the proposed technique of selecting a centroid job is effective. The decrease of the feature vector here enhanced the precision of the proposed work while the recall value was also enhanced by two-phase learning of the suggested TLBO algorithm.

Expression Class	Khorsheed [15]	Wu [11]	Proposed Work
Angry	0.6500	0.8889	1.0000
Disgust	0.3338	0.6222	1.0000
Fear	0.4555	0.5666	0.6667
Happy	0.3422	0.6222	0.6667
Neutral	0.5633	0.3333	0.6667
Sadness	0.2333	0.3333	0.3333
Surprise	0.3555	0.4111	0.6667

Tab. III Recall value comparison of previous and proposed work.

Tab. IV shows that compared to previous algorithms, the F-Measure value of the proposed work was high. Here proper leaning feature with pre-processing binary feature increase the efficiency of the work. It has been observed that proposed work two-phase population updates increase efficiency.

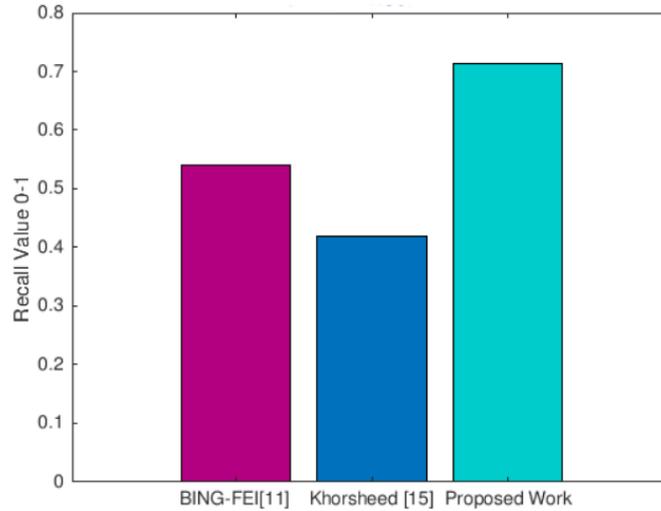


Fig. 4 Average recall value comparison of previous and proposed work.

Expression Class	Khorsheed [15]	Wu [11]	Proposed Work
Angry	0.5999	0.8000	0.9474
Disgust	0.2332	0.3077	1.0000
Fear	0.4111	0.6500	0.8000
Happy	0.6115	0.2857	0.7500
Neutral	0.3443	0.4615	0.7500
Sadness	0.1444	0.4000	0.4615
Surprise	0.3443	0.2000	0.8000

Tab. IV *F*-Measure value comparison of previous and proposed work.

Fig. 5 and Tab. V shows that execution time in proposed work is less compared to the previous algorithm. As we saw that proposed work one too many different picture sets classification is better than previous information. Reducing the vector feature decreased the time of execution in proposed work.

Number of Iteration	Khorsheed [15]	Wu [11]	Proposed Work
5	1.9744	2.4744	1.2135
10	3.7932	3.4932	2.6667
15	4.5111	6.5111	3.3334

Tab. V Average execution time in seconds comparison of previous and proposed work.

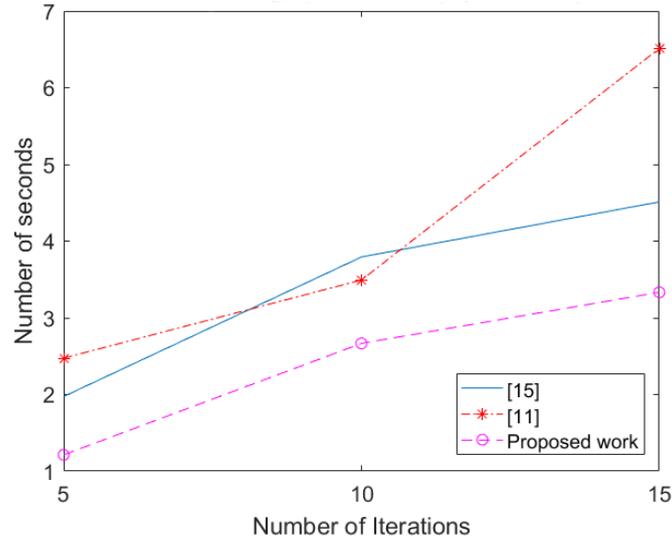


Fig. 5 Average execution time in seconds comparison of previous and proposed work.

5. Conclusion

In a scene with little or no effort, human easily detects and identifies faces and facial expressions. It is rather hard to develop an automated system that performs this job. There have been various approaches to robust facial expression recognition, distinct picture acquisition, and extraction, analysis, and classification techniques. This work presents a study of methods for the detection of user facial expression detection using a binary format of edge feature in the genetic algorithm. Here fitness value of the chromosome was an estimate and two-phase population updating increases the cluster selection accuracy of the work Results show that the proposed genetic algorithm-based clustering technique has increased the precision value by 35.36 % as compared to the previous approach. While recall value was also increased by 57.78 %, at the same time accuracy of the expression detection was also increase by 50.86 %.

Acknowledgement

The authors wish to acknowledge the support for the research work from the National Natural Science Foundation of China under grant No. [61572351], [61772360], and [61876125].

References

- [1] EKMAN P., KELTNER D.: Universal facial expressions of emotion: An old controversy and new findings, Lawrence Erlbaum Associates, Inc., 1997.

- [2] MATSUMOTO D., KUPPERBUSCH C. Idiocentric and allocentric differences in emotional expression, experience, and the coherence between expression and experience, *Asian Journal of Social Psychology*, Wiley Online Library, 2001, 4, pp. 113–131.
- [3] LUCEY P., COHN J.F., KANADE T., SARAGIH J.A., ZARA M.I. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression, 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, IEEE, 2010, pp. 94–101.
- [4] KULKARNI K., CORNEANU C., OFODILE I., ESCALERA S., BARO X., HYNIEWSKA S., ALLIK J., ANBARJAFARI G. Automatic recognition of facial displays of unfeared emotions, *IEEE transactions on affective computing*, IEEE, 2018.
- [5] ZHOU X., JIN K., SHANG Y., GUO G. Visually interpretable representation learning for depression recognition from facial images, *IEEE Transactions on Affective Computing*, IEEE, 2018.
- [6] BARSOUM E., ZHANG C., FERRER C.C., ZHANG Z.: Training deep networks for facial expression recognition with crowd-sourced label distribution, *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, ACM, 2016, pp. 279–283.
- [7] LI S., DENG W. Deep emotion transfer network for cross-database facial expression recognition, 2018 24th International Conference on Pattern Recognition (ICPR), IEEE, 2018, pp. 3092–3099.
- [8] QI C., LI M., WANG Q., ZHANG H., XING J., GAO Z., ZHANG H. Facial expressions recognition based on cognition and mapped binary patterns, *IEEE Access*, IEEE, 2018, 6, pp. 18795–18803.
- [9] UDDIN M.Z., KHAKSAR W., TORRESEN J. Facial expression recognition using salient features and convolutional neural network, *IEEE Access*, IEEE, 2017, 5, pp. 26146–26161.
- [10] DING Y., ZHAO Q., LI B., YUAN X. Facial expression recognition from image sequence based on LBP and Taylor expansion, *IEEE Access*, IEEE, 2017, 5, pp. 19409–19419.
- [11] WU B., LIN C. Adaptive feature mapping for customizing deep learning based facial expression recognition model, *IEEE access*, IEEE, 2018, 6, pp. 12451–12461.
- [12] AMAJAMA J. Effect of air pressure on the output of photovoltaic panel and solar illuminance (or intensity), *International Journal of Scientific Engineering and Applied Science (IJSEAS)*, 2016, 2, pp. 139–144.
- [13] JAFFE Database, <http://www.kasrl.org/jaffe.html>.
- [14] GHIMIRE D., JEONG S., LEE J., PARK S.H. Facial expression recognition based on local region specific features and support vector machines, *Multimedia Tools and Applications*, Springer, 2017, 76, pp. 7803–7821.
- [15] KHORSHEED J.A., YURTKAN K. Analysis of local binary patterns for face recognition under varying facial expressions, 2016 24th signal processing and communication application conference (SIU), IEEE, 2016, pp. 2085–2088.