Intelligent Facial Emotion Recognition

Shahana A., Harish Binu K. P.

Computer Science and Engineering, MEA Engineering College, Perinthalmanna, Kerala, India

ABSTRACT

The system introduces an intelligent facial emotion recognition using artificial neural network (ANN). The concept first takes modified local binary patterns, which involve horizontal vertical and neighborhood pixel comparison, to produce initial facial representation. Then, a microgenetic algorithm(mGA) embedded Particle Swarm Optimization(PSO), is proposed for feature optimization. It incorporates a nonreplaceable memory, a small-population secondary swarm, a new velocity updating strategy, a sub dimension based in-depth local facial feature search, and a cooperation of local exploitation and global exploration search mechanism to mitigate the premature convergence problem of conventional PSO. Artificial Neural Network is used as a classifier for recognizing seven facial emotions. ANN is implemented as classifier for pattern recognition. Based on a comprehensive study using within- and cross-domain images from the extended Japanese database. The empirical results indicate that our proposed system outperforms other state of-the-art PSO variants, conventional PSO, classical GA, and other related facial expression recognition models reported in the literature by a significant margin.

Keywords : Particle Swarm Optimization, Local Binary Patterns

I. INTRODUCTION

Facial emotion recognition has come up with a new view for human computer interaction, and has given asset to a wide range of computer vision applications, such as medical field, surveillance, event detection, personalized learning, and robotics. Proper emotion classification commit heavily on effective facial representation. but, it is still a challenging task of recognising significant facial features that could represent the characteristics of each person. The Way of representing the same emotion varies with person to person. It is because of the personal variation and emotional intensity.

This paper means to manage such difficulties to deliver compelling and streamlined discriminative facial portrayals to benefit continuous outward appearance acknowledgment. In examination with other element choice strategies, evolutionary computational(EC) calculations show amazing world wide inquiry capacities, and have been generally acknowledged as efficient procedures for features election[1]. Among various EC calculations, the particle swarm optimization (PSO) calculation is inspired by the flocking practices of birds, and has been broadly utilized for highlight enhancement with the benefits of a low computational expense and a fast convergence speed. However, conventional PSO tends to converge prematurely and therefore be trapped in local optima. Therefore, in this paper, a PSO variation implanted with the idea of a microgenetic algorithm(mGA) is proposed Known as mGA-embedded PSO, the proposed algorithm incorporates a nonreplaceable memory, a small-population secondary swarm, a new velocity

updating strategy, a subdimension-based regional facial feature search strategy, and a cooperation of local exploitation and global exploration search strategy to overcome both premature convergence and local optimum problems encountered by conventional PSO.

The proposed facial emotion recognition framework comprises of three stages including feature extraction, feature optimization and emotion recognition. First of all, use modified local binary patterns (LBPs), i.e., horizontal and vertical neighbourhood comparison LBP, to extract the initial facial representation. Then, the proposed mGA-embedded PSO algorithm is utilized to distinguish the most discriminative and critical distinct facial expressions. Different classifiers are connected to perceive seven feelings like happiness, sadness, anger, fear, surprise, disgust neutral. The framework is assessed with a facial expression database, a Japanese database which include variety of facial images. State-of-the-art PSO variants, conventional PSO, and classical genetic algorithm (GA) are used to compare with the proposed mGA-embedded PSO algorithm in feature optimization[2]. The experimental outcomes show that the proposed framework beats state-of-the-art optimization techniques and other related facial emotion recognition research work reported in literature review by a great significant edge.

The fundamental commitments of this framework are abridged as pursues. An modified LBP operator that conducts horizontal and vertical neighborhood pixel comparison is proposed, in request to beat the disadvantages of original LBP by recovering the missing difference data inserted in the area to produce the underlying discriminative facial expressions. A mGA-embedded PSO algorithm is introduced for feature optimization, so as to mitigate the premature convergence and local optimum problems of conventional PSO. It gives incredible adaptability to permit the features election procedure

to not just separate facial expressions into explicit regions for in-depth local, but in addition to consolidate facial features for overall global search. The introduced algorithm consists a new velocity updating strategy by employing the personal average experience to produce the individual best, pbest, and Gaussian mutation to produce the global best, gbest, to increases warm
diversity.

The proposed algorithm likewise applies the decent diversity maintenance methodology of mGA to keep the original swarm in a nonreplaceable memory[3], which stays flawless during the lifetime of the algorithm, so as to diminish the probability of premature convergence. So as to accelerate the evolution speed, the little population size idea of mGA is utilized to produce an secondary swarm with five particles. The secondary swarm comprises of the swarm head and four following particles from the nonreplaceable memory with the minimal maximum relationship with the leader to increment local exploitation and global exploration. These local search, global search systems work in a collaborative way to control the search toward global optima. A subdimension-based search strategy is also implemented, in order to point out optimal features for each facial area. The proposed system is evaluated with Japanese database. It out performs state-of-the-art LBP and PSO variants, and other facial expression recognition methods reported in the literature significantly.

II. METHODS AND MATERIAL

Pre-processing of Image

- **Resizing of image**

Here, input to the system is an image of any emotion, which is to be recognized. Resizing of image is very important since the input of the system is an image. The system requires standard images to be processed,

and is not compulsory that every image subjecting to the system is unique in size. So, resizing is the first step to be considered to make the image in standard size so that the system can be processed easily.
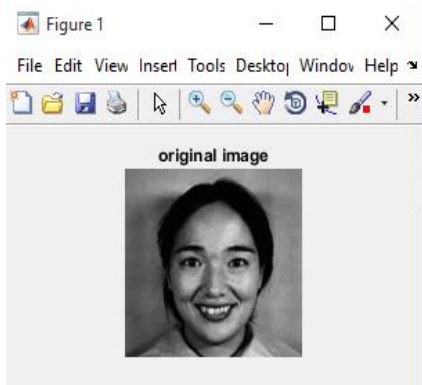

Figure 1. Image to be Recognized

- **Detection of face**

The image inputting might contain backgrounds which are avoidable for the purpose. So, it is important to find out the face area from the image. For that Harr Feature algorithm is used. Haar-like features are digital image features used in object recognition. They owe their name to their instinctive comparability with Haar wavelets and were utilized in the main ongoing face detector. A Haar-like feature thinks about adjoining rectangular region at a particular area in a location window, sum up the pixel powers in every region and figures the contrast between these totals. This difference is then used to categorize subsections of an image. For example, let us say we have an image database with human faces. It is a typical perception that among all faces the area of the eyes is darker than the area of the cheeks. In this manner a typical Haar feature for face detection is a group of two adjoining rectangles that lie over the eye and the cheek region. The position of these square shapes is characterized with respect to a location window that demonstrations like a bouncing box to the objective item (the face for this situation).
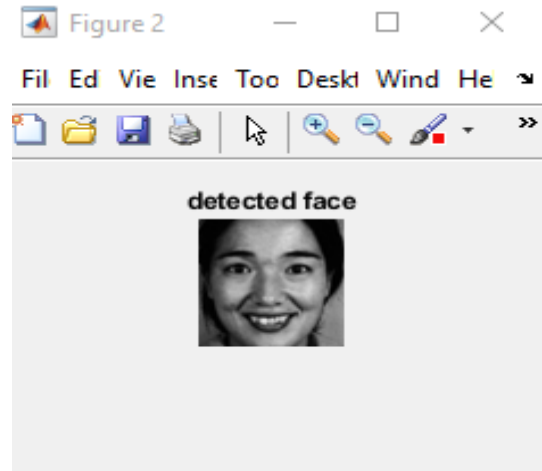

Figure 2. Detected Face

- **Resizing of face**

Since the face of every person differs in size and shape, the detected faces will not having the standard size. So, resizing should be done again to make the detected face image to a standard size.

- **Removal of intensity variation**

The same image taking in and out of a room differs in the light intensity. This variation should be removed before the system processing. Here Histogram Equalization technique is adopted to remove the intensity variation over the image. Histogram equalization is a method for adjusting image intensities to enlarge contrast. This technique usually enhance the global contrast of many images, specially while the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the histogram. This makes the areas of lower local contrast to attain a higher contrast.

Histogram balance achieves this by viably spreading out the most continuous power esteems. The technique is beneficial in images with backgrounds and foregrounds that are bright or both dark. In particular, the idea can lead to good views of bone structure in x-ray images, and to better information

in photographs that are over or under-exposed. A key benefit of the idea is that it is a fairly straightforward technique and an invertible operator. So in theory, if the histogram equalization function is transparent, then the original histogram can be recovered. The estimation isn't computationally concentrated. A weakness of the strategy is that it is aimless. It might build the difference of foundation clamor, while diminishing the usable sign.
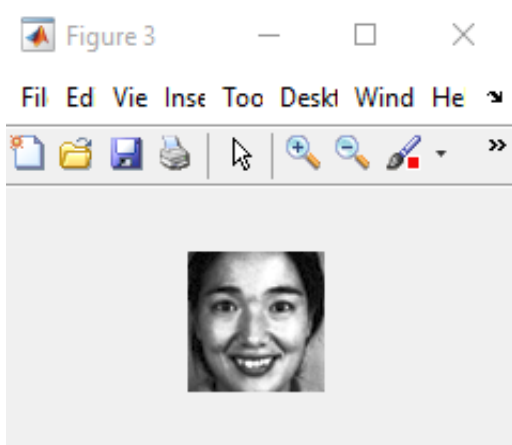

Figure 3. Equalising Intensity Variation

- **Removal of noise**

Noises might include in an image because of many reasons such as dust in the camera etc. Removing the noise from the image is useful for further processing of the image without any interruption. Here Bilateral Filter is used for removing the noise. A bilateral filter is a non-linear, edge-protecting, and noise-diminishing smoothing filter for images. It replaces the intensity of each pixel with a weighted average of intensity values from neighbourhood pixels. After these all preprocessing image undergoes feature extraction and all.
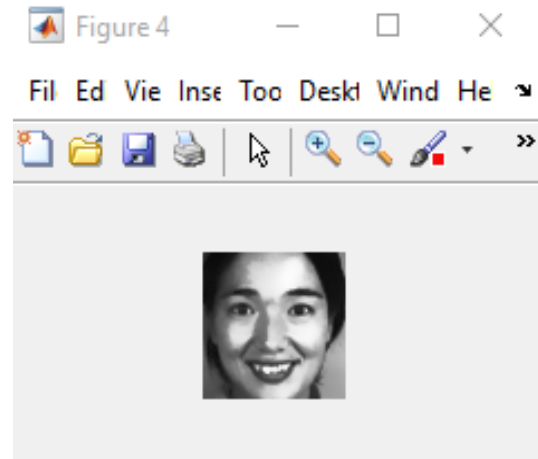

Figure 4. Removing Noise

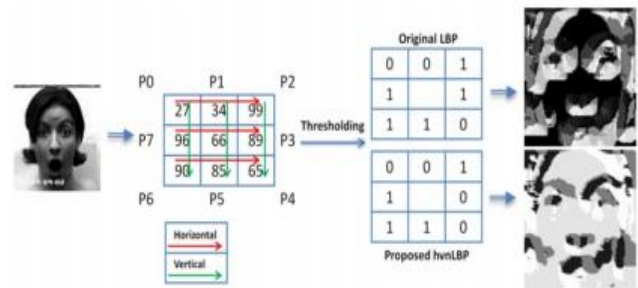**Facial Feature Extraction Using the Proposed LBP**

Figure 5. Proposed hvnLBP operator in comparison with that of the original LBP

$$hvnLBP_{p,r} = \{S(\max(l0,l1,l2)),S(\max(l7,l3)), S(\max(l6,l5,l4)),S(\max(l0,l7,l6)), S(\max(l1,l5)),S(\max(l2,l3,l4))\}$$

Where,

$S(\max(lj, lk, lm)) = 1$; if maximum
  0; if non maximum

So as to improve the discriminative capacities of LBP, propose horizontal and vertical neighborhood pixel comparison LBP (hvnLBP). It is coordinated with the Gabor filter for delivering the discriminative facial representation. There are four stages in the feature extraction process: preprocessing for light changes and noise invariance; face identification; Gabor magnitude image generation; the proposed hvnLBP-based textural depiction. First of all, apply

equalization and bilateral filter to remunerate enlightenment varieties and reduce noise in the input picture, separately. At that point utilize a Haar-cascade face detector to recognize faces. A 2-D Gabor filter is also used to produce magnitude pictures. At last, the proposed hvnLBP operator is used to create the textural description of facial images.

As a known texture descriptor, LBP takes a circular neighborhood for feature extraction[4]. This original LBP operator process a comparison between the central pixel and the eight neighborhood pixels, so there is no chance to lose the contrast information among the neighborhood pixels. To solve this problem, propose hvnLBP to capture missing contrast information among the neighborhood pixels. Instead of comparing with the central pixel as in original LBP, hvnLBP uses horizontal and vertical neighborhood pixels for direct comparison to generate the resulting textural descriptions. As an example, employ P = p0, p1, p2, p3, p4, p5, p6, p7 to represent the eight neighborhood pixels in LBP. In either vertical or horizontal comparison, the values of the vertical or horizontal neighbouring pixels are compared with one another. A 1 is assigned to the pixel with the highest value and a 0 is assigned to the remaining pixels. This horizontal and vertical comparison process can be conducted in any order, i.e., horizontal comparison followed by vertical comparison, or vice versa.

Overall, in comparison with the original LBP operator, the entire results indicate that hvnLBP is more capable of capturing discriminative contrast information such as corners and edges among neighborhoods to inform subsequent PSO based feature selection and facial expression analysis than conventional LBP.
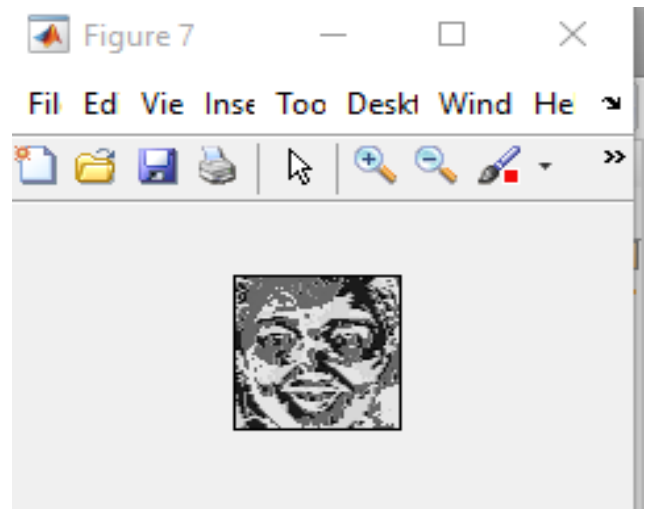

Figure 6. Image after hvnLBP

Proposed PSO Algorithm for Feature Optimization

Particle swarm optimization (PSO) is a computational strategy that advances an issue by iteratively attempting to improve a candidate solution with regard to a given measure of quality. It takes care of an issue by having a population of candidate solutions, here named particles, and moving these particles around in the search-space according to simple mathematical formulae over the particle's position and velocity. Every particle's development is impacted by its its local best known position, but at the same time is guided toward the best known positions in the search-space, updated as better positions are found by other particles. This is expected to move the swarm toward the best solutions.

To identify the discriminative characteristics of each expression, we propose a PSO variant embedded with the concept of mGA for feature optimization, this proposed PSO algorithm mitigates the premature convergence problem of conventional PSO, and shows superior capabilities of discriminative feature selection. The proposed mGA-embedded PSO algorithm employs personal average experience and Gaussian mutation for velocity updating.

Furthermore, it integrates the diversity maintenance strategy of mGA to keep the original swarm in a nonreplaceable memory, which remains intact during the lifecycle of the algorithm to increase swarm diversity. Inherited from the concept of mGA, a secondary swarm with a small population size of five particles is employed. The swarm comprises a leader and four follower particles with the highest or lowest correlation to the leader from the nonreplaceable memory to increase local and global search capabilities and avoid premature convergence. Moreover, the algorithm separates facial features into specific areas for in-depth local subdimension-based search. Overall, the local exploitation and global exploration search strategies of the algorithm work cooperatively to lead the search process to the global optima.

Besides the velocity updating mechanism, the proposed PSO algorithm integrates the concepts of mGA and a secondary swarm, as well as the cooperation of local exploitation and global exploration search strategies to balance between convergence speed and swarm diversity. In summary, the proposed algorithm employs the diversity maintenance strategy of mGA using a non-replaceable memory. This non replaceable memory comprises the initialized swarm to sustain search diversity. Motivated by the small population size concept of mGA, a secondary swarm with five particles comprising the swarm leader and four follower particles from the non-replaceable memory with the highest or lowest correlation with the leader is constructed to increase local exploitation and global exploration. A sub dimension-based search in the secondary swarm is also conducted, in order to identify the discriminative regional facial features. Moreover, the local exploitation and global exploration search strategies of the secondary swarm work in a collaborative manner to avoid stagnation and overcome premature convergence.
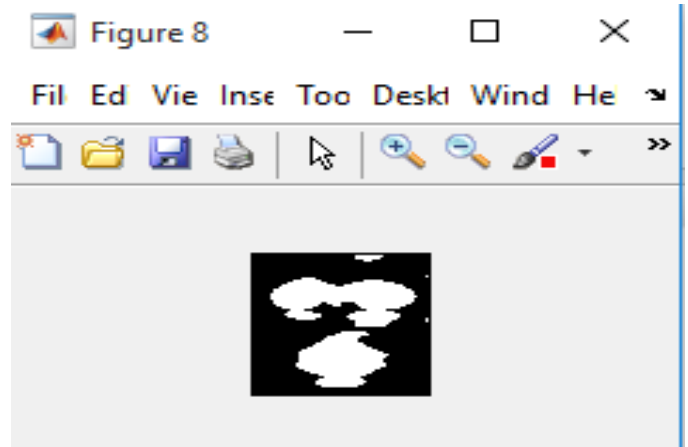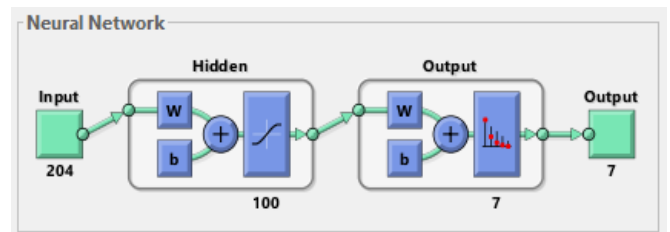

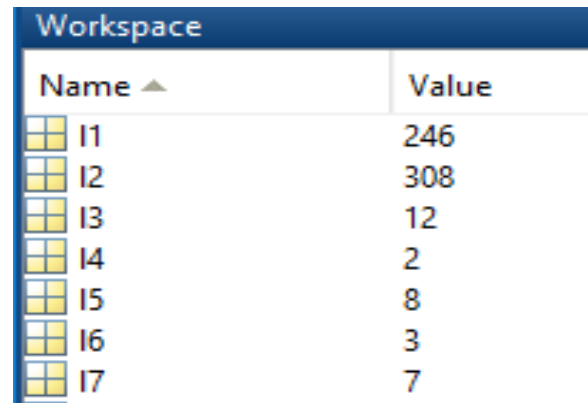**Figure 7.** Image after PSO

Emotion Recognition by ANN

Lead an investigation of seven-class facial emotion recognition utilizing the features automatically created by the mGA-embedded PSO. ANN can be used as a classifier and also for training and testing in this facial emotion recognition system. Artificial neural networks (ANN) or connectionist systems are computing systems vaguely inspired by the biological neural networks and astrocytes that constitute animal brains. Such frameworks "learn" to perform tasks by considering examples, for the most part without being customized with any undertaking task-specific rules. For instance, in image recognition, they may figure out how to distinguish pictures that contain cats by analysing example pictures that have been physically named as "cat" or "no cat" and using the results to identify cats in other images. They do this without having any initial knowledge about cats, for example, that they have fur, tails, whiskers and cat-like faces. Instead, they automatically create

recognition features from the learning material that they process.

An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal from one artificial neuron to another. An artificial neuron that receives a signal can process it and then signal additional artificial neurons connected to it. In common ANN implementations, the signal at a connection between artificial neurons is a real number, and the output of each artificial neuron is computed by some nonlinear function of the sum of its inputs. The connections between artificial neurons are called 'edges'. Artificial neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Artificial neurons may have a threshold such that the signal is only sent if the aggregate signal crosses that threshold. Typically, artificial neurons are aggregated into layers. Different layers may perform different kinds of transformations on their inputs. Signals travel from the first layer (the input layer), to the last layer (the output layer), possibly after traversing the layers multiple times.

The system has trained by using ANN with number of images so that the system conclude with better result. Here ANN is act as a tool for pattern recognition. In this case features of every emotion is of pattern based. Seven emotion will be with seven distinct patterns. Identifying that pattern is happening in the case of testing. The feature distribution is shown in the following figure. Here l1, l2, l3, l4, l5, l6, l7 represents surprise, happy, sad, neutral, fear, disgust and angry respectively.

Figure 8. Matching Features

Model Implementation

The system is implemented using MATLAB. MATLAB is a multi-paradigm numerical computing environment and proprietary programming language developed by Math Works developed by Cleve Moler. MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces and interfacing with programs written in other languages, including C, C++, C#, Java, Fortran and Python. A Japanese database is used in the system for testing and training.

III. RESULTS AND DISCUSSION

Conduct a study of seven-class facial emotion recognition using the features automatically generated by the mGA-embedded PSO and Artificial Neural Network. Horizontal, Vertical and neighbourhood pixel comparison is proposed for extracting the features. Artificial neural network is used for classification and pattern recognition. Here, seven emotion can easily recognize by inputting an image. Training is one of the most essential part of the system. Training with more images will positively reflect on the output. The system guarantees a level of accuracy as possible which is closely depends on the training data set. If any plan exists to

commercialize the product, it is compulsory to enlarge the training data set as much as possible.
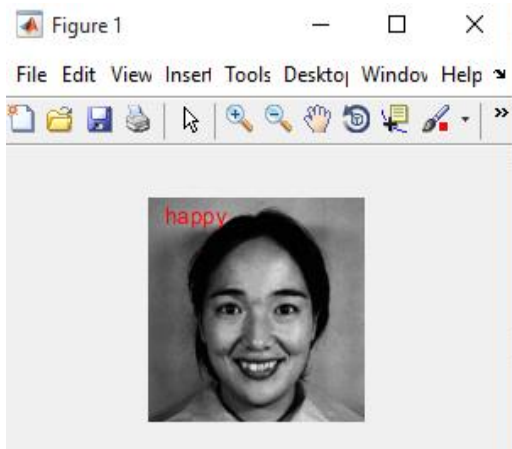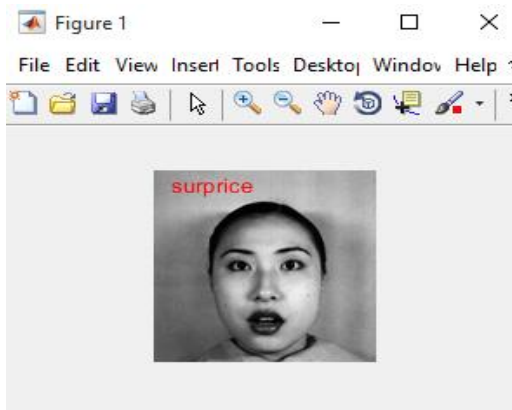
A. Results of the System

Figure 9. Happy
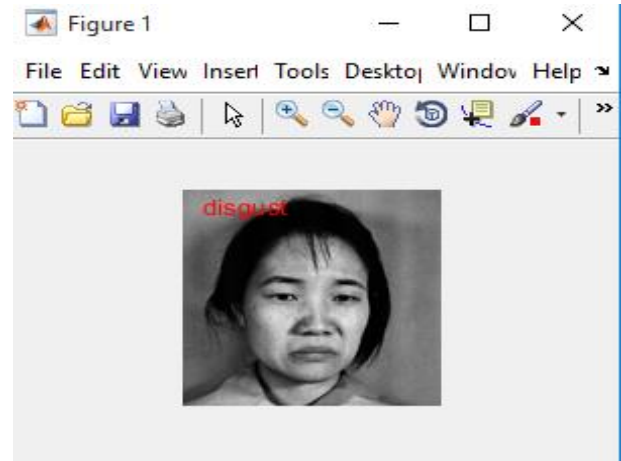

Figure 10. Surprise

Figure 11. Angry

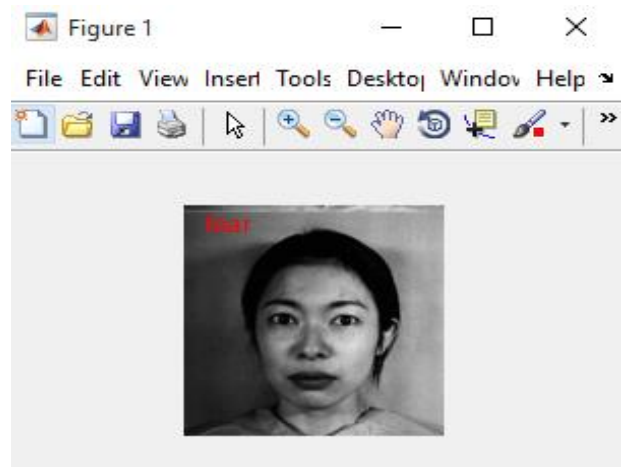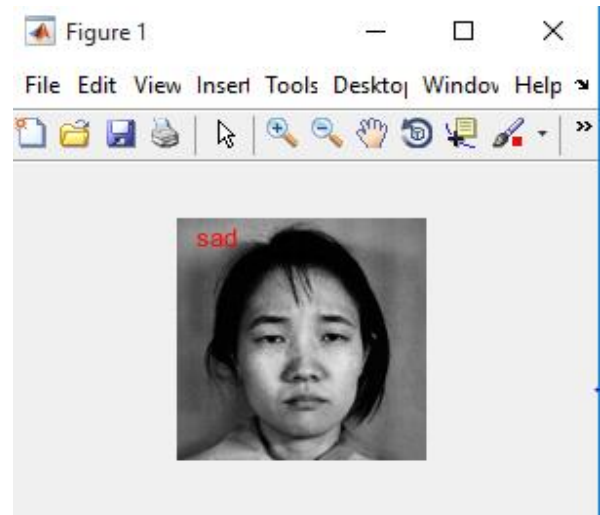**Figure 12.** Disgust

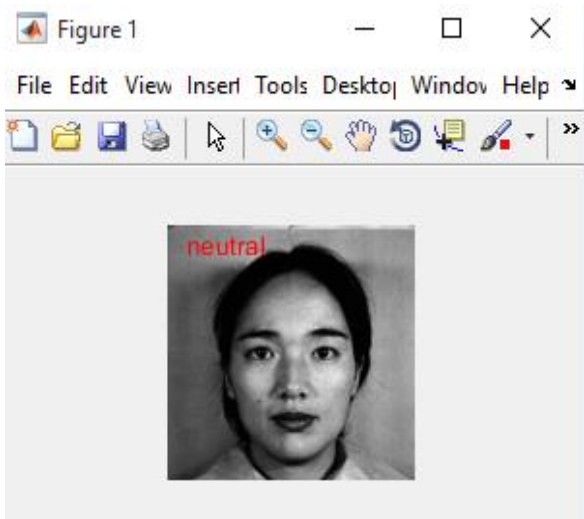**Figure 13.** Fear

Figure 14. Sad

Figure 15. Neutral

The above figure shows the identification of seven emotions.

B. Performance of Proposed Classifiers

The performance graph of an emotion smile that got at the end of the system running. Here at the X-axis is marked as cross entropy and Y-axis is denoted as number of epochs. Cross entropy can be described as the error occurred while running the system. Here performance is measured in terms of error. Lesser error; Better result. The figure shows best validation performance of 0.10711 at epoch 138. The ideal facial emotion recognition system will result with the best validation performance of 0.00 at epoch n.
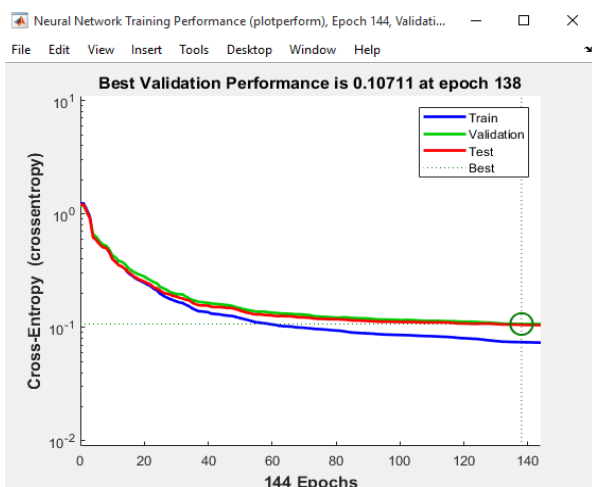

Figure 16. Performance Graph

IV. CONCLUSION

Proposed a facial expression recognition framework with hvnLBP based feature extraction, mGA-embedded PSO-based feature optimization and diverse analyser based expression recognition. The proposed hvnLBP administrator performs horizontal and vertical neighborhood pixel correlation with recover the underlying discriminative facial features. It outperforms state-of-the-art LBP, LPQ, and conventional LBP for texture analysis. Moreover, a new PSO algorithm, i.e., mGA-embedded PSO, has been proposed to solve the premature convergence problem of conventional PSO in terms of feature optimization. The mGA-embedded PSO algorithm incorporates personal average experience and Gaussian mutation for velocity updating as well as employs the diversity maintenance strategy of mGA by keeping the original swarm in a nonreplaceable memory, which remains intact during the lifecycle of the algorithm to increase swarm diversity. Furthermore, it also keeps up an secondary swarm with a little population size of 5 to host the swarm head and four adherent particles with the most elevated/least connection with the head from the nonreplaceable memory to increase local and global search capabilities. The algorithm subsequently separates facial features into special areas for in-depth local sub dimension based search. Overall, the local exploitation and global exploration search mechanisms of the algorithm work cooperatively to guide the search toward the global optimal solutions. The empirical results indicate that our PSO algorithm outperforms other state-of-the-art PSO variants and conventional PSO and GA for optimal feature selection.

V. REFERENCES

[1]. Xue, B., Zhang, M., and Browne, W. N. (2012). Particle swarm optimization for feature selection in classification: A multi-objective

approach. IEEE transactions on cybernetics, 43(6):1656–1671.

[2]. Mistry, K., Zhang, L., Neoh, S. C., Lim, C. P., and Fielding, B. (2016). A microga embedded pso feature selection approach to intelligent facial emotion recognition. IEEE transactions on cybernetics, 47(6):1496–1509. (ICB). IEEE, 2015.

[3]. Coello, C. A. C. C. and Pulido, G. T. (2001). A micro-genetic algorithm for multiobjective optimization. In International Conference on Evolutionary Multi-Criterion Optimization, pages 126–140. Springer.

[4]. Heikkilä, M., Pietikäinen, M., and Schmid, C. (2009). Description of interest regions with local binary patterns. Pattern recognition, 42(3):425–436.

Cite this article as :