

Multiple Object Tracking Using Hybrid Neuro Fuzzy Network Applied to Face Recognition from Feed Forward Neural Network

Mr. Chethan Chandra S Basavaraddi¹, Dr. Subhas Singh Parihar²

Department of Computer Science, North East Frontier Technical University,
Sibu-Puyi, Aalo (P.O), West Siang (Distt.), AP, India^{1,2}

Abstract: In the present study we present an innovative approach towards countering the problem of partial occlusion in face recognition scenario. The partial occlusion can be caused by various objects such as scarfs, sunglasses etc., and its effects are confounding in the performances of the recognition rates. The advantage that the adopted pre-processing algorithm poses before face recognition steps is to eliminate the distortions due to the variance in light illumination field at the given instance when the facial image is recorded or captured. The framework tends to mathematically model the curvature and other essential features of the face such as micro-expression and the curves of the facial regions. This, significantly enhances the probability of matching the parent image to that of the occlude image that is how multiple object recognition using hybrid approach. The presented algorithm is tested over Extended Yale B & CMU PIE standardized datasets. Over the years biometrics has gained unparalleled popularity in digital medium and has proven its usefulness for several applications concerned with the threats and crime or security purposes. Face Recognition is a widely emerging biometric for automating the surveillance, as it has aid in strengthening the security from several types of terrorist or criminal threats. Though, there are several face recognition techniques which are categorized based on its error rates in recognition but there are few that gives the marginal rate for sufficient and validated recognition rates for occlude images.

Keywords: Hybrid Neuro Fuzzy Network, Face Recognition, Neural Network, Normalized Rapid Descriptor (NBD), PCA, LDA, NMF, LNMF, ICA.

I. INTRODUCTION

A Computer vision is connected to industry computerization, for example, review of parts and protest acknowledgment. With progression in assembling innovation, creation has turned out to be more mechanized and speedier than any other time in recent memory. For continuous applications, the examination speed and precision are significant worries in protest review. Machine vision has turned out to be well known in new assembling enterprises, where picture arrangement and acknowledgment calculations are helpful for some process. Object recognition builds robots for performing assignments. Object recognition explore to an overlook the problems in household environment. The present study addresses the issue related to the application of these procedures to household situation. Here technique related to object recognition selects reasonable model to be adjusted and assessed on a testing the datasets. Evaluation compares the false positives and detected objects for both approaches. Detection of salient region features region of intrigue in whole objects and suppresses the background areas. Visual consideration is considered to include two components: task driven and stimulus driven. Object recognition identifies an object in an image in potential applications that ensures the process to find the surrounding. It contributes the realization of human computer interaction. Object recognition is used for retrieval of information related to the object identified. The appearance of object varies w.r.t illumination, viewpoint or occlusion changes. The variation of object appearances is larger than the appearances between objects in an image. The recognition of object handles variances between or within class. It creates more challenging task when sample size available is small for training.

The object recognition is classified into classification and recognition. The methods in [1], [2] classifies the objects in an image into classes and it is separated roughly into temoccluded region [3]–[5] and feature- based methods [6]–[8]. The latter is used in a portable application [8] [9] [4] to acquire the recognition rate with improved accuracy that recognizes the objects even in lower resolutions.

However, object recognition for HD applications necessities large computations in GOPS but for real-time environment, it requires only a low power consumption. The integration of heterogeneous core makes the energy efficiency possible in recognition an object for object recognition processors [9] [10]. However, communication between the number of



cores becomes difficult when the number of cores increases and it leads to low bandwidth, propagation delay and high-power consumption [11]. The heterogeneity in multi core processes is utilised for high energy utilization, where NoC is adopted widely in core architectures [4] [8].

Why Face Recognition?

Occlusion in facial imagery alludes to the obstacle due to scarfs, glasses, different accessories or hairs. Endeavours on recognition of face has just been accomplished in investigations upto this point like enlightenment field, outward appearance, present estimation and fractional occultents has been the research for successful face recognition under changing condition. We have arranged the writing in wonderment of a similar issue and hence enroll the few of the powerful endeavors that have been made to determine the issue. This techniques like Part Based Methods which includes strategies like PCA, LDA, NMF, LNMF, ICA and alternate varieties of it has demonstrated better recognition rates. Likewise, different strategies like component-based strategy or fractal technique for the most part considers the highlights around the fundamental facial parts is fused in calculation; however, such techniques demonstrate assorted arrangement of examination results. Occultents is separated into two classes i.e., characteristic occultents and manufactured occultents. Regular occultents alludes to the non-purposeful barricade of perspective of the products while the manufactured occultents to the eagerly presenting of fake obstruction between the perspectives. Fractional occultents has now and then extremely influence the picture handling of biometrics as it has a tendency to upset the distinguishing proof of the specific picture. For instance: in instances of retinal biometrics eyelashes blocked district the iris, hoops may deform the ear cartilage and along these lines actuating the blunder rates in the biometric recognition process [1]. Presently, with regards to the face recognition there are a few frills that may be deliberately or accidentally used to impede the pictures [2].

Presently, there have been huge ascent in reports where the wrong doing committer has presented camouflaged for which it comes up short the computerized security framework to recognize and stop the danger at primer stages itself or even in circumstances where individuals change their hope to trap the face recognition frameworks. This face recognition frameworks bombed totally in this criterion to defeat such troubles [3]. In spite of the fact that the face recognition PC technology has accomplished different accomplishments yet the issues of enlightenment, present estimation, fractional occultents has remain yet settled with precision higher for rates of recognition [4- 9]. Henceforth, there are couple of procedures received in the past examinations to defeat the circumstance [10-15]. Strategies like LDA [17], neural systems [18], PCA [16], and a few varieties conquer this circumstance however none has demonstrated its adequacy for the goals of this restrictive situations like non-straight cases has tasks occultents. There are different techniques by the preferences neural systems [13], KDA [14], FDA [15], GDA [16] which are considered suitable for non-straight cases yet have not developed as a result of its high computational expense of activity for both preparing and dataset testing.

The face recognition depends essentially on two capabilities i.e., worldwide and neighborhood highlights [17,18]. It highlights are effectively affected by occultents or clamors of specific kinds and in this manner corrupting the heartiness of such strategies. There are considers which reports the high viability of neighborhood highlights when utilized for face recognition [15-19]; and yet these highlights are for the most part inclined to the effects by incomplete occultents. Be that as it may, whenever utilized comprehensibly such techniques demonstrate surprising outcomes. Endeavors made by Martinez which utilizes his vigorous recognition framework by consolidating the neighborhood highlights in light of similitudes [20]; or of the endeavors on SVM ordered the nearby highlights in to bunch that arrangements with the highlights of facial parts like eyes, ear, and nose to set up a geometrical translation of relating highlight sets. [21-23]. Presently, there are different classes of face recognition which centers around a greater amount of the all- encompassing highlights of facial areas i.e., depending on worldwide highlights. Subsequently, in the accompanying area we introduce the survey of acknowledgeable strategies which refers to better outcomes in execution test particularly for somewhat impede pictures.

Steps in Face Recognition

Face recognition preparing method in applications with impacts of impediment, enlightenment, and condition for imaging on the real time images. The mixture of discovery and recognition of faces in investigates. Application for identification is used for discovering the position of the countenances. Recognition algorithm groups the given images with properties used generally as a part in vision applications. These images have some properties like same determination, including same comparable eye and facial component segments. These images are referred as standard. Recognition uses standard images and locationalgorithms identify confronts and concentrate face images which incorporate eyes, eyebrows, nose, and mouth. That makes the algorithm more confounded than single recognition or recognition algorithm. Second step is discovery of face from the image. As a third step, recognition of face that takes the images of face from identification. Last step is individual way of life as a recognition part consequence. An outline is given in Figure Securing pictures to PC from camera and computational medium (environment) by means of edge grabber is the initial phase in face recognition framework applications.

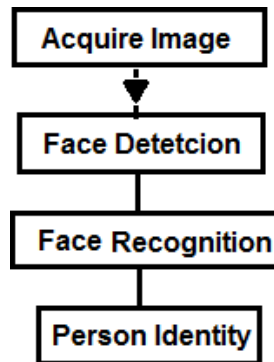


Figure 1.1: Steps of Face Recognition System Applications.

The information picture, as computerized information, is sent to face identification algorithm some piece of a product for extricating every face in the picture. Numerous systems are accessible for identifying faces in the pictures in the writing. The accessible systems could be arranged into two fundamental gatherings as; information based and appearance-based strategies. Quickly, information-based routines are gotten from human learning for elements that grimaces. Appearance-based techniques are gotten from preparing and/or learning strategies to discover countenances. The insights about the strategies will be outlined in the following section. After appearances are identified, the confronts ought to be perceived to distinguish the persons in the face pictures. In the writing, a large portion of the routines utilized pictures from an accessible face library, which is made of standard pictures. After countenances are identified, standard pictures ought to be made with a few techniques. While the standard pictures are made, the countenances could be sent to recognition algorithm. In the writing, systems can be separated into two gatherings as 2D and 3D based strategies. In 2D strategies, 2D pictures are utilized as info and some learning/preparing techniques are utilized to arrange the distinguishing proof of individuals. In 3D strategies, the three dimensional information of face are utilized as a data for recognition. Distinctive methodologies are utilized for recognition, utilizing comparing point measure, normal half face, and 3D geometric measure. Insights about the systems will be clarified in the following area.

II. DESCRIPTION OF SURVEY SECTIONS

Vision is a rich source of information for robots deployed in the real world. Although considerable research as performed in the area of robot vision, existing algorithms are still inadequate for accurate scene understanding and object recognition in the real world. Robots frequently find it difficult to recognize objects and successfully complete the assigned task in challenging scenarios with a significant amount of clutter. Improving the ability of robots to fully utilize the information encoded in visual inputs is hence crucial for the widespread deployment of robots. Contextual cues are very important for object recognition in humans. Recent research in computer vision and robot vision has hence focused on using context to improve object recognition on robots. Contextual cues can enable robots to use the known information about some object in the domain to locate other related objects more effectively. This research project describes the context of objects in images using color histograms and local image gradient (SIFT) features of the neighboring image segments. Once the robot has learned the typical context of desired objects, these objects can be recognized effectively in test images by comparing the context of candidate image regions with the learned context, using the nearest neighbor algorithm. This approach is evaluated on a set of images captured by a camera mounted on a mobile robot.

Jiann-Der Lee [12] presented optimal feature extraction using neural networks for recognizing a 2D objects using two stages approach: Normalized Rapid Descriptor (NRD) and linear extraction process. The former process determines the unknown 2-D objects coefficients from the region boundaries and the latter process is a learning process that extracts the NRD coefficients. The feedforward neural network is trained using object recognition.

Nikhila and Rawat [13] presents an object recognition and classification method for region detection using different visual cues. It uses bottom-up method for salient object recognition to acquire local contrast map and compactness.

Choeng & Chanchareon [14] presents feature classification and recognition algorithms using a suitable pattern. The correlation-base matching algorithms identifies the cylinder, cone, spheres and boxes.

Harjono and Kusuma [15] present a Best Increasing Subsequence object recognition method. This method finds the feature pairs and increases the similarity score between images.

Liu et. al. [16] studies video datasets for object recognition using Latent Bi-constraint SVM with framework for object recognition. This method uses structured-output SVM for handling the noisy data and further it ensures output decision consistency with reduced time.

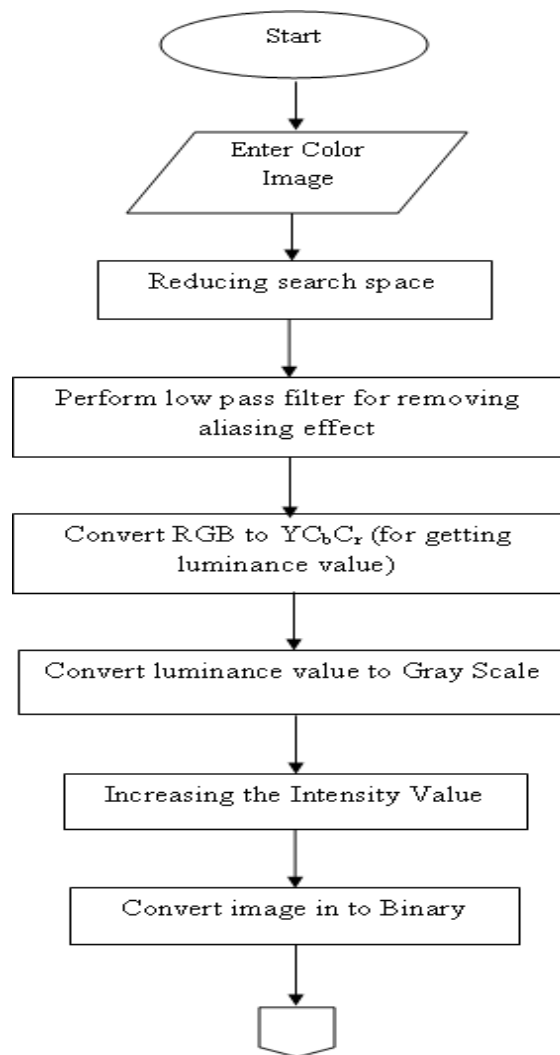


Figure 1.2: Algorithm for Extracting facial features from Occluded region Images.

The general occluded region recognition consists of following two processing phases:

- Facial feature Extraction and
- Face Recognition.

In the first phase of processing, the facial features extracted from the subjected facial image. Numerous works have been done in the area of extracting the occluded region. In this section, we present the methodology for recognizing the faces from the extracted features from occluded region. The presented methodology is depicted in the following figure in a detailed manner; however, it is comprised of the following stages of processing.

Algorithm of Face Recognition from Feed Forward Neural Network

1. Form network according to the specified topology parameters
2. Initialize weights with random values within the specified \pm weight bias value.
3. Load trainer set files (both input image and desired output text)
4. Analyze input image and map all detected symbols into linear arrays
5. Read desired output text from file and converts each face to a binary Unicode value to store separately
6. For each face:
 - a. Calculate the output of the feed forward network
 - b. Compare with the desired output corresponding to the symbol and compute error
 - c. Back propagate error across each link to adjust the weights
7. Move to the next face and repeat step 6 until all faces are visited
8. Compute the average error of all faces
9. Repeat steps 6 and 8 until the specified number of epochs

- a. Is error threshold reached? If so abort iteration.
- b. If not continue iteration.
 - Face segmentation.
 - Generation of training dataset and fuzzy rules
 - Recognizing the unknown faces
 - Face Segmentation

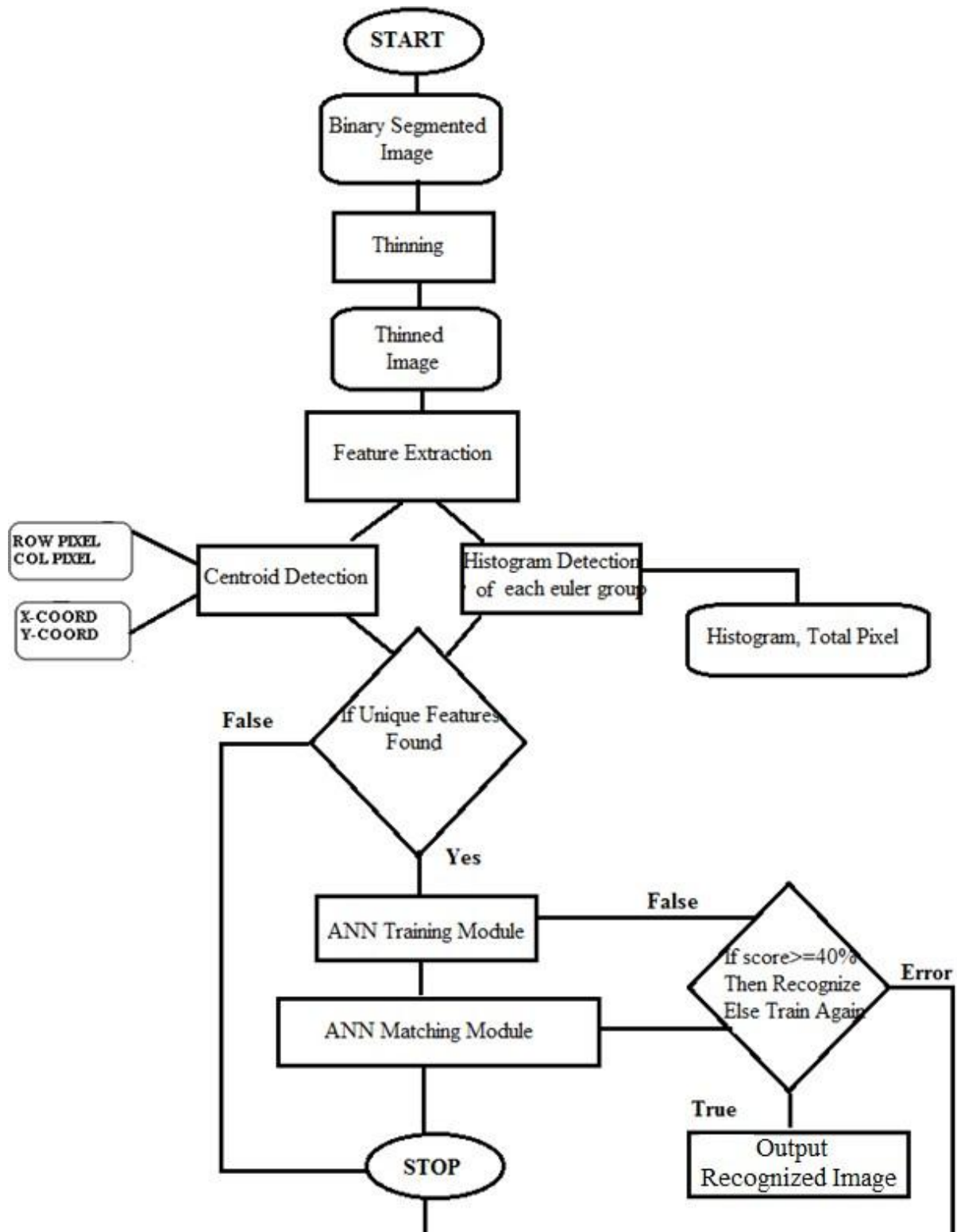


Figure 1.3: Flowchart of ANN face recognition module

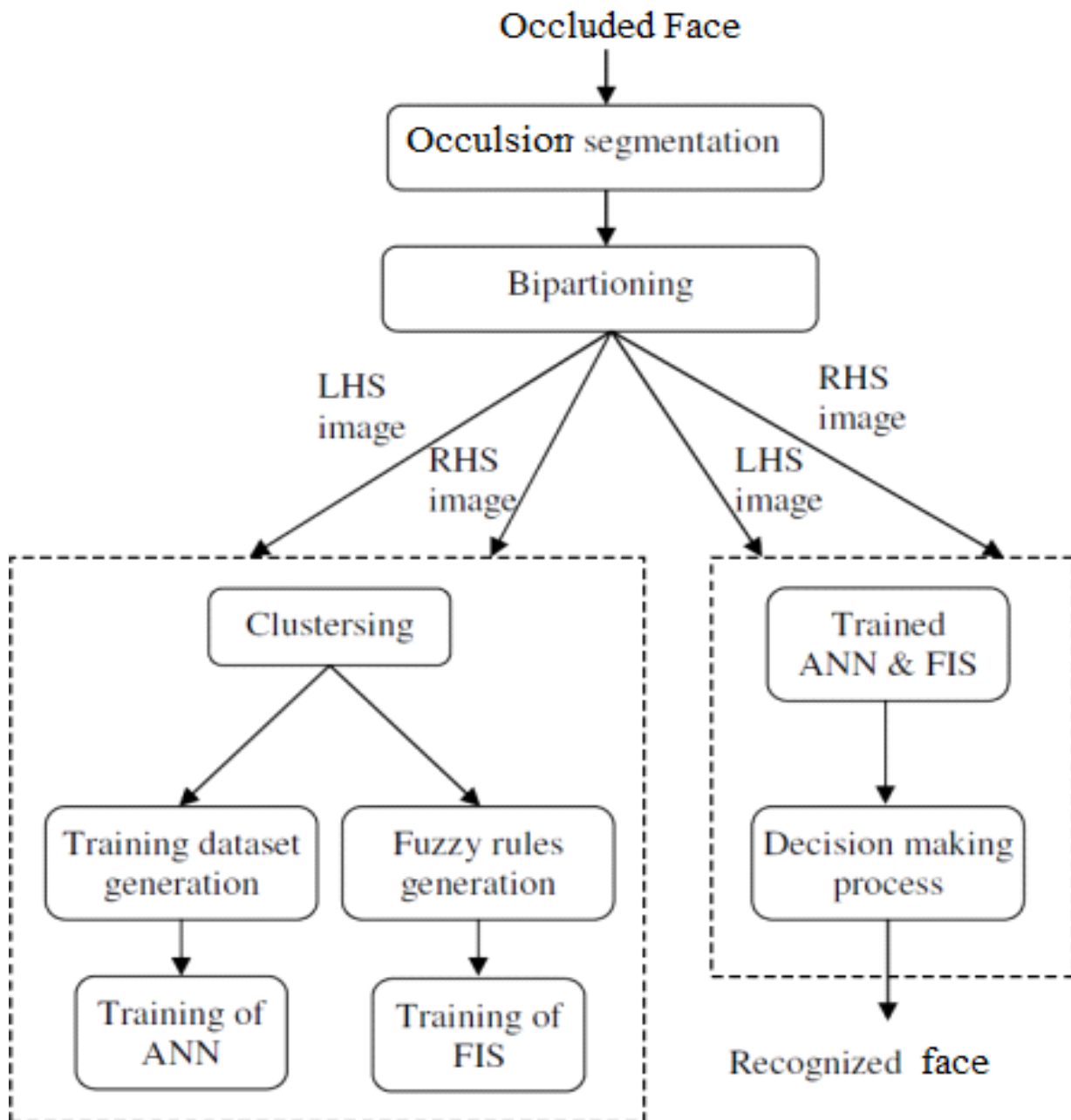


Figure 1.4: The detailed view of the presented methodology in recognizing the occluded region of faces.

The face segmentation is the initial process to be performed in the phase of recognizing the faces from the extracted occluded region. To perform this, the occluded region portion, which is in RGB color space is converted to gray scale image. Then the gray scale image is converted to binary image and then inverted. The binary image is morphologically opened, which is nothing but erosion of the image and then dilation of the eroded image. The every row of faces are extracted from the image are extracted. In every row of faces, the connected components are extracted. The connected components are nothing but the faces that are to be extracted and to be recognized. Thus obtained faces are then subjected to the presented recognition methodology.

Generation of the Training dataset

The occluded region will work in our system as an input. The first step that is followed is generation of the training dataset for the use of neuro fuzzy. To generate the training dataset, we present a bipartitioning technique and a clustering technique. They extract the features from the faces that are used to train the neural network as well as to generate fuzzy rules.

III. ALGORITHM FOR ENCODING FACIAL FEATURES FROM OCCLUDED REGION IMAGES

The idea relies over the idea to map the topological features from the given facial image and thereby coding it in logical sequence with the help of cascaded neural network; since in previous study such attempt has already been quite successful in decoding the face recognition patterns involved with human brain and then training the neural nets to mimic such process [28]. Therefore, the underlying pre-processing steps to reduce the effects posed any poor or uncontrolled lighting environment. The presented pre-processing steps begin with the color transformation of the given facial image in order to normalize the light illumination field. The latter steps dynamically adjust the contrast of the image by breaking down the wavelets in logical association with spatial colorized pixel field. This ensures the optimization of the reflectance field derivative from the pixel values. Hence, making it easy to perform computational operations over optimized values. Such that, the given RGB standard image is to be normalized into *rg* scheme by using the following pre-processing algorithm:

Algorithm: Pre-Processing Algorithm

Input: RGB standard Image

Where, *m* & *n* are the row & column of the given image.

Output: *rg* normalized color scheme

Loop: *for i to m*

Loop: *for j:n* //RGB normalization to *rg* scheme

end

end

Check & Segment:

if() //To adjust contrast

{

//for sampling skin texture

else

//for sampling background light

//illumination field

}

Where, the value *M* represents the model of skin color, which is embarked as low intensity pixels after the pre-processing. & are the mean & covariance of the pixel distribution based on intensities in *rg* color scheme after pre-processing.

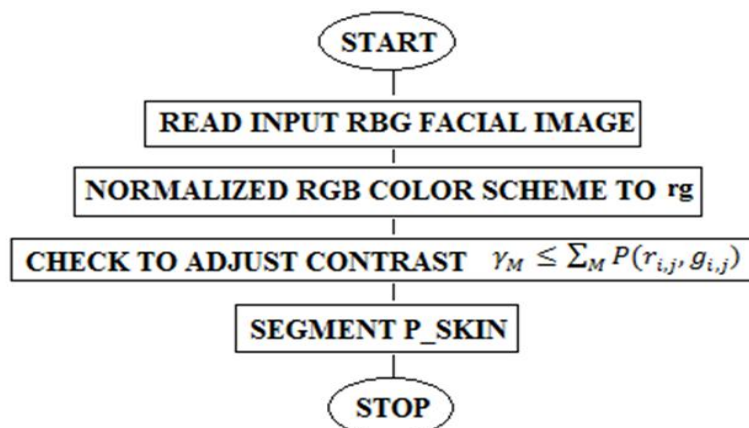


Figure 1.6: Flowchart for Encoding Facial Features

Bipartitioning technique

The Bipartitioning technique splits every face into two subparts. The first part is Right hand side (RHS) and second part is Left hand side (LHS). The subparts generated will belong to other faces also. For this we can consider an example of face B. Face B consists of two subparts. The RHS part of B is similar to that of the RHS part of 3, and also LHS part of B is similar to that of the LHS part of E. So our system will first consider the LHS and then calculate the probability of faces that match with the subpart. And then the same will be followed for RHS. After it face which will be common for both LHS and RHS will be generated on the output. A more detailed exemplary case is depicted in the following figure. Then considering both LHS and RHS the system will make an output which will contain face and will be the ad joint of LHS and RHS. Lather that, after all calculation and considering all the faces the output will be reflected dback.

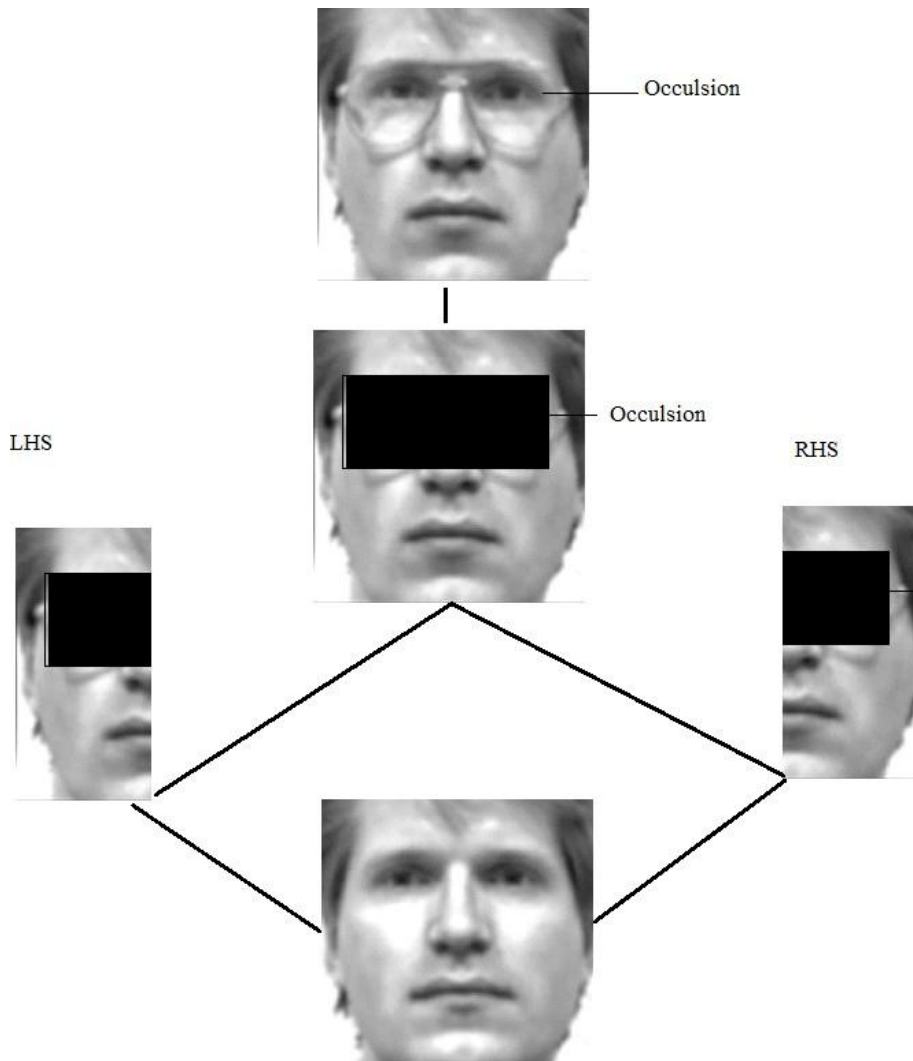


Figure 1.7: An explanatory case of Bipartitioning technique with test samples

Once the faces are bipartitioned, the partitions are subjected to clustering. The clustering technique is described below. Mathematical equation & steps given in section...

Clustering

Prior to clustering technique, the obtained subparts of every face need to be resized because of the fact that while bipartitioning the partitioned image gets vary. The resizing is performed as follows.

All the images that are bipartitioned belongs to same size. For example the size of the matrix of [I] can be 36×36 and for [M] it can be of 36×50 because of the partitioning operation. So, the aforesaid model converts the matrix [I] and [M] to a size of 36×50 .

Let be the matrix obtained after bipartitioning and resizing of a face, where represents the number of clusters and represents the number of elements in a cluster. Let be the mean of the given data, which can be determined as where, is the number of elements in the cluster and is the matrix of the cluster element. The clustering process is detained in the following pseudo code.

```
for every binary matrix  
set first binary matrix as  
for  
    if  
        end if  
    end for  
end for
```

After clustering, we obtain the LHS based face clusters as well as the RHS based facial feature cluster. Based on the clustered elements, the training dataset for neural network and fuzzy rules for fuzzy inference system are generated. Multilayer feed forward neural network is utilized in our methodology. The input layer has neurons i.e. number of matrix elements, the hidden layer has neurons and the output layer has neurons i.e. the number of faces. Back propagation algorithm is used to train the neural network, which is described below.

Step 1: Generate arbitrary weights within the interval [0, 1] and assign it to the hidden layer neurons as well as the output layer neurons. Maintain a unity value weight for all neurons of the input layer.

Step 2: Input the training dataset I to the classifier and determine the BP error as follows

Assign cluster

Convert to non- binary matrix (5)

In Eq. (1), is the target output and is the network output, which can be determined as, are the network outputs.

Step 3: Adjust the weights of all neurons as, where, is the change in weight which can be determined as is the learning rate, usually it ranges from 0.2 to 0.5.

Step 4: Calculate the outputs from the hidden layer: Where X_j is the net input to the i th input unit, O^h_j is the output of the j th hidden layer neuron, is the gain of the activation function and W_{ij} is the weight on the connection from the i th input unit to the j th hidden unit.

(i) Calculate the actual outputs: Where O_k is the actual output for the k th output unit, W_{jk} is the connection weight from the j th hidden unit to k th output unit.

ii) Calculate the error terms for the output units and hidden units: Where T_k is the target output for the k th output unit and $\delta_o k$ is the signal error term for the k th output unit is the signal error term for the j th hidden unit.

iii) Update weights on the output layer η where is learning rate.

iv) Update weights on the Hidden layer

v) As per the following expression, we used to calculate the error: if New else New

Step 4: Repeat the process from step 2, until BP error gets minimized to a least value. Practically, the criterion to be satisfied is.

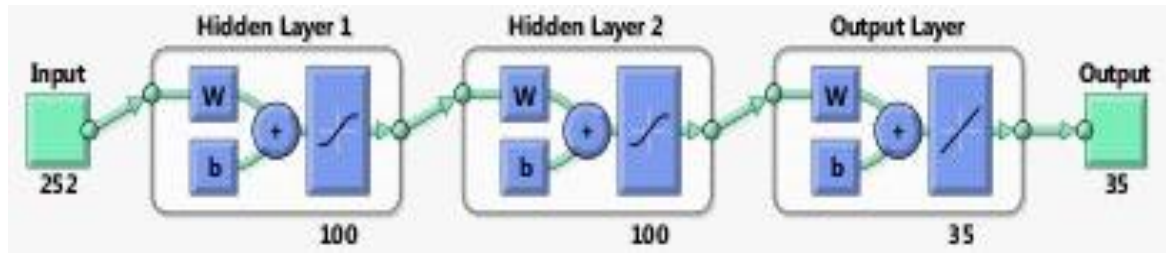


Figure 1.8: The structure of the multilayer feed forward neural network model

After the completion of process, the well- trained network is suitable for face details of LHS or RHS.

Step 5: Stop

The gain parameter value influences the activation function slope for large value of gain. The the function for activation acts as a step function and the sigmoid function is used for estimates a linear function for smaller value of gain. The learning rate accelerates the learning of BP convergence. The LR sets appropriately inorder to dominate the BP performance. The BP converges for a longer instance of time or if LR is small it does not converge. The convergence rate is accelerated by the network however it oscillates the path if the value of LR is high. Results shows that LR, momentum constant and gain of the activation function affectsthe training speed.

IV. RESULTS AND DISCUSSION

The presentd technique has been implemented in MATLAB 7.10 and to process the occluded region, we have used Image Processing Toolbox. For the evaluation of the presented methodology, we have considered many occluded regions as input to our system. We have considered the following occluded region image as an example to the system. For the training of neuro – fuzzy system we have created a database which include 0 to 9 letters and then A to Z faces as input database. Then as stated above, the bipartitioning technique has been done and then the clustering is performed for the considered database. As in this technique the grouping of all the subparts of a face and then all the faces that match with each other are placed in the same cluster. Hence the database has been generated and the neural network and the FIS have been trained well. The training performance of the neural network. Considering all of the faces and their subparts, the technique has recognized the subjected face.



(A)



(B)



(C)

Original Image

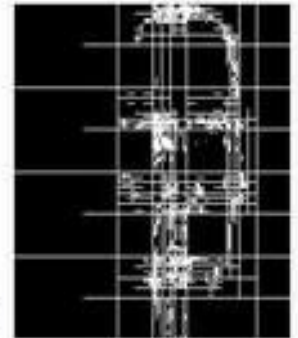
Logically Associative Nodes



All round Lighting



Poor lighting Environment
light source at left



(D)

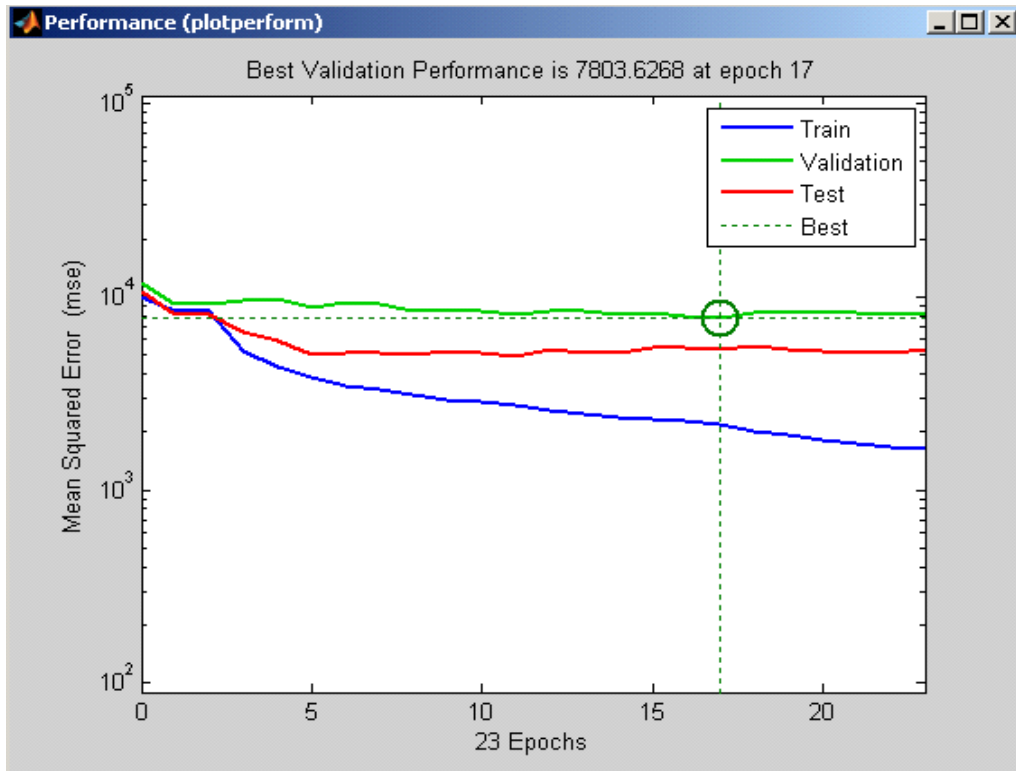


Lighting Source Beneath
The Image

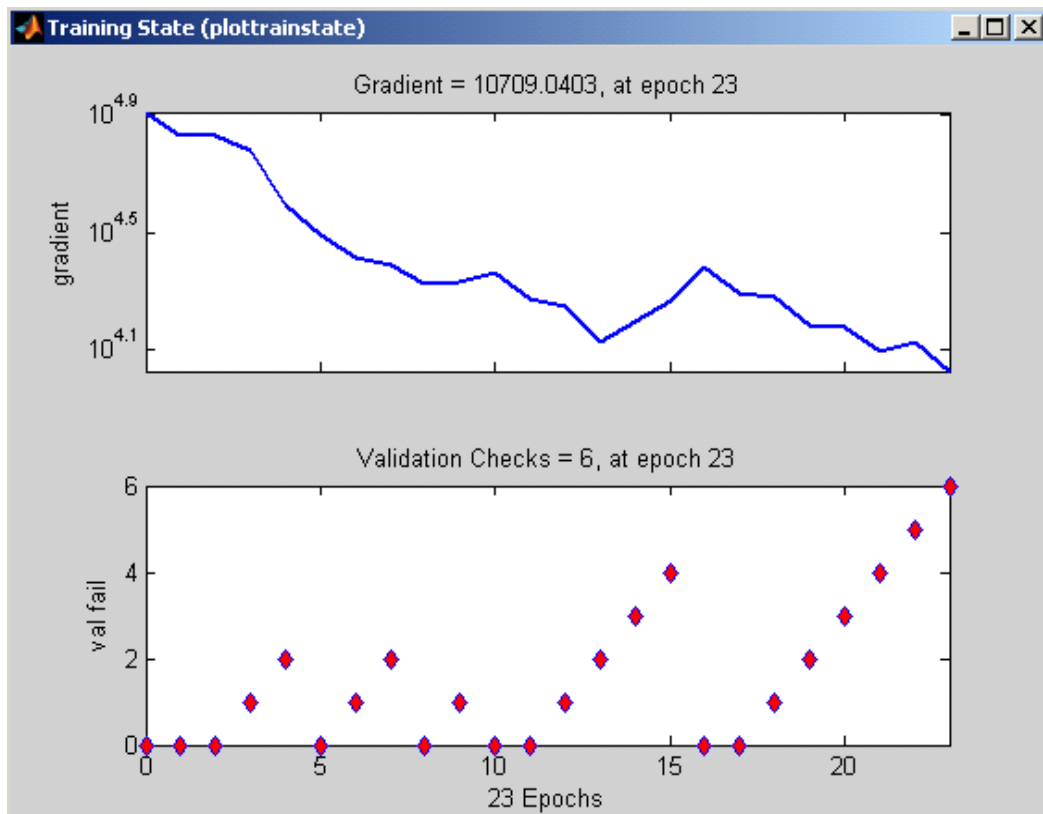


Figure 1.8:(A) The sample results of the PSA algorithm showing the recognized IDs of the test images with occluded region sub spaces during the training period of the cascaded neural network. (B) Sample images from 4 test categories. (C) Images from disguise and sunglasses category which fails in the PSA cascading network in aligning with the trained datasets resulted in inappropriate classification. (D) Results of the formation of associatively cascaded logical blockswith the curvature of the face in the given facial images with different positioning of light source.

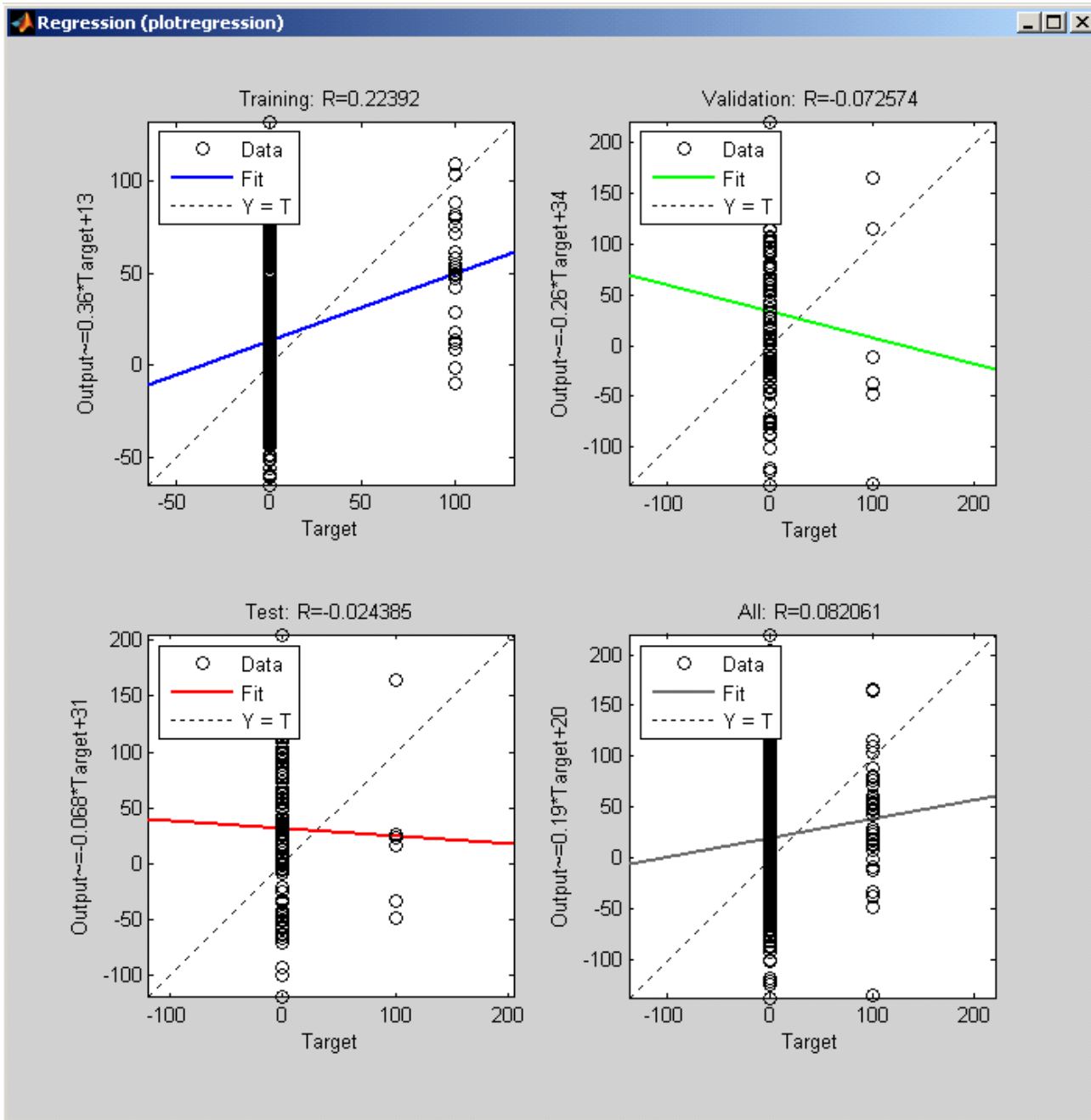
After the face segmentation, the segmented faces have been given to the well-trained Neuro-fuzzy system. The bipartitioning results obtained for the subjected segmented faces.



(a)



(b)



(c)

Figure 1.9: Training performance of the neural network for the given training dataset
(a) Performance (b) Training and (c) Regression

V. CONCLUSION

This paper deals with the problem of face recognition for occluded region images using the presentd PSA algorithm and performance is compared with the other technique that promises to overcome the same problem but failed under uncontrolled conditions, mainly varying light illumination fields. There are attempts which effectively operates in controlled environments to make the system adaptive and robust uncontrolled condition which is way challenging. Existing face recognition methods is classified into three divisions. Among this, the less work is carried out in fractal approaches, however feature approach eliminates loopholes in partial occlusion imagery that attains better results in the near future. However, it has also has least cotribution. The occlusion is found in distorted pixels and part based method is more effective that the field is saturating within region. Face recognition is considered effective with highcomputational complexity that is considered undesirable for a simple real-time application. This is where the presentd algorithm comes

in the play and gives the enhanced results for the uncontrolled environment of varying light illumination and occlusion problem. We hope that the current recognition rate of 86% for occluded region images can find real time applications. A few moderations can also resolve the pose variance problem in future studies.

REFERENCES

- [1]. Benjamin C., David B., Philip M., and Jitendra M., "A Real-Time Computer Vision System for Facial Tracking and Traffic Surveillance," *Transportation Research Part C: Emerging Technologies*, vol. 6, no. 4, pp. 271-288, 1998.
- [2]. Sharif M., Sajjad M., Jawad J., Younas J., and Mudassar R., "Face Recognition for Disguised Variations using Gabor Feature Extraction," *Australian Journal of Basic and Applied Sciences*, vol. 5, no. 6, pp. 1648- 1656, 2011.
- [3]. Yang M., "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition using Kernel Methods," in *Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition, USA*, vol. 14, pp. 215-220, 2002.
- [4]. Chellappa R., Wilson C., and Sirohey S., "Human and Machine Recognition of Faces: A Survey," in *Proceedings of IEEE*, vol. 83, no. 5, pp. 705-740, 1995.
- [5]. Zhao W., Chellappa R., Phillips P., and Rosenfeld A., "Face Recognition: A Literature Survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399-458, 2003.
- [6]. Zisheng L., Jun-ichi L., and Masahide K., "Block- Based Bag of Words for Robust Face
- [7]. Zhengjue Wang, Yinghua Wang, Hongwei Liu and Hao Zhang, "Structured Kernel Dictionary Learning with Correlation Constraint for Object Recognition", DOI 10.1109/TIP.2017.2718187, *IEEE Transactions on Image Processing*, 2017.
- [8]. Yuli Chen, Yide Ma, Dong Hwan Kim, and Sung-Keek Park, "Region-Based Object Recognition by Color Segmentation Using a Simplified PCNN", *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, IEEE, 2014.
- [9]. Xuesong Zhang, Yan Zhuang, Wei Wang, and Witold Pedrycz, "Online Feature Transformation Learning for Cross-Domain Object Category Recognition", *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, 2017.
- [10]. Yu Xiang, Wogun Choi, Yuanqing Lin, and Silvio Savarese, "Subcategory-aware Convolutional Neural Networks for Object Proposals and Detection", *IEEE Winter Conference on Applications of Computer Vision*, 2017.
- [11]. Baudat G. and Anouar F., "Generalized Discriminant Analysis using a Kernel Approach," *Neural Computation*, vol. 12, no. 10, pp. 2385- 2404, 2000.
- [12]. Guo G., Li S., and Chan K., "Face Recognition by Support Vector Machines," in *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble*, pp. 196-201, 2000.
- [13]. Pontil M. and Verri A., "Support Vector Machines for 3D Object Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 6, pp.637- 646, 1998.
- [14]. Christian Szegedy, Alexander Toshev, Dumitru Erhan, "Deep Neural Networks for Object Detection", *The Neural Information Processing Systems Conference*, 2017.
- [15]. Soren Goyal and Paul Benjamin, "Object Recognition Using Deep Neural Networks: A Survey", arXiv 1412.3684v1, 10th dec 2014.
- [16]. Johnny Mauricio Barreno Oñate, Dario José Mendoza Chipantasi, Nancy del Rocío Velasco Erazo, "Tracking Objects using Artificial Neural Networks and Wireless Connection for Robotics", e-ISSN: 2289- 8131 Vol. 9 No. 1-3.
- [17]. Umar Asif, Mohammed Benamoun, and Ferdous A. Sohel, "RGB-D Object Recognition and Grasp Detection Using Hierarchical Cascaded Forests", *IEEE TRANSACTIONS ON ROBOTICS*, IEEE- 2017.
- [18]. Bjo rnBrowatzki, Vadim Tikhonoff, Giorgio Metta, Heinrich H. Bu lthoff, and Christian Wallraven, "Active In-Hand Object Recognition on a Humanoid Robot", *IEEE TRANSACTIONS ON ROBOTICS*, VOL. 30, NO. 5, OCTOBER 2014.
- [19]. Hieu Minh Bui, Margaret Lech, Cheng, Katrina Neville, and Ian S. Burnett, "Object Recognition Using Deep Convolutional Features Transformed by a Recursive Network Structure" DOI 10.1109/ACCESS.2016.2639543, *IEEE Access*.
- [20]. IAN MODEL AND LIOR SHAMIR, "Comparison of Data Set Bias in Object Recognition Benchmarks", 10.1109/ACCESS.2015.2491921, *IEEE*, October 16, 2015.
- [21]. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster R-CNN: towards Real-Time Object Detection with Region Proposal Networks", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10.1109/TPAMI.2016.2577031, *IEEE* 2016.
- [22]. Keke Tang, *Student Member, IEEE*, Peng Song, and Xiaoping Chen, "3D Object Recognition in Cluttered Scenes with Robust Shape Description and Correspondence Selection", DOI 10.1109/ACCESS.2017.2658681, *IEEE Access*-2017.
- [23]. Wai Keung Wong, Zhihui Lai, Yong Xu, *Member, IEEE*, Jiajun Wen, and Chu [21]Po Ho, "Joint Tensor Feature Analysis For Visual Object Recognition", *IEEE TRANSACTIONSON CYBERNETICS*, VOL. 45, NO. 11, NOVEMBER 2015.
- [24]. Jimmy Ba, Volodymyr Mnih, Koray Kavukcuoglu, "Multiple Object Recognition with Visual Attention", [Submitted on 24 Dec 2014 (v1), last revised 23 Apr 2015 (this version, v2)] by Borjain 2020.
- [25]. Carlos R. del-Blanco, Fernando Jaureguizar, and Narciso Garcia, "An Efficient Multiple Object Detection and Tracking Framework for Automatic Counting and Video Surveillance Applications", April 2020, DOI: 10.1186/s13673-020-00219-9.
- [26]. Yanni Zhang, "Object Detection Based on Multiple Information Fusion Net", *Appl. Sci.* 2020, 10, 418; doi:10.3390/app10010418
- [27]. Chinthakindi Baram Murthy, "Investigations of Object Detection in Images/Videos Using Various Deep Learning Techniques and Embedded Platforms—A Comprehensive Review", Published: 8 May 2020. *Appl. Sci.* 2020, 10, 3280; doi:10.3390/app10093280.