

---

## Skin Color Segmentation in RGB Color Space by Adaptive Network Based Fuzzy Inference System (ANFIS)

Mohammad Hussein Fakhravari <sup>1</sup>, Marzieh Dadvar <sup>2</sup>

<sup>1</sup>Computer Engineering Graduate Student, Ciências e Filial Research, Islâmica Azad Universidade, Bushehr, Irã

<sup>2</sup>Conferente Department of Computer Engineering, Ciências e Filial Research, Islâmica Azad Universidade, Bushehr, Irã

### Abstract

*Skin color detection is a popular and useful technique because of the wide range of application in both human computer interactions and analyses based on diagnostic. Therefore, providing an appropriate method for pixel-like skin parts can solve many problems. The presented color segmentation algorithm works directly in RGB color space without having to convert the color space. Using Genfis3 function, we formed the Sugeno fuzzy network and clustered the data using fuzzy C-Mean (FCM) clustering rule and for each class and cluster we considered a Rule. In the next step, the output resulting from pseudo-polynomial data mapping is provided as the input to Adaptive Network Based Fuzzy Inference System (ANFIS).*

**Keywords:** Skin Detection, Fuzzy Inference System Structure, Fuzzy C-Mean (FCM) Clustering, Fuzzy Neural Network, Adaptive Network, Morphological Processing.

## 1 Introduction

Fuzzy neural networks and systems are general known estimators that can estimate any given non-linear function with intended accuracy under the condition of the presence of sufficient neurons in the middle layer and fuzzy rules. Recent studies in the field of neural networks and fuzzy systems indicate that the combination of these two methods can be very effective in identifying nonlinear systems.

Fuzzy modeling and detection was analyzed by Takagi-Sugeno and obtained practical applications in the field of control, detection and prediction. Fuzzy neural networks are obtained by combining fuzzy structures with artificial neural networks that are used to detect systems, predict time series and other various cases.

Adaptive Network Based Fuzzy Inference System (ANFIS) structure presented in 1993 is the result of integrating adaptive neural networks and fuzzy logic the parameters of which can be set based on the input and output data to model the systems. The basic principle used in training the adaptive networks is the rule of reducing gradient. Using gradient reduction alone is not desirable due to the slowness and tendency of the local minima to set network parameters. However new hybrid methods that are the result of combining gradient reduction and other similar methods lead to increased speed of least square error (LSE).

There is a high level of uncertainty to achieve the desired automatic segmentation in each image segmentation layout. This fact also can be extended into face detection especially skin color segmentation. Therefore, fuzzy theory is a good way to achieve the basic detection because fuzzy model provides a mechanism that presents image uncertainty. According to the fuzzy rules extracted from different color spaces used in training phase, each pixel can be divided into the skin or non-skin pixels. One of the features of pixels-based color is that it required no space ground, therefore fixed and fast orientation and size for processing, detection and tracking applications such as human body parts

detection, face detection and detection of naked people all benefit from skin detection. Also, skin color detection helps to block the abusive image or video content. In addition to its use in computer-related technologies, skin color plays an important role in human and human relationships.

## 2 Basic Concept

Dealing with black and white colors does not provide any satisfaction with image. The degree of membership (suggested by Professor Zadeh in 1965) is a new way to solve the problems and issues to deal with vague fuzzy data or in a fuzzy form through which the fuzzy systems are allowed to control certain random levels without causing damage and endangering performance the system. Fuzzy neural system is one of the most successful and most tangible aspects of this effort. Linking and fuzzy neural system integration is done in two ways: an equipped neural network with the ability to handle and control fuzzy information (which is called fuzzy neural network) and a fuzzy system augmented by a neural network to promote some of its features such as flexibility, speed and the ability to adapt known as neural-fuzzy system NFS or ANFIS. An adaptive neural-fuzzy system is designed to realize the fuzzy reasoning processes where network connections are consistent with the parameters of fuzzy reasoning.

Color Image Segmentation using fuzzy clustering is one of the methods of segmentation based on pixels, where the fuzzy system determines each pixel belongs to which category. Therefore, the purpose is to create a fuzzy system that can classify more colors. To do so, an expert person is required that sets the rules and membership functions according to the training data which is very time-consuming and cumbersome and the final rules might not be the best ones. Therefore an automatic method is required to create fuzzy rules and membership functions according to the training data. Many techniques have been developed for this purpose. In this article ANFIS method has been

devised to produce fuzzy rules and membership functions automatically.

The main advantages that provide explicit meaning for skin models for example : their simple decision making rules and integrating them with the main advantages of nonparametric skin models that is said they have lower decision making time for training and clustering. In order to perform this method a fuzzy model for detecting skin in color images is presented.

### 2.1 Tsk Fuzzy Inference System

Fuzzy inference systems are formed from a series of if-then fuzzy rules. An example of TSK model is as follows:

$R_j$  : if  $x_1$  is  $A_{1j}$  and  $x_2$  is  $A_{2j}$  and ... and  $x_n$  is  $A_{nj}$

Then  $y = g_j(x_1, x_2, \dots, x_n)$  ( $j=1, 2, \dots, R$ )

Whre: N: the number of input variables, R: the number of fuxxr rules.  $A_{ij}$ : the fuzzy set corresponding to the input variable  $i$  for fuzzy rule  $j$  and  $g_j$  is a fixef function of  $x_i$  that has a simple linear form:  $g_j(x_1, x_2, \dots, x_n) = q_0 + q_1x_1 + \dots + q_nx_n$ . The final output of the above fuzzy system is expressed as follows:

$$y = \frac{\sum_{j=1}^R g_j(.) T_{i=1}^{m_j} \mu_{ij}(x_i)}{\sum_{j=1}^R T_{i=1}^{m_j} \mu_{ij}(x_i)}$$

Where:  $\mu_{ij}$ : the membership function for fuzzy set  $A_{ij}$ ,  $m_j$  ( $1 \leq m_j \leq n$ ): the number of input variables in the in the initial part of fuzzy rules and T is a fuzzy T - Norm operator. TSK fuzzy system is a single fuzzy system. Partitioning the input space using clustering, networking, etc. is essential for the better efficiency of fuzzy system. Also some adaptable methods are presented in order to estimate the form and number of membership functions in the initial part and free parameters in the successor part of system.

### 2.2 Color Space

RGB color space is mostly used in digital images. This color code is composed of three primary colors: red (R), Green (G) and blue (B). RGB space is usually observable as a 3D cube in

which R, G and B are three perpendicular visible axes. One of the main advantages of this space of RGB is its simplicity. In calculations, the value of a member is stored the range of 0 to 255.

Color space is a provider of analytical mathematical models of color. All color spaces can be obtained by the RGB information stored by devices like cameras and scanners. A wide range of color has been applied in modeling skin color. In order to make a skin color system, two main steps are required: (1) to show the image pixels in a suitable color space, (2) to select an appropriate classification. In this paper, RGB color space is widely used by computer vision goals because of its simplicity. Red, green and blue are three primary additive colors presented by 3D Cartesian coordinate system (Figure 1). Cube diameter with equal amounts of each primary component represents gray levels. With this color space, there is no need to convert RGB color value of the input image into other coordinates. It is clearly observed that the skin color RGB is distributed in linear model [20, 21]. A sample of 3D plan for skin and skin pixels in this space is presented in figure 2.

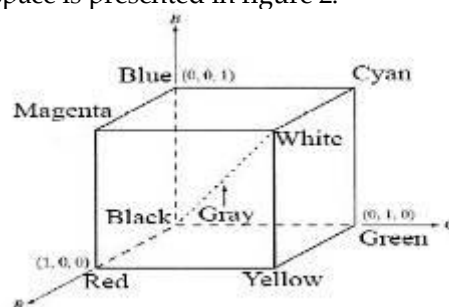


Figure 1: RGB color cube

#### 2.2.1 Classification under Skin Control

When RGB color space is chosen as skin feature, it is important to segment skin pixels. Image classification is a useful area of image processing and machine learning, which is trying to tag a picture with proper ID. Training and testing stage: classification system is divided into two main sections. Training set is selected manually and used to represent the set of the class in a common image. Classification and analysis of training sets for that specific class is based on specific features of training set.

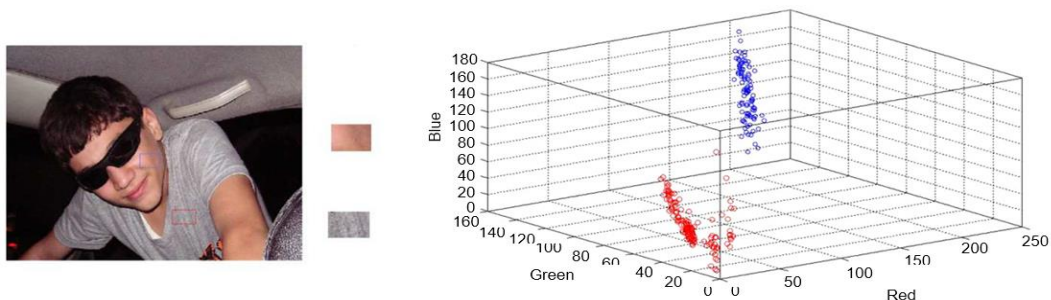


Figure 2: Analysis of color in RGB color space: 3 D skin design (blue), 3D - Non-skin design (red). Skin color is distributed almost linearly in RGB color space

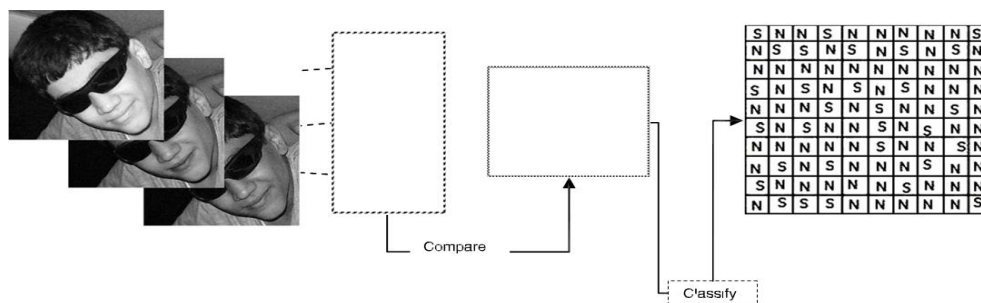


Figure 3: Other images of the intended database present the similarities between the classes in the selected image. After training stage, feature vector of an image works as the input of the test; this stage is known as the testing phase. Classification determines training model and works based on its own specific classification rules as a class which belongs to the vector ability. Figure 3 shows the classification process.

### 2.3 Anfis Structure

Neural networks and fuzzy systems are estimators independent from the model and represent similar features in dealing with uncertainty and noise. Therefore it is possible to convert the fuzzy inference system to a trainable network. The network obtained in this way can use neural network learning methods to train its parameters. In addition, the structure obtained in this way is not left as a black box and based on segmentation advantages of interpretation of fuzzy systems, it will have more advantages and the final result can be expressed in the form of linguistic rules. Such as fuzzy systems, ANFIS structure also consists of two parts. The first part is the initial part and the second one is called successor and the two parts are connected by fuzzy rules in the form of a network. Figure 4 shows the structure of an ANFIS network in five layers, wherein the first layer performs fuzzification, the second layer implements T-Norm action for the fuzzy rules phase, the third layer is applied for normalization, the fourth layer of the successor part creates fuzzy rules and finally the fifth layer calculates the final

output of the system. Feedforward relations for the network structure shown in Figure 4 are as follows:

$$w_{jk} = \mu_{M_j^1}(x_1) \cdot \mu_{M_k^2}(x_2), \quad j, k = 1, 2, \dots, m \quad (2)$$

$$w_{jk} = \frac{w_{jk}}{\sum_{i=1}^m \sum_{i_2=1}^m w_{i_1} w_{i_2}}, \quad j, k = 1, 2, \dots, m \quad (3)$$

$$f_{jk} = q_{0,jk} + q_{1,jk} \cdot x_1 + q_{2,jk} \cdot x_2 \quad (4)$$

In these two equations  $m$  represents the number of membership functions for each input variable and  $\{q_0, q_1, q_2\}$  are real numbers corresponding to linear weights in successive part of ANFIS system. The final output  $y$  in ANFIS network can be calculated as follows:

$$y = \frac{\sum_{j=1}^m \sum_{k=1}^m f_{jk} \mu_{m_j^1}(x_1) \mu_{m_k^2}(x_2)}{\sum_{j=1}^m \sum_{k=1}^m \mu_{m_j^1}(x_1) \mu_{m_k^2}(x_2)} = \sum_{j=1}^m \sum_{k=1}^m f_{jk} w_{jk} \quad (5)$$

In order to model complex nonlinear systems, ANFIS model divides input space into different parts, in other words the input space is divided into many local areas. ANFIS network uses fuzzy membership functions to divide each input dimension. The membership functions overlap

each other, in other words, a single entry will activate two membership functions at the same time. ANFIS network capability depends on to the number of membership functions considered for each input dimension. Usually membership functions used are Gaussian bell-shaped functions with the maximum value of 1 and the minimum value equal to 0 that can be defined as follows:

$$\mu_{M_i}(x) = \exp\left(-\left(\frac{x - x_i}{\sigma_i}\right)^2\right) \quad (6)$$

Where:  $\{x_i, \sigma_i\}$  are parameters of membership functions that affect its shape.

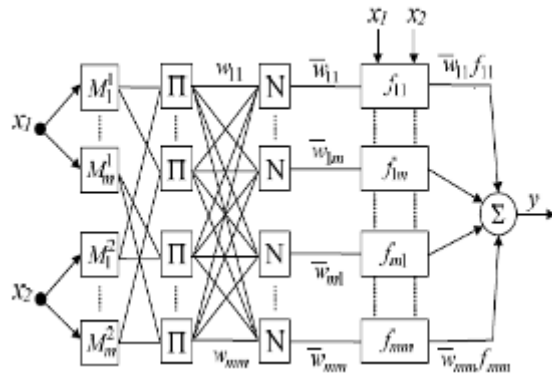


Figure 4: ANFIS structure with two inputs

### 3 Data Mapping

In order to carry out face image segmentation operations based on skin color in RGB color space. First among the 10 selected image we select 10-12 pixels related to skin and equal number of pixels related to non-skin parts in each image and put them in an  $N * 3$  vector orderly i.e. first in the vectors related to the face and then in vectors related to other parts,  $N$  is the total number of selected pixels and in this project it equal 216. It should be noted that the input vector is divided by 255 to be in the range  $[0, 1]$ . (Normalized)

$$\text{TrainInputs} = \begin{bmatrix} R_1 & R_2 & R_3 & R_m & R_n \\ G_1 & G_2 & G_3 & G_m & G_n \\ B_1 & B_2 & B_3 & B_m & B_n \end{bmatrix}^T$$

*skin skin skin nonskin nonskin*

The output vector is as follows:

$$\text{TrainTargets} = [1 \quad 1 \quad 1 \quad \dots \quad 0 \quad 0]^T$$

*skin skin skin nonskin nonskin*

Using function Genfis3, we have formed fuzzy network of Sugeno and clustered the data using the fuzzy C-Mean clustering rule and considered a rule for each cluster (clusterings and creating Sugeno network and rule were done by Genfis3 function). C-Means can put the data in a generalized space in a desired number of clusters so that, unlike the K-Means, each data can be a member in any number of clusters. It is clear that the membership of data in each cluster can be different.

### 4 The Structure of the Network

The created network was given to function's input and MSE error was minimized in iterations of ANFIS training algorithm. Then the network error was calculated for the training vectors and then was drawn. After the completion of training, the selected images were converted into input vector and then applied in the network input and if the output was less than the threshold 5 or 128.256, the pixels related to other parts (other than face) in case of being above the threshold is related to the face. At the end output binary image of the network underwent closing and opening morphological operations with structural disk member with a radius of 4 pixels that led to the disappearance of small and undesirable objects and the final image was selected as the mask on the original image cut.

Table 1: Parameters Network

Inference Method
<b>type:</b> 'sugeno'
<b>andMethod:</b> 'prod'
<b>orMethod:</b> 'probor'
<b>defuzzMethod:</b> 'wtaver'
<b>impMethod:</b> 'prod'
<b>input:</b> [1x3 struct]
<b>output:</b> [1x1 struct]
<b>rule:</b> [1x2 struct]

### 4.1 Fuzzy c mean Clustering Algorithm (FCM)

Similar to Classical C mean algorithm in this algorithm, the number of clusters (c) has

already been determined. The objective function defined for this algorithm is as follows:

$$J = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m d_{ik}^2 = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m \|x_k - v_i\|^2$$

In the above formula  $m$  is a real number greater than 1 the value of which is selected to be 2 in most cases.  $x_k$  is the sample  $k$  and  $v_i$  represents cluster  $i$ .  $u_{ik}$  presents the degree of membership of sample  $i$  in cluster  $k$ .

$\| * \|$  indicates the similarity (distance) of the sample with (from) the center of cluster that any given function presenting the similarity between the sample and cluster center can be used.

$$\sum_{i=1}^c u_{ik} = 1, \forall k = 1, \dots, n$$

From  $u_{ik}$  it is possible to define a  $U$  matrix that has  $c$  rows and  $n$  columns the components of which can take any value between 0 and 1. If all the components of the matrix  $U$  are 0 or 1 the algorithm will be similar to classic  $c$  mean. Although  $U$  matrix components can take any value between 0 and 1 but the total components in each column must be equal 1, then:

The meaning of this condition is that the total membership of each sample (divided) by  $c$

cluster must equal 1. Using the above condition and minimizing the objective function:

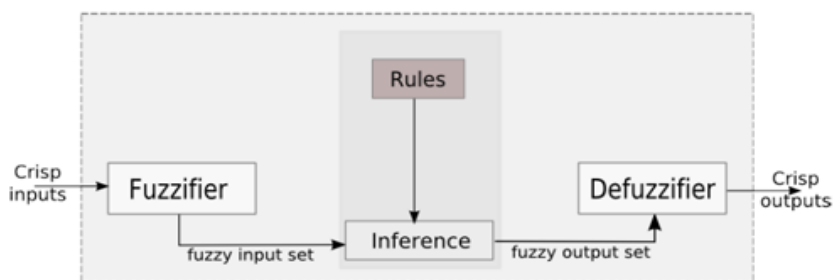
$$u_{ik} = \frac{1}{\sum_{j=1}^c \left( \frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)}} \quad v_i =$$

Levels of algorithm:

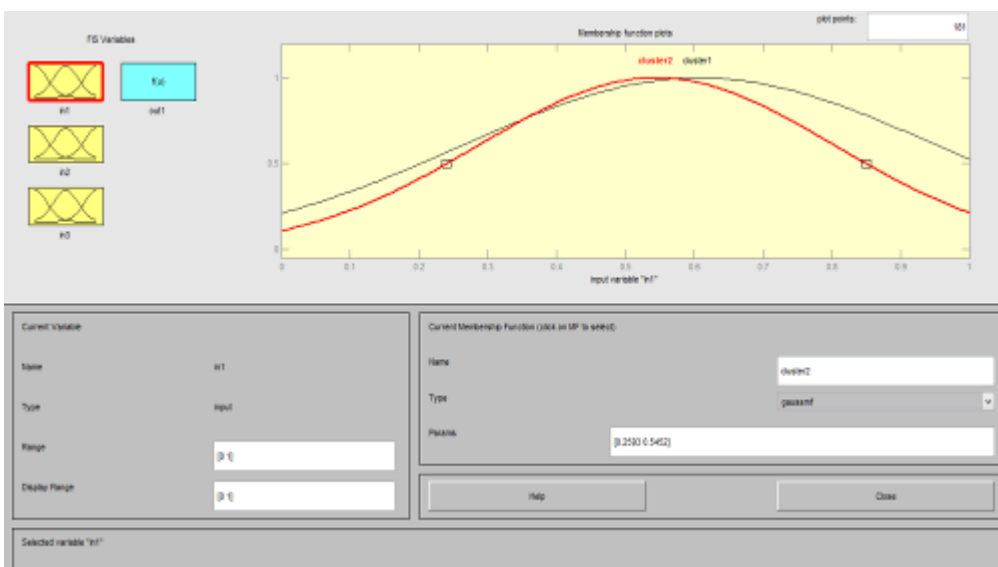
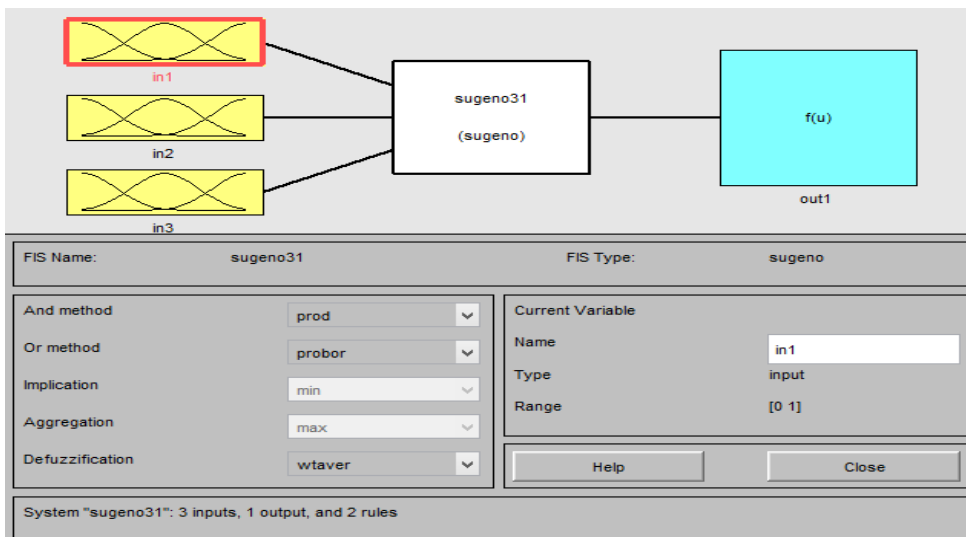
1. Initialization for  $c, m$  and  $U^0$ . Initial clusters are guesses.
2. Cluster centers are calculated ( $v_i$  calculation)
3. Membership matrix calculation by clusters calculated in 2
4. If  $\|U^{i+1}-U^i\| \leq \epsilon$ , the algorithm ends, otherwise go to step 2.

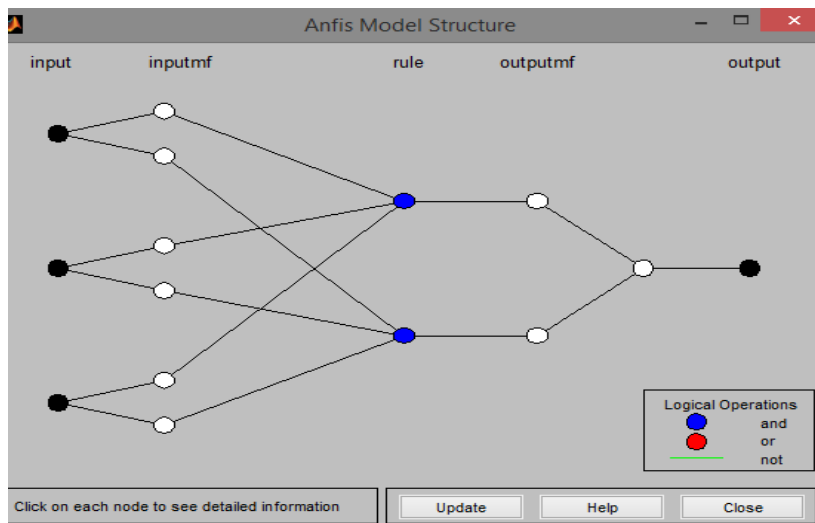
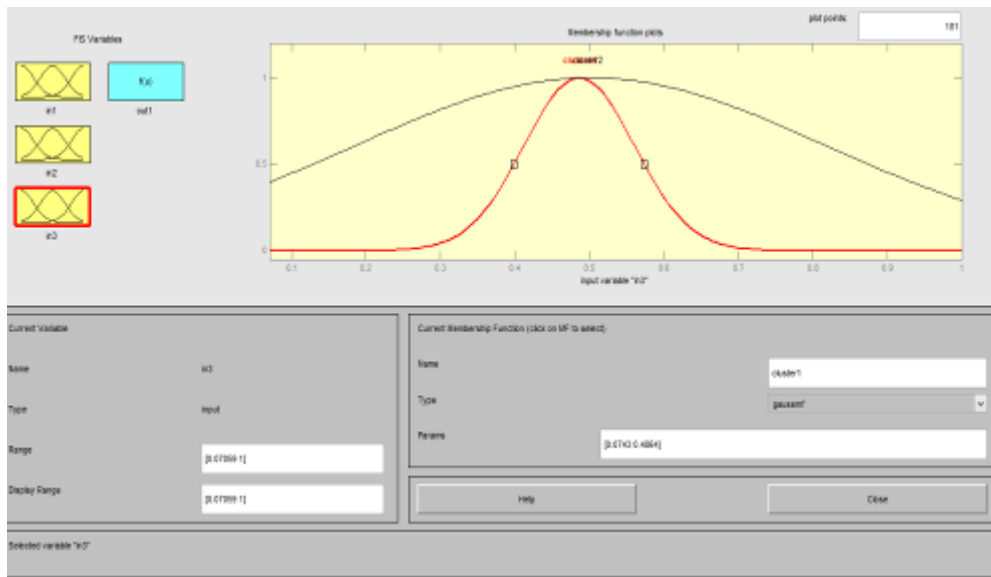
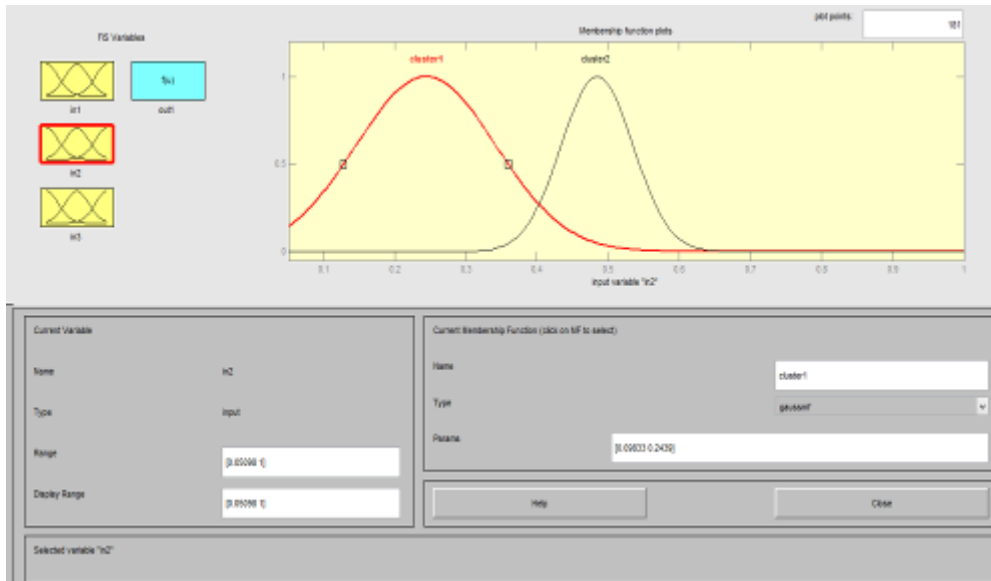
### 4.2 Fuzzy Logic of a System With Certain Parameters

The answer is to convert the system inputs from certain state into fuzzy variables. This is called Fuzzification. Then it is possible to implement fuzzy rules on the fuzzified sets the output of which will be fuzzy again. Now these fuzzy outputs should be converted into certain outputs which is called defuzzification.

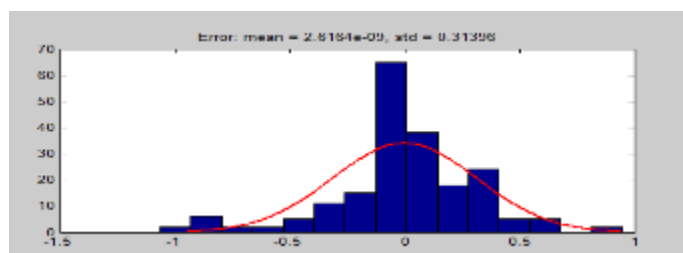
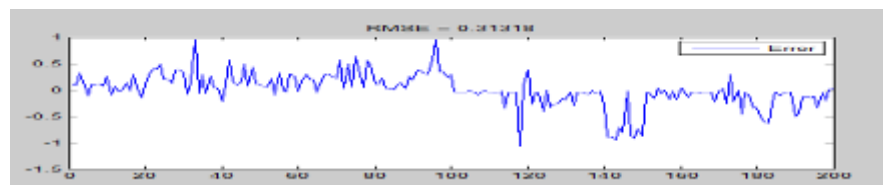
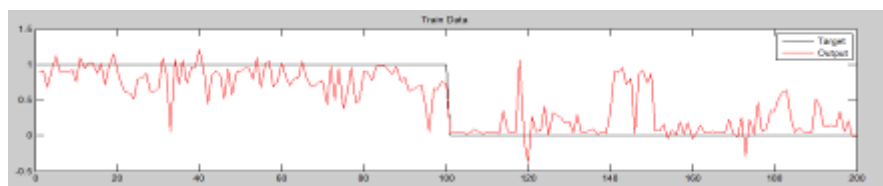
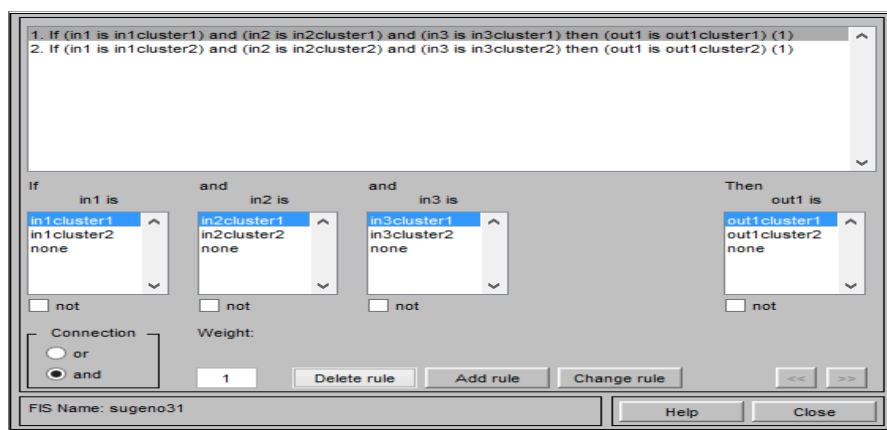
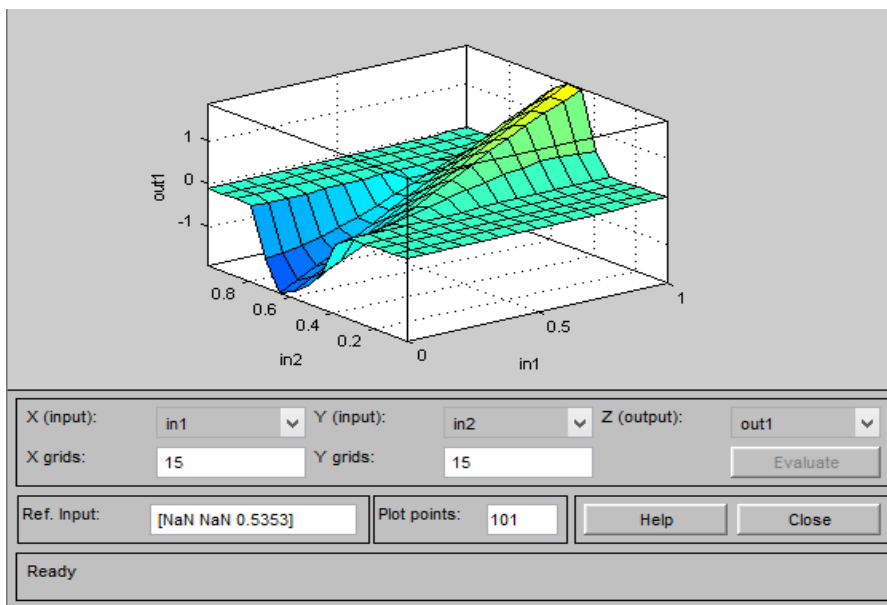


### 4.3 Network outputs



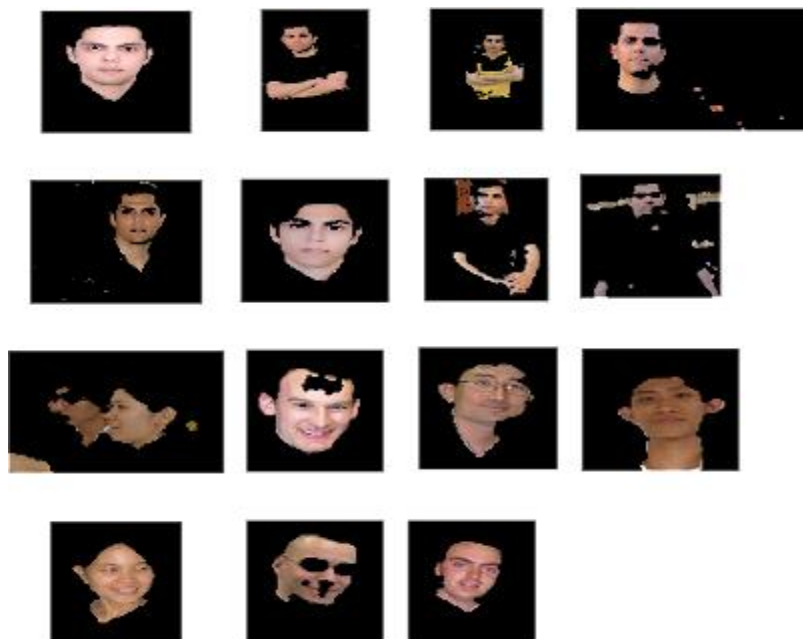






## 5 Simulation and Results





One of the methods used to detect a face is to use skin detector designer that is proved to have great impact in many cases. Here the fuzzy skin color detector uses the proposed RGB color space. In this study, a segmentation into two zones (skin or non-skin as separately) is used for skin detection; segmentation is performed based on pixel to classify skin and non-skin pixels

independent of their neighbors. In this application, the search for objects of interest, such as the face or hands, can be reduced by identifying areas of the skin. Fore this purpose, skin segmentation is very effective because it usually contains limited amount of calculations that can be performed regardless of the gesture.

<b>CDR</b>	$\frac{\text{The number of correct pixels}}{\text{Total pixels in the test data set}}$
<b>FAR</b>	$\frac{\text{The number of non-skin pixels classified as skin pixels}}{\text{Total pixels in the test data set}}$
<b>FRR</b>	$\frac{\text{The number of skin pixels classified as skin pixels}}{\text{Total pixels in the test data set}}$

Table. 2. Results

82.6667	8.1121	9.2213	CDR = 77.39 FRR = 1.47 FAR = 9.36
97.8402	0.0201	2.1397	
93.1473	0.2907	6.5620	
97.8216	0.0611	2.1173	
63.9541	2.3684	33.6774	
91.5353	0.0620	8.4027	
97.8348	0.2025	1.9627	
90.9704	4.1754	4.8542	
71.1081	0.0496	28.8422	
94.7000	0.5697	4.7303	
89.3795	0.0215	10.5990	
87.9285	7.1850	4.8865	
76.0833	0.1500	23.7667	
86.7099	1.2543	12.0359	
93.4255	0.4131	6.1614	

## 6 References

- Navid Razmjooya, B. Somayeh Mousavib, F. Soleymani, A hybrid neural network Imperialist Competitive Algorithm for skin color segmentation, *Mathematical and Computer Modelling* 57 (2013) 848 - 856
- Mohammad Saber Iraj, Azam Tosinia, Skin Color Segmentation in YCBCR Color Space with Adaptive Fuzzy Neural Network (Anfis)
- Arash Sharifi, Mehdi Aliari, Mohamad Teshnelab,(2010) ,Introducing fuzzy system quasi polynomial Takagi-Sugeno- Kang System Identification and classification model
- Akshay Bhatia, Smriti Srivastava, Ankit Agarwal, Face Detection using Fuzzy Logic and Skin Color Segmentation in Images, *Third International Conference on Emerging Trends in Engineering and Technology*
- Rihane Sepehr, Hassan Moradi, Gonche Mashayekhi, Lale Khardar, Atie Bamdadyan, Evaluate and compare different methods of fuzzy clustering technique Based on segregation FCM Standard fuzzy clustering
- Masoud Yaghini, Maryam Ranjpour, Farid Yousefi, A Survey of Fuzzy Clustering Algorithms