Face Recognition Using Hyper Sausage Neuron Networks

Shaffan MADANNY[1], SAMSURYADI[2*], and Novi YUSLIANI[3]

[1]*Department of Informatics, Faculty of Computer Science, Universitas Sriwijaya, Indonesia*
[2]*Master Program of Computer Science, Faculty of Computer Science, Universitas Sriwijaya, Indonesia*
[3]*Department of Informatics, Faculty of Computer Science, Universitas Sriwijaya, Indonesia*
*Corresponding author: syamsuryadi@unsri.ac.id

Abstract

Biometric authentication systems are generally utilized in different manners. One of the most generally utilized biometric systems is the face. The face is one path for every person to know one another and has their very own attributes or is extraordinary to every individual. The face recognition system is divided into three phase's, namely pre-processing, feature extraction, and recognition. This paper combines Linear Discriminant Analysis (LDA) for feature extraction and Hyper Sausage Neuron Networks (HSNN) for the face recognition system. LDA is used to overcome the resistant to changing image conditions, and HSNN can overcome high false acceptance problems on Traditional Pattern Recognition (TPR). The research used the face94 database with LDA-HSSN could enhance the accuracy of face recognition from the previous methods such as SVM with RBF kernel and KNN respectively 2.03% and 35.54%.

Keywords: *face recognition, Hyper Sausage Neuron Networks, Histogram Equalization, Linear Discriminant Analysis*

Introduction

Biometric authentication systems are generally utilized in different manners. One of the most generally utilized biometric systems is the face. The face is one path for every person to know one another and has their very own attributes or is extraordinary to every individual [1]. The face recognition system is divided into three phase's, namely pre-processing, feature extraction, and classification [1].

The pre-processing stage has the role of removing noise and blurred images. Several studies have employed pre-processing methods such as: Pose Correction AAM Model [2], Edge Detection [3], Discrete Wavelet Transform [4,5], and Histogram Equalization (HE) [6]. HE is used to overcome low or high lighting images.

Several related studies were to extract of features in face images such as: Local Binary Pattern (LBP) [2], Scale Invariant Feature Transform (SIFT) [2], Principle Component Analysis (PCA) [3,4,7], Contourlet Transform [8], and Fuzzy LBP [9]. The PCA method is not good at changing image conditions such as lighting, expression, and pose on face images [4]. This study proposed Linear Discriminant Analysis (LDA) for sample images with changing conditions [10] to overcome the problem of feature extraction.

In addition, previous studies on face recognition such as: SVM Multi-Class [2], Back-Propagation Neural Networks (BPNN) [3], Support Vector Machine (SVM) [4,8,9], li Norm Nearest Neighbor Convex Hull (liNNCH) Method [7], K-Nearest Neighbor (KNN) Classifier [9], and Hidden Markov

Model (HMM) [11]. All Face recognition methods above are Traditional Pattern Recognition (TPR) methods. The traditional patterns based on the idea of "division" so that it affects two things, namely not being able to solve high false acceptance problems in untrained samples and retraining the entire system or re-distributing all sample types when new types of face samples are added [12]. According to these problems, the proposed of biomimetic patterns were to overcome problems in TPR [13]. Face recognition studies on Hyper Sausage Neuron Network (HSNN) [12,14,15] and Three Weight Neuron Networks (TWNN) [16].

The research proposed the combination of Linear Discriminant Analysis (LDA) and Hyper Sausage Neuron Networks (HSNN) to overcome the lower accuracy of face recognition. The research used the face94 database with LDA-HSSN could enhance the accuracy of face recognition from the previous methods such as SVM with RBF kernel and KNN respectively 2.03% and 35.54%.

Face Recognition Systems

Face Images Data

This research uses face94 database with JPEG format with 2280 images from 114 persons. The JPEG format on face recognition has better recognition accuracy compared to other formats. This has been proven in the study [5] with the results of the study stating the JPEG format is good for the process of reducing dimensions and increasing recognition

close to 8%. Images data is divided by 1596 for training and 684 for testing. In the learning phase, the training data is divided into 5 sections such as: 280, 560, 840, 1120, and 1596.

Resizing

In this part is to change the size of original face images resolution become 30x30 pixels. The resize phase is to make small size data and speed up the pre-processing process. The resize phase can affect the increase in accuracy in face recognition [17].

Grayscale

Gray image is an image that has brightness information on images rather than color information on images [14]. Therefore, the color image will be converted into an image with a gray value, from the black to white level (0-225). The conversion of color image to gray image can be stated in Equation (1).

$$\text{Gray Value} = (R + G + B) / 3 \tag{1}$$

R is the red gamma correction intensity of the color image, G is the green gamma correction value of the color image, and B is the blue intensity gamma correction value of the color image. The gamma correction parameter for this study uses the values $A = 15$ and $\gamma = 1 / 2.1$.

Histogram Equalization

In general, face images have some weaknesses in lighting, especially images with dim or too bright brightness. Histogram Equalization (HE) is the basic idea for converting histograms of real face images into histograms by dividing good proportions so that the quality of the images obtained becomes better [6]. HE is an algorithm that is used to overcome lighting in the image.
HE works using histograms and uses information from histograms of all parts of the image to transform functions. Based on [18], the description of the HE enlightenment method is explained as follows.
An input image is denoted by X = {X (i, j)}, arranged based on the gray level L with a variable value [0, L-1], denoted by $\{X_0, X_1, \ldots, X_{L-1}\}$ and X (i, j) are representations of the intensity values in the location (i, j) that meet conditions (i, j) $\epsilon \{X_0, X_1, \ldots, X_{L-1}\}$. H histogram is defined as the frequency of appearance of the gray level of an image and described in Equation (2).

$$h(X_k) = n_k, k = 1, 2, \ldots, L - 1 \tag{2}$$

X_k is the gray level k-*th* and n_k is the number of occurrences of the gray level in an image. Next, calculate Probability Density Function (PDF) using Equation (3).

$$P(X_k) = \frac{h(X_k)}{(MxN)}, k = 1, 2, \ldots, L - 1 \tag{3}$$

M x N is a measure of image X. The next step is to calculate the Cumulative Distribution Function (CDF) using Equation (4).

$$C(X_k) = \sum_{j=0}^{k} P(X_j), k = 1, 2, \ldots, L - 1 \tag{4}$$

HE made improvements to the quality of image X by using CDF as a transformation function. The $T(X_k)$ transformation function is defined using Equation (5).

$$T(X_k) = (L - 1)x\, C(X_k), k = 1, 2, \ldots, L - 1 \tag{5}$$

At the end of the process, a new image is obtained Y = {Y(i, j)} using the result of transformation $T(X_k)$. Y images is obtained using Equation (6).

$$Y = f(X) = \{ f(X(i, j)) | \forall\, X(i, j) \epsilon\, X \} \tag{6}$$

Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a dimension reduction method for classification problems. LDA is used to obtain matrix projections with maximized intra-class and inter-class ratios. LDA also considers class membership for dimension reduction. The main idea of the LDA is to separate the class average from the projection direction and achieve small variations in the class average [10]. The characteristics of the LDA are linear combinations of the original sample data.

The basic steps on the LDA are:
Calculate the intra-class distribution matrix, SW with Equation (7).

$$S_w = \sum_{i=1}^{c} (x_i - \mu_{k_i})(x_i - \mu_{k_i})^T \tag{7}$$

Calculate the inter-class distribution matrix, SB with Equation (8).

$$S_B = \sum_{i=1}^{c} n_i (\mu_i - \mu)(\mu_i - \mu)^T \tag{8}$$

Calculate the eigenvector of the projection matrix with Equation (9).

$$W = eig(S_t^{-1}, S_B) \tag{9}$$

Compare the projection matrix from the projection matrix test image from each training image using the same measurement. S_B is an intra-class spread matrix, S_W is an inter-class distribution matrix, $S_T = S_B + S_W$ is the total distribution matrix, notation c is the total number of samples in all image data, x_i is the characteristic vector of a sample, and μ_{k_i} is the image vector class of x_i, μ_i is the average feature vector of class *i*, and this is the number of samples from class *i*.
Intra-class describes variations in the condition of the appearance of images of the same individual due to differences in lighting and Face expressions. Inter-classes

describe variations in the image display conditions due to differences in identity.

Hyper Sausage Neuron Networks

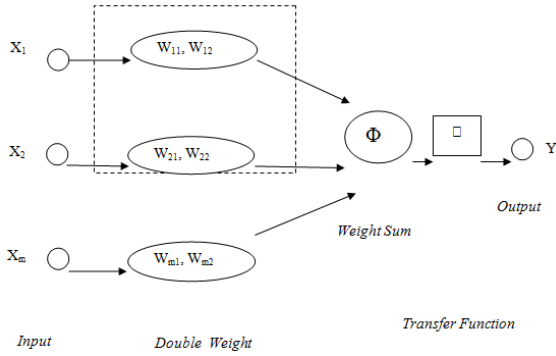Hyper Sausage Neuron Networks (HSNN) architecture was showed in Fig. 2.

Figure 1. Hyper Sausage Neuron Networks (HSNN) Architecture [19].

HSNN forms a characteristic space shaped like sausage as a new class of neural networks. Formula of HSNN can be seen in Equation (10).

$$Y = f[\Phi(W_1, W_2, X) - Th] \qquad (10)$$

$$\Phi(W_1, W_2, X) = \|X - \theta_{(W_1,W_2)}\| \qquad (11)$$

$\theta(W_1, W_2))$ is a sausage shaped area, bordered by a sausage-shaped space covered by W1, and W2.

$$\theta(W_1, W_2) = \{Y|Y = \alpha W_1 + (1 - \alpha)W_2, \alpha \in [0,1]\} \qquad (12)$$

$\Phi(W_1, W_2, X) - \theta$ is the Euclidean distance between the sample X and the sausage containing the neurons from the sausage.

In BPR, the determination of subspace in a particular sample class is based on the sample class itself. The training of the sample uses that sample, no need to use the entire sample. The steps in sample training can be seen as follows [19]:

Step 1: Perform extraction of features from the sample and get N feature vectors;

Step 2: Practice the scoping neurons $P_1, P_2, ..., P_i, ..., P_{N-1}$;

Step 3: Save the parameters of the N-1 neurons and stop the training.

Threshold is needed for the introduction of biomimetic patterns because they are useful for learning and training in face recognition. The threshold value greatly determines the best scoping form for the testing process in each sample. This study uses threshold values based on Mean Absolute Error (MAE) used by [7] in Equations (13).

$$Th = \frac{1}{n}\sum_{i=1}^{n}|x_i - r_i| \qquad (13)$$

n is the number of images, x_i is the image that is processed, and r_i is the reference image or first image. This study uses the second threshold value in the research conducted by Samsuryadi [19]. The second threshold value uses the difference of each sample characteristic (S_{ij}) minus the average feature of all samples in group (S) and divided (nd) in Equation (14).

$$Th_2 = \frac{1}{nd}\sum_{i=1}^{n}\sum_{j=1}^{d}|S_{ij} - \bar{S}|, \ \bar{S} = \frac{1}{nd}\sum_{i=1}^{n}\sum_{j=1}^{d}S_{ij} \qquad (14)$$

n is the number of face samples in group (S), d is the number of features in the face sample, and S_{ij} is the characteristic $(j$-$th)$ of the sample $(i$-$th)$.

The algorithms of HSNN scoping room are as follows.

Step 1: Calculate the Euclidean distance for each of the two sample points at S and find the two points with the smallest distance, denoted as B_{11} and B_{12}. L_1 is a line segment $\overline{B_{11}B_{12}}$. The scope of HSNN B_{11} and B_{12} is denoted as H_1, the scope of C_1 is:

$$C_1 = \{ X \mid \rho(X, L_1) < Th, X \in R^d\} \qquad (15)$$

$$L_1 = \{ Y \mid Y = \alpha B_{11} + (1 - \alpha)B_{12}, \alpha \in [0,1]\} \qquad (16)$$

$\rho(X, L_1)$ is the distance between sample point X and the circumference part of L_1, Th is the radius of dimension d.

Step 2: Know S '= S - {B_{11}, B_{12}}. Find the point S 'which is closest to denoted as B_{12} and denoted B_{13}, Make the second segment $\overline{B_{12}B_{13}}$, denoted as L_2. The scope of B_{12} and B_{13} is denoted as H_2, the coverage of C_2 is:

$$C_2 = \{ X \mid \rho(X, L_2) < Th, X \in R^d\} \qquad (17)$$

$$L_2 = \{ Y \mid Y = \alpha B_{12} + (1 - \alpha)B_{13}, \alpha \in [0,1]\} \qquad (18)$$

Step i: every remaining points in C_1, C_2, C, ..., C_{i-1}. Make point B_1 $(i + 1)$ at the remaining points, then determine the closest distance to B_1 and the line segment $\overline{B_{1i}B_{1(i+1)}}$, denoted as L_i. HSNN covers B_{1i} and $B_{1(i+1)}$ denoted as H_i, and the scope of C_i is:

$$C_i = \{ X \mid \rho(X, L_i) < Th, X \in R^d\} \qquad (19)$$

$$L_i = \{ Y \mid Y = \alpha B_{1i} + (1 - \alpha)B_{1(i+1)}, \alpha \in [0,1]\} \qquad (20)$$

Step (n-1): Form line segments for the last point, B_{1n} denoted as $\overline{B_{1(n-1)}B_{1n}}$, or L_{n-1}. HSNN covers $B_{1(n-1)}$ and B_{1n} is denoted as $H_{(n-1)}$, and the coverage of C_{n-1} is:

$$C_{n-1} = \{ X \mid \rho(X, L_{n-1}) < Th, X \in R^d\} \qquad (21)$$

$$L_{n-1} = \{ Y \mid Y = \alpha B_{1(n-1)} + (1 - \alpha)B_{1n}, \alpha \in [0,1] \qquad (22)$$

At the end of the process, get as much as possible HSNN and the scope of all sample training areas in face recognition cases. Scoping all areas of this network is represented as follows.

$$C = \bigcup_i^m C_i \qquad (23)$$

Sample recognition is done using a discriminant function, this function is used to determine the shortest distance of X face samples for many types of faces (many K faces). The discriminant function is illustrated in Equation (24).

$$class = min_{i=1}^{K}\rho_i, \text{ i} = 1,2, ..., \text{K}. \qquad (24)$$

ρ_i is the distance between sample X to the scope of class i. In other words, ρ_i is the shortest distance X and HSN in class i using Equation (25).

$$\rho_i = min_{j=1}^{M_i}\rho_{ij}, \text{ i} = 1,2, ..., \text{C}. \qquad (25)$$

M_i is the number of HSNN in class i, C is many classes, and ρ_{ij} is the distance of the sample you want to recognize and the scope of j HSNN of class i.

RESULTS AND DISCUSSION

Face image data is divided into 2 categories, namely 1596 training and 684 test data. The research experiment was carried out using α parameters, lots of training data, and lots of test data. Testing with 684 test data is used by 4 variants α values. The results of testing face images in the form of a percentage level of classification accuracy of face images grouped correctly are calculated using Equations (26). The research results are discussed on Table 1.

Table 1. The Results of Experiment Scenario with 684 Images Test Data Based on Various α Values

Scenario	Number of Images Training	Average of accuracy (%)			
		α = 0.1	α = 0.05	α = 0.025	α = 0.01
1	280	92.69	92.69	92.69	92.83
2	560	88.01	88.01	88.15	88.15
3	840	84.21	84.21	84.21	84.21
4	1120	79.38	79.38	79.38	79.38
5	1596	75.14	75.00	75.14	75.14

Table 1 showed that the highest accuracy was obtained at 92.83% with the number of 280 images training data and the number of 684 images testing data. The results show that the increase in accuracy is inversely proportional to many training data. This is because the more image data that is trained and tested; the greater the network formed from the image data. Large-scale neural networks will provide sufficient knowledge to get better accuracy. However, this research cannot occur because there are many factors that influence the accuracy of recognition such as the unfavorable pre-process stage and the weakness of LDA feature extraction in the form of the Small Sample Size (SSS) problem. SSS problems in LDA are caused by the number of dimensions greater than many samples per class and this is stated in the study [20]. Based on the research [20], face recognition in this study experienced SSS problems because the number of dimensions of the image was larger than many samples of each class, namely the image dimension of 900 and the number of samples for each class of 20.

The test results show that there is a decrease in the level of accuracy based on the determination of the parameter α. The increase in accuracy of this study is inversely proportional to the parameter α. This is because the smaller α parameter used, the greater the likelihood that neurons can be received in an HSNN scoping. This scoping of HSNN was obtained using

Equation (22) and the determination of whether or not the test sample was received was done using Equation (21). Thus, increasing accuracy will be better if α parameters used are getting smaller. While the increase in the amount of training data will affect the level of accuracy as indicated by the explanation of the three experiments in this study and get the results that the greater the amount of training data the smaller the accuracy obtained.

Based on the results of testing experiments it can be concluded that the combination of LDA and HSNN methods in Face recognition is able to provide the best accuracy of 92.83% with parameters α = 0.01 and the amount of training data as more as 280. The results of the introduction of LDA-HSNN are able to provide better accuracy than some methods studied by [21] using the Faces94 database with the best accuracy ratio of 90.8% for the classification of SVM with RBF kernel and 57.29% for the KNN classification.

CONCLUSION

The combination of Linear Discriminant Analysis (LDA) and Hyper Sausage Neural Networks (HSNN) can be used for face recognition. The increasing number of training images reduces the accuracy values in face image recognition

because of the weakness influence of LDA method for Small Sample Size (SSS). Therefor the dimensions need to be reduced to meet the requirements of LDA that affect the features of face images. The use of a small α parameter affects the increase accuracy in face image recognition.

REFERENCES

[1] S. Umer, B. C. Dhara, B. Chanda, Biometric Recognition System for Challenging Faces, Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graph, India: IEEE, (2015).

[2] R. Pal, A. K. Gautam, Age Invariant Face Recognition Using Multiclass SVM, India: IEEE, (2015).

[3] R. Toufiq, Md. R. Islam, Face Recognition System using PCA-ANN Technique with Feature Fusion Method, International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), Bangladesh: IEEE, (2014).

[4] Xianwei, L and Guolong, C. (2012). Face Recognition Based on PCA and SVM, China: IEEE.

[5] I.A. Abdul-Jabbar, J. Tan, Z. Hou, Wavelet Based Image De-Noising to Enhance The Face Recognition Rate, International Journal of Computer Science Issues (IJCSI), China, (2013), 648-653.

[6] C. HaiYan, L. JunHong, The Image Process Technologies in Face Recognition, China: IEEE, (2010).

[7] X. Zhou, Y. Shi, P. Zhang, G. Nie, W. Jiang, A New Classification Method for PCA-based Face Recognition, International Conference on Business Intelegence and Financial Engineering, IEEE, (2009), 445-449.

[8] Y. Wang, J. Li, J. Lin, L. Liu, The Contourlet Transform and SVM Classification for Face Recognition, China: IEEE, (2008), 208-211.

[9] Gubbi, M. Hanmandlu, M. F. Azeem, A Novel LBP Fuzzy Feature Extraction Method for Face Recognition, India: Annual IEEE India Conference (INDICON), (2013).

[10] E. Hidayat, A. F. Nur, A. K. Muda, C. Y. Huoy, S. A. Ahmad, Comparative Study of Feature Extraction Using PCA and LDA for Face Recognition, 7th International Conference on

Information Assurance and Security (IAS), Malacca, Malaysia: IEEE, (2011), 354-359.

[11] D. Boruah, K. K. Sarma, A. K. Talukdar, Different Face Regions Detection Based Face Expression Recognition, 2nd International Conference on Signal Processing and Integrated Networks (SPIN), India: IEEE, (2015), 459-464.

[12] W. Zhi-hai, M. Hua-yi, L. Hua-xiang, W.A. Shou-jue, Method of Biomimetic Pattern Recognition for Face Recognition, China: IEEE, (2003), 2216-2221.

[13] W. Shou-jue, C. Xu, W. Yan, Biomimetic Pattern Recognition – a New Model of Pattern Recognition Theory and Its Applications, Acta Electronica Sinica, 30 (10), (2002).

[14] Y. Xu, F. Luo, Y. Zhai, J. Gan, Joint Iris and Face Recognition Based on Feature Fusion and Biomimetic Pattern Recognition, Proceedings of the 2013 International Conference on Wavelet Analysis and Pattern Recognition, 14-17 July, Tianjin, China: IEEE, (2013), 202-208.

[15] Z. You-zheng, F. Hao, D. Li-jun, Face Recognition Based on Biomimetic Pattern Recognition by High Dimensional Geometry, China: IEEE, (2009).

[16] L. Ding, H. Feng, Quaternion K-L Transform and Biomimetic Pattern Recognition Approaches for Color-face Recognition, China: IEEE, (2009).

[17] N. H. Barnouti, Improve Face Recognition Rate Using Different Image Pre-Processing Techniques, American Journal of Engineering Research (AJER), Baghdad, Iraq, (2016), 46-53.

[18] N. Sharma, S. Saurav, S. Singh, R. Saini,. A. K. Saini, A Comparative Analysis of Various Image Enhancement Techniques for Face Images, International Conference on Advances in Computing, Communications and Informatics (ICACCI), India: IEEE, (2015), 2279-2284.

[19] Samsuryadi, Biomimetic Pattern Recognition for Writer Identification Using Geometrical Moment Invariants, Ph.D. Thesis. Universiti Teknologi Malaysia, (2013).Tharwat, T. Gaber, A. Ibrahim, A. E. Hassanien, Linear Discriminant Analysis: A Detail Tutorial, Egypt: AI Communication, (2017), 1-22.

[20] Bouzalmat, J. Kharroubi, A. Zarghili, Face Recognition Using SVM Based on LDA. Marocco: IJCSI, (2013), 171-179.