
The Application Of Artificial Neural Network To Detect The Position Of Human Face In Digital Image

Dr. Salman Abd Kadum

Abstract

One of the important parts in human face recognition is detecting face position. In this paper a implement face position detection experiment using (Artificial Neural Network, ANN) to give outputs of human face number, position and dimension as found in a digital image. The system is trained using available ace samples. Quickprop algorithm and active learning method are used to speed the system training process up. And also indicate the comparison of the training time with standard Backpropagation algorithm and the training with Quickprop algorithm. The experiment is conducted using 200, 300 and 400 data. For each of the trainings, the iteration is stopped when the error value reaches 0.05. It is observed that the bigger number of the training data of the Quickprop algorithm causes significant increase in the training rate. Based on the results of the experiment with 14 files containing 273 face images, the face detection system gives 70.24% detection rate and 62 false positives.

Key words: Face Detection, Artificial Neural Network, Quickprop, BackPropagation, and Active Learning

I. Introduction

The increasingly high ability of computer to process a big number of data at high speed results in the increasingly wider application of

biometrics in human life. Biometric system.is an identifying system of physical characteristics of human bodies such as finger prints, eye retina and speech recognition. The One of the methods used in the biometric system is face recognition. The face recognition aims at identifying individual face by comparing the face with the existing database. In the face recognition system, the detection of face position represents one of the important stages because in a real world a face may appears in an image of various dimension and position with various backgrounds [1].biometric technology has some advantages such as long storage, free of forgetting, hard to falsify and the uniqueness of human individual In this paper we implement a human face detection experiment using ANN that is able to give outputs of human face number, position and dimension as found in a digital image.

The ANN is an information processing system with its operation similar to that of biological nervous system [2]. The ANN is developed as a mathematical modal of biological nerves based on the assumption that [3]:

The process goes at simple elements called neurons.

- a. Signals are transmitted between neurons through connectors.

- b. Each of the connectors has a weight that will multiply the passing signals.
- c. Each of the neurons serves an activating function that will determine the output signal value.

The Neural Network can be classified into various kinds based on its architecture, which is the pattern of relations between neurons and its training algorithm that is the method to determine the weight of the connectors. Each neuron i in a network represents a simple processing unit that calculates its activation value s_i of excitation input that is also referred to as net input net_i , as shown weight of the units j and i , and θ_i denotes the biased value of the unit i . The biased value is

$$net_i = \sum_{j \in pred(i)} s_j w_{ij} - \theta_i \quad \dots\dots(1)$$

often substituted by the weight connected to the unit with the value 1 to make the representation easier. Thus, the biased value is considered to be similar to the connection weight. in eq. (1). [4].

The objective of the supervised learning is to determine the connection weight value in the network that the network can operate input-output mapping as expected. The mapping is done using a set of sample patterns or training data set. Each of the pattern pair p consists of input vector x^p and target vector t^p . Once the training has been in completion, the input x^p into the network should give the output value t^p . The magnitude of the difference between the target vector and the actual output is measured using error value, which is referred to as cost function [9-11], which is illustrated in eq.(2) :

$$E = \frac{1}{2} \sum_{p \in P} \sum_n (t_n^p - s_n^p)^2 \quad \dots\dots(2)$$

which n is the number of the units in the output layer. Essentially, the training aims at finding a global minimum value E .

The Backpropagation Algorithm is one of the algorithms of ANN that is widely used in pattern recognition. The algorithm is in generally used in multi-layer feed-forward ANN consisting of some layers and signals are siphoned in unidirectional way from input to output [6 -15] .

Essentially, the Backpropagation training algorithm consists of three steps, which are:

- a. Feed the training data value into the algorithm that it gives its output.
- b. Inverse propagation of error value is obtained.
- c. The adjustment of the connection weight is necessary to minimize error value.

The three steps are continuously repeated till the expected error value is obtained. Once the training has been in completion, it is only the first step useful in the ANN.

Mathematically, the basic idea of the backpropagation algorithm is the application of chain rule to calculate the impact of each of the weights on error function such as shown in eq.(3) and eq.(4).

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial s_i} \frac{\partial s_i}{\partial w_{ij}} \quad \dots\dots(3)$$

∂E Denotes to the partial deferential of error function, ∂s_i denotes to the partial deferential of activation value, and ∂w_{ij} denotes to the partial deferential of weights, and

$$\frac{\partial s_i}{\partial w_{ij}} = \frac{\partial s_i}{\partial net_i} \frac{\partial net_i}{\partial w_{ij}} = f'_{log}(net_i) s_j \dots\dots\dots(4)$$

In which $f'_{log}(net_i)$ is the first derivative of net_i , and w_{ij} is the connecting weight of the neuron j and the neuron i , s_i is the output, and net_i is the multiplication result of the input of neuron i .

The Quickprop algorithm is another algorithm of ANN that uses the assumption that the error function of each of the connecting weights is upward open parabola and the error curve gradient of a weight is not influenced by other weights. Thus, the change in the weight is calculated using local information of each of the weights only [1-3][6]. The change in the weight in the *existing* Quickprop algorithm is as follow in eq. (5).

$$\Delta w(t) = -\varepsilon \frac{\partial E}{\partial w}(t) + \frac{\frac{\partial E}{\partial w}(t)}{\frac{\partial E}{\partial w}(t-1) - \frac{\partial E}{\partial w}(t)} * \Delta w(t-1) \dots(5)$$

where

$\Delta w(t)$: The change in weight,

$\Delta w(t-1)$: The change in weight in prior epoch,

ε : Learning rate,

II. Mathematical Model

Multi-layer perceptron is a feed-forward ANN consisting of a number of neurons connected by connecting weights. In general, the neurons are arranged in layers consisting of an input layer, a hidden layer and an output layer to the three layers as shown in Figure (1), in general, this figure exits to explain the layers in ANN. The input layer receives external signals and passes the signals to the first hidden layer that will be forwarded to the output layer [5].

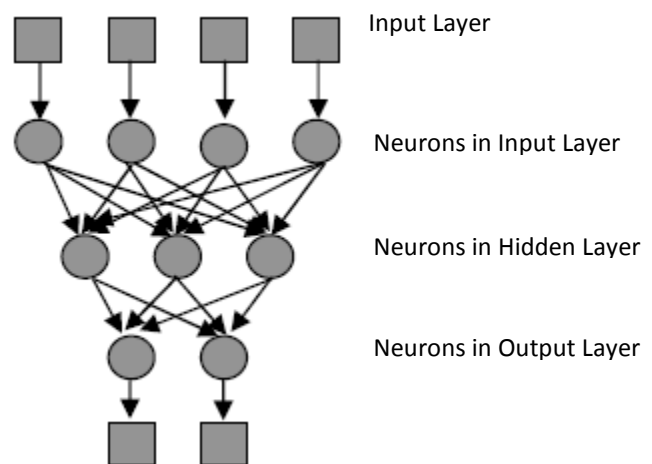


Figure (1): ANN with three layers

In which $pred(i)$ denotes the collection of the predecessors of the unit i , s_i denotes to activation value, w_{ij} denotes the connection in

the flow chart of Active Learning Technique is shown in Figure (2) below.

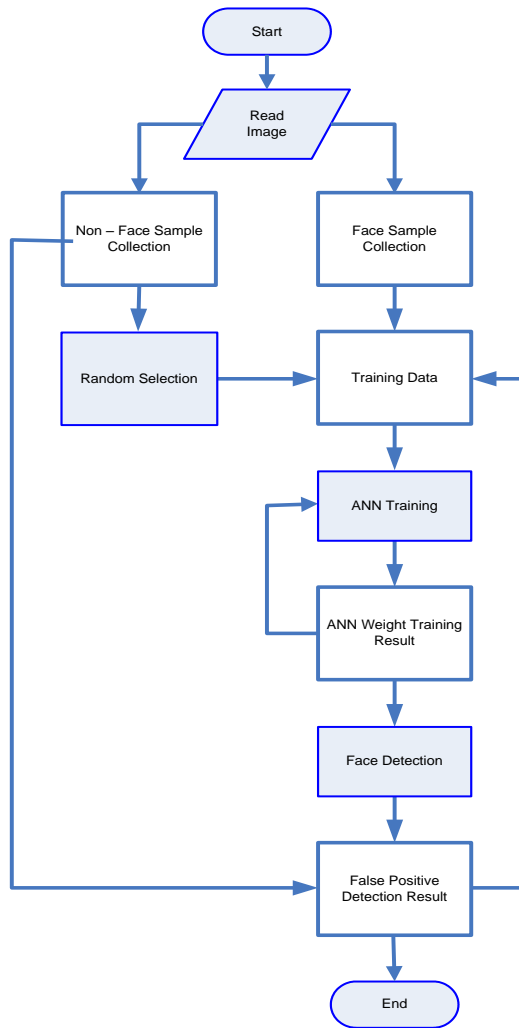


Figure (2). The Flow Chart of the Active Learning Technique for Face Detection System

The Active Learning Technique is organized stepwise using active learning method. At the first step, it begins with a few non-face data. In the next step, a few of the non-face training data is added. However, the addition data is selected from certain data,

$\frac{\partial E}{\partial w}(t)$: Derivative error, and

$\frac{\partial E}{\partial w}(t-1)$: The derivative error in prior epoch.

III. Method of Analysis and Results

The data of the study consists of: a set of images as data set for training and a set of images as data set for testing.

The training data set contains 300 face images of (60 x 60) pixels, while the non-face images are obtained from the image files that do not contain human faces. The system uses an ANN of the kind of multi-layer perceptron. The architecture is adapted from the results of the study by with more simplification. The input layer consists of 400 input units that receive the inputs of (60 x 60) pixels gray scale value of the detected sub-images. The gray scale of about (0 to 255) is normalized into -1 and 1 before it is fed as inputs. The output layer consists of a unit with the output value of about -1 and 1. The value 1 in the training data set is defined for face data and -1 for non-face data. The hidden layer consists of 25+16 = 41 units. The first part is connected to the input layer and forms 25 areas of 8 x 8 pixels. The second part is connected to the input layer and forms 16 areas of (10 x 10) pixels. In general, the network contains 883 connecting weights, including biases. It is simpler than Rowley's system [2] in which it is connecting weights amounts to 437,

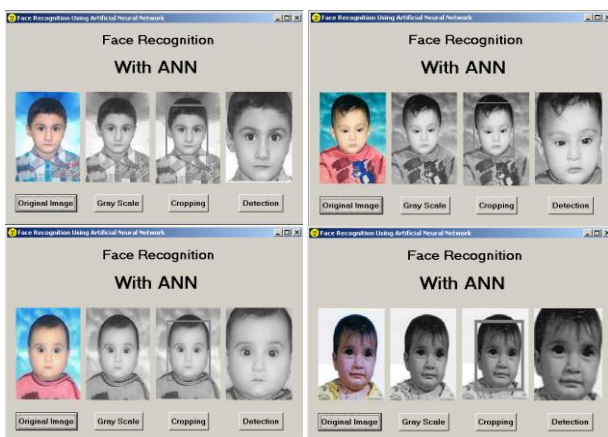


Figure (3). The Processed Samples of Children Face Data.



Figure (4). The Processed Samples of Human Face Data.

During the face detection of the image, faces are detected in some contiguous locations. The locations are referred to as face candidates. Therefore, it is necessary to combine the locations in a merging process. It is the process of combining the candidate locations of the contiguous faces as case study as shown in Figure (5).



which is detected as face (i.e., false positive) in the results of the prior training step. Thus, the number of the training data used in the ANN will be smaller and the time will be shorter. Figure (3) above illustrates the active learning technique used in the face detection system.

The face detector uses the neural network architecture similar to that used for training. The connecting weight is the last one resulting from the training process. The detecting result will be considered to be a face if the output of the ANN is more than 0, and as non-face if it is less than or equals to 0.

Face position may be found anywhere in the detected image and hence (60 x 60) pixels window is used because it enables dragging over the whole area of the image. The image areas through which the window is dragged will be examined one by one in order to find out whether there is a face or not. Anticipating various face dimensions in the detected image, the image is minimized gradually at the scale 1:1, 2 as done by [3]. In each of the minimized image, the 60 x 60 window will be dragged over the whole area of the image.

Before the training data set is used, the image will be pre-processed through the following steps [16-18]:

- Histogram equalization to improve its contrast.
- Masking, this is removing image angles to reduce image variations and to minimize data dimension.
- Normalization, this is converting gray scale intensity value of the image into the range of -1 to 1.

The preprocessing steps are also used in the face detection process using these faces images illustrated in Fig. (3) and Fig. (4).

Table (1) illustrates the detection results that have been found by other authors using the data test of MIT-23 [5]. The comparison is not considered as an absolute standard to infer that a method is better than others because of the following factors:

- The data set used in the training is not the same, and
- The number of the data used in the training is not the same.

Table (1): Some face detection results of the test data of MIT-23 [5]

METHOD	RATE DETECTION	FALSE POSITIVE
Support vector machines	74.2%	20
Distribution-based and Clustering	79.9%	5
Neural Networks	84.5%	8
Kullback relative information	94,1%	64

Table (2) summarizes the impact number of the training data on the detection results. It contains the detection results of an image with 15 faces and the total number of windows is 790, 17. It is observed that the higher is the number of the non-face training data, the smaller the number of the false

Figure (5). The detecting results before and after merging process.

The performance of the face detector is generally measured using two parameters, which are detection rate and false positive rate. The detection rate is the comparison between the number of the faces successfully detected and the number of all of the existing faces, while the false positive rate is the number of the non-face sub-images that are detected as face.

The samples of the detection results of some testing images are found in Figure (6). The image collection is firstly published and the test gives the detection rate of 70.24% and 62 false positives. The results are obtained from the training using 300 face data and 520 non-face data that are obtained using active learning method.



Figure (6). The samples of the face detection results of some image tests.

IV. Discussion

100	400	500	15/15	1
-----	-----	-----	-------	---

Table (3) indicates the comparison of the training time with standard Backpropagation algorithm and the training with Quickprop algorithm. The experiment is conducted using 200, 300 and 400 data. For each of the trainings, the iteration is stopped when the error value reaches 0.05. It is observed that the bigger number of the training data of the Quickprop algorithm causes significant increase in the training rate.

Table (3). The Impact of the Quickprop Algorithm on Training Rate.

THE NUMBER OF TRAINING DATA	ERROR	TRAINING TIME (SECOND)	
		Backprop Standard	Quickprop
200	0.05	49	12
300	0.05	69	96
400	0.05	464	467

Table (4) summarizes the comparison between the training results with active learning method to select the non-face data samples and the training results with non-face data that is selected randomly. In the first experiment, 600 data containing 300 face data and 300 non-face data is used. Meanwhile, in the second experiment, 820 data containing 300 face data and 520 non-face data is used. It is seen that the

positive results. It indicates that the more complete is the data, the better the training results of the system will be.

Table (2). The Impact Number of the Training Data on the Face Detection Performance.

THE NUMBER OF TRAINING DATA			DETECTION RATE	FALSE POSITIVE
Face	Non-Face	Total		
100	100	200	15/15	42
100	120	220	15/15	32
100	140	240	15/15	31
100	160	260	15/15	24
100	180	280	15/15	23
100	200	300	15/15	19
100	220	320	15/15	10
100	240	340	15/15	5
100	260	360	15/15	4
100	280	380	15/15	2
100	300	400	15/15	2
100	320	420	15/15	2
100	240	440	15/15	2
100	360	460	15/15	1
100	380	480	15/15	1

88-162). Pittsburgh, PA: Carnegie Mellon University, pp: 88-162,

URL

<http://www.cs.cmu.edu/afs/cs.cmu.edu/user/sef/www/publications/qp-tr.ps>.

- [2] R. Battiti, 1989, "Accelerated backpropagation learning: Two

THE NUMBER OF TRAINING DATA	RANDOM DATA		ACTIVE LERANING	
	Detection Rate	False Positive	Detection Rate	Fals e Positive
600	62.42%	160	70.24%	54
820	63.76%	232	70.24%	62

optimization methods", Complex Systems, Vol.(3), pp: 331–342.

- [3] A. C., Veitch, G., Holmes, 1991, "A modified quickprop algorithm", *Neural Computat* Vol. (3), pp:310–311.

- [4] W. Schiffmann, M. Joost, R. Wernor, 1993, "Comparison of Optimized Backpropagation Algorithms", *Proc. Of the European Symposium on Artificial Networks (ESANN) '93 Brussels*.

- [5] H. Rowley, S. Baluja, T. Kanade, 1998, "Neural Network-Based Face Detection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. (20), no. (1).

active learning technique gives better results. It means that the active learning technique can select the truly required data that it minimizes the number of the training data used.

Table (4). The results of the Comparison between Active Learning Method and Random Data.

V. Conclusion

From the discussion above, we conclude that the ANN of the kind of multi-layer perceptron can be applied to detect faces in digital image. In the sample-based face detection, the face detection performance is highly dependent of the quality and the quantity of the existing sample data. For a big number of data, Quickprop algorithm causes significant increase in the training rate. The active learning method can be used to select more appropriate samples that it minimizes the number of the training data used, and gives better results.

VI. Reference

- [1] S. E., Fahlman, 1988, " An empirical study of learning speed in back-propagation networks", (*Tech. Rept. No. CMU-CS-*

- [12] A. S. Tolba, A.H. El-Baz, and A.A. El-Harby, 2006, "Face Recognition: A Literature Review" *International Journal of Information and Communication Engineering*, Vol. (2). pp: 88-103.
- [13] J. Kumar Basu, D. Bhattacharyya, K. Taihoon, 2010, "Use of Artificial Neural Network in Pattern Recognition", *International Journal of Software Engineering and Its Applications*, Vol. (4), No. 2, pp: 23-34.
- [14] M. Agarwal, N. Jain, M. Kumar, and H. Agrawal. 2010 "Face Recognition Using Eigen Faces and Artificial Neural Network ", *International Journal of Computer Theory and Engineering*, Vol. 2, No. 4, pp: 624-629.
- [15] N. P., Kalavdekar, 2010 "Face Detection using Neural Network", *International Journal of Computer Applications*, Vol. (14), pp: 36-39.
- [16] T. Kumar, K. V. Singh , S. Malik 2011 , " Artificial Neural Network in Face Detection", *International Journal of Computer Applications*, Vol. (14) No. (3), pp: 5-7.
- [16] S. Kakarlapudi, M.R.Narasinga Rao, 2011, "A Novel Facial Recognition Model Using An Artificial Neural Network", *International Journal of Science and Advanced Technology (ISSN 2221-8386)*, Vol. (1) No. (5), pp: 242-245.
- [17] A. Bouzalmat, N. Belghini, A. Zarghili, and J. Kharroubi, 2011, "Face Detection And Recognition Using Back Propagation Neural Network And Fourier Gabor
- [6] A., Nikov, S., Stoeva, 2001, "Quick fuzzy backpropagation algorithm", *Neural Network* Vol. (14), pp:231-244.
- [7] E., Hjelmas, B. k., Low, 2001, "Face Detection: A Survey ", *Computer Vision and Image Understanding*, No. (83), pp: 236-274.
- [8] R., Caruana, S., Lawrence, and L., Giles, 2001, "Over fitting in Neural Networks: Backpropagation, Conjugate Gradient, and Early Stopping.", *In Advances in Neural Information Processing Systems, Denver, Colorado, MIT Press*, Vol. (13), pp: 402-408.
- [9] M. H., Yang, D.J., Kriegman, and N., Ahuja, 2002, "Detecting Faces in Images: A Survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24 (1), pp: 34-58.
- [10] M. Shin-yo, U. Haruki, 2003 "Face Detection and Facial Feature Localization for Human-machine Interface", *NII Journal* No. (5), pp: 25-39.
- [11] S. Wiegand, C. Igel, and U. Handmann, 2004 "Evolutionary Multi-Objective Optimization Of Neural Networks For Face Detection", *International Journal of Computational Intelligence and Applications, World Scientific Publishing Company*, Vol. (4), No. (3), pp: 1-16.

Filters”, *Signal & Image Processing: An International Journal (SIPIJ)* Vol. (2), No. (3), pp: 15-21.

- [18] V. Pushpaja, D.S. Chaudhari, 2012, “Facial Expression Recognition using Neural Network – An Overview” *International Journal of Soft Computing and Engineering (IJSCE)* ISSN: 2231-2307, Vol. (2), Issue-(1), pp: 224-227.

الخلاصة

احد اهم الامور في تمييز الوجه البشري هو تحديد مركز الوجه, في هذا البحث تم تنفيذ تجربة لتحديد موقع الوجه باستخدام الشبكة العصبية الذكية ANN حيث تم الحصول على الوجه من حيث الموقع والابعاد للصورة الرقمية.

طبقت التجربة على نماذج اوجه معطاة بمساعدة خوارزمية QuickProp [طريقة التعليم النشطة Active Learning Method لتسريع عمل التجربة. وتم ايضا تأشير وملاحظة فروقات وقت التنفيذ بأستخدام خوارزمية الانتشار الخلفي القياسية.

نفذت التجربة باستخدام 200, 300, 400

صورة رقمية لكل تطبيق, تتوقف عملية التكرار عندما تصل نسبة الخطأ الى 0.05 مع ملاحظة ان الرقم الاكبر من البيانات لخوارزمية [QuickProp] تعطي تزايد ملحوظ في نسبة التدريب بالاعتماد على نتائج التجربة وباستخدام 14 ملف و 273 صورة رقمية تم تحديد موقع الوجه بنسبة 70.24 مع نسبة 62 لتحديد الصورة التي لا تتضمن وجه

