Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 32 (2007) 458–468
Expert Systems with Applications
www.elsevier.com/locate/eswa

Glaucoma detection using adaptive neuro-fuzzy inference system

Mei-Ling Huang [a,*], Hsin-Yi Chen [b], Jian-Jun Huang [c]

[a] Department of Industrial Engineering and Management, National Chin-Yi Institute of Technology, 35, Lane 215, Sec. 1, Chuang San Road., Taiping, Taichung 411, Taiwan
[b] Department of Ophthalmology, Glaucoma Service, China Medical University Hospital, Taiwan
[c] Institute of Production System Engineering and Management, National Chin-Yi Institute of Technology, Taiwan

Abstract

Purpose. To develop an automated classifier based on adaptive neuro-fuzzy inference system (ANFIS) to differentiate between normal and glaucomatous eyes from the quantitative assessment of summary data reports of the Stratus optical coherence tomography (OCT) in Taiwan Chinese population.

Methods. This observational non-interventional, cross-sectional, case–control study included one randomly selected eye from each of the 341 study participants (135 patients with glaucoma and 206 healthy controls). Measurements of glaucoma variables (retinal nerve fiber layer thickness and optic nerve head topography) were obtained by Stratus OCT. Decision making was performed in two stages: feature extraction using the orthogonal array and the selected variables were treated as the feeder to adaptive neuro-fuzzy inference system (ANFIS), which was trained with the back-propagation gradient descent method in combination with the least squares method. With the Stratus OCT parameters used as input, receiver operative characteristic (ROC) curves were generated by ANFIS to classify eyes as either glaucomatous or normal.

Results. The mean deviation was -0.67 ± 0.62 dB in the normal group and -5.87 ± 6.48 dB in the glaucoma group ($P < 0.0001$). The inferior quadrant thickness was the best individual parameter for differentiating between normal and glaucomatous eyes (ROC area, 0.887). With ANFIS technique, the ROC area was increased to 0.925.

Conclusions. With Stratus OCT parameters used as input, the results from ANFIS showed promise for discriminating between glaucomatous and normal eyes. ANFIS may be preferable since the output concludes the if–then rules and membership functions, which enhances the readability of the output.

Keywords: Adaptive neuro-fuzzy inference system; Orthogonal array; Glaucoma; Stratus OCT

1. Introduction

A method for detecting and documenting optic nerve structure should be part of routine clinical management of glaucoma. However, according to limited evidence available sensitivity and specificity of imaging instruments for detecting of glaucoma are comparable to that of expert evaluation of stereo color-photography. Recently, digital imaging, such as scanning laser tomography, scanning laser polarimetry and optical coherence tomography, is recommended as a clinical tool to enhance and facility the assessment of the optic disc and nerve fiber layer in glaucoma management. Structural measurement of the ONH and RNFL offers the prospect of improving the early detection of glaucoma (Choplin, Lundy, & Dreher, 1998; Cioffi, Robin, & Eastman, 1993; Hoh, Greenfield, & Mistlberger, 2000; Greany, Hoffman, & Garway-Heath, 2002; Sanchez-Galeana, Bowd, & Blumenthal, 2001; Uchida, Brigatti, & Caprioli, 1996; Wollstein, Garway-Heath, & Hitchings, 1998).

The ability to early glaucoma detection using these instruments has been widely described and discussed

* Corresponding author. Tel.: +886 4 23924505; fax: +886 4 23934620.
E-mail address: huangml@ncit.edu.tw (M.-L. Huang).

0957-4174/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2005.12.010

(Chen, Hung, & Hung, 2005; Chen, Wang, Lee, & Hung, in press; Choplin et al., 1998; Cioffi et al., 1993; Hoh et al., 2000; Greany et al., 2002; Sanchez-Galeana et al., 2001; Uchida et al., 1996; Wollstein et al., 1998; Hoh, Ishikawa, & Greenfield, 1998; Horn, Jonas, & Martus, 1999; Schuman, Hee, & Puliafito, 1995; Weinreb, Shakiba, & Sample, 1995). Automatic analysis of results using appropriate databases is helpful for identifying abnormalities consistent with glaucoma. Structural assessment should preferably include such a biostatistical analysis (Greany et al., 2002). Classification quantitative structure–activity relationship studies are used for both the selection of principal physicochemical characteristics (descriptors) and relating them to biological activities and the derivation of mathematical models that involve these multivariate data in order to be used for predictive purposes in drug design (Loukas, 2001). Neural networks and fuzzy logic are two complimentary technologies. Neural networks can learn from data and feedback; however, the understanding of the knowledge or the pattern learned by the neural networks has been difficult. Conversely, fuzzy rule-based models are easy to comprehend because it uses linguistic terms and the structure of if–then rules. Unlike neural networks, fuzzy logic does not come with a learning algorithm. The combination of neural networks and fuzzy logic has created a new term as a neuro-fuzzy system (Yen & Langari, 1999). Neural networks, fuzzy systems, and the combination of both have been successfully applied to computer aided diagnosis such as microcalcification detection (Cheng, Lui, & Freimanis, 1998; Yu & Guan, 2000), automatic detection of distorted plethysmogram pulses in neonates and paediatric patients (Belal et al., 2002), detection of erythemato-squamous diseases (Übeyli & Güler, 2005) and lung nodule detection (Penedo, Carreira, Mosquera, & Cabello, 1998). Fuzzy inference systems have been successfully applied in fields such as automatic control, data classification, decision analysis, and computer vision (Fuzzy Logic Toolbox, 2005). Since slight variations in the parameters resulting from measurement errors might change the classification decision, fuzzy set theory plays an important role in dealing with uncertainty when making decisions in medical applications (Belal et al., 2002).

The purpose of this study was to evaluate the performance of fuzzy logic systems for glaucoma detect using Stratus OCT data.

2. Methods

This observational non-interventional, cross-sectional study included one randomly selected eye from each of the 341 study participants (135 patients with glaucoma and 206 healthy controls). All recruited cases were examined in the glaucoma service, Department of Ophthalmology, China Medical University Hospital. China Medical University Hospital is the ophthalmic referral center in mid-Taiwan area. Informed consent was obtained from all participants, and the study was approved by the Institutional Review Board of the China Medical University Hospital. This research follows the tenets of the Declaration of Helsinki.

All subjects underwent a complete ophthalmic examination, including slit lamp biomicroscopy, measurement of intraocular pressure, stereoscopic fundus examination and standard full threshold automated perimetry (30-2 mode, Humphrey Field Analyzer, model 750, HFA; Carl–Zeiss Meditec, Inc.). The inclusion criteria were best corrected visual acuity not worse than 20/40, spherical refractive error within the range of -6.00 D to $+3.00$ D, and reliable Humphrey visual field test results, defined as having fixation loss $<20\%$ and false-positive and false-negative errors $<25\%$. Subjects were excluded if they had a history of previous ocular surgery or other retinal disease, optic disc anomalies such as coloboma or optic disc drusen, or any kind of neurologic diseases that might cause visual field defects.

Inclusion criteria for normal subjects included no history of eye disease, no family history of glaucoma, intraocular pressure lower than 21 mmHg when measured by Goldmann applanation tonometry and normal optic disc appearance based on clinical stereoscopic examination. A normal result on the glaucoma hemifield test and corrected standard deviation (HFA, program 30-2) within normal limits were required. Subjects with normal eyes were volunteers from the staff at the China Medical University Hospital. Inclusion criteria for the patients with glaucoma were an early reproducible glaucomatous visual field defect in the absence of any other abnormalities to explain the defect, and a mean deviation of more than -7 dB. The mean age was 44.6 ± 14.3 years in the normal group and 42.8 ± 13.3 years in the glaucoma group. There was no significant difference in age between the two groups (t-test, $P = 0.228$). There was a significant difference in mean deviation of visual field between the normal group (-0.67 ± 0.62 dB) and the glaucoma group (-5.87 ± 6.48 dB) (t-test, $P < 0.0001$). Among the 135 glaucoma patients, 75 with high-tension open angle glaucoma, 25 with low-tension open angle glaucoma, 25 with primary angle closure glaucoma, and 10 with secondary glaucoma. All the glaucoma patients had received at least 12 months regular follow-up at the glaucoma service at China Medical University Hospital between March 2004 and March 2005. In this study, the patient with open-angle glaucoma and no history of documented intraocular pressure >21 mmHg was defined as having low-pressure glaucoma, whereas the patient with intraocular pressure >21 mmHg and no identifiable secondary causes was classified as having primary open-angle glaucoma. Glaucoma with a closed angle on gonioscopy with no evidence of secondary angle closure detected in the first visit was defined as primary angle-closure glaucoma. Second glaucoma included five uveitic glaucoma cases and five traumatic angle recession glaucoma cases.

2.1. Stratus OCT measurements

The Stratus OCT (Version A 4.0, Carl Zeiss Meditec Inc., Dublin, California, USA) consisted of an infrared-sensitive video camera to provide a view of the scanning probe beam on the fundus, a low-coherence interferometer as light source and detection unit, a video monitor, a computer and an image analysis system. The basic principles and technical characteristics of Stratus OCT have been described extensively (Schuman, 2004; Zafar, Gürses-Özden, & Makornwattana, 2004). The Stratus OCT delineates intraretinal and cross-sectional anatomy with axial resolution of $\leqslant$10 μm and transverse resolution of 20 μm. The Stratus OCT software package includes 18 scan acquisition protocols and 18 analysis protocols. Together they enable the optic disc, RNFL and the macula to be analyzed with a single instrument. The OCT protocol in our study included a regular 3.4 mm circular scan to determine RNFL thickness and a fast ONH radial scan to measure optic disc topography. All scans were completed in a single session by a trained operator after pupil dilatation with tropicamide 1% to achieve a minimal pupillary diameter of 6 mm.

The OCT protocol in our study included regular 3.4 mm circular scan to determine RNFL thickness. The RNFL thickness 3.4 mm scan protocol consisted of three separate circular scans with a diameter of 3.4 mm centered on the optic disc, each of which consisted of 512 A-scans obtained in 1.28 s. The total scan time for the three scans is between 1 and 2 min as each B-scan is briefly reviewed and saved before a subsequent scan can be recorded. The results were obtained from the mean of three scans. The NFL thickness was determined by the difference in distance between the vitreoretinal interface and a posterior boundary based on a predefined reflectivity signal level. A good-quality scan was one with a signal-to-noise ratio of >35, 100% accepted A-scans and good delineation of the anatomic boundaries. Subjects were not included in the study if the quality of the OCT image was suboptimal. The RNFL thickness report included the OCT image, the fundus image, and the thickness chart. Circular diagrams showed quadrant (temporal, superior, nasal, inferior) thickness and clock-hour RNFL thickness (11 o'clock = superior temporal: 45–75°; 1 o'clock = superior nasal: 105–135°; 7 o'clock = inferior temporal: 285–315°; 9 o'clock, temporal: 345–15°). The average RNFL thickness was the average thickness along the entire circumference of the optic disc.

The fast ONH radial scan protocol consisted of six linear scans crossing the optic scan. This protocol acquires six 4 mm radial scans in 1.92 s. The machine automatically determined the edge of the ONH as the end of the retinal pigment epithelium/choriocapillaris layer. This could be manually corrected in cases where the machine did not identify the edge correctly. A straight line connected the edges of the retinal pigment epithelium/choriocapillaries, and a parallel line was constructed 150 μm anteriorly. The structure below this line was defined as the disc cup and the structure above this line was defined as the neuroretinal rim. OCT ONH analysis measured the following: vertically integrated rim area volume, horizontally integrated rim width, disc area, cup area, rim area, cup/disc (C/D) area ratio, C/D horizontal ratio, and C/D vertical ratio.

Quality assessment of Stratus OCT scans was determined by an experienced examiner. Good quality scans had to have focused ocular fundus images, the signal strength needed to greater than six and a centered circular ring around the optic disc had to be present. Patients with unacceptable Stratus OCT scans were excluded from further analysis. ONH images were excluded when the machine incorrectly determined the edge of the ONH as the end of the retinal pigment epithelium/choriocapillaris layer in automatic mode; the ONH image was unacceptable; and images which could not be analyzed in the software A 4.0 version, such as those with very small C/D ratios. Therefore, the disc margin in our study was automatically defined in all eyes.

We selected the average RNFL thickness, quadrant thickness (temporal, superior, nasal, inferior), 12 clock hour (30° sector) RNFL thicknesses, and ONH analysis results (vertical integrated rim area, horizontal integrated rim area, disk area, cup area, rim area, cup/disk area ratio, cup/disk horizontal ratio, cup/disk vertical ratio) as our 25 input parameters. The perimetry and OCT examinations were all performed within a maximum period of 2 weeks. If the tests were done on the same day, the perimetry examination was done first.

2.2. Two-stage procedures

The collection of sufficient and accurate input data is the basic requirement to obtain an accurate model. Decision making was performed in two stages: feature extraction using the orthogonal array and the ANFIS trained with the back-propagation gradient descent method in combination with the least squares method.

2.3. Properties of orthogonal arrays

Orthogonal arrays have been developed to accomplish experiment designs with a number of arrays. Orthogonal arrays are designated by the notation L (L for Latin squares) with a subscript. The subscript refers to the number of rows in the table, which indicates the number of combinations in the design. The word orthogonal means in terms of array is that the columns of the arrays are balanced not only within themselves but also in any two columns in the array. That is, there are an equal number of levels within a column and the combinations of the levels between the columns considered are also equal in number.

2.4. Feature extraction using orthogonal arrays

Selection of the ANFIS inputs is crucial of designing the classifier based on pattern classification since even the best classifier will perform poorly if the inputs are not selected well. The selection of inputs is to find which sets of the input components can best represent a given pattern. There are many ways to perform feature selection. In this study, the orthogonal array (L_{32}) was chosen. Since each variable can be the candidate as one of the input components, two levels for each variable is set to present "chosen" (level 1) and "not chosen" (level 2) in the orthogonal array. The larger-the-better signal-to-noise (S/N) ratios were calculated according to the formula below:

$$\mathrm{S/N} = -10\log_{10}\left(\frac{1}{n}\sum_{i=1}^{n} y_i^2\right) \quad (1)$$

where y_i is the ith observation in each experiment combination.

Variables with higher S/N ratios are considered as useful variables and will be treated as input parameters for ANFIS. From the result of analysis of variance (ANOVA, Table 1) and response table (Table 2), the parameter combinations used as the input of ANFIS were inferior quadrant thickness, 2, 5, 10 and 12 o'clock segment thickness, vertical integrated rim area and horizontal integrated rim width.

2.5. Adaptive neuro-fuzzy inference system (ANFIS)

Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping then provides a basis from which decisions can be made, or patterns discerned. The process of fuzzy inference includes membership functions, fuzzy logic operators, and if–then rules (Fuzzy Logic Toolbox, 2005).

ANFIS provides a method for the fuzzy modeling procedure to learn information about a dataset, in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given input/output data (Fuzzy Logic Toolbox, 2005).

2.6. Architecture of ANFIS

Fig. 1 shows the five layers in ANFIS (Yen & Langari, 1999). In the first layer, all the nodes are adaptive nodes. The outputs of layer 1 are the fuzzy membership grade of the inputs, which are given by

Table 1
Analysis of variance of S/N ratios for feature selection

Parameter	Factor	Df[a]	SS[a]	MS[a]	F
Temporal quadrant	T	(1)	(10.66)	–	–
Superior quadrant	S	1	166.3	166.3	2.06
Nasal quadrant	N	(1)	(9.6)	–	–
Inferior quadrant	I	1	1088.9	1088.9	13.48#
Average RNFL thickness (μm)	Average	(1)	(98.29)	–	–
O'clock hour segment thickness					
12	X_1	1	1042.71	1042.71	12.91#
11 (superior temporally)	X_2	1	331.61	331.61	4.11
10	X_3	1	437.61	437.61	5.42#
9 (temporal)	X_4	(1)	(43.17)	–	–
8	X_5	(1)	(55)	–	–
7 (inferior temporally)	X_6	(1)	(21.25)	–	–
6	X_7	(1)	(46.59)	–	–
5	X_8	1	608.82	608.82	7.54#
4	X_9	(1)	(37.51)	–	–
3 (nasal)	X_{10}	(1)	(2.75)	–	–
2	X_{11}	1	657.38	657.38	8.14#
1	X_{12}	(1)	(2)	–	–
ONH analysis result					
Vertical integrated rim area	Y_1	1	489.43	489.43	6.06#
Horizontal integrated rim width	Y_2	1	353.76	353.76	4.38#
Disc area	Y_3	1	343.17	343.17	4.25
Cup area	Y_4	1	208.14	208.14	2.58
Rim area	Y_5	(1)	(6.33)	–	–
Cup/disk area ratio	Y_6	(1)	(0.16)	–	–
Cup/disk horizontal ratio	Y_7	(1)	(55.55)	–	–
Cup/disk vertical ratio	Y_8	(1)	(24.45)	–	–
	Error	20	1615.52	80.776	
	Total	31	7343.38		

"#" denotes statistical significance, and those parameters with "#" will be used for further processing.
[a] df: degree of freedom, SS: sum of square, MS: mean square.

Table 2
Response table of S/N ratios for features selection

Level	Factor								
	T	S	N	I	Average	X_1	X_2	X_3	X_4
1	10.143	8.441	11.268	16.554	12.473	16.429	13.940	14.419	9.559
2	11.298	13.000	10.173	4.887	8.968	5.012	7.501	7.023	11.882
Difference	1.154	4.559	1.096	11.667	3.505	11.417	6.438	7.396	2.323
	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	Y_1
1	12.032	11.535	9.514	15.082	9.638	11.014	15.253	10.971	14.631
2	9.410	9.906	11.927	6.359	11.803	10.427	6.188	10.470	6.810
Difference	2.622	1.630	2.413	8.724	2.165	0.586	9.065	0.500	7.822
	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8		
1	14.045	7.446	13.271	10.276	10.650	9.403	11.595		
2	7.396	13.995	8.170	11.165	10.791	12.038	9.846		
Difference	6.650	6.549	5.101	0.890	0.141	2.635	1.748		

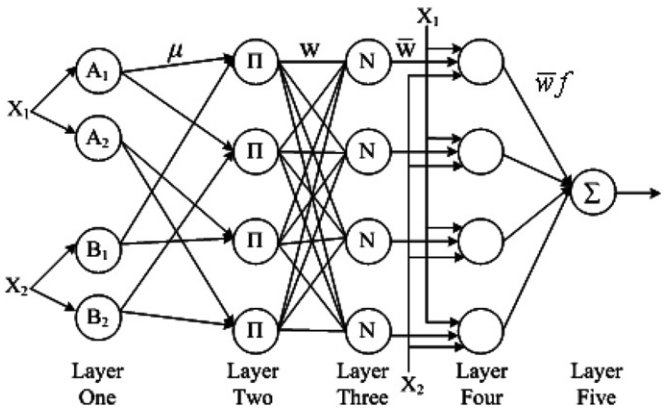

Fig. 1. The architecture of adaptive neuro-fuzzy inference system: five layers in adaptive neuro-fuzzy inference system.

$$y_i^1 = \mu_{A_i}(x_1), \quad i = 1,2 \tag{2}$$

$$y_i^1 = \mu_{B_{(i-2)}}(x_2), \quad i = 3,4 \tag{3}$$

where $\mu_{A_i}(x_1)$, $\mu_{B_i}(x_2)$ can adapt any fuzzy membership function.

In the second layer, every node in this layer is a fixed node, whose output is the product of all the incoming signals:

$$y_1^2 = W_1 = \mu_{A_1}(x_1) \cdot \mu_{B_1}(x_2) \tag{4}$$

$$y_2^2 = W_2 = \mu_{A_1}(x_1) \cdot \mu_{B_2}(x_2) \tag{5}$$

$$y_3^2 = W_3 = \mu_{A_2}(x_1) \cdot \mu_{B_1}(x_2) \tag{6}$$

$$y_4^2 = W_4 = \mu_{A_2}(x_1) \cdot \mu_{B_2}(x_2) \tag{7}$$

In the third layer, each node in this layer is a fixed node. The ith node calculates the ratio of the ith rule's firing strength to the sum of all rules firing strength:

$$y_i^3 = \overline{W}_i = \frac{W_i}{W_1 + W_2 + W_3 + W_4}, \quad i = 1,2,3,4 \tag{8}$$

which are so-called normalized firing strengths.

Every node i in the fourth layer is an adaptive node with a node function:

$$y_i^4 = \overline{W}_i f_i = \overline{W}_i \cdot (p_i x_1 + q_i x_2 + r_i), \quad i = 1,2,3,4 \tag{9}$$

where $\overline{W}_i$ is a normalized firing strength from layer 3 and $\{p_i, q_i, r_i\}$ is the parameter set of this node. Parameters in this layer are referred to as consequent parameters.

In the fifth layer, the single node in this layer is a fixed node, which computes the overall output as the summation of all incoming signals:

$$y^5 = \sum_{i=1}^{4} \overline{W}_i f_i \tag{10}$$

Thus, the ANFIS was constructed and can be functionally equivalent to a first-order Sugeno fuzzy model. It will be used in the present glaucoma study due to its transparency and efficiency.

2.7. Parameter design of ANFIS

Since there are many possible levels for factor including range of influence, initial step size, step size decrease rate, and step size increase rate in ANFIS, a L_9 orthogonal array (Table 3) was performed to figure out the best combination of parameters used in ANFIS. From factor effect analysis and ANOVA, the step size for parameter adaptation had an initial value of 0.005, the range of influence was 0.9,

Table 3
Parameter design of ANFIS

Factor	Level		
	1	2	3
Range of influence	0.3	0.6	0.9
Initial step size	0.005	0.01	0.015
Step size decrease rate	0.3	0.6	0.9
Step size increase rate	1.1	1.4	1.7

the step size decrease rate was 0.3, and the step size increase rate was 1.4.

The adequate function of ANFIS depends on the sizes of the training set and the testing set. In this study, the training set and testing set were formed by 230 and 111 data. The 230 data (140 normal cases and 90 glaucomatous cases) was used for training and the remaining 111 data (66 normal cases and 45 glaucomatous cases) was used for testing. In order to enhance the generalization capability of the ANFIS, the training and the testing sets were collected from different individuals. The ANFIS shown in Fig. 1 was implemented by using Matlab software package (Matlab version 6.5 with fuzzy logic toolbox).

Table 4
Subject demographics

	Normal	Glaucoma	p-Value[a]
Gender			
Male, n (%)	93(45.15%)	89(65.93%)	
Female, n (%)	113(54.84%)	46(34.07%)	
Age (years), mean ± SD	44.6 ± 14.3	42.8 ± 13.3	0.228
Mean deviation, mean ± SD (dB)	−0.67 ± 0.62	−5.87 ± 6.48	< 0.0001
Refraction, mean ± SD (Diopters)			

[a] Compared by t-test.

3. Results

3.1. Stratus OCT

The mean age was 44.6 ± 14.3 years in the normal group and 42.8 ± 13.3 years in the glaucoma group. There was no significant difference in age between the two groups (t-test, $P = 0.228$). There was a significant difference in mean deviation of visual field between the normal group (−0.67 ± 0.62 dB) and the glaucoma group (−5.87 ± 6.48 dB) (t-test, $P < 0.0001$) (Table 4).

Table 5 lists the statistical results in both groups of 25 glaucoma variables measured from Stratus OCT. There were 25 comparisons of variables; for $\alpha < 0.05$, the Bonferroni adjustment required $P < 0.002$ ($0.05/25 = 0.002$) for the difference to be considered significant. The t-test revealed significant differences between both groups on all of the parameters used in the study except disk area ($P = 0.725 > 0.002$).

3.2. Classification results

Table 6 summarizes the sensitivities, specificities and areas under the ROC curves for 25 individual parameters. The inferior quadrant thickness was the best individual parameter for differentiating between normal eyes and glaucomatous eyes (ROC area, 0.887). The average RNFL

Table 5
Stratus OCT glaucoma variables included in the 25 input set

Parameter	Factor	Normal	Glaucoma	p-Value[a]
Average RNFL thickness (μm)	Average	112.2 ± 12.2	83.9 ± 22.1	<0.0001
Temporal quadrant	T	84.2 ± 16.2	72.1 ± 25.2	<0.0001
Superior quadrant	S	137.0 ± 17.4	103.7 ± 29.5	<0.0001
Nasal quadrant	N	84.2 ± 19.6	62.0 ± 19.0	<0.0001
Inferior quadrant	I	143.7 ± 18.6	97.9 ± 32.6	<0.0001
O'clock hour segment thickness				
12	X_1	134.4 ± 25.8	103.2 ± 33.7	<0.0001
11 (superior temporally)	X_2	147.8 ± 21.7	112.38 ± 38.0	<0.0001
10	X_3	99.7 ± 20.0	82.1 ± 30.9	<0.0001
9 (temporal)	X_4	68.2 ± 15.5	60.4 ± 22.1	<0.0001
8	X_5	88.8 ± 20.3	74.0 ± 30.5	<0.0001
7 (inferior temporally)	X_6	160.0 ± 26.7	109.0 ± 45.0	<0.0001
6	X_7	152.2 ± 26.1	102.3 ± 40.5	<0.0001
5	X_8	119.3 ± 24.3	83.6 ± 27.1	<0.0001
4	X_9	82.9 ± 21.7	61.1 ± 20.9	<0.0001
3 (nasal)	X_{10}	73.5 ± 20.6	54.9 ± 18.1	<0.0001
2	X_{11}	97.0 ± 25.2	72.5 ± 24.0	<0.0001
1	X_{12}	129.6 ± 24.2	96.3 ± 29.5	<0.0001
ONH analysis result				
Vertical integrated rim area	Y_1	0.509 ± 0.409	0.253 ± 0.210	<0.0001
Horizontal integrated rim width	Y_2	1.804 ± 0.269	1.443 ± 0.356	<0.0001
Disc area	Y_3	2.53 ± 0.434	2.510 ± 0.571	0.725
Cup area	Y_4	0.717 ± 0.432	1.223 ± 0.679	<0.0001
Rim area	Y_5	1.807 ± 0.402	1.307 ± 0.528	<0.0001
Cup/disk area ratio	Y_6	0.275 ± 0.135	0.481 ± 0.226	<0.0001
Cup/disk horizontal ratio	Y_7	0.552 ± 0.154	0.702 ± 0.164	<0.0001
Cup/disk vertical ratio	Y_8	0.469 ± 0.128	0.684 ± 0.554	<0.0001

[a] Compared by t-test.

Table 6
The sensitivity and specificity of individual parameters

Parameter	Sensitivity		ROC area (%)
	Specificity of 80%	Specificity of 90%	
Average RNFL thickness (μm)	0.800	0.696	87.6
Temporal quadrant	0.474	0.274	64.4
Superior quadrant	0.689	0.526	83.2
Nasal quadrant	0.711	53.7	79.7
Inferior quadrant	0.807	0.733	88.7
O'clock hour segment thickness			
12	0.578	0.459	76.0
11	0.570	0.504	77.8
10	0.481	0.452	66.9
9	0.459	0.356	64.2
8	0.444	0.378	62.9
7	0.674	0.570	82.4
6	0.756	0.689	84.1
5	0.726	0.637	84.1
4	0.615	0.430	77.0
3	0.578	0.452	76.3
2	0.644	0.474	76.8
1	0.711	0.563	80.9
ONH analysis result			
Vertical integrated rim area	0.607	0.504	78.4
Horizontal integrated rim width	0.622	0.504	78.8
Disc area	0.274	0.163	53.8
Cup area	0.570	0.481	73.1
Rim area	0.593	0.496	77.6
Cup/disk area ratio	0.622	0.570	77.3
Cup/disk horizontal ratio	0.548	0.481	74.1
Cup/disk vertical ratio	0.644	0.541	77.7

thickness was the second best parameter (ROC area, 0.876). In addition, both 5 o'clock and 6 o'clock hour segment thickness ranked third (ROC area, 0.841).

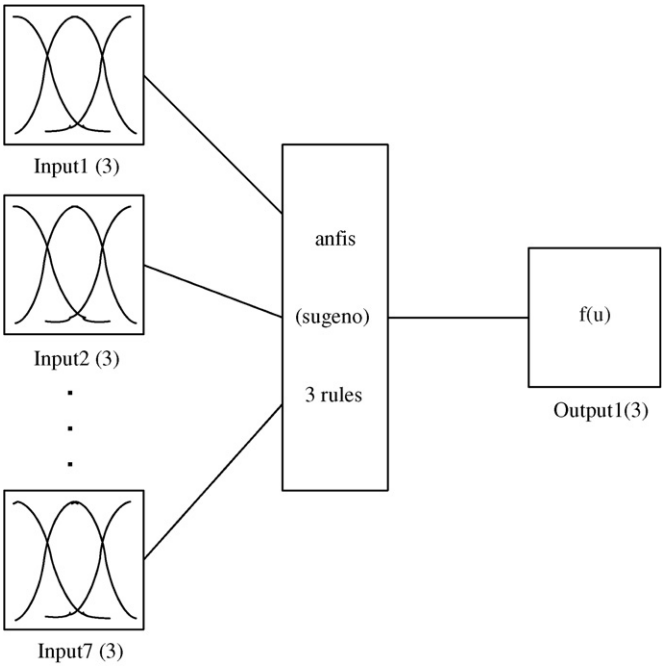

Fig. 2. The fuzzy rule architecture of the ANFIS using the generalized bell-shaped membership function.

Fig. 2 shows the fuzzy rule architecture of the ANFIS using the generalized bell-shaped membership function. There are three fuzzy rules in the architecture. The ANFIS shown in Fig. 3 used 230 training data in 100 training periods and the final mean square error was 0.3315. The membership functions for individual parameters are almost consistent before and after training in ANFIS except parameter horizontal integrated rim width. The before and after training membership functions for horizontal integrated rim width were displayed in Fig. 4. For the train-

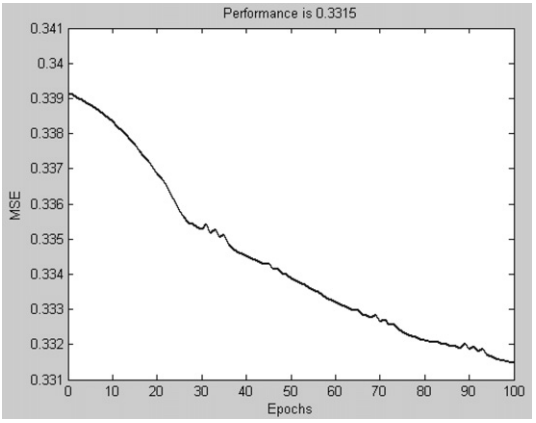

Fig. 3. The convergence plot after training of ANFIS. The final mean square error was 0.3315.

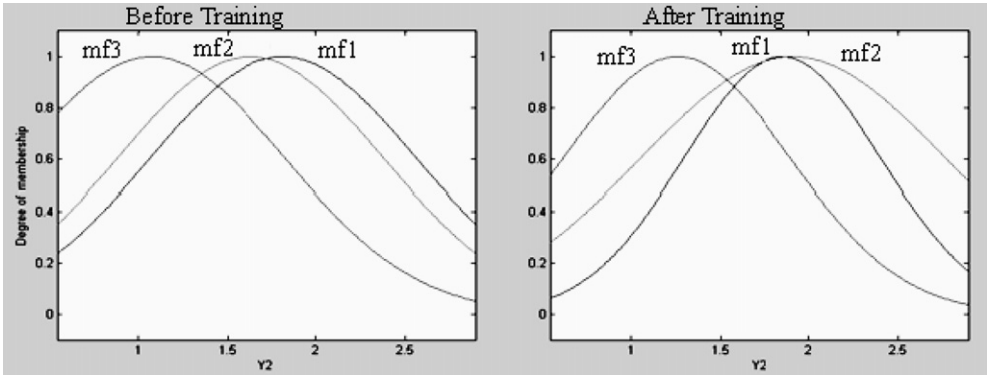

Fig. 4. The before and after training membership functions for horizontal integrated rim width.

Table 7
ANFIS classification results

Training set				Testing set			
Test	Disease		Total	Test	Disease		Total
	Present	Absent			Present	Absent	
Positive	78	11	89	Positive	35	6	41
Negative	12	129	141	Negative	10	60	70
Total	90	140	230	Total	45	66	111
Accuracy		90.00%		Accuracy		85.6%	

ing set, the area under the ROC curve from ANFIS was 0.942 with 0.889 and 0.878 sensitivities at 80% and 90% specificities. The overall accuracy was 90.0% for the training set (Table 7). For the testing set, the area under the ROC curve from ANFIS was 0.925 with 0.867 and 0.800 sensitivities at 80% and 90% specificities (Fig. 5). The overall accuracy was 85.6% for the testing set (Table 7). The if–then rules generated from ANFIS were shown in Fig. 6.

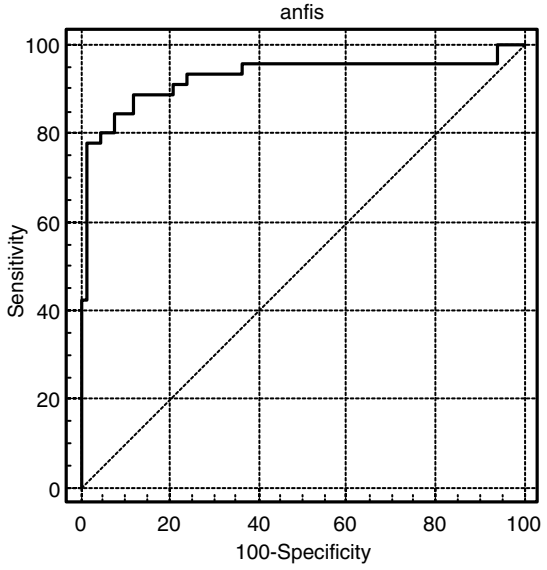

Fig. 5. Area under the ROC curve for ANFIS was 0.925 with 0.889 and 0.882 sensitivities at 80% and 90% specificities for the testing set.

4. Discussion

The previous cross-sectional studies have revealed that RNFL thickness measured by OCT can discriminate glaucomatous from normal eyes (Chen et al., 2005; Chen et al., in press; Hoh et al., 1998; Horn et al., 1999; Schuman et al., 1995; Weinreb et al., 1995). Our recent work (Chen & Huang, 2005; Huang & Chen, 2005; Huang, Chen, & Hung, 2005) also confirmed that combined both RNFL thickness and ONH parameters in Stratus OCT showed promise in discriminating glaucomatous from normal eyes using automated classifiers. There have been some reports using traditional and commonly used classification methods to help glaucoma diagnosis such as artificial neural networks, supported vector machine, and linear discriminant analysis (Bowd, Chan, & Zangwill, 2002; Brigatti, Hoffman, & Caprioli, 1996; Brigatti, Nouri-Mahdavi, & Weitzman, 1997; Goldbaum, Sample, & Chan, 2002; Goldbaum, Sample, & White, 1990, 1994; Henson, Spenceley, & Bull, 1997; Huang et al., 2005; Iester, Jonas, & Mardin, 2000; Mutlukan & Keating, 1994; Wollstein, Ishikawa, & Wang, 2005). However, there is no research available on glaucoma diagnosis with ANFIS application. Therefore here we have presented a new method for detecting glaucoma obtained from the summary data reports from Stratus OCT. We chose fuzzy logic in this system due to the uncertainty in the glaucoma classification, which is a result of imprecise boundaries between two classes "normal" and "abnormal". Using fuzzy logic enabled us to use this uncertainty in the classifier design and consequently to increase the credibility of the system output.

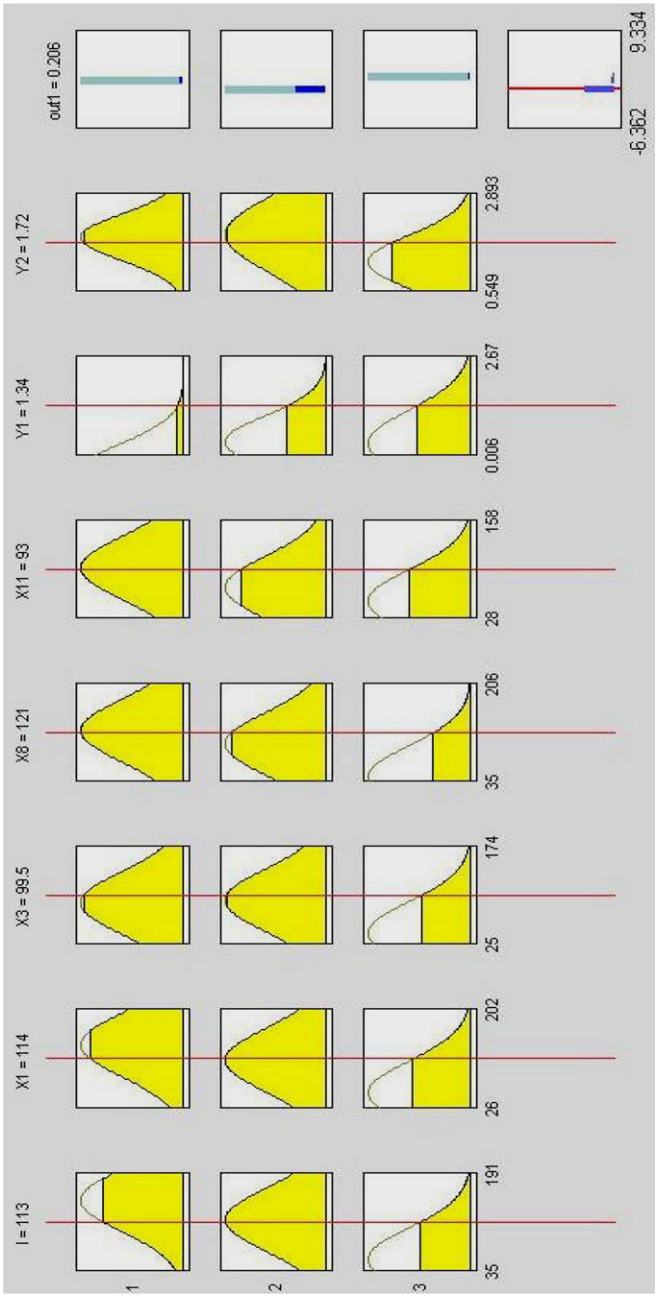

Fig. 6. The if–then rules generated from ANFIS results.

Neuro-fuzzy systems are fuzzy systems which apply artificial theory to determine the fuzzy sets and fuzzy rules by processing data samples. ANFIS utilizes the mathematical properties of artificial in tuning rule-based fuzzy systems that approximate the way human beings process information. Further investigations are required to determine if an appropriate fuzzy inference system can be designed using generalized membership functions applicable to a wide range of patient group. A question is always encountered when ANFIS is using: readability or precision. If a readable result is desired, we might be willing to accept a sub-optimal performance of the whole system. A Mamdani-type fuzzy system is a typical example of this. Conversely, if the linguistic interpretations of the fuzzy rules and more precise results are interested, a Sugeno-type fuzzy system is suitable (Nauck & Kruse, 1999). In our study, A Sugeno-type fuzzy system aiming to obtain better performance was chosen.

Compared ANFIS with ANN, unlike the black-box approach of ANN, the ANFIS approach is more transparent. Interpreting the ANN is very difficult. The problem-specific qualitative modeling representation can be easily understood and can be translated into medical terms to assist the day-to-day clinical diagnosis. Experts can incorporate into the selection of the inputs and the manipulation of model rules. The main features and advantages of the feature selection and the ANFIS developed in this paper are described as follows: (1) the optimal input set is determined based on the response table and ANOVA through orthogonal array experiments; (2) a general framework that combines ANN and fuzzy logic system is developed; (3) the non-linear ability in ANFIS; (4) fast processing time for decision making; (5) ANFIS beyond its classification performance increases the comprehensibility of the system.

Recently, there were some studies evaluating Stratus OCT in glaucoma diagnosis. Wollstein et al. (2005) reported that the highest area under ROC curves for distinguishing between groups were for ONH parameters (rim area = 0.97, horizontal integrated rim width = 0.96, vertical integrated rim area = 0.95) and peripapillary nerve fiber layer (NFL) thickness (ROC area = 0.94) followed by macular volume and thickness (both 0.80). A statistically significant difference existed in ONH and NFL areas under ROC curves when compared with macular area under ROC curves ($P \leqslant 0.007$, for both). The result showed that OCT ONH and NFL parameters provided similar discrimination capabilities between healthy eyes and those of glaucoma patients and superior discrimination capabilities when compared with macular parameters. Meanwhile, Medeiros, Zangwill, and Bowd (2005) also reported that no statistically significant difference was found between the area under ROC curves for the RNFL thickness parameter with the largest area under ROC curve (inferior thickness, ROC area = 0.91) and the ONH parameter with largest area under ROC curve (cup/disk area ratio, ROC area = 0.88) ($P = 0.28$). The RNFL parameter inferior thickness had a significantly larger area under ROC curve than the macular thickness parameter with largest ROC area (inferior outer macular thickness, ROC area = 0.81) ($P = 0.004$). However, a combination of selected RNFL and ONH parameters resulted in the best classification function for glaucoma detection with ROC area of 0.97 when applied to the independent sample. The best individual parameter was inferior quadrant thickness with ROC area of 0.887 in our study, and the result is similar to other study (Medeiros et al., 2005). Although the classification capability of single individual parameters is not good enough, the discriminating power increased after ANFIS classification was applied. To review the currently available

evidence on the use of OCT for diagnosis and evaluation of glaucoma, several studies have shown that RNFL assessment with OCT is able to discriminate patients with glaucomatous visual field loss from normal subjects. The area under the ROC curves for the earlier versions of the OCT to the latest model of Stratus OCT ranged from 0.79 to 0.97 depending on the parameters and characteristics of the population evaluated (Chen & Huang, 2005; Guedes, Schuman, & Herzmark, 2003; Huang & Chen, 2005; Huang et al., 2005; Kanamori, Nakamura, & Escano, 2003; Medeiros et al., 2005; Williams, Schuman, & Gamell, 2002; Zangwill, Bowd, & Berry, 2001). Besides, recent evidence also points toward the use of optic disc topography measurements for glaucoma evaluation and to a possible improvement in the diagnostic accuracy when RNFL and optic nerve head parameters are combined (Chen & Huang, 2005; Huang & Chen, 2005; Huang et al., 2005; Medeiros et al., 2005). Therefore, Stratus OCT did show promise in glaucoma diagnosis and evaluation. However, because of the changing software in the new machine and the lack of normative database in Oriental population especially the huge population of the Chinese, an efficient and reliable classification method is strongly urgent for Chinese population at this time.

Although the result is good in our study, there still leaves some more rooms on variable selection before the application of ANFIS. In this study, the orthogonal array was used to determine which and how many variables should be used as the input of the ANFIS. However, the ANFIS result might be inconsistent when different variables are used as the input of ANFIS. Moreover, the orthogonal array was used again as the parameter design for factor levels of ANFIS. If different factor levels are used in the model, the result will not be guaranteed. While ANFIS can provide useful information and support to the medical experts through identification of patterns that may not be readily apparent, human intervention to exploit the extracted knowledge is strongly recommended. The physician's intuition and interpretive skills are un-replaceable. Our results demonstrate promise that ANFIS technique can provide assistance in making decision on glaucoma diagnosis. We hope that our efforts in development of the automated classifiers can make strides in glaucoma diagnosis in Chinese population.

Acknowledgments

The authors would like to thank for the National Science Council of the Republic of China for financially supporting this research under Contract No. NSC-94-2213-E-167-002.

References

Belal, S. Y., Taktak, A. F. G., Nevill, A. J., Spencer, S. A., Roden, D., & Bevan, S. (2002). Automatic detection of distorted plethysmogram pulses in neonates and paediatric patients using an adaptive-network-based fuzzy inference system. *Artificial Intelligence in Medicine, 24*, 149–165.

Bowd, C., Chan, K., Zangwill, L. M., et al. (2002). Comparing neural networks and linear discriminant functions for glaucoma detection using confocal scanning laser ophthalmoscopy of the optic disc. *Investigative Ophthalmology and Visual Science, 43*, 3444–3454.

Brigatti, L., Hoffman, B. A., & Caprioli, J. (1996). Neural networks to identify glaucoma with structural and functional measurements. *American Journal of Ophthalmology, 121*, 511–521.

Brigatti, L., Nouri-Mahdavi, K., Weitzman, M., et al. (1997). Automatic detection of glaucomatous visual field progression with neural networks. *Archives of Ophthalmology, 115*, 725–728.

Chen, H. Y., & Huang, M. L. (2005). Discrimination between normal and glaucomatous eyes using Stratus Optical Coherence Tomography in Taiwan Chinese subjects. *Graefes Archives for Clinical and Experimental Ophthalmology*. online April 15.

Chen, H. Y., Hung, P. T., & Hung, T. J. (2005). Optical coherence tomography study on retinal nerve fiber layer thickness in open angle glaucoma. *American Journal of Ophthalmology, 139*(suppl S53). Abstract.

Chen, H. Y., Wang, T. H., Lee, Y. M., & Hung, T. J. (in press). Retinal nerve fiber layer thickness measured by optical coherence tomography and its correlation with visual field defects in early glaucoma. *Journal of the Formosan Medical Association*.

Cheng, H. D., Lui, Y. M., & Freimanis, M. R. (1998). A novel approach to microcalcification detection using fuzzy logic technique. *IEEE Transactions on Medical Imaging, 17*, 442–450.

Choplin, N. T., Lundy, D. C., & Dreher, A. W. (1998). Differentiating patients with glaucoma from glaucoma suspects and normal subjects by nerve fiber layer assessment with scanning laser polarimetry. *Ophthalmology, 105*, 2068–2076.

Cioffi, G. A., Robin, A. L., Eastman, R. D., et al. (1993). Confocal laser scanning ophthalmoscope: reproducibility of optic nerve head topographic measurements with the confocal laser scanning ophthalmoscope. *Ophthalmology, 100*, 57–62.

Fuzzy Logic Toolbox (2005). *For use with Matlab, user's guide, Version 2* (pp. 2–20). The MathWorks, Inc.

Goldbaum, M. H., Sample, P. A., Chan, K., et al. (2002). Comparing machine learning classifiers for diagnosing glaucoma from standard automated perimetry. *Investigative Ophthalmology and Visual Science, 43*, 162–169.

Goldbaum, M. H., Sample, P. A., White, H., et al. (1990). Discrimination of normal and glaucomatous visual fields by neural network. *Investigative Ophthalmology and Visual Science, 31*, S503. Abstract nr 2471.

Goldbaum, M. H., Sample, P. A., White, H., et al. (1994). Interpretation of automated perimetry for glaucoma by neural network. *Investigative Ophthalmology and Visual Science, 35*, 3362–3373.

Greany, M. J., Hoffman, D. C., Garway-Heath, D. F., et al. (2002). Comparisons of optic nerve imaging methods to distinguish normal eyes from those with glaucoma. *Investigative Ophthalmology and Visual Science, 43*, 140–145.

Guedes, V., Schuman, J. S., Herzmark, E., et al. (2003). Optical coherence tomography measurement of macular and nerve fiber layer thickness in normal and glaucomatous human eyes. *Ophthalmology, 110*, 177–189.

Henson, D. B., Spenceley, S. E., & Bull, D. R. (1997). Artificial neural network analysis of noisy visual field data in glaucoma. *Artificial Intelligence in Medicine, 10*, 99–113.

Hoh, S. T., Greenfield, D. S., Mistlberger, A., et al. (2000). Optical coherence tomography and scanning laser polarimetry in normal, ocular hypertensive, and glaucomatous eyes. *American Journal of Ophthalmology, 129*, 129–135.

Hoh, S. T., Ishikawa, H., Greenfield, D. S., et al. (1998). Peripapillary nerve fiber layer thickness measurement reproducibility using scanning laser polarimetry. *Journal of Glaucoma, 7*, 712–715.

Horn, F., Jonas, K., Martus, J. B., et al. (1999). Polarimetric measurement of retinal nerve fiber layer thickness in glaucoma diagnosis. *Journal of Glaucoma, 8*, 353–362.

Huang, M. L., & Chen, H. Y. (2005). Development and comparison of automated classifiers for glaucoma diagnosis using stratus optical coherence tomography. *Investigative Ophthalmology and Visual Science, 46*, 4121–4129.

Huang, M. L., Chen, H. Y., & Hung, P. T. (2005). Analysis of glaucoma diagnosis with automated classifiers using stratus optical coherence tomography. *Optical and Quantum Electronics, 37*, 1239–1249.

Iester, M., Jonas, J. B., Mardin, C. Y., et al. (2000). Discrimination analysis models for early detection of glaucomatous optic disc changes. *British Journal of Ophthalmology, 84*, 464–468.

Kanamori, A., Nakamura, M., Escano, M. F. T., et al. (2003). Evaluation of the glaucomatous damage on retinal nerve fiber layer thickness between normal and glaucomatous eyes using the Heidelberg retina tomography, GDx nerve fiber analyzer, and optical coherence tomography. *American Journal of Ophthalmology, 135*, 513–520.

Loukas, Y. L. (2001). Adaptive neuro-fuzzy inference systems: an instant and architecture-free predictor for improving QSAR studies. *Journal of Medicinal Chemistry, 44*, 2772–2783.

Medeiros, F. A., Zangwill, L. M., Bowd, C., et al. (2005). Evaluation of retinal nerve fiber layer, optic nerve head, and macular thickness measurements for glaucoma detection using optical coherence tomography. *American Journal of Ophthalmology, 139*, 44–55.

Mutlukan, E., & Keating, K. (1994). Visual field interpretation with a personal computer based neural network. *Eye, 8*, 321–323.

Nauck, D., & Kruse, R. (1999). Obtaining interpretable fuzzy classification rules from medical data. *Artificial Intelligence in Medicine, 16*, 149–169.

Penedo, M. G., Carreira, A., Mosquera, A., & Cabello, D. (1998). Comuter-aided diagnosis: A neural-network-based approach to lung nodile detection. *IEEE Transactions on Medical Imaging, 17*, 872–880.

Sanchez-Galeana, C., Bowd, C., Blumenthal, E. Z., et al. (2001). Using optical imaging summary data to detect glaucoma. *Ophthalmology, 108*, 1812–1818.

Schuman, J. S. (2004). In C. A. Puliafito & J. G. Fujimoto (Eds.), *Optical coherence tomography of ocular disease* (2nd ed.). Thorofare, NJ: Slack.

Schuman, J. S., Hee, M. R., Puliafito, C. A., et al. (1995). Quantification of nerve fiber layer thickness in normal and glaucomatous eyes using optical coherence tomography. *Archives of Ophthalmology, 113*, 586–596.

Übeyli, E. D., & Güler, I. (2005). Automatic detection of erythematosquamous diseases using adaptive neuro-fuzzy inference systems. *Computers in Biology and Medicine, 35*, 421–433.

Uchida, H., Brigatti, L., & Caprioli, J. (1996). Detection of structural damage from glaucoma with confocal laser image analysis. *Investigative Ophthalmology and Visual Science, 37*, 2393–2401.

Weinreb, R. N., Shakiba, S., Sample, P. A., et al. (1995). Association between quantitative nerve fiber layer measurement and visual field loss in glaucoma. *American Journal of Ophthalmology, 120*, 732–738.

Williams, Z. Y., Schuman, J. S., Gamell, L., et al. (2002). Optical coherence tomography measurement of nerve fiber layer thickness and the likelihood of a visual field defect. *American Journal of Ophthalmology, 134*, 538–546.

Wollstein, G., Garway-Heath, D. F., & Hitchings, R. A. (1998). Identification of early glaucoma cases with the scanning laser ophthalmoscope. *Ophthalmology, 105*, 1557–1563.

Wollstein, G., Ishikawa, H., Wang, J., et al. (2005). Comparison of three optical coherence tomography scanning areas for detection of glaucomatous damage. *American Journal of Ophthalmology, 139*, 39–43.

Yen, J., & Langari, R. (1999). *Fuzzy logic: Intelligence, control and information*. Prentice-Hall Inc., p. 444.

Yu, S., & Guan, L. (2000). A CAD system for the automatic detection of clustered microclacification in digitized mammogram films. *IEEE Transactions on Medical Imaging, 19*, 115–126.

Zafar, S., Gürses-Özden, R., Makornwattana, M., et al. (2004). Scanning protocol choice affects optical coherence tomography (OCT-3) measurements. *Journal of Glaucoma, 13*, 142–144.

Zangwill, L. M., Bowd, C., Berry, C. C., et al. (2001). Discriminating between normal and glaucomatous eyes using the Heidelberg retina tomography, GDx nerve fiber analyzer, and optical coherence tomography. *Archives of Ophthalmology, 119*, 985–993.