

A hybrid visual field classifier to support early glaucoma diagnosis

M^a Aránzazu Simón¹, Luis Alonso², Alfonso Antón³

¹Department of Computer Science, E.T.S.I. Informática, University of Valladolid, Campus Miguel Delibes s/n, 47011 Valladolid, Spain

²Department of Computer Science, University of Salamanca, Salamanca, Spain

³IOBA (Institute of Ophthalmology and Visual Sciences), University of Valladolid, 47005 Valladolid, Spain

arancha@infor.uva.es, lalonso@tejo.usal.es, aanton@ioba.med.uva.es

Resumen

Primary Open Angle Glaucoma is an eye disease that can, eventually, cause irreversible damage to the optic nerve. Because of this an accurate diagnosis at early stages of the disease is necessary to stop or delay its progression. Perimetry, one of the most important tests to detect glaucoma, gives a large amount of numerical data that is difficult to analyze. A number of approaches are described in the literature to overcome this problem, some of them using artificial neural networks, mainly MLP with BP. In this paper, a Hybrid Visual Field Classifier System is proposed, comprising a Self-Organizing Map (SOM) and a rule based expert system, integrating the knowledge that the SOM discovers with the expertise of the ophthalmologist. With this association, individual results of each component are improved up to a diagnostic precision of 97%.

Keywords: Glaucoma Diagnosis, Hybrid Classifier, Expert System, Self-Organizing Map (SOM).

1 Introducción

Nowadays, one of the main research approaches in glaucoma is devoted to early diagnosis, given that this disease is the second cause of blindness in developed countries. Since glaucoma irreversibly damages the optic nerve affecting the visual function, it is very important to detect it as soon as possible and stop, or delay, its progression.

Glaucoma diagnosis in the advanced stages is not difficult for the general ophthalmologist, since it is generally associated with a high ocular pressure, a clear optic disc cup and some typical visual field damage. Nevertheless, the diagnosis is much more difficult in the initial stages of the disease [3].

One of the most important tests is the perimetry, or exploration of the visual field (VF). This test allows visual field assessment and is continuously improving by supplying better and more complete data. To facilitate the interpretation, some statistical indexes are defined: mean sensitivity (MS), mean defect (MD), loss variance (LV), etc. Analysis of results is complicated since there is no unanimity on the criteria that define a VF as glaucomatous, but early diagnosis is even more difficult. The fact that many of the most prevalent eye diseases (cataract, diabetic or hypertensive retinopathy, myopia, etc.) can damage the VF must be taken into account.

Different studies found that the field defects due to glaucoma follow some defined morphological, topographic and evolutive patterns and this fact

led to the development of expert systems (Ocosmart [4, 10, 12]) and to the application of pattern recognition methods based on statistical approaches or on artificial neural networks [2, 7, 13, 8, 9] to facilitate the interpretation of VF data.

In this paper we develop a visual field classifier that supports early diagnosis of glaucoma. It is a hybrid classifier system, symbolic-connectionist, made by a SOM based classifier and a rule based expert system. We developed a tool, GESL, that facilitates the training and evaluation of the SOM maps and implies a modification of Kohonen's original software.

The paper is organized as follows: Section 2 describes the methods and patient data used; in Section 3 a tool developed to help in the training and labelling of the neural networks used for visual field classification is presented, and the best results are showed and compared with other works. Section 4 presents a hybrid symbolic-connectionist system to classify visual fields; finally, section 5 is devoted to the conclusions.

2 Methods and Patients

2.1 Data

2.1.1 Training Sample

A sample of 180 patients over 40 years of age was studied. The training sample included glaucoma patients, normal subjects, and patients with other ophthalmic diseases. All had previous perimetric experience, and their VFs had less than 15% false positive and false negative responses.

Group 1 (Glaucoma): Ninety-six consecutive patients with primary open-angle glaucoma and no other ophthalmic pathology were selected from the clinical records of the Glaucoma Unit. Patients were included whose progression to the advanced stages of the disease eliminated possible misdiagnosis: intraocular pressure ≥ 21 mmHg on at least two occasions, open angle, glaucomatous disc (cup/disc ratio ≥ 0.6 and/or cup/disc asymmetry ≥ 0.2 and/or notching of neuroretinal rim and/or disc hemorrhages), and pathologic VF (at least two contiguous defects of 10 decibels (dB) or more, or three contiguous defects of 5 dB or more). These fields were only considered for

inclusion criteria, but were not used for data processing. From these patients, previous VFs were chosen with less than 10 locations with defect > 4 dB, and mean defect ≤ 6 dB. On the date the VF selected for data processing was carried out, the rest of the inclusion criteria had to be satisfied (IOP ≥ 21 mmHg on at least two occasions, open angle, and glaucomatous disc as described above). Exclusion criteria for the glaucoma group were ocular hypertension, normal tension glaucoma, aphakia, or any other disease that could alter the VF.

Group 2. These patients were consecutively selected from the outpatient clinic and included 37 people without any ophthalmic disease (Group 2A, normal) and 47 patients with a high-prevalence ophthalmic disease (Group 2B, other diseases). Group 2B comprised patients with the following pathologies: 22 with mild to moderate cataract (visual acuity $\geq 10/20$), 12 cases of diabetic background retinopathy (microaneurysms, microhemorrhages and/or hard exudates in no more than one third of the macular surface and/or up to three retinal quadrants), and 13 with mild hypertensive retinopathy (stage I or II of the Keith-Wagener-Barker classification). Patients with more than one disease, glaucoma, ocular hypertension, aphakia and any other disease which could alter the VF, were excluded.

2.1.2 Test Sample

Neural networks were evaluated with an independent test sample of 48 subjects (Groups 3, 4A, and 4B) not included in the training sample. Group 3 included 20 patients with primary open-angle glaucoma and the same characteristics as Group 1. Group 4A comprised 9 people with no ophthalmic disease. Group 4B included 19 patients with other ophthalmic diseases and the same characteristics as Group 2B: seven with cataract, seven with diabetic background retinopathy, and five with mild hypertensive retinopathy.

2.1.3 Patient Examination

In all cases included in the training and testing samples, the clinical history, complete ophthalmic examination (visual acuity, refraction, anterior segment biomicroscopy, and fundus visualization with indirect ophthalmoscopy, and biomicroscopy

	Glaucoma	Normal	Cataract	Diabetes	HTA	TOTAL
Training sample	96	37	22	12	13	180
Test sample	20	9	7	7	5	48

Table 1: Training and test sample distribution.

with 90 diopter lens), and at least two tests with phases 1 and 3 of the G-1 program from Octopus 500 (Interzeag AG, Switzerland) were recorded. The first VF was excluded, and the second or subsequent examinations were selected. Only one eye per patient was included in the study. If both eyes met the inclusion criteria, one was selected at random.

The VF distribution in the different classes is shown in Table 1.

2.2 Preprocessing

Zone Field Distribution In order to reduce the number of variables, the VF was divided into 7 zones, taking into account the distribution of nerve fibres in the retina, clinical experience in the Glaucoma Unit of the University Hospital of Valladolid, and those used in the Glaucom-Easy expert system [2, 3, 16]. The zones are shown in Figure 1. The 59 sensitivity data, averaged for each zone, were used to train the neural networks. In this way, 59 data were reduced to 7.

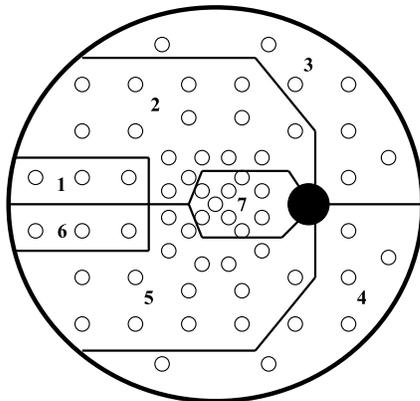


Figure 1: VF zone distribution.

3 GESL Development

In order to facilitate the training of a Self-Organizing Map (SOM), an experiment generator was developed. Although a SOM that dis-

tinguishes glaucoma from non-glaucoma is sufficient for the VF Classifier System (VFCS), it was considered to be of interest to carry out a more complete survey to find out if it is possible to distinguish between defects produced by different diseases. Self-Organizing Maps seem very appropriate for this task, since they can cluster data in a non-supervised way.

As a starting point, the SOM_PAK software [11] was used. This requires a number of parameters whose influence on the map performance is not clear. Thus, it is necessary to train several maps with different parameters to get satisfactory results. To this aim, we developed a tool, GESL [5], that generates sequences of training experiments covering the range of values of the parameters, and performs the training of the maps thus generated. Once trained, GESL labels the resulting maps, minimizing non-labelled neurons.

With all the labelled maps, a second filtering step is performed to discard those maps whose sensitivity (percentage of examples of a category correctly classified) and specificity (percentage of examples not belonging to a category and not classified as such) values are not good enough. GESL is implemented in Visual C++, taking the SOM_PAK 3.1 as reference and modifying some of its algorithms.

The selection of parameters for training was carried out as explained in [14].

3.1 Modification of SOM_PACK algorithms

Initially, we tried to classify the data (180 training and 48 test examples) by using SOM_PACK [11] directly. With five categories (Glaucoma, Normal, Cataract, Diabetes and HTA) a diagnostic precision of 46% was achieved, which increased to 72% when only two categories were left: glaucoma, non-glaucoma [1]. These results were not good enough, so it was decided to make some modifications to SOM_PACK.

SOM_PACK works in five steps: initialization, training, evaluation of mean quantization error,

labelling and map evaluation. The three initial steps were left untouched but we designed our own labelling and evaluation method to improve the above mentioned results. Finally, the filtering step with sensitivity and specificity was added.

3.1.1 Map labelling

In this step, a set of already classified examples are used with a label assigned.

1. Euclidean distance between characteristic vector of a training example and all the weight vectors of the trained map is computed.
2. The winner neuron is assigned the label of the example.

Steps 1 and 2 are repeated until the training sample is finished. At the end, each neuron of the map has the number of examples of each class that activated it.

3. Each neuron is labelled with the class that has more examples on it.
4. If there are two or more classes with the same number of examples, a minimum distance criterion is used to select the label.
5. A never winning neuron is left with no label.
6. The distance between non labelled neurons and the rest is computed, and the minimum distance is used to label the neuron.

3.1.2 Map Evaluation

We have to get a classifier whose success ratio is maximum. So the choice criterium is not the quantization error but the sensitivity (S) and specificity (E), as is usual in medical practice. GESL determines the best maps according to their diagnostic precision (DP).

3.2 Results

With the GESL tool, maps were trained to classify VF. Networks classifying in two (glaucoma, non glaucoma), three (glaucoma, normal, other

diseases) and five (glaucoma, normal, diabetes, cataract, HTA) categories were considered. Experiments were made with the 59 original data and with the reduced vector of dimension 7, as explained above. In this way results can be compared.

The results, detailed in [14, 15], are summarized in Table 2. The best ones were obtained using the average defects in the 7 zones of Figure 1 as input since the expert knowledge is implicitly present in the zone division. Besides, we proved that the SOMs cannot classify correctly when five classes are used for the visual fields, giving sensitivity and specificity values that are too low to be considered reliable. Table 3 shows the best SOM obtained for classifying in two categories (glaucoma and non-glaucoma).

Number of categories	DP (%)	
	59 points	7 zones and MD
2	86	94
3	73	79
5	67	69

Table 2: Best diagnostic precision using SOM for 59 points and 7 zones according to the number of categories.

3.3 Comparison with other works

In order to have a more complete view, a summary of the most relevant results obtained in VF classification is presented, especially those using ANN or the same data set as in our work. This summary is given in Table 4. It is apparent that our results are rather significant with respect to previous ones. We got sensitivity and specificity values of 95% and 93%, when using two categories. These figures are similar to the ones obtained in the Ph.D. thesis by Reyes [13], with the same data set.

It is important to note that we used VF with very incipient defects in order to make an early diagnosis, as Antón and Reyes did previously. In some other works, classes were Glaucoma and Normal, and data sets from normal patients did not have abnormal VF. In our case, non-glaucoma cases included VF with defects from other diseases, which made the classification task more difficult. To sum up, it could be said that ANN, and in particular SOMs, are a good tool to help in the early diagnosis of glaucoma.

Topology*	$\alpha 1$	$\alpha 2$	Variables	S(%)	E(%)	DP(%)
8×3	0.03	0.01	7 zones and MD	95	93	94

Table 3: Best SOM with two categories and 7 points.

Technique	Groups	Glaucoma Grade	S (%)	E (%)
Neural networks MLP with BP (Goldbaum [7])	Glaucoma vs. Normal	Any	65	72
Neural networks MLP with BP (hierarchical network) (Zahlmann [17, 20])	Glaucoma vs. Normal vs. Other Pathologies	?	80-89	?
Neural networks RBF (hierarchical network) (Zahlmann [18, 19])	Glaucoma vs. Normal vs. Other Pathologies	?	72-98	70-95
Logistic Discrimination(7 zones, MD and LV) (Antón [2])	Glaucoma vs. Normal	MD \leq 6dB, less than 10 defective points	80	92
Neural networks MLP with BP (7 zones, hierarchical network) (Reyes [13])	Glaucoma vs. No glaucoma	MD \leq 6dB, less than 10 defective points	100	82
Kohonen Networks (SOM) (7 zones) (Simón [14, 15])	Glaucoma vs. No glaucoma	MD \leq 6dB, less than 10 defective points	95	93

Table 4: Summary of results of previous works.

4 Visual Field Classifier System

In this section we present the Visual Field Classifier System (VFCS). It was developed to be integrated into an expert system for global diagnosis of glaucoma. VFCS classifies by using two subsystems: an expert system to diagnose the visual field and a neural network. The output of both subsystems are integrated through production rules. The whole system was implemented in Visual C++.

In our opinion, combining the symbolic and connectionist approaches leads to an improvement in diagnostic precision. The architectural design of the VFCS is shown in Figure 2. It includes three blocks:

- Visual Field Expert System.
- SOM to classify VF.
- Integration by production rules.

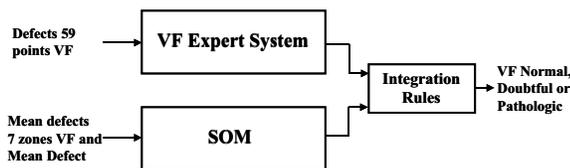


Figure 2: Architectural scheme of the VFCS.

4.1 Expert System for VF diagnosis

The objective of this expert system is to help in the classification of visual fields to diagnose incipient glaucoma. This point is essential to justify some decisions in the design of production rules. In a previous work by Gaasterland et al [6], a quantitative method was developed to measure the reliability of the test and the gravity of the VF defects on a Humphrey campimeter. The quantification of the glaucoma defects were based on the number and depth of clusters on the upper and lower halves of the field and on the nasal area. Their range goes from 0 (no defect) to 20 (all points with deep defects). They start from a value 0 and, by applying a set of rules, compute the final value.

Our system evaluates the sensitivity defects from

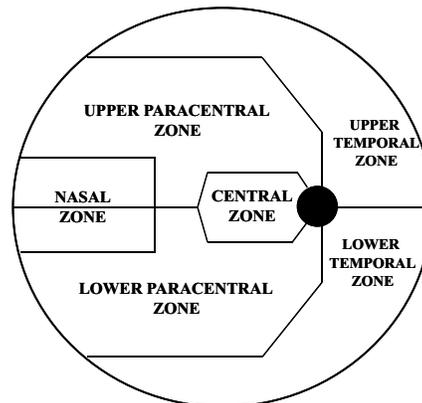


Figure 3: Visual Field Zones.

Doubtful(Glaucoma)	Doubtful (Non-glaucoma)	Total Doubtful
5 (25%)	12 (43%)	17 (35%)

Table 8: Examples of each class classified as doubtful by SOM.

S (%)	E (%)	DP (%)	FP	FN	Doubtful (%)
100	87	94	2	0	35

Table 9: Diagnostic figures for the class glaucoma with SOM.

	S (%)	E (%)	DP (%)	FP	FN	Doubtful (%)
Expert System	100	84	90	3	0	40
Neural Network	100	87	94	2	0	35
VFCS	100	94	97	1	0	31

Table 10: Diagnostic figures for VFCS.

4.2 SOM modification to classify VF

We used the SOM that produced the best results, as described in section 3 of this work. Note that the rules applied to label the map, used by the GESL tool, helped to improve the results. Two of these rules were 3 and 4, repeated here for clarity:

- Each neuron is labelled with the class that has more examples on it.
- If there are two or more classes with the same number of examples, a minimum distance criterion is used to select the label.

By a careful analysis of training examples, and winner neurons for each case, we found that some neurons made nearly the same number of errors and successes, since they are activated by very similar examples belonging to different classes. These neurons are labelled as "doubtful". The original map is shown in Figure 5, where glaucoma labelled neurons are shown in black.

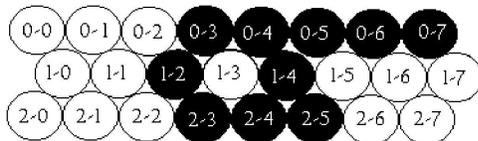


Figure 5: Two categories SOM: glaucoma, non-glaucoma.

As the theory says, this SOM acted as a topology preserving map, since input vectors of similar characteristics appear near to each other. By labelling some neurons as "doubtful", as explained above, the map of Figure 6 was obtained, where doubtful neurons are shown in grey. As expected, grey neurons are situated between glaucoma and non-glaucoma ones.

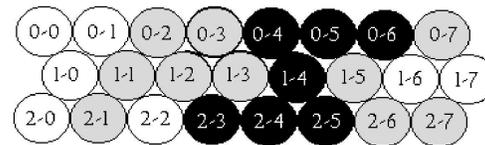


Figure 6: Previous SOM with doubtful cases shown in grey.

4.2.1 Results

The results of applying this SOM to the test set are shown in Tables 8 and 9. Now, sensitivity and specificity values are over the recommended figures (80%) and more balanced than in the expert systems. The percentage of doubtful cases fell to 35%.

4.3 Hybrid symbolic connectionist system

To integrate both approaches, expert system and SOM, a small set of rules was designed taking into account the results from Tables 7 and 9. With this integration, the final results of the hybrid system are shown in Table 10, together with the results of Tables 7 and 9 for easy comparison.

From the observation of this table, it can be concluded that:

- Sensitivity is maximum (100%).
- Specificity of the hybrid system improves to 94%.
- Diagnostic precision improves to 97%.
- The percentage of doubtful cases diminishes to 31%.

- The number of false positive cases is reduced.

The performance of this system was compared with the classification made by an expert ophthalmologist. The human expert obtained 91% of diagnostic precision, with 50% of doubtful cases.

5 Conclusion

Early diagnosis of Primary Open Angle Glaucoma is not an easy task to approach and solve. Many efforts have been made, and are being made, to design tools and techniques to help diagnose this disease in the incipient stages. In this paper we present a Visual Field Classifier System (VFCS) to aid in the early diagnosis of glaucoma. The VFCS offers a diagnostic precision of 97% with only 31% of doubtful cases. These results compare favourably with that of the human expert (91% and 50%).

This hybrid system combines symbolic and connectionist techniques. By integrating neural networks and a rule based system, we are able to make the whole system manage the knowledge that a Self-Organizing Map discovers in already classified cases, as well as the knowledge from the human expert. It also has the flexibility of a neural network to manage complex information coming from computerized perimetry. We show that the mixture of classifiers allows the creation of systems that are more precise than any of the components.

The Kohonen self organizing maps have demonstrated an ability to give good classification results in spite of the class overlappings, present in almost all medical problems, that limit other classification techniques. The modifications made to the original SOM labelling algorithms improved the map performance.

The results supplied by any Artificial Neural Network depend directly on the data, on the way they were selected, the distribution of training examples between classes, etc. In our case, the selection of patients was strictly and carefully made, by establishing a set of exclusion and inclusion criteria determined by the Institute of Applied Ophthalmobiology of the University of Valladolid (IOBA). Even though the results we got in this work are good, or even better than those of previous works, the experiments should be repeated with larger training and test samples. Given

the available samples, however, it could be said that the developed techniques and tools are adequate to aid in early glaucoma diagnosis, but there should be a much more exhaustive phase of validation and refinement. The reduction of the input parameters by dividing the visual field into zones based on the ophthalmologist's experience, resulted in an improvement of the results in all the experiments.

References

- [1] L. Alonso, A. Simón, and A. Antón. Use of Self Organizing Methods for Visual Fields Classification. In *Actas de la VII Reunión de Trabajo en Procesamiento de la Información y Control, San Juan, Argentina, 17-19 de Septiembre*, volume I, pages 171–173, 1997.
- [2] A. Antón. *Valor de las Redes Neuronales y la Discriminación Logística en el análisis de los defectos del campo visual*. Tesis doctoral, Universidad de Valladolid. (Spain), 1995.
- [3] A. Antón, J. Jordano, and J. A. Maquet. Sistema Experto de diagnóstico de glaucoma: GLAUCOM-EASY. *Arch Soc Esp Oftalmol*, 69:23–28, 1995.
- [4] H. Bebie. Computer-assisted evaluation of visual fields. *Graefe's Arch Clin Exp Ophthalmol*, 228:242–245, 1990.
- [5] E. S. de la Fuente Gallego and D. Bahillo Pérez. Clasificación de campos visuales mediante redes neuronales. Proyecto fin de carrera, EUP. Universidad de Valladolid. (Spain), 1998.
- [6] D. E. Gaasterland, F. Ederer, E. K. Sullivan, J. Caprioli, and M. N. Cyrlin. Advanced Glaucoma Intervention Study. 2. Visual Field Test Scoring and Reliability. *Ophthalmology*, 101, n^o 8:1445–1455, 1994.
- [7] M. H. Goldbaum, P. A. Sample, H. White, B. Côté, P. Raphaelian, R. D. Fechtner, and R. N. Weinreb. Interpretation of Automated Perimetry for Glaucoma by Neural Network. *Investigative Ophthalmology and Visual Science*, 35(9):3362–3373, Aug. 1994.
- [8] D. B. Henson, S. E. Spenceley, and D. R. Bull. Spatial Classification of Glaucomatous Visual Field Loss. *British Journal of Ophthalmology*, 80:526–531, 1996.

- [9] D. B. Henson, S. E. Spenceley, and D. R. Bull. Artificial neural network analysis of noisy visual field data glaucoma. *Artificial Intelligence in Medicine*, 10:99–113, 1997.
- [10] H. P. Hirsbrunner, F. Fankhauser, A. Jenni, and A. Funkhouser. Evaluating a perimetric expert system: experience with octosmart. *Graefe's Arch Clin Exp Ophthalmol*, 228:237–241, 1990.
- [11] T. Kohonen, J. Kangas, and J. Laaksomen. SOM_PAK. The Self-Organizing Map Program Package V1.2. *SOM Programming Team of the Helsinki University*, 1995.
- [12] C. E. T. Krakau. Artificial intelligence in computerized perimetry. *Doc Ophthalmol Proc Series*, 49:169–174, 1987.
- [13] M. Reyes. *SoftComputing É Aplicaciones Biomédicas. Estudio del Glaucoma en estado incipiente*. Tesis doctoral, Universidad Autónoma de Barcelona. (Spain), 1998.
- [14] A. Simón, L. Alonso, D. Bahillo, E. de la Fuente, and A. Antón. Clasificación de campos visuales mediante redes neuronales. In *Actas del I TALLER AIRENE sobre Reconocimiento de Patrones con Redes Neuronales, Antofagasta, Chile, 24-28 de Mayo*, pages 90–94, 1999.
- [15] M. A. Simón. *Desarrollo de un Sistema Experto Simbólico Conexionista de Ayuda al Diagnóstico del Glaucoma*. Tesis doctoral, Universidad de Valladolid. (Spain), 2002.
- [16] M. A. Simón and L. F. de la Fuente. *Sistema experto Glaucom-Easy v 1.0*. Universidad de Valladolid, 1993.
- [17] G. Zahlmann, M. Obermaier, C. Ritzke, and M. Scherf. Knowledge-based monitoring of glaucoma patients - a connectionist's approach. In *Proceedings MIE96, 19-22.8.96*, 1996.
- [18] G. Zahlmann, M. Scherf, and A. Wegner. A neuro-fuzzy-classifier for a knowledge-based glaucoma monitor. *Lecture Notes in Artificial Intelligence, Keravnou et al.*, 1211:273–284, 1997.
- [19] G. Zahlmann, M. Scherf, A. Wegner, M. Obermaier, and M. Mertz. Situation Assessment of Glaucoma Using a Hybrid Fuzzy Neural Network. A Differential Diagnostic-Decision Support System to Aid in Early Detection in Primary-Care Environments. *IEEE Engineering in Medicine and Biology*, pages 84–91, January/February 2000.
- [20] G. Zahlmann, M. Schubert, Obermaier, M., and G. Mann. Concept of a knowledge based monitoring system for glaucoma and diabetic retinopathy using a telemedicine approach. In *Proc. of the 13th Annual International Conf. IEEE-EMBS*. IEEE Press, 1996.