View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

ACADEMIC PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Biomedical Informatics 35 (2003) 151–159

Journal of Biomedical Informatics

www.elsevier.com/locate/yjbin

Neural network-based system for early keratoconus detection from corneal topography

P. Agostino Accardo[a,*] and Stefano Pensiero[b]

[a] *Dipartimento di Elettrotecnica, Elettronica e Informatica (DEEI), Università degli Studi di Trieste, Facoltà di Ingegneria, via Valerio 10, I-34100 Trieste, Italy*

[b] *U.O. di Oculistica, Istituto per l'Infanzia I.R.C.C.S., via dell'Istria 65/1, I-34100 Trieste, Italy*

Received 17 December 2001

Abstract

Some automatic methods have been proposed to identify keratoconus from corneal maps; among these methods, neural networks have proved to be useful. However, the identification of the early cases of this ocular disease remains a problem from both a diagnostic and a screening point of view. Another problem is whether a keratoconus screening must be performed taking into account both eyes of the same subject or each eye separately; hitherto, neural networks have only been used in the second alternative. In order to examine the differences of the two screening alternatives in terms of discriminative capability, several combinations of the number of input, hidden and output nodes and of learning rates have been examined in this study. The best results have been achieved by using as input the parameters of both eyes of the same subject and as output the three categories of clinical classification (normal, keratoconus, other alterations) for each subject, a low number of neurons in the hidden layer (lower than 10) and a learning rate of 0.1. In this case a global sensitivity of 94.1% (with a keratoconus sensitivity of 100%) in the test set as well as a global specificity of 97.6% (98.6% for keratoconus alone) have been reached.
© 2002 Elsevier Science (USA). All rights reserved.

Keywords: Neural network; Corneal topography; Early keratoconus; Computer-aided diagnosis

1. Introduction

Keratoconus (KC) is a frequent non-inflammatory corneal ectasia characterised by a localised conical protrusion with stromal thinning. This thinning appears to result from the loss of structural components in the cornea, probably of collagen fibrils [5,16]. KC becomes manifest at puberty and can progress either slowly, stabilising over the course of approximately 10 years, or relatively rapidly, requiring keratoplasty in 10–25% of cases [23]. It is a bilateral corneal pathology, but it is asymmetrical, presenting a different severity in the two eyes. A monolateral KC is present in 4% of patients.

Diagnosis of KC is usually performed when some pathognomonic signs are visible in the cornea using a slit-lamp. They are: vertical stress lines deep in the affected stroma (Vogt's striae); distorted lower lid in down-gaze (Munson's sign); sub-epithelial scarring of the cone; increased visibility of the corneal nerves; Fleischer's iron ring at the base of the cone [6]. In the early cases (early KC), however, these clinical signs are not present, if we exclude a slight inconstant astigmatism or a slight distortion of keratometry mires.

Keratometry is conventionally used as an indicator of disease severity, distinguishing among mild (<45 diopters, D), moderate (45–52 D) and advanced (>52 D) keratoconi [6]. The CLEK Survey [23] indicated that a majority of keratoconic eyes do not show individual pathognomonic signs of KC (they are early cases) and that higher keratometric readings are associated with a higher prevalence of one or more of the slit-lamp signs. In fact many ophthalmologists consider sufficient for diagnosis the presence of other factors, the most important of which is a documented increase in corneal curvature over time [6,9,16].

Given the spread of refractive surgery especially among young people (where KC could be present in the

* Corresponding author. Fax: +39-040-558-3460.
E-mail address: accardo@deei.univ.trieste.it (P.A. Accardo).

1532-0464/02/$ - see front matter © 2002 Elsevier Science (USA). All rights reserved.
PII: S1532-0464(02)00513-0

early stage and only in one eye), it is necessary to have methodologies capable of identifying or screening early KC, which has hitherto been the main contraindication to the execution of this type of surgery.

Computer-assisted videokeratoscopes, which generate colour-coded maps and topographic indices, are currently the most sensitive and sophisticated devices for confirming the diagnosis of KC and for following in time its evolution [16]. Ultrasonic pachymetry may be useful to confirm corneal thinning in patients with suspected KC, but it is less used because of the large range of central and paracentral variation of pachymetry readings in the normal population [16].

Corneal topography is therefore indispensable to detect early KC, but, in some cases, doubts remain about the correct diagnosis, so much so that some authors [9,17,20,22] define Keratoconus Suspects (KCS) as the presence of an abnormal map without clinical signs. In fact, topographic maps of eyes with KC display a variety of different patterns [22] that may be confused with other corneal irregularities, defined as corneal warpage [17,20]. The term KCS should be used if only one map for each eye is available. In fact, by following up these suspect cases it is often possible to identify the true keratoconi, if a corneal power progression is verified [23]. Because of similar topographic patterns have been noticed in clinically normal family members of keratoconus patients and in the clinically normal fellow eye of patients with clinically unilateral KC [12,13], in addition to the corneal power progression, other criteria (family history and presence of a KC in the other eye) can be considered to produce, with a good probability, a correct diagnosis of KC in absence of slit-lamp findings.

In order to help the clinician in cases of difficult map interpretation and for a screening purpose, numerical methods have been developed to distinguish the keratoconus cornea from the normal cornea [2,6], in particular the Rabinowitz–McDonnell test [11] is the most widely used. It employs the following indices obtained from a sagittal topography: the central corneal power (CCP), the I–S values (inferior–superior dioptric asymmetry), and the difference between right and left CCP (ΔCCP). Later on, Rabinowitz [15] added the SRAX (skewed radial axes) index that quantifies the skewing of the steepest radial axes above and below the horizontal meridian.

Many other quantitative parameters, calculated from corneal maps, have been proposed to better characterise corneal patterns [7,21]: simulated keratometry readings (SimK1 and SimK2), surface asymmetry index (SAI), surface regularity index (SRI), differential sector index (DSI), opposite sector index (OSI), centre/surround index (CSI), analysed area (AA).

More recently, Maeda and colleagues [7,8] have proposed the use of expert systems. At first, they used the discriminant analysis [7], then a neural network [8]. Afterwards, Smolek and Klyce [18] compared five videokeratographic methods for keratoconus detection demonstrating the validity of the neural network approach. The ability to respond with the same terminology used by clinicians, without requiring them to learn a new terminology, represents a particular advantage of this approach.

During a topographic screening, all the maps with a keratoconus-like pattern should be classified as KCS. But we believe it should also be possible to identify cases of early KC among the KCS maps using suitable methods of map analysis. To correctly perform such a type of study, the reference maps should be KCS maps successively evolved to KC with clinical signs. In practice, this is a very difficult task to perform, as it requires many years of follow up. Therefore, we used the previously described criteria to obtain, with a high probability, early KC maps.

Another unsolved problem is whether keratoconus screening must be performed taking into account both eyes of the same subject (as proposed by Rabinowitz) or each eye separately (as proposed by Maeda and Klyce); hitherto, neural networks have only been used in the second alternative. In 1989 Rabinowitz and McDonnell [11] used the ΔCCP as an important parameter to identify the keratoconus, following the clinical observation of the asymmetric presentation (different stage in the two eyes) of this pathology.

In this paper we describe a new expert system based on a neural network which utilises parameters that differ from those used by Maeda et al. [8] to detect the KC corneal patterns. At first, the parameters evaluated from each eye were considered separately, then we used the data taken from both eyes. A comparison among the network architectures that were tested was carried out.

2. Materials and methods

2.1. Subjects

A study on KC progression must take into account the repeatability of any corneal curvature measure, particularly of corneal topography, which however has not been determined for KC specifically [23]. That is why, we followed the suggestions of Di Lorenzo [1] who observed a better repeatability for sagittal cone apex powers <53 D. This repeatability is extremely important if a corneal power increase is to be detected.

We divided our maps into normal maps (N), keratoconus (KC) and other non-keratoconus conditions (O). Only KC of mild and moderate severity was considered, with a sagittal cone apex power <53 D, where no clinical sign was present (early KC) or only the Vogt's striae were detected (mild or moderate KC, 12 eyes).

In order to obtain a correct classification of our early KC cases, we took into consideration all the maps that

we had previously classified as KCS, and we defined as KC only those in which one of the following criteria was satisfied: (1) the pathology was already present (Vogt's striae) in the other eye with the same topographic pattern, (2) a family history of KC was present, (3) the apex progression of the corneal protrusion was >1 D after 1 or 2 years of follow up investigation. Some KCS were included in the O group if a contact lens warpage was demonstrated by the disappearance of the corneal warpage after 1–3 months of corneal lens abstention. All the other KCS-type maps were excluded.

As far as the first criterion is concerned, we also included in our study the 12 eyes with clinical KC (with the highest corneal powers of our series), used to identify maps of early KC in the second eye.

Thus, 396 corneal topographic maps, obtained with a videokeratoscope (EyeSys), were selected from the cases recorded over a three-year period at the Ophthalmological Unit of the Children's Hospital 'Burlo Garofolo' of Trieste. All the KCS cases had a follow up: if the KCS case became Normal in both eyes for the interruption of contact lens wearing, the first registered map would be classified as contact lens corneal warpage and it would be included in an 'Other group' (O, 66 maps: 33 subjects, 9 males, and 24 females); if, instead, the KCS could be classified as early KC, following the above-mentioned criteria, the first registered map would be included in a 'KC group' (KC, 120 maps: 60 subjects, 35 males, and 25 females, 50 both eyes with KC and 10 only one eye with KC); all the other KCS maps were discarded because it was impossible to classify them differently from KCS. Moreover, we considered the first 110 maps belonging to Normal subjects in both eyes that were included in a Normal group (N, 110 maps: 55 subjects, 28 males, and 27 females) and another 100 maps (50 subjects, 18 males, and 32 females) of subjects with various bilateral non-keratoconus conditions, especially congenital astigmatism, which were included in the Other group together with the previous contact lens corneal warpage cases.

The mean ages in the three groups were 22 ± 13 (N), 35 ± 12 (KC) and 21 ± 14 (O) years. The mean age is higher in the KC group because we considered all the cases recorded during 3 years where a keratoconus-like pattern was present, in order to identify a sufficient number of early cases. In particular we considered also cases of all ages where the corneal map was recorded for an increment of the astigmatic correction with glasses. Instead, the other two groups are constituted by the first recorded patients for a pre-surgical KC screening. However we underline that changes in corneal map are age-linked only if a corneal pathology is present.

2.2. Neural networks: architecture

The neural network was constructed with three layers and the backpropagation method was used in the training process. To understand which network architecture provided the best KC identification, we examined six different combinations of neuron numbers both in the input and output layers; for each of these situations we considered a different number of neurons in the hidden layer, three possible learning rates and a suitable number of epochs.

In particular (Table 1), the number of input neurons could be 9 (corresponding to the topographic indices of each eye listed below), 18 (corresponding to the same parameters when considering both eyes of the same subject simultaneously) or 10 or 19, by adding to the previous indices the absolute difference between the central corneal powers of the two eyes (ΔCCP). The output layer could consist of three neurons, one for each topographic category (N, KC, and O) or of six neurons, one for each topographic category of the two eyes considered simultaneously.

Thus, as regards the input, in the first situation (Table 1) each eye was considered by itself (monocular approach), while in the remaining approaches (situations II-VI) at least one information from both eyes of the same subject was considered (binocular approach).

As regards the output, the goal was to detect the presence of KC in a specific eye (eye screener—ES) or to verify the presence of KC at least in one eye of a subject (subject screener—SS). Since the clinical aim is the identification of patients affected by KC, subject classifications were also extracted from the results of the ES (situations I, II, V, and VI) and compared to the SS of situations III and IV.

The parameters used in the input layer were: simulated K1 (SimK1), simulated K2 (SimK2), asphericity (Q), and corneal uniformity (CU) indices, supplied by the EyeSys Holladay diagnostic summary, and central corneal power (CCP), inferior–superior asymmetry (I-S), differential sector index (DSI), opposite sector index (OSI) and centre/surround index (CSI) calculated from

Table 1
Possible combination of input and output neuron numbers considered in this study

Situation number	Number of input neurons	Number of output neurons	Input approach	Output approach
I	9	3	Monocular	Eye screener
II	10	3	Binocular	Eye screener
III	18	3	Binocular	Subject screener
IV	19	3	Binocular	Subject screener
V	18	6	Binocular	Eye screener
VI	19	6	Binocular	Eye screener

the numeric maps. These nine parameters were chosen from a larger number of indices taken from the literature or supplied by the videokeratoscope after a comparison, by means of the Student t test, of the mean values detected in the N and KC groups.

The sum of the weighted inputs to a given neuron of the hidden layer was transformed through a non-linear transfer function (sigmoidal curve) to modulate the response characteristics. The non-linear transfer function (sigmoidal curve) for the neurons in the output layer was normalised to provide an output ranging from 0 to 1. As a result, a value of 0 for a particular output neuron indicated a complete lack of response to the corresponding category, whereas an output strength approaching 1 indicated maximum response.

The network weights and biases were initially set to suitable values so that the active regions of the layer neurons were more or less evenly distributed over the input space (Nguyen–Widrow initialisation method [10]).

With the aim of gaining the best accuracy and optimising the network, the number of input and output layers was changed as previously described and also the number of hidden neurons was varied between 4 and 120 (from 4 to 20 with unitary step and from 25 to 120 with a step of 5, for a total of 37 steps). The maximum value was selected considering the number of the training cases. Finally, the learning rate was varied. The values considered were 0.1, 0.01, and 0.001. Thus, for each of the six different input/output situations, $37 \times 3 = 111$ possible combinations were examined.

2.3. Neural networks: training

The learning process was performed by using the maps of the right and left eyes, either separately or together, of 25 normal subjects, 30 subjects with KC (25 both eyes, 5 only one eye) and 40 subjects with other alterations (55 normal maps, 55 KC maps, and 80 other alterations maps when the eyes were considered separately). In each group the maps were randomly selected.

The training of the neural network was performed by using the backpropagation method, which was considered completed when either the value of the training performance, measured as the mean square error of output approximation (mse), went below 0.005 or when a minimum gradient (fixed to 1E-7) in the performance curve was reached. The mse was calculated as the mean sum of squares of the network errors:

$$\text{mse} = \sum_{i=1}^{N} (E_i - O_i)^2 / N \tag{1}$$

being N the output neuron number (3 or 6), E_i and O_i respectively, the expected and the observed outputs of the ith neuron.

2.4. Neural networks: test

The efficacy of the neural network classifier was evaluated with a test set of 30 normal subjects, 30 subjects with KC (25 both eyes, 5 only one eye) and 43 subjects with other alterations. A winner-takes-all threshold criterion of 0.5 was used to determine the classification made by the neural network, thus each map or each pair of maps was associated to the category with a value which exceeded the threshold. If no category satisfied this requirement, it was classified as *unknown*.

The results for each category were described in terms of sensitivity (true positive/[true positive + false negative]), specificity (true negative/[true negative + false positive]), and accuracy ([true positive + true negative]/total number of maps). These statistical measures were computed both separately for each category and globally.

We computed the Akaike Information Criterion (AIC) score to determine the best network in order to reward a network with low mean square error (mse) but to penalize networks with a large number of weights. Since the contribution of the mse in the test set was negligible at all (or became negligible already with eight neurons in the situations I and II) in respect to the contribution of the number of free parameters that linearly rises with the increasing of hidden neurons number, the AIC score only states that, when the global sensitivity values are equivalent, the best network is that has got the minimum number of hidden neurons. Then, in each situation, the network configuration that produced the highest global sensitivity was considered. These six *best* neural networks were used to compare the six situations. To this aim, the simple bootstrapping method (with $N = 300$ resampling with replacement of the test data) was applied [4,19] in order to obtain, for each considered network, a distribution of the bootstrap estimates so as to determine possible significant differences in sensitivity, sensibility and accuracy, among the six situations. In this work the bootstrap estimate without any modification correction was used because the contribution of the training set error (considered in the so-called 632 bootstrap) was about constant and less than 1%.[1]

[1] Efron [3] states that the bootstrap estimate we used (Rb) could result upward biased, so he proposed to use the bootstrap 632 estimate (R632) where the weighted contribution of the training set error E was introduced to correct the upward bias in Rb as $\text{R632} = .368 * E + .632 * \text{Rb}$.

Since in our NN the E values for different situations and number of hidden neurons was very low and about constant (from 0.03 to 0.01 in situations I and II, less than 0.01 in the other situations), we used the simple Rb neglecting the correction that does not affect the final results.

The neural network model developed in this study was implemented on a PC using the MATLAB v.5 software package.

3. Results

3.1. Parameters mean values

Table 2 shows that in KC the mean values of all the used parameters are very different from those obtained in the remaining two groups; the difference between the KC and N groups and between the KC and O groups is always significant ($p < 0.00001$ at the Student's t test) for all the parameters. The difference between the N and O groups is also significant for all the parameters ($p < 0.05$ for CCP, ΔCCP and I–S; $p < 0.01$ for all the other parameters) although some variables presented a high standard deviation.

3.2. Neural network results

In situations III–VI shown in Table 1, the tolerance threshold in most training cases was reached after a few hundreds of epochs (see the example in Fig. 1a), while in situations I and II the minimum gradient of the performance curve was often reached without satisfying the preset performance (see in Fig. 1b an example of the situation I).

With the aim of identifying for each situation the network that could produce the highest global sensitivity (GS), the number of neurons in the hidden layer and the three learning rates (0.1, 0.01, and 0.001) were modified during the training phases. The results are reported in Fig. 2a–f.

The behaviour of the GS was similar both in the training and test sets for each couple (I–II, III–IV, V–VI) of situations, differing only for the absence/presence of the ΔCCP parameter.

Table 2
Mean values and standard deviations of the 10 considered indices in the three groups

	Normal ($n = 120$ maps)	Keratoconus ($n = 110$ maps)	Other alterations ($n = 166$ maps)
CCP	43.47 ± 1.47	47.13 ± 5.12	43.96 ± 1.96
ΔCCP	0.28 ± 0.22	3.48 ± 3.70	0.58 ± 1.67
I-S	-0.12 ± 0.54	4.22 ± 3.36	0.06 ± 0.89
DSI	0.88 ± 0.41	6.99 ± 3.98	2.47 ± 1.29
OSI	0.66 ± 0.42	6.39 ± 3.79	1.10 ± 0.95
CSI	0.23 ± 0.28	2.13 ± 3.42	0.44 ± 0.69
SimK1	43.71 ± 1.47	47.14 ± 4.05	45.09 ± 2.06
SimK2	43.20 ± 1.45	44.89 ± 3.34	42.64 ± 1.85
Q	-0.08 ± 0.15	-0.63 ± 0.87	-0.14 ± 0.25
CU Index	100	84.00 ± 16.66	98.86 ± 6.73

CCP, central corneal power; ΔCCP, absolute value of the difference in CCP between the two eyes of a given patient; I–S, inferior–superior asymmetry; DSI, differential sector index; OSI, opposite sector index; CSI, centre/surround index; SimK1, simulated K1; SimK2, simulated K2; Q, corneal asphericity; CU index, corneal uniformity index.

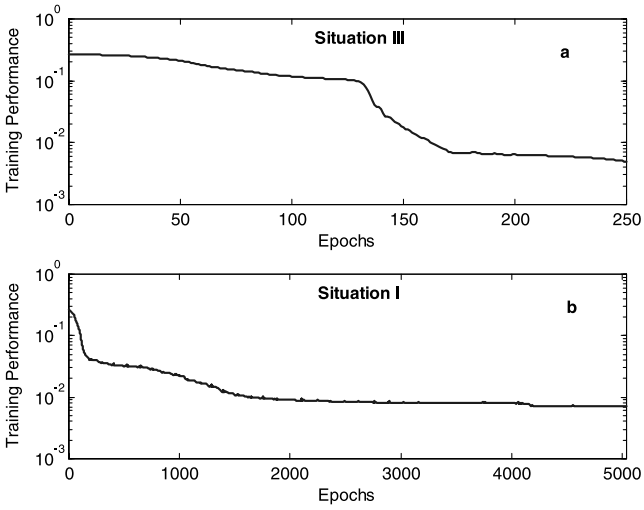

Fig. 1. Examples of training performance behaviour as a function of the number of epochs: (a) situation III, 7 hidden neurons, learning rate = 0.1, global sensitivity = 94.2%, required tolerance error met; (b) situation I, 8 hidden neurons, learning rate = 0.01, global sensitivity = 87.2%, minimum gradient of the performance curve met.

In situations I and II, the performance never reached 100% (ranging between 97% and 99%) during the training, while in the test set (Fig. 2a–b) the maximum value of GS was obtained with a low number (about 20) of hidden neurons. The highest sensitivity values were reached in situation II (with ΔCCP).

In situations III and IV, a GS of 100% was always reached in the training set; in the test set (Fig. 2c–d), after maximum GS values for 4–8 hidden neurons, a quick decrease was present.

In situations V and VI, the GS during the training set almost always reached 100% (99% at minimum), while in the test (Fig. 2e–f), the GS was independent of the number of hidden neurons showing intermediate values between situations I–II and III–IV, respectively.

In all the cases, the comparison of the results obtained by using the three learning rates (0.1, 0.01, and 0.001) showed that even if a significant difference was not present, the 0.1 learning rate appeared to be the most convenient, being the value producing the quickest convergence.

Table 3 reports the combinations of the hidden neuron numbers and learning rates that produced the highest GS in the six situations. Situations III and IV (18 or 19 inputs, 3 output categories) presented the highest GS values. Table 4 shows the results obtained for each neural network of Table 3 when the complete test set was used.

Tables 5 and 6 show the 5th and the 95th percentiles together with the mean values of the global sensitivity, specificity, and accuracy, obtained from the 300 test set

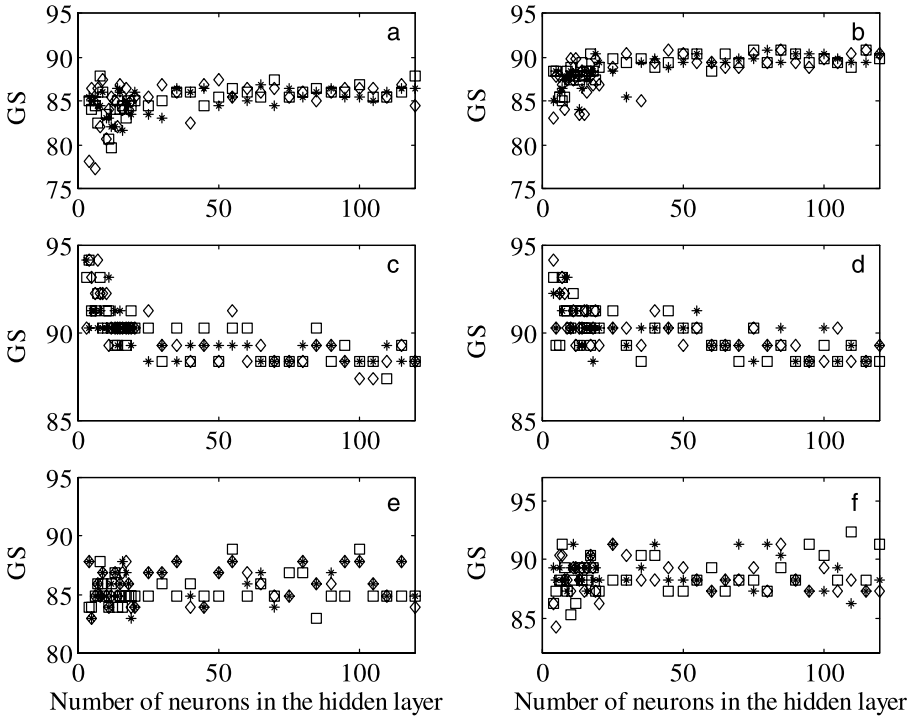

Fig. 2. Global sensitivity (GS) vs. number of neurons in the hidden layer in the six considered situations for three different learning rates. Panels (a)–(f) progressively correspond to situations from I to VI. Diamond: learning rate = 0.1; square: learning rate = 0.01; star: learning rate = 0.001.

Table 3
Hidden neuron numbers and learning rates that produce the highest global sensitivities in the six considered situations

Situation	Number of hidden neurons	Learning rate	Global sensitivity
I	8	0.01	87.86
II	45	0.1	90.78
III	7	0.1	94.17
IV	4	0.1	94.17
V	55	0.01	88.83
VI	7	0.01	92.23

resampling combinations, for both the ES (situations I, II, V, and VI) and the SS, respectively. However, from the results of ES it is possible to obtain a SS considering the classification of the two eyes jointly. Thus, with the aim to identify the situation that bring to the best SS, in Table 6, we reported together with the two SS conditions (situations III and IV), the Subject Screeners built by using the data of situations I, II, V, and VI of Table 5. In these cases we classified as KC a subject with at least one eye with keratoconus and as UK a subject whose eyes were classified in different categories by the ES (or at least one eye was classified as *unknown*).

In the ES there were no significant differences between the four situations for each statistic parameter (sensitivity, specificity and accuracy). In the SS the differences were significant only between situation I and situations III, IV, and VI for sensitivity and between situations II and V for specificity.

4. Discussion

Corneal topography is an objective and quantitative support to the clinical diagnosis of KC. In fact, many quantitative parameters can be evaluated from the maps and also be used in some automatic methods, like the neural network, to identify the patterns of corneal pathologies.

The ability to screen automatically KC corneal topographic patterns is very important if large populations are to be studied, in particular in the propaedeutic examination for refractive surgery. For a screening purpose that today has a practical application (refractive surgery and genetic studies in families), the most important information is the presence or absence of an early KC.

Maeda et al. [8] and Smolek and Klyce [18] already demonstrated the quality and value of the neural network approach in identifying the topographic patterns of KC; our work agrees with their results. Unlike other authors, our work considered only cases of early KC and compared different network architectures to evaluate the results of both ES and SS. We chose nine parameters with a remarkable difference in mean values among the three categories (Table 2), we tested the six situations shown in Table 1 and we used different numbers of hidden neurons and learning rates that allowed us to reach the preset goal (Tables 5 and 6).

While normally each eye is considered as an independent case (ES based on a monocular approach), we

Table 4
Classification results of the six neural networks described in Table 3, by using the complete test set

Actual category	Estimated category				Total
	N	O	KC	UK	
Situation I					
N	55	8	1	1	65
O	8	77	1	0	86
KC	1	5	49	0	55
Total	64	90	51	1	206
Situation II					
N	54	10	1	0	65
O	5	80	0	1	86
KC	1	1	53	0	55
Total	60	91	54	1	206
Situation III					
N	30	0	0	0	30
O	4	37	1	1	43
KC	0	0	30	0	30
Total	34	37	31	1	103
Situation IV					
N	30	0	0	0	30
O	4	37	2	0	43
KC	0	0	30	0	30
Total	34	37	32	0	103
Situation V					
N	60	3	2	0	65
O	10	75	1	0	86
KC	0	5	48	2	55
Total	70	83	51	2	206
Situation VI					
N	63	1	0	1	65
O	10	75	1	0	86
KC	0	3	52	0	55
Total	73	79	53	1	206

N, normal; KC, keratoconus; O, other alterations; UK, unknown. The numbers represent eyes (situations I, II, V, and VI) or subjects (situations III and IV) classified by NN compared to the actual (clinical) categories.

Table 5
Mean values and 5th and 95th percentiles of the global sensitivity, specificity, and accuracy distributions, obtained from the 300 test set resampling combinations, for the eye screening (situations I, II, V, and VI)

Situation	Mean (%)	5th percentile	95th percentile
Global sensitivity			
I	85.4	81.5	89.1
II	88.6	84.3	91.9
V	86.1	81.2	89.7
VI	91.4	87.4	94.9
Global specificity			
I	93.7	92.0	95.5
II	96.1	94.6	97.4
V	94.2	91.7	95.9
VI	96.2	94.1	97.6
Global accuracy			
I	90.9	88.6	93.1
II	93.6	91.2	95.3
V	91.5	88.1	94.0
VI	94.6	92.1	96.7

Table 6
Mean values and 5th and 95th percentiles of the global sensitivity, specificity, and accuracy distributions, obtained from the 300 test set resampling combinations, for the subject screening (all situations)

Situation	Mean (%)	5th percentile	95th percentile	Mean (%)
Global sensitivity				KC sensitivity
I	81.7	75.8	86.9	96.6
II	87.4	81.4	91.9	100
III	94.1	90.4	97.8	100
IV	93.2	89.4	95.7	96.7
V	88.4	82.3	92.5	93.3
VI	92.3	88.4	95.7	100
Global specificity				KC specificity
I	98.6	97.0	99.9	98.6
II	99.0	98.1	99.9	98.6
III	97.6	95.7	99.4	98.6
IV	97.6	95.7	98.9	100
V	96.1	93.7	97.8	100
VI	97.6	95.6	98.9	98.6
Global accuracy				KC accuracy
I	92.9	90.6	95.0	98.0
II	95.1	92.9	97.0	99.0
III	96.4	93.9	98.6	99.0
IV	96.1	93.9	98.0	99.0
V	93.5	90.7	95.9	98.0
VI	95.8	93.3	98.1	98.0

The data of situations I, II, V, and VI were calculated classifying as KC a subject with at least one eye with keratoconus and as UK a subject whose eyes had been previously classified in different categories or who had at least one eye that was classified as unknown. In the last column, the KC sensitivity, specificity, and accuracy mean values are also reported.

also matched the information obtained from both eyes of the same subject (ES and SS based on a binocular approach). Moreover, we also compared the performances of the ES and the SS, as, for clinical purposes, the identification of the patient is more important than the identification of the sick eye. In fact, a patient suffering from a monolateral pathology must be periodically checked to assess the eventual incoming of KC in his/her sound eye and moreover, KC represents the major contraindication to refractive surgery in both eyes even if one is (still) sound.

The introduction of the ΔCCP parameter slightly improves the neural network performance in situations II and VI with respect to I and V, both in ES and SS (Tables 5 and 6 and Fig. 2b–f compared to 2a–e) while in situation IV vs. III this parameter is ineffective (Table 6 and Fig. 2d vs. 2c). However, such differences are generally not significant, although situation I has a GS mean value significantly lower than that of situations III, IV, and VI. It is evident that the monocular approach, normally used in clinical procedures, is less

effective than a binocular approach for a correct identification of the cases of each group (N, O, and KC). Moreover, the binocular approach generally appears to be slightly better than the monocular one for the detection of KC sensitivity (Table 6).

In any case, the best results in terms of global sensitivity and accuracy as well as KC screening power, are obtained with situation III (slightly better than in situation IV).

However, the subject screener presented in situation II (Table 6) showed the best performance in terms of mean global specificity, even if this higher value was gained with a lower global sensitivity. This was mainly due to the classification of subjects with other alterations. In fact, when the SS was evaluated in situations I and II, with respect to the ES, there was a cancellation of the false positives (O cases no longer classified as N) and a corresponding increment of the false negatives (O cases classified as UK).

The KC specificity value (linked the false positive cases) in the SS (Table 6) was high in all the situations, standing at 100% in situations IV and V. The KC sensitivity value (linked the false negative cases) in the SS, was low in situation V. The input/output configuration used as well as the low number of monolateral KC cases considered during the NN training could explain this low value of sensitivity. On the contrary, in the other SS situations, with the exception of situations I and IV, all the KC subjects were correctly identified (sensitivity of 100%), pointing out an optimal discriminative capability, for KC subjects, of the various architectures used, in spite of the reduced presence of KC in the NN training.

Therefore, one of the main results of this study is that the utilisation of the nine parameters from both eyes (binocular approach) and of the three output categories improves the discriminative capability of the neural network: in situation III the global accuracy was 96.4% and the KC sensitivity was 100% (Table 6). Moreover, if the O group is divided into subgroups, this system would be suitable for detecting other specific patterns in corneal topography, choosing the appropriate quantitative indices. Finally, the low number of hidden neurons, which is necessary to obtain the best result, produces a network that is easy to implement and quick in both the training and test phases.

Maeda et al. [8], who reported global results similar to ours, divided their KC maps into two groups, with only eight maps of mild KC.

Our results are slightly different with respect to those obtained by Smolek and Klyce [18], who reported 100% right classification using their neural network. It is however important to remember that there are substantial differences in the choice of cases. Their maps were divided into three groups (KCS, KC, and O): the KCS group was constituted by six maps in the training and test sets. The KC group was separated in three subgroups: mild (up to 55 D; 11 maps), medium and severe (a total of 22 maps). In any case the maximum total number of early maps was 17, but in the classical classification [6] the mild and moderate cases present a corneal power lower than 52 D.

In our opinion for a KC screening, the training set cannot be performed using KCS maps, where in many cases the use of a contact lens is responsible for the KC-like pattern. Rabinowitz [14] took into account only KC with clinical signs to obtain the reference values (he did not use neural networks), and he required that screening be used in a noncontact lens wearing population.

It is evident that the aim of Smolek and Klyce [18], using a KCS group as the input and output of their neural network, was to obtain an indication of the cases of the screened population to be followed up. This is an important goal but, in the case of a KC screening, an increase of both the false positives (due to the large number of persons wearing contact lenses) and (since these persons produce an increment of contact lens corneal warpages) of the screening costs is expected. We believe that it is more correct to use an expert system to take KC out of the KCS: this is what we have tried to do, using a neural network to differentiate KC and contact lens corneal warpage, inserted in the O group.

We used 110 KC maps (55 in the test set), especially of early cases, recorded on 60 patients, of whom 10 presented a monolateral pathology.

In our work we decided to identify three categories (N, KC, and O), with a category of normal maps (with the rule astigmatism < 1 D) because, being KC an asymmetrical (rarely monolateral) pathology, at its beginning one eye is affected while the other is still sound. The classifier must consequently know the characteristics of normality.

It is clinically easy to identify KC with clinical signs, but it is also normally identified by a screener using only corneal maps. A KC screening is usually performed in young patients; therefore it is important to consider especially early cases and try to obtain a diagnosis of KC as soon as possible.

For this reason, a screener classification in which a KCS group appears should not be accepted. We believe that the various types of maps constituting KCS could be separated directly by a well-trained screener, with few errors. Furthermore, a good KC screening has to identify the largest number of cases, with the minimum possible false positives. The capability of our screener to differentiate at least KC and contact lens warpage is another important result of this study.

Finally, even if in the situations II and V the high number of hidden neurons (Table 3) could suggest an overfitting of data, we could accept the second highest values for GS (situation II: GS = 89.81, 10 neurons, learning rate = 0.1; situation V: GS = 87.84, 8 neurons,

learning rate = 0.01) without changing of the final result of the work (that the simultaneous utilisation of topographic parameters from both eyes improves the discriminative capability of the neural network).

5. Conclusion

This study expands the value of the neural network approach to the automatic screening of early keratoconus and states that the simultaneous utilisation of topographic parameters from both eyes improves the discriminative capability of the neural network.

Acknowledgments

This work was partially supported by the University of Trieste (MURST60%) and by IRCCS "Burlo Garofolo" of Trieste.

References

[1] Di Lorenzo G, Traverso CE, Zingirian M. L'esame della topografia corneale nel cheratocono. In: Atlante di topografia corneale; a cura di Paolo Vinciguerra; Fogliazza Editore, Milano; 1995. p. 224–42.
[2] Dingeldein SA, Klyce SD, Wilson SE. Quantitative descriptors of corneal shape derived from computer-assisted analysis of photokeratographs. J Refract Corneal Surg 1989;5:372–8.
[3] Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Assoc 1983;78:316–31.
[4] Efron B, Tibshirani R. An introduction to the bootstrap. London: Chapman and Hall; 1993.
[5] Kenyon KR, Hersh PS, Starck T, Fogle JA. Corneal dysgeneses, dystrophies and degenerations. In: Tasman W, Jager EA, editors. Duane's clinical ophthalmology, vol. 4. Philadelphia, PA: Lippincott Williams & Wilkins; 1998 [Chapter 16].
[6] Krachmer JH, Feder RS, Belin MN. Keratoconus and related non-inflammatory corneal thinning disorders. Surv Ophthalmol 1984;28:4, 293–322.
[7] Maeda N, Klyce SD, Smolek MK, Thompson HW. Automated keratoconus screening with corneal topography analysis. Invest Ophthalmol Vis Sci 1994;35:2749–57.
[8] Maeda N, Klyce SD, Smolek MK. Comparison of methods for detecting keratoconus using videokeratography. Arch Ophthalmol 1995;113:870–4.
[9] Maguire LJ, Lowry LC. Identifying progression of subclinical keratoconus by serial topography analysis. Am J Ophthalmol 1991;112:41–5.
[10] Nguyen DD, Lee JS. New LMS-based algorithm for rapid adaptive classification in dynamic environments. Neural Networks 1989;2(3):215–28.
[11] Rabinowitz YS, McDonnell PJ. Computer-assisted corneal topography in keratoconus. Refract Corneal Surg 1989;5:400–8.
[12] Rabinowitz YS, Garbus J, McDonnell PJ. Computer-assisted corneal topography in family members of keratoconus. Arch Ophthalmol 1990;108:365–71.
[13] Rabinowitz YS, Nesburn AB, McDonnell PJ. Videokeratography of the fellow eye in unilateral keratoconus. Ophthalmology 1993;100:181–6.
[14] Rabinowitz YS. Videokeratographic indices to aid in screening for keratoconus. J Refract Surg 1995;11:371–9.
[15] Rabinowitz YS, Yang H, Brickman Y, Akkina J, Riley C, Rotter JI, Elashoff J. Videokeratography database of normal human corneas. Br J Ophthalmol 1996;80(7):610–6.
[16] Rabinowitz YS. Keratoconus. Surv Ophthalmol 1998;42(4): 297–319.
[17] Smolek MK, Klyce SD, Maeda N. Keratoconus and contact lens-induced corneal warpage analysis using the keratomorphic diagram. Invest Ophthalmol Vis Sci 1994;35:4192–204.
[18] Smolek MK, Klyce SD. Current keratoconus detection methods compared with a neural network approach. Invest Ophthalmol Vis Sci 1997;38:2290–9.
[19] Tibshirani R. A comparison of some error estimates for neural network models. Neural Comput 1996;8:152–63.
[20] Wilson SE, Lin DTC, Klyce SD, Reidy JJ, Insler MS. Topographic changes in contact lens-induced corneal warpage. Ophthalmology 1990;97:734–44.
[21] Wilson SE, Klyce SD. Quantitative descriptors of corneal topography. A clinical study. Arch Ophthalmol 1991;109:349–53.
[22] Wilson SE, Lin DTC, Klyce SD. Corneal topography of keratoconus. Cornea 1991;10:2–8.
[23] Zadnik K, Barr JT, Gordon MO, Edrington TB. Biomicroscopic signs and disease severity in keratoconus. Cornea 1996;15(2):139–46.