ORIGINAL ARTICLES—Clinical science

Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool

G G Gardner, D Keating, T H Williamson, A T Elliott

Abstract

Aims—To determine if neural networks can detect diabetic features in fundus images and compare the network against an ophthalmologist screening a set of fundus images.

Methods—147 diabetic and 32 normal images were captured from a fundus camera, stored on computer, and analysed using a back propagation neural network. The network was trained to recognise features in the retinal image. The effects of digital filtering techniques and different network variables were assessed. 200 diabetic and 101 normal images were then randomised and used to evaluate the network's performance for the detection of diabetic retinopathy against an ophthalmologist.

Results—Detection rates for the recognition of vessels, exudates, and haemorrhages were 91.7%, 93.1%, and 73.8% respectively. When compared with the results of the ophthalmologist, the network achieved a sensitivity of 88.4% and a specificity of 83.5% for the detection of diabetic retinopathy.

Conclusions—Detection of vessels, exudates, and haemorrhages was possible, with success rates dependent upon preprocessing and the number of images used in training. When compared with the ophthalmologist, the network achieved good accuracy for the detection of diabetic retinopathy. The system could be used as an aid to the screening of diabetic patients for retinopathy.

(*Br J Ophthalmol* 1996;80:940–944)

Department of Clinical Physics and Bio-Engineering, Tennent Institute of Ophthalmology, Glasgow
G G Gardner
D Keating
T H Williamson
A T Elliott

Correspondence to:
Dr Glen G Gardner, Department of Clinical Physics and Bio-Engineering, Tennent Institute of Ophthalmology, Western Infirmary, 38 Church Street, Glasgow G11 6NT.

Accepted for publication 28 June 1996

The screening of diabetic patients for the development of diabetic retinopathy potentially reduces the risk of blindness in these patients by 50%[1] and in the USA would provide considerable cost savings to public health systems.[2] It has been estimated that 30 000 individuals per million total population need to be examined to implement such a programme.[3] This is beyond the scope of currently existing ophthalmology departments and would produce a heavy clinical burden if left to diabetic physicians.
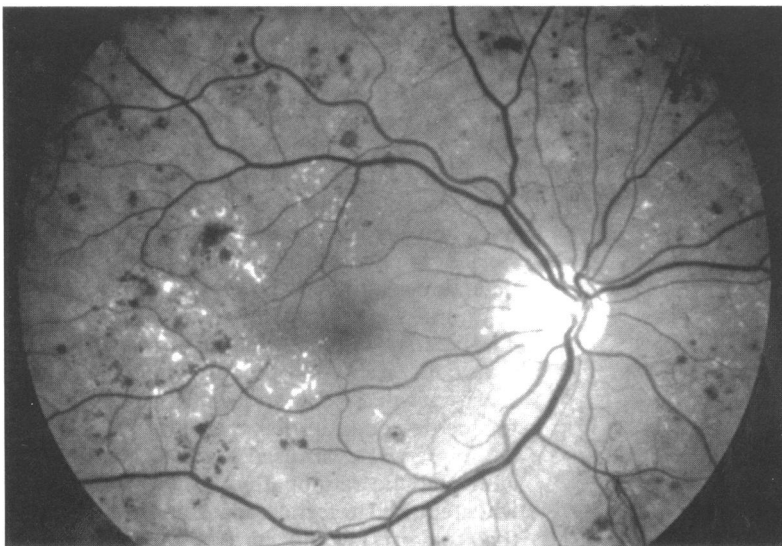
Photographing the fundus of such patients with later assessment of the photographs has been tried with some success[4-7] but still requires the relatively expensive time of the specialist ophthalmologist to provide a classification of the retinopathy. This may delay referral of the patient for further examination. A wholly automated approach involving fundus image analysis by computer could improve the efficiency of the assessment of the image by providing an immediate classification of the fundus of the patient at the time of acquisition of the image. The operator of the camera could be provided with an indication of whether to refer the patient to an ophthalmologist.

Conventional digital image analysis is hampered in the examination of images of the retina because of the inherent variability of the fundi between individuals. Although this has been applied to the detection of specific features such as microaneurysms, exudates, and oedema in diabetic retinopathy[8-12] with some success, an application in diabetic screening has not been determined.

Artificial neural network (ANN) analysis may be more able to cope with the variability because this system employs mathematical 'weights' to decide the probability of input data (for example, a feature on a digitised fundus image) belonging to a particular output. The 'weights' are adjusted by training the network with data of known output. Subsequently, the network can be tested with new data, providing a probability of such data belonging to a particular output. In medicine, such computer programs have been used to assess chest radiographs[13] and images from histopathology.[14] In screening of patients by mammography, the use of a neural network provided better predictability for the detection of malignancy than inexperienced radiographers,[15] and in ophthalmology, neural networks have been used to diagnose visual field defects.[16-19]

In this study a neural network was trained to recognise features of diabetic fundus images and tested for the sensitivity and specificity of recognition of these features. The predictive capability of the system to correctly detect retinopathy in fundus images was examined to determine whether a trained neural network could be used in the screening of diabetic patients.

Figure 1   Example of the field of view examined by the neural network.

Methods

A back propagation neural network (Neural Works Professional II/Plus) was employed to analyse images of the fundi of 147 patients with diabetic retinopathy and 32 images from normal subjects. Sixty degree red-free fundus photographs of the posterior pole of the fundus were obtained by a Canon fundus camera (60UV) (Fig 1). The images were scanned by a 'Nikon Coolscan' slide reader to provide 700 by 700 pixel images; these were divided into squares of 30 × 30 and 20 × 20 pixels, depending on which feature was being detected, by applying grids using a specially designed program in C language. Eight bit data from the green plane of the digitised colour image were used to simulate red-free photographs.

The squares of retina were displayed on the computer screen as grey scaled images and classified by a trained observer into normal retina without blood vessels (normal), or normal retina containing normal blood vessels (vessel), or retina with exudates (exudate), or retina with haemorrhages or microaneurysms (haemorrhage) (Fig 2). The squares were divided randomly into training data and test data. Analysis for neovascularisation was not performed at this stage. An assumption was made that neovascularisation was unlikely to be present in the absence of exudates or haemorrhages.

The training data were used to 'teach' the network to classify the squares according to the known outputs (normal, vessel, exudate, and haemorrhage). The network was then tested with the data from the previously 'seen' training data and the 'unseen' test data. The maximum sensitivity and specificity of the network for the correct classification of the square of retina was obtained from receiver operator coefficient (ROC) curves. In this analysis sensitivity and specificity were assumed to be of equal importance and were produced for both the training and the test data; the convergence of the curves providing an indication of the maximum sensitivity and specificity obtainable for the particular feature. In this way it was possible to determine the optimum preprocessing filters in the image analysis and the learning rule, transfer function, number of training iterations, and hidden layer size in the neural network. The maximum size of the training sets required to reach maximum predictive values was also determined in this way.

In the discrimination of the features by the neural network, vessel squares were tested against normals; exudates were tested against normals or vessel; and haemorrhages were tested against normals or vessels.

Once the optimum software protocols had been determined, the detection of diabetic retinopathy by the neural network was compared with that of a gold standard. This standard was considered to be those cases where an ophthalmologist, on examination of the images, agreed with the original diagnosis for the presence of retinopathy. Two hundred diabetic and 101 normal fundus images were randomised and presented to an experienced retinal specialist (THW) for classification as: (1) normal retina, (2) fundus with diabetic retinopathy requiring referral for an ophthalmic opinion, (3) diabetic retinopathy not requiring referral.

The decision for referral was subjective but was made if signs of preproliferative retinopathy, neovascularisation, and treatable maculopathy—for example, circinate exudation were present. Very mild background retinopathy was regarded as not requiring referral.

The images were presented in a masked fashion and the sensitivity and specificity of the neural network system tested.

Results

The optimum computer software protocols and training data sizes after ROC analysis are shown in Table 1. Using these protocols the detection rate for the classification of vessels
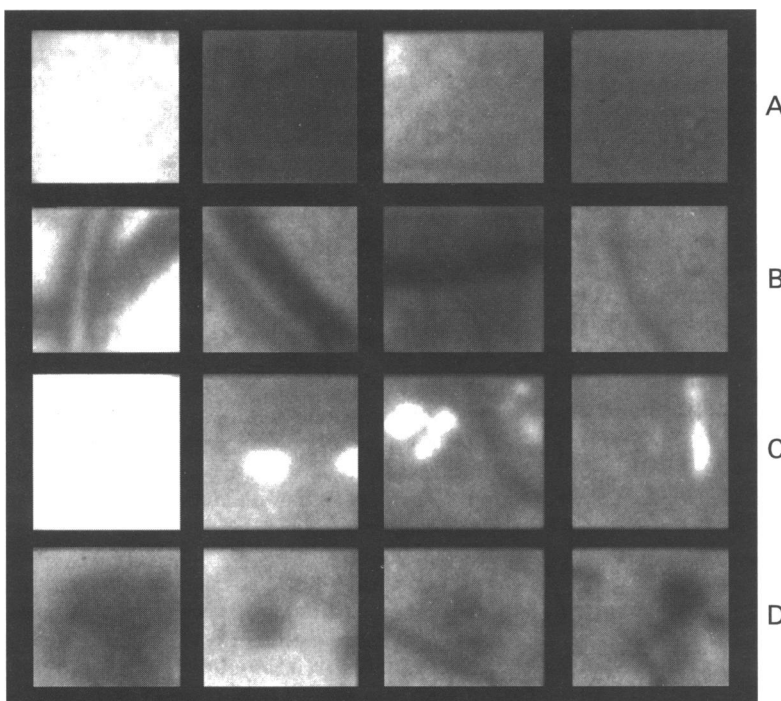

Figure 2   Example squares of (A) normal background retina, (B) normal vessels, (C) exudates, and (D) haemorrhages.

Table 1   The image preprocessing and neural network protocols used for each of the detected features on diabetic fundi

Retinal feature	Blood vessels	Exudates	Haemorrhages
Learning rule	Delta	Delta	Delta
Transfer function	Sigmoid	TanH	TanH
Hidden layer elements	80	80	300
Grid size (pixels)	20 × 20	20 × 20	30 × 30
Squares per image	1225	1225	544
Preprocessing filter	Histogram equalisation, remove small spots (smooth), edge detection	Maximum median smoothing	Minimum median smoothing
Number of training iterations	120 000	300 000	380 000
Training time (hours) (60 MHz Pentium PC)	3	7.5	79
Number of training squares Normal/abnormal	4500/4500	4454/4454	5950/5744
Number of test squares Normal/abnormal	600/600	1000/1000	342/342
Training set specificity/sensitivity	93.7/93.7	94.1/94.1	89.6/89.6
Test set specificity/sensitivity	91.7/91.7	93.1/93.1	73.8/73.8

from normal retina was 91.7%; detection rates for exudates and haemorrhages from normal retina (or retina with vessels) were 93.1% and 73.8% respectively.

The ophthalmologist classified 13 images as of unsure diagnosis and disagreed with the original diagnosis attributable to the image in four 'normal' images and six diabetic images. These images were removed from the comparison with the neural network. For the detection of diabetic retinopathy the network achieved a sensitivity of 88.4% and specificity of 83.5% in comparison with the ophthalmologist (Table 2). If the neural network was asked to identify all of the patients with retinopathy then nine of the 100 patients requiring referral would not have been detected.

The neural network's determination of retinopathy for a whole image is based on a statistical threshold since, in all images, the network will get a certain percentage of squares wrong. By altering the threshold, it is possible to change the network's detection rate to increase the sensitivity; this will of course lead to a lower specificity. For example, by adjusting the threshold for the diagnosis of retinopathy, a sensitivity of 99% was achieved with a specificity of 69%. At this threshold all of the patients whom the ophthalmologist referred were detected.

Discussion

This study has shown that it is possible to train a neural network to recognise common features of diabetic retinopathy on fundus images. The network can be used to identify the presence of vessels, exudates, and haemorrhages with high predictive values. Retinal haemorrhages were most difficult to recognise because of their similar pixel values to vessels. Different image preprocessing and neural network variables were found to provide optimum

protocols for the detection of the various features and by employing these protocols a diagnosis of diabetic retinopathy was made possible. In comparison with the ophthalmologist, the neural network was able to detect patients with retinopathy with sensitivity and specificity values comparable to other screening methods such as optometrists, Polaroids from non-mydriatic cameras, and diabetic physicians for which sensitivity values for the detection retinopathy have varied from 61% to 97% and specificities from 78% to 96%.[6 20-24] The results with the neural network were at least comparable to other systems providing a sensitivity of 88.4% and specificity of 83.5% for the detection of diabetic retinopathy. If the neural network was used to detect the patients with diabetic retinopathy, 9% of the patients who required referral (according to the ophthalmologist) would have been missed. By optimising the sensitivity of the system to 99%, by adjusting the threshold of the ANN, the specificity was reduced to 69% and none of the patients requiring referral were missed

Using commercially available digitised fundus imaging systems it will be possible to produce a system whereby images are obtained from a fundus camera, digitised via a video (or digitised camera), and processed by the neural network. An operator, who need not be trained in the assessment of diabetic retinopathy, could be instructed by the program in the computer to send the patient to an ophthalmology department for further examination if required. Although training of the network in some circumstances required up to 5 days of computing, once trained the network has the capability to analyse an image in a few seconds; therefore, it would not be difficult to perform a rapid assessment of the fundi of large numbers of individuals.

Digitisation of fundus images is likely to become more common as a means of storing and viewing data. Indeed, the digitisation of fundus images in the community for transfer to a central assessment site is a likely way forward for diabetic screening. The neural network program described in this study could be used either to screen for diabetic retinopathy in the community or be used as a filter to remove images with no retinopathy thereby reducing the quantity of the images that need to be assessed centrally by an ophthalmologist. Further improvements in the programming may be possible by using colour images which will require increased computing power[12] or by including the input of previous fundus photographs and medical details in the data set for the neural network.[25] Screening for other common diseases such as glaucomatous optic disc cupping and macular degeneration may also prove possible.

In conclusion, this study has demonstrated that a neural network program can be trained to recognise different features on a fundus image of diabetic retinopathy. The system can screen fundi with retinopathy with accuracy comparable to other systems and is potentially a powerful tool for the screening of diabetic retinopathy.

Table 2   Results from comparison of the artificial neural network (ANN) with ophthalmologist's decision to refer or not. Results are given for the optimum trade off between sensitivity and specificity, and also for near maximum sensitivity

ANN threshold	TP	FP	TN	FN	Specificity/sensitivity
Optimum sensitivity and specificity	160	16	81	21	88.4/83.5
Maximum sensitivity	179	30	67	2	99/69

TP=true positives, FP=false positives, TN=true negatives, FN=false negatives.

1 Singer DE, Nathan DM, Fogel HA, Schachat AP. Screening for diabetic retinopathy. *Ann Intern Med* 1992;**116**:660–71.
2 Javitt JC. Canner JK. Sommer A. Cost effectiveness of current approaches to the control of retinopathy in type I diabetics. *Ophthalmology* 1989;**96**:255–64.
3 Retinopathy Working Party. A protocol for screening for diabetic retinopathy in Europe. *Diabetic Med* 1991;**8**: 263–7.
4 Williams R, Nussey S, Humphrey R, Thomson G. Assessment of non-mydriatic photography in detection of diabetic retinopathy. *BMJ* 1986;**293**:1140–2.
5 Ryder RE, Vora JP, Atiea JA, Owens DR, Hates TM, Young S. Possible new method to improve detection of diabetic retinopathy: Polaroid non-mydriatic photography. *BMJ* 1985;**291**:1256–7.
6 Schachat AP, Hyman L, Leske C, Connell S, Hiner C, Javornik N, *et al.* Comparison of diabetic retinopathy detection by clinical examinations and photograph gradings. *Arch Ophthalmol* 1993;**111**:1064–70.
7 Higgs ER, Harney BA, Kelleher A, Reckless JP. Detection of diabetic retinopathy in the community using a non-mydriatic camera. *Diabetic Med* 199;**8**:551–5.
8 Ward NP, Tomlinson S, Taylor CJ. Image analysis of fundus photographs. The detection and measurement of exudates associated with diabetic retinopathy. *Ophthalmology* 1989; **96**:80–6.
9 Spencer T, Phillips RP, Sharp PF, Forrester JV. Automated detection and quantification of microaneurysms in fluorescein angiograms. *Graefes Arch Clin Exp Ophthalmol* 1992;**230**:36–41.
10 Phillips RP. Spencer T. Ross PG. Sharp PF. Forrester JV. Quantification of diabetic maculopathy by digital imaging of the fundus. *Eye* 1991;**5**:130–7.
11 Katz N, Goldbaum M, Nelson M, Chaudhuri S. An image processing system for automatic retina diagnosis. *SPIE* 1988;**902**:131–7.
12 Goldbaum ME, Katz N, Nelson MR, Haff LR. The discrimination of similarly coloured objects in computer images of the ocular fundus. *Invest Ophthalmol Vis Sci* 1990;**31**:617–23.
13 Lo SC, Freedman MT, Lin JS, Mun SK. Automatic lung nodule detection using profile matching and back-propagation neural network techniques. *Journal of Digital Imaging* 1993;**6**:48–54.
14 Astion ML, Wilding P. The application of backpropagation neural networks to problems in pathology and laboratory medicine. *Arch Path Lab Med* 1992;**116**:995–1001.
15 Wu Y, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE. Artificial neural networks in mammography: application to decision making in the diagnosis of breast cancer. *Radiology* 1993;**187**:81–7.
16 Spenceley SE, Henson DB, Bull DR. visual field analysis using artificial neural networks. *Ophthal Physiol Opt* 1994;**14**:239–48.
17 Goldbaum MH, Sample PA, White H, Cote B, Raphaelian P, Fechtner RD, *et al.* Interpretation of automated perimetry for glaucoma by neural network *Invest Ophthalmol Vis Sci* 1994;**35**:3362–73.
18 Keating D, Mutlukan E, Evans A, Mc Garvie J, Damato B. A back propagation neural network for the classification of visual field data. *Phys Med Biol* 1993;**38**:1263–70.
19 Mutlukan E, Keating D. Visual field interpretation with a personal computer based neural network. *Eye* 1994;**8**: 321–3.
20 Taylor R, Lovelock L, Tunbridge WM, Alberti KG, Brackenridge RG, Stephenson P, *et al.* Comparison of non-mydriatic retinal photography with ophthalmoscopy in patients: mobile retinal camera study. *BMJ* 1990;**301**: 1243–7.
21 Pugh JA, Jacobson JM, Van Heuven WA, Watters JA, Tuley MR, Lairson DR, *et al.* Screening for diabetic retinopathy. The wide-angle retinal camera. *Diabetes Care* 1993;**16**: 889–95.
22 Lee VS, Kingsley RM, Lee ET, Lu M, Russell D, Asal NR, *et al.* The diagnosis of diabetic retinopathy. Ophthalmoscopy versus fundus photography. *Ophthalmology* 1993;**100**: 1504–12.
23 Kalm H, Egertsen R, Blohme G. Non-stereo fundus photography as a screening procedure for diabetic retinopathy among patients with type II diabetes. Compared with 60D enhanced slit-lamp examination. *Acta Ophthalmol* 1983;**67**:546–53.
24 Wareham NJ. Cost-effectiveness of alternative methods for diabetic retinopathy screening. [Letter] *Diabetes Care* 1993; **16**:844.
25 Kalm H. Non-stereo photographic screening in long-term follow-up for detection of proliferative diabetic retinopathy. *Acta Ophthalmol* 1992;**70**:228–34.

Appendix

THE ARTIFICIAL NEURAL NETWORK

The artificial neural network (ANN) computer program in some ways mimics the analytical processes of the human brain by being able to learn to recognise patterns of data. In this study the ANN consisted of three layers: an input layer and output layer connected by a hidden layer (Fig 3). The hidden layer consists of a number of nodes connected to the input and output nodes by mathematical algorithms (weights). Training is achieved by presenting the network with many examples of data of known output (each time an example is presented equals one iteration). Once the weights in the hidden layer have been adjusted by training, the network can be shown a previously unseen input and categorise this into the appropriate output. For example, an ANN may be provided with the digitised data of various scotomata from visual field analysis such as arcuate defects, hemianopias, etc. Once the network has been trained to recognise the digitised visual fields, an example of a scotoma which has not previously been seen can be presented. Even though this may not be exactly the same as any of the scotomata in the training set the neural network will be able to provide a probability of the scotoma belonging to one of the outputs—for example, arcuate defect. The ANN can therefore categorise data which are inherently variable as most images of biological systems are, for example, fundus photographs.

COMPARISON OF ANN WITH CONVENTIONAL COMPUTER PROGRAMMING

There are key differences in the way conventional programs and artificial neural networks work. The former require programming whereas ANN are trained. A conventional program uses serial processing, ANN uses parallel processing. ANN are fault tolerant—that is, removal of some of the elements in the program still allows it to function rather like excision of a small part of the human cortex may allow continued overall function. This is not possible with conventional computer systems.

Conventional programs are accurate and good at numerical processing but poor at pattern recognition and prediction of trends from complex data. The opposite is true of ANN which can be used for recognition of patterns of hand writing or speech, or predicting financial or economic trends.

THE ANN PROGRAM

The network used in this study was a back propagation network in which there was a feedback mechanism allowing the network to retrain itself on data for adjustment of the internal weights. Transfers between each layer (the input layer, output layer, and hidden layer) are governed by a transfer function and the back propagation controlled by a learning function both of which can be varied to optimise performance.

The memory of a network is stored in the connection weights, the number of which are determined by the number of elements in the hidden layer. Increasing the number of elements in the hidden layer increases the number of possibilities that the network can learn, while reducing the number forces the network to try to generalise. The optimum number of weights for a particular problem must be determined, for example, in this study by alter-
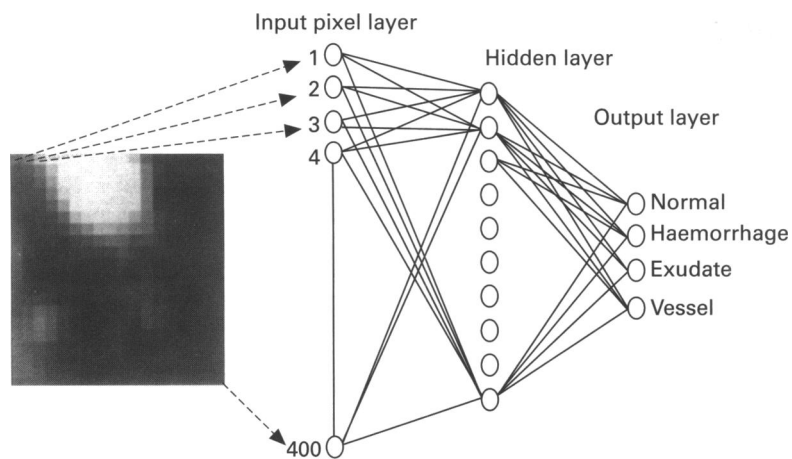
Figure 3 The structure of a simple neural network is represented with input, hidden, and output layers

ing the number of elements in the hidden layer and testing by ROC curve analysis.

The number of examples used in training greatly affects the outcome of the neural network process. By training with too few examples, the network can learn all of the training set by memorising separate details from each example. The network performs poorly when tested with data it has not yet seen because of lack of generalisation. Many examples are required for optimum performance—for example, in this study thousands of squares containing the retinal features were used.

IMAGE PROCESSING

Image preprocessing filters are software programs which allow manipulation of digitised images. Those utilised in this study were from standard commercially available software.

Median smoothing was used for the detection of exudates and haemorrhages. This was a 9 point neighbourhood median filter. In this filter, for each pixel the nearest 8 neighbouring pixels were compared and the median value substituted into the centre pixel. Maximum median filtering was used for exudates where the maximum neighbourhood value was substituted and minimum median filtering was used for haemorrhages where the minimum pixel value was used. For the detection of vessels, a 9 point averaging filter was first used followed by a Sobel edge detection filter.

BJO

Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool.

G G Gardner, D Keating, T H Williamson and A T Elliott

Updated information and services can be found at:

http://bjo.bmj.com/content/80/11/940

These include:

Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes