Bayesian Classification and Forecasting of Visual Field Deterioration

Allan Tucker, Xiaohui Liu
Department of Computer Science
Brunel Univeristy, Middlesex,
UK. UB8 3PH
allan.tucker@brunel.ac.uk
xiaohui.liu@brunel.ac.uk

David Garway-Heath
Glaucoma Unit,
Moorfield's Eye Hospital,
London, UK.
david.garway-heath@moorfields.nhs.uk

Abstract

Recently, there has been an explosion in the amount of data being stored on patients who suffer from visual deterioration including field test data, retinal image data and patient demographic data. In this paper we document some preliminary work in what will be a large-scale investigation of visual field deterioration in conditions such as glaucoma. We explore the use of Bayesian network models to classify and forecast VF data as well as predict the conversion of VFs from non-glaucomatous to glaucomatous. Bayesian networks can easily be interpreted by non-statisticians and queried in order to discover interesting characteristics of visual field deterioration in different conditions such as glaucoma. Initial results are promising for both the classification and forecasting of visual field tests but leave room for improvement. Analysis of the models reveals the potential of using such models for knowledge discovery within ophthalmic databases.

1 Introduction

The Visual Field (VF) test assesses the sensitivity of the retina to light. It is typically measured by automated perimetry, a technique in which the subject views a dim background as brighter spots of light are shone onto the background at various locations in a regular grid pattern. The brightness at which the subject sees the spots of light is related to the retinal sensitivity. There are many diseases and conditions that affect the VF, the most common being neurological disease and glaucoma. Early detection of glaucoma as well as other conditions and diseases that cause visual impairment are invaluable as early intervention can slow VF deterioration.

There has been much research on modelling visual field data starting with the exploration of the distribution of point-by-point light sensitivity, at a single point in time, in normal (Katz and Sommer 1986, Heijl et al. 1987) and glaucomatous populations (Weber and Rau 1992). However, much remains unknown about the behaviour of the visual field test, such as the light sensitivity relationship between adjacent and distant visual field test points, the relationship between light sensitivity and other ocular parameters (such as optic nerve appearance and intraocular pressure level), and how stable and deteriorating visual fields behave over time. Historically, various approaches have been made to identify deterioration within a visual field series. These include clinical judgement (expert opinion), and the use of classification systems (Lee et al. 2002), trend analysis (Heijl et al. 1990, Fitzke et al. 1996) and event analysis (Heijl et al. 1990). Goldbaum et al. document a comprehensive comparison of machine learning classifier systems for the classification of glaucomatous visual fields (Goldbaum et al. 2002). This does not include any Bayesian methods of classification, which have recently shown excellent properties (Langley and Sage 1994). Ibanez and Simo have investigated spatio-temporal statistical models with the aim of forecasting visual field deterioration (Ibanez and Simo 2003) but to date have only looked at visual fields of normal eyes. We know of little research in using probabilistic models to understand VF data. Previously, a state space model has been used to classify glaucomatous patients (Andersen and Jeppesen 1998) and Bayesian statistics have been proposed to record VF data (Bengtsson et al. 1997).

In this paper we look at Bayesian network methods for combining both the classification and the forecasting of visual field data. The models can easily be interpreted by non-statisticians and queried in order to discover interesting characteristics of visual field deterioration in different conditions such as glaucoma. Visual field data has both temporal and spatial characteristics. For this reason, we make use of spatial operators that we developed for efficiently learning Bayesian networks from spatio-temporal data such as visual field data (Tucker et al. 2003). The results documented in this paper are very promising: the classification and forecasting all achieved reasonable cross validation errors (approximately 75 percent correctly classified), though there is still room for improvement through the use of more data and model refinement.

In addition, Bayesian networks allow us to analyse the discovered network structures in order to try to understand characteristics of visual field deterioration. In this paper, the networks have demonstrated the classic 'nasal step' field defect which is known to be an early sign of glaucoma.

2 Background

Bayesian Networks (BNs) are probabilistic models that can be used to combine expert knowledge and data. They also facilitate the discovery of complex relationships in large datasets. A BN consists of a directed acyclic graph, made up of links between nodes that represent variables in the domain. The links are directed from a parent node to a child node, and with each node there is an associated set of conditional probability distributions. Learning the structure of a BN from data (Cooper 1992) is a non-trivial problem due to the large number of candidate network structures and as a result there has been substantial research in developing efficient algorithms within the optimisation communities. The Dynamic Bayesian Network (DBN) is an extension of the BN that can model time series (Friedman 1998).

Previously we have developed algorithms for efficiently learning DBN structures (Tucker et al. 2001). We have also investigated learning DBNs from VF Data in order to explore the VF relationships discovered within the DBN structure. Due to the spatial as well as the temporal nature of VF data we developed spatial operators to efficiently learn DBN structures. Full documentation of the algorithm and operators can be found in (Tucker et al. 2003). In this paper we focus on forecasting VF data, classifying VF data as glaucomatous or non-glaucomatous and predicting the classifications of VF data in order to try and preempt when a VF is about to convert to glaucomatous.

3 Experiments and Results

The dataset used in this paper involves 24 patients with 623 measurements in all, concerning only the right eye of patients who are converters (from normal to early glaucomatous). Two points in the VF correspond to the blind spot and should not contain any useful data. We have included these points to check for spurious relationships. See Figure 1 for an example VF test. The data were discretised into four states using a frequency-based method where bin sizes are determined such that there are equal numbers of each state per variable in the dataset. Discretisation was performed on a point-wise basis. The data are slightly imbalanced - approximately 60 percent of the tests are classified as normal whilst 40 percent are classified as glaucomatous.

This paper covers three different sets of experiments:

1. The classification of glaucomatous VF.
2. The forecasting of VF data.
3. The predicting of glaucomatous VF conversion.

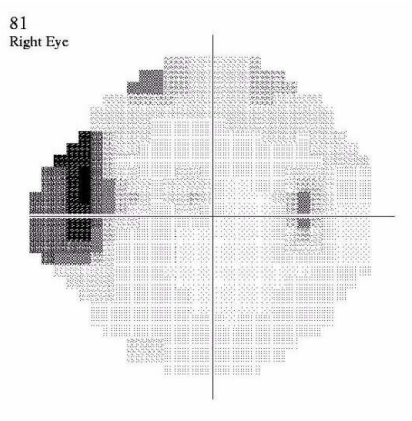

Figure 1: A Typical VF Test. Note the Large Dark Patch which Represents deterioration and the Small Dark Patch on the Horizontal Axis which Represents the Blindspot

In the first set of experiments the state of a previously unseen VF must be classified as either glaucomatous or non-glaucomatous, given the state of the VF. This involves learning several classes of Bayesian classifier and applying leave one out cross validation in order to score the forecast quality without overfitting. We look at learning BNs and DBNs using the algorithm in Tucker et al. 2003. We also look at other Bayesian classification methods including the naive Bayes' classifier (Langley and Sage 1994) and the Tree Augmented Network (TAN) classifier (Friedman 1997) both of which impose a structure over the variables. The naive Bayes Classifier assumes all features are independent and consists of a network where the class node is the parent of every feature node. TAN on the other hand relaxes this assumption and so involves learning a tree structure over the variables. Table 1 shows the Cross Validation (CV) error for each system as well as the sensitivity of the classifications (the percentage of correctly classified glaucomatous fields) and the specificity (the percentage correctly identified non-glaucomatous fields). It also shows the results for predicting future classifications with a DBN which we will discuss later. First of all we look at the different Bayesian classifier methods for modelling the

Table 1: Results for Classification and Predicting Conversion

Method	Sens.	Spec.	Error
Naive Classifier	0.60	0.81	0.28
TAN Classifier	0.60	0.79	0.28
BN Classifier	0.63	0.80	0.27
DBN Classifier	0.65	0.79	0.26
DBN Converter Prediciton	0.65	0.81	0.25

clinician's decision of whether a VF is glaucomatous or not. In Table 1 it can be seen that the Naive Bayes classifier does well with respect to error and specificity but less well with respect to sensitivity. This appears to be common across all methods. Surprisingly, the TAN, BN and DBN methods do slightly worse than Naive Bayes with respect to specificity (though this difference is minimal). However, BN and DBN improve on sensitivity and error. The best method across all classifiers for error and sensitivity is the DBN classifier. This is most likely due to its ability to model changes in VF points over time, though the difference between classifier results is very small. It should be worth noting that the definition of conversion is quite strict (Heijl et al. 2003) possibly resulting in earlier fields displaying glaucomatous features that do not yet reach the threshold of the conversion criteria. This will be investigated in further research.

The second set of experiments involves learning DBNs in order to forecast the future states of a VF. Networks are scored according to various pieces of expert knowledge from clinicians including the expected development of retinal nerve fibre bundle defects. In addition the forecasts are scored using leave one out cross validation. A typical discovered DBN structure for forecasting VF data is illustrated in Figure 2. Notice that the links show in general a spatial arrangement with most link's parents and children being close to one another. Notice also that there are no links associated with the blind spot as should be expected.

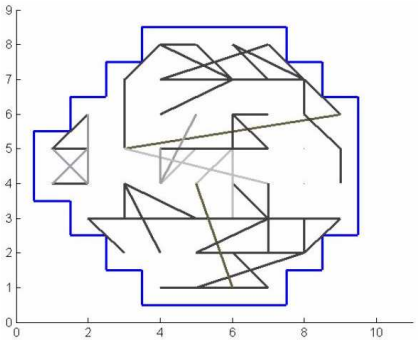

Figure 2: A Typical Dynamic Bayesian Network structure Used for Forecasting VF Data. Note the Lack of Links at the Blind Spot

The quality of the discovered DBN structures used to forecast future states of the VF data is illustrated in Table 2 below, which shows two measures of final network quality based upon clinical knowledge of the eye. Points on the VF should be related if they sit on the same Nerve Fibre Bundle (NFB). Figure 3 shows the expected NFB layout (demarcated by lines) according to (Garway-Heath et al. 2000). Indeed the discovered networks contained 78 percent of links within the same bundle.

Another expected characteristic of the VF is that points that are closely related should have a similar angular distance from the Optic Nerve (ON). The mean ON distance between parents and children of links in the DBN was found to be 19 degrees. This is relatively low, bearing in mind that the maximum distance between the angle of two points to the ON is 180 degrees. Using the discovered DBN to forecast future VF values using leave one out cross validation generated a 28 percent error. This error is not extremely low but is an encouraging first attempt to forecast using a DBN model with no expert knowledge whatsoever. We intend to look at ways at improving this forecast through the use of more information such as retinal images as well as incorporating expert prior knowledge, to which the DBN is extremely suited.

Table 2: Mean Quality of DBNs learnt from VF Data

Links in Same Bundle	78.3
Mean ON Distance	19.2
Mean Forecast Error	0.28

The third set of experiments involve trying to predict the classification of the next unseen time point of VF data in order to pre-empt the glaucomatous diagnosis. Again, leave one out cross validation is used to score these experiments. Table 1 shows the CV results for learning DBNs for predicting VF conversion. The specificity was good at 80 percent but sensitivity not so good at 65 percent. We intend to improve this prediction in various ways including using more data such as intraocular pressure.

A useful characteristic of Bayesian networks is that we can easily interpret the structure and parameters of the models. Figure 3 shows the percentage of times a VF point was discovered as a parent node (during cross validation) for predicting whether a VF converts in the next time point. It is interesting to notice that a cluster of VF points appears along the horizontal axis. It appears that these points were the most predictive for VF conversion to glaucomatous and are also those involved in the classic 'nasal step' field defect, which is known to be an early sign of glaucoma (Hart and Becker 1982). Another individual point also appears to be useful in forecasting conversion (marked in bold in Figure 3). This point is only 10 degrees away (ON distance) from the nasal points, so might be an expected finding if the nasal step is an early sign.

4 Conclusions and Future Work

In this paper we have begun what will be a large investigation into the modelling of patient visual field data in order to understand better different visual field conditions and diseases such as glaucoma. We have used probabilistic models to model the clinician's classification decision of whether a visual field is glaucomatous or not with some success. We have also used temporal models to forecast future visual field states based on previous visual field data, as well as made an attempt

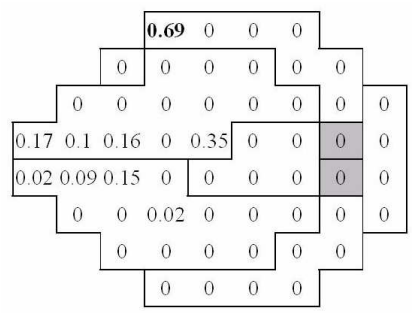

Figure 3: Nerve Fibre Bundle Layout and The Proportion of Networks Learnt Containing a Link to the Classification Node. In Other Words, the Most Influential Points for Forecasting VF Classification. Note the Shaded Region Representing the Blind Spot

in forecasting the conversion of healthy visual fields to glaucomatous. It has been informative to investigate which visual field points have proved the most predictive for glaucomatous conversion. This may inspire future visual field test strategies where more weight, or testing time, can be given to these more informative regions of the visual field.

We intend to focus on improving the classification and forecasts by including more visual field data from both eyes as well as demographic information, intraocular pressure and retinal image information. For example, we may be averaging over several different classes of glaucoma, which can be separated out along demographic lines. In addition, we may look at using continuous forms of Bayesian networks (Geiger and Heckerman 1994). Another issue we must investigate is possible bias due to the criteria used to define conversion. We may make use of retinal image data to redefine conversion.

5 Acknowledgements

We would like to thank Nick Strouthidis for his assistance in preparing the datasets. This work was funded by the EPSRC grant GR/R35018/01.

References

[1] K.E. Anderssen, V. Jeppesen, Classifying Visual Field Data, Technical Report, MSc Thesis, Aalborg University, Denmark, 1998.

[2] B. Bengtsson et al., A new generation of algorithms for computerized threshold perimetry, SITA. Acta Ophthalmologica Scandinavica 75, (1997), 368-375.

[3] Cooper, G.F., Herskovitz, E., A Bayesian Method for the Induction of Probabilistic Networks from Data, Machine Learning, Vol. 9, (1992), 309-347.

[4] Fitzke, F.W., Hitchings, R.A., Poinoosawmy, D., McNaught, A.I., Crabb, D.P. Analysis of visual field progression in glaucoma. Br J Ophthalmol. 1996 Jan;80(1):40-8.

[5] Friedman, N., Geiger, D., Goldszmidt, M. Bayesian Network Classifiers, Machine Learning 29, No.s 2-3, (1997), 131-163.

[6] Friedman, N., Learning the Structure of Dynamic Probabilistic Networks, Proceedings of the 14th Annual Conference on Uncertainty in AI, (1998), 139-147.

[7] Garway-Heath, D.F., Fitzke, F., Hitchings, R.A., Mapping the Visual Field to the Optic Disc, Opthalmology 2000, 107, (2000), 1809-1815.

[8] Geiger, D., Heckerman, D. Learning Gaussian Networks, Technical Report, MSR?TR-94-10, Microsoft Research, (1994).

[9] Haley, M.J. (ed.), The Field Analyzer Primer, Allergan Humphrey, San Leandro, California, (1987).

[10] Hart, W. M., Jr. and Becker, B. The onset and evolution of glaucomatous visual field defects, Ophthalmology, (1982), 89, 268-79

[11] Heijl, A., Lindgren, G., Olsson, J. Normal variability of static perimetric threshold values across the central visual field. Arch Ophthalmol. (1987);105:1544-1549.

[12] Heijl, A., Lindgren, G., Lindgren, A. Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2, in Mills R and Heijl A (eds): Perimetry Update 1990/1. Amsterdam/New York, Kugler Publications, 1990, 303-15.

[13] Heijl, A., Bengtsson, B., Leske, M.C., Bengtsson, B., Bergstrm, A., sman, P. Comparison of AGIS, CIGTS and EMGT Visual Field Glaucoma Progression Criteria. A Pilot Study, Investigative Ophthalmology and Visual Science, (2003).

[14] Ibanez, M.V., Simo, A. Spatio-Temporal Modelling of Perimetric Test Data, Statistics in Medicine 00, (2003), 1-41.

[15] Katz J, Sommer A. Asymmetry and variation in the normal hill of vision. Arch Ophthalmol. (1986);104:65-68.

[16] Langley, P., Sage, S. Induction of Selective Bayesian Classifiers, Proceedings of the 10th Annual Conference on Uncertainty in AI, (1994), 399-406.

[17] Lee, A.C., Sample, P.A., Blumenthal, E.Z., Berry, C., Zangwill, L., Weinreb, R.N., Infrequent confirmation of visual field progression. Ophthalmology. 2002 Jun;109(6):1059-65.

[18] Spry, P.G., Johnson, C.A. Identification of progressive glaucomatous visual field loss. Surv Ophthalmol. (2002);47(2):158-73.

[19] Tucker, A., Liu, X., Ogden-Swift, A., Evolutionary Learning of Dynamic Probabilistic Models with Large Time Lags, International Journal of Intelligent Systems, Vol. 16, No. 5, (2001), 621-646.

[20] Tucker, A., Garway-Heath, D., Liu, X., Spatial Operators for Evolving Dynamic Bayesian Networks from Spatio-Temporal Data, GECCO (2003), 2346-2357.

[21] Weber, J., Rau, S. The properties of perimetric thresholds in normal and glaucomatous eyes. : Ger J Ophthalmol. (1992);1:79-85.