A Neural Network Chromosome Classifier

JIM GRAHAM, PHIL ERRINGTON AND ANNE JENNINGS

Wolfson Image Analysis Unit Department of Medical Biophysics University of Manchester Stopford Building Oxford Road Manchester M13 9PT UK

Chromosome Analysis/Automated Karyotyping/Classification/Neural Networks

We present a chromosome classifier for automated karyotyping of banded chromosomes which uses a multi-layer perceptron neural network. Two network configurations have been investigated. The resulting classifiers have been trained and tested on data from three different data sets covering a range of data quality. Classification results compare very favourably with those obtained using a highly optimised parametric classifier.

KARYOTYPING

Figure 1 shows an image of a dividing cell at metaphase as it might be seen on a microscope slide at high magnification. At this stage in cell division, the genetic material of the cell has contracted into chromosomes which can be seen as individual ribbon-shaped objects which have been stained to show a characteristic pattern of bands along their lengths (G-banding). The clinically important procedure of karyotyping consists of grouping the chromosomes into 24 distinct pairs (22 pairs of autosomes and two sex chromosomes, XX or XY for female or male individuals respectively) and identifying deviations from the normal appearance of this classification which are associated with certain genetically determined clinical syndromes.

Karyotyping, along with other chromosomal investigations, such as aberration detection, has been seen as a useful application area for automation using computer vision techniques, and in recent years there has been some success in introducing automated karyotyping systems into clinical laboratories\(^1\). For automatic classification, features are measured from the images corresponding to the features used by human karyotypers, namely chromosome length, centromere position and banding pattern. Chromosome lengths vary dramatically from cell to cell, but the relative lengths within a cell are reliable features. Banding patterns are generally similar for the same chromosomes in different cells, but there is considerable variability in detail. Length and centromere position are straightforwardly measured to provide classification features. The centromere position is usually expressed in terms of the centromeric index: the ratio of the length of the p-arm to the total chromosome length. A number of representations of the banding pattern have been proposed to be used as features in statistical classifiers. They typically involve the extraction of a one dimensional density profile (Fig. 2) or the identification...
Fig. 1. An image of a metaphase cell showing G-banded chromosomes

Fig. 2. Representing the banding pattern. (a) A banded chromosome with fitted curved axis. (b) The one dimensional profile obtained by averaging across the axis. (c) The low resolution profile, obtained by local averaging along the profile, for input into the network.
of particular bands in the image. The most successful of these methods has been that based on weighted density distributions originally proposed by Granum and subsequently much refined (Piper and Granum).

NEURAL NETWORKS

Artificial Neural Networks have recently been the object of a great deal of interest for computational tasks, such as pattern recognition, in which the performance of traditional algorithmic methods fall well short of the performance of human experts (or even non-experts). The essence of a neural network is that a calculation, such as a classification based on a set of input features, is distributed over a set of simple calculations. Figure 3 shows a commonly used network organisation, known as a multi-layer perceptron (MLP). Each node in this figure represents the execution of a simple calculation—a nonlinear function of the sum of a set of inputs—to create a single output. This output is sent along links to form the input to a number of other nodes (Fig. 4). The important point is that each link has associated with it a weight by which the output of the previous node is multiplied before becoming the input to the next node. This pattern of weights determines the mapping from the inputs to the outputs, i.e. specifies the classification. Networks are generally highly connected. In the case of the

Fig. 3. A multi-layer perceptron network showing the connectivity between nodes. In each layer send their outputs along links to every node in the next layer. All the nodes Inputs to the network are classification features and outputs are the classification results.
multi-layer perceptron, the nodes are connected in layers, each node in a given layer receiving inputs from each node in the preceding layer and sending its outputs via weighted links to each node in the next layer. The number of weights to be specified is therefore large. Fortunately the pattern of weights appropriate for a particular task can be determined by training. Features for classification are presented as inputs and the difference between the observed output and the desired output (the correct class) is used to adjust the pattern of weights\(^6\). A number of parameters need to be specified for the network, among which are the number of nodes and the number of layers required to achieve the best classification performance. These, together with parameters which control training are generally determined empirically.

EXPERIMENTS

Neural networks have the capacity to be highly adaptable due to the fact that the classification task is specified by training. Here we present some experimental investigations into the applicability of neural networks to the chromosome classification problem.

Data

Our classification experiments have used databases of annotated chromosome measurements originating in Copenhagen, Edinburgh and Philadelphia and available within the EC Concerted Action on Automated Cytogenetics. These data sets are summarised in Table 1.

In the case of the Copenhagen data set, chromosomes were carefully measured from

\[ x_i = f(\sum w_{ij} \times x_i) \]
selected cells of high quality. In particular, no overlapped chromosomes were included. The other two data sets were taken from routine material and include measurements errors arising from overlapped or bent chromosomes. The nature of the slide preparation methods results in chorionic villus samples providing cells of significantly poorer visual quality than in the case of peripheral blood. The data sets therefore represent a range of data quality.

In tests of classification performance we have employed the procedure normal in pattern recognition of training on half of the data set and classifying the "unseen" half. Classification results presented are average results for unseen classification of both halves of each data set.

A Simple Network

In our first approach, we have used a MLP network and attempted to optimise the classification performance by varying the numbers of hidden nodes, and the number of layers over which they are distributed. We excluded the sex chromosomes from this study, since they occurred in smaller numbers than the autosomes in the data sets. Our network therefore had 22 output nodes. The density profiles used were of variable length up to 140 samples. For input to the net, these were reduced to 15 samples for all chromosomes by local averaging (Fig. 2). There were therefore 15 input nodes representing the banding profile, together with two inputs for the normalised length and centromeric index. Best classification was obtained with two hidden layers of 51 and 22 nodes respectively. Table 2 shows the classification performance as measured by the misclassification rates on the unseen data compared with the Weighted Density Distribution (WDD) classifier described by Piper and Granum⁴. It can be seen from table 2 that the performance of the network classifier approaches, but does not improve on the performance of the parametric classifier.

### Table 1. The three chromosome data sets used.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Tissue of origin</th>
<th>Digitisation method</th>
<th>Number of chromosomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copenhagen</td>
<td>Peripheral blood</td>
<td>Densitometry from photographic negatives</td>
<td>8106</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>Peripheral blood</td>
<td>T.V. Camera</td>
<td>5469</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>Chorionic villus</td>
<td>CCD line scanner</td>
<td>5817</td>
</tr>
</tbody>
</table>

### Table 2. Misclassification rates for the three data sets using a simple multi-layer perceptron compared with the performance of the Weighted Density Distribution (WDD) classifier.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Simple Network</th>
<th>WDD classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copenhagen</td>
<td>6.9%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>19.3%</td>
<td>18.3%</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>25.6%</td>
<td>22.8%</td>
</tr>
</tbody>
</table>
A Compound Network

A more complex classification strategy was implemented in the form of a compound network, consisting of a preclassifier using length and centromeric index alone, producing outputs which, along with the banding profile data, are input into a second classifier (Fig. 5). Both classifiers are two layer (one hidden layer) MLP’s. The first MLP acts as an unbanded classifier, whose outputs correspond to the Bayesian likelihoods of belonging to ten classes approximating to the Denver groups\(^7,8\). The second stage classifier, therefore, has 25 inputs and 24 outputs since sex chromosomes were included in this case. The rationale behind the use of this classifier was to give greater weight to the size and centromeric index features than was

![Diagram of compound network](image)

Fig. 5. The compound network. A pre-classifier, consisting of a two layer MLP with two input nodes and ten output nodes, produces an output approximating to the Denver classification based on size and centromeric index. These outputs, together with the density profile, form the inputs to a second stage classifier, also a two layer MLP.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Simple Network</th>
<th>WDD classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copenhagen</td>
<td>6.2%</td>
<td>6.5%</td>
</tr>
<tr>
<td>Edinburgh</td>
<td>17.1%</td>
<td>18.3%</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>23.0%</td>
<td>22.8%</td>
</tr>
</tbody>
</table>

Table 3. Misclassification rates for the three data sets using a compound network compared with the performance of the Weighted Density Distribution (WDD) classifier.
the case for the simple network. The performance of this classifier is shown compared to the WDD classifier in table 3.

The addition of the preclassification using the centromere and length inputs produced a significant improvement in classification of all three data sets, particularly the Edinburgh and Philadelphia sets. This performance proved best with a two layer net with 50 hidden nodes.

DISCUSSION

There are several aspects of the chromosome classification problem which make it attractive for the application of neural networks. The one dimensional nature of the density profiles representing the banding pattern provides a source of data of tractable dimension from real, useful images. The variability of the profiles provides scope for the generalising properties of networks.

The comparison of our network classifier with the Weighted Density Distribution classifier is useful for two reasons. Firstly, the WDD classifier is highly successful, both in terms of its classification results and its widespread application. Secondly, identical classification tests to those described here have been performed on the WDD classifier using the same data sets. The results of these tests provide the comparisons shown in tables 2 and 3. These comparisons are therefore very direct indeed.

The network classifier shows some improvement overall in classification performance over the WDD classifier. However the similarity in performance of the two classifiers is striking. It may indicate that for these data sets a limit has been reached in the ability to effect a classification on the basis of the data available.

The WDD classifier, in common with most existing statistical classifier for karyotyping, was designed to deal with low resolution banding patterns typical of those obtained from rather contracted metaphase chromosomes which have been used for routine analysis in the recent past. It is now much more common for karyotyping to be performed using more elongated metaphase or prometaphase chromosomes or, for the detection of subtle abnormalities, with prophase chromosomes. These preparations provide much higher resolution banding than that previously used, and it is not at all clear that existing classifiers will be easily modified to deal with this type of data. Our experience shows that good classification performance can be achieved rather quickly with a trainable neural network, and we are encouraged that adaptation to the requirements of different types of chromosome preparations should be straightforward.

ACKNOWLEDGEMENTS

This work was greatly facilitated by the exchange of materials and ideas available within the Concerted Action of Automated Cytogenetics Groups supported by the European Community, Project No. II.1.1/13. We are particularly grateful to Jim Piper of the MRC Human Cytogenetics Unit, Edinburgh, for the provision of classification results using the WDD classifier.
A NEURAL NETWORK CHROMOSOME CLASSIFIER

We are grateful for the support of the United Kingdom Science and Engineering Research Council for PE and AJ.

REFERENCES


