

A Neural Network Approach to Recognition of Structural Aberrations in Chromosomes¹

M. Turner^a, J. Austin^a, N. Allinson^b, P. Thompson^c.
Departments of Computer Science^a, Electronics^b and Psychology^c,
York University,
York, UK.

Abstract

Neural networks are applied to the problem of detecting structural aberrations in chromosomes from shape. We present a simple technique for initial location of scattered chromosomal objects within multi-resolution images of human blood cells. A system for classifying located objects is also described. It is proposed that the system be applied to multi-resolution images. Application to low resolution images is illustrated.

1. Introduction

The rapid development of human cytogenetics in recent years has produced an increased awareness of the relationships between chromosomal abnormalities and medical disorders, which in turn has lead to a large growth in the volume of clinical microscope work. There is now an great demand for rapid, accurate and cheap methods for analysing chromosome spreads. Currently, semi-automated systems are available for clinical karyotyping (1), prevalent in pre-natal screening, where the task is to determine the characteristic chromosome constitution of an individual. However, no systems are available for chromosome aberration scoring - the search for specific structural abnormalities within a cell.

A number of physical and chemical agents can induce chromosome structural aberrations - breakage and rearrangement of chromosomal material. Affected cells may contain unusually shaped objects, or have abnormal chromosome constitutions, and this may be spotted by a skilled cytotechnician using a microscope at high magnification. There is, however, a shortage of skilled technicians and their training - which for the most part consists of viewing many examples of damaged cells - is time consuming. Attempts at automation have focused on image segmentation, measurement of predefined features, and classification based upon these measurements (2-4). However, this approach has failed to cope with the general variability in chromosome appearance. Chromosomes are often bent, or they may touch or overlap. Diffraction effects make objects appear fuzzy under a light microscope. In addition, the existence of staining gradients and artefacts is troublesome.

We use neural network systems acting on multi-resolution images in an attempt tackle the problem of variability. Neural networks are adaptive systems of interconnected units which acquire knowledge through experience rather than preprogramming. Knowledge is

1. Supported by SERC grant 06R00174

stored implicitly in the connection strengths, and recalled in response to the presentation of cues. We hope that by presenting many examples of image portions to neural networks that they can learn through experience to locate chromosomal objects, and to extract image features which permit centring and classification of located objects. Simple, supervised, single-layer networks are employed in the initial location of scattered chromosomal objects within a cell. Kohonen self-organising maps (5) are used to extract salient image features within windows placed in the neighbourhood of located objects. Feature extraction forms the first stage in a centring and classification system.

We have been fortunate to have had assistance from a cytogenetics company in our work. As a result we have been able to learn much about how cytotechnicians recognise structural aberrations, and compare their classification of microscope cell images and digitised (computer) images. In addition, access to a large set of mutagen-treated blood cells has been provided.

2. Initial Location of Chromosomal Material

2.1 Data Set.

Photographs of mutagen-treated human blood cells at metaphase were taken using a microscope-camera system, with enlargements produced and scanned. The resolution of digitised images obtained ranged from 750x750 to 1200x1200 pixels. 256 grey levels were available. Figure 1 shows a portion from a cell image.

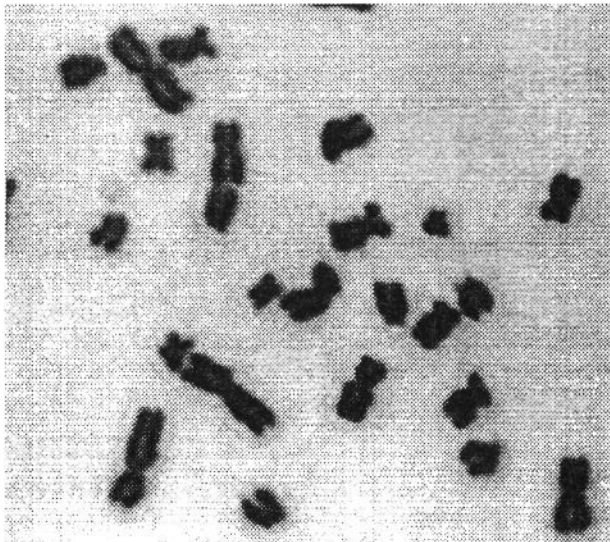


Figure 1. Portion from a cell image. Cells usually contain 46 chromosomes of varying size and orientation.

A number of images were chosen from the collected set and rescaled by linear interpolation to be 512x512 (the originals were approximately square), and a lowpass circular filter of diameter 128 pixels applied (in the frequency domain). Each resultant image formed the base for a 3-layer Gaussian pyramid. We convolved 4x4 neighbourhoods of base level images with Gaussians of standard deviation 1.0, to produce pixel values for the pyramid layer above. By only choosing neighbourhoods centred on every other pixel in every other row a reduction of 2 in each dimension for the next layer was achieved. By repeating the process 512-256-128 Gaussian pyramids were created.

2.2 Network Training.

One simple, supervised, single layer neural network was associated with each of the three levels in our pyramids. Units in a network were arranged as a 2-D sheet. All units had an associated adjustable reference vector $\mathbf{m}(t)$, where t is a time step, and all received identical inputs $\mathbf{x}(t)$, namely the set of normalised image portions grabbed by a moveable window on an image. Each unit in a network was designated responsibility for signalling the presence of chromosomal material within a particular subregion of the window; if a network consisted of $n \times m$ units then the window would be divided into $n \times m$ regions, and unit (i,j) in the network would be responsible for signalling the presence of material centred within window subregion (i,j) . For each pyramid level many image portions containing chromosomal material were presented to a network. Each time, the reference vector of the responsible, or required winning unit, w , and its neighbours, (following Kohonen, see section 3), were adjusted to be closer to the image input $\mathbf{x}(t)$,

$$m_i(t+1) = \frac{m_i(t) + \alpha(t) \mathbf{x}(t)}{\|m_i(t) + \alpha(t) \mathbf{x}(t)\|} \quad (i \in N_w(t)) \dots\dots\dots (1)$$

$$m_i(t+1) = m_i(t) \quad (i \notin N_w(t))$$

Each of the three networks contained 5x5 units and input images were 64x64 pixels. During training the neighbourhood, $N_w(t)$, a square centred on the required winning unit, decreased from 5x5 units to 1x1 units (i.e., just the required winner) over the first 500 steps. Thereafter it remained at 1x1 units. Since neighbourhoods encompassed different subsets of units the net effect of adjustments for each unit in the network tended to be smoothed out over time. (So, for example, if a unit was seldom chosen as the required winner early in training, its reference vector would tend towards an averaged version of the vectors of units in its neighbourhood).

We took $\alpha(t) = 100/t$, so that as the neighbourhood decreased so did the magnitude of reference vector adjustments. Over 1000 steps were carried out for each network. Units in each network developed into something like blob detectors whose receptive fields overlap. Network reference vectors for each pyramid level are shown in figure 2.

2.3 Recall.

Once trained the three networks could be used in a recall phase to locate chromosomal material. Gaussian pyramids were constructed from cells not in the training set. Units were assigned activations according to the inner product between their reference vectors and the input (both normalised), and the unit with the greatest activation was deemed to best indicate the position of a chromosomal object. The best guess location of an object was taken as the centre of the window region which had the same row-column coordinates relative to the window as the winning unit had relative to the network. Once found, the best guess location at one level determined the position at which to centre the window at the level below. As we moved down from the top pyramid level it was possible to locate chromosomal material with increased precision. Furthermore, by gradually reducing the window's size at the pyramid base, and only matching with reference vectors over this region, we could land on chromosomal material with great reliability even in the presence of clutter. A test on 8 pyramids constructed from 8 cell images not in the training set produced one "missed" chromosome at the base level from a total of 374 start positions at the top pyramid level. Here the size of base level window was gradually shrunk from 64x64 to 9x9 (about a chromosome's width) in steps of 5 pixels.

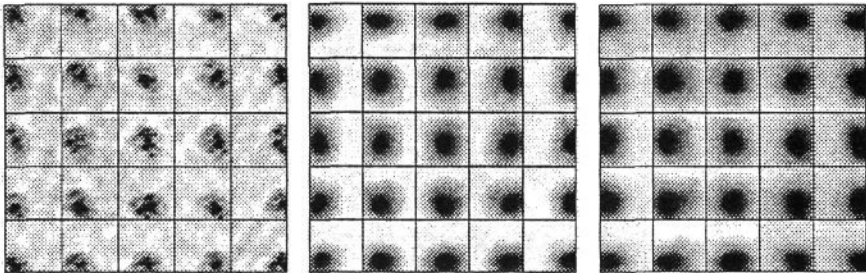


Figure 2. Reference vectors learnt by the networks for the top, middle and bottom pyramid levels.

3. Classification of Low Resolution Images

Given that we can land on chromosomal objects we would like to centre on them (the pyramid system lands on chromosomal objects but does not necessarily find their centres), and perform some classification. Due to the variability in chromosome appearance we are initially only interested in estimating the rough position, size and orientation of objects within low resolution cell images. This information may then be used to constrain chromosome appearance at higher resolutions. The centring and classification system employed on low resolution images (128x128 pixels) is shown schematically in figure 3. The system is centred around the use of a Kohonen network for feature extraction.

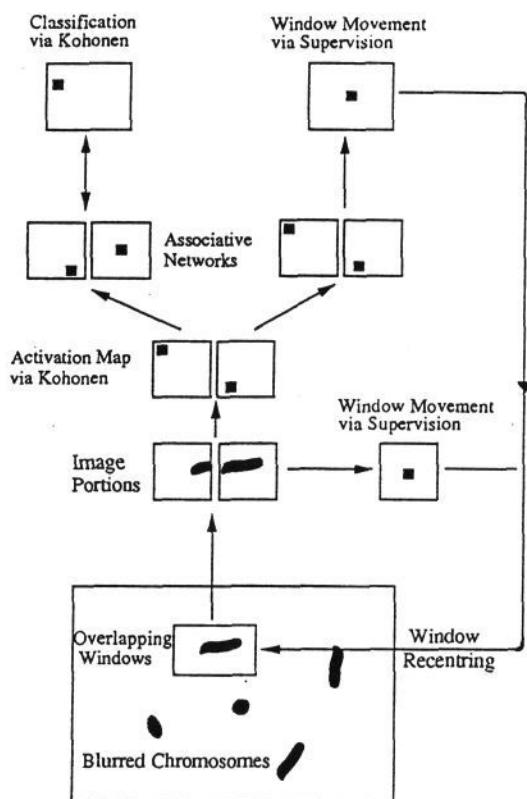


Figure 3. Schematic diagram of centring and classification system.

3.1 Feature Extraction.

The Kohonen Network or Self-Organising Map is a self-organising neural network whose units become specifically tuned to various features present in a set of input signals through a period of unsupervised training. Moreover, the locations of responses on the network tend to become ordered such that nearby units respond to similar input features.

Our supervised networks in section 2 were based upon the Kohonen network. During training we specified a required winning unit for each input image, and adjusted the reference vector of the required winner and its neighbours accordingly. As stated, the Kohonen network itself is unsupervised. Each time an input vector is presented to the network the winning unit is not specified by the user, rather it is taken as that unit whose reference vector is closest by some distance measure to the input vector. When input and reference vectors are normalised a suitable updating algorithm for reference vectors is as given in

equation (1) for our supervised networks, but with the winning unit, w , now such that

$$\mathbf{x}^T(t) \mathbf{m}_w(t) = \max \{ \mathbf{x}^T(t) \mathbf{m}_i(t) \} \dots\dots\dots (2)$$

The effect of equation (2) is that, over many time steps, units tend to become specifically tuned to particular domains of the input space. Furthermore, because reference vectors in the neighbourhood of the winner are adjusted, nearby units on the Kohonen network respond to similar input features. At small time steps the neighbourhood is chosen to be relatively large in size, encouraging a rough global ordering of reference vectors. Gradually the neighbourhood is decreased in size, usually until just the best-matching vector is updated.

Our Kohonen network is arranged as a 2-D sheet, 8x8 units in size. Inputs are normalised image portions grabbed from each window within a set of 5x5 overlapping windows placed on low resolution cell images. The window set is centred on positions located by the pyramid system described in section 2. Each window is 8 x 8 pixels. Neighbouring windows overlap by 60%. (The characteristics of the window set are chosen after consideration of: the need to encode similarity between features without losing the ability to express differences; the requirement that the window set covers chromosomes; the need for computations to be performed in a reasonable time).

During training we took $\alpha(t)=100/t$. The square neighbourhood, $N_w(t)$ decreased from 7x7 units to 1x1 units over the first 1000 steps, remaining at 1x1 thereafter. Over 10,000 steps were performed in all, where a step corresponds to the presentation of a single image portion. The feature map for the trained network is shown in figure 4. It consists of blob and bar-like features. (The map provides for good reconstruction of images from their extracted features, indicating that the window set chosen is reasonable).

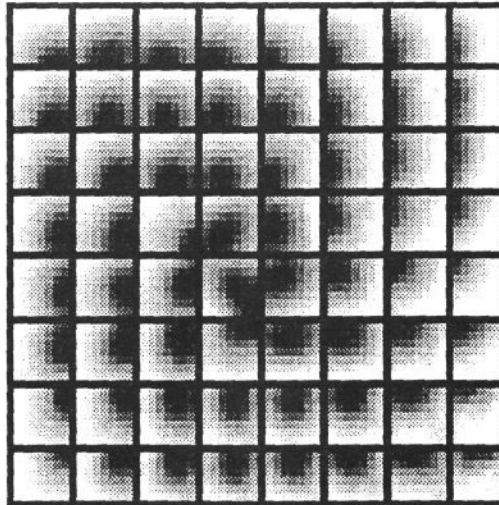


Figure 4. Feature map for low resolution images.

It is the responses of the trained Kohonen network to each image portion that determine the values on our activation map. The activation map is a 2-D sheet of binary units. The sheet is divided into 5x5 regions to mirror the organisation of the window set. We transform the Kohonen network responses produced for each image portion, setting the most responsive Kohonen unit to 1, all others to 0, and copy these responses onto the appropriate region of the map. Thus, the activation map identifies the best-matching feature for each of the image portions grabbed by the window set.

3.2 Auto-associative Networks.

The activation map is the input layer of two auto-associative neural networks - one involved in centring, the other in classification of centred objects (see figure 3). Associative networks learn to store input-output pairs of patterns during a period of training. When a noisy version of an input pattern is then presented to a network, the associated (noise-free) output pattern is retrieved (under certain conditions relating to storage capacity of the network). For auto-associative networks the input and output patterns are the same - networks learn to associate patterns with themselves. When a noisy pattern is presented to such a network the noise-free pattern is retrieved.

As stated, our auto-associative networks (which are similar to the Willshaw network (6)) have the activation map as their input layer. Output layers are identical in structure to the activation map (i.e., 2-D sheets of 40x40 binary units, divided into 5x5 regions). We taught networks to store activation map states produced from images of "good" chromosomes (i.e., straight, uncluttered and non-overlapping). During training we simply clamped a network's output layer into the state of its input layer (i.e., the activation map), and created connections between an input unit-output unit pair if the units were on together.

On recall, we wish a network to output a stored activation map state given a noisy version as input. We set the output state of unit i which is in region S , to 1 if

$$\sum_j T_{ij} I_j > \sum_j T_{kj} I_j \quad (k \in S) \dots\dots\dots (3)$$

where I_j is the state of input unit j and $T_{ij}=1$ if input unit i and output unit j are connected, $T_{ij}=0$ otherwise. If the inequality is not satisfied the output of unit i is set to 0.

In effect, each of the best-matching features encoded in the input state votes for compatible features to be present in the output (or retrieved) state. For each region of the output layer the feature with the most votes is chosen. While our auto-associative memories can retrieve states which permit centring on, and classification of, relatively straight chromosomes in clutter, we are interested in their use as part of a technique for bent and overlapping chromosomes (see section 4).

3.3 Window Centring.

Initially, the activation map sets the input state of the associative network employed in centring. The output state of the network passes directly as input to a single-layer, supervised network of the type described in section 2. In this case each unit in the supervised network is responsible for signalling the presence of a chromosome's central region within a particular window. The network works in tandem with a second supervised network taking input directly from the image portions. This second network tries to ensure that the new centre for the window set suggested by the first is actually on chromosomal material rather than just to one side. Attempts are made to locate chromosome centres on new images by repeatedly cycling through the Kohonen, associative and supervised networks until a stable position is found (see figure 3). Each time the window set is moved a new activation map is produced, and fed on to the associative network.

3.4 Classification.

When centring has been achieved the current activation map state forms the input to the associative network associated with classification. The network output state is passed on as input to a "semi-supervised" Kohonen network for classification. By semi-supervised we mean that for each presented input we specify which row we would like the most responsive unit in the network to belong to. The specification is based upon the length of the chromosome within the window set. The network is 6x6 units. (We require classification into one of six lengths at this resolution). During training the network self-organised such that the locations of the responses within a row became ordered with respect to chromosome orientation.

4. Discussion.

Our system is capable of making good estimates of chromosome position, length and orientation from low resolution images of chromosomes provided chromosomes are reasonably straight. This applies in the presence of clutter. However, cells examined for structural aberrations usually contain at least some bent chromosomes, and possibly overlaps. We are currently addressing the problems of overlapping and bent chromosomes. We are investigating a technique which prevents features from appearing in the output of an associative network unless they are "similar enough" to the features extracted directly from the image. The similarity measure is a distance on the Kohonen map, and increases over time. Also, we are looking at ways of suppressing appropriate image features once one classification of an overlap has been performed in order that a second classification may be made.

We intend to apply the centring and classification system described here to higher resolution images. The appearance of chromosomal objects at such resolutions would be constrained by the estimates of position, size and orientation obtained at the low resolution. We would expect a Kohonen network to extract features associated with structurally aberrant chromosomes, such as gaps and breaks.

References

- [1] M.G. Daker. Automation in cytogenetics. *Medical Laboratory Sciences*, V45, 324-332, 1988.
- [2] M.Turner, J.Austin, N.Allinson, P.Thompson. Automated Chromosome Analysis. ACAG Memo 36, Computer Science Department, York University, 1991.
- [3] J.Bille, A.Erhardt, G.Johannsen and B.Ueberreite. Biological dosimetry by aberration scoring using POLYP. Proceedings of the 6th International Conference on Pattern Recognition, 1200, IEEE Computer Society, Munich, 1982.
- [4] R.K.Aggarwal and K.S.Fu. Automated Recognition of Irradiated Chromosomes. *Journal of Histochemistry and Cytochemistry*, V22, N7, 561-568, 1974.
- [5] T.Kohonen. The Self-Organising Map. Proceedings of the IEEE, V78, N9, 1464-1480, 1990.
- [6] D.J.Willshaw, O.P.Buneman, H.C.Longuet-Higgins. Non-Holographic Associative Memory. *Nature* V222, 960-962, 1969.