

2005

Artificial neural networks : A comparative study of implementations for human chromosome classification

Nancy Akl
Edith Cowan University

Follow this and additional works at: https://ro.ecu.edu.au/theses_hons



Part of the [Artificial Intelligence and Robotics Commons](#)

Recommended Citation

Akl, N. (2005). *Artificial neural networks : A comparative study of implementations for human chromosome classification*. https://ro.ecu.edu.au/theses_hons/1048

This Thesis is posted at Research Online.
https://ro.ecu.edu.au/theses_hons/1048

Edith Cowan University

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.
- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author's moral rights contained in Part IX of the Copyright Act 1968 (Cth).
- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

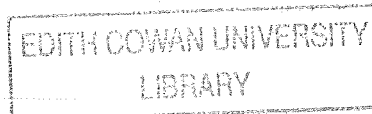
**ARTIFICIAL NEURAL NETWORKS: A
COMPARATIVE STUDY OF IMPLEMENTATIONS
FOR HUMAN CHROMOSOME CLASSIFICATION**

**Nancy Akl
Bachelor of Science (Computer Science)**

**This thesis is presented in fulfilment of the requirements for the
degree of Bachelor of Computer Technology (Honours)**

**Faculty of Regional Professional Studies
Edith Cowan University**

November, 2005



USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

ABSTRACT

Artificial neural networks are a popular field of artificial intelligence and have commonly been applied to solve many prediction, classification and diagnostic tasks. One such task is the analysis of human chromosomes. This thesis investigates the use of artificial neural networks (ANNs) as automated chromosome classifiers. The investigation involves the thorough analysis of seven different implementation techniques. These include three techniques using artificial neural networks, two techniques using ANNs supported by another method and two techniques not using ANNs. These seven implementations are evaluated according to the classification accuracy achieved and according to their support of important system measures, such as robustness and validity. The results collected show that ANNs perform relatively well in terms of classification accuracy, though other implementations achieved higher results. However, ANNs provide excellent support of essential system measures. This leads to a well-rounded implementation, consisting of a good balance between accuracy and system features, and thus an effective technique for automated human chromosome classification.

DECLARATION

I certify that this thesis does not, to the best of my knowledge and belief:

- (i) incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution of higher education.
- (ii) contain any material previously published or written by another person except where due reference is made in the text; or
- (iii) contain any defamatory material.

I also grant permission for the Library at Edith Cowan University to make duplicate copies of my thesis as required.

Signature:

Date: 15/11/03

ACKNOWLEDGEMENTS

Firstly, I would like to thank my family. My parents have been truly wonderful in helping me through this year. Their support and encouragement has been invaluable. My sisters also deserve a big thank you. They have at times seemed distracting, but in hindsight have helped in keeping me grounded and spirited. My grandparents: who have been visiting from overseas yet have been content with only watching me type all day and night.

Sam and Angelo: I would not have made it through, with my sanity intact, without you guys. Thanks for sharing the fun, the stress, "the closet" and most importantly the friendship.

Also, I would like to thank God. This journey has been difficult, at the best of times, yet my faith has helped me overcome the obstacles of the every day.

I would like to thank two lecturers who have been an enormous support during my tertiary education. Firstly, Lindsay, for his supervision of this research. His comments, feedback and guidance were essential in shaping the research and this thesis. Secondly, Viv, as her initial guidance and advice had set me into this course and for her continued support and encouragement throughout the years.

And finally, a big thank you to all my friends, family, lecturers and colleagues not mentioned here.

TABLE OF CONTENTS

USE OF THESIS.....	II
ABSTRACT.....	III
DECLARATION	IV
ACKNOWLEDGEMENTS	V
1. INTRODUCTION.....	1
1.1. INTRODUCTION TO THE RESEARCH.....	1
1.2. PURPOSE OF THE STUDY.....	1
1.3. HYPOTHESIS AND RESEARCH QUESTIONS	2
1.4. SCOPE	3
1.5 DOCUMENT STRUCTURE	3
2. BACKGROUND	4
2.1. CHROMOSOME CLASSIFICATION	4
2.1.1. <i>History</i>	4
2.1.2. <i>Introduction to Chromosomes</i>	5
2.1.2.1. Structure.....	6
2.1.2.2. Chromosome Groups	6
2.1.3. <i>Chromosome Databases</i>	8
2.2. ARTIFICIAL NEURAL NETWORKS	9
2.2.1. <i>Introduction to Artificial Neural Networks</i>	9
2.2.2. <i>Neural Network Structure</i>	9
2.2.3. <i>Neural Network Activation</i>	11
2.2.4. <i>Neural Network Learning</i>	11
2.2.4.1. Supervised Learning.....	11
2.2.4.2. Unsupervised Learning	12
3. LITERATURE REVIEW	13
3.1. CHROMOSOME KARYOTYPING	13
3.1.1. <i>History</i>	13
3.1.2. <i>Process</i>	14
3.1.3. <i>Problems with Manual Labour</i>	16
3.2. COMPUTERISED CHROMOSOME ANALYSIS	17
3.2.1. <i>Process</i>	17
3.2.1.1. Image Enhancement	17
3.2.1.2. Chromosome Segmentation	18
3.2.1.3. Feature Selection.....	19
3.2.1.4. Chromosome Classification	20
3.2.2. <i>Context-Free and Context-Dependant Classification</i>	20
3.3. ARTIFICIAL INTELLIGENCE.....	21
3.3.1. <i>Introduction to Artificial Intelligence</i>	21
3.3.2. <i>Medical Artificial Intelligence</i>	22
3.4. ARTIFICIAL NEURAL NETWORKS	23
3.4.1. <i>ANNs in Medical Decision Support</i>	23
3.4.2. <i>ANNs in Chromosome Classification</i>	24
3.4.3. <i>Strengths and Limitations</i>	25
3.5. LIMITATIONS	26
3.5.1. <i>Limitations of current literature</i>	26
4. METHODOLOGY	28
4.1. PURPOSE	28
4.2. RESEARCH FRAMEWORK.....	28
4.3. RESEARCH DESIGN.....	29
4.4. DATA COLLECTION	30

4.4.1. <i>First Case Study Group</i>	31
4.4.2. <i>Second Case Study Group</i>	31
4.4.3. <i>Third Case Study Group</i>	32
4.5. DATA ANALYSIS	32
4.6. VALIDITY	33
4.4.1. <i>Internal Validity</i>	33
4.4.2. <i>External Validity</i>	34
5. ANALYSIS	35
5.1. ARTIFICIAL NEURAL NETWORKS	35
5.1.1. <i>First Case Study</i>	35
5.1.1.1. Implementation Details	36
5.1.1.2. Results.....	37
5.1.2. <i>Second Case Study</i>	38
5.1.2.1. Implementation Details	38
5.1.2.2. Results.....	38
5.1.3. <i>Third Case Study</i>	39
5.1.3.1. Implementation Details	40
5.1.3.2. Experiments	41
5.1.3.3. Results.....	43
5.2. ARTIFICIAL NEURAL NETWORKS SUPPORTED BY OTHER TECHNIQUES.....	43
5.2.1. <i>First Case Study</i>	44
5.2.1.1. Implementation Details	44
5.2.1.2. Results.....	46
5.2.2. <i>Second Case Study</i>	47
5.2.2.1. Implementation Details	48
5.2.2.2. Results.....	49
5.3. NON-ANN TECHNIQUES	51
5.3.1. <i>First Case Study</i>	51
5.3.1.1. Implementation Details	51
5.3.1.2. Results.....	53
5.3.2. <i>Second Case Study</i>	53
5.3.2.1. Implementation Details	54
5.3.2.2. Results.....	54
6. RESULTS	56
6.1. PERFORMANCE COMPARISONS.....	56
6.1.1. <i>Implementations using ANNs</i>	56
6.1.1.1. Context-Free Classification.....	56
6.1.1.2. Context-Dependant Classification.....	57
6.1.2. <i>Implementations using ANNs with a Supporting Technique</i>	58
6.1.2.1. First Case Study	58
6.1.2.2. Second Case Study.....	59
6.1.3. <i>Implementations not using ANNs</i>	60
6.1.3.1. First Case Study	60
6.1.3.2. Second Case Study.....	60
6.1.4. <i>Outcome</i>	61
6.2. SYSTEM MEASURES COMPARISONS	63
6.2.1. <i>Ability to Generalise</i>	63
6.2.2. <i>Robustness</i>	64
6.2.3. <i>Efficiency in Computation Burden and Speed</i>	65
6.2.4. <i>Validity in Real-World Data</i>	66
6.2.4.1. Validity of Testing Data.....	66
6.2.4.2. Validity of Testing Experiments	67
6.2.5. <i>Degree of Human Reliance</i>	68
6.2.6. <i>Outcome</i>	69
6.3. BEST CLASSIFIER PERFORMANCE.....	70
6.3.1. <i>Best Classifier Implementation</i>	70
6.3.2. <i>Strengths and Limitations of this Technique</i>	72
7. CONCLUSIONS AND RECOMMENDATIONS	74
7.1. RESEARCH OUTCOMES.....	74
7.1.1. <i>Classification Accuracy</i>	74
7.1.2. <i>System Measures</i>	74
7.1.3. <i>Overall Assessment</i>	75

7.2. LIMITATIONS	75
7.3. RECOMMENDATIONS AND FURTHER WORK	76
7.3.1. <i>Imaging Techniques</i>	76
7.3.2. <i>Image Processing Techniques</i>	76
7.3.3. <i>Feature Selection</i>	77
7.3.4. <i>Network Architecture</i>	77
REFERENCES	79
APPENDIX A	85
A.1. RATINGS OF ANN IMPLEMENTATIONS	86
A.1.1. <i>First Case Study</i>	86
A.1.2. <i>Second Case Study</i>	88
A.1.3. <i>Third Case Study</i>	90
A.2. RATINGS OF IMPLEMENTATIONS USING ANNS SUPPORTED BY OTHER TECHNIQUES.....	92
A.2.1. <i>First Case Study</i>	92
A.2.2. <i>Second Case Study</i>	94
A.3. RATINGS OF NON-ANN IMPLEMENTATIONS	96
A.3.1. <i>First Case Study</i>	96
A.3.2. <i>Second Case Study</i>	98
APPENDIX B	100
B.1. GLOSSARY	100

1. INTRODUCTION

1.1. Introduction to the Research

Cytogenetics is defined as the study of chromosomes and their abnormalities (Jorde, Carey, Bamshad & White, 2000, p. 108) and is an important process in the diagnosis and treatment of human diseases (Keller, Gader, Sjahputera, Caldwell, & Huang, 1995, p. 125; Wang, et al., 2005, p. 2536). With cytogenetics being such a crucial and beneficial study, it has “evolved into a specialized discipline with widespread applications in both research and clinical practice, including prenatal screening, genetic counselling, oncology, radiation dosimetry and toxicology” (Carothers & Piper, 1994, p. 161). Keller, et al. (1995) support Carothers and Piper (1994) and state “human genetic investigations have provided some of the most dramatic progress in medicine in recent times” (p. 125).

Chromosomes store the ‘blueprints’ of all features of every individual. Graham and Errington (2000) identify some important applications of chromosome analysis by observing that “analysis of the appearance of chromosomes is routinely undertaken in hospital laboratories, for example, for diagnosis of inherited, or acquired, genetic abnormality or the monitoring of cancer treatment” (p. 249). The smallest error or abnormality within chromosomes often results in a larger and much more serious human irregularity. In order to identify these errors within chromosomes, cytogeneticists must often retrieve cell samples and organise the given chromosomes into their pre-determined groups. Aberrations are often identified by abnormalities in the structure of a chromosome or in the number of chromosomes found in the cell (Snustad & Simmons, 2000, p. 142).

1.2. Purpose of the Study

The traditional method of manual classification of chromosomes by a human expert presents several difficulties. These include the shortage of experts leading to an increase in workload for existing experts, the large amount of time required to perform such a tedious and detailed task and the costs associated with such manual classifications, (Lerner, 1998, p. 544).

Computerised decision support systems aim to solve many of the problems outlined above. “The automatic chromosome classification is an essential component of such systems, since it helps to reduce the tedium and labour-intensiveness of traditional methods of chromosome analysis” (Martínez, Juan & Casacuberta, 2002, p. 565). By using a computerised decision support system, it is therefore arguable that the time of classification is significantly reduced and the workload of experts is decreased thus effectively decreasing costs. One method for automating chromosome analysis is through the use of artificial neural networks (ANNs). ANNs are a subset of the artificial intelligence (AI) field of computer science and have been widely applied in problems involving prediction, classification and image recognition (Patterson, 1996). This study will investigate the use of artificial neural networks as automatic human chromosome classifiers.

1.3. Hypothesis and Research Questions

This study has been based on the hypothesis: *Artificial neural networks are an effective technique for classifying human chromosomes. They perform better than implementations that do not use artificial intelligence.*

The following research question has framed this study: *Are artificial neural networks a suitable implementation technique for automated chromosome analysis?* To add further depth and structure to the research, two sub-questions have been identified:

1. *How do ANNs perform in classification accuracy as compared to other implementations?, and*
2. *How do ANN classifiers perform in system measures as compared to classifiers based on other processing methods?* System measures, in this case, refer to factors such as the ability to generalise, robustness, efficiency in computational burden and speed, validity in real-world data and degree of human interaction required.

By addressing these questions, the research will conduct a rigorous analysis of the different implementations of artificial neural networks in human chromosome classification.

1.4. Scope

To thoroughly inform the research and to test the above hypothesis, this research will consider different implementations of techniques using ANNs and of techniques not using ANNs. The case studies to be considered will include:

- three implementations of artificial neural networks;
- two implementations using artificial neural networks supported by another technique; and
- two implementations of a technique not using artificial intelligence.

The case studies provide a general representation of the various implementation techniques available for chromosome classification. As such, this research aims not only to explore the use of artificial neural networks, but also the use of other contending techniques.

1.5 Document Structure

An introduction to the basic concepts of chromosome analysis and artificial neural networks will be presented in Chapter 2. Chapter 3 presents a literature review examining the previous and current issues of computerised chromosome analysis. The methodology adopted for this research will be discussed in Chapter 4. The analysis of the chosen case studies will be presented in Chapter 5, followed by a discussion of the results in Chapter 6. Chapter 7 brings to light the conclusions gained from this research and recommendations for future work in this area.

2. BACKGROUND

This chapter presents an introduction to the basic concepts of chromosome classification and neural networks. The chapter is organised into two sections addressing these important topics. Each section will present an introduction to the topic and discuss the main characteristics of both chromosomes and neural networks.

2.1. Chromosome Classification

2.1.1. History

Although interest in the science of genetics and trait inheritance has existed for thousands of years, significant observations only came about in the middle of the 19th century (Emery & Mueller, 1988, p. 1; Snustad & Simmons, 2000, p. 4; Jorde, Care, Bamshad & White, 2000, p. 1). In 1865, Gregor Mendel, an Austrian monk, achieved the first scientifically valid discovery of inheritance in living beings (Snustad & Simmons, 2000, p. 4). Emery and Mueller (1988) report that “Mendel made his far-reaching discoveries through careful and painstaking analysis of the results of crossing varieties of garden pea” (p. 2). These experiments led Mendel to suggest that “every cell contained pairs of ‘factors’ and that each pair determined a specific trait” (Snustad & Simmons, 2000, p. 4). These factors represent what is now known as genes (Snustad & Simmons, 2000, p. 4).

However, Mendel’s results were not recognised until 1900, when further understanding was gained on cell structure and division, which in turn facilitated the interpretation of Mendel’s results (Snustad & Simmons, 2000, p. 4). From that time, the study of genetics was enhanced and several developments followed. One such development occurred in 1944, when “Oswald Avery showed that genes are composed of DNA (deoxyribonucleic acid)” (Jorde, et al., 2000, p. 3). Following this breakthrough, James Watson and Francis Crick identified the physical structure of DNA in 1953 and completed the picture of inheritance and molecular genetics (Jorde, et al., 2000, p. 3). Another important development was the identification of the correct number of chromosomes in a normal human cell; it was believed that there were 48 chromosomes until 1956, when the correct number of 46 was established (Emery & Mueller, 1988, p. 12). The process of chromosome classification became very popular and was facilitated by technological developments in the 1960s (Jorde, et al., 2000, p. 3).

2.1.2. Introduction to Chromosomes

Every detail of a living being is represented by material called DNA (deoxyribonucleic acid), which is arranged and stored in sections referred to as genes (Snustad & Simmons, 2000, p. 17). These genes are arranged in bodies known as chromosomes. The name *chromosome* arises from the Greek words *chromo*, meaning colour, and *soma*, meaning body, thus representing a coloured body (Jorde, et al., 2000, p. 6; Snustad & Simmons, 2000, p. 27).

The molecular substance of a chromosome consists of chromatin, which contains DNA material, chromosomal proteins and other constituents from the cell nucleus (Snustad & Simmons, 2000, p. 235). Chromatin also gives the chromosomes its structure; Jorde, et al. (2000) state “just before a cell undergoes division, the chromatin condenses to form discrete, dark-staining bodies called chromosomes” (p. 6). Figure 2.1 gives a visual representation of a highly magnified chromosome.

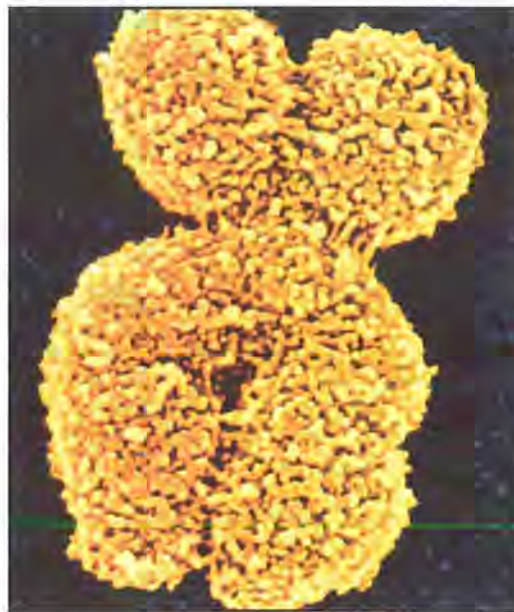


Figure 2.1. A highly magnified chromosome (Snustad and Simmons, 2000, p. 190)

Apart from common external factors, such as hair colour, eye colour and other physical features, genes within a chromosome may also represent abnormalities (Jorde, et al., 2000, p. 6). Abnormalities generally occur due to an anomaly in a single chromosome structure, chromosome number, or in a cluster of chromosomes (Snustad & Simmons, 2000, p. 142). By studying chromosomes within cell samples, cytogeneticists are able to identify possible abnormalities and where available

recommend treatments. In analysing chromosomes within a cell, cytogeneticists focus on the structure of the chromosomes and assigning the chromosomes into groups.

2.1.2.1. Structure

The features visible on the chromosomes play an important role in chromosome analysis. Levitan (1988) indicates that “chromosomes can generally be differentiated in three ways by (1) length, (2) position of the centromere, and (3) staining characteristics” (p. 24). The centromere represents the area of the chromosome where the two chromatid sisters overlap, thus forming a constriction. The centromere divides the chromosome into a shorter length and a longer length, commonly referred to as p-arm and q-arm respectively (Snustad & Simmons, 2000, p. 141). Figure 2.1, above, also shows the two chromatids, the central restriction representing the centromere and the short and long arms of each chromatid.

The banding pattern, or staining characteristic, is another important distinguishing feature of a chromosome as it “helps greatly in the detection of deletions, duplications and other structural abnormalities, and it facilitates the correct identification of individual chromosomes” (Jorde, et al., 2000, p. 111). Levitan (1988) defines a band as “a part of a chromosome that is clearly distinguishable from its adjacent segments by appearing darker or lighter as a result of the new staining methods” (p. 32). The banding pattern generally identifies the number of bands in the chromosome, the distance between each band, the distance between the bands and the centromere region and the density of each band (Keller, Gader, Sjahputera, Caldwell, & Huang, 1995, p. 127). By using the chromosome length, centromere position and banding patterns, cytogeneticists are able to facilitate the process of chromosome classification into groups.

2.1.2.2. Chromosome Groups

Before appropriate image analysis and dyeing techniques were available, it was difficult to identify matching chromosomes and thus the chromosomes within a cell were first organised into seven groups by their sizes (Snustad and Simmons, 2000, p. 141). Snustad and Simmons (2000) further describe the difficulty of chromosome analysis by stating:

Cytogeneticists could only arrange the chromosomes into groups according to size, classifying the largest group as A, the next largest as group B, and so forth. Although they could recognize seven different groups, within these groups it was nearly impossible to identify a particular chromosome. (p. 141)

These seven groups are commonly referred to as the Denver groups, as they were first acknowledged at a medical conference in Denver in 1960 (Levitan, 1988, p. 28). Nowadays, cytogeneticists may still arrange the chromosomes into their seven size groups and then determine the matching, or homologue, chromosomes within each group. A human somatic cell (a non-reproductive cell) contains 46 chromosomes arranged into 23 pairs where one of the 23 pairs consists of the sex chromosomes, which are an X and a Y chromosome in males, or two X chromosomes in females (Jorde, et al., 2000, p. 6). Table 2.1 presents the seven Denver groups and the chromosome classes belonging to each group.

Table 2.1. The seven Denver groups and the chromosome classes belonging to each group

Chromosome Group	Chromosome Classes
Group A	1 - 3
Group B	4 - 5
Group C	6 - 12 and X chromosome
Group D	13 - 15
Group E	16 - 18
Group F	19 - 20
Group G	21 - 22 and Y chromosome

Figure 2.2 gives a visual representation of the chromosomes arranged into their respective Denver groups and in chromosome classes within these groups.

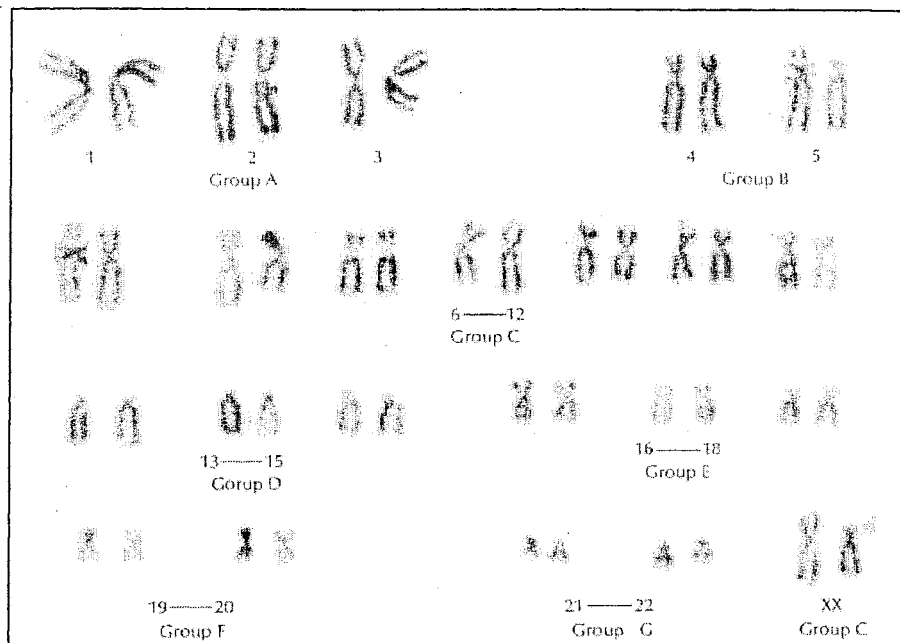


Figure 2.2. Chromosome karyotype showing Denver groups and classes within these groups (Levitan, 1988, p. 27)

2.1.3. Chromosome Databases

Three common databases are used for testing chromosome karyotyping systems. These are the Copenhagen database, the Edinburgh database and the Philadelphia database. These databases are used in several case studies presented by this research and thus will be briefly discussed.

The Copenhagen database was collected and developed at the Rigshospitalet, in Copenhagen, by Lundsteen and Granum in 1976 – 1978 (Piper & Granum, 1989, p. 243; Sweeney, et al., 1994, p. 19 – 20). Graham and Errington (2000) note that the images of the Copenhagen database were developed from “photographic negatives of selected cells of good appearance. Chromosomes involved in touches or overlaps were rejected from the data-set, so the visual ‘quality’ of the chromosomes was high” (p. 251).

The Edinburgh database was developed by Piper in Edinburgh in 1984 (Piper & Granum, 1989, p. 243; Sweeney, et al., 1994, p. 20). Graham and Errington (2000) claim that the images in the Edinburgh database were digitised from photographic images of cell material and were selected to have few overlapping chromosomes, thus resulting in good quality data (p. 251).

The final database is the Philadelphia database collected at the Jefferson Medical College in Philadelphia in 1987 (Piper & Granum, 1989, p. 243; Sweeney, et al., 1994, p. 20). Graham and Errington (2000) argue that the preparation techniques used for cell culture in the Philadelphia database have led to poor visual quality in the chromosome images (p. 251). Table 2.2 shows the three different chromosome databases and presents the number of cells, number of chromosomes and quality of images for each database.

Table 2.2. Chromosome databases and their contents (adapted from Sweeney, et al., 1994, p. 20)

Database	Number of Cells	Number of Chromosomes	Quality of Images
Copenhagen	180	8106	Good
Edinburgh	125	5548	Fair
Philadelphia	130	5847	Poor

2.2. Artificial Neural Networks

2.2.1. Introduction to Artificial Neural Networks

Patterson (1996) defines artificial neural networks as “simplified models of the central nervous system. They are networks of highly interconnected neural computing elements that have the ability to respond to input stimuli and to learn to adapt to the environment” (p. 1). Over the years, researchers have been evolving artificial neural networks based on their biological counterparts. Patterson (1996) affirms that “much of the research work in ANNs has been inspired and influenced by our knowledge of biological nervous systems” (p. 6). However, artificial neural networks have not yet achieved full similarity to a human neural network. Negnevitsky (2002) states “a present-day artificial neural network (ANN) resembles the human brain much as a paper plane resembles a supersonic jet” (p. 165). Despite their limitations in resembling biological networks, artificial neural networks have been successfully applied to several complex problems including forecasting, diagnosis, scheduling and pattern and image recognition (Patterson, 1996). Several key factors of artificial neural networks will now be discussed, including their structure, activation and learning methods.

2.2.2. Neural Network Structure

Negnevitsky (2002) defines an artificial neural network as “a model of reasoning based on the human brain” (p. 164). The structure of an artificial neural network is

based on the structure of a biological neural network. Figure 2.3 illustrates a biological neural network containing two neurons. Figure 2.4 depicts a three-layer artificial neural network. The resemblance between the two networks is not easy to discern. However, these diagrams show that the flow of information in a neural network resembles the flow of signals in a biological neural network.

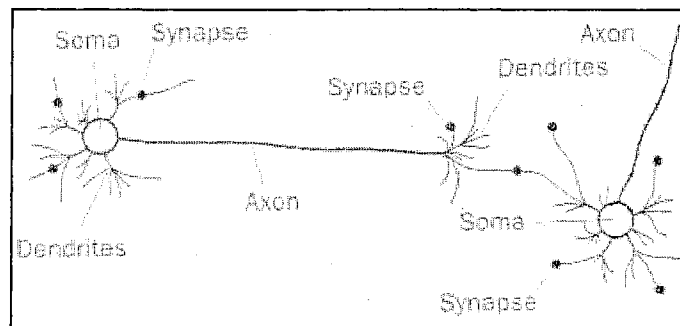


Figure 2.3. A biological neural network (Negnevitsky, 2002, p. 164)

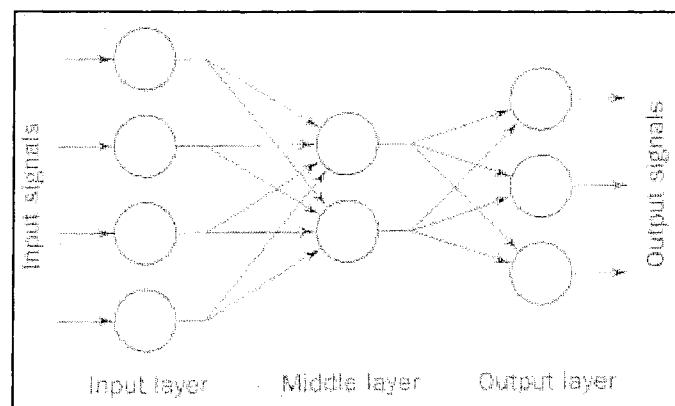


Figure 2.4. An artificial neural network (Negnevitsky, 2002, p. 165)

An artificial neural network consists of several main processing nodes called neurons. These neurons are typically arranged in at least three layers (Negnevitsky, 2002, p. 173):

1. Input layer: the purpose of the input layer is to accept input signals and to redistribute these signals to the neurons in the hidden layer.
2. Hidden layer(s): a neural network architecture often contains one hidden layer; however, some complex functions require more than one hidden layer. It is customary to keep the number of hidden layers to a minimum since “each additional layer increases the computational burden exponentially” (Negnevitsky, 2002, p. 174). The hidden layer is required to detect the features from the input signals and propagate these features to the output layer.

3. Output layer: the purpose of the output layer is to present the output of the neural network's computations.

2.2.3. Neural Network Activation

Artificial neural networks function by accepting input signals. Each input is weighted by the connection strength before reaching the processing neuron (Kartalopoulos, 1996, p. 40). The neuron then calculates the sum of these weighted input signals and the result is compared to a threshold value (Kartalopoulos, 1996, p. 40). The output of the neuron is then dependant upon whether the sum of input signals is greater or less than the threshold value (Negnevitsky, 2002, p. 167). An artificial neural network does not automatically know the correct output to produce when faced with different input stimuli. Instead, the network must gradually learn the output required through a series of small adjustments of the neuron weights (Negnevitsky, 2002, p. 169). This process depicts the learning process of an artificial neural network and is described in detail below.

2.2.4. Neural Network Learning

The term *learning* is commonly used to represent the process that an artificial neural network undertakes when it is faced with new input stimuli. "Learning is the process by which the neural network adapts itself to a stimulus, and eventually (after making the proper parameter adjustments to itself) it produces a desired response", (Kartalopoulos, 1996, p. 43). There are two main methods of learning: supervised and unsupervised learning.

2.2.4.1. Supervised Learning

Supervised learning involves the use of a desired output, or correct answer. The neural network must continually adjust its outputs until the actual output reaches the desired output. To begin learning, the weights of the connections in a network are randomly assigned from a predetermined range of values (typically between -0.5 to 0.5). The network learning is done by "making small adjustments in the weights to reduce the difference between the actual and desired outputs" (Negnevitsky, 2002, p. 169). The weight is either negatively or positively adjusted depending on the variance between the actual and desired outputs. This weight adjustment value is always a pre-set value known as the learning rate. The learning rate of a neural network plays a crucial role in

the time taken for the network to correctly learn its tasks. A large learning rate value may allow the network to learn quicker, but the network may never arrive at the desired response. On the other hand, a small learning rate value would result in a longer training period but would produce better results (Negnevitsky, 2002, p. 184).

The most popular network training method is the back-propagation method (Negnevitsky, 2002, p. 174). The back-propagation algorithm consists of two phases: forward flow of signals from input to output neurons, and a backward flow of weight adjustments from output to input neurons. Negnevitsky (2002) describes this process as:

First, a training input pattern is presented to the network input layer. The network then propagates the input pattern from layer to layer until the output pattern is generated by the output layer. If this pattern is different from the desired output, an error is calculated and then propagated backwards through the network from the output layer to the input layer. The weights are modified as the error is propagated. (p. 174)

2.2.4.2. Unsupervised Learning

In unsupervised learning or self-organised learning, the network is not presented with the desired output. Kartalopoulos (1996) describes this learning method as:

During the training session, the neural net receives at its input many different excitations, or input patterns, and it arbitrarily organizes the patterns into categories. When a stimulus is later applied, the neural net provides an output response indicating the class to which the stimulus belongs. If a class cannot be found for the input stimulus, a new class is generated. (p. 44)

However, Kartalopoulos (1996) notes that although no desired target is set, the network is still given guidelines on how to discriminate between signals and how to form groups (p. 44). Kartalopoulos (1996) continues “if no guidelines have been given as to what type of features should be used for grouping the objects, the grouping may or may not be successful” (p. 45). Negnevitsky (2002) “unsupervised learning algorithms aim to learn rapidly. In fact, self-organising neural networks learn much faster than back-propagation networks, and thus can be used in real time” (p. 198).

The following chapter presents a literature review on several important facets of automated chromosome analysis, artificial intelligence in medicine and specifically the use of artificial neural networks.

3. LITERATURE REVIEW

This chapter presents a discussion of the history and process of chromosome analysis and investigates the limitations of manual karyotyping. The process of automated karyotyping is introduced and described in detail. The computer science field of artificial intelligence is introduced, with a focus on artificial neural networks and what this offers to the task of chromosome classification.

3.1. Chromosome Karyotyping

3.1.1. History

A fundamental task of chromosome analysis is karyotyping. “The visual analysis of chromosome images, known as karyotyping, involves counting the chromosomes and examining them for structural abnormalities” (Graham & Errington, 2000, p. 250). The first successful attempt at chromosome analysis was in 1882, by Flemming, who used basic dyes on human tissue to view the chromosomes (Winchester & Mertens, 1983, p. 8). An improvement in the karyotyping technique came about in 1956, when Tjio and Levan conceived the method of pressing cells to flatten and spread the cell contents and increase visibility of the individual chromosomes (Winchester & Mertens, 1983, p. 9). The full potential of this technique was eventually realised when Tjio in association with Puck were able to develop cell culturing techniques, which facilitated the access to human chromosomes (Winchester & Mertens, 1983, p. 10). Figure 3.1 shows a highly magnified image of chromosomes from a human metaphase cell.



Figure 3.1. Image of chromosomes from a human metaphase cell (Lerner, 1998, p. 544)

The traditional method of karyotyping involves culturing metaphase cells, photographing these cells, making paper cut-outs of the individual chromosomes and then arranging these chromosomes into pairs and assembling in order by size (De Robertis & De Robertis, 1980, p. 439; Winchester & Mertens, 1983, p. 11). Wang, et al. (2005) claim that “karyotyping is the most common procedure for analysing and classifying banded chromosomes from images of a metaphase cell” (p. 2536). This is due to the end product of karyotyping, which “defines the number and arrangement, size and structure of the chromosomes and assigns each chromosome to one of the 24 human chromosome classes” (Wang, et al., 2005, p. 2536-2537). This karyotype displays the chromosomes arranged in pairs and by size and therefore helps cytogeneticists identify missing or abnormal chromosomes. Martínez, Juan and Casacuberta (2002) emphasise that “producing a karyotype of a cell is of practical importance since it greatly facilitates the detection of abnormalities in the chromosome structure” (p. 565). A karyotype of human chromosomes is shown in figure 3.2.

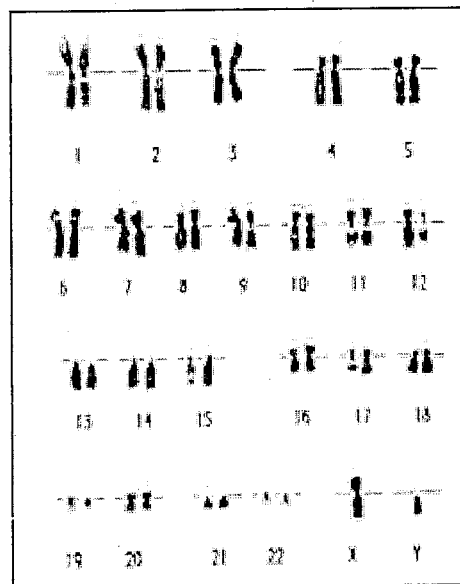


Figure 3.2. A karyotype of human chromosomes from a metaphase cell (Lerner, 1998, p. 545)

3.1.2. Process

This section discusses the process of karyotyping and the common abnormalities found in chromosome karyotypes. The phase at which the cells are most suitable for karyotyping is the metaphase stage (Lerner, 1998, p. 544). The metaphase stage is the second main step in cell division. During this stage, the chromatids have been duplicated and are now attached through the centromere. The popular use of this cell stage for karyotyping is due to the structure of the chromosomes at that stage. Snustad

and Simmons (2000) note that “metaphase chromatids are tightly coiled and discrete, thus facilitating accurate chromosome counts and gross structural analysis” (p. 31). To prepare cell samples for analysis, cytogeneticists commonly stimulate the cells to start division until the metaphase stage is reached. The cell division is then arrested through the use of specified chemicals to prevent further division (Snustad & Simmons, 2000, p. 140). Once the cell culture has been prepared, chromosome analysis begins with the identification of the required chromosome features, such as banding patterns.

The banding patterns of a chromosome generally become visible through the use of certain dyes. Levitan (1988, p. 32) explains:

In the late 1960s and early 1970s new staining techniques were discovered that have made it possible for human cytogenetics not only to specify every chromosome but even, in many cases, to identify exactly parts of chromosomes that had been moved to unusual locations in the genome.

This ability is supported by the banding techniques now available, which identify that each chromosome has a unique banding pattern (Levitan, 1988, p. 32). The several different banding techniques available include:

1. Q-Banding: This technique uses quinacrine, a fluorescent compound that highlights chromosome bands when exposed to ultraviolet light (Snustad & Simmons, 2000, p. 140).
2. Giemsa Banding: This technique uses the Giemsa dye, which also produces visible bands on the chromosome. Snustad and Simmons (2000) state that “the nature of the banding pattern depends on how the chromosomes were prepared prior to staining” (p. 141). The different banding methods include G-banding (Giemsa banding), which highlights dark bands similar to the Q-banding technique; R-banding (Reverse banding), which reverses the patterns seen in G-banding and Q-banding; and C-banding, which stains the centromere region of the chromosome (Jorde, Carey, Bamshad & White, 2000, p. 111; Snustad & Simmons, 2000, p. 141).

Each of these banding techniques highlights different patterns on the chromosomes. This allows cytogeneticists to “analyse fine details of chromosome structure (Snustad & Simmons, 2000, p. 141). Figure 3.3 displays human chromosomes stained using the R-banding technique.



Figure 3.3. R-Banding staining technique on human chromosomes (Snustad & Simmons, 2000, p. 141)

3.1.3. Problems with Manual Labour

The importance of chromosome analysis is illustrated by Cho (2000), who states:

Cytogenetic analysis of chromosomes is widely used in many hospitals for genetic diagnosis of fetuses, pregnant women, and nursing mothers, as well as in many genetic laboratories for research with animals and plants. Therefore, automatic chromosome analysis has attracted much attention due to its potential wide application and its importance. (p. 28)

Apart from the importance and applications of karyotyping, another difficulty in manual chromosome classification arises from its complex process. Carothers and Piper (1994) argue that “the need for automation arises from the fact that the ‘traditional’ (i.e. manual) methods of analysis are tedious and labour-intensive” (p. 161). Another reason behind the complexity in chromosome analysis lies in the method of collecting sufficient data. Carothers and Piper (1994) explain:

Because chromosomes are frequently lost or obscured during preparation, several cells must usually be analysed until the observer is satisfied as to their chromosome constitution, or ‘karyotype’. However, cells at the stage of division (metaphase) when the chromosomes are most easily analysed are relatively sparse, so that finding the required number may take time. (p. 161)

Therefore, analysing chromosomes frequently involves examining several different cells and creating multiple karyotypes in order to gain a full understanding of any abnormalities. This process is repetitive and extremely time-consuming; Carothers

and Piper (1994) allege that even an experienced cytogeneticist could take about an hour to carry out typical karyotype analysis (p. 161). Therefore, cytogeneticists have turned to computerised chromosome analysis systems to facilitate karyotyping. Martínez, Juan and Casacuberta (2002) identify the need for automatic karyotyping systems by arguing that “automatic chromosome classification is an essential component of such systems, since it helps to reduce the tedium and labour-intensiveness of traditional methods of chromosome analysis” (p. 565).

3.2. Computerised Chromosome Analysis

3.2.1. Process

The automated process of karyotyping was one of the earliest pattern recognition techniques to be computerised (Charters & Graham, 2002, p. 2080). The process of performing computerised chromosome analysis draws from the manual procedure for karyotyping. Wang, et al. (2005) identify the four main processing tasks involved in computerised karyotyping as “(1) image enhancement, (2) chromosome segmentation (detection) and alignment, (3) feature computation and selection and (4) chromosome classification” (p. 2537). Figure 3.4 gives a visual illustration of these tasks and these will be further discussed in separate sections below.

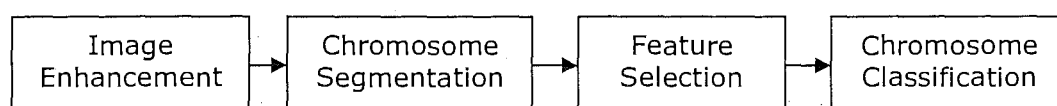


Figure 3.4. Four main tasks of automated karyotyping systems, adapted from Wang, et al. (2005, p. 2537)

3.2.1.1. Image Enhancement

The culturing of cells to the metaphase stage and the use of various staining and imaging techniques often add noise and external data to the cell image (Wang, et al., 2005, p. 2538). The classification accuracy of automated karyotyping systems is dependant on the quality of the data supplied. Thus, image enhancement is a vital task for improving image quality and therefore improving classification accuracy. Wang, et al. (2005) note “the aim of image enhancement is to improve visibility of low-contrast chromosomes (or related features) while suppressing noise” (p. 2538). Lerner (1998) supports the above by stating “the preprocessing stage aims to improve the quality of

the cell image by techniques of noise removal, edge enhancement and/or contrast improvement” (p. 545). Wang, et al. (2005) argue that “image enhancement improves not only the display and visualization of chromosome images but also the recognition rate and accuracy of chromosome classification” (p. 2538).

3.2.1.2. Chromosome Segmentation

Chromosome images commonly contain touching or overlapping chromosomes (Wang, et al., 2005, p. 2538). Therefore, chromosome segmentation is vital; however, researchers have found difficulty in fully automating this task. Wang, et al. (2005) point out that “finding solutions for automated separation of chromosomes is difficult yet vital” (p. 2538). Lerner (1998) identifies the difficulty by stating:

Most conventional image segmentation methods are based on either threshold selection, adaptive thresholding, edge detection or matching with a set of prototype shapes. However, almost all of these methods tend to fail or lose accuracy when considering complicated images or those of partially occluded objects as in the case of a chromosome image. (p. 546)

Lerner (1998) continues:

Consequently, it is not surprising that in most of the published works concerning chromosome analysis, *manually* segmented databases are used. Neither is it surprising to find that almost all the commercial ‘automatic’ chromosome analysis systems are in fact ‘semiautomatic’ and require a continuous interaction of the cytotechnician. (p. 546)

Popescu, et al. (1999) support Lerner (1998) and acknowledge that “commercially available automated karyotyping systems (AKS) are semiautomatic, requiring human intervention to perform certain tasks. These systems are typically unable to perform well with chromosomes that are overlapped” (p. 62).

One successful technique for chromosome segmentation is the use of knowledge-based chromosome contour searching (Wang, et al., 2005, p. 2538). This method uses edge detection, to remove random noise while preserving the chromosome edges, and contour tracking to identify the contours of connected segments (Wang, et al., 2005, p. 2538). Wang, et al. (2005) report that from a total of 124 touching and overlapping chromosomes, 82% of the clusters were successfully separated (p. 2539).

Lerner (1998) presents a varied technique, named the classification-driven partially occluded object segmentation (CPOOS) method (p. 547). This method consists of three stages: firstly identifying the pixels within the image; secondly, identifying clusters of chromosomes based on their size; and finally creating potential separating lines in the chromosome clusters (Lerner, 1998, p. 547-548). Lerner (1998) argues that this method is superior to the thresholding technique of edge detection as it eliminates the “tedious, usually unreliable experimentation with threshold selection” (p. 547). The results produced to show an improvement over edge detection; the CPOOS method correctly segmented 90% of clustered chromosomes with an 8.7% rejection rate when tested on 46 human cell images (Lerner, 1998, p. 550).

Another technique for chromosome segmentation is presented by Ji (1994, who proposes a recursive rule based segmentation procedure, “in which the rules adapt classification and segmentation parameters for each cell” (p. 197). Ji (1994) validates this approach by explaining:

In manual segmentation techniques, it is usually possible to split a big cluster into individual chromosomes in ‘one go’. By contrast, a single split in an automatic system will typically divide a cluster into just two new objects, and full decomposition will require recursive application of the algorithm. (p. 198)

This technique proposed by Ji (1994) achieved 95.2% correct segmentation accuracy when tested on 256 human cells and rejected only five cells.

3.2.1.3. Feature Selection

Following successful chromosome segmentation, the features required from each chromosome are collected. Wang, et al. (2005) define this stage as “a search, among all possible transformations (or extracted features), for the best subspace that preserves class separability as much as possible in the lowest possible dimensional space” (p. 2539). The common features used are length, centromeric index and banding profile. The length of a chromosome is often retrieved by extracting the skeleton of the chromosome image, and then calculating the length (Wang, et al., 2005, p. 2539). From the extracted skeleton, the centromeric index can also be computed (Wang, et al., 2005, p. 2539). The banding profile of a chromosome can be extracted by determining variances in the grey-level pixels of the chromosome image, which portray the density profile (Wang, et al., 2005, p. 2539).

3.2.1.4. Chromosome Classification

The final task in an automated karyotyping system is that of chromosome classification. The performance of this task is directly dependant on the performance of previous tasks. Chromosome classification uses the features extracted in the previous task to assign chromosomes to their respective groups. Wang, et al. (2005) state:

In order to improve the performance of automated chromosome classification (including recognition of disordered chromosomes), artificial intelligence and machine learning methods have been widely used in the computer-assisted chromosome detection and classification systems. (p. 2540)

3.2.2. Context-Free and Context-Dependant Classification

Two possibilities exist for chromosome classification: context-free and context-dependant classification. Context-free classification is defined by Carothers and Piper (1994) as “individual chromosomes are considered as independent objects, without regard to their context as components of a karyotype” (p. 164). Lerner (1998) supports the above by describing context-free classification as “the data set is classified as is and without *a posteriori* rearrangement of the chromosomes” (p. 550). This technique does not consider the fact that there should be two chromosomes in each class, and therefore assigns the chromosomes to their classes without considering matching or homologue chromosomes.

Context-dependant classification, on the other hand, takes into account the *a priori* knowledge that there should be two chromosomes in each class in a normal cell (Lerner, 1998, p. 550). This technique is usually applied as a global constraint, which is commonly referred to as the karyotyping constraint. Graham and Errington (2000) state that “it is possible to effect significant improvement on the classification of individual chromosomes by application of the *karyotyping constraint*, namely that there are exactly two chromosomes in (almost) all classes” (p. 258). This technique not only reduces error rates but also mimics the karyotyping process used in manual classification. Rutovitz (1977) and Piper, et al. (1980), cited by Tso and Graham (1983), observe that a human operator takes into account all chromosomes within a cell and knows at the outset how many chromosomes should be in each class (p. 489). Carothers and Piper (1994) support the above and point out that “human karyotypers rely strongly on between-chromosome comparison, and this has been shown to reduce error rates” (p. 165).

A popular method of implementing context-dependant classification is through the use of the transportation algorithm. The transportation algorithm is commonly used for finding the most economical route passing through predetermined destinations and is applied in cases such as the Travelling Salesman Problem (Patterson, 1996, p. 298), and in this case, context-dependant chromosome classification. Graham and Errington (2000) explain that “the chromosome classification problem is a special case of the Transportation Problem, in that the destinations (the individual chromosomes) all have a demand of unity on the sources (the chromosome classes)” (p. 258). The transportation algorithm is not limited to only normal cells but can also analyse cells with missing or extra chromosomes (Tso & Graham, 1983, p. 491).

The next section will provide a brief discussion on artificial intelligence. This discussion will include popular definitions and association of artificial intelligence in medical decision support, thus leading to medical artificial intelligence.

3.3. Artificial Intelligence

3.3.1. Introduction to Artificial Intelligence

A precise definition of AI is elusive, due to the fact that related terms are somewhat ambiguous themselves. Patterson (1990, p. 2) argues that a full understanding of artificial intelligence would require a precise explanation of related terms, such as intelligence, knowledge, reasoning etc., and that such precise scientific definitions are elusive (Patterson, 1990, p. 2). Patterson’s definition of artificial intelligence is presented as:

AI is a branch of computer science concerned with the study and creation of computer systems that exhibit some form of intelligence: systems that learn new concepts and tasks, systems that can reason and draw useful conclusions ... and systems that perform other types of feats that require human types of intelligence. (1990, p. 2)

Boden (1997) cited by Negnevitsky (2002, p. 2) presents a similar definition of artificial intelligence as “the goal of artificial intelligence as a science is to make machines do things that would require intelligence if done by humans”. The above authors all provide a common thread in the definition of AI: a computer application mimicking human intelligence. This constitutes a main difference between conventional computer applications and those using artificial intelligence. Patterson (1990) asserts:

AI is not the study and creation of conventional computer systems. Even though one can argue that all programs exhibit some degree of intelligence, an AI program will go beyond this in demonstrating a high level of intelligence to a degree that equals or exceeds the intelligence required of a human in performing some task. (p. 3)

This ability of artificial intelligence to mimic human intelligence has set it apart from other computing techniques. Several different AI techniques have been developed; artificial neural networks (ANNs) are one such technique.

3.3.2. Medical Artificial Intelligence

Clancey and Shortliffe (1984), cited by Coiera (1996), provide an early definition of medical artificial intelligence as “medical artificial intelligence is primarily concerned with the construction of AI programs that perform diagnosis and make therapy recommendations” (p. 363). However, Coiera (1996) claims that “today this definition would be considered narrow in scope and vision” (p. 363). This arises from the fact that medical intelligence today covers a much larger field than just diagnosis and recommendations. Therefore, although the above definition was appropriate for its time, it now appears to limit the full power of medical intelligence (Coiera, 1996, p. 363).

The following definition provides a more inclusive vision of medical informatics. Perry, Roderer and Assar (2005) paraphrase Frisse, Braude, Florance and Fuller (1995) and define medical informatics as:

Being at the crossroads between biomedical science and information technology, with a focus on developing and delivering information systems that support health care, decision making, databases for outcomes analysis and health sciences research and administration. (p. 220)

Although the above definition is geared towards medical informatics in general, it does apply for medical artificial intelligence. Medical AI aims to achieve all the goals defined above by using procedures similar to those used by human experts. Artificial neural networks are one of the AI techniques commonly used in medical applications.

3.4. Artificial Neural Networks

3.4.1. ANNs in Medical Decision Support

Medical decision support is inherently complicated, due to two main sources of difficulty, identified by Dybowski (2000, p. 26) as:

1. Workload: the number of experts within each specialised domain is not enough to manage the large load of complex data provided.
2. Complexity: medical data can be increasing complex, so that even an experienced specialist may overlook certain vital details.

Dybowski (2000) argues that the use of artificial neural networks is a natural choice for solving and alleviating these problems (p. 26).

Medical decision support systems aim to act and mimic the performance of a human expert. Using artificial intelligence for medical decision support systems has been popular due to the knowledge handling characteristics of AI systems. Patterson (1990) describes the importance of knowledge in AI systems, and stresses that the acquisition of knowledge, knowledge representation, knowledge organisation and knowledge manipulation are all important features of any AI system (p. 14-17). Patterson (1996) states “much of the research work in ANNs has been inspired and influenced by our knowledge of biological nervous systems” (p. 6). Apart from their biological influences, ANNs have several other characteristics lending to their use in chromosome classification. These characteristics include the ability to generalise and handle data it has not been previously exposed to, the ability to learn and retain new knowledge and the ability to handle uncertainty and noise in data (Negnevitsky, 2002, p. 250).

Artificial neural networks have been applied to many different facets of medical decision support. Popular implementations of neural networks include:

- Outcome prediction: Baxt (1995) notes “a major area of interest in health care policy is outcome prediction, and [artificial neural] networks have been used extensively for this purpose” (p. 1137). One such application is that of tumour behaviour prediction (Azuaje, et al., 1999; Catto, et al., 2003);
- Signal processing: Artificial neural networks have been used for analysing signal data for over a decade (Baxt, 1995, p. 1137). Signal processing

implementations include analysing electroencephalograph (EEG) signals and electrocardiograph (ECG) signals (Silipo & Marchesi, 1998; Kangas & Keller, 2000); and

- Image processing: artificial neural networks have been used in image processing applications such as cancerous cell classification (Zhou, Jiang, Yang & Chen, 2002) and analysis of myocardial infarction images (Lo, Lin, Freedman & Mun, 1998). The next section will discuss the use of artificial neural networks for the classification of chromosome images.

3.4.2. ANNs in Chromosome Classification

Computer science generally attempts to solve and automate problems that require extensive complex and repetitive processes. The classification of human chromosomes is one such problem. Carothers and Piper (1994) emphasise that the task of chromosome classification can be complex and tedious, due to the necessity of classifying several cells in order to complete a full karyotype when cell images are incomplete or unclear (p. 161).

Lisboa, Ifeachor and Szczepaniak (2000) support the above and argue that “automated image analysis and understanding is one of the most challenging areas in biomedical engineering, since there is usually considerable patient-to-patient variation in images pertaining to similar medical conditions, adding to the other sources of noise already present” (p. 211). Artificial neural networks are well suited for these problems since they have the ability to handle incomplete or imprecise data (Negnevitsky, 2002, p. 259), which is common in images with low clarity. Baxt (1995) supports this by stating “one of the areas to which artificial neural networks were first adapted was imaging, using both features extracted by human assistance and raw data from different radiological techniques” (p. 1136).

Between the various artificial intelligence techniques available, artificial neural networks have been very popular for the task of chromosome classification. Wang, et al. (2005) state:

Among them [AI techniques], artificial neural network is the most popular tool owing to its capability of modelling the human brain decision making process to recognize objects based on incomplete or

partial information, as well as its simple topographic structure and easier training process. (p. 2540)

3.4.3. Strengths and Limitations

The use of neural networks presents several advantages. Patterson (1996) argues that neural networks:

Exhibit a number of desirable properties not found in conventional symbolic computation systems including robust performance when dealing with noisy or incomplete input patterns, a high degree of fault tolerance, high parallel computation rates, the ability to generalize, and adaptive learning. (p. 2)

Negnevitsky (2002) supports Patterson (1996) by stating that neural networks perform well in areas involving imprecision and uncertainty in data, are easily adapted to incorporate new knowledge, have a good learning ability and are easily maintained when changes are necessary (p. 259).

Wasserman (1993) notes that “for a neural network to be useful, it must accommodate this variability, producing the correct output vector despite insignificant deviations between the input and test vectors. This ability is called generalization” (p. 3). Patterson (1996) paraphrases Sietsma and Dow (1991) and suggests that “networks generalize well when trained with noise distorted training patterns” and that “training with random noise can dramatically improve a network’s ability to correctly classify noisy inputs” (p. 207).

Although powerful in their processing, artificial neural networks do have several limitations. One such limitation is overfitting. Patterson (1996) explains that overfitting can develop when “a limited training set has been used repeatedly too many times in the training process” (p. 190). When this occurs, the neural network memorises its training examples and produces incorrect outputs when presented with new data (Negnevitsky, 2002, p. 223). This leads to a lack of generalisation. However, overfitting can be prevented through proper network architecture and training. Negnevitsky (2002, p. 323) states that “the practical approach to preventing overfitting is to choose the smallest number of hidden neurons that yields good generalisation”. This approach involves additional computations, since the network performance must be analysed with several different network architectures, but it does produce good results. Other approaches to prevent overfitting include terminating the training before the network begins to

memorise training data and using a sufficiently large training set (Patterson, 1996, p. 208).

Another problem arising from the structure of a neural network is the computational time required. By including more neurons and layers within the network, results produced can be more accurate but the training and execution time increases exponentially. Negnevitsky (2002) argues that “complex patterns cannot be detected by a small number of hidden neurons; however, too many of them can dramatically increase the computational burden” (p. 323). The back-propagation training algorithm also adds to the time consumption. Negnevitsky (2002) acknowledges that using back-propagation leads to extensive calculations, which in turn cause long training periods (p. 183).

Although the structure of the neural network is behind the ambition to resemble a biological neural network, it also presents a major limitation of this technique. A neural network is unable to explain or validate the outputs produced; this is known as the ‘black box’ characteristic. Dybowski (2000) notes “the manner in which a neural network derives an output value from a given feature vector is not comprehensible to the non-specialist, and this lack of comprehension makes the output from neural networks unacceptable” (p. 31). However, Dybowski (2000) does not discredit the use of ANNs altogether by arguing that the acceptance of the results produced by the artificial neural network would depend on the area in which it is used (p. 31).

3.5. Limitations

3.5.1. Limitations of current literature

As demonstrated throughout this literature review, artificial neural networks have a long history of being applied for automated chromosome classification. Although many of these research endeavours consider the results produced by other implementations, they do not perform a complete and unbiased comparison. The article presented by Carothers and Piper (1994) presents a review into automated chromosome analysis. However, the focus of their research is on the separate tasks involved in chromosome analysis rather than the different implementations available (Carothers & Piper, 1994, p. 161). Wang, et al. (2005) also present a study into automated systems for chromosome classification. This research is similar to that provided by Carothers &

Piper (1994) in that it reviews the methods involved in chromosome classification tasks, rather than providing an overview of the different implementation techniques used. The aim of this research is to examine several different implementation techniques currently in use for chromosome classification, to conduct a thorough investigation into the results produced by each implementation and to identify outstanding issues in automated chromosome analysis. The methodology used to frame this research is discussed next.

4. METHODOLOGY

4.1. Purpose

This chapter presents the methodology used in conducting this research. This discussion covers the research framework used, the research design implemented, the data collection and analysis strategies developed and strategies used for maintaining the validity of the research. The data selected for this research are also briefly introduced, followed by a thorough analysis of this data in Chapter 5.

4.2. Research Framework

This research has used aspects from both the qualitative and quantitative research frameworks, thus leading to a mixed method approach. Punch (1998) supports the use of a mixed method framework by stating that “at a general level, the reasons for combining are to capitalize on the strengths of the two approaches, and to compensate for the weaknesses of each approach” (p. 246). There are different strategies for combining the two theoretical approaches; in this research, the two theoretical designs will be combined in sequence. This approach is referred to by Punch (1998) as the *stage in the research process* and is defined as “quantitative and qualitative research may be appropriate to different stages of a longitudinal study” (p. 247). Creswell (2003) presents a similar approach named the *sequential exploratory strategy*, which details the strategy factors including the implementation strategy, priority of each design, integration of data analysed and the overall theoretical perspective (p. 211 – 213). Creswell (2003) defines this strategy as follows:

It is conducted in two phases, with the priority generally given to the first phase, and it may not be implemented within a prescribed theoretical perspective...this model is characterized by an initial phase of qualitative data collection and analysis, which is followed by a phase of quantitative data collection and analysis... The findings of these two phases are then integrated during the interpretation phase. (p. 215)

These similar approaches from Punch (1998) and Creswell (2003) illustrate the importance of identifying the correct framework in conducting a research project at both the theoretical and applied levels. This chapter will now consider the qualitative and quantitative design strategies used within this mixed-method approach.

4.3. Research Design

This research used the qualitative case study design followed by a quantitative observation study. In particular, a multiple (or collective) case study approach is used. A single case study involves an in-depth study of a particular event or individual for a specified period of time (Leedy & Ormrod, 2005, p. 135). A multiple case study consists of several individual cases, which “may be similar or dissimilar, redundancy and variety each important” (Stake, 2003, p. 138). Leedy and Ormrod (2005) describe the purpose of a multiple case study as “to make comparisons, build theory, or propose generalizations” (p. 135). Figure 4.1 shows a diagrammatical representation of the design used in this research.

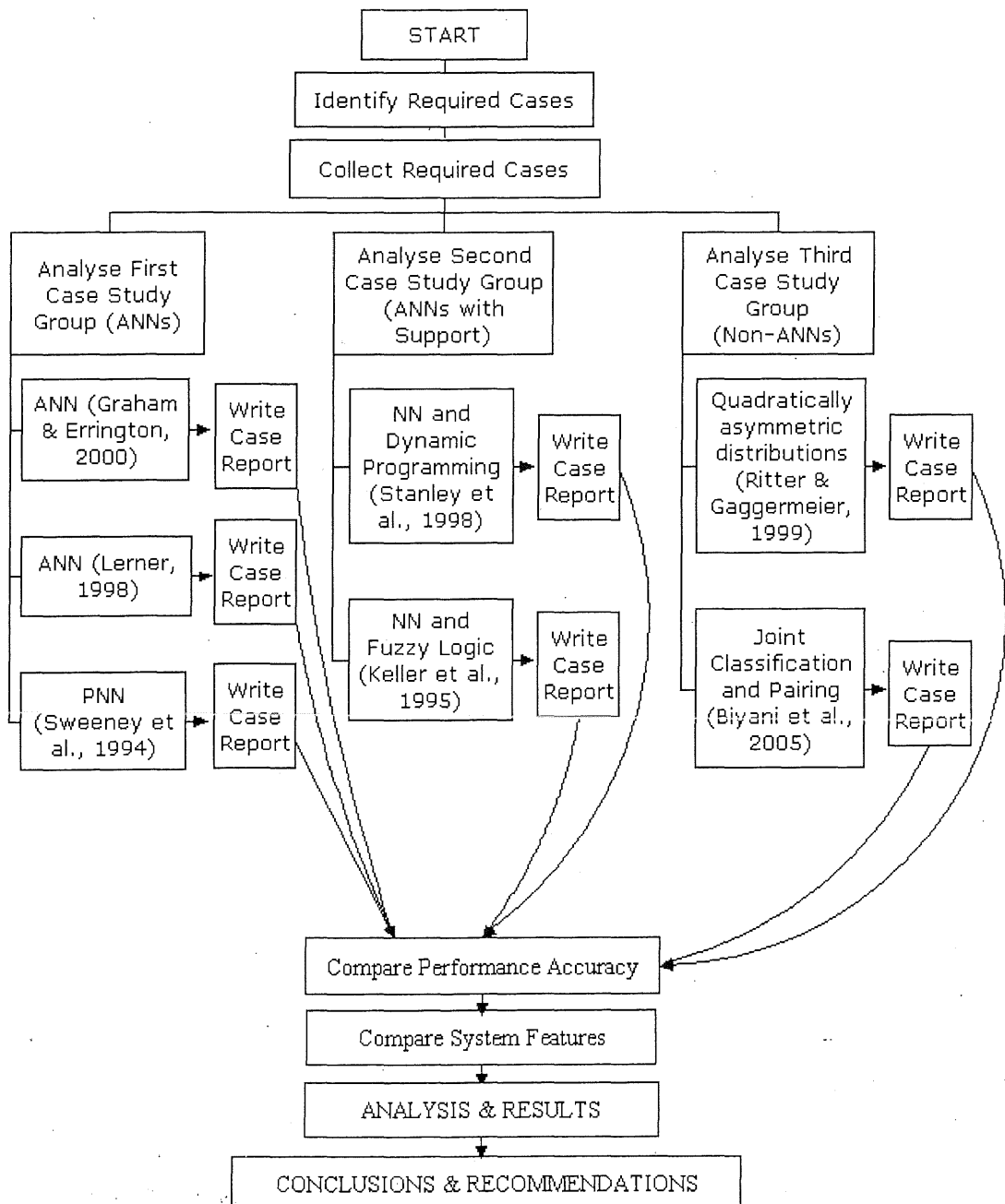


Figure 4.1. The applied research design, adapted from Yin (1994, p. 49)

The application of a multiple case study method in this research has followed the approach outlined in Figure 4.1. The tasks involved identifying the cases to analyse, determining appropriate data collection requirements, analysing each case study, writing reports on each analysis and then finally comparing the data collected. The data collection and analysis on each case study has revolved around investigating the implementation method, the results produced by each implementation on a variety of data sets and the process of achieving optimum performance. The different experiments conducted and the effects of these experiments on the overall accuracy were also considered.

Following the case studies, an observation study was conducted, which is described by Leedy and Ormrod (2005) as “the focus is on a particular aspect of behaviour. Furthermore, the behaviour is quantified in some way” (p. 180). In respect to this research, the observation study was used to quantify the performance of each implementation technique in terms of system features. The main system features considered are robustness, ability to generalise, validity, speed and degree of automation. These features are discussed in detail in Section 4.5 Data Analysis.

Upon completing the multiple case studies and the observation study, this research led to the integration phase, where the data were combined and studied. This phase focused on identifying the most effective automated chromosome classification technique based on results from performance and system feature criteria. The results obtained are presented and discussed in Chapter 6.

4.4. Data Collection

The data collected for this study are obtained from secondary data. Leedy and Ormrod (2005) define primary data as “the most valid, the most illuminating, the most truth-manifesting” (p. 89). They go on to define secondary data as data not collected from the source itself but from the primary data instead (Leedy & Ormrod, 2005, p. 89). Although secondary data is considered less valid, it does have advantages associated with its use. These include less cost for collection, easy accessibility, higher quality, and less time involved in collection (Punch, 1998, p. 107). These factors (cost, time, quality and accessibility) have high importance in this research of limited time span. Therefore, the use of secondary data has been appropriate for this study. The secondary data

identified for this study has been divided into three distinct case study groups, which are discussed below.

4.4.1. First Case Study Group

The first group of case studies consists of implementations describing the use of artificial neural networks as automated chromosome classifiers. Three case studies are identified:

1. *Classification of chromosomes: A comparative study of neural network and statistical approaches* (Graham & Errington, 2000). This article presents a detailed description of an artificial neural network implementation to human chromosome classification and also describes the process undertaken in achieving the optimal network architecture. The implementation was tested on the three popular chromosome databases and the performance of the system was compared to that of a statistical classifier.
2. *Toward a completely automatic neural-network-based human chromosome analysis* (Lerner, 1998). This article also presents a detailed implementation of artificial neural networks with a focus on creating a completely automated system, where the implementation is able to handle overlapping chromosomes and requires little or no human interaction.
3. *Classification of chromosome using a probabilistic neural network* (Sweeney, Musavi & Guidi, 1994). This article presents a varied implementation of neural networks: probabilistic neural networks. It also discusses the different testing experiments conducted and describes in detail the results produced.

4.4.2. Second Case Study Group

The second case study group includes implementations using artificial neural networks supported by another technique. Two case studies were chosen:

1. *Data-driven homologue matching for chromosome identification* (Stanley, Keller, Gader & Caldwell, 1998). This article presents an implementation of artificial neural networks supported by dynamic programming to the task of automated chromosome classification.

2. *A fuzzy logic rule-based system for chromosome recognition* (Keller, Gader, Sjahputera, Caldwell & Huang, 1995). This article merges fuzzy logic with neural networks to implement a chromosome classifier. The implementation is tested on its accuracy of identifying chromosomes from a selected class.

4.4.3. Third Case Study Group

The last case study group presented here discusses implementations not using artificial neural networks. Two cases studies are identified and these include the following articles:

1. *Automatic classification of chromosomes by means of quadratically asymmetric statistical distributions* (Ritter & Gaggermeier, 1999). This implementation uses statistical techniques to achieve chromosome classification. A novel implementation strategy is presented as different implementations are used for the different forms of chromosomal abnormalities.
2. *Joint classification and pairing of human chromosomes* (Biyani, Wu & Sinha, 2005). This case study also presents the use of statistical techniques for chromosome classification but attempts to merge the tasks of classification and pairing of chromosomes to achieve better performance.

4.5. Data Analysis

The analysis of the case studies considers two main factors: the classification accuracy, and the support of additional system features. The classification accuracy includes the different experiments conducted in testing the implementation technique and highlights the best results produced. Apart from achieving a high accuracy in chromosome classification, an effective implementation should support other important system features. The system should be robust and able to handle incomplete or imprecise data without loss of performance. Chromosome images commonly contain indistinct or unreliable information (Wang, et al., 2005, p. 2538) and therefore an automated chromosome analysis system should be able to effectively manage such data. In addition, the system should also be able to generalise and accept data it has not been trained with. Another important factor is the speed of classification. Piper, et al. (1980), cited in Tso and Graham (1983), note that the speed of an implementation is just as important as the classification accuracy (p. 495). This factor presents a dilemma of

sorts: a classification may be slow but produce better results, or may be very fast but have low accuracy. An effective technique should find a balance between speed and accuracy. The accuracy of a system may be greatly influenced by the training and testing data used. Therefore, in selecting the most effective implementation technique, valid training and testing data should have been used and the system should have been testing using experiments that apply to real-life karyotyping tasks. The degree of reliance on human interaction is another important feature to be considered. Automated karyotyping systems attempt to reduce the load on human experts and thus should aim to produce a system requiring little or no human interaction to function.

4.6. Validity

The integrity of the results produced by this research are directly related to the validity of the research. Validity, in this context, is defined as “the accuracy, meaningfulness, and credibility – *of the research project as a whole*” (Leedy & Ormrod, 2005, p. 97). In considering methods to achieve overall research validity, both internal and external research validity must be addressed.

4.4.1. Internal Validity

Leedy and Ormrod (2005) state “the internal validity of a research study is the extent to which its design and the data it yields allow the researcher to draw accurate conclusions about cause-and-effect and other relationships within the data” (p. 97). Leedy and Ormrod (2005) present several strategies for obtaining internal validity (p. 98-99); of these four approaches, the triangulation strategy has been used in this research. Patton (1987), cited in Yin (1994), identifies four different types of triangulation, and of these four strategies, data triangulation is used. Leedy and Ormrod (2005) define data triangulation as “multiple sources of data are collected with the hope that they will all converge to support a particular hypothesis or theory” (p. 99). This strategy applies to the research at hand as different case studies are analysed and the results produced by these analyses are compared. If the results converge, an effective technique for automated chromosome classification will be identified. Leedy and Ormrod (2005) support the use of triangulation in mixed method approaches by stating “triangulation is also common in mixed-method designs, in which both quantitative and qualitative data are collected to answer a single research question” (p. 99).

4.4.2. External Validity

External validity is described by Punch (1998) as the extent to which the study's findings can be generalised (p. 30). In obtaining external validity, the *replication in a different context* technique has been considered. Leedy and Ormrod (2005) define this method as "another researcher who conducts a similar study in a very different context reaches the same conclusion" (p. 100). By applying this technique to the research study, the results of other researchers will be compared to the results produced by this research. If the results converge, the external validity of the research will be supported.

5. ANALYSIS

This section provides an analysis of the case studies identified in this research. These case studies include applications of automated chromosome analysis systems using artificial neural networks, applications using artificial neural networks supported by other techniques and applications not using artificial neural networks. The articles presenting the different implementation methods are analysed and their techniques are presented and compared here. A detailed discussion and comparison of the performance of each technique is presented in the following chapter, Chapter 6 – Results. The analysis of each group of case studies is presented below.

5.1. Artificial Neural Networks

Artificial neural networks have been a popular choice for the implementation of automated karyotyping systems. Wang, et al. (2005) support this by noting that for chromosome classification:

Artificial neural network is the most popular tool owing to its capability of modelling the human brain and decision making process to recognize objects based on incomplete or partial information, as well as its simple topographic structure and easier training process. (p. 2540)

The articles chosen to represent the application of ANNs to chromosome classification are:

- *Classification of Chromosomes: A Comparative Study of Neural Networks and Statistical Approaches* (Graham & Errington, 2000);
- *Toward a Completely Automatic Neural-Network-Based Human Chromosome Analysis* (Lerner, 1998); and
- *Classification of Chromosomes Using a Probabilistic Neural Network* (Sweeney, Musavi & Guidi, 1994).

The analysis of these articles will discuss the neural network topology, training and testing methods and the results produced.

5.1.1. First Case Study

The article by Graham and Errington (2000), *Classification of chromosomes: A comparative study of neural networks and statistical approaches*, presents a discussion on the use of artificial neural networks as a system for chromosome classification. In their article, Graham and Errington (2000) use an ANN implementation as a

chromosome classifier and then compare the results obtained to that of a statistical chromosome classifier. Additionally, they experiment with different network architectures and different input information to achieve the optimum performance from the network. Their experiments included varying the chromosome features used as inputs, varying the network architecture and implementing the karyotyping constraint.

5.1.1.1. Implementation Details

In their final optimal network architecture, Graham and Errington (2000) combined two neural networks to conduct the chromosome classification. The first neural network acts as a pre-classifier. The role of this network was to accept two inputs representing the size and centromeric index features and to produce a Denver classification of the chromosomes; it therefore had seven outputs, referring to the seven Denver groups of chromosomes. The results of this network were then fed into the main neural network, which had 22 input nodes: seven inputs representing the Denver groups obtained from the pre-classifier and 15 inputs describing the banding features of the chromosome in question. This network consisted of one hidden layer containing 100 hidden nodes and produced 24 outputs, representing the 24 chromosome classes. Figure 5.1 portrays this network architecture.

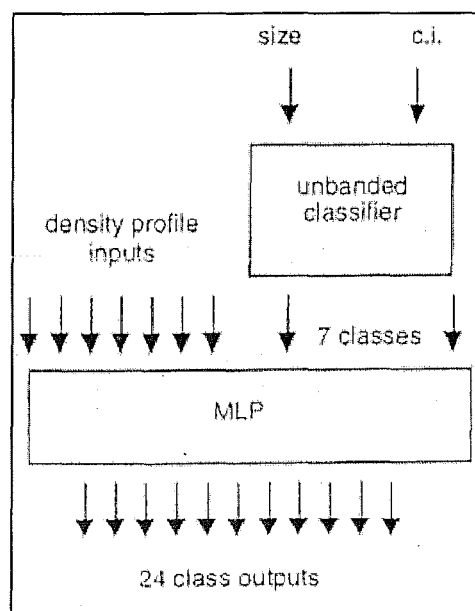


Figure 5.1. Graham & Errington's neural network architecture (Graham & Errington, 2000, p. 254)

5.1.1.2. Results

The neural network implementation was trained and tested on the three popular chromosome databases: Copenhagen, Edinburgh and Philadelphia. The training followed the hold-out cross validation training technique, in which half the data set was used for training and the other half used for testing (Graham & Errington, 2000, p. 255). The best classification results using context-free classification were 94.2% classification accuracy for the Copenhagen database, 83.0% classification accuracy for the Edinburgh database and 77.5% classification accuracy for the Philadelphia database (Graham & Errington, 2000, p. 255). Table 5.1 displays these results. The considerable difference in classification accuracy between databases is due to the significant difference in image quality within the three databases.

Graham and Errington (2000) also experimented by adding the karyotyping constraint to conduct context-dependant classification. The karyotyping constraint specifies that “there are exactly two chromosomes in (almost) all classes” and is implemented as a global constraint using the transportation algorithm (Graham & Errington, 2000, p. 258). By adding the karyotyping constraint to the network, Graham and Errington (2000) were able to produce higher classification accuracy; “transportation rearrangement achieves misclassification rates which have not been bettered by any other approaches” (Graham & Errington, 2000, p. 258). The results produced were 95.8% misclassification for the Copenhagen database, 85.6% misclassification for the Edinburgh database and 81.1% misclassification for the Philadelphia database (Graham & Errington, 2000, p. 260). These results are also shown in Table 5.1.

Table 5.1. Context-free and context-dependant classification accuracy

	Context-Free	Context-Dependant
Copenhagen Database	94.2%	95.8%
Edinburgh Database	83.0%	85.6%
Philadelphia Database	77.5%	81.1%

In comparing the neural network classifier to a statistical classifier, Graham and Errington (2000) concluded that the neural network approach “can give a higher classification accuracy than a classical parametric method” (p. 261). Yet they note that

“while the improvement was statistically significant, however, it was still small in absolute terms” (Graham & Errington, 2000, p. 261). Therefore, the improvement was not of a significant value to make a real difference. However, other factors pertaining to the neural network system advocate its use over a statistical classifier. Graham and Errington (2000) claim that the neural network development costs were less, time involved was less, less manpower was used, and that the neural network is likely to be more adaptable and more stable when dealing with data of different quality (p. 261). This is supported by Patterson (1996) who identifies several valuable characteristics of neural networks as learning, generalisation, robustness and parallel processing capabilities (p. 24-27).

5.1.2. Second Case Study

The second article involving the use of neural networks is: *Toward a completely automatic neural-network-based human chromosome analysis* (Lerner, 1998). In this article, Lerner (1998) presents his research in attempting to completely automate the process of chromosome classification.

5.1.2.1. Implementation Details

The network architecture used by Lerner (1998) is similar to that of Graham and Errington (2000) as it also involves more than one classifier. The first classifier, a ‘group classifier’, produces a Denver classification of the chromosomes. Seven ‘type classifiers’ are then used to classify the chromosomes within each of the Denver groups. Each of these type classifiers is trained and tested on a particular Denver group only, therefore acting as specialised classifiers. Lerner (1998) supports this approach by stating

Chromosome identification by first classifying the patterns into groups followed by a classification in the groups and into types yields both a desired task decomposition and a compatibility with the common cytogenetic methodology, which partitions the twenty-four chromosome types into seven groups. (p. 547)

5.1.2.2. Results

Lerner (1998) also experimented with using different features in his neural network implementations. In one experiment, Lerner (1998) used sixty-six chromosome features, which consisted of the chromosome length and centromeric index and 64

density profile features. This experiment was conducted using a two-layer neural network trained using the back-propagation algorithm. The network was trained and tested on a private database, named the Soroka5 database. However, the network was only required to classify five types of chromosomes, out of the total of 24 chromosome classes. The results produced were extremely accurate, with the network achieving an average of 99.3% classification accuracy.

In another experiment, the length and centromeric index were also used but only 15 out of the 64 density profile features were included. These features were input into a 'group classifier', which was responsible for classifying the chromosomes into their respective Denver groups. This structure therefore had 17 input nodes and contained seven output nodes. Following this group classification, a 'type classifier' was implemented to classify the chromosomes into their classes within a specific Denver group. The inputs to this network were also the 17 chromosome features but the outputs ranged between two to eight nodes, depending on the number of chromosomes within the group under analysis. This system was trained and tested on the Edinburgh dataset and achieved an 83.6% classification accuracy using this implementation (Lerner, 1998, p. 549).

By incorporating the transportation algorithm into this implementation, the results produced led to 84.5% classification accuracy (Lerner, 1998, p. 550), a slight improvement from the 83.6% of context-free classification. Table 5.2 displays the classification accuracy of context-free and context-dependant classification on both the Soroka5 and Edinburgh databases.

Table 5.2. Context-free and context-dependant classification accuracy

	Context-Free	Context-Dependent
Soroka5	99.3%	-
Edinburgh	83.6%	84.5%

5.1.3. Third Case Study

A different approach to applying neural networks to chromosome classification is that of using probabilistic neural networks. The article by Sweeney, Musavi and Guidi (1994), *Classification of chromosomes using a probabilistic neural network*,

describes the application of probabilistic neural networks (PNNs) to the task of chromosome classification.

Sweeney, Musavi and Guidi (1994) describe PNNs as “the combination of a kernel based estimator for estimation of probability densities and a Bayes rule for the classification decision” (p. 18). Patterson (1996) supports the use of PNNs in classification tasks by referencing Mood and Graybill (1962) and stating “the PNN models the popular Bayesian classifier, a technique which minimizes the expected risk of classifying patterns in the wrong category” (p. 350). Sweeney, et al. (1994) emphasise that the advantages of a PNN implementation include fast processing time, simple training process and the ability to generalise without requiring extensive training (p. 18). Wasserman (1993) supports this by noting “the PNN process is as much as five orders of magnitude faster than backpropagation” (p. 35). The fast processing time of the PNN is supported by large memory; this requirement is not a limitation due to memory being “abundant and affordable” (Sweeney, et al., 1994, p. 18).

5.1.3.1. Implementation Details

The structure of the PNN used by Sweeney, et al. (1994) consisted of 30 input features, including the normalized area, size, density, length and centromeric index (Sweeney, et al., 1994, p. 20). Figure 5.3 shows the architecture of the PNN used in this implementation. The values shown as $X_1 - X_{30}$ represent the input values. These values are accepted by the pattern units, shown as class 1 to class 24. Each class produces an output represented by Y_1 to Y_{24} . These outputs are then sent to a summation node (the maximum selector) which then produces the final output the network, represented by Y . Patterson (1996, p. 353) presents a general architecture for probabilistic neural networks, which is shown in Figure 5.4. The figure provided by Patterson (1996) provides further understanding of the architecture described by Sweeney, et al. (1994).

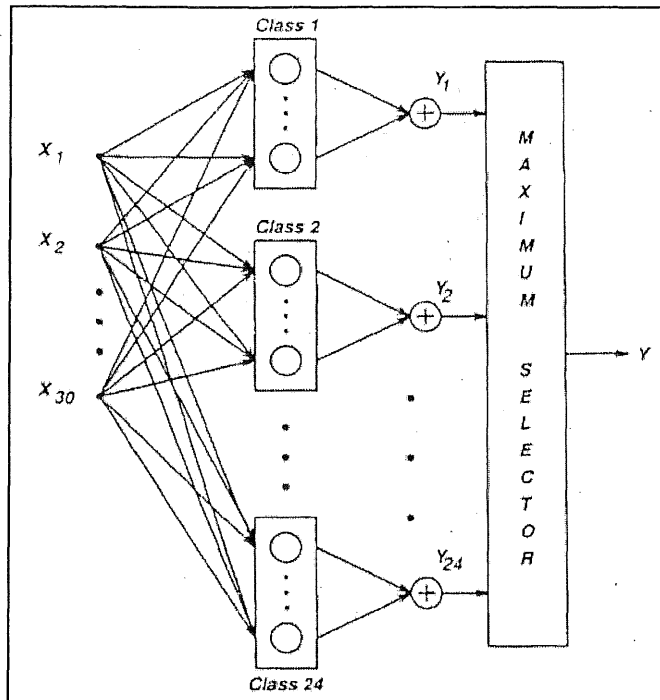


Figure 5.3. Probabilistic neural network architecture used in this implementation (Sweeney, et al., 1994, p. 19)

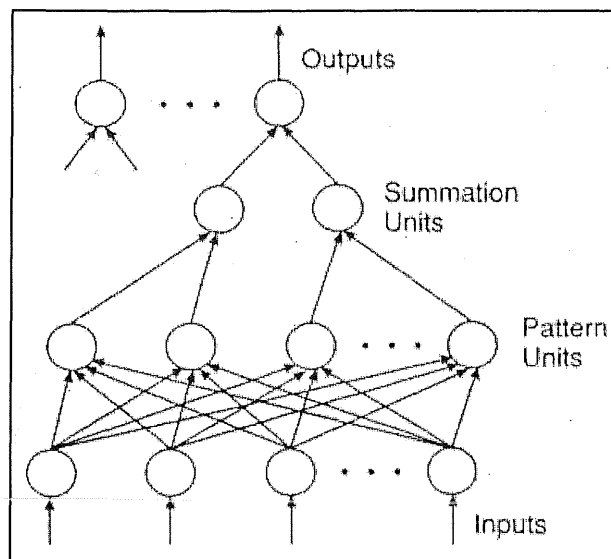


Figure 5.4. General probabilistic neural network architecture (Patterson, 1996, p. 353)

5.1.3.2. Experiments

The system was trained and tested using the three common databases: Copenhagen, Edinburgh and Philadelphia. However, the data was filtered, as “chromosomes from each of the databases that were either touching, overlapping, or unclassifiable were excluded from the experiments” (Sweeney, et al., 1994, p. 20). This severely limits the real-life applicability and generalisation ability of the system as real-life chromosome images commonly contain overlapping, touching or clustered

chromosomes (Wang, et al., 2005, p. 2538). This requirement of pre-segmented data presents a significant reliance on human interaction and therefore does not propose a completely automated system for chromosome classification. Additionally, the use of isolated chromosomes affects performance accuracy as removing complex data facilitates the classification process and consequently leads to lower error rates.

Two different testing and training methods were used:

1. the hold-out technique (also known as cross-validation); and
2. the leave-one-out technique where “one cell from the database is removed, the remaining cells are used for training and then the isolated cell is used for testing. This process is repeated for every cell in the databases” (Sweeney, et al., 1994, p. 20).

The authors experimented with different training and testing techniques to determine the most effective performance. In addition to the two training methods described above, Sweeney, et al. (1994) introduced an update procedure, which “gives the network knowledge that there can be a maximum of two chromosomes assigned to each class” (p. 19). The process of the update procedure is described as:

If a class has more than 2 chromosomes assigned to it, then the 2 chromosomes with the highest estimates are kept in that class, while the others are assigned to a new class, one to which they were not assigned before, corresponding to their next highest estimates. (Sweeney, et al., 1994, p. 19)

This procedure is repeated a set number of times and as with the karyotyping constraint applied to artificial neural networks, the update procedure helps improve the classification accuracy of the probabilistic neural network (Sweeney, et al., 1994, p. 19).

Sweeney, et al. (1994) used combinations of the two training techniques and the update procedure to conduct five different experiments for chromosome classification. These experiments were:

1. PNN using the hold-out technique for training;
2. PNN using the hold-out technique with the update procedure;
3. PNN using the leave-one-out training technique;
4. PNN using the leave-one-out technique with the update procedure; and finally,
5. Inter-database classification, in which the network was trained with one database and tested with the remaining two databases. This experiment did not incorporate the use of the update procedure (p. 20-21).

5.1.3.3. Results

The results achieved from conducting the above experiments showed that using a PNN classifier trained with the leave-one-out technique and using the update procedure (experiment #4) gave the best classification accuracy. This performance gave a 97.0% classification accuracy rate for the Copenhagen database, 84.7% classification accuracy for the Edinburgh database and a 78.8% classification rate for the Philadelphia database. These results are portrayed in Table 5.3. Sweeney, et al. (1994) had anticipated this experiment to outperform the rest and argue that “this is expected because the maximum possible number of training sets was used and the network is forced to assign a maximum of two chromosomes to a class” (p. 22).

The best performance achieved without the use of the update procedure was that of experiment #3, using the PNN with the leave-one-out training method. The results produced were 95.6% accuracy for the Copenhagen database, 83.4% accuracy for the Edinburgh database and 77.8% accuracy for the Philadelphia database (Sweeney, et al., 1994, p. 22). Table 5.3 shows these results.

Table 5.3. PNN classification accuracy

	PNN using leave-one-out (No Update)	PNN using leave-one-out (With Update)
Copenhagen Database	95.6%	97.0%
Edinburgh Database	83.4%	84.7%
Philadelphia Database	77.8%	78.8%

5.2. Artificial Neural Networks Supported by Other Techniques

This section will discuss implementations that use artificial neural networks supported by other techniques to perform chromosome classification. The two articles chosen to represent these case studies are:

- *Data-driven homologue matching for chromosome identification* (Stanley, Keller, Gader & Caldwell, 1998); and
- *A fuzzy logic rule-based system for chromosome recognition* (Keller, Gader, Sjahputera, Caldwell & Huang, 1995).

5.2.1. First Case Study

The article by Stanley, Keller, Gader and Caldwell (1998), *Data-driven homologue matching for chromosome identification*, presents the use of dynamic programming and neural networks to classify chromosomes. Many chromosome classification implementations assume two chromosomes per class, which cannot be applied when dealing with abnormal chromosomes (Lerner, 1998, p. 550). Stanley, et al. (1998) attempt to address this problem and therefore the focus of their paper is on “the development of image analysis techniques that are directly applicable to evaluating numerical aberrations evolving from structural abnormalities” (p. 452).

5.2.1.1. Implementation Details

To conduct chromosome classification, Stanley, et al. (1998) focus on identifying matching homologues and assigning them to the representative class. The technique is implemented using an iterative process, described by Stanley, et al. (1998) as:

For the selected class, the best representative or primary chromosome is found within the metaphase spread. Homologue candidates are obtained using simple criteria. The candidates are matched to the primary chromosome for homologue determination. (p. 452)

The process of chromosome classification uses a neural network confidence assignment and then dynamic programming to determine matching homologues. The process commences by automatically extracting chromosome features from images of metaphase cell spreads (Stanley et al, 1998, p. 454). The features extracted are the chromosome size (including length and area), the centromeric index, the banding pattern features and other chromosome profile features (Stanley, et al., 1998, p. 454). These values were entered into the neural network which then produced a confidence value representing the likelihood of the chromosome belonging to a certain class. “The initial candidates chosen were the chromosomes with confidence values greater than zero in the desired class” (Stanley, et al., 1998, p. 454). Candidates were then eliminated if their features, such as banding patterns and centromeric index ratios, were not representative of chromosomes belonging to the class in question (Stanley, et al., 1998, p. 456). “From the remaining candidates, the chromosome with the greatest margin of victory in neural-network confidence was chosen as the reference, prototype, or primary chromosome” (Stanley, et al., 1998, p. 456). The remaining candidates were then

inspected and the matching homologue was chosen using dynamic programming (Stanley, et al., 1998, p. 456).

Figure 5.5 provides a summarised version of the algorithm used by Stanley, et al. (1998) in conducting homologue matching. This algorithm conducts two checks when performing homologue matching: firstly, the algorithm identifies the primary chromosome for a selected class and then determines the matching homologue. The second check uses the identified homologue and selects a new candidate pool from which the matching chromosome is found. If the matching chromosome found is the original primary chromosome, then the matching is complete and both chromosomes are assigned to the selected class. If the matching chromosome found is not the primary chromosome, then only the original primary chromosome is assigned to the class under analysis. In essence, this algorithm conducts a two-way matching, to ensure the homologue chromosome matches the primary chromosome and that the primary chromosome matches the homologue chromosome.

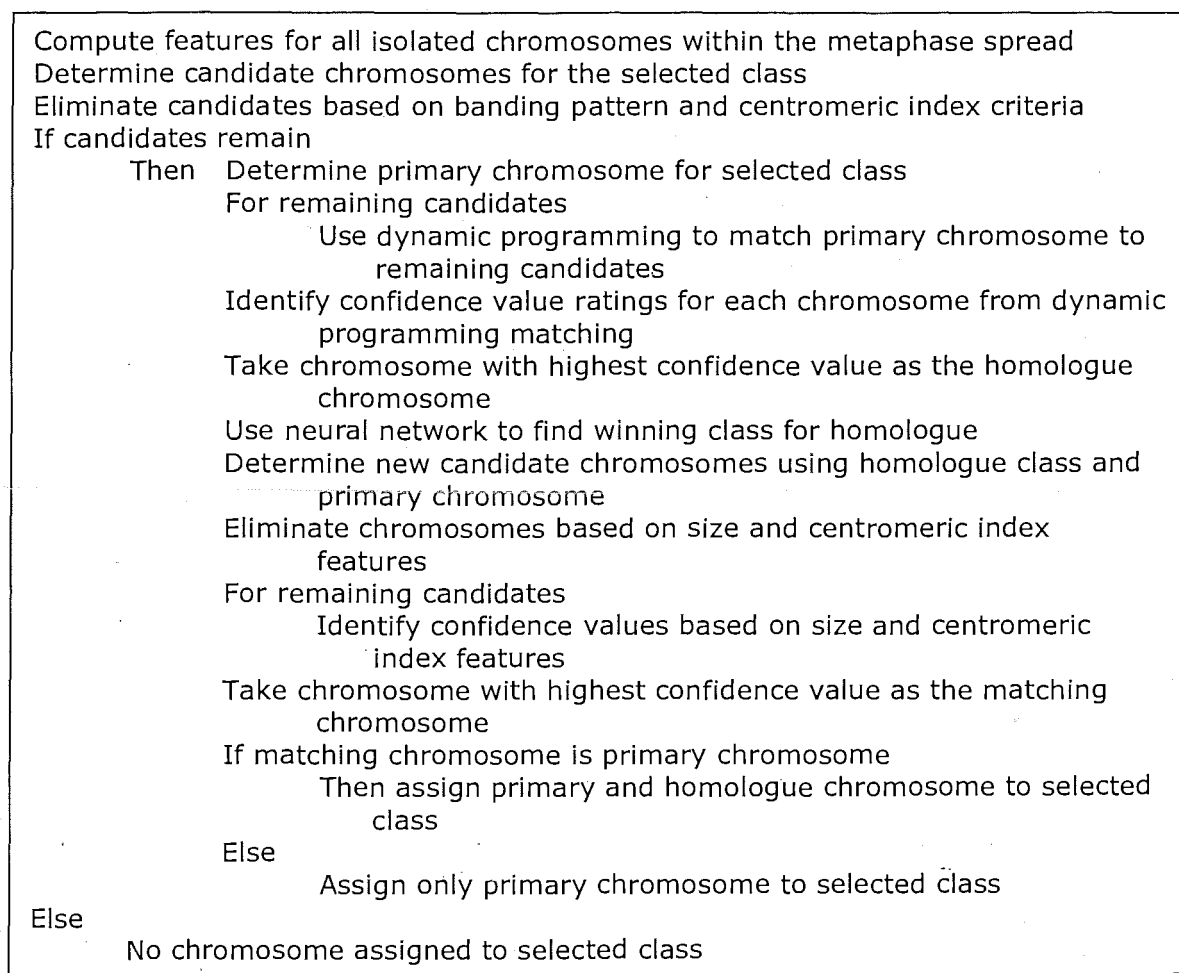


Figure 5.5. Summarised algorithm used for homologue matching (adapted from Stanley, et al., 1998, p. 454).

The neural network was used only to produce confidence values of chromosome assignment to each class and was not used for any image processing, feature extraction or final chromosome assignment. In effect, the neural network was only applied to identify a select group of chromosome candidates for each chromosome class. The neural network was given inputs representing the density and shape profile distribution features of each chromosome and the network produced an output representing the confidence values for each chromosome class.

This application is limited in its correlation with real world data as “only isolated chromosomes within metaphase spreads were of interest for this study” (Stanley, et al., 1998, p. 454). Metaphase spreads commonly contain chromosomes that are overlapping or touching (Wang, et al., 2005, p. 2538). By eliminating such data, the application eliminates a large source of uncertainty and also eliminates a large set of actual real-life data.

5.2.1.2. Results

Three experiments were performed to test the accuracy of this implementation:

1. The first experiment involved identifying matching chromosomes from a selected class. Stanley, et al. (1998) illustrate the importance of such a classification by arguing that “the ability to find chromosomes coming from a specific class is important for karyotyping and anomaly detection” (p. 459). In this experiment, sixteen chromosome 17s were placed within a metaphase spread. Of these sixteen chromosomes, four pairs were homologous. The application was able to correctly identify all four matching homologues, thus leading to a 100% accuracy rate.
2. The second experiment tested the ability of the system to identify the matching homologues for a selected class from a metaphase spread. The system was tested in identifying the class 17 chromosomes using 55 metaphase spreads. Of these 55 metaphase spreads, 53 had two chromosome 17s and the remaining 2 spreads had only one chromosome 17 (Stanley, et al., 1998, p. 459). The best performance of this system achieved a correct identification of the chromosomes of class 17 in 49 of the 55 metaphase spreads, thus leading to an 89.1% accuracy rate (Stanley, et al., 1998, p. 459-460).

3. The third experiment used a neural-network supported by the transportation algorithm to identify the homologues of class 17 from within a metaphase spread. This experiment is essentially the same as the second experiment but the only variance is in the implementation technique. This technique correctly found the homologues in 44 of the 55 metaphase spreads, leading to an 80% accuracy rate (Stanley, et al., 1998, p. 460).

Table 5.4. Percentage of classification accuracy of the three experiments described above

Test	Classification Accuracy
Experiment 1	100%
Experiment 2	89.1%
Experiment 3	80%

Although this implementation achieves a high rate of classification accuracy, the testing was limited to identifying only one chromosome class out of the total of 24 classes and only 55 metaphase images were used. This presents a rather small data set as other implementations have used large databases containing over three times the amount of cell images (Piper & Granum, 1989, p. 244; Sweeney, et al., 1994, p. 20).

5.2.2. Second Case Study

Another implementation of chromosome classification uses an artificial neural network supported by a fuzzy logic system. This implementation is described by Keller, Gader, Sjahputera, Caldwell and Huang (1995) in their article: *A fuzzy logic rule-based system for chromosome recognition*.

Keller, et al. (1995) acknowledge that “uncertainty abounds in every phase of computer vision” and that this uncertainty is commonly due to noise, imprecise computations and ambiguous interpretations (p. 126). As chromosome classification is a highly sensitive technique, a small additional noise introduced in a cell image could lead to different representation of the chromosome structure and features (Ritter & Gaggermeier, 1999, p. 1001).

Negnevitsky (2002) notes that real-life data is often “incomplete, inconsistent, uncertain, or all three. In other words, information is often unsuitable for solving a problem” (p. 55) Fuzzy systems are capable of handling such uncertainty (Nguyen &

Walker, 1997, p. 11; Negnevitsky, 2002, p. 259). This ability to handle uncertainty and imprecise data makes a fuzzy logic implementation suited to that of chromosome classification.

5.2.2.1. Implementation Details

Keller, et al. (1995) identify two possible ways in which to merge the fuzzy logic system with the neural network classifier:

1. an independent check on the results of the neural network classifier; or
2. a pre-classifier to place the chromosome image into its Denver group, and then allow specially devised neural networks to resolve within group ambiguity. (p. 129-130).

The article by Keller, et al. (1995) describes the implementation of the first approach, an independent check on the classification of two classes of chromosomes: class 16 and class 18. The features used for chromosome classification include the centromeric index, relative length, and three banding pattern values for the number of bands, band spacing and band intensity (Keller, et al., 1995, p. 127). For this preliminary test, Keller, et al. (1995) required a total of seventy-four (74) rules:

- 25 rules representing the class 16 confidence using centromeric index and length;
- 25 rules representing class 18 confidence using centromeric index and length; and,
- three sets of 8 rules (total of 24 rules) representing the confidence of class 16 and 18 based on three different band density values. (p. 131)

Figure 5.6 displays the fuzzy sets used to represent various membership functions for the centromeric index ratio. Keller, et al. (1995) describe their usage of the centromeric index as “the ratio of the short arm to the long arm of a chromosome, and so, is a value scaled into the interval [0, 1]” (p. 128). Keller, et al. (1995) have not identified the bounding limits of each fuzzy set but the total range would of necessity be between 0 and 1.

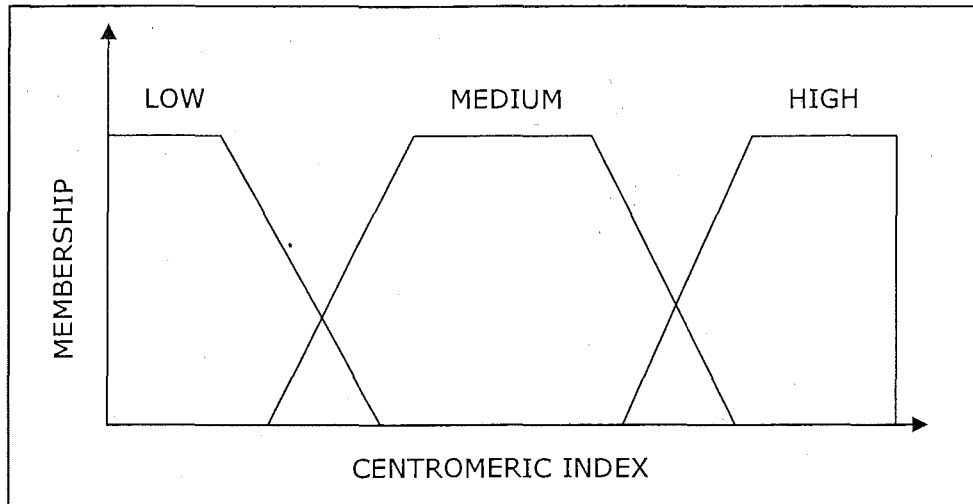


Figure 5.6. Fuzzy sets representing centromeric index (adapted from Keller, et al., 1995, p. 128).

Figure 5.7 provides an example of one rule used to identify the confidence of class 18 chromosomes based on the values of various chromosome features. As described above, Keller, et al. (1995) used seventy-four such rules; each rule produces its individual confidence value and “the fuzzy inference mechanism aggregates these values to produce final results for each class” (Keller, et al., 1995, p. 129).

IF	relative length is	SMALL	AND
	subtelocentric confidence is	HIGH	AND
	P-band is	ONE	AND
	Q-band is	TWO	AND
	distance(P1) is	MEDIUM	AND
	length(P1) is	LARGE	AND
	distance(Q1) is	SMALL	AND
	length(Q1) is	MEDIUM	AND
	distance(Q2) is	LARGE	AND
	length(Q2) is	MEDIUM	
THEN	chromosome 18 confidence is	LARGE	

Figure 5.7. Fuzzy rule for identify class 18 chromosome confidence (Keller, et al., 1995, p. 129).

5.2.2.2. Results

The system correctly classified all chromosome 16 images, and 87% of chromosome 18 images (Keller, et al., 1995, p. 131). By integrating these values to achieve overall classification accuracy, Keller, et al. (1995) were able to achieve a 100% reliability of classification with a 23% rejection rate (p. 131).

Table 5.5. Classification accuracy in classifying class 16 and class 18 chromosomes

Chromosome Tested	Classification Accuracy
Chromosome 16	100%
Chromosome 18	87%

Although this application achieved a high rate of correct classification, it does have several limitations. The tests experiment with the classification of only two chromosome classes and do not test the system performance in classifying the full 24 chromosome classes and producing a karyotype. Additionally, the testing data used pre-processed information, which would require a human expert to manually extract all necessary features for each individual chromosome. This still places a large reliance on the human expert.

Another possible area of improvement lies within the merger of neural networks and fuzzy logic. In the system presented by Keller, et al. (1995), the two techniques are not merged but rather executed in sequence with no collaboration between data. Although such an implementation still takes advantage of the characteristics of each individual AI technique, it does not take advantage of the enhanced performance offered by integrating the two techniques. Negnevitsky (2002, p. 266) states that “fuzzy logic and neural networks are natural complementary tools in building intelligent systems”. The reasoning behind this argument is that fuzzy logic supports the weaknesses in neural networks and neural networks supplement the weaknesses of fuzzy logic. Negnevitsky (2002) argues that:

Integrated neuro-fuzzy systems can combine the parallel computation and learning abilities of neural networks with the human-like knowledge representation and explanation abilities of fuzzy systems. As a result, neural networks become more transparent, while fuzzy systems become capable of learning. (p. 267)

Given the strong advantages of fully merging fuzzy logic and neural networks, it is expected that an integrated system would perform better than a system using only one technique. This expectation is supported by Catto, et al. (2003) in their study. Catto, et al. (2003) investigated the use of an integrated neuro-fuzzy system for the prediction of tumour behaviour. Their results showed that the integrated neuro-fuzzy system achieved higher accuracy than a stand-alone artificial neural network in most of the test cases (Catto, et al., 2003, p. 4175).

5.3. Non-ANN Techniques

Several automated karyotyping implementations use techniques that do not involve any form of artificial intelligence. These techniques commonly rely upon complex statistical distributions and mathematical equations to identify chromosomes within a metaphase spread. This section will discuss two such implementations:

- *Automatic classification of chromosomes by means of quadratically asymmetric statistical distributions* (Ritter & Gaggermeier, 1999); and
- *Joint classification and pairing of human chromosomes* (Biyani, Wu & Sinha, 2005).

5.3.1. First Case Study

The first article not using AI is *Automatic classification of chromosomes by means of quadratically asymmetric statistical distributions* (Ritter & Gaggermeier, 1999). The aim of this article is to “study whether algorithms can achieve human performance in a complex, clear-cut, and highly specific image-recognition task as the present one [of chromosome classification]” (Ritter & Gaggermeier, 1999, p. 998).

5.3.1.1. Implementation Details

In designing their system, Ritter and Gaggermeier (1999) decided to implement three different classifiers, one for each of the possible structural abnormalities in chromosomes. They support this decision by noting that “some cells may contain abnormal constellations and sometimes there are artefacts of preparation and culture. These aberrations usually cause a cell to contain fewer or additional chromosomes”, (Ritter & Gaggermeier, 1999, p. 1001). Given the common abnormalities found within human chromosome cells, Ritter and Gaggermeier (1999) then conclude that “it is clear that we need three classifiers: one for cells with 46 chromosomes, and two classifiers for cells with one missing and one extra chromosome, respectively” (p. 1001). The three individual classifiers will be briefly described below:

1. The first classifier deals with cells containing the correct number of chromosomes (46) and it is assumed that these cells have the correct homologue pairing or two chromosomes per class (Ritter & Gaggermeier, 1999, p. 1001). “After numbering the 46 chromosomes in an arbitrary way, the classification task consists in finding a correct assignment [of chromosomes to their respective classes]”, (Ritter & Gaggermeier, 1999, p. 1002). The process of assignment is

achieved using probability and likelihood estimation (Ritter & Gaggermeier, 1999, p. 1002).

2. The second classifier works on cells with one missing chromosome, therefore having a total of 45 chromosomes. In this case, a dummy chromosome is introduced in order to increase the total chromosome number to 46 (Ritter & Gaggermeier, 1999, p. 1002). The classification is also based on a probability and likelihood estimation (Ritter & Gaggermeier, 1999, p. 1002).
3. The third classifier is used for cells with one extra chromosome, thus having 47 chromosomes in total. This classifier functions on the assumption that an extra chromosome would indicate either one of the five well-known anomalies. To represent these known anomalies, an additional class is introduced to each of the chromosome numbers which represent these abnormalities (Ritter & Gaggermeier, 1999, p. 1003). The likelihood of chromosomes belonging to a particular class is also calculated using probability functions (Ritter & Gaggermeier, 1999, p. 1003).

While such a tailored approach would provide a more specified and detailed procedure for each of the three numerical aberrations, it does come with its limitations. The abnormality in chromosome number would have to be previously identified in order to activate the correct classifier. This requires additional computations to be performed by the cytogenetic expert. Additionally, the three classifiers described above do not cater for all possible chromosomal abnormalities. Other chromosomal abnormalities exist, including cases where individuals have an additional set of chromosomes or several additional chromosomes leading to 48 or 49 total chromosomes (Emery & Mueller, 1988, p. 125-139; Jorde, Carey, Bamshad & White, 2000, p. 112-121; Snustad & Simmons, 2000, p. 142-151).

In this implementation, all chromosomes are assumed to be independent and therefore only context-free classification is conducted. Ritter & Gaggermeier (1999) acknowledge that using context-free classification does not improve misclassification probability but they also argue that catering for homologous (matching) chromosomes “makes only a small difference” (p. 1002).

5.3.1.2. Results

To train and test their system, Ritter and Gaggermeier (1999) used the Copenhagen chromosome data set. The system used twenty-four (24) features for each chromosome; these features included the size, density, centromeric index, banding pattern and others (Ritter & Gaggermeier, 1999, p. 1005). These features were pre-extracted and therefore this implementation does not offer any automatic image manipulation functions. This also places a restriction on the application of this implementation as before being able to classify data, the cytogeneticist must first extract features from all chromosomes within the cell. As mentioned previously, Ritter and Gaggermeier (1999) identify one main goal of their research as to “study whether algorithms can achieve human performance in a complex, clear-cut, and highly specific image-recognition task as the present one [of chromosome classification]” (p. 998). However, from the descriptions provided, it is clear that the implementation presented by Ritter and Gaggermeier (1999) does not present an image recognition system but rather a pattern recognition application as the implementation does not accept chromosome images but chromosome features.

Despite these limitations, the research does present reasonable results. The system was tested using the cross-validation approach and the results obtained showed that 17.5% of chromosome cells were misclassified, leading to an 82.5% correct classification rate (Ritter & Gaggermeier, 1999, p. 1005-1006).

5.3.2. Second Case Study

Another application of a non-artificial intelligence technique to the task of chromosome classification is presented by Biyani, Wu and Sinha (2005) in their article *Joint classification and pairing of human chromosomes*. The main aim of this article is to attempt to improve the classification and pairing of chromosomes by combining the two tasks. “Better performance can be expected for both classification and pairing if one can combine the two properties, or jointly optimize the statistical decisions of chromosome classification and homologue pairing” (Biyani, et al., 2005, p. 105). The individual process of classification and pairing are first discussed before examining the integrated approach.

5.3.2.1. *Implementation Details*

Classification of chromosomes to their respective classes uses maximum likelihood estimation with the transportation algorithm. For cells with less than 46 chromosomes, dummy values are introduced into the data to represent these missing chromosomes. This requirement is necessary as using the transportation algorithm assumes that all classes have two chromosomes and in the case of missing chromosomes, the dummy values are needed to equalise the data. The classification process is based on two types of chromosome features:

1. Scalar features including “chromosome size, length, intensity, centromeric ratios, the number of bands in the banding profile” (Biyani, et al., 2005, p. 103), and;
2. A vector feature that represents the banding profile of the chromosome (Biyani, et al., 2005, p. 103).

The pairing of homologues involves identifying two matching chromosomes for all chromosome classes. The process of pairing homologous chromosomes is conducted using maximum likelihood estimation with a graph matching algorithm (Biyani, et al., 2005, p. 104).

Biyani, et al. (2005) differentiate between the process of classification and pairing by stating that “although the transportation algorithm for chromosome classification and the maximum-weight graph matching algorithm for homologue pairing are both based on maximum likelihood estimation, they rely on different statistical properties of chromosome data” (p. 104). This difference in data is portrayed in the representation of chromosome features. Biyani, et al. (2005) explain:

The transportation algorithm utilizes the property that the features ... of a given class fall within an expected range of variations, whereas the graph matching algorithm exploits the property that within a cell two chromosomes of a given class have similar features. (p. 104)

5.3.2.2. *Results*

The system was tested on two databases: a private database consisting of 350 cells and the popular Copenhagen database (Biyani, et al., 2005, p. 108). The chromosome features used include chromosome length, area, density, centromeric index and others (Biyani, et al., 2005, p. 108). These features were pre-extracted and thus

require extensive image processing by a human expert or external automated system before classification.

Testing the implementation followed the hold-out technique, where the data set was split into two halves and training was conducted on the first subset of data, while the other was used for testing and then visa versa. Several experiments were conducted to test the performance of this implementation in the individual tasks of classification and pairing and then the combined classification and pairing process. The classification of chromosomes from the Copenhagen database using the transportation algorithm achieved 98.1% classification accuracy, which is the highest result produced from all implementations discussed in this research.

The following chapter will examine and compare the results produced by each implementation. Other system features will also be considered and the most effective technique for chromosome classification will be identified and discussed.

6. RESULTS

This chapter presents an evaluation of the case studies analysed according to the performance accuracy in chromosome classification and various system factors such as robustness and validity. The most suitable implementation will then be identified and discussed.

6.1. Performance Comparisons

A main factor in identifying the most effective implementation technique is the performance accuracy. This section discusses the performance accuracy of all identified case studies and compares the results retrieved.

6.1.1. Implementations using ANNs

The three case studies using artificial neural networks have tested their systems using both context-free and context-dependant classification. Therefore, a distinction is made between the two different methods and each technique is discussed in a separate section below.

6.1.1.1. Context-Free Classification

As previously defined in the Literature Review, context-free classification is conducted when "individual chromosomes are considered as independent objects, without regard to their context as components of a karyotype" (Carothers & Piper, 1994, p. 164). The results produced by each case study are presented in Table 6.1. This table shows that two different implementations achieved comparable results. The ANN technique provided by Lerner (1998) achieved the highest classification accuracy on the Edinburgh database while the PNN implementation by Sweeney, Musavi and Guidi (1994) achieved the highest classification accuracy on the remaining two databases.

Table 6.1. Context-free classification accuracy

Implementation	Copenhagen	Edinburgh	Philadelphia
Graham & Errington (2000)	94.2%	83.0%	77.5%
Lerner (1998)	-	83.6%	-
Sweeney, et al. (1994)	95.6%	83.4%	77.8%

The accuracy results shown in Table 6.1 show that there is only a small margin of difference between each of the three implementations. This accuracy difference ranges from 0.2% to 1.4% and is a very small margin. Graham and Errington (2000) had compared their results to those obtained by Lerner (1998) and they propound that this small improvement in accuracy could be due to the more carefully chosen density features used for classification (Graham & Errington, 2000, p. 262). This could lead to the assumption that artificial neural network as chromosome classifiers have achieved their optimal performance, and any further improvements on this performance would be based on improvements in image processing and feature selection and extraction techniques (Ritter & Gaggermeier, 1999, p. 1007).

6.1.1.2. Context-Dependant Classification

The second popular method of classification is context-dependant classification. This technique gives the network knowledge that there must be two chromosomes per class in a normal cell (Lerner, 1998, p. 550). It is generally applied in the form of a global constraint, thus affecting all assignments of chromosomes to classes. Using context-dependant classification results in better chromosome assignment and therefore higher classification accuracy (Tso, Kleinschmidt, Mitterreiter & Graham, 1991, p. 118; Graham & Errington, 2000, p. 258).

Table 6.2 shows the misclassification error rates achieved for context-dependant classification. These results show that the PNN implementation by Sweeney, et al. (1994) achieved the best classification accuracy for two out of the three databases, the Copenhagen and Edinburgh databases, while the ANN approach by Graham and Errington (2000) obtained the best performance for the Philadelphia database.

Table 6.2. Context-dependant classification accuracy

Implementation	Copenhagen	Edinburgh	Philadelphia
Graham & Errington	95.8%	85.6%	81.1%
Lerner	-	84.5%	-
Sweeney, et al.	97.0%	84.7%	78.8%

For context-dependant classification, there is also only a small degree of variation in the classification accuracy, with the range being between 0.2% to 2.3%. By comparing the results shown in Table 6.1 to those of Table 6.2 above, it is evident that using context-dependant classification has indeed improved the classification for all

implementations on all testing data. This improvement varies between 0.9% (a rather insignificant improvement shown by Lerner's ANN implementation) to 3.6% (a much more substantial improvement shown by the Graham and Errington's ANN).

6.1.2. Implementations using ANNs with a Supporting Technique

The second group of case studies presents the use of artificial neural networks supported by other techniques. The two case studies from this group have conducted different experiments in testing their implementation and therefore each case study is discussed in a separate section below.

6.1.2.1. First Case Study

Stanley, Keller, Gader and Caldwell (1998) present a data-driven technique supported by neural networks and test their system using three experiments. The experiments are reviewed below and the results produced are shown in Table 6.3.

1. The first experiment involved identifying four homologous chromosome pairs from a total of 16 chromosomes. The chromosome class selected was class 17 and the implementation, using neural networks supported by dynamic programming, correctly identified all four pairs, thus achieving a 100% accuracy rate.
2. The second experiment involved finding chromosomes from a selected class from a completely metaphase spread image using neural networks supported by dynamic programming. Again, the class selected was class 17 and the implementation identified the chromosomes of class 17 in 49 of the 55 metaphase spreads, thus leading to an 89.1% accuracy rate.
3. The third experiment was similar to that of experiment two, except a neural network implementation supported by the transportation algorithm was used. This implementation identified chromosomes of class 17 in 44 out of the 55 metaphase spreads, resulting in 80% classification accuracy.

Table 6.3. Classification accuracy of each experiment

Test	Classification Accuracy
Experiment 1	100%
Experiment 2	89.1%
Experiment 3	80%

These experiments show that the classification accuracy of neural network implementations can be improved by using dynamic programming to assist in homologue matching. The neural network and dynamic programming implementation system achieved 89.1% classification accuracy compared to only 80% accuracy by the neural network with the transportation algorithm. However, the tests conducted are not thorough as they only test the performance of the implementation in identifying class 17 chromosomes. It cannot be assumed that the implementation will achieve the same error rates if applied to identifying chromosomes of a different class or identifying a complete karyotype from a metaphase spread. Additionally, only 55 cell images were used; other case studies have tested their implementations using databases consisting of between 125 to 180 metaphase cells (Piper & Granum, 1989, p. 244; Sweeney, et al., 1994, p. 20; Graham & Errington, 2000, p. 251).

6.1.2.2. Second Case Study

Keller, Gader, Sjahputera, Caldwell and Huang (1995) used a fuzzy logic rule-based system as an independent check on the results produced by a neural network. The implementation was tested on its ability to identify chromosomes of a selected class when presented with the given features. The data consisted of features extracted from 23 chromosomes of class 16 and 30 chromosomes of class 18 (Keller, et al., 1995, p. 131). The system was able to correctly classify all class 16 chromosomes and 87% of class 18 chromosomes (Keller, et al., 1995, p. 131). These results are shown in Table 6.4.

Table 6.4. Classification accuracy of class 16 and class 18 chromosomes

Chromosome Tested	Classification Accuracy
Chromosome 16	100%
Chromosome 18	87%

As these experiments involved only identifying chromosomes from a selected class; it is assumed that context-free classification is conducted as no mention of

homologue matching is given. From the results shown in Table 6.4, the classification accuracy of this implementation is higher than most results produced by the ANN implementations. However, the results cannot be directly comparable as the tests conducted on this implementation are rather limited. The system was only tested in identifying two out of a total 24 chromosome classes. The performance of the system in producing a full karyotype from metaphase cells cannot be generalised based on the performance achieved in these experiments.

6.1.3. Implementations not using ANNs

The last case studies presented did not use artificial neural networks in their implementations of chromosome classification. Again, each case study will be discussed individually as different experiments and tests were conducted.

6.1.3.1. First Case Study

This implementation by Ritter and Gaggermeier (1999) used probability and likelihood estimations to conduct chromosome classification. A variety of experiments were conducted and all were tested using the Copenhagen database. The best implementation achieved 82.5% classification accuracy.

This performance accuracy is positive and is comparable with results achieved by other implementations. It is important to note that all experiments conducted only context-free classification, and as compared with the previous case studies, this accuracy result is only 0.5% - 1.0% lower than the previously discussed ANN approaches. However, this implementation was reliant upon pre-extracted features from chromosome databases and therefore did not present any technique to automatically segment chromosome images and extract the required features (Ritter & Gaggermeier, 1999, p. 1005).

6.1.3.2. Second Case Study

Biyani, Wu and Sinha (2005) used maximum likelihood estimation as the basis for jointly classifying and pairing human chromosomes. They conducted experiments on two different data sets in testing their implementation. The first experiment,

conducted on a private database, tested the accuracy of the implementation in the individual tasks of classification and pairing. The results are shown in Table 6.5.

Table 6.5. Classification and pairing accuracy for female and male data sets

Data	Correct Classification Rate	Correct Pairing Rate
Female Set	94.25%	89.56%
Male Set	94.1%	90.1%

The data from this data set was divided into two individual sets for the female and male categories separately. This distinction is necessary due to the difference in chromosome classes between male and female cells (Biyani, et al., 2005, p. 108). As shown in Table 6.5, the difference between the gender sets is negligible; however the difference between the classification and pairing tasks is significant.

The second experiment tested the ability of the system to perform a complete karyotype and used the Copenhagen database of chromosome images. The best classification accuracy achieved was 98.1%, which was obtained by using the transportation algorithm. This result is outstanding and is comparable with the results produced by human expert cytogeneticists which usually lie in the range of 0.1 to 3.0% misclassification (Lundsteen, Lind & Granum, 1976, cited by Jennings & Graham, 1993, p. 959)

Although these results show the highest classification accuracy from all presented case studies, the implementation presented here used only pre-extracted features from the database and therefore did not offer any image segmentation and feature extraction procedures. This also places a large reliance on the human expert. Additionally, by omitting these tasks, the implementation removes a large area of error, as using pre-extracted data removes a large margin for error.

6.1.4. Outcome

This section discusses all the results presented in previous sections and compares performance accuracy in order to determine the best classifier implementation. As several implementations used different databases and different classification techniques, an accurate comparison is difficult to obtain. However, four

main cases are considered, including context-free classification using ANNs, context-dependant classification using ANNs, classification using ANNs supported by another technique and finally classification using a non-ANN technique. Since several different databases were used, only one database is chosen in order to level the comparisons. The Copenhagen database was used in most implementations and therefore all results presented here are based on classification accuracy achieved using this database.

As the case studies presented using artificial neural networks used both context-free and context-dependant classification to test the implementations, both techniques will be considered. For the artificial neural network context-free classification, Lerner's (1998) method produced the best results with 83.6% misclassification. For context-dependent classifications, the probabilistic neural network performed best on the Copenhagen database, with a 97.0% misclassification rate. The investigation into implementations using artificial neural networks supported by another technique showed that the data-driven homologue matching technique presented by Stanley, et al. (1998) performed best, with an 89.1% correct classification rate. From the non-neural network techniques, the best classification performance achieved a 98.1% correct classification rate, using the joint classification and pairing method presented by Biyani, et al. (2005). These results are all shown in Table 6.6.

Table 6.6. Context-dependant classification accuracy using Copenhagen database

Case Study Group	Implementation Technique	Results
Context-Free (NN)	ANN by Lerner (1998)	83.6%
Context-Dependant (NN)	Probabilistic Neural Network	97.0%
ANN with Support	Data-Driven Dynamic Programming	89.1%
Non-ANN	Joint Classification and Pairing	98.1%

By assessing only performance criteria as presented in these case studies, the best classification accuracy is achieved by the joint classification and pairing technique by Biyani, et al. (2005). However, as these case studies have all relied upon several different databases using images of varying quality and have used different testing experiments, the classification accuracy is not sufficient in assessing the most effective

chromosome classifier. Therefore, other system measures must be considered; these are discussed in the following section.

6.2. System Measures Comparisons

Although the accuracy of the classifiers is a crucial feature, other system qualities are important and must be considered when determining the most effective chromosome classifier. These characteristics include the ability to generalise, robustness, efficiency in computation burden and speed, validity in real-world data and degree of reliance on human interaction. These are discussed in separate sections below.

6.2.1. Ability to Generalise

An important feature of artificial neural networks is the ability to generalise when faced with new data. Patterson (1996) asserts that “generalization is an essential trait of intelligent behavior” (p. 25). Patterson (1996) describes generalisation as “ANNs generalize when they compute or recall full patterns from partial or noisy input patterns, when they recognize or classify objects not previously trained on, or when they predict new outcomes from past behaviors” (p. 25). Generalisation, however, can be limited by poor network architecture or training methods. The result is overfitting, which occurs when the neural network becomes specialised and limited within its training data and produces incorrect responses when faced with new data. Overfitting can be avoided by using proper training techniques and introducing noisy data (Patterson, 1996, p. 208).

Probabilistic neural networks are also capable of generalising to new data. Sweeney, et al. (1994) state “the network generalizes to any new incoming training patterns without having to repeat an extensive training process” (p. 18). Wasserman (1993) supports this by stating “inputs that are similar, but not identical to those in the training set will, within limits, be correctly classified” (p. 36).

However, fuzzy systems on their own do not generalise and adapt well when faced with new data (Negnevitsky, 2002, p. 267). The use of a neural network with a fuzzy system can overcome this limitation since, as described above, neural networks have good generalisation abilities, if set up correctly.

When implemented and trained correctly, maximum likelihood estimators (MLE) are also able to generalise. Eliason (1993) indicates that “as the sample size grows large, the MLE tends toward the properties of an unbiased estimator” (p. 20). However, the NIST/SEMATECH e-Handbook of Statistical Methods (n.d. a) argues that “maximum likelihood estimates can be heavily biased for small samples”. Therefore, in determining the extent of generalisation given by maximum likelihood estimation, the sample size plays a crucial role. Therefore, given the large size of chromosome samples in the popular chromosomes, it is arguable that the MLE implementations have sufficient data to reach a satisfactory unbiased state.

6.2.2. Robustness

The practice of chromosome analysis will often deal with uncertain or incomplete data. This is due to the process of chromosome culturing, which may often lead to extra particles among the chromosomes, and of imaging techniques, which can create images of low clarity (Wang, et al., 2002, p. 2538). Cho (2000) supports the above by noting that “it is difficult to get a clear microscopic chromosome image due to the variation of cell culturing conditions, chromosome staining, and microscope illumination” (p. 28).

Patterson (1990) defines robustness as “the ability of a learning system to function with unreliable feedback and with a variety of training examples, including noisy ones” (p. 364). As chromosome images will unavoidably contain indistinct areas, chromosome analysis systems must be equipped to handle these ambiguities. Among the case studies presented in this research, the use of artificial neural networks is appropriate for handling incomplete or uncertain data. Patterson (1996) reports that ANNs “continue to perform well when part of the network is disabled or presented with noisy data” (p. 27). Negnevitsky (2002) also affirms that artificial neural networks are capable of tolerating uncertainty and imprecision in data (p. 259). This ability is provided by the structure of the neural networks; Patterson (1996) describes:

This is possible because the ‘knowledge’ stored in an ANN is distributed over many neurons and interconnections, not just a single or a few units. Consequently, concepts or mappings stored in an ANN have some degree of redundancy built in through this distribution of knowledge. (p. 27)

Probabilistic neural networks are also robust; Patterson (1996) states “PNN networks also tolerate noisy samples and they can work with sparse samples too” (p. 354). Wasserman (1993) supports the above by claiming “erroneous, noisy or incomplete training or data inputs do not have a disproportionate effect on the classification accuracy” (p. 36).

Another technique effective in terms of its robustness, is fuzzy logic. Negnevitsky (2002) defines fuzzy logic as “logic that is used to describe fuzziness” (p. 87). It is therefore evident that fuzzy systems should perform well when dealing with uncertain or ambiguous data. Merging fuzzy logic with neural networks leads to a high powered system. Negnevitsky (2002) states “fuzzy logic and neural networks are natural complementary tools in building intelligent systems” (p. 266).

The use of maximum likelihood estimation also allows for interpretation of abnormal data. The NIST/SEMATECH e-Handbook of Statistical Methods (n.d. b) notes that apart from transforming the abnormal data into normal ranges, the main alternative is to “use a fitting criterion that directly takes the distribution of the random errors into account when estimating the unknown parameters. Using these types of fitting criteria, such as maximum likelihood, can provide very good results”. However, MLE does present complications as it is harder to use than other techniques (NIST/SEMATECH e-Handbook of Statistical Methods, n.d. b)

6.2.3. Efficiency in Computation Burden and Speed

A manual classification by a cytotechnician is a long and tedious process (Carothers & Piper, 1994, p. 161). Therefore, one of the aims of automated chromosome analysis is to lessen the computational requirements and accelerate the process.

The artificial neural network technique for chromosome analysis presents some problems with computational burden. The popular network training technique, the back propagation algorithm, performs well but is slow and complex. Kartalopoulos (1996) notes that “the algorithm suffers from extensive calculations and, hence, slow training speed” (p. 81). This computational burden has limited the applicability of neural networks and Kartalopoulos (1996) argues that the back-propagation algorithm is not suitable for many real-time applications (p. 82). Sweeney, et al. (1994) acknowledge the disadvantages of the back-propagation algorithm and claim “since chromosome

classification takes very large data sets with high dimensional input and output spaces, the time to train a BP [back propagation] network could take many hours of computing time (pg. 18).

The probabilistic neural network technique includes the advantages of neural network robustness while easing the training process. Sweeney, et al. (1994) argue that “the significant advantages of the PNN classifier are its speed and simplicity of the training process” (p. 18). Patterson (1996) supports the above by stating “one of the main advantages of the PNN is the speed with which it can be trained. No iterative procedures are used and no feedback paths are required in the training process” (p. 354). Kartalopoulos (1996) states that a “PNN simply stores the training patterns, avoiding the iterative process. It therefore learns very fast, but large data sets require large networks” (p. 105). A comparison given by Specht (1990), cited in Patterson (1996), shows that a probabilistic neural network was trained 200,000 times faster than a multilayer feedforward neural network trained with the back propagation algorithm (p. 354).

The maximum likelihood estimation implementations are based on complex algorithms and therefore can be computationally expensive. The NIST/SEMATECH e-Handbook of Statistical Methods (n.d. c) acknowledges this factor by noting that the procedure of MLE is “complicated and computationally intensive”.

6.2.4. Validity in Real-World Data

Although previous factors focused on the characteristics of the implementation techniques used, this characteristic is dependant upon the process of training and testing. An important characteristic of a classifier is the ability to associate with real-life situations and still perform well. Validity in this purpose can be seen as validity of testing data and validity of testing experiments

6.2.4.1. Validity of Testing Data

To correctly assess the performance of the classifier, the system should be tested with a wide range of data. The data should represent the various cases appearing in real-life data.

Several implementations did not consider data containing touching or overlapping chromosomes. These implementations include:

- the PNN implementation by Sweeney, et al. (1994),
- the neural network and fuzzy logic technique by Keller, et al. (1995),
- the data-driven technique by Stanley, et al. (1998),
- the statistical distribution approach by Ritter and Gaggermeier (1999),
- the neural network approach by Graham and Errington (2000), and;
- the joint classification and pairing technique by Biyani, et al. (2005).

The only technique to offer a system to manage touching or overlapping chromosomes was the ANN approach provided by Lerner (1998). All other implementations excluded data which contained chromosomes that were touching or overlapping. This severely skews the results produced by the testing experiments. Logically, if given data of high quality, the system should generally perform better than when dealing with data of low quality. This is confirmed in the classification accuracy results. When systems were tested with the Copenhagen database, which is considered to be a database of high quality, all implementations performed better than when tested with the Philadelphia database, a database of low quality. Only Lerner's (1998) technique offers an implementation capable of dealing with real-life data, which commonly consist of chromosomes that are touching or overlapping (Wang, et al., 2005, p. 2538).

6.2.4.2. Validity of Testing Experiments

Another factor in testing the performance of a classifier is the variety of testing experiments conducted. The system implemented should be capable of identifying chromosome homologues and producing a full karyotype. Therefore, the system testing should cover a range of experiments and assess the performance in analysing chromosomes from full metaphase spreads, not just isolated images.

Stanley, et al. (1998) conducted two different experiments on their system. The first was identifying matching homologues and the second was identifying chromosomes of a selected class from a metaphase spread. However, in both experiments, the performance of the system was only tested in identifying chromosomes of one class; class 17. No reference is made on the performance of the system in

classifying other chromosome classes. Additionally, Stanley, et al. (1998) did not test the system's ability in carrying out a full analysis and producing a full karyotype, which is the most common chromosome analysis technique (Wang, et al., 2005, p. 2536). Keller, et al. (1995) conducted similar experiments in identifying chromosomes of class 16 and 18 only and did not test the functionality of the system in producing a full karyotype. The remaining case studies all conducted experiments on testing the systems' performance in achieving a correct complete karyotype.

6.2.5. Degree of Human Reliance

One major criticism of automated chromosome analysis is the reliance of these systems on human intervention in many phases of the analysis. Wang, et al. (2005) state "although several commercialized software and systems have been developed, they are mostly semi-automatic products ... the interaction of a skilled laboratory technician is required to check the results and manually complete the karyotyping" (p. 2541). Piper and Granum (1989) support this important feature by stating that several error rates should "be treated with some caution since they depend substantially on the extent of prior interaction" (p. 242). Several case studies presented in this research have a large reliance on human interaction to facilitate data processing whether in feature extraction or chromosome segmentation.

The case studies relying on pre-extracted features and thus not incorporating any automated chromosome segmentation or feature selection are:

- the probabilistic neural network technique by Sweeney, et al. (1994),
- the fuzzy logic implementation by Keller, et al. (1995),
- the statistical distribution implementation by Ritter & Gaggermeier (1999),
- the neural network approach by Graham and Errington (2000), and;
- the joint classification and pairing approach by Biyani, et al. (2005).

All above case studies deal with the values of the chromosome features and thus involve a great deal of data pre-processing which must be done by a human expert. Given that these systems cannot perform a fully automated chromosome analysis from metaphase spread image to karyotype, they can be seen as cytogenetic aids, rather than complete systems.

The case study by Stanley, et al. (1998), using the data driven technique, presents an automation of the feature extraction task. However, only isolated chromosomes were used in that study and therefore no attempt is made at automating chromosome segmentation (Stanley, et al., 1998, p. 454). Lerner (1998) presents the only case study to provide a fully automatic chromosome segmentation system, which is capable of independently segmenting chromosomes and extracting features as well as conducting the actually chromosome analysis.

6.2.6. Outcome

In order to effectively compare the presented implementations, it has been necessary to construct a suitable framework for evaluation of system measures. This framework is presented using a rating scale consisting of the values of 0, 1 and 2, where 0 indicates a lack of the system measure under comparison and 2 indicates a strong possession of this characteristic. Each implementation was rated according to the scale described above; the rating was based solely on the information provided in each article and is used simply to facilitate the comparison and applies to this research and these case studies only. The implementation with the highest overall score was found to be the best implementation in terms of system features.

Table 6.7 presents the results of this comparison. There are a total of five system features under comparison and each feature may be assigned a maximum of two points leading to an overall total of 10 points. The rating scale and detailed results of each implementation are presented in detail in Appendix A.

Table 6.7. System measures ratings

Implementation	Generalisability	Robustness	Efficiency	Validity (Data)	Validity (Experiments)	Human Reliance	TOTAL
ANN (Graham & Errington, 2000)	2	2	0	0	1	0	5/10
ANN (Lerner, 1998)	2	2	0	1	1	2	8/10
PNN (Sweeney, et al., 1994)	2	2	2	0	1	0	7/10
Data-Driven Homologue Matching (Stanley, et al., 1998)	2	1	0	0	0	1	4/10
Fuzzy Logic Rule- Based System (Keller, et al., 1995)	1	2	1	0	0	0	4/10
Quadratically Asymmetric Statistical Distributions (Ritter & Gaggermeier, 1999)	2	2	0	0	1	0	5/10
Joint Classification and Pairing (Biyani, et al., 2000)	2	2	0	0	1	0	5/10

By assessing each implementation according to the various system factors, the artificial neural network implementation presented by Lerner (1998) produced the best overall results. This implementation performed well in most of the system features including generalisability, robustness, validity, human operator interaction but only suffered from expensive computational burden.

6.3. Best Classifier Performance

6.3.1. Best Classifier Implementation

The seven implementations have been assessed in terms of accuracy of chromosome classification and on five additional system measures. In order to achieve

an overall rating indicating the most effective implementation technique, the classification accuracy of each implementation has been combined with the ratings assigned from the system measure framework. Both factors are given equal weighting, thus leading to an averaged total score. These scores are shown in Table 6.8.

Table 6.8. Overall implementation scores

Implementation	Classification Accuracy	System Measure Rating	OVERALL TOTAL
ANN (Graham & Errington, 2000)	85.6%	5 (50%)	67.80%
ANN (Lerner, 1998)	84.5%	8 (80%)	82.75%
PNN (Sweeney, et al., 1994)	84.7%	7 (70%)	77.35%
Data-Driven Homologue Matching (Stanley, et al., 1998)	89.1%	4 (40%)	64.55%
Fuzzy Logic Rule-Based System (Keller, et al., 1995)	87%	4 (40%)	63.50%
Quadratically Asymmetric Statistical Distributions (Ritter & Gaggermeier, 1999)	82.5%	5 (50%)	66.25%
Joint Classification and Pairing (Biyani, et al., 2000)	98.1%	5 (50%)	74.05%

Although the most accurate chromosome classifier was the joint classification and pairing technique presented by Biyani, et al. (2005), its lack of support of important system measures has limited its effectiveness. The main limitation of this technique is the validity of system testing and reliance on human interaction. The data used in training and testing this implementation was already pre-processed, as only values for chromosome features were used. This implementation does not offer automation of any of the image processing techniques, such as feature extraction and chromosome segmentation and therefore places a large reliance on human expert interaction. The

PNN implementation presents a very effective technique for classifying human chromosomes but is only partially automated and therefore still requires extensive time and effort from cytogeneticists. The other implementations did have their individual strengths and limitations but did not have a high overall rating.

The artificial neural network implementation presented by Lerner (1998) presents a completely automated approach to classify chromosomes in which feature extraction and chromosome segmentation as well as final classification are all computerised. This ANN approach out-performed other implementations in regard to system measures. Lerner's ANN approach accepts chromosome images containing overlapping and touching chromosomes, and consequently the data is more valid. Additionally, the ANN approach provides good generalising abilities and performs well when data is incomplete or uncertain, which is common in metaphase images (Cho, 2000, p. 28; Wang, et al., 2002, p. 2538). Overall, the ANN approach by Lerner (1998) provides a well-rounded implementation, offering a good balance between classification accuracy and system features and therefore portraying the most effective technique for automated chromosome classification.

6.3.2. Strengths and Limitations of this Technique

The use of an ANN for chromosome classification presents several advantages including robustness and the ability to generalise. Additionally, the approach used by Lerner (1998) resulted in a fully automatic system with little or no reliance on human experts and a well-trained system which could easily be applied within the medical field without much need for adaptation.

The ANN approach by Lerner (1998), however, does present several limitations. One such limitation is computational burden. A neural network trained using a back-propagation algorithm requires extensive training and consequently takes up much time (Kartalopoulos, 1996, p. 81). However, this limitation is outweighed by the strong support of additional system features and satisfactory classification performance.

Another limitation of artificial neural networks is their 'black box' structure, thus preventing them from explaining or validating the given outputs (Dybowski, 2000, p. 31). For many implementation examples, a system should be able to explain the given results and display the reasoning behind the outputs. Again, this characteristic does not

pose a large limitation as the reasoning behind assigning chromosomes to classes is not as important as producing a correct karyotype.

The following chapter, Conclusions and Recommendations, presents a discussion on the possible improvements for automated chromosome systems. Also, the conclusions drawn from this research are discussed and the original research questions are re-assessed as related to the results achieved.

7. CONCLUSIONS AND RECOMMENDATIONS

7.1. Research Outcomes

This research has conducted an analysis of various implementation techniques for automated human chromosome classification. The different implementations have included techniques using artificial neural networks, techniques using artificial neural networks supported by another method and techniques not using artificial neural networks. The main research question structuring this investigation was: *Are artificial neural networks a suitable implementation technique for automated chromosome analysis?* In order to address this research question, the two sub-questions identified must be considered first.

7.1.1. Classification Accuracy

The first sub-question framing this research was *How do ANNs perform in classification accuracy as compared to other implementations?* In comprehensively addressing this question, this research has analysed the various experiments conducted on the identified implementations and has compared the accuracy results. The outcome of this investigation found that artificial neural networks performed well as automated chromosome classifiers, but did not perform as well as other techniques. The best classification accuracy achieved by a neural network approach was 97% from the probabilistic neural network. However, the statistical approach offering a joint classification and pairing technique achieved a 98.1% misclassification. Therefore, in addressing this question, this research has found that artificial neural networks do not perform as well as other techniques when only classification accuracy is considered. However, the implementations presented have been based on various data sets and differing testing experiments and comparing only classification accuracy does not present a complete analysis of the systems. Hence, to conduct a well-rounded comparison, the system measures offered by each implementation must also be compared; these are discussed next.

7.1.2. System Measures

The second sub-question identified for this research was *How do ANN classifiers perform in system measures as compared to classifiers based on other processing methods?* The system measures that were analysed include the ability to generalise,

robustness, efficiency in computational burden and speed, validity in real-world data and degree of human interaction required. In terms of these system measures, artificial neural networks proved to be an effective implementation as compared to the other techniques presented in this research. Artificial neural networks offer good generalising abilities and stability and robustness even when dealing with incomplete data. The ANN implementation by Lerner (1998) offered a completely automated approach to chromosome classification and thus required little or no human interaction and maintained validity with real-world data. Several other implementations presented in this research did not put forward completely automated systems and relied heavily on pre-extracted chromosome features and consequently human interaction.

7.1.3. Overall Assessment

This research has attempted to investigate whether the artificial neural networks are an effective computing technique for human chromosome classification. The investigation has found that artificial neural networks offer acceptable classification accuracy while maintaining high support of desirable features. Wang, et al. (2005) claim that artificial neural networks are a popular tool for detecting and classifying chromosomes (p. 2540) This research has supported the above claim; it is found that artificial neural networks do indeed present an effective technique for human chromosome classification.

7.2. Limitations

This research has investigated several different implementation approaches for automated chromosome classification. However, the research presented does have several limitations.

This research has considered a small representative of the various implementation techniques available for chromosome analysis. From these case studies, it has been shown that artificial neural networks are an effective technique for human chromosome classification. This gives room for more research work, as the various implementation techniques not considered in this study could be investigated and compared to the effectiveness of ANNs.

Although all case studies were concerned with chromosome classification, the implementations used different testing data, different testing experiments and different classification processes. This introduces a measure of complexity into the comparisons as a level evaluation is not directly applicable. However, this research has shed light on the use of both classification accuracy and system measures as criteria for thoroughly evaluating performance of automated chromosome classifiers.

7.3. Recommendations and Further Work

The most effective artificial neural network implementation identified through this study has produced satisfactory results in both accuracy and system measures. However, there are several recommendations and possible improvements, not only for this implementation but for the field of automated chromosome classification as a whole. These improvements include advances in the imaging techniques and in implementation techniques.

7.3.1. Imaging Techniques

Automated chromosome analysis begins with an image of a metaphase cell. Therefore, the quality of that image will affect the entire classification process. In order to facilitate and improve classification accuracy, enhancements in imaging techniques are required. Wang, et al. (2005) claim “the performance of the systems can be improved when the slides are well-prepared, the microscope has good optical quality and the camera can digitize the image with sufficient clarity and resolution” (p. 2540). A clear, high resolution image will provide a more accurate representation of the chromosomes in that cell and thus will lead to a more accurate classification due to lower occurrences of noise and uncertainty within that image.

7.3.2. Image Processing Techniques

Another area for improvement in automated karyotyping systems lies within the task of processing the metaphase cell image. As chromosomes are commonly overlapping or touching, improvements in these image processing techniques will consequently lead to improvements in classification accuracy. Carothers and Piper (1994) identify “poorly segmented or severely distorted chromosomes” as a cause of high error rates (p. 169). Ritter and Gaggermeier (1999) support the above and confirm that “in order to remove more classification errors it will, however, be necessary to take

another look at image processing” (p. 1007). Wang, et al. (2005) confirm that the performance of automated karyotyping systems is directly influenced by the results of chromosome segmentation (p. 2540) by observing that error rates were substantially increased when classification involved touching or overlapping chromosomes (p. 2541).

7.3.3. Feature Selection

The selection of chromosome features also plays an important role in final classification accuracy. Tso and Graham (1983) suppose that “better discrimination might be achieved by including more chromosome measurements” (p. 495). This is supported by Piper and Granum (1989) who also note that possible enhancements in classification accuracy can be achieved through improved feature selection (p. 254). These assumptions have been supported by Graham and Errington (2000) who observe that the improvement of Lerner’s (1998) ANN to their ANN implementation “might be due to the more carefully chosen density features” (p. 262).

7.3.4. Network Architecture

A further area of improvement for automated classification accuracy lies within the implemented neural network structure. Cho (2000), in reference to the back-propagation training algorithm, claims “better training algorithms to reduce training times are needed” (p. 32). A faster and less complex training process would improve the usefulness of artificial neural networks in real-time applications.

Additionally, a change in the method of reporting results could be investigated. Several implementations presented in this research produce a karyotype of all chromosome cells and rely on the cytogeneticist to review the outputs in order to identify possible wrong classifications. A practical improvement on this method would be the generation of a system that classifies all possible chromosomes and alerts the human operator to potential errors or abnormalities when faced with a difficult or unlikely classification. Stanley, et al. (1998) support the above by arguing that “with the purpose of aiding a cytogenetic expert, making no decision for chromosome assignment is better than an incorrect assignment” (p. 452).

Finally, as technology constantly advances and improves, the applications using it will grow alongside it. Artificial neural networks have been presented as a viable and

effective technique for the classification of human chromosomes. Enhancements in the technology of imaging techniques and further research into the artificial neural network frameworks will arguably result in improvements in their applicability.

REFERENCES

- Azuaje, F., Dubitzky, W., Lopes, P., Black, N., Adamson, K., Wu, X., & White, J.A. (1999). Predicting coronary disease risk based on short-term RR interval measurements: A neural network approach [Electronic Version]. *Artificial Intelligence in Medicine*, 15, 275-297.
- Baxt, W.G. (1995). Application of artificial neural networks to clinical medicine. [Electronic Version] *The Lancet*, 346, 1135-1138.
- Biyani, P., Wu, X., & Sinha, A. (2005). Joint classification and pairing of human chromosomes [Electronic Version]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2(2), 102-109.
- Carothers, A., & Piper, J. (1994). Computer-aided classification of human chromosomes: A review. *Statistics and Computing*, 4, 161-171.
- Catto, J.W.F., Linkens, D.A., Abbod, M.F., Chen, M., Burton, J.L., Feeley, K.M., & Hamdy, F.C. (2003). Artificial Intelligence in Predicting Bladder Cancer Outcome: A Comparison of Neuro-Fuzzy Modeling and Artificial Neural Networks [Electronic Version]. *Clinical Cancer Research*, 9, 4172-4177.
- Charters, G.C., & Graham, J. (2002). Disentangling chromosome overlaps by combining trainable shape models with classification evidence [Electronic Version]. *IEEE Transactions on Signal Processing*, 50(8), 2080-2085.
- Cho, J.M. (2000). Chromosome classification using backpropagation neural networks: A process that overcomes nonlinearity problems to correctly classify giemsa-stained human chromosomes [Electronic Version]. *IEEE Engineering in Medicine and Biology*, 19(1), 28-33.
- Coiera, E.W. (1996). Artificial intelligence in medicine: The challenges ahead. [Electronic Version] *Journal of the American Medical Informatics Association*, 3, 363-366.

- Creswell, J. (2003). *Research design: Qualitative, quantitative and mixed method approaches* (2nd Ed). Thousand Oaks, California: Sage Publications Limited.
- De Robertis, E.D.P., & De Robertis, E.M.F. (1980). *Cell and molecular biology* (7th Ed). Philadelphia: Saunders College.
- Dybowski, R. (2000). Neural computation in medicine: Perspectives and Prospects. In: Malmgren, H., Borga, M., & Niklasson, L. (Eds). *Artificial neural networks in medicine and biology: Proceedings of the ANNIMAB-1 Conference, Göteborg, Sweden, 13-16 May 2000* (p. 26-36).
- Eliason, S.R. (1993). *Maximum likelihood estimation: Logic and practice*. (Sage University Paper series on Quantitative Applications in the Social Sciences, 07-096). Newbury Park, California: Sage Publications Limited.
- Emery, A.E.H., & Mueller, R.F. (1988). *Elements of medical genetics* (7th Ed). Edinburgh: Churchill Livingstone.
- Graham, J., & Errington, P.A. (2000). Classification of chromosomes: A comparative study of neural network and statistical approaches. In: Lisboa, P.J.G., Ifeakor, E.C., & Szczepaniak, P.S. (Eds). *Artificial neural networks in biomedicine* (p. 249-265). London: Springer-Verlag Limited.
- Ji, L. (1994). Fully automatic chromosome segmentation [Electronic Version]. *Cytometry*, 17, 196-208.
- Jorde, L.B., Carey, J.C., Bamshad, M.J., & White, R.L. (2000). *Medical genetics* (2nd Ed). St. Louis, Missouri: Mosby Inc.
- Kangas, L.J., & Keller, P.E., (2000). Neurometric assessment of adequacy of intraoperative anaesthetic. In: Lisboa, P.J.G., Ifeakor, E.C., Szczepaniak, P.S. (Eds). *Artificial neural networks in biomedicine* (p. 81-91). London: Springer-Verlag Limited.

- Kartalopoulos, S.V. (1996). *Understanding neural networks and fuzzy logic: Basic concepts and applications*. New Jersey: IEEE Press.
- Keller, J.M., Gader, P., Sjahputera, O., Caldwell, C.W., Huang, H-M.T. (1995). A fuzzy logic rule-based system for chromosome recognition [Electronic Version]. *Eighth IEEE Symposium on Computer-Based Medical Systems (CBMS'95)* (p. 125-132).
- Leedy, P.D., & Ormrod, J.E. (2005). *Practical research: Planning and design* (8th Ed). Upper Saddle River, New Jersey: Pearson Education Limited.
- Lerner, B. (1998). Toward a completely automatic neural-network-based human chromosome analysis [Electronic Version]. *IEEE Transactions on Systems, Man and Cybernetics*, 28(4), 544-552.
- Levitan, M. (1988). *Textbook of human genetics*. New York: Oxford University Press.
- Lisboa, P.J.G., Ifeakor, E.C., & Szczepaniak, P.S. (Eds). (2000). *Artificial neural networks in biomedicine*. London: Springer-Verlag Limited.
- Lo, S-C.B., Lin, J-S.J., Freedman, M.T., & Mun, S.K. (1998). Application of artificial neural networks to medical image pattern recognition: Detection of clustered microcalcifications on mammograms and lung cancer on chest radiographs [Electronic Version]. *Journal of VLSI Signal Processing*, 18, 263-274.
- Martínez, C., Juan, A., & Casacuberta, F. (2002). Using recurrent neural networks for automatic chromosome classification [Electronic Version]. In J.R. Dorronsoro (Eds). *Artificial neural networks: International conference on artificial neural networks 2002*, Lecture notes in computer science (p. 565-570). Springer-Verlag.
- NIST/SEMATECH e-Handbook of Statistical Methods (n.d. a). Retrieved November 3, 2005, from <http://www.itl.nist.gov/div898/handbook/apr/section4/apr412.htm>

- NIST/SEMATECH e-Handbook of Statistical Methods (n.d. b). Retrieved November 3, 2005, from <http://www.itl.nist.gov/div898/handbook/pmd/section4/pmd453.htm>
- NIST/SEMATECH e-Handbook of Statistical Methods (n.d. c). Retrieved November 3, 2005, from <http://www.itl.nist.gov/div898/handbook/apr/section4/apr422.htm>
- Negnevitsky, M. (2002). *Artificial intelligence: A guide to intelligent systems*. Essex, England: Pearson Education Limited.
- Nguyen, H.T., & Walker, E.A. (1997). *A first course in fuzzy logic*. Boca Raton, Florida: CRC Press.
- Patterson, D.W. (1990). *Introduction to artificial intelligence and expert systems*. Englewood Cliffs, New Jersey: Prentice Hall Inc.
- Patterson, D.W. (1996). *Artificial neural networks: Theory and applications*. Singapore: Prentice Hall Limited.
- Perry, G.J., Roderer, N.K., Assar, S. (2005). A current perspective on medical informatics and health sciences librarianship [Electronic Version]. *Journal of the Medical Library Association*, 93, 199-205.
- Piper, J., & Granum, E. (1989). On fully automatic feature measurement for banded chromosome classification [Electronic Version]. *Cytometry*, 10, 242-255.
- Popescu, M., Gader, P., Keller, J., Klein, C., Stanley, J., & Caldwell, C. (1999). Automatic karyotyping of metaphase cells with overlapping chromosomes [Electronic Version]. *Computers in Biology and Medicine*, 29, 61-82.
- Punch, K.F. (1998). *Introduction to social research: Quantitative and qualitative approaches*. London: Sage Publications Limited.
- Ritter, G., & Gaggermeier, K. (1999). Automatic classification of chromosomes by means of quadratically asymmetric statistical distributions. *Pattern Recognition*, 32, 997-1008.

- Silipo, R., & Marchesi, C. (1998). Artificial neural networks for automatic ECG analysis [Electronic Version]. *IEEE Transactions on Signal Processing*, 46(5), 1417-1425.
- Snustad, D.P. & Simmons, M.J. (2000). *Principles of genetics* (2nd Ed). New York: John Wiley & Sons.
- Stake, R.E. (2003). Case studies. In: Denzin, N.K., & Lincoln, Y.S. (Eds). *Strategies of qualitative inquiry* (2nd Ed., p. 134-164). Thousand Oaks: Sage Publications Limited.
- Stanley, R.J., Keller, J.M., Gader, P., Caldwell, C.W. (1998). Data-driven homologue matching for chromosome identification [Electronic Version]. *IEEE Transactions on Medical Imaging*, 17(3), 451-462.
- Sweeney, W.P., Musavi, M.T., & Guidi, J.N. (1994). Classification of chromosomes using a probabilistic neural network [Electronic Version]. *Cytometry*, 16, 17-24.
- Tso, M., & Graham, J. (1983). The transportation algorithm as an aid to chromosome classification. *Pattern Recognition Letters*, 1, 489-496.
- Tso, M., Kleinschmidt, P., Mitter, I., & Graham, J. (1991). An efficient transportation algorithm for automatic chromosome karyotyping. *Pattern Recognition Letters*, 12, 117-126.
- Wang, X., Zheng, B., Wood, M., Li, S., Chen, W., & Liu, H. (2005). Development and evaluation of automated systems for detection and classification of banded chromosomes: current status and future perspectives [Electronic Version]. *Journal of Physics D: Applied Physics*, 38, 2536-2542.
- Wasserman, P.D. (1983). *Advanced methods in neural computing*. New York: Van Nostrand Reinhold.
- Winchester, A.M., & Mertens, T.R. (1983). *Human genetics* (4th Ed). Ohio: Charles E. Merrill Publishing Company.

Yin, R.K. (1994). *Case study research: Design and methods* (2nd Ed). Thousand Oaks, California: Sage Publications Limited.

Zhou, Z., Jiang, Y., Yang, Y., & Chen, S. (2002). Lung cancer cell identification based on artificial neural network ensembles. [Electronic Version] *Artificial Intelligence in Medicine*, 24, 25-36.

APPENDIX A

This section describes and presents the rating framework used in this research. The rating framework was developed to serve as a method of standardising the comparisons of all presented case studies. The system measures under comparison are

1. Ability to generalise
2. Robustness
3. Efficiency in computational burden and speed
4. Validity
5. Degree of Human Reliance

In regards to validity, two main kinds of validity are considered. These are:

1. Testing data validity: this considers whether the data used in testing the implementation is viable and applicable with real-world data.
2. Testing experiments validity: this factor is considers whether the experiments conducted on the implementation are the type of functionality required for a typical real-world automated chromosome classifier.

The rating values are in the range of [0,2]. A rating of 0 implies that the implementation does not offer good support of the measure in question while a rating of 2 indicates a good presence of that system measure. Given that there are a total of five system measures under comparison, the total possible score is ten (two points for each measure; this includes validity, where each sub-factor is rated out of only one point, giving a total of two points).

After allocating a complete rating for each implementation, the system measure ratings are converted into a percentage (e.g. 8/10 converts to 80%). That percentage is combined with the performance accuracy percentage and each factor is given equal weighting. The sum of these percentages is then divided in half to give an overall rating of the implementation under comparison.

A.1. Ratings of ANN Implementations

A.1.1. First Case Study

Case Study: Classification of chromosomes: A comparative study of neural network and statistical approaches (Graham & Errington, 2000).

FACTOR	RATING
Ability to Generalise	2
Robustness	2
Efficiency in Computational Burden & Speed	0
Validity <i>Testing Data</i>	0
<i>Testing Experiments</i>	1
Degree of Human Reliance	0
TOTAL	5/10

Explanations:

Each of the features provided above will be briefly discussed and validated here.

- A. Ability to generalise: The ANN implementation by Graham & Errington (2000) has been assigned a 2 for its ability to generalise. Patterson (2000) claims that generalisation is an important characteristic of artificial neural networks (p. 25) and Negnevitsky supports this claim by noting that ANNs are efficient at generalising and adapting to new input data (p. 259).

- B. Robustness: A score of 2 has also been assigned to the measure of robustness. This is validated by Patterson (1996), who reports that ANNs “continue to perform well when part of the network is disabled or presented with noisy data” (p. 27). Negnevitsky (2002) also confirms that artificial neural networks are capable of tolerating uncertainty and imprecision in data (p. 259).

- C. Efficiency in Computational Burden and Speed: Kartalopoulos (1996) notes that “the algorithm suffers from extensive calculations and, hence, slow training

speed” (p. 81). Therefore, a score of 0 has been assigned to this factor as training neural networks using the back-propagation algorithm is computationally expensive.

D. Validity:

- a. Testing Data: This ANN implementation has relied upon the use of pre-extracted chromosome features and therefore has been assigned a 0 in regard to the use of valid testing data.
- b. Testing Experiments: Testing this implementation focused on the task of analysing a complete set of chromosomes and consequently producing a karyotype, which is the most popular chromosome analysis technique (Wang, et al., 2005, p. 2536). This implementation was allocated a score of 1 for its support of valid testing experiments.

E. Degree of Human Reliance: This ANN implementation has not presented any techniques for automated image processing or automated chromosome segmentation and is largely reliant upon human operators to conduct these tasks. It is therefore assigned a 0 for this feature.

In total, the ANN implementation provided by Graham and Errington (2000) has scored a total of five points out of a possible ten. In completing the comparison, this score is converted to a percentage and then combined with the classification accuracy to determine an overall rating. This is depicted below:

$$\begin{aligned} \text{Overall Rating} &= (\text{Classification Accuracy} + \text{System Measures Rating}) \div 2 \\ &= (85.6\% + 50\%) \div 2 \\ &= 67.80\% \end{aligned}$$

Therefore, the ANN implementation by Graham and Errington (2000) has been assigned an overall rating of 67.80%.

A.1.2. Second Case Study

Case Study: Toward a Completely Automatic Neural-Network-Based Human Chromosome Analysis (Lerner, 1998);

FACTOR	RATING
Ability to Generalise	2
Robustness	2
Efficiency in Computational Burden & Speed	0
Validity: <i>Testing Data</i>	1
<i>Testing Experiments</i>	1
Degree of Human Reliance	2
TOTAL	8/10

Explanations:

Each of the features provided above will be briefly discussed and validated here.

- A. Ability to generalise: The ANN implementation by Lerner (1998) has also been assigned a 2 for its ability to generalise. For a discussion on the reasoning behind this rating, please refer to Section A.1.1 in Appendix A.
- B. Robustness: A score of 2 has also been assigned to the measure of robustness. Again, the reasoning behind this rating has been previously discussed; please refer to Section A.1.1 in Appendix A.
- C. Efficiency in Computational Burden and Speed: This ANN implementation has been given a score of 0 for the factor of efficiency in computational burden and speed. The reasoning behind this rating is also previously discussed in Section A.1.1 in Appendix A.
- D. Validity:
 - a. Testing Data: This ANN implementation has used images of human metaphase cells and therefore is valid in terms of real-life data. It has been given a score of 1 for this factor.

- b. Testing Experiments: This implementation was allocated a score of 1 for its support of valid testing experiments. The reasoning behind this rating is also provided in the explanations given to the ANN implementation by Graham and Errington (2000).

E. Degree of Human Reliance: This ANN implementation has presented a completely automated system for human chromosome classification and therefore has provided techniques for automatic feature extraction and chromosome segmentation. It has therefore been assigned a rating of 2 points.

In total, the ANN implementation presented by Lerner (1998) has scored a total of eight points out of a possible ten. In completing the comparison, this score is converted to a percentage and then combined with the classification accuracy to determine an overall rating. This is depicted below:

$$\begin{aligned} \text{Overall Rating} &= (\text{Classification Accuracy} + \text{System Measures Rating}) \div 2 \\ &= (84.5\% + 80\%) \div 2 \\ &= 82.75\% \end{aligned}$$

Therefore, the ANN implementation by Lerner (1998) has been assigned an overall rating of 82.75%.

A.1.3. Third Case Study

Case Study: Classification of Chromosomes Using a Probabilistic Neural Network (Sweeney, Musavi & Guidi, 1994)

FACTOR	RATING
Ability to Generalise	2
Robustness	2
Efficiency in Computational Burden & Speed	2
Validity: <i>Testing Data</i>	0
<i>Testing Experiments</i>	1
Degree of Human Reliance	0
TOTAL	7/10

Explanations:

Each of the features provided above will be briefly discussed and validated here.

- A. Ability to generalise: The PNN implementation by Sweeney, et al. (1994) has been assigned a 2 for its ability to generalise. This is due to the fact that Sweeney, et al. (1994) state “the network generalizes to any new incoming training patterns without having to repeat an extensive training process” (p. 18).
- B. Robustness: A score of 2 has also been assigned to the measure of robustness. This is validated by Patterson (1996) who states that “PNN networks also tolerate noisy samples and they can work with sparse samples too” (p. 354).
- C. Efficiency in Computational Burden and Speed: This PNN implementation has been given a score of 2 has been assigned for the factor of efficiency in computational burden and speed. Sweeney, et al. (1994) argue that “the significant advantages of the PNN classifier are its speed and simplicity of the training process” (p. 18). Patterson (1996) supports the above by stating “one of the main advantages of the PNN is the speed with which it can be trained” (p. 354).

D. Validity:

- a. Testing Data: This PNN implementation has relied upon the use of pre-extracted chromosome features and therefore has been assigned a 0 for its use of valid testing data.
- b. Testing Experiments: This implementation was allocated a score of 1 for its support of valid testing experiments as it was focused on assessing the performance of the system in analysing a complete set of chromosomes and producing a full karyotype.

E. Degree of Human Reliance: This PNN implementation has not presented any techniques for automated image processing or automated chromosome segmentation and is largely reliant upon human operators to conduct these tasks. It is therefore assigned a 0 for this feature.

In total, the PNN implementation presented by Sweeney, et al. (1994) has scored a total of seven points out of a possible ten. In completing the comparison, this score is converted to a percentage and then combined with the classification accuracy to determine an overall rating. This is depicted:

$$\begin{aligned} \text{Overall Rating} &= (\text{Classification Accuracy} + \text{System Measures Rating}) \div 2 \\ &= (84.7\% + 70\%) \div 2 \\ &= 77.35\% \end{aligned}$$

Therefore, the PNN implementation by Sweeney, et al. (1994) has been assigned an overall rating of 77.35%.

A.2. Ratings of Implementations Using ANNs Supported by Other Techniques

A.2.1. First Case Study

Case Study: Data-driven homologue matching for chromosome identification (Stanley, Keller, Gader & Caldwell, 1998)

FACTOR	RATING
Ability to Generalise	2
Robustness	1
Efficiency in Computational Burden & Speed	0
Validity: <i>Testing Data</i>	0
<i>Testing Experiments</i>	0
Degree of Human Reliance	1
TOTAL	4/10

Explanations:

Each of the features provided above will be briefly discussed and validated here.

- A. Ability to generalise: The use of dynamic programming allows the implementation to adapt to the data at hand and neural networks are effective in generalising (Negnevitsky, 2002, p. 259). Therefore, this implementation has been given a score of 2 for generalisability.

- B. Robustness: A score of 1 has also been assigned to the measure of robustness. Neural networks are capable of accepting incomplete or ambiguous data (Patterson, 1996, p. 27), while algorithmic programmes generally require exact parameters in order to function.

- C. Efficiency in Computational Burden and Speed: This implementation has been given a rating of 0 for efficiency as training neural networks is computationally expensive (Kartalopoulos, 1996, p. 81)

D. Validity:

- a. Testing Data: This implementation has relied upon the use of pre-extracted chromosome features and therefore has been assigned a 0 in regard to the use of valid testing data.
- b. Testing Experiments: Stanley, et al. (1998) only tested their system on the ability to identify one chromosome class and did not assess the performance of producing a complete karyotype. Therefore, the rating for this factor is 0.

E. Degree of Human Reliance: This implementation has not presented any techniques for automated image processing or automated chromosome segmentation and is largely reliant upon human operators to conduct these tasks. It is therefore assigned a 0 for this feature.

In total, the data-driven homologue matching technique presented by Stanley, et al. (1998) has scored a total of four points out of a possible ten. In completing the comparison, this score is converted to a percentage and then combined with the classification accuracy to determine an overall rating. This is depicted:

$$\begin{aligned} \text{Overall Rating} &= (\text{Classification Accuracy} + \text{System Measures Rating}) \div 2 \\ &= (89.1\% + 40\%) \div 2 \\ &= 64.55\% \end{aligned}$$

Therefore, the data-driven homologue matching implementation by Stanley, et al. (1998) has been assigned an overall rating of 64.55%.

A.2.2. Second Case Study

Case Study: A fuzzy logic rule-based system for chromosome recognition (Keller, Gader, Sjahputera, Caldwell & Huang, 1995)

FACTOR	RATING
Ability to Generalise	1
Robustness	2
Efficiency in Computational Burden & Speed	1
Validity: <i>Testing Data</i>	0
<i>Testing Experiments</i>	0
Degree of Human Reliance	0
TOTAL	4/10

Explanations:

Each of the features provided above will be briefly discussed and validated here.

- A. Ability to generalise: Fuzzy systems do not generalise and adapt well when faced with new data (Negnevitsky, 2002, p. 267). However, neural networks do generalise well. As only half the implementation technique supports the ability to generalise, this implementation has been assigned a score of 1 for generalisability.
- B. Robustness: The neural network approach supported by fuzzy logic has been given a score of 2 for robustness. Both neural network and fuzzy logic are well equipped to handle incomplete or ambiguous data (Patterson, 1996, p. 27; Negnevitsky, 2002, p. 87)
- C. Efficiency in Computational Burden and Speed: Training a neural network can be computationally expensive; therefore, this fuzzy-neural implementation has been given a score of 1 for the factor of efficiency in computational burden and speed.

D. Validity:

- a. Testing Data: This implementation has relied upon the use of pre-extracted chromosome features and therefore has been assigned a 0 in regard to the use of valid testing data.
- b. Testing Experiments: In testing this implementation, the experiments involved identifying chromosomes from only two out of the total 23 classes and did not test the functionality of the system in producing a full karyotype. Therefore, a rating of 0 has been assigned to this factor.

E. Degree of Human Reliance: This fuzzy-neural implementation has not presented any techniques for automated image processing or automated chromosome segmentation and is largely reliant upon human operators to conduct these tasks. It is therefore assigned a 0 for this feature.

In total, the fuzzy-neural implementation presented by Keller, et al. (1995) has scored a total of four points out of a possible ten. In completing the comparison, this score is converted to a percentage and then combined with the classification accuracy to determine an overall rating. This is depicted:

$$\begin{aligned} \text{Overall Rating} &= (\text{Classification Accuracy} + \text{System Measures Rating}) \div 2 \\ &= (87\% + 40\%) \div 2 \\ &= 63.50\% \end{aligned}$$

Therefore, the fuzzy logic implementation supported by neural networks as presented by Keller, et al. (1995) has been assigned an overall rating of 63.50%.

A.3. Ratings of Non-ANN Implementations

A.3.1. First Case Study

Case Study: Automatic classification of chromosomes by means of quadratically asymmetric statistical distributions (Ritter & Gaggermeier, 1999)

FACTOR	RATING
Ability to Generalise	2
Robustness	2
Efficiency in Computational Burden & Speed	0
Validity: <i>Testing Data</i>	0
<i>Testing Experiments</i>	1
Degree of Human Reliance	0
TOTAL	5/10

Explanations:

Each of the features provided above will be briefly discussed and validated here.

- A. Ability to generalise: The use of maximum likelihood estimation, as presented by Ritter and Gaggermeier (1999), supports the ability to generalise. Eliason (1993) indicates that “as the sample size grows large, the MLE tends toward the properties of an unbiased estimator” (p. 20). Therefore, this implementation has been given a rating of 2 for this factor.
- B. Robustness: This implementation has been assigned a score of 2 for its support of robustness. The NIST/SEMATECH e-Handbook of Statistical Methods (n.d. b) claims that the use of maximum likelihood estimation can accept unknown or incomplete information.
- C. Efficiency in Computational Burden and Speed: The maximum likelihood estimation implementations are based on complex algorithms and therefore can be computationally expensive (The NIST/SEMATECH e-Handbook of

Statistical Methods, n.d. c). Therefore, this implementation has been given a score of 0 for efficiency.

D. Validity:

- a. Testing Data: This implementation has relied upon the use of pre-extracted chromosome features and therefore has been assigned a 0 in regard to the use of valid testing data.
- b. Testing Experiments: Testing this implementation focused on the task of analysing a complete set of chromosomes and consequently producing a karyotype, which is the most popular chromosome analysis technique (Wang, et al., 2005, p. 2536). This implementation was allocated a score of 1 for its support of valid testing experiments.

E. Degree of Human Reliance: This statistical implementation has not presented any techniques for automated image processing or automated chromosome segmentation and is largely reliant upon human operators to conduct these tasks. It is therefore assigned a 0 for this feature.

In total, implementation presented by Ritter and Gaggermeier (1999) has scored a total of five points out of a possible ten. In completing the comparison, this score is converted to a percentage and then combined with the classification accuracy to determine an overall rating. This is depicted:

$$\begin{aligned} \text{Overall Rating} &= (\text{Classification Accuracy} + \text{System Measures Rating}) \div 2 \\ &= (82.5\% + 50\%) \div 2 \\ &= 66.25\% \end{aligned}$$

Therefore, the implementation presented by Ritter and Gaggermeier (1999) has been assigned an overall rating of 66.25%.

A.3.2. Second Case Study

Case Study: Joint classification and pairing of human chromosomes (Biyani, Wu & Sinha, 2005)

FACTOR	RATING
Ability to Generalise	2
Robustness	2
Efficiency in Computational Burden & Speed	0
Validity: <i>Testing Data</i>	0
<i>Testing Experiments</i>	1
Degree of Human Reliance	0
TOTAL	5/10

Explanations:

Each of the features provided above will be briefly discussed and validated here.

- A. Ability to generalise: Biyani, et al. (2005) use maximum likelihood estimation as the basis for chromosome classification and pairing. This supports the ability to generalise and thus has been assigned a score of 2. Please refer to Section A.3.1 in Appendix A for a validation of this rating.
- B. Robustness: This implementation has been assigned a score of 2 for its support of robustness. For validation of this rating, please refer to Section A.3.1 in Appendix A.
- C. Efficiency in Computational Burden and Speed: The use of maximum likelihood estimation does present problems in efficiency and therefore this implementation has been assigned a rating of 0 for this factor. Please refer to Section A.3.1 in Appendix A for a validation of this rating.

D. Validity:

- a. Testing Data: This implementation has relied upon the use of pre-extracted chromosome features and therefore has been assigned a 0 in regard to the use of valid testing data.
- b. Testing Experiments: Testing this implementation focused on the task of analysing a complete set of chromosomes and consequently producing a karyotype, which is the most popular chromosome analysis technique (Wang, et al., 2005, p. 2536). This implementation was allocated a score of 1 for its support of valid testing experiments.

E. Degree of Human Reliance: This statistical implementation has not presented any techniques for automated image processing or automated chromosome segmentation and is largely reliant upon human operators to conduct these tasks. It is therefore assigned a 0 for this feature.

In total, the joint classification and pairing technique presented by Biyani, et al. (2005) has scored a total of five points out of a possible ten. In completing the comparison, this score is converted to a percentage and then combined with the classification accuracy to determine an overall rating. This is depicted:

$$\begin{aligned} \text{Overall Rating} &= (\text{Classification Accuracy} + \text{System Measures Rating}) \div 2 \\ &= (98.1\% + 50\%) \div 2 \\ &= 74.05\% \end{aligned}$$

Therefore, the implementation presented by Biyani, et al. (2005) has been assigned an overall rating of 74.05%.

APPENDIX B

B.1. GLOSSARY

- Artificial Intelligence:*** “The goal of artificial intelligence (AI) as a science is to make machines do things that would require intelligence if done by humans” (Boden, 1997, cited in Negnevitsky, 2002, p. 2)
- Artificial Neural Networks:*** “Simplified models of the central nervous system. They are networks of highly interconnected neural computing elements that have the ability to respond to input stimuli and to learn to adapt to the environment” (Patterson, 1996, p. 1)
- Banding Pattern:*** The banding pattern generally identifies the number of bands in the chromosome, the distance between each band, the distance between the bands and the centromere region and the density of each band (Keller, Gader, Sjahputera, Caldwell, & Huang, 1995, p. 127) – see also Chromosome Band
- Centromere:*** The centromere represents the area of the chromosome where the two chromatid sisters join together (Snustad & Simmons, 2000, p. 27)
- Centromeric Index:*** “The ratio of the length of the short arm to the whole length of the chromosome” (Cho, 2000, p. 29)
- Chromosome:*** Thread-like bodies consisting of DNA material arranged in sections called genes. Humans have 46 chromosomes in a normal cell (Jorde, Carey, Bamshad & White, 2000, p. 6)
- Chromosome Band:*** “A part of a chromosome that is clearly distinguishable from its adjacent segments by appearing darker or lighter

as a result of the new staining methods” (Levitan, 1988, p. 32)

- Cytogenetics:*** The study of chromosomes and their abnormalities (Jorde, Carey, Bamshad & White, 2000, p. 108)
- Denver Groups:*** Seven chromosome groups (Group A – Group G) first identified at a medical conference in Denver, in 1960. Chromosomes are arranged in these groups in decreasing order of size (Levitan, 1988, p. 28)
- Homologue:*** “A chromosome pair” (Stanley, Keller, Gader & Caldwell, 1998, p. 451)
- Karyotype:*** “A layout of chromosome images organised by decreasing size in pairs” (Lerner, 1998, p. 544)
- Karyotyping Constraint:*** The karyotyping constraint specifies that “there are exactly two chromosomes in (almost) all classes” (Graham & Errington, 2000, p. 258)
- Metaphase:*** “The stage of a cell at which the chromosomes are most suitable for analysis” (Lerner, 1998, p. 544)
- Transportation Algorithm:*** The transportation algorithm is commonly used for finding the most economical route passing through predetermined destinations (Patterson, 1996, p. 298) and is applied in chromosome classification to implement the karyotyping constraint – See also Karyotyping Constraint.