Automated Analysis of Mammograms using Evolutionary Algorithms

Yuan Zhang

MPhil

University of York

Electronics

September 2012

Abstract

Breast cancer is the leading cause of death in women in the western countries. The diagnosis of breast cancer at the earlier stage may be particularly important since it provides early treatment, this will decreases the chance of cancer spreading and increase the survival rates. The hard work is the early detection of any tissues abnormal and confirmation of their cancerous natures. In additionally, finding abnormal on very early stage can also affected by poor quality of image and other problems that might show on a mammogram.

Mammograms are high resolution x-rays of the breast that are widely used to screen for cancer in women. This report describes the stages of development of a novel representation of Cartesian Genetic programming as part of a computer aided diagnosis system. Specifically, this work is concerned with automated recognition of microcalcifications, one of the key structures used to identify cancer. Results are presented for the application of the proposed algorithm to a number of mammogram sections taken from the Lawrence Livermore National Laboratory Database.

The performance of any algorithm such as evolutionary algorithm is only good as the data it is trained on. More specifically, the class represented in the training data must consist of the true examples or else reliable classifications. Considering the difficulties in obtaining a previously constructed database, there is a new database has been construct to avoiding pitfalls and lead on the novel evolutional algorithm Multi-chromosome Cartesian genetic programming the success on classification of microcalcifications in mammograms.

Contents

List of Figures

List of Tables

Acknowledgments

I would like thank my supervisor Dr, Steve Smith and Dr. Julian Miller for their supports and guidance. Some of the work would not have been possible without the help of you.

I would like to thank Dr.Eddy Munday and Dr. Anna Murphy for providing the mammograms for the research.

I would like to thank my family for their support, especially my parents, for her patient, encourage, and support always without any conditions. I really miss them.

Finally I would like to thank all the friends who have been supported and helped me in my tuff period.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

1 Introduction

Each year, breast cancer accounts for one third of cancers in women in the western world. In 2005, it caused the death of 10,500 women in the UK and therefore is one of the most common cancers in women and the second most common death cancer. [1]

In 1988, breast cancer screening was introduced and deaths caused by breast cancer have declined. The survival rate for people diagnosed as having breast cancer has increased, since detection of breast cancer in the early stages. Screening is therefore essential in reducing deaths from breast cancer and it is important to make this accurate. It has been shown in research that [2]screening could reduce the mortality rate of the regularly screened women by 35% in the UK.

There are several modalities of screening patients, such as ultrasound, magnetic resonance imaging and mammograms, and the main method of screening patients is the mammogram, which is a high resolution X-ray of the breast. Refer to [2] [3] [4]the basis of several clinical trials, the World Health Organization announced in [5] in 2002 that screening mammography for women aged between 50 and 69 years reduced the rate of death from breast cancer by 25%. Mammography is X- ray imaging which allows the radiologist to see the details of the breast. It uses does ironing radiation to create images with lower energy X-ray compared to regular radiography screening used in bones. This is high resolution, high contrast film and the goal of mammography is to achieve early detection of breast cancer. For the average woman, according to [5], the US Food and Drug Administration (FDA) pointed out that mammography could detect 85 to 90 percent of breast cancer for women aged 50 and over, the FDA also showed the evidence that mammography could discover a lump up to 2 years before it could be felt.

[7] shows about 15% of cancers have not been detected on a mammogram film. There are several reasons for this, such as the differentiations of the tissue between the appearance of cancerous and normal parenchyma; the morphology findings which are not related with the cancer and possible miss diagnosis from radiologist caused by visual fatigue. Generally, each mammogram image is checked by one or two radiologists for improving the sensitivity. Nevertheless, double reading makes the process inefficient from a medical point of view because the cost to train a radiologist is high just to ensure a small group of specialists is available at each given medical institution, thus reducing productivity. Therefore several researchers have tried to use the Computer Aided Detection/Diagnosis System (CAD) to detect or classify cancer in digital mammograms as a second opinion for radiologists. CAD systems help scan digital mammograms to support preventive medical check-ups. Consequently, CAD systems focus on specific aspects of the diagnostic process, such as identification of microcalcifications clusters or the detection of irregular tissue structures, where either of these suggests the specific abnormalities. It is fundamentally based on pattern recognition, normally many images are required to optimize the algorithm, and digital mammograms are copied to a CAD server. Therefore, a mature mammogram data set is really important for CAD to analyse in several steps.

The process of identifying and evaluating signs of cancer from mammograms is a very difficult and time consuming task, usually requiring skilled and experienced radiologists. Two fully trained radiologists are usually required for each diagnosis. However by its nature, this assessment is highly subjective and susceptible to errors and can easily lead to cancers being missed and patients being misdiagnosed. CAD systems have the potential to help in two main ways: (i) the detection of suspicious areas in the

mammogram that require further investigation and (ii) the classification of such areas as cancerous (malignant) or non-cancerous (benign) [5].

The radiologist is looking for several things when reviewing a mammogram such as clusters of white specks called microcalcifications, distortions, special patterns or tissue density, any mass and its shape, also the differences between the images of both breasts. The two most powerful indicators of cancer can be seen on mammograms are spiculated masses and microcalcifications. It is difficult to distinguish between malignant and benign and malignant microcalcifications. Although most processes are benign, it is also found that malignant processes could be identified easier by looking at microcalcifications that occur as a result of the malignant process.

Over recent years, there has been much research into the application of CAD to breast cancer using numerous different approaches. However the CAD systems have a few limitations that are most CAD systems are based on classical image processing operations that are limited to the general understanding of how malignancies are found, the limitation of practical implementation and limitation of the classical image processing operations. One method to overcome these problems is to use an evolutionary algorithm. Evolutionary Algorithms have the great advantage that they are based on learning how to do a task rather than just following a sequence of orders. The ultimate aim of our work is to assess the potential benefit of a new representation of evolutionary algorithm called Multiple chromosomes Cartesian genetic programming in the classification of mammograms as part of a CAD system and determine whether further development of such algorithms will lead to a more confident diagnosis.

To support this investigation, there are 2 sets of data (mammograms) taken from the Lawrence Livermore National Laboratory database (LLNL) and York District Hospital (YDH). The set of data in LLNL is used to train and test the evolutionary algorithms, and then another set of data from YDH will be used to verify the multi-chromosome CGP network. This gives strong evidence in developing an algorithm that could reliably distinguish between those microcalcifications indicating cancer or those indicating benign.

1.1 Thesis Organisation

This thesis is organised into three main segments. Chapters 1 to 4 introduce the concepts of background, related work in computer detected system, and genetic programming. Chapters 6 to 7 describe the novel contributions of this research. Specifically:

Chapter 1: Overall introduction

Chapter 2: Reviews of breast cancer screening Modalities especially focusing on mammography

Chapter 3: Reviews of related work in CAD system in past years

Chapter 4: Introduces Cartesians Genetic Programming; discussing the motivation behind the approach and its implementation.

Chapter 5: Reviews image data sets especially mammography data in research

Chapter 6: Constructs a dataset and designed Multi chromosome Cartesian Genetic Programming network for automatic analysis of mammograms

Chapter 7: Presents experimental results on Multi-chromosome Cartesians Genetic Programming in two stages: detection of microcalcifications and classification of microcalcifications

Chapter 8: summarises, draw conclusions, and offers some speculative suggestions for future research.

1.2 Hypothesis

The work presented in this thesis addresses the following hypothesis

"Multi-chromosome Cartesian Genetic programming is an effective evolutionary algorithm to facilitate automatic analysis of mammograms."

In this automatic analysis system, Multi-chromosome CGP network classify microcalcifications by using the raw pixel values, this self-learning algorithm has great advantages as they are based on learning how to do a task rather than just following a sequence of orders. This learning algorithm relies on its training dataset, therefore, a highly objective and dependable dataset has been constructed with standard protocol.

1.3 Contribution

The current study was based on these issues:

- To show that Multi-chromosome Cartesian Genetic Programming as a representation of Evolutionary Algorithm has potential benefit on automated analysis of mammograms.
- To propose a specified mammographic image database, containing clearly identified and verified cases of benign or malignant tumours. All cases were

evaluated and annotated in relation to the previously defined mammographic dataset properties.

- Therefore, a general reviews of properties of medical dataset for diagnosis of breast cancer will be organised, with disadvantages of the existing datasets , at the end, a suggestion of a new dataset for diagnosis and classification of microcalcifications will be provide.

2   Breast Cancer screening mammography

Breast cancer is one of the leading causes of death of women in western countries, it is also one of the leading common malignant tumours found in women around the world. Based on statistics from the world Health organisation(WHO), cancer is becoming one of the leading causes of death worldwide, accounting for 7.6 million deaths which is about 13% of all deaths in 2008 [8]. And there is an estimated about 12 million people dying from cancer by 2030 [9] . In all known cancers, breast cancer is currently the top cancer in women world wise and one of the leading deaths among women in high income countries. According to published statistics around the world, breast cancer has become a major health problem in women worldwide. In 2008, it was estimated that about 1.38 million women worldwide were diagnosed with breast cancer. In the UK, there were 48,034 new cases of breast cancer diagnosed and 47,693 cases appears in women [3]. In 2008, there were 12.116 deaths from breast cancer and 12,047 deaths among women [3].

The number of deaths was in a direct ratio with age-specific mortality rates for female breast cancer. The majority of breast cancers are diagnosed in women over 50 years of age and in younger women aged from 35 to 54. Breast cancer is the most common cause of all deaths from cancer [3]. However, mortality rate fell by 44% in women aged 40 to 49 between 1989 and 2008 in Britain [3]; by 44% in women aged 50-64[3]; by 37% in women aged 65-69[3]; by 39% in women aged 15-39[3]; and by 19% in women over 70[3]. There are several different causes for the reduction of breast cancer mortality rate including NHS screening policy, personal education of specialisation of care and the widespread adoption of tamoxifen treatment since 1992[4]. So far, the exact cause of

breast cancer is unknown so there is no effective prevention for breast cancer. However, efficient diagnosis of breast cancer at an early stage could lead to better chances of survival rate for women. Therefore, an early detection of breast cancer could reduce breast cancer mortality and morbidity rate and increases the chance for successful treatment. There are two major components for early detection of breast cancer in Britain: screening and online personal education and risk assessment (OPERA) [5]. Those at risk increases with ages, all women aged between 50 and 70 are invited for breast cancer screening every 3 years and also women aged over 70 and over are entitled to be screened [5].

2.1   Introduction of breast cancer

2.1.1   The Breast

Breast cancer is a type of cancer originating from breast tissue, it occurs in humans and other mammals, while in the majority of cases, it overwhelmingly occurs in women, although breast cancer can also occur in men. The breast lies on the upper ventral region of the torso of the primate, in both the left and right side. In women, the budding breasts is one of the secondary significant characteristics.

In human female breast anatomy, it belongs to the reproductive system. It is a porcine gland which can produce milk to feed babies or little children. The minor's breast shape is like a half ball, firm and persistently rich. During the latter half of gestation and lactation period, breasts enlarges due to hyperplasia of the mammary glands; after breast feeding, mammary gland starts to degenerate, while the breasts gets smaller. Older women have pendulous breasts due to loss of elastic fibre.

In women, the breast overlays the pectoralis major muscles and cover much of the chest area in front of the human rib cage from the level of the second rib to the sixth rib. Breast is also one of the significant organs of a woman. In women's breast, the breast has apocine glands which produce milk to feed babies. The centre of a breast is a nipple which is surrounded by an areola, the colour of that varies from pink to dark brown and it has many sebaceous glands. Women's breast is made up of connective tissues, fats, and many glands known as lobules. Lobules produces the fatty breast milk, which are then delivered through from the duct to nipple. From a functional aspect of breast, it is a gland that produces milk to feed offspring. The basic units of the breast are the terminal duct lobular units which could produce fatty breast milk as a mammary gland. The mammary glands are situated in the breast in humans and in ruminants' udders

Figure 2.1 shows the anatomic structure of a human female breast cross section scheme with the mammary gland. In women, the breast overlays the pectorals major muscles and cover much of the chest area in the front of the human rib cage from the level of the second rib to the sixth rib. One functional aspect of the breast is it is a gland that produces milk to feed offspring. The nipple of the breast is centred and surrounded by an areola. The basic units of the breast are the terminal duct lobular units which could produce fatty breast milk as a mammary gland. The basic components of a mammary gland are the alveoli, these alveoli join up together to form as lobules. Each lobule has a lactiferous duct which drains into the nipple. Nearly all breast cancer originates from the lobules or ducts of the mammary glands. One cause of a tumour in a mammary gland can be caused by an abnormal expression level of circulating hormones, or from a mechanical mammary stoma change. Under either of these two circumstances, mammary epithelial cells would grow out of control and eventually become as cancer.

Figure 2-1 Human female breast structure from [10]
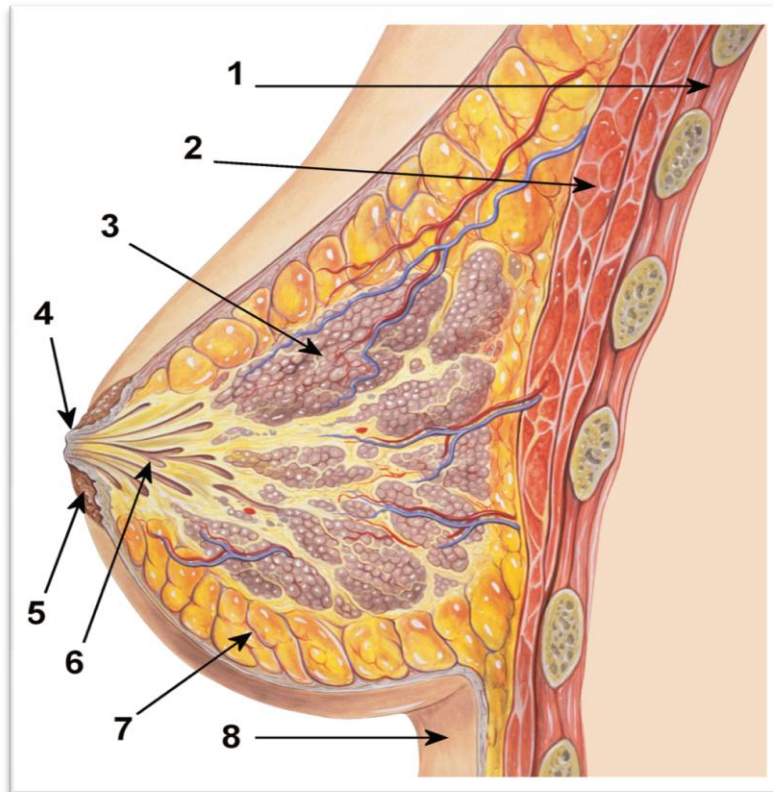
| 1. Chest wall |
| 2. Pectoralis muscle |
| 3. Lobules |
| 4. Nipple |
| 5. Areola |
| 6. Milk duct |
| 7. Fatty tissue |
| 8. Skin |

2.1.2   Symptoms of breast cancer

There are several different types of cancer in the world and they all have in common a growth of abnormal cells, which grow out of control. When the malignant tumour shows up in the breast, it forms breast cancer. Generally, patients are encouraged to do self-checking for lumps in the breast. Patients could also check for changes to the size and shape, or feel of their breasts to find whether there are any changes inside. The first noticeable symptom of breast cancer is a lump or mass which is painless, hard, uneven edge from the rest of the breast tissues. More than 80% of breast cancer cases are discovered when the woman feels a lump; however UK cancer research centre showed that although many women have breast lumps but around 90% are benign. Generally, benign breast lumps are:

- cysts : containing sacs of fluid in breast tissue;

- Lumpiness: before a period, areas of breast cell changes causing the lumpiness to be more obvious, this could happen particularly in women over 35.

- Fibro adenoma: it is a collection of fibrous glandular tissue and more common in young women.

American Cancer Society [11]listed the following unusual changes with could be symptoms of breast cancer:

- A lump or thickening in the underarm area.

- A change in the size or shape of a breast.

- Breast or nipple Pain.

- Skin irritation or dimpling.

- A change in the shape of nipple such as if it turns inward sinking into the breast or has an irregular shape

- A redness, scaliness or thickening of the nipple or the surround area.

- A nipple discharge other than breast milk, especially a blood.

- Swelling of all or part of the breast.

Like breast lumps, these symptoms do not necessarily mean cancer. Different types of breast cancer have different symptoms and patients are encouraged to see their GP once they find any of these symptoms.

2.1.3   Stages of cancer development

Initially, breast cancer may not present any signs or symptoms or problems as the irregularities are too small to be noticed. A lump or any changes may be too small to cause any noticeable changes during the self-examination. Thus, mammography is successful if it finds out any irregularities earlier than a self examination. It is crucial for mammography needs to detect the tumour before it becomes palpable.

The source of breast cancer is as a cell which multiplies like a regular cell does, however the speed at which the cancer cell multiplies might be different because it depends on how the cell has been damaged. The process of growing a tumour could be divided into three stages. Kopans [12] divides the whole process of growing a tumour into three stages. The first stage is called the occult stage. At this stage, cancer is not detectable, indicating that "until there are well over a million cells" in a tumour, then the tumour enters the second stage. During the second stage, the tumour starts to be recognised on mammograms; however it is still hard to be detected by self -examination. Therefore, at this time, the tumour is still in its preclinical phase and is not palpable. This period has been called the sojourn time. At the final stage, the tumour becomes big enough to be palpable and detectable in self-examination.

Therefore, we need make sure we have a screening programme which works well to detect tumours before they becomes palpable to enable the earlier treatment for patients. Having said the above, the time difference between finding the tumour by screening programme and finding the tumour by self examination is very little. And tumours do not have a fixed growth rate in their sojourn time, every tumour has its own growth rate and even change the growth rate during its development. This problem has become a

good discussion point for researchers who are concerned about the screening intervals. [12] describe a model to for better understanding of screening intervals. He mentioned that although the speed of at which the tumour grow varies and is never the same for different tumours, but it is generally possible to assume that tumours constantly start growing and assume that the speed of growth is same for all tumours. This means that mammography could only be used as an appropriate method of diagnosis when it is performed at regular interval. The intervals not only should be short enough that no tumour could grow from its occult stage to palpable stage without being detected but also should be long enough to avoid any unnecessary radiation [13] pointed out especially that there is another benefit that mammographic screening will especially save unnecessary cost.

In fact the breast cancer does not kill the patient, instead it is the metastatic spread to other vital organs which results in possible deaths. Metastasis means "next placement", it means the spread of cancer from the affected part to unaffected part, most used for the spread of cancerous cell to other organs in the body. The spread of cancer typically happens through lymphatic or blood vessels [3]. There is an area of breast tissue which leads into the armpit. The armpit contains a lot of lymph glands known as lymph nodes. These lymph glands around the breast form as a network and are connected by tiny tubes, all of these together are known as part of the body's lymphatic system. One of the chains of the lymph gland goes into the centre of chest, known as the internal mammary chain. Lymphatic system helps to get rid of waste products from the body. Lymph glands are important for breast cancer, based on the characteristic that lymph glands drain into the lymphatic system and goes into circulation. If any cancer cells have broken off from the tumour they could be carried out by the lymph fluid to the nearest

lymph glands. The general problem of metastatic spread is it makes the treatment even harder because cancerous cells could go to anywhere over the body. The single metastases may only be discovered when they are already affecting the organ. In order to increase the likelihood of successful treatment, we need mammography to find the likelihood of tumours as soon as possible before metastatic spread occurs. [14] stated that when the tumour has been detected, the diameter is about 1cm. However [4] explained that once a tumour is in its occult stage "until there are well over a million cells" that would be a diameter about 0.5cm. Also, he mentioned that we should be trying to find the invasive cancer before it goes to 1cm and "Invasive cancers are potentially lethal, and prognosis is improved if they can be treated while they are smaller than 1cm in size", that is because the metastasis begins to grow faster and spreads when they are bigger than 1 cm. Detection needs to be noticed early because [15] states when a tumour goes into the palpable stage, it has an average size of about 2cm and approximately 50% of these cases have turned to metastases.

2.1.4   Diagnosis methods for breast cancer

There are different methods about how to detect breast cancers, all these methods have their advantages and disadvantages. There is no one perfect method which is suitable for all the patients with all the conditions. However, all these used methods have given their advantages which can result in relatively accurate diagnosis.

Firstly, for detection of breast cancer, this would mean that to find a cancer by self examination when the tumour is palpable. Self-examination is firstly recommended as the first breast cancer screening in 1930 [16]. The major problem about this method is

that the cancer may have already caused metastasis and started to spread to other part of the body.

The second method is called lymphography. It is shown on radiography of the lymphatic channels after a radiopaque material has been injected. This method is rather old and the effect is not really successful when comparing with other methods. Therefore, lymphography is rarely used in recent time.

The third method is Computer Tomography (CT), which is an x-ray technology and gives information about the internal organs or cross-section in 2-dimensional slice. Currently, it is out of the market for breast cancer detection. CT scan is not used routinely to evaluate the breast currently or for an early detection of breast cancer. This is because the CT scan could not give information about the internal structures of the glands. However, if the patients have large breast cancer, doctors still order CT scans to check whether the cancer has spread to other parts of the body such as the lymph nodes, lungs, liver, brain or spine. There are other advantages of using a CT scan during the treatment. This allows the doctor to see whether the treatment is working or not, or to see if there is the breast cancer and if it has spread or recurred outside of the breast. Some researchers are investigating whether CT scan could do better or as good as traditional mammography as a screening tool. CT scan uses higher energy X-ray which gives it an ability to indicate microcalcifications for breast cancer. [2] also indicated during the breast CT, if the patients lie face down on the table to let the CT scanner rotate around the breast, there would be less compression to the breast than mammography and the total dose of radiation is the same as mammogram. However, research on breast CT scanning for screening is still in its an early stage.

2.1.4.1 Magnetic Resonance Imaging (MRI)

Another method which has been used for detection and distinction between breast cancer which is Magnetic Resonance Imaging (MRI). MRI is a technology which uses magnets and radio waves to produce cross-sectional images in details of the body. It does not use X-rays, so there is no radiation exposure worries. As figure 2-2 shows, patients need to lie on the stomach on a padded platform with cushioned opening for the breasts, so the breasts hang down. This position gives a better picture and reduces the negative effect of breathing. Before the test, patients need to have a contrast solution which is called dye injected into their arm. The dye could give the contrast solution which help any potential cancerous breast tissue show up more clearly. The best advantage for MRI breast scan is that it can show scars, cysts, and even invasive tumours well. And MRI could perform really well on women with really dense breast tissue; it could get more accurate information than mammograms. Also, doctors use MRI scans for women with lobular breast cancer to help them decide whether lumpectomy is an option. However, there are also disadvantages of this method.

 Not only is that MRI a very expensive method and requires highly specialized equipment and highly trained experts it could also could miss out some cancers that would be detected by mammography. The dye could affect the kidneys, so the doctor may perform kidney function tests before injecting dye. There are certain conditions to the dye being absorbed better like a tumour in the duct lead to the dye being stopped from going further and the dye builds up in one spot. This means that the general procedure of an MRI is rather a time consuming. The patients need to stay still during the test while takes around 30 to 45 minutes. Plus an MRI is expensive compared to the other methods which makes MRI less attractive than other mass screening programmes.

An MRI is better for distinguishing between scars and a malignant tumour and could be used for more detailed examination of cysts. After treatment, MRI can be used for checking scar tissues in women who have undergone lumpectomy and any significant changes could be noticed by MRI, suggesting a return of breast cancer. This means that MRI is suitable for additional examination. However, MRI is not recommended as a routine screening tool for all women due to its high cost, and limitations about dislodging certain metal devices. Even at its best, MRI produces many uncertain findings which are called unidentified bright objects (UBOs) by radiologists. Finally, MRI cannot detect calcifications which is one of the main indicators for breast cancer.
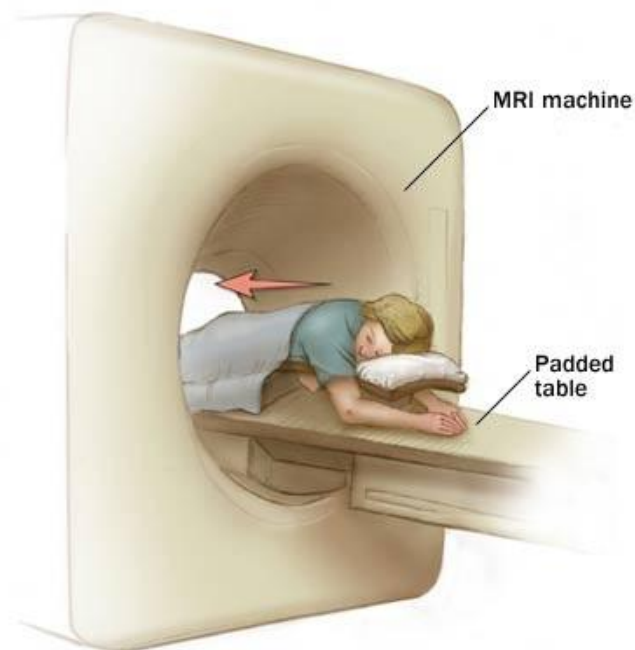
Figure 2-2 MRI breast screening from

2.1.4.2 Ultrasound

The Ultrasound method has been used for detection of breast cancer as well. Its main advantage is that Ultrasound is widely used for other examination and it is considered to

be more economical. This makes ultrasound very attractive for mass screening. [**17**] showed about 90% of small tumours which can be monitored by ultrasound have a size of about 8mm in diameter, therefore ultrasound works just as good as mammography with the tumours' size about 1cm. Ultrasound works very well on detection of cysts. When the patient has been suspected of having a cyst during mammography, the patient is sent to have an ultrasound examination to confirm the diagnosis of a cyst or to exam the wall around the cyst to have a look whether the cyst is in a dangerous condition. Ultrasound is also a main method for examination of younger women. Younger women usually have very dense breast tissue which joins together as a big white mass on mammogram. Research by [**18**] showed that mammograms caused high "false positives" that require follow-up to rule out cancer for women under the age of 40 and did not detect any tumours among women under the age 25. There are two strong evidences that breast ultrasound is very good at distinguishing between lumps and cancer in women aged younger than 40. It showed that ultrasound was 100% effective at classifying a lump as benign without a biopsy for women aged younger than 39. The studies suggest that ultrasound monitoring of the breast is really good at distinguishing between benign lumps and cancer in women younger than 40. There is another useful point about breast ultrasound for the radiologist to help perform a biopsy. In order to take a sample from a suspicious area of the breast, the radiologist needs ultrasound to guide the positioning of the needle, this is because ultrasound is a very safe, easy and quick method. The disadvantage about ultrasound is it cannot show microcalcifications which could be seen as a strong indicator for breast cancer. Additionally, the visualization of ultrasound is relatively poorer than in mammograms, therefore it requires the highly skilled and really careful experts to work on in order to avoid any missing indicators.

2.1.4.3 Mammography

The most important method of diagnosis for breast cancer is mammography. It is an imaging technique which uses low-doses of amplitude-X-rays (usually around 0.7 mSv) to examine the human breast. Mammogram is the examination of mammography and it is used as a diagnostic and a screening tool in the early detection of breast cancer, in order to reduce the death rate. It goes through the detection of characteristic masses or microcalcifications. The x-rays are a form of radiation; it passes through most objects including the body. Different parts of the body absorb the x-rays in varying degrees. Bone absorbs much of the radiation while soft tissue allows more of the x-rays to pass through. As a result, bones appear white on the x-ray, soft tissue shows up in shades of grey and air appears black. It is important for mammography to distinguish between the different structures within the tissue. Research in [**19**] showed that typically 28-30kV on does for a mammography X-ray system could increase the contrast by 5-10%.
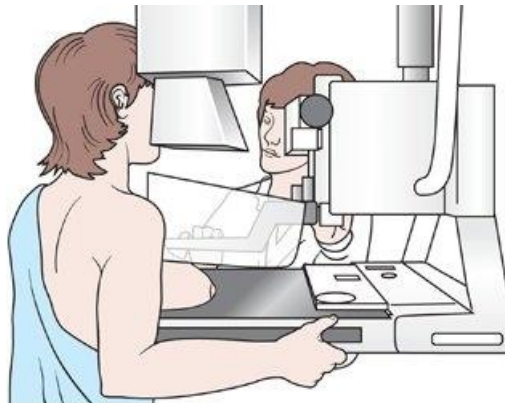
Figure 2-3 The Process of mammography from [2]

There are reasons to choose mammography as a main screening for early detection of breast cancer. The main advantage is it is a very economical method that is easily and

quickly carried out without any injection compared to other methods. Additionally, mammography gives a good image of the breast and it is the only examination method showing both microcalcifications and mass which are two main important indicators for malignant cancer. However in some cases, the radiologists cannot be sure whether the condition is benign or malignant, therefore an extra call back mammography or other methods of examinations like ultrasound or MRI will be used. In the past several years, some women may have areas of concern which cannot be resolved with the only the information available from the mammogram. They would be called back for a diagnostic mammogram. At this stage, radiologist will monitor each of the additional films as they are taken from the technologist. Generally, the unusual appearance is found to be benign, if the cause cannot be determined certainly, a biopsy will be recommended, and it will be used to get the real tissue from the site of pathology to examine microscopically. Depending on the nature of the findings; an ultrasound may be applied at this stage as well.

The mammography examination process will take approximately 30 minutes and it is usually performed on an outpatient basis. During examination, a qualified radiological technologist will position one breast in a mammography unit which looks like a rectangular box that contains a tube for producing x-ray. The breast will be placed on a special platform and compressed with a paddle. Breast compression is then operated by the technologist, it is necessary in order to:

- Even out the thickness of the breast, so all the tissue can be pictured.

- Hold the breast, so the blurring of the image caused by motion is minimized.

- Spread out the tissue so that small abnormalities are less likely to be obscured by overlying the breast tissue.

- Allow the use of a lower does of x-ray for imaging the thinner part of breast tissue

- Reduce X-ray scatter in order to increase the sharpness of the image.

The patient will be asked to change positions between images. The routine views are from top to bottom and side angle.

Views of screening mammograms are done as a regular screening to get an overview picture of the breasts and make sure that everything is well. Different views increase the likelihood of detection. The X-ray mammograms are typically made in medio-lateral oblique (MLO) view and cranio-caudal (CC) views. The names of the directions are from Latin and are used worldwide. The term of ' medio-lateral ' means x-ray direction is from the medial (inner) part of the breast to the lateral (side) part of the breast. This is to give a horizontal orientation view. The term of ' cranio-caudal' means the radiation goes from the cranium (head) to cauda (feet). This means a vertical direction of the radiation coming from the top.
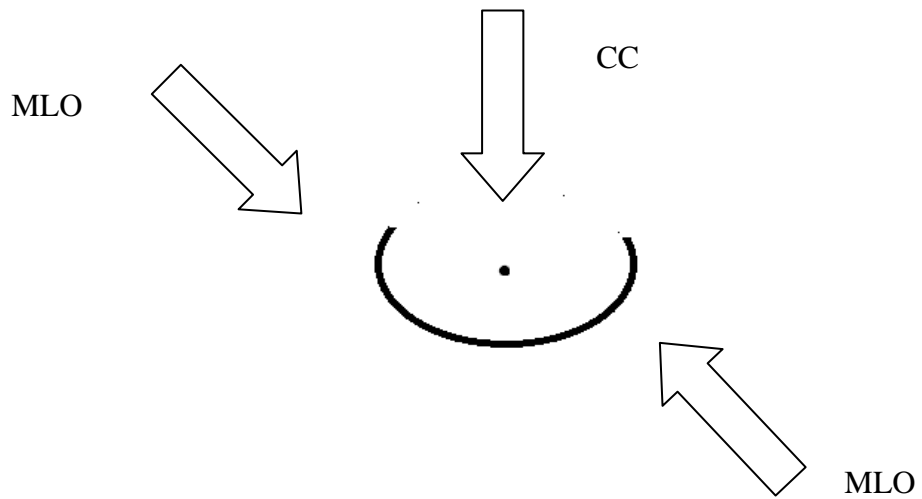
Figure 2-4 Screening orientations for mammograms

Abbreviation	Position	Direction from X-ray
MLO	Medio-Lateral Oblique	From the medial (inner) part to the lateral (side)
CC	Cranio-Caudal	From the cranium (head) to cauda (feet)

Table 2-1 views of screening mammograms

However, different areas have different protocols which are dependent upon the specific facilities. In America, there are four films which are required for breast, two for each breast. In Europe, most countries do one view of each breast, if a suspicious area is noticed, then additional views are taken.

A radiologist, a physician specially trained to supervise and interpret radiology examinations, will analyze the images and send a signed report back. Figure 2-4 and 2-5 shows the example of the mammography images, on the left is a benign case versus a malignant case on the right. When evaluating an image the radiologist would first get old images as reference to see if there is any change in the overall structure of the breast. If radiologists find anything of concern in the images, then they could see whether the conditions have appeared during the last screening. If it has, then the radiologists are able to compare the images to state the change of the size and to give a report about the rate of growth of the tumour. When the radiologists are evaluating the new image, they will first step back to look through asymmetries. Then they could get a bit closer and look for masses. After that, the radiologists will zoom in on the image to examine the mammogram. This means they are looking for microcalcifications which could indicate a malignancy. They even need a magnifying glass if the condition is not that clear.
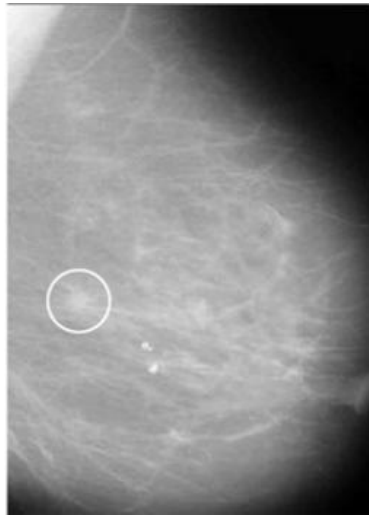
Figure 2-5 mammogram with a spiculated mass (Ozekes S, Osman O, Camurcu AY, 2005) from
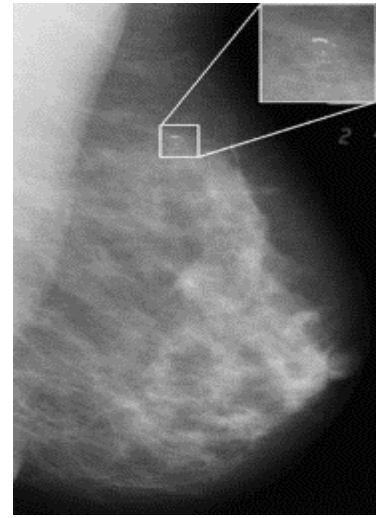
Figure 2-6 mammogram containing microcalcifications (J. Suckling, J. Parker, D.R. Dance, S.M. Astley, I. Hutt, C.R.M.

Mammography is believed to reduce the modality; there are several benefits for the patient:

- Imaging the breast improves the ability of physicians to detect small tumour, therefore if the cancers are smaller, the women have more treatment options and a higher cure rate.

- No radiation remains in the patient's body after an x-ray examination.

- X-ray usually has no side effects in the diagnostic range.

The use of screen mammography increase the ductal carcinoma in situ (DCIS), which is the small abnormal tissue growths confined to the milk ducts in the breast. These early tumours will not harm the patient if they are removed at the early stage and mammography is the only proven method that reliably detects these tumours.

While mammography is the only breast cancer screening method that has been proved to save lives, however it is not perfect.

- The radiation exposure is a potential risk of screening.

The effective radiation dose from a mammogram is 0.7 mSv, this is the same as the average person receiving from radiation in the background for three months, and it appears to be greater in younger women. There are a lot of studies about radiation from mammogram which shows that the women aged 40 years or older have a small risk of radiation induced breast cancer. Therefore, special care is taken during x-ray examination to use the lowest radiation dose possible while producing the best images for evaluation. Organizations such as the national cancer research institute take such risk into account and update the technical standards when formatting screening guidelines. [20]

- False Positives

The screening exam is intended to have a high sensitivity to make sure it does not miss any cancers, the cost of this high sensitivity is that a large number of results which would be regarded as suspicious in patients without cancer. The patient is called back for further testing and are sometimes referred to as "false positives", implying an error. Also, paper [21]shows that women who receive false positive results may be affected in their well-being and behaviour. Some of them may intend to return for routine screening more frequently, and some of them become anxious, worried about the possibilities of having breast cancer.

- False negatives

Mammograms also have a rate of missed tumours – "false negative". This is   partly due to dense tissues obscuring the cancer and the fact that the appearance of cancer on mammograms have a large overlap with the appearance of normal tissues. [22]for this reason, in the UK, the screening program does not start calling women for screening mammograms until the age of 50. Dr Samuel S. Epstein reported in 'the politics of cancer' that in women aged 40 to 49; a quarter of instances of cancer are missed at each mammography [23].

There is always a slight chance of cancer from inordinate exposure to radiation, however the benefits of an accurate diagnosis far outweighs these risks.

2.1.5   Mammogram Indicators

When examining a mammogram, there is a recommendation to examine the mammogram in three steps which include to look at the mammogram at a certain distance. Radiologists are typically looking for asymmetries or architectural distortion, mass and clusters of calcifications which are the main indicators for mammograms.

First step the radiologists will look for is the distance to allow them to recognise any large structural changes. In this step it is important to look out for asymmetries of the two breasts. Asymmetries are normally being seen as differences in structure between two equal views of the breasts. They do not contain any information about cancerous tissue, however they provide an indication about whether there are any structural changes in the breast so the radiologists could take a closer look into the changed areas to see if the structure has changed and is part of a benign process or if the change

indicates a malignant process. [**24**] gave some cases that symmetries are closely measured as they present a likelihood of cancerous changes than normal tissue.

Two of the powerful indicators: masses and microcalcifications are commonly used in evaluating mammograms.

2.1.5.1 Spiculated Masses

Masses are the larger of the two indicators and can be either benign or malignant. Characteristics such as the borders and density of the mass are greater for malignant cases; this could be used as classifications. Comparing to asymmetry, masses don't just have structure changes in the breast tissue, but they also have dense irregular lumps. On mammograms, the denseness masses will exist as bright lumps and are not easy to identify. An example of a mass is shown in Figure 2-7.Mass is the larger of the two indicators and can be either benign or malignant. Characteristics such as the border and density of the mass, which is greater for malignant examples, can be used for classification.
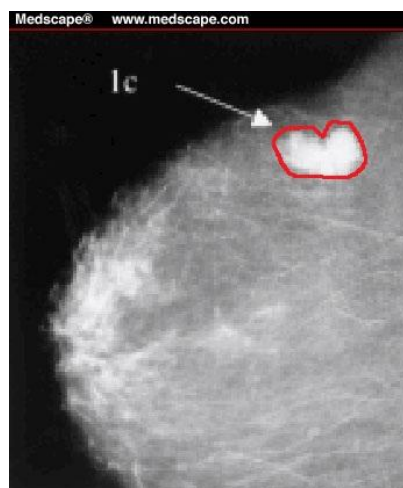
Figure 2-6 mammogram showing a mass from [137]

2.1.5.2 Microcalcifications

Only mammography screening could show the microcalcifications in cluster, they look like a small group of white dots and look like almost noise, and to organize them will depend on their size, shape, distributions and texture and of course the quality of mammogram.

Generally, masses are more difficult to classify than microcalcifications. Microcalcifications are small calcium deposits which occur as secretion from the ducal structures which have been thickened and dried. They can have the most benign cases but also can indicate a malignancy. They are common on mammograms and their appearance increases with age so that they could be found in 8% of mammograms of women in their 20s and 86% in late 70s. (D.B.KOPANS, Third Edition,2007)

One condition that is often not recognized as a cancer but is thought to be a preliminary stage to invasive cancer is DCIS which will be mentioned in Section 5. The only indication of DCIS leaves on a mammogram are the typical clusters of microcalcifications. [12]introduced that DCIS is virtually only found by finding calcifications on a mammogram.

Microcalcifications could be commonly to seen depending on the age group. [12] states that there is a likelihood of having increased microcalcifications during the life, women in their late 20s have 8% chance of having microcalcifications, while women in their late 70s, have 86% chance to have microcalcifications. However, calcification found in the breast is due to a benign process, therefore, microcalcifications can be an indication for malignancy but mostly they are benign.

The literature available to help successfully classify microcalcifications gives different results. Other showed that about half of malignant tumours were found by mammography because of the presence only of suspicious calcifications. [12] gave an opposite result that "17% of invasive cancers were found by mammography only because of the calcifications that had in the intraductal portion of these cancers".

The main criteria of microcalcifications can be classified Microcalcifications which indicate malignancy are usually less than 0.5mm in size and often grouped into clusters of five or more, any microcalcifications larger than 1mm are almost always benign. Therefore, the smaller the microcalcifications are, the harder to see them on mammograms. There are several researches that talk about how large the microcalcifications are to make them visible in mammograms. mentioned about 0.1mm-0.2mm microcalcifications are the smallest range that could be seen in mammograms. [25] mentioned that the limited detect ability is 0.2mm to 0.3mm. Since the maximum size for calcification can indicate a malignancy is 0.5mm, it could be said that microcalcifications that indicate a malignancy can be seen on a mammogram are between 0.2mm and 0.5mm.

The distribution of microcalcification is important as the size. Microcalcifications only appear in or around the centre of the malignancy. Therefore, it will be many tumours that would go along with other structure changes in the breast. When microcalcifications have been found in clusters in mammogram, it will request a further investment. [12] says that if a cluster of microcalcifications is considered to be malignant, the number of calcifications has to be 5 or above per cm^3. The clusters contain microcalcifications less than 5 will be considered as benign. This regularly happens on routine checking for the mammogram.

The morphology would include the attribution of shape, texture or density . All those attributes hold some important information to be able to a correct classification. It can only be sais that round microcalcifications with a smooth surface are typically benign while rough surfaces as well as irregular shapes are typically malignant. [25]described malignant microcalciofications as wild, unordered or with fine linear branching, while benign microcalcifications are rounds solid rods or with lucent centres.

The location can be important as calcifications due to cancer usually build up in around a malignancy. They keep around a certain centre and that centre can possibly tell us the point of origin. In addition, microcalcifications that are caused by malignant processes are almost always located in ducts.

2.2   Screening Mammography

2.2.1   Breast Cancer Screening

The aim of breast cancer screening is to detect any cancerous tumours before they become palpable and clinical evidence, which means in their sojourn time. Generally, most statistics show a steady increase of occurrence with most cases in the group of women aged from 50 to 70. EU recommends breast cancer screening for women between 50 and 69. Breast screening was first introduced in UK in 1988 and now offers tests to women over the age of 50 in every 3 years. Older women could also apply the screening, and the National Health Service (NHS) says that if a women who has reached their upper age of routine invitation, they are encouraged to make their own appointment [26]. If the women are over 70, they do get an invitation to attend screening, but they may go if they want to. These ages limits are about to change. The

UK government has decided to extend the screening programme to women aged from 47 to 73 by 2012. For the average women, according to , the US Food and Drug Administration (FDA) pointed out that mammography could find 85 to 90 percent of breast cancer for women aged 50 and over, FDA also showed the evidence that mammography could discover a lump up to 2 years before it can be felt. According to [27]in 2009 that the US Preventive Services Task Force recommended mammography every two years for women from the age of 50 to 74. In [28], the U.S. Food and Drug administration (FDA) has approved that mammography is the only exam to screen breast cancer in women who cannot be tested for any signs or symptoms of the disease by physical examination. [29] Also pointed out that in 2010 the European Cancer Observatory suggested mammography every 2-3 years between 50 and 69. In UK, the early detection of breast cancer is often done by mammogram, it forms the basis of a NHS breast screening programme [20]. The NHS Breast Screening programme provides free mammography every three years for women aged between 50 and 70 and it is expected that in 2016 the age extension range of women for mammography is from 47 to 73.

As already mentioned in Section 2.1, mammography is the perfect method for screening as an early detection of breast cancer. It is reliable, cheap, accurate, relatively comprehensive, not very time consuming and the images are reproducible. All these could make mammography appear attractive and to be taken into account for governments implementing a screening programme.

2.2.2 Mammography Screening

In many countries routine mammography for older women is encouraged as a screening method to diagnose early breast cancer, as risk of breast cancer increases with age. In UK, all women aged from 50 to 70 have been routinely invited to have a test to look for early breast cancer [30]; in US, the USPSTF (United States Preventive Services Task Force) recommends screening mammography, with or without clinical breast examination, every 2 years for women aged 50 to 74 [31].

In 1913, Albert Salomon who is a surgeon reported his investigations; he had used radiography of mastectomy specimens to demonstrate the spread of tumour to the axillaries lymph nodes. [20] In Europe, in the 1950s and 1960s, Charles Gros pointed out the value of mammography in the diagnosis of breast disease and its potential for screening asymptomatic women. From 1963 to 1966, Philip Strax, Louis Venet and Sam Shapiro under the sponsor of the Health Insurance Plan of new York, they organized the first randomized physical examination and mammography to determine whether it could reduce breast cancer death rate. After 5 years, the results showed that compared with the control group, the mortality rate was reduced by almost one third for the women who took screening mammography. The decrease remained a significant level which was through 18 years of follow-up. [32]They proved that combined mammography and physical examinations of asymptomatic women could significantly reduce the breast cancer mortality rate.

2.2.2 Film mammography and digital mammography

Conventional film screening mammography uses low energy x-rays that pass through a compressed breast during a mammographic examination. The exiting x-rays are absorbed by film which is then developed into a mammographic image that can be held and looked at by the radiologist. With digital mammography, low energy x-rays pass through the breast exactly like conventional mammograms but are recorded by means of an electronic digital detector instead of the film. This electronic image can be displayed on a video monitor like a TV or printed onto film. Again, this is similar to digital cameras that produce a digital picture that can be displayed on a computer screen or printed on paper. The radiologist can manipulate the digital mammogram electronically to magnify an area, change contrast, or alter the brightness. Mammography remains the best method of early breast cancer detection. However, traditional film-screen mammography is limited in its ability to detect some cancers, especially those occurring in women with radiographically "dense" breasts. For this reason, extensive research efforts to improve mammography have occurred. Digital mammography offers theoretical advantages compared to film-screen mammography for cancer detection. [33]

The differences are in the way the image are recorded, viewed by a doctor and stored. Several studies [34] showed that digital mammograms have no significant differences .

Equipment costs for digital mammography systems is 3-5 times the cost of film screen mammography. Additional on-going costs of maintenance and image storage compound the price differential. Some insurers, such as the federal government, have recognized these cost differences and reimburse digital mammography at a slightly higher rate than film screen mammography. Different researches [35] [36] have shown no significant

difference in cancer detection between film screen mammography and digital mammography. While some critics have considered these similar results to be a "negative" for digital mammography, one must remember that to achieve similar cancer detection performance with first generation equipment is a major accomplishment. One study showed the call-back rate from screening, needing to return for additional diagnostic mammography after a screening mammogram. It was better with full field digital mammography than screen-film systems.

3   Computer Aided Diagnosis System

Over the last decade there have been a lot of researches in computer aided diagnosis system (CAD), it contains numerous different ways to achieve an operation. The different methods are designed specifically for the task range from the classical image processing to biological inspired evolutionary methods such as neural network. The aim of this chapter is to highlight the differences between widely used techniques and compare them in the area of the classification of microcalcifications in medical imaging in particular.

3.1 Overview of Computer Aided Diagnosis System

At the beginning of this century, early detection mammography increases breast cancer treatment options and reduced the breast cancer mortality rate by approximately 30% [37]. However, mammography is not perfect, on a screening mammogram leading to further recall are identified in approximately 5% - 10% of patients in America , even though breast cancer is confirmed at only 3 to 10 cases in every 1000 women screened [38,39]. To identify and evaluate the signs of cancer requires a skilled and experienced radiologist. This assessment may also lead to the cancer being missed and the patient being misdiagnosed, by its high susceptible to error in nature. This could happen by viewing of a large number of images to detect a small number of cancers. The complex radiologic structure of the breast, the nature of mammographic characteristics of early breast cancer and even radiologists fatigue. For every thousand cases analyzed by a radiologist, only 3 to 4 are cancerous and thus an abnormality may be overlooked. As a result, radiologists fail to detect 10-30% of cancers [40,41]   in these false-negative results, most of them caused by missing lesions which are evident retrospectively. To overcome the known limitation of human observes, a second reading of screening mammograms by another radiologist has been implemented at many sites. Rapid and continuing advantages in computer technology, computer prompting technology have increased the interest to enable the radiologist to act as their own second reader and the final decision is made by the radiologist. One example of this is computer aided detection (CAD) system in screening mammography. Previous studies have shown that CAD detection systems have improved radiologists accuracy of detection in breast cancer. Computer aided diagnosis (CAD) system for aiding in the decision between

follow up and biopsy are still being developed. Mammographic image analysis is a challenging task for the following reasons:

- It needs to be the nearly perfect since the efficacy of CAD systems have very serious implications.
- This is an extremely challenging image analysis task because the large variability in the appearance of abnormal condition.
- The abnormalities are often hidden in dense breast tissue, and this makes detection difficult.

3.2   Classical Imaging Processing

Digital imaging processing has been widely used in a variety of application areas, it is used for tasks such as pattern recognition or classification schemes.

Imaging processing is a special 2D version of signal processing; therefore the methods and structures are established similar to those of signal processing. As this area has a long history, there are a number of methods that need to be described. Generally in most cases there are several steps which are undertaken; these steps are shown in figure 3-1.

The imaging processing will first start with pre-processing, after that the image might undergo segmentation which is typically followed by analysis feature extraction. There is a feature selection step to select the extracted features from the last step. Finally, there is a recognition classification depending on what the aim the image processing is.

```
┌─────────────────────────────┐
│   Sensing / Pre-processing  │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│        Segmentation         │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│      Feature Extraction     │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│       Feature Selection     │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│        Classification       │
└─────────────────────────────┘
```
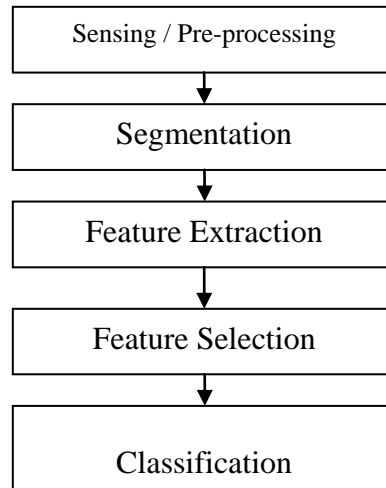
Figure 3-1A typical flow of pattern recognition of image processing

3.2.1   Pre-processing

In this stage, data is often complex and time consuming; they are not quite fit for the system to process. Especially the medical imaging, it is common that they are not stored in a conventional image format as the data needs to be processed by programs. In the meanwhile, it might need to define a region of Interest (ROI), or then it may be necessary to perform adjustment on brightness and contrast or to rotate the image. It may be necessary to rescale the image size to fit with the critical certainty.

In CAD system, how the image is obtained is an important part of imaging processing of mammograms. The digitization techniques vary from different departments. For examples, Digital Database for Screening Mammography has used three different digitizer (described in Section 5). Each of the digitizer uses a different mapping from grey level to intensity. It needs to be noticed that a good performance using one capture scheme may give bad results when applied to an alternative technology. As well as to capture the image, some pre-processing might be performed. In CAD system, the key

aim of this step is to improve the contrast so that the microcalcifications can be seen in dense breast tissue and the masses can be differentiated. To achieve this, different approaches include histogram equalization, filtering, curve fitting and contrast stretching at both local and global views have been used in different pieces of research. Chan et al. [**42**] investigated the application of unsharp masking for digital mammography. Receiver operating characteristics (ROC) studies showed that unsharp masking improved the detestability of calcifications in digital mammograms. However, the method increased noise and caused some artefacts in the images.

3.2.2   Segmentation

The aim of this step is to assign one pixel into one or a certain number of regions. In this step, pixels are partitioned for making the further steps easier. In medical images, it might simplify the image representation to help locate the objects or boundary. Typically, it tried to divide the foreground and background. For mammograms, it tries to separate the target of suspicion from the imaging. To achieve this, there are different methods. The most common one is thresholding of region based approach. It creates a binary image by thresholding and using the results to represent background or foreground, and therefore, two areas have been achieved. The disadvantage is the threshold which needs to be adapted, as colour shading in images varies a lot. There is some statistics approach to achieve segmentation stage as well. The segmentation area could be chosen with certainty by selecting certain areas from the image histogram. This principle works similarly to spatial domain and frequency domain so the general procedures are the same. Additionally, other methods have been developed such as edge detection with subsequent operation to close the gap between the edges. Region based approaches such as seeding with a single pixel and growth in a region to be segmented.

For example, to provide a brief explanation of the technologies. Certain locations are selected as seed points, and without any information these seed points would be all pixels in the image. For example, a microcalcification might be expected to have a grey level over a certain threshold in order to remove some pixels from consideration. Starting from the seed regions growing, if the neighbouring pixels are examined for some similar properties to the seed, or they both followed some pattern from the seed, then they are included in the region. The same is also done for their neighbours. A seed can then be included when it has successful grown in a region around it. There is another method called k-nearest neighbour algorithm. It is a clustering algorithm that partitions an image into k clusters, which is similar to the region growing method but involving an extra step of using statistics to confirm the regions are well formed.

3.2.3   Feature Extraction

Once segmentation is completed there might be a selection of microcalcifications which needs to be described in terms of features. These features are collected in the stage of feature extraction. Classification stages applies on these features and statistical properties, the right features and properties are carefully selected because a good quality of classification is the key to classifying any given data. Features in image processing means real numbers obtained by applying mathematical expression to image data. From the simple examples is the mean grey level amongst a set of pixels. The reason for extracting features in image processing is that they can provide powerful measures of the properties of the image beyond what can obviously be seen by looking at the pixels. A large number of pixels can be reduced to a smaller number of features which provides a concise measurement of the properties of the image. Features can then be extracted from the spatial domain or from the frequency domain though here the former will be

the focus. Features could be divided into two catalogues: textures based features and morphological based features.

- Texture based feature

A number of texture features are described in [43]based on simple statistics. One of these is the nth moment of the grey level z around the mean. Variance, the second moment and the third moment were widely used:

$$\mu(z) = \sum_{i=0}^{L-1}(z_i - m)^n p(z_i)$$
Equation 3-3-1

The mean is defined as normal as:

$$m = \sum_{i=0}^{L-1} z_i p(z_i)$$
Equation 3-3-2

Normalization by dividing by $(L-1)^n$.

3.2.4 Morphological features

Morphological features are often referred to as shape feature and are useful in the classification of microcalcifications. [44] provide several pieces of information about various features of benign and malignant calcifications. Such as, benign calcification has a round ring like shape and has well defined borders. Malignant ones have varying shapes and poorly defined borders.

Morphological or shape features are an alternative way of describing a microcalcification or any object in an image that we are trying to classify. Some simple features based on the boundary are as follows:

- Length of the boundary

- Diameter of the boundary = max[D(pi, pj)] where D is distance and pi and pj are 2 points, given that most shapes are not perfect circles.

- Orientation is another useful feature and taken with magnitude makes up the major axis of the microcalcification

- Compactness takes the length of a shape and divides it by the number of pixels in the shape.

- Statistical moments can be taken by using a segment of a boundary. A straight line is taken from one end of the segment to the other and statistics are taken based on the distances of the pixels on the boundary from this straight line.

3.2.5 Using Pixel Values

Instead of getting and applying the features into the network, it is also possible to apply raw pixels to the network and train on them. One of the strengths of an evolutionary algorithm is that it can discover new solutions that a conventional design would never uncover and it does this by its random nature. If conventional features are extracted to be fed into the network, then this puts a major limit on what the network can achieve in an absolute sense. Feeding the raw pixel values in, essentially gives the network less limitation to extract the information it decides to be the most important. However, in

order to converge on a solution, there is the possibility of a much larger network and many generations may be needed, which takes a longer runtime.

3.3   Feature selection

Feature selection can be regarded in the same way as image enhancements, this stage is not required as a complete pattern recognition system works. Many neural network based papers [45] did not use a feature selection stage. At the end of the feature extraction stage there might be a very large number of features. If a statistical classifier has been used, it is not helpful to have so many features. It may make a longer running time and the likelihood is that the features may not be relevant to the benign and malignant lesions. Thus, the advantage is in selecting the features that are most relevant. Genetic Algorithms have been used in the feature selection method successfully. In [46] it states that with a chromosome, which is the length of the total number of features available, each gene in the chromosome represents a part of which is 1 or 0 where 1 indicates a particular chromosome is   included. There is a population of random chromosome and for each one classification is performed. A new population is generated using parent selection, crossover and mutation. When the parents are selected it is designed so that the ones which resulted in a more accurate classification are more likely to be chosen. This continued for either a certain number of population generations or for a certain level of classification eventually obtained.

3.4 Classification of microcalcifications

Microcalcifications in mammograms appear as relatively bright regions due to the higher X-ray attenuation coefficient (or density) of calcium, as compared with normal breast tissue. Microcalcifications present within dense masses or are superimposed by dense tissues. In the process of acquisition mammograms these present as low grey-level differences or contrast with respect to their local background. On the other hand, Microcalcifications present against a background of fat or low-density tissue and would possess higher differences and contrast. Malignant Microcalcifications grow to be numerous, clustered, small, varying in size and shape, angular, irregularly shaped, and branching in orientation [47] [48].On the other hand, Microcalcifications for benign cases are generally larger, more rounded, smaller in number, more diffusely distributed, and more homogeneous in size and shape. The detection and classification of microcalcifications has been extensively studied, with many authors reporting on several successful approaches to this task. A recent survey by Cheng et al. [49]lists almost 200 references on computer-aided detection and classification of microcalcifications, including methods for the visual enhancement of microcalcifications, segmentation, detection, analysis of malignancy, and strategies for the evaluation of detection algorithms.

El-Naqa et al. [50] used support vector machines to detect microcalcification clusters. The algorithm was tested using 76 mammograms, containing 1120 microcalcifications. A sensitivity of 94% was reported, at one false positive per image. An improvement of the method was published by Wei et al. [51] using a relevance vector machine. A database of 141 mammograms containing microcalcifications was used to test the

algorithm. The method achieved a sensitivity of 90% at one false positive per image. The statistical performance of the method was similar to that of the method of El-Naqa et al. [50] , but the authors reported a 35-fold improvement in computational speed. Yu et al. [52]used a wavelet filter for the detection of microcalcifications, and a Markov random field model to obtain textural features from the neighbourhood of every detected calcification. The Markov-random-field-based textural features, along with three auxiliary textural features (the mean pixel value, the gray-level variance, and a measure of edge density), were used to reject false positives. The method was evaluated using 20 mammograms containing 25 areas of clustered microcalcifications. A sensitivity of 92% was obtained, at 0.75 false positive per image.

[52] developed a technique for the detection of clustered microcalcifications which is comprised of two parts: detection of potential microcalcification pixels, and delineation of individual microcalcifications by the elimination of false positives. The first part involves the extraction of features based on wavelet decomposition and gray-level statistics, followed by a neural-network classifier. The detection of individual objects requires a vector of 31 features related to gray-level statistics and shape factors, followed by a second neural-network classifier. A database of 40 mammograms containing 105 clusters of calcifications was used to assess the performance of the proposed algorithm: a sensitivity of 90% was attained with 0.5 false positive per image.

[53]compared four groups of features according to their discriminate power in separating microcalcifications into the benign and malignant categories. The microcalcifications were segmented using an automated method, and several features were extracted. Each feature belonged to one of the following four categories: multi-wavelet-based features, wavelet-based features, Haralick's texture features [54]and

shape features. Within each group, a feature-selection procedure based on genetic algorithms was employed to identify the most-suitable features for use with a k-nearest-neighbour classification scheme. The classification performance of each group of features was then determined using ROC analysis. The area under the ROC curve obtained ranged from 0.84 to 0.89, and it was observed that the multi-wavelet features gave the best performance, followed by the shape features.

3.5   Neural Network

Artificial Neural Networks (ANNs) are computational models inspired in the natural neurons. ANNs have been used in the areas such as machine learning and pattern recognition. When speaking about ANN it is necessary to distinguish between ANN and Biological Neural Networks (BNN). For a better understanding how ANN works, it is helpful to understand the basic principles of BNN, although BNN is not understood in its full complexity in this days.

A human brain is a piece of grey and jelly style material. It does not working as CPU which is only working by using single or a few processing units. The surface of a brain looks like an outside layer of walnuts, and there are approximately 10^{10} of neurons have been compressed in the grey matter of human cerebral cortex, even a brain of ant has 10^4 neurons.

At the first 9 months of human life, neurons at brain has been produced by 25,000 per minutes , an individual neuron can be described as consisting of three sections : the cell body, the axon and the dendrites, has been showed in Figure 3-2. The cell body is the main part of the neuron. The dendrites are originating in the cell body and are

branching into the surroundings like fibres. The main task is that dendrites have to fulfil

to act as inputs to the cell so that the cell body can receive signals from other cells.

Unlike the other cells in other parts of human body, each neuron has an axon is

connected to the cell body as a fibre, the length of each axon could be extended to

several centimetres, and then branches at its end into many small outputs called axon

terminals in order to transfer the neurons output signal to different directions. The

axons and dendrites of different neurons can then make connections (synapses) and fire

their signals through those synapses.
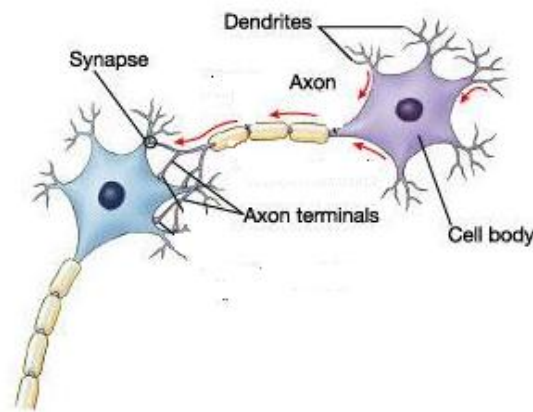
Figure 3-2 Natural neurons

It is not entirely clear that how neurons work, however as far as they have been

understood the sum of input signals might trigger the synapse to fir an output signal.

The output signal is an electrical spike of very short duration, this might trigger some

other neurons inputs which again cause another neuron to fire a spike of its own. It has

been noticed that not all inputs of a neuron have the same significance. During a process

of learning, the neurons simply connect to the other neurons by using their synapses,

however the process is completely random and complex. When comparing the ANN

and BNN it could be said that in both cases neurons are arranged in some sort of

62

network where individual neurons fire signals depending on their inputs signals. ANN's individual neuron does not have a particular shape, it is also a great difference between BNN and ANN. To simulate the physical properties of BNN as well as possible, , neurons of ANN are usually designed to have a number of different inputs, these inputs are weighted to simulate different significance. These inputs are summed up and if the resulting value is a above pre determined threshold, then the neuron fire a signal. The output of this particular neuron will be used by any other neurons or even the entire output of the network. As the two states of a brain neuron : excitory or inhibitory, the practical implementation of ANN means that the signal of neuron would be binary as there either is a signal or there is not. Figure 3-3 shows the structure of an artificial neuron. It basically consists of inputs like synapses in natural neuron; the inputs are multiplied by weights which could be seen as the strength of the respective signals; and then computed by a mathematical function which determines the activation of the neuron. Another function computes the output of the artificial neuron. If the network has n inputs in total, and the input could be x_1, x_2, x_3, ..., x_n,, ad each of the input should have a corresponding weight $w_1,w_2,w_3,...w_n$, the transfer function could be the sum of all weight and inputs: $= \sum_{j=0}^{n} xn \times wn.$
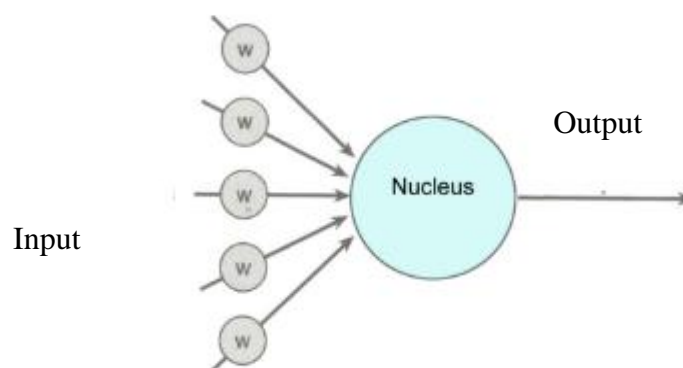
Figure 3-3 An artificial neuron

The higher a weight of an artificial neuron is, the stronger the input which is multiplied by it will be. The desired output could be obtained by adjusting the weights of artificial neurons. If the ANN contains hundreds or thousands of neurons, algorithms could be found to adjust weights of the ANN in order to obtain the desired output from the network. This process of adjusting the weights is called learning or training.

One major difference of ANN and BNN is : neurons of ANN are positioned in a grid of structures while there is no recognizable BNN structure. It means that ANN are made up by different layers, with each layer having a defined number of neurons. Figure 3-4 shows the first layer is called the input layer, which is connected to the input signals that are to be processed. The last layer of the ANN is the output layer, which is providing the processed signals. These are two layers in neurons may be connected to any signals on the outside of network. Any inter medium columns would not be accessible for neither the overall inputs nor outputs for using in ANN. These middle layers are called hidden layers. The hidden layers are required in order to allow the NN to learn a special behaviour, but cannot be influenced from the outside. The number of neurons in one layer may differ from the number of neurons in another layer, although for simplicity the number of neurons are kept constant in all hidden layers. Figure 3-4shows an example of the construction of a neural network, it is called feedforward neural network which was the first and simplest type of ANN. In this network, the signal moves in only one direction, forward, from the input nodes, through the hidden nodes and to the output nodes. There are no cycles or loops in the network.
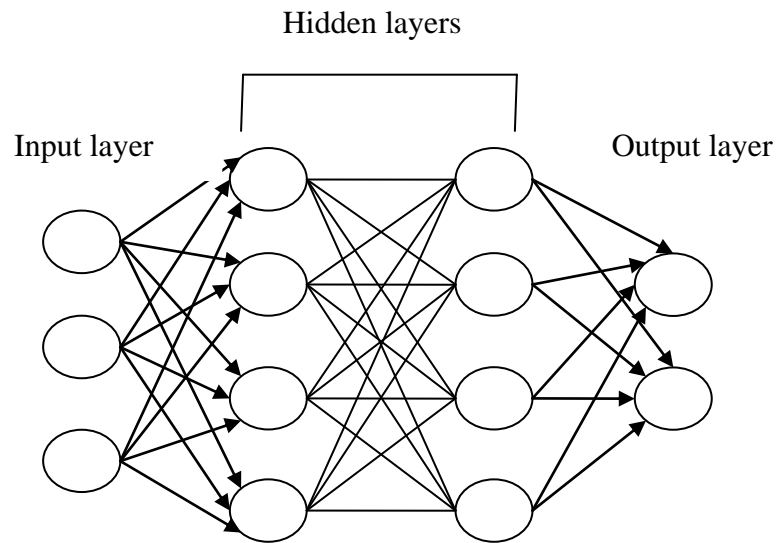
Figure 3-4 An example of an ANN

The most interest in NN is the possibility of learning. The learning process in ANNs is implemented by examples, and is achieved by adjusting the connection weights and the network function.. Generally, there are three types of learning processes: supervised, unsupervised and reinforcement. Supervised learning is based on the direct comparison between the actual output and the expected output. The optimization algorithm could be used to adjust the connection weights in order to minimise the error. Reinforcement learning is a special case of supervised learning; it is o only based on whether the output is correct. Unsupervised learning is only based on the correlation of the input data.

NNs are capable of pattern recognition in computer aided diagnosis, particularly the applications of mammography. There are a number of researches which have been used ANN based on its advantage.

ANN could be used as a part of CAD system and used an ANN combination with classifier. ANN can be used for the detection of microcalcifications. Some other examples have using ANN are as a full CAD system.

Paper	Procedure	Results
[55]	ANN was used for a whole CAD system as well as only parts of the system.	Reported a better approach where ANN was used in conjunction of other methods
[56]	Applied ANN with error-back -propagation algorithm	Sensitivity was increased from 90%b to 95%.
[57]	ANN was used for detection system.	A 87% detection rate .

4 Evolutionary Algorithm

Evolutionary Computation (EC) is based on nature observations and it involves two main subsets: Swarm intelligent and Evolutionary Algorithm.

Swarm Intelligent (SI) is a technique which is based on the studies around the behaviour of self organised systems, the natural example is ant colonies where the groups are working with simple rules and there is no centralised control structures, which means the individuals are only interacting locally with each other within the environment. The interactions between the agents of individuals are often leads to an emergence of global behaviour. The expression of SI was introduced by Gerardo Beni and Jing Wang in 1989 in the context of cellular robotic system [58] .

The other area is Evolutionary Algorithm (EA). In Artificial Intelligent, EA is a generic population based metaheuristic optimization algorithm. The interested area about EA as a technique that develops an ability to solve a problem by learning, especially for the problems which are extremely complex or with unknown solutions. It is suitable for the development of solutions in areas such as image processing, Roger Alsing used genetic programming to create specific painting of Mona Lisa [59]. Based on the implementation details and the nature of particular applied problems, EA has been differed into four major fields Evolutionary Programming (EP), Evolutionary Strategies (ES), Genetic Algorithm (GA), Genetic Programming (GP).

4.1 Biological principles

Evolutionary algorithm is based on the principle of Darwin's theory of survival of the fittest as well as a number of biological principles. Therefore, an EA is inspired by biological evolution such as mutation, selection and crossover. These methods are principally the same for the algorithms of EA however they might implement varies in each technique.

4.1.1 Darwin's Theory

British naturalist Charles Robert Darwin has first made the discovery of evolutionary theories. The book "The Origin of Species" in 1859 has explained Darwin's theory in details. He mentioned that all species on earth evolved over the time from one or few ancestors. There was a widely nowadays accepted idea that human descended from apes that descended from other mammals. With all human life this could be traced back to a simple form in water.

There are five observations which he made in nature and these observations have been summarised by Ernst Mayr [60]. Firstly, Species have great fertility and produce more offspring than they can grow. Secondly, the size of the populations roughly remain the same, although loads of the offspring are born, with a significant amount of them not surviving. Thirdly, food resources are limited but are relatively constant most of the time. All three of these observations show an inference that in such an environment, individuals have struggled for survival. The other two observations are : in sexually reproducing species, generally no two individuals are identical, variation is spreading out, much of these variations is inheritable.

From these, there may be inferred that in a world of stable populations, each individual must struggle to survive, those with the best characteristics will be survived and the desirable traits will be passed over on to their offspring. These good characteristics are inherited followed by generations and become dominant around populations through the time. It means that all only the individuals which fulfil certain criteria have a good chance of survival rate. The concept is called natural selection. By the definition of fitness, individuals with greater fitness are more likely to contribute to offspring for the next generation, while individuals with lesser fitness are more likely to die early or fail to reproduce. It is also referred to this as survival of the fittest. When Darwin mentioned these observations, he did not know what the genetics were, however he made basic observations about genetic fundamentals. Therefore Darwin's theory of evolution are summarised as follows:

Variation	There is Variation in Every Population.
Competition	Organisms Compete for limited resources
Offspring	Organisms produce more Offspring than can survive
Genetics	Organisms pass Genetic traits on to their offspring
Natural Selection	Those organisms with the Most Beneficial Traits are more likely to Survive and Reproduce.

Table 4-1The summarised Darwin's theory of evolution

Some variations are helpful. For example, any variation that increases an antelope's speed may help it elude predators; any variation that increases water retention in a desert plant will favour survival rate of this plant in order to reach maturity. Those animals and plants which survive to maturity and are able to reproduce become the parents of the next generation, passing on the genes for the successful variation.

Darwin called the process by which benefit variations are passed from generation to generation natural selection. He made many important observations on the relationship of individual variation to survival. During his stay on the Galapagos Islands [61], Darwin noted that the populations of tortoises on each island had physical features so that people could often tell which island an animal came from just simply by looking it.

Natural selection is commonly referred to as survival of the fittest. Fittest means that organisms must not only survive to adulthood, but they must also actually reproduce. If they do not reproduce, their genes are not passed on to the next generation. Evolution occurs only when advantageous genetic variations are passed along and become represented with increasing frequency in succeeding generations.

4.1.2   Biological Background

A particular characteristic of an organism is evolution and this occurs through changes in heritable traits. Inherited traits are controlled by genes and these complete set of genes in an organism's genome is called genotype. This complete set of observable traits which show up the behaviour of an organism is called phenotype. Many aspects of an organism's phenotype are not inherited, because these traits come from the interaction of its genotype with the environment.

Heritable traits are passed from one generation to the next through DNA which is a molecule and it encodes genetic information. [62] Introduced that DNA is a long polymer which composed of 4 bases, and the sequence of bases in a particular DNA molecule will specify the genetic information. Before a cell divides, the DNA is copied, so that both of the two cells inherit the DNA sequence. Genes are a portions of a DNA molecule and are used to specify a functional unit , therefore different genes have different sequences of bases. In cells, the long string of DNA forming the condensed structures is called chromosomes. In a chromosome, locus is the specific location of a DNA sequence. If the DNA sequence at a locus varies between individuals, the different forms of this sequence are called alleles. DNA sequences can change through mutation to produce new alleles. If a mutation occurs in a gene, the new allele may affect the trait and finally change the phenotype of the organism.

4.1.3 Biological Methods

The natural inspired selection, crossover and mutation are the three main methods to achieve evolution. They are also giving benefits to achieve minor genetic changes in offspring; it means it does not need to produce the exact copies from parents to enlarge the search spaces. The success of these methods are dependent on how and which technique they are used. Usually selection will be the first to apply, then combination and finally mutation.

4.1.3.1 Selection

Selection is the process which allows the best individuals to be chosen for mating. This follows Darwin's idea of survival and competition. The unit of selection can be the individual or it can be another level within the hierarchy of biological organisation, such as genes, cells or even organisms.

In nature, a certain selection of individuals is taking place, according to Darwin's theory, the strongest or fittest individuals survive and the weak individual disappears. One of the natural selection examples is choosing the king of lions. For example, the male lion will fight to each other to be the king of their group. Once the young male lions reach to a certain age, the male adult lion will attack or kill the young lions to make sure he is the only male lion in the group.

Inspired by this concept, selection works similar in programming; it is also the first method out of three to be applied to an evolution algorithm. Basically, it takes a certain number of individuals from a population, by calculating the best fitness through a fitness function, select the individual which has the best fitness score. The fitness function is based on what results are expected for. In some case, selection is used to find individuals that survive and to get rid of the individuals that are not fit enough. However, the chosen individuals are usually selected for reproduction which also concludes the mutation and crossover.

4.1.3.2 Mutation

In biology, mutation is "the process or an instant changing of a small part of an individual's structure" [63]. In genetics, mutation "changes the structure of a gene, results a variant form and may be transmitted to subsequent generations" [63]. In nature, mutations occur spontaneously and can be induced. It could happen as a result of radiation under certain environmental factors or as an error that occurs in DNA replication.

4.2 Evolutionary Algorithms (EA)

An EA is a computational optimisation algorithm which is strongly inspired by Darwin's biological evolution concepts. Each EA consists of a population of potential solutions to the optimisation problem which alternated at generations on an artificial timeline via mutation and recombination. At each generation, potential solutions are selected from population, according to their fitness , so that better population ones are to be obtained in the new population. This artificial evolution cycle continues until a potential solution has meet the requirements. This cycle is formed in Figure 4-1 as follows:

```
┌─────────────────────────┐
│                         │
│   Generate population   │
│                         │
└─────────────────────────┘
             │
             ▼
┌─────────────────────────┐
│                         │
│   Evaluate population   │
│                         │
└─────────────────────────┘
        ╱           ╲
       ▼             ╲
┌──────────────────┐  ┌──────────────────┐
│                  │  │                  │
│ Selection Process│─▶│ Produce offspring│
│                  │  │                  │
└──────────────────┘  └──────────────────┘
```
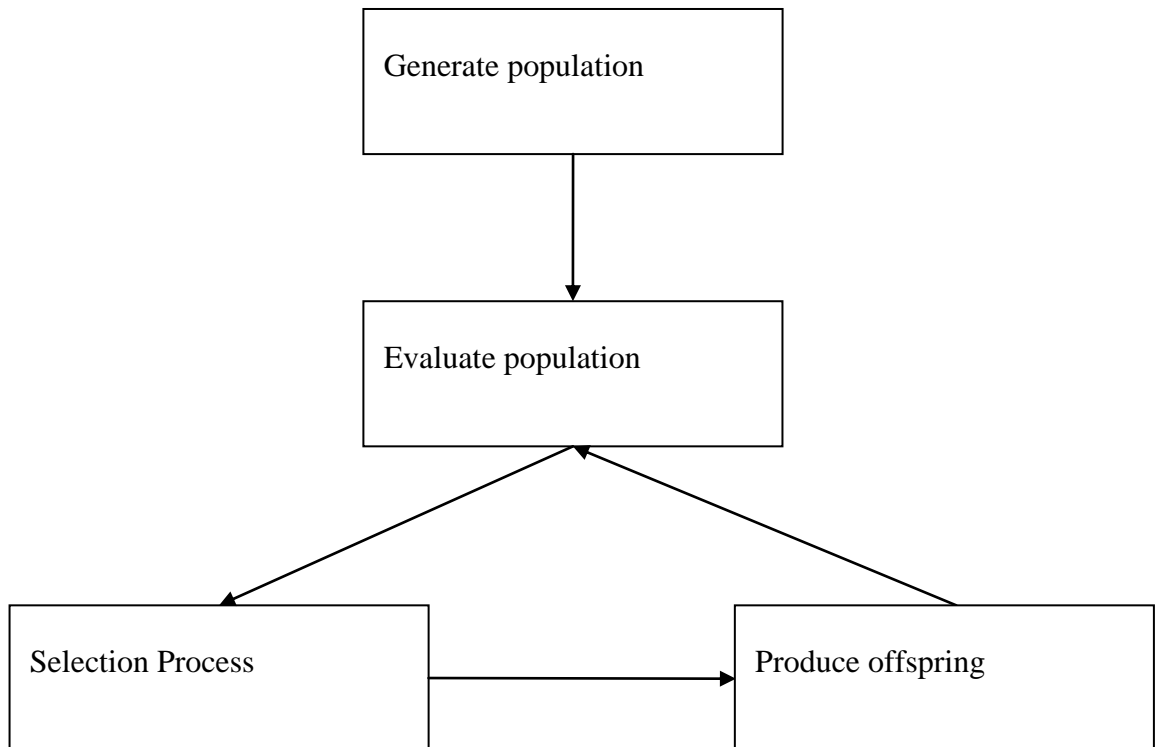
Figure 4-1Evolutionary cycle of an EA

4.2.1 Population

The population in EA is like a number of individuals in biological evolution. Each individual of population consist of a genotype. A genotype represent a potential solution to the problem. Each genotype consists of a number of genes, each gene consists of a parameter or part of the solution. Depending on the types of EA, the population size could be different and the number of genes in a genotype can be binary numbers, integers, real numbers or possible combination.

4.2.2 Evaluation

The evaluation stage in an EA is that the individuals of the population are assessed to see if they present good solutions to the problem. When using EAs, fitness function is used to determine the evaluations and give a reward to an individual, this individual could form a solution to the problem.

4.2.3 Selection

Once the population has been evaluated and each individual is assigned fitness. The selection process takes over. It is inspired by natural evolution, and select member of individuals and these individuals will be promoted to the next generation. The actual selections varies between EAs but all the selection tries to promote sub-standard individuals.

4.2.4 Reproduction

Once the individual has been chosen to be parents, they are used to produce a number of offspring's so the population is always of an equal size. In order to produce offspring the genotype s of the parents undergo crossover/mutation operators.

The crossover operator exchanges genetic material from two parents to produce two offspring. Crossover operator is a dependent on the genotype representation which vary between EAs.

The mutation operator is also very dependent on the genotype representation used by the EA, and therefore can vary in its effect on the genotype. In most EAs, both crossover and mutation operators are probabilistic and happen at a centre rate in genotype. .

4.3   Genetic Algorithm

In the computer science field of artificial intelligence, a genetic algorithm (GA) has been mentioned by J.H.Holland from Michigan University, USA in 1975. It is a search heuristic which mimics the process of natural evolution. GA is a subset of evolutionary algorithm (EA) which generate solutions to optimization problems which using techniques inspired by biological evolution: reproduction, mutation, recombination, selection and crossover.

In a typical genetic algorithm, a population of string encodes individuals or phenotype to an optimization problem, that's because GA is a search based on Darwin's theory of evolution:

- Darwin's theory of evolution. The most important part of Darwin's theory of evolution is the survival of the fittest and surviving individuals reproduce, propagate favourable characteristics. Each species starts getting more adapted to the environment in their development process. The basic characteristics were passed over to the further generations, and the off springs mutate to be different from their parents. When the surviving environment changes, only those who can adapt to the individual characteristics of the environment in order to remains.

- Mendelian inheritance. It is a scientific theory of the genetics information that are encoded in the cells which known as genes in a chromosome. The different position of gene is represented in different characteristics. Therefore, the individuals which were produced by each genes have some kind of 'resilience' to the environment. Mutation and recombination also produced offsprings who

could fit the environment better. Followed by the law of survival of the fittest, the fitter genetic structure could stay over.

Therefore, there are several conceptions from evolution and genetic theory.

- String. This represents as chromosomes or the genotype of the genome in genetics.

- Population. The set of individuals is population, the number of individuals in a set is called population size, string is the elements in population.

- Gene. Gene is the element in the string, it represents the characteristics of individual. For example, if a string = 1011, therefore 1, 0, 1, 1 theses four elements are called Gene, and their values are Alletes.

Flow Chart explanation of GA:

- Create a random initial population: An initial population is created by a random selection of solutions. These solutions have been seen as represented by chromosomes as in living organisms. A chromosome is a packet of genetic information organized in a standard way that defines completely and individual solution. The genetic structure enables the solution to be manipulated. The genetic operands enables the solutions to reproduces and evolve.

- Evaluate fitness: a value for fitness is assigned to each solution depending on how close it actually is to solve the problem. Therefore the problem is needed to define and model it, simulate it or have a data set as sample answers. Each possible solution has to be tested in the problem and the answer is evaluated on

how good it is. The overall mark of each solution relative to all the marks of all solutions produces a fitness ranking.

- Produce next generation: Those chromosomes with a high fitness value are more likely to reproduce offspring; the population of next generation will be produced using the genetic operators. Reproduction by copy or crossover and mutation will be applied to the chromosomes according to the selection rule. This rule states that the fitter and individual is , the higher the probability it has to reproduce.

- Next generation or termination: if the population in the last generation contains a solution that produces an output that is close enough or equal to the desired answer then the problem has been solved. If this is not the case, then next generation will go through the same process as their parents did, and the evolution will continue. A termination criterion that always must be included is time out. Since one drawback of Evolution Programming is that is very difficult to know if the ideal termination criterion is going to satisfied or when .

4.3.1 Genetic programming

Genetic programming (GP) is a biological evolution inspired methodologically that uses computer programs to perform a user defined task, it is a specification of Genetic Algorithm where each individual is computer programmed.

Genetic programming (GP) is the newest group member of EA. GP started to become more widely known after publication of John Koza's book in 1992 Stanford

University,USA. He mentioned that programs are expressed in GP as syntax trees rather than code line, e.g. as LISP and used this LISP expression for solving a differential equation. [64]Nowadays LISP is still widely used by researchers in AI.

In 1981, Forsyth use GP as artificial intelligence, he used the primitive functions such as AND, OR, NOT and the arithmetic operations +,-, times and multiple to predicting British soccer results. In a syntax tree, programs can be written in the form of trees data structure. For example, the simple expression min((a+b),(a-(c*d)) becomes:

$$(min (+ a b)(- a (* c d)))$$

The basic tree program representation used in genetic program is shown in figure 4-2. In the notation , it could see the expression and their syntax trees could be seen. In this example the program represented by the tree is equal to $x^3 + 2x^2 + 3x + 4$. The terminals "x" and "1" are in square background, the nodes "+" and "×" have a square shaped background.
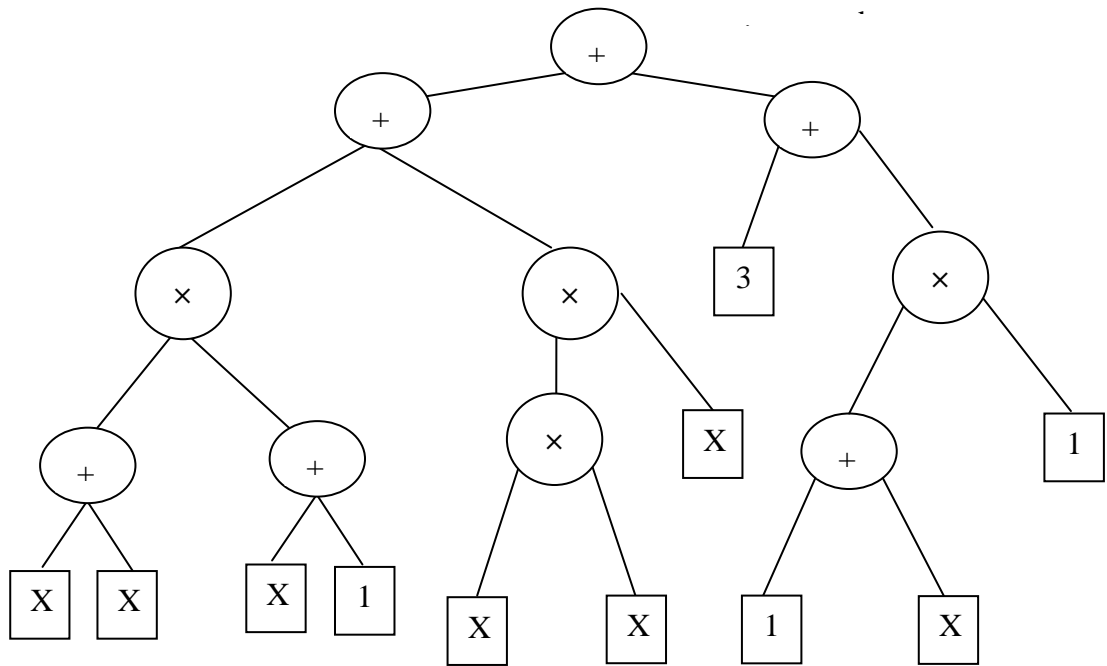
Figure 4-2basic tree as program representation used in genetic programming

The tree includes nodes and links, the node indicates the instructions to execute such as +,-. The links indicate the arguments for each instruction. In more advanced forms of GP, programs can be composed of multiple components therefore the representation used in GP will be a set of trees grouped together under a node called root, all the individual tree will be called branch.

Just like in any other Evolutionary Algorithm, Genetic Programming works according to evolutionary principles. This includes using techniques like selection, mutation and crossover. In term of selection it needs to be differentiated between parental selection and survival selection. While the survival selection is based on a generational replacement, the parent selection is strictly executed on the bases of their according fitness. The excitation of mutation in GP has been introduced as a random change within the tree structure, based on fitness, parents have been selected probabilistically

for mutation. On each node of parent, mutation operation randomly pick up a node and detect sub tree at the picked node, this node will be seen as mutation point and the then new sub tree at this point has been grow in same way as generated trees for initial random population. The offspring produced by mutation is a syntactically valid executable program. It has turned out that selection and mutation work quite well for GP. Crossover has been implemented as an exchange of sub-trees, there are two versions of crossover operations: sexual recombination and two offspring version. sexual recombination operation occurred on two parents which have been selected probabilistically based on fitness, then randomly pick a node from each of the two parental programs. Then these two sub trees have been switched and rooted at the two picked points. The other crossover operation happened on two offsprings. There is a downside of crossover which is it allows any branch of a tree to be exchanged with any other brand of another tree. Therefore the resulting children are not of the same size as their parents. The effect of the growing size of trees have been described as bloat and is the most commonly described side effect of crossover in GP. Many researchers have been trying to find solutions to avoid bloat while introduce crossover.

4.4 Cartesian Genetic Programming

Cartesian Genetic programming was first developed by Julian Miller [**65**][3] for the purpose of evolving digital circuits. The name of Cartesian Genetic Programming first came up from [**66**] in 1999 by representing a program of two dimensional array as 'Cartesian' and was proposed as a general programming form of genetic programming in 2000 [**67**]. Rather than using trees representing computer program it uses n-dimensional array (typically n is 1 or 2) or network of functions, it could be seen as

genetic programming generalized from trees to acyclic graphs where the nodes and edges connections are functions. The functions are not in the form of programming language, they can be the simple logic gates such as AND gate. The advantages of CGP are: implicit reuse of expression due to its graphical representation; reuse of functional redundancy. [**68**]

4.4.1   General form of CGP

CGP representation consists of a network of nodes, where each node is having the same number of inputs and one output. The dimensions of the network have to be specified to fulfil special needs if necessary. For a successful application it is, however , advisable to have more than just one column as the network nodes will have to use the outputs of previous nodes, which is not possible if all nodes are arranged in one column. However, if all nodes are arranged in one single row the nodes have the maximum number of possible inputs. Although some say this arrangement is the idea one for CGP, there are also arguments for using multiple rows. One advantage of having multiple rows is for example an easier implementation in hardware should one decide to built some application.

Each node consists of a basic set of information: inputs and functions. Each network of nodes has a certain number of inputs and usually one output, although multiple outputs are possible. For each component it is possible to bind to any other component that is further towards to the input. It is not possible to bind any component in the same column because this would lead to potential recursive loops, this is forbidden in CGP. An example of such a network is showed below:

inputs: 0 1 2

1 2 III (3) 0 0 I (4) 0 1 III (5) 3 4 III (6) 4 5 II (7) 2 7 I (8) 6 7 I (9) 8 9 II (10) 6 8 II (11)

output: 10

Figure 4-3 An example of CGP genotype

In this example, there are 3 inputs to the overall network, which in this case is one dimensional. The outputs of the nodes are normally not specifically stated but were added in brackets in this example. The outputs are then used as inputs for later components. The first two numbers are the two inputs that are used in each node. The third number in roman number is the function which needs to be defined in a separated function set. It is known that the function set of this example must has at least 3 functions. It is possible that not all nodes or inputs are used to produce the output. In fact. in some cases it can happen that the input is connected directly to the output. I the given example the node providing output 11 is not used for producing the network output and is therefore redundant. One major difference between GP and CGP is that CGP has a limited number of available nodes and has the ability to reuse data rather than processing the same data twice. In GP, not possible to reuse data because this means that smaller branches would be used by some bigger branches ones. This reuse in code can reduce bloat as one operation only implemented once, rather than once for every branch. CGP thus remains a constant size and presents a certain flexibility. Because of this, a simple change in CGP could have a much wider impact on the result than GP.

A mutation operator used could alter both the present function within a grid cell and the connection between components. For example, Mutations can make active genes become inactive and inactive genes becomes one of the several things that can happen. This is beneficial in evolutionary search. [**74**][10]

In CGP either the inputs to a node might be mutated or the function used to calculate the nodes output from these inputs. The mutation impact is therefore dependent on the position and the kind of the mutation. The type of mutations shows as below ;

Type of genes:	Type of mutations:
Function gene	Gene will be mutated to a new random chosen function from valid functions.
Connection gene	Gene will be mutated to a new valid random chosen connection gene.
Output gene	Gene will be mutated to a new valid output connection.

An example is showed below in Figure 4-5 and the equivalent tree structure of CGP in figure 4-6

inputs: 0 1 2

1 2 III (3) 0 0 I (4) 0 1 III (5) 3 4 III (6) 4 5 II (7) 2 7 I (8) 6 7 I (9) 8 9 II (10) 6 8 II (11)

output: 11

Figure 4-4 An example of the genotype of CGP
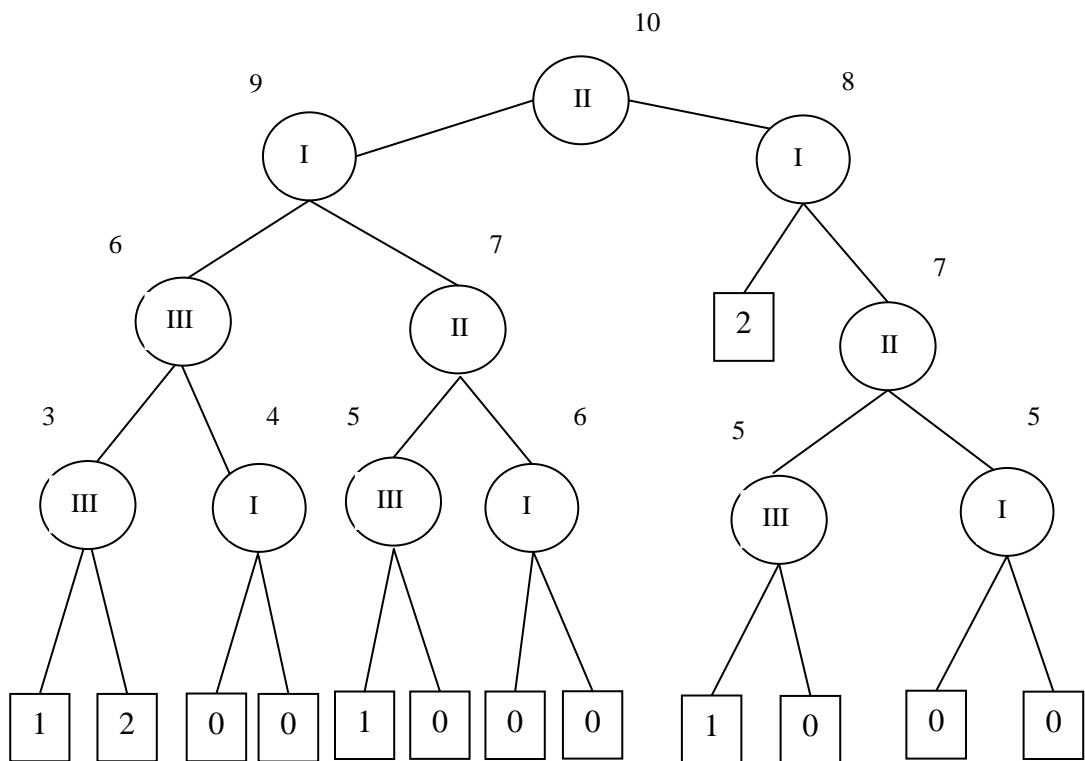
Figure 4-5 Equivalent tree structure of CGP

In this example, the only parameter that has been mutated and has changed is the networks output, which has been mutated from node 10 to 11. Node 11 has not been used previously and has changed from being redundant to being in use while node 9 and 10 have been made redundant. It is also very obvious that one simple mutation had a very severe impact on the function structure. It shows that mutation can have minor as well as major influence on the networks structure and therefore has ability to produce rather diverse offspring. Therefore, mutation is a powerful operator of CGP.

4.4.2 Decoding a CGP genotype

In the decoding process, inactive genes in non-coding genes are not processed. Figure 4.6 shows how the CGP genotype is decoded to produce a phenotype.

(a) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

(b) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

(c) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

(d) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

(e) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

(f) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

(g) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

(h) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

(i) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

(j) 0 0 2 | 0 0 3 | 3 4 5 | 0 1 2 | 0 1 3 | 2 5 7 | 2 6 9 | 0 5 7 | 2 11 8 | 0 11 8 | 4 | 9 | 12 | 13
    4   5   6   7   8   9   10   11   12   13 O_A O_B O_C O_D

Figure 4-6 an example of decoding procedure for a CGP genotype for the two bit multiplier problem from [8]

(a) Output O_A connects to the output of node 4, move to node 4.

(b) Node 4 connects to the program inputs o and 2, therefore output O_A has been decoded, and move to output O_B.

(c) Output O_B connected to the output of node 9, move to node 9.

(d) Node 9 connected to the output of node 5 and 7, move to node 5 and 7.

(e) Node 5 and 7 connected to the program inputs 0, 3, 1 and 2, therefore output O_B has been decoded, and move to output O_C.

(f) Output O_C connected to the output of node 12, move to node 12.

(g) Node 12 connect to the output of node 8 and 11, move to the node 8 and 11.

(h) Output of node 11 connected to the output of node 5 and 7, node 5 , 7 and 8 connected to the program inputs of 0,1,2 and 3, therefore output O_C has been decoded and move to output O_D.

The procedures continue until output O_D has been decoded (step (i) and (j) respectively, when all outputs are decoded, the genotype is fully decoded.

4.4.3 Evolutionary Strategy

In CGP, a form of $\mu+\lambda$ evolutionary strategy (in general $\mu=1$, $\lambda=4$) is often used, this give a population size of $\mu+\lambda$. The μ value indicates the number of individuals promoted to the next generation as parents and the λ value indicates the number of offspring generated from the promoted parents. [9][7] [**75**][11]

This has the form: (For example of 1+4)

- randomly generate the population size of 5 programs

- calculate the fitness of the population

- select the fittest of the population as the parent

- Mutate the parent to generate 4offspring to form the new population

- Select a new parent, by following:

If any offspring have the best fitness which is better than the parent, this offspring becomes the new parent

Else if there are many offspring have the same fitness as the parent, one of them has been randomly chosen as the new parent. Due to CGP chromosome have large number of inactive genes, so with small numbers of mutation different, population members could decode to the same phenotype.

Else the current parent remains as the parent.

A Recombination (Crossover) doesn't seem to add anything [7], however if there are multiple chromosomes with independent fitness assessment then it helps a LOT. [9]

Each functions are chosen to use on the CGP network and will have an important effect on the fitness achieve. In this project, considering that the Pattern recognition is normally evolving mathematical techniques; therefore a good selection of mathematical functions would be necessary. There is no way to be sure which function is the ideal without experiments but the functions available could make a lot of potentially varied and powerful networks to be evolved.

CGP has been shown to perform well within a wide range of problem domain. The applicants of CGP so far: Digital Circuit Design, Mathematical functions, Control systems, Image processing, Medical diagnosis, Bio-informatics, Developmental Neural Architectures, Evolutionary Art, Artificial Life, Optimization problems.[7]

4.4.4 Example of CGP

CGP can represent many different kinds of computational structures. The example is the initial aim of CGP, using CGP genotype to encode a digital circuit. Figure 5.4 shows a CGP genotype and the corresponding phenotype that across in the evolution of a 2-bit parallel multiplier.
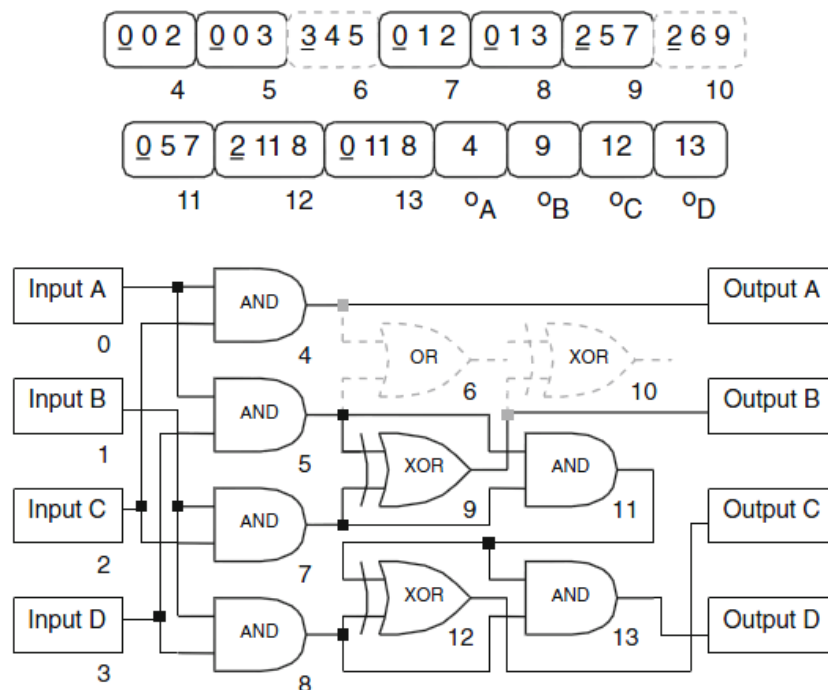
Figure 4-7A CGP genotype-phenotype for a 2 bit multiplier circuits [copy from [76][12]]

In this graph, CGP genotype parameters are:

$n_c = 10, \quad n_r = 10, \quad l = 10.$

The CGP genotype represents a digital circuits for two bits parallel multiplier, it multiplies two by two bits number together, therefore, this circuit contains four inputs and four outputs, there are four logic gates which are represented as primitive functions in function set. The underline function genes in the genotype encodes the function of each node, from the look-up function table, the functions are: AND(0), AND with one input invert (1), XOR(2) and OR(3). The addresses are shown in each program input and node in the genotype and phenotype. Node 6 and 10 are inactive areas of the genotype and have been shown as grey.

4.4.5   Strengths and weaknesses of CGP

CGP has the strength that no assumptions are made, since it is GA based, it works by randomly searching the full solution space allowing solutions to occur that could not occur with conventional techniques. One of the most powerful techniques at the moment in image processing is the use of wavelet. Effectively a transformation similar to Fourier Transform but in this offer, both scale and frequencies information instead of just spatial domain information or just frequency, proven by M.Sifuzzamanln's article called 'Application of Wavelet Transform and its Advantages Compared to Fourier Transform'. However whilst being sophisticated it is still quite conventional with implementation involving blocks of filter banks. For the basis or mother wavelet only one of 3 or 4 families (such as Haar wavelet) tend to be used because a lot is known about them, not just because it is the best for the problem. CGP is far less conventional and solutions could involve what appears to be random arrangements of adders, multipliers, filters, comparators, functions to calculate means or else.

CGP has been shown to perform well within a wide range of problem domain. In Miller's lecture notes about CGP, it shows that the applicants of CGP so far: Digital Circuit Design, Mathematical functions, Control systems, Image processing, Medical diagnosis, Bio-informatics, Developmental Neural Architectures, Evolutionary Art, Artificial Life, Optimization problems.

CGP has proven useful in image processing [71]already, for example, it has been found useful effectively in removing noise from an image, and it is quite likely that it can be used in processing mammogram.

4.5 Multi-Chromosome CGP Network

Multi chromosomes have been used in a number of ways within GP. One of the first was by Hillis [77] . He co-evolved genotypes comprising 15 pairs of chromosomes to produce minimal sorting networks that were capable of outperforming human designs. Cavill [78] discovered that the use of multi-chromosomes and also having multiple copies of chromosomes within the representation is advantageous to evolution on symbolic regression problems. Using a 2-stage crossover operator , similar chromosomes from 2 parents are paired using chromosome shuffling and then a n-point crossover is used to exchange material between the pairs of chromosomes.

4.5.1 Multi-Chromosomes Representation

The difference between CGP and a multi-chromosome CGP network is that the multi-chromosome CGP genotype is divided into a number of equal length chromosomes, each chromosome could see as an individual single chromosome CGP network. The Number of chromosomes present in the genotype of an individual is depended upon by

the number of program outputs required by the problem, as each chromosome is connected to a single program output. This allows that a large difficult problem with multiple outputs to be broken down into many smaller problem with an individual output. The idea of it is making the problem easier to solve. By allowing each small problem to be encoded in a chromosome, the whole problem is encoded in a single genotype. Each chromosome contains an equal number of nodes, and been treated as a genotype of an individual with single program output. As an example of multi-chromosome CGP genotype showed in Figure 4-6:
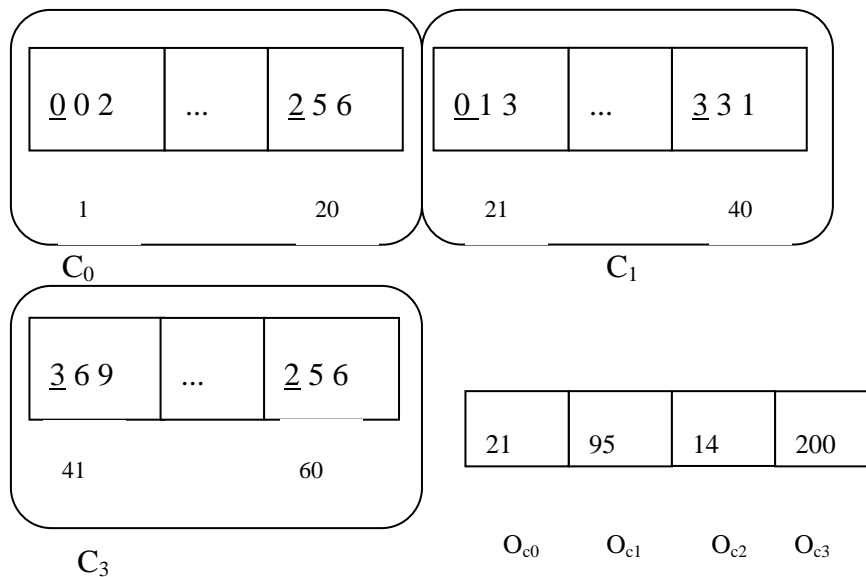
Figure 4-8 an example of a Multi-chromosome CGP network encoding a 2 bit multiplier with four outputs (O_{c0} - O_{c3}) containing four chromosomes (C_0 - C_3), each consisting of 20 nodes.

4.5.2   Multi-Chromosome Evolutionary Strategy

Rather than assigning a single fitness value to a number of program outputs, as in single chromosome CGP, a fitness value is assigned to the output of each chromosome in multi-chromosome CGP, as each chromosome's output is also a program output.

93

In [**77**], it shows that the (1+4) multi-chromosome evolutionary strategy selecting the best chromosome at each position from all of the individuals in the population and generates a new best generation of individuals, containing the fittest chromosome at each position. The new best of generation individuals may not exist in the population, as it is a combination of the best chromosomes from all the individuals. This is so it could be thought of as a super individual. The multi-chromosome (1+4) evolutionary strategy behaves as it selects the best parts from all the individuals. The overall fitness of the new individual also is better or equal to the fitness of any individual in the population from which it was generated.

4.6   Advantage of EA in mammography

Classical image process are capable of dealing with the tasks like detection and classification. The major difference between EA and image processing is their structures.

Classical image processing is based on the sequential use of existing filters and functions, providing solution to problems have already been solved by means of mathematical manipulation. Programmer needs a full understanding of the problem as well as the processes in which the functions and filters are used. So programmer can use the functions or filters in a correct order with effective parameters and all these are needed to solve the given problem.

EA learns to perform the tasks and build their own internal structures with no consequences to the programmer. The programmer may have no understanding of how the algorithm processes from the given data, giving the algorithm some variability. EA

could deal with complete different problem. It has the ability to learn to deal with new situation. It may be able to change the application from one to another.

EA is suitable for complex structure problems since it learn to evolve from a solution to a problem.

The particular area of mammography is extremely complex and is therefore ideal for evolutionary methods. There is a main restriction for those evolutionary methods is presented by the available database. As they have major influence on an algorithms performance great care needs to be taken when making a decision for a database to be worked with .

5   Mammography Datasets

The performance of any algorithm is only as good as the data it is trained on. More specifically the classes represented in the training set must consist of true examples or else a reliable classification will be difficult, if not impossible to achieve. This presents a challenge for medical applications especially where the condition under investigation is difficult to diagnose. Therefore, an understanding of mammograms dataset has been considered in this research. This chapter introduces the properties, limitations of medical dataset, especially for mammograms. Then the dataset which is used for training and testing the algorithm has been introduced.

5.1   Medical Dataset

Medical dataset is an organised collection of data, typically in digital form for machine learning. The data is typically organised to model relevant aspects of reality, they are applied to support processes requiring the information.

It is essential to have data sets to support the research on the detection of breast cancer. There are four main reasons that datasets are not only needed but also required to be used in this particular thesis, shown in the following.

- Training of learning algorithms

    Machine learning algorithm is a branch of artificial intelligent, it focuses on prediction based on known properties learned from training data. For example, as evolutionary algorithm, the performance of any EA is only as good as the data

it trained on, therefore a completely clear, fully detailed image data set could lead a better training performance of evolution algorithm.

- Evaluation of new classifiers:

  Datasets are commonly used to evaluate new classifiers. Also to investigate specific properties including existing conditions and clinical features of the mammographic datasets which are used to evaluate a new classifier - Cartesian Genetic Program (CGP) in CAD systems.

- Characterisation of different medical conditions

  Different diseases have different medical conditions with the specific symptoms and signs. Therefore, different datasets are more like a container to model different specific medical conditions. A lot of scientific research is based on the gathering and analysis of measurement data, especially for medical image research. Datasets including models and parameters are important and sometimes can be seen as the primary intelligent input of the research because they cannot be reproduced and will be necessary for longitudinal research or to test or check for further insights in the future.

Medical imaging widely accepted as a special case due to its reliability. Medical decisions are based on the clinical tests which provide huge amount of data on patients and their medical conditions. For example, a pathologist analyzing biopsies to decide whether they are malignant or not, the radiologists planning a sequence of radiation doses is searching for the harms. The sizes of medical datasets varies and depends on several conditions such as spatial distributions, collection strategies, time consuming issues, etc.

Medical dataset drives from reality and it is not easy to simply build up. The medical dataset usually involve subjective clinical assessment, and often there is no definitive clinical test for every condition under investigation.

This section provides an overview of medical image datasets for early detection of breast cancer. Four publically available datasets of mammograms have been introduced in section 5.2. Finally, a suitable new dataset for developing a novel evolutionary algorithm for classification of microcalcifications in mammogram is constructed in section 5.3.

5.2   Medical Datasets for Detection of Breast Cancer

There are a varieties of datasets for clinical breast cancer diagnosis and research. These includes information of stages, ages, pathologic record, results of physical exams and recent images are included. Research shows that early detection of breast cancer could improve survivals. Breast cancer screening is one kind of cancer screening which has been provided for an early diagnosis of breast cancer. A number of screening tests have been employed including: clinical and self-breast examinations, mammography, ultrasound, and magnetic resonance imaging. Following the procedure of breast screening, there is a wide range of dataset modalities commonly used in clinical practice including, classic x-ray based mammographic image, magnetic resonance imaging which is a three-dimensional imaging modality more sensitive to soft tissue, ultrasound which is using high frequency sound wave to produce images; the parameters which have been used to classify breast cancer and to indicate cancer stages; the biopsies for getting tissue sample with information about fine-needle aspiration, nipple aspirate and ductal

lavage; The traditional modalities as case notes from radiologists, oncologists and pathology report are also included.

5.2.1 Characteristics of Breast Cancer

Once the cancer has been confirmed, the information includes the stage of the breast cancer in pathology report and it is important for doctors to make a plan of treatment. The stages cover the range of whether breast cancer is limited in one area of breast or if it has already spread into healthy tissues inside the breast or other parts of the body. There are several characteristics that can be assigned as local, regional and distant to determine the stages of breast cancer: invasive and non-invasive cancer, the size of the cancer, cancer is in or not in the lymph nodes, whether the cancer has spread to other parts of the body.

- **The size of the cancer:** size indicates how large the tumour is at its widest point. Doctor measures the tumour in millimetres or centimetres. The size is used to help the radiologist to determine the stages of the breast cancer. However it should be noted that the size of the tumour is not everything, because a small size tumour could be aggressive while the large one not.

- **Cells in Lymph nodes:** as mentioned in section 2, they are easily affected by breast cancer cells, therefore, before the surgery to remove an invasive breast cancer, an examination of the lymph nodes will be involved. The more lymph nodes that contain cancer cells, the more serious the cancer might be.

- **Spreading:** Whether the cancer has grown to other parts of the body beyond the breast.

- **Invasive or non-invasive breast cancer**: Breast cancer normally starts to grow from two parts, either in the lobules cells which are the milk producing glands, or the ducts which are the passages which drains milk from the lobules to the nipple. Non-invasive cancer sometimes is called carcinoma in situ, it stays in the milk ducts or lobules in the breast and do not grow into normal tissues or cross the edge of the breast. Invasive cancer does grow into healthy tissues. Most of the breast cancers belong to the invasive cancer type, however it really depends on the patients' treatments and how the patients respond to the treatments. Figure 5-1 and 5-2 shows the differences of the normal cells, non-invasive cells and invasive cells of breast cancer.

Figure 5-2Invasive cell from [127]   **Figure 5-2Non-invasive cell from** [127]

Cancers do not always remain as single invasive or non-invasive. That means part of the cancer has been grow in normal breast tissues, part of the cancer and part of tissue has been spread in other parts of the body. In most cases, breast cancer can be classified as follows:

DCIS (Ductal Carcinoma In situ): DCIS is the most common non-invasive breast cancer which stays inside the milk duct. Ductal carcinoma in situ (DCIS) is a sign that cells in some ducts of the breast have started to transform as cancer cells. It is a preliminary stage to invasive cancer. 'In situ' means that cancer is 'remaining confined to the duct'. DCIS develops from a normal duct that would typically have one layer of cells as an outer boundary. If the cell developed more than should and form an additional regular layer around the boundary the condition would be called Hyperplasia. If additional cells of a hyperplasia takes an irregular form, the condition is called an atypical hyperplasia. If the condition keeps growing uncontrolled, it is called DCIS. It is important because it is close to the stage on early detection of breast cancer. Even it is not developed into cancer yet, most cancer grows in ducts, if untreated, DCIS leads I about 30%-50% of all cases turns into invasive cancer [**12**].

LCIS (Lobular Carcinoma In Situ): LCIS is an overgrowth of cells that stay inside the lobule. It is not a true cancer; more like a warning sign of an increased risk for developing an invasive cancer in the future.

IDC (Invasive Dutal Carcinoma): The most common type of breast cancer, invasive ductal carcinoma begins in the milk duct, has grown into the surrounding normal tissue inside the breast.

ILC (Invasive Lobular Carcinoma): ILC starts inside the lobule but grows into the surrounding normal tissue inside the breast.

Inflammatory Breast Cancer: Inflammatory breast cancer is a fast-growing form of breast cancer that usually starts with the reddening and swelling of the breast, instead of a distinct lump.

[**78**] estimates that 90% of all breast cancers are ductal cancer and only 10% are lobular cancer.

The stages of cancer represent how big the cancer tumour is and whether the cancer has spread. The stages are important because they help the professionals to make the best treatment plan for the patients. The main types of staging system are the number of systems and the TNM (Tumour, Node, and Metastasis) system which is the system for providing more details about how the cancer looks and behaves. These two types of staging systems are shown on table 5-1 and 5-2.

Table 5-1TNM staging system

Letters:	Category:	Description
T : Size	TX	Tumour cannot be measured or found.
	T0	There is no evidence to show the primary tumour.
	Tis	Cancer is 'in situ', means that the tumour has not started to grow into other parts of the body.
	T1-4	1 to 4 represent the size of the tumours and the chance that tumour could extend to chest wall or skin.
N: Lymph node involvement	NX	Tumour cannot be measured or found.
	N0	There is no evidence to show the primary tumour.
	N1-3	1 to 4 represent the size of the tumours and the chance that tumour could extend to neighbouring

		breast tissue.	104
M: Metastasis	MX	Metastasis cannot be found.	
	M0	There is no distant metastasis.	
	M1	Distant metastasis is present.	

Table 5-2Number staging system

N umbers:	Category:	Description
Stage 0		Non-invasive breast cancer
Stage 1		The tumour is smaller than 2cm; No cancer cells in lymph node; The cancer has not spread from breast.
Stage 2	2A	Tumour is smaller than 2cm and has spread to lymph node; or tumour is bigger than 2cm but smaller than 5cm and has not spread to lymph node.
	2B	Tumour is smaller than 5cm and has spread to lymph node; or tumour is bigger than 5cm but has not spread to lymph node.

Stage 3	3A	there is no tumour in breast, and cancer in lump node is smaller than 5cm ad has spread to lymph nodes;
	3B	There are cancer cells in lymph nodes and the cancer has spread to the tissue nearby the breast and may attach the skin or muscle which is surrounded.
	3C	The cancer has spread to lymph nodes and below the breastbone, near the neck or under the collarbone.
Stage 4		Metastatic invasive breast cancer which has spread over breast and lymph nodes to other organs of the body.

5.2.2　Mammogram Datasets Properties

A number of range of screening methods commonly used in clinical practice include mammography, MRI, Ultrasound which to achieve an earlier diagnosis. The dataset of it is an important dataset in the whole breast datasets. These varieties of breast screen give different properties of breast screening dataset in which the mammogram dataset plays a very important role. Therefore mammogram dataset plays a very important role in breast screening datasets. Most of the public mammogram datasets are digitized, because digital mammography overcomes several technologies limitation associated with screen film mammography. This section identifies several aspects of the image characteristics of mammogram data, including intensity properties such as contrast, spatial properties such as texture and structure properties such as breast density.

5.2.2.1　Full Frame Digitized Mammogram

- Overview

The images in mammogram dataset are in digital form, it could be achieved by either digitizing the screen film mammograms or using the digital mammogram machine. Digital mammogram is and will rapidly be forming the basic of modern radiology process to replace the traditional filmed radiography in many health institutions. Digital mammogram data is a capture in digital form, it means that the digital mammogram data are sampled spatially and in signal level.

The most significant property for digital mammogram is it separates the process of image acquisition from the image storage, image display. This helps to optimize of each

of the separated functions and give flexibilities in the adjustment of image display characteristics. A well optimized digital mammogram data provide these advantages:

These dataset has more efficient acquisition.

- Digital mammogram data is captured in numerical form.

- The amount or characteristics of X-ray exposure does not affect the control of display brightness and contrast.

- Digital mammogram could use image processing to adapt data for matching visual performance such as eyes, and also could overcome limitation of the display devices.

- The digital mammogram data could be used to evaluate and compare the performance of a number of technologies such as quantitative imaging techniques, Computer aided diagnosis system, contrast imaging, telemammography.

- Digital mammography or full frame digitized mammography has the abilities to remove other structural pattern such as fixed noise.

To record digitized mammogram, the equipment need special x-ray detectors to perform digitized mammogram. [**79**] stated that digital detectors have properties on increasing efficiency on absorbing the incident x-ray photons, a linear response over a wide range of radiation has enhancement on visibility of subtle contrast differences between the normal background tissue and tumour.

In several public datasets, they use different digitizer to digitize the screen film mammograms. For example, The Digital Database for Screening Mammography

(DDSM) database used four different digitized scanners with different sampling rates to digitize the screen filmed mammograms. They have different grey levels:

Table 5-3 The sampling rate, number of gray levels (M. Heath, 2000)

Digitizer	Sampling Rate (microns)	Gray Levels (bits)
DBA M2100 ImageClear	42	16
Howtek 960	43.5	12
Lumisys 200 Laser	50	12
Howtek MultiRad850	43.5	12

- Spatial Sampling Properties

Both the spatial and intensity distribution of the X-ray transmission pattern are sampled to form the digital mammogram. For digitized mammography, in spatial domain, the interval between pitch samples and the response profile of the detector element (del) will determine the spatial resolution of the imaging system [**80**]. The basic difference between the detectors used for both screen film mammography and digitized

mammography is that the signal varies more or less continuously in spatial and intensity domains for screen film mammograms, however the analogue signal is sampled from the detector of a digital mammography system. In mammogram dataset, the data spatial sampling has been achieved by different detectors using different approaches; however there are some important common concepts to all systems. For example, a detector has been divided into detector elements which are also called del. In image acquisition, every signal detector element provides one or a set of discrete X-ray measurements. These measurements are used to contribute to the image data. Figure 5-3 has showed a simplest example, the signal from one del will provide the information displayed in one pixel of the final image. The del will array at centre to centre distance or pitch p. A detector element contains an active region whose dimension is d. Dels are spaced at one pitch p, non-sensitive edge area on del, the fraction of the area which is sensitive to X-rays (d^2/p^2), can be less than 1.
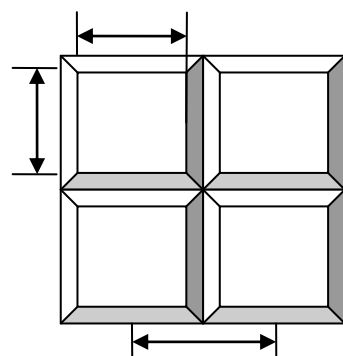
Detector element pitch p

Figure 5-3 The concept form of detector element and spatial sampling from [80]

[**81**] mentioned that digitization improved abilities of detection larger, low contrast subject, but did not change on small objects. His experiment also approved that to a relatively low spatial resolution of 0.1 mm /pixel does not achieve high quality diagnosis performance in digital mammography.

5.2.2.2 Image Properties of Mammogram Data

- Image properties

Spatial resolution is measured commonly in line-pairs per millimetre (p/mm) or cycles per mm, and is used for indicating the size of the smallest visible object structure by using a low kilo voltage and high mAs value. In the meanwhile, it increases the contrast while reducing noise which are the conditions that are very different from the exposure parameters used in diagnostic setting. Spatial resolutions improve the sharpness of the images and allow better detection performance on a really small detailed image. Spatial resolution is measured by a X-ray phantom through the alternating bars and spaces. A high spatial resolution improves the morphological analysis of microcalcifications [**82**].

5.2.2.3 Ages

In the case of breast cancer, most statistics show that there is a steady increase of incident cases in the group of women from 50 to 70 years old. Therefore, EU recommends that breast cancer screening for women between the age of 50 to 69 [**83**] .In the member states of EU, the target age for breast cancer screening varied from 47 to 74 years old. In the UK, the National Health Service (NHS) invite all women between 50 and 70 years old for breast screening on every three years [**31**].Also, the women aged over 70 are not explicitly invited to screen, however they still can arrange

screening if they want. It needs to noticed that, in UK, NHS has started to extend their breast screening programme to invite women in their late 40s and up to 73 years old, and this is expected to be completed by 2016.
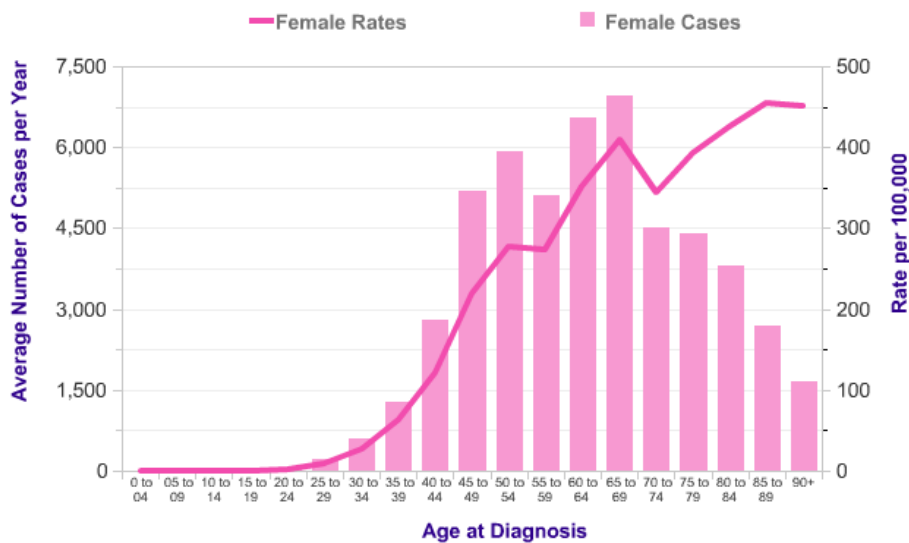
Figure 5-4 Average number of cases per year and Age specified incident rate per 100,000 in UK 2011-2013 copy from [85]

According to Figure 5-4, age groups from 50 to 70 years old still remains at the main peak values. Other age groups around the NHS screening age group also have a high risk in breast cancer development, therefore it is reasonably understandable for NHS to rearrange the age group from 50-70 years old to 47-73 years old, since the age group between 45 to 49 years old also has a high average number of breast cancer cases and incident rate. In addition, the age group of female above 70 years old are still welcome to attend the screening programme and there is no age limitation for older women on their breast screening. That is because it is easier to catch the disease in older women,

therefore, there is also a disagreement about the usefulness and benefit of screening for women older than 70. There is a discussion about whether women around 40s should join the screening routine all over the world. There is no significant evidence to show that women in their 40s should be screened, however [84] used two different databases and reported that mammography screening reduces breast cancer mortality by 15% for women aged 39 to 49.

The world health organisation (WHO) compared the 10 most common causes of death by country and income group for women between 20 to 59 years old [85] . There were 223,000 deaths of women who were aged 20 to 59 caused by breast cancer worldwide. It was the sixth most common cause of death worldwide for women aged 20 to 59. The further statistics followed up by comparing the 10 most common causes of death in a country by income catalogue. In women aged 20 to 59, breast cancer was the fourth cause of death for women in mid-income countries; and the first common cause of death in high-income countries in other words western countries. [29] show that breast cancer mortality in the geographical distribution of the EU, where breast cancer occurrences appears mainly in Western, Middle and Eastern parts of the EU. UK has the highest mortality rates of cancer occurrences while Netherlands, North-Western France, Western Germany, the Czech Republic, Hungary, and parts of Romania and Bulgaria also present high rates; whereas, north and south of that band has lower mortality rate.

5.2.2.4   Reliability

Mammogram is a high contrast, high resolution film and designed specifically for creating a detailed image of the breast. There are eight typical kinds of abnormalities revealed with a mammogram which have been mentioned in Chapter 2. Usually, a

mammographic abnormality is followed up first by additional imaging studies such as Ultrasound, and if the area is suspicious a sample of tissue may be sent for biopsy. The leading risk factor for the development of breast cancer is being a female. The second important risk factor is age.

- Sensitivity

Sensitivity of mammogram is in the proportion of breast cancer that can be detected when the breast cancer exist. Sensitivity relies on a number of factors, such as lesion size, breast tissue density, and age, the hormone state of the tumour, image quality and skill of the radiologists. According to the Cancer stats report about breast screening in UK in 2003 [86], the sensitivity of mammograms for women aged between 50 and 69 have been estimate to be from 69% to 90%, for women aged 40 to 49, the sensitivity is lower at about 62% to 71%.

However, mammography is less effective in finding cancers in women younger than 50 because the breast tissue is still denser compared to the fatty breast tissue for women over 50. Therefore, the early stage breast cancers in young women are harder to find by using screened mammograms. [87] shows that young women have more false positive mammograms and additional imaging but fewer biopsies than older women. [88] did one study about the examination of 576 women who had invasive breast cancer following a mammogram to determine whether it was denser tissue, or that faster growing tumours for young women could cause lower sensitivity. They found that "greater breast density explained most (68%) of the decreased mammographic sensitivity in younger women at 12 months. Whereas at 24 months, rapid tumour growth and breast density explained approximately equal proportions of the interval

cancers."Research shows that the use of computer aided diagnosis systems may help to diagnose more breast cancer using mammography.

- Positive Predictive Value

Although the use of the mammogram has shown its reliability in the early stages of breast cancer detection, a positive predictive value has been recognized as the number of cancers diagnosed per recommended number of biopsies. In the UK, 6% to 8% of women are recalled for further tests after their first screening indicates cancer [86].

- Investigator Bias

Different trials have been introduced to document the capacity of mammography to the diagnosis of breast cancer, or the early detection of breast cancer. These trials also show that the survival rate of cancer is higher in screened women than in non-screened women [32] [86]. However, these comparisons are susceptible to a number of important biases.

Lead time bias:

Survival time from the diagnosis of cancer in mammography includes the time between detection and the time in which breast cancer would have been detected by clinical symptoms. However, this time does not include the survival time of finding breast cancer as a result of symptoms.

Length bias

Mammography detects a cancer when the cancer is preclinical and this duration is not fixed. If breast cancer has a long preclinical time, this cancer probably could be detected by screening, and this kind of cancer is more likely to grow slowly and has a good prognoses.

Over-diagnosis bias:

Screening may find a cancer which grows really slow and maybe never needs to be clinical. This is an extreme form of length bias [89].

Healthy volunteer bias:

The screening population may be healthier than the general population.

- Impartiality

BI-RADS stands for Breast Imaging-Reporting and Data System and it is established by American College of Radiology. BI-RADS are a quality assurance tool designed to standardize mammography reporting where it reduces the confusion in breast imaging. BI-RADS system is now adapted for use with Ultrasound and MRI. It is a quality assurance tool originally designed for mammography. Terms have been developed to describe breast density, lesion features, impression, and recommendations [90]. This system forces the radiologist to assign each case into a different catalogue, and then accurate statistics will be counted to evaluate the radiologists' work. BI-RADS classifications have also helped in monitoring breast cancer diagnosis, treatment and the supporting of breast cancer research, are also easy to calculate statistically.

In UK, the Royal College of Radiologists Breast Group standard scoring system is for improving the communications between the referrals and the radiologists. It helps avoid ambiguity which may cause mismanagement of patients [91]. It is recommended that this standard be used in the reporting of all breast imaging examinations in the U.K. In the national institution in the UK, the patients' management is based on the principle of the 'triple test' which are clinical examination, imaging and needle sampling. If the clinical results are suspicious or the overall imaging results are 3, 4 or 5, the needle sampling will be paramount.

Table 5-4 shows the standard U.K.RCRBG scores system and BI-RADS catalogues with associated meanings and there is third similar five points system developed by the Australian National Breast Cancer Centre. BI-RADS are widely used through North America and some parts of Europe, RCRBG scoring system is applicable in UK practices, NBCC is in collaboration with the Royal Australian and New Zealand College Radiologists.

Table 5-4 Comparison of imaging classification systems from (A.J. Maxwella, 2009)

Category	BI-RADS	NBCC	RCRBG
0	Incomplete assessment:		

	Need to view prior studies and complete additional imaging		
1	Negative, continue routine screening.	No significant abnormality.	No significant abnormality.
2	Benign, continue routine screening.	Benign finding. No further image is request.	Benign finding.
3	Probably Benign, malignancy chance <2%	Indeterminate findings. Requires further investigation usually FNA core biopsy.	Indeterminate/probably benign findings. There is a small risk of malignancy, further investigation is indicated.
4	Suspicious abnormality, coming with biopsy or needle biopsy.	Suspicious findings of malignancy. Requires further	Findings suspicious of malignancy. There is a moderate

		investigation. May require excision biopsy	risk of malignancy. Further investigation is indicated.
5	High suspicious malignant: Biopsy and treatment are necessary.	Malignant findings, requires further investigation, even if non-excision sampling is benign.	Findings highly suspicious of malignancy. There is a high risk of malignancy. Further investigation is indicated.
6	Biopsy proven malignancy, treatment and pending		

5.2.3 Existing Mammogram Datasets

Within mammographic research, there are difficulties with the datasets which cannot be simply generated; therefore the new dataset has to be taken from an existing dataset. Between different datasets, there are a great many differences which have to be taken into account. There are several different publically available datasets and many privately constructed datasets of mammograms for breast cancer research. There are currently four public datasets which have been widely used for different researches. They are Nijmegen database, Mammographic Image Analysis Society (MIAS) database, University of South Florida Digital Mammography Database for Screening (DDSM) and Lawrence Livermore/University of California (LLNL/UCSF) Database. The advance could provide broad range information for CAD system. It remains to be difficult to compare the performance of different algorithms without meaningful datasets. It is not meaningful to compare different systems if the systems are not tested in the same dataset.

5.2.3.1 Nijmegen Dataset

Nijmegen dataset is known as the first public database for screening mammography. It consists of 40 images, since it is no longer available, therefore it cannot be invested any further.

5.2.3.2 The Mammography Imaging Analysis Society Dataset (MIAS)

The MIAS database has been build up by the University of Manchester and was completed in 1994. It has 322 images of different sizes and has a later adapted to form the Mini-Mias dataset. The Mini-MIAS has a number of selected images scaled in

1024×1024 pixels and have been centred in the matrix. It needs to be noticed that the Mini-MIAS have a lower quality than original ones.

The MIAS database is no longer available, and the Mini-MIAS database has been consistently cited for many years but it has failed to represent all the cases of mammograms.

5.2.3.3   *Digital Database for Screening Mammography Dataset (DDSM)*

DDSM first completed on 1999. The examinations of the present cases have been taken between 1988 and 1999. There are 2620 cases in total; four standard views from each case were digitized on one of four different mammography screen exams. The DDSM contains mammograms obtained from Massachusetts General Hospital, Wake Forest University School of Medicine, Sacred Heart Hospital and Washington University of St. Louis School of Medicine.

(d) Left in CC   (a) right in CC   (b) Left in MLO   (c) Left in MLO
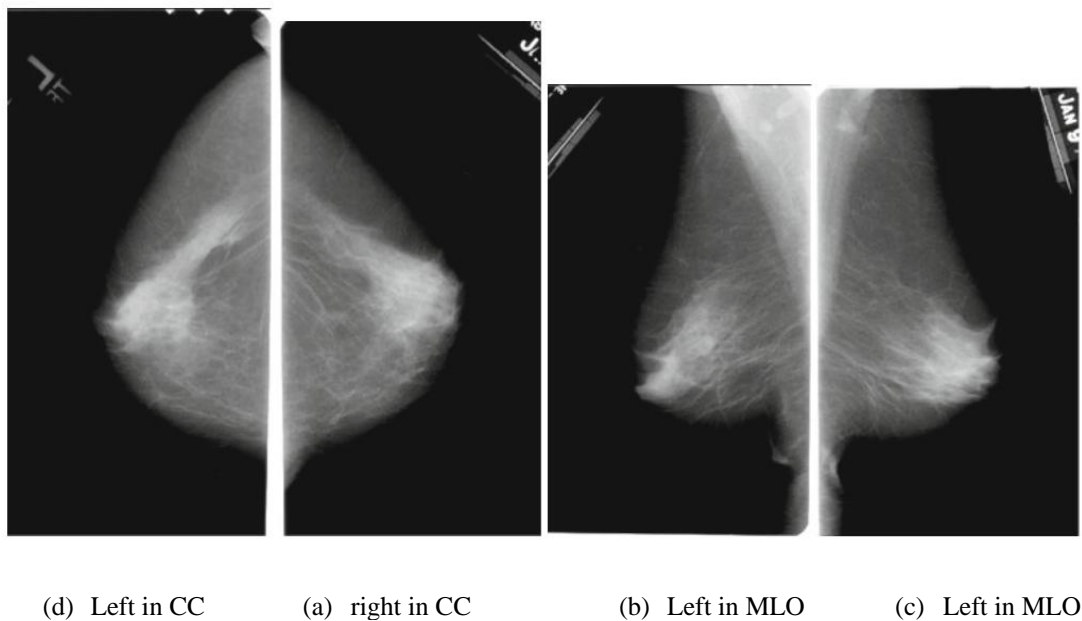
Figure 5-5 A DDSM example case containing mammograms show left and right standard views :cranio-caudal (CC) and mediolateral-oblique (MLO) [7]

Each case in DDSM contains the patient age, the screening exam date, the date on which the mammograms were digitized and ACR breast density which was marked by an expert radiologist. Each marking contains a subtlety value and a radiologist' description using the BI-RADS™ lexicon.

Table 5-5 Contents of DDSM dataset [92]

Institution	Digitizer	Number of cases by most sere finding				Total
		Normal	Benign without Call-back	Benign	Malignant	
MGH	DBA M2100 Image Clear	430	0	0	97	527
	Howtek 960	78	0	446	323	847

WFU	Lumisys 200	82	93	126	159	490
SH	Laser	0	48	202	234	484
WU	Howtek MultiRad850	105	0	96	101	302
Total		695	141	870	914	2620

Several software tools are available for the user to download from the database website, these tools simplify the most common tasks, such as case selection, data extraction and performance evaluation.

DDSM database is an infinitely larger database and is stored in a format which is not conductive to use quickly.

5.2.3.4 The Lawrence Livermore National Laboratory Database (LLNL)

The LLNL has been built in cooperation with the University of California in San Francisco. It has been completed in 1996 and contained 50 cases. Furthermore, the LLNL dataset has been specialised in microcalcifications. In these 50 cases, 20 cases

show benign microcalcifications and 12 are suspicious calcifications and 8 are malignant calcifications.

However, the LLNL dataset is no longer publically available.

5.2.3.5 Other Datasets

In additional to those existing datasets, other datasets exist which have not been made public or are part of partly published project.

- eDiaMonD project

  eDiaMonD project project is a grid computing project for the distribution information on breast cancer treatment promising a vast dataset of digital mammogram images. However, this project has been terminated in 2004 due to a lack of success.

  The IRMA project [93]

  This database is the dataset containing the texture pattern extracted from digitized mammograms of different BI-RADS classes. It collects mammograms image data from DDSM, MIAS, LLNL and RWTH databases and extracted patches and resized the image data to 128×128 pixel. The dataset could be catalogued into two different slots:

  12er_patches contains 2796 patches of 12 classes, each with 233 images;

  20er_patches contains 880 patches of 20 classes, each with 44 images.

  They are used for image retrieval and computer-aided diagnosis.

5.2.3.6 Limitation of Mammograms datasets

There are available summarized essentials of mammography datasets when subjective evaluation is involved as multiple assessors. Medical datasets are more sensible than other datasets based on their natures and effects. A good mammography dataset requires multiple assessors to provide greater confidence. Since the diagnosis of screening mammogram has a high false positive rate and the high risk of false positive, in some areas, there are two radiologists who are required to diagnose one mammogram to try to avoid any misdiagnosis and later may even use a biopsy to provide more diagnostic confidence.

However, a huge amount of raw information has disadvantages as these may confuse the system and decrease decision accuracy. These datasets have uniformity, but actual performance evaluation methodologies related to data selection is lacking consistency in for example:

- Classification confidence.

  Classification confidence represents the level of suspicion about malignancy on imaging. These levels are numerically categorised, however there are significant variations in the definitions of each numerical category between different areas.

- demographic spread

  There are significant variations between countries in breast cancer [7], this is also true of diagnosis and survival rates. Figure 1 shows the rate in age-standardised per 100000 women between high-income and low-middle income countries in 2004.

- Diagnosis

The radiologists in different areas have been trained differently to classify mammograms; this may lead to confusion and potential error when staff moves between areas.

6 Design the Multi-Chromosomes Cartesian Genetic Programming (MCGP) network for automated analysis mammograms

The aim of project is to investigate the suitability of the new algorithm called multi-chromosomes Genetic Programming network (MCGP) as described previously to ensure a reliable diagnosis of mammograms as part of a CAD system. For automated analysis of mammograms, it is possible to apply raw pixels to the network and train on them. One of the strengths of an evolutional algorithm is that can discover new solutions a conventional design would never uncover and it does this by its random nature. If conventional features are extracted to be fed into the network, then this puts a major limit on what the network can achieve in an absolute sense. Putting the raw pixel values in essentially free the network to extract whatever information it decides is most important. The downside is that in order to cover on a solution there is the possibility a much larger network may be needed. This increases the runtime which could already be a problem which MCGP has the potential to overcome.

This chapter describes the experimental methods in detail that will be used in automated analysis of microcalcifications. Methods include dataset used for experiment, details of the new representation of Evolutionary Algorithm, which is also used for the application and details of performance measurement.

6.1 Experimental dataset

In the entire research, an algorithm of MCGP (as described in Chapter 4) has been used as a solution to automated assessment of microcalcifications in mammograms.

Therefore the microcalcifications mammogram images have been obtained. The data for both the training and testing the algorithm was obtained from the LLNL database. As described in Section 5 compared to other databases, specialising in microcalcifications of the LLNL database is outbalance by the disadvantages of this database. No information has been found as to exactly how many malignancy calcification cases were found in the other three public databases. The other three public databases also contains masses and asymmetries.

In total 31 images were created, of which 13 contained malignant microcalcifications images and 18 benign microcalcifications images from 5 separate patients.

6.1.1   Region of Interests (ROIs)

As a result of image restriction of the available data, there is a reduction of image size as a region of interest (ROI) has been marked from original mammograms and supplied by Dr. Eddy Munday. Each ROI has a size of 128 ×128 pixels . If the ROI has to be chosen to be 256 ×256 pixels, that is because the ROI of 256 ×256 pixels would have contain too many relatively large areas of surrounding tissue and extremely smaller areas of malignant microcalcifications. As a description in [94], one critical feature of the development of malignant microcalcifications is that when the surrounding tissue under these conditions grows to an approximate radius of 180 μm, then necrosis or cell death might be reached. This is because the tumour has grown beyond this size and any tissue outside this radius could be diagnosed as benign, a making the situation not suitable for a classification stage. Therefore a smaller ROI with 180×180 pixels haves relatively the best conditions which still contain a significant amount of healthy tissues, as microcalcifications takes about 5% of the ROI. One ROI is featuring at least one

microcalcification, this selection provides 31 ROIs from the dataset of 18 containing benign microcalcifications and 13 containing malignant microcalcifications.

Using raw pixels values to classify the mammograms approach, it can therefore be appreciated that any processing using MCGP network on full field mammograms will take an unacceptably long time. Subsequently, the mammograms will have to be partitioned in to smaller part images to make feasible processing.

6.1.2 Grid Structure for Segmentation Representation

The segmentation is developed as a sequential representation providing a different approach for filtering an image. The segmentation representation theoretically allows for parallel processing, the aim is to break down the overall problem into a number of sub-problems which offers potential to be processed in parallel.

Images containing different information are split into different parts. Based on this representation of segmentation, some areas contain microcalcifications which may differ greatly. There are also some areas which regularly holds fairly similar information since they contain only a breast tissue background.

Therefore, the whole image has been divided into smaller parts, using a MCGP network, it was decided that one image is divided up into equally sized parts of non-overlapping rectangular shapes. The basic principle is shown in Figure 6-1.
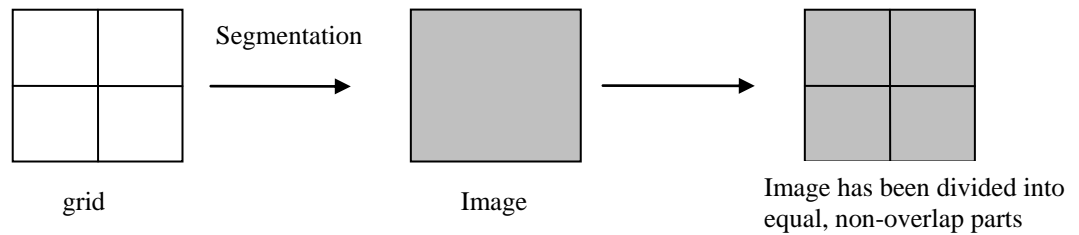
Figure 6-1Process for grid structure of segmentation representation

Since the image size is fixed and the number of image parts depends on the size of the sub-image. For the particular experiment image of 128×128 pixels, the size of the sub-image needs to be rectangular or squared so the whole image could be divided exactly into an integer number of parts. A greater number of image segmentation means that the size of individual image segments becomes smaller. Having too many sub-images will result in a large network; therefore [**95**] did experiments to find out the best size for a sub-image by using the CGP network. The results show that a sub-image sized 8×8 pixels gives the best performance between fitness and best average fitness. As a result, 128×128 pixels ROI have been divided into 256 non-overlapping 8×8 pixels areas or parts. Figure 5-8 shows the process from choosing a ROI and then logically divided into 256 parts and the status of each part labelled as either being benign or malignant according to the radiologist.
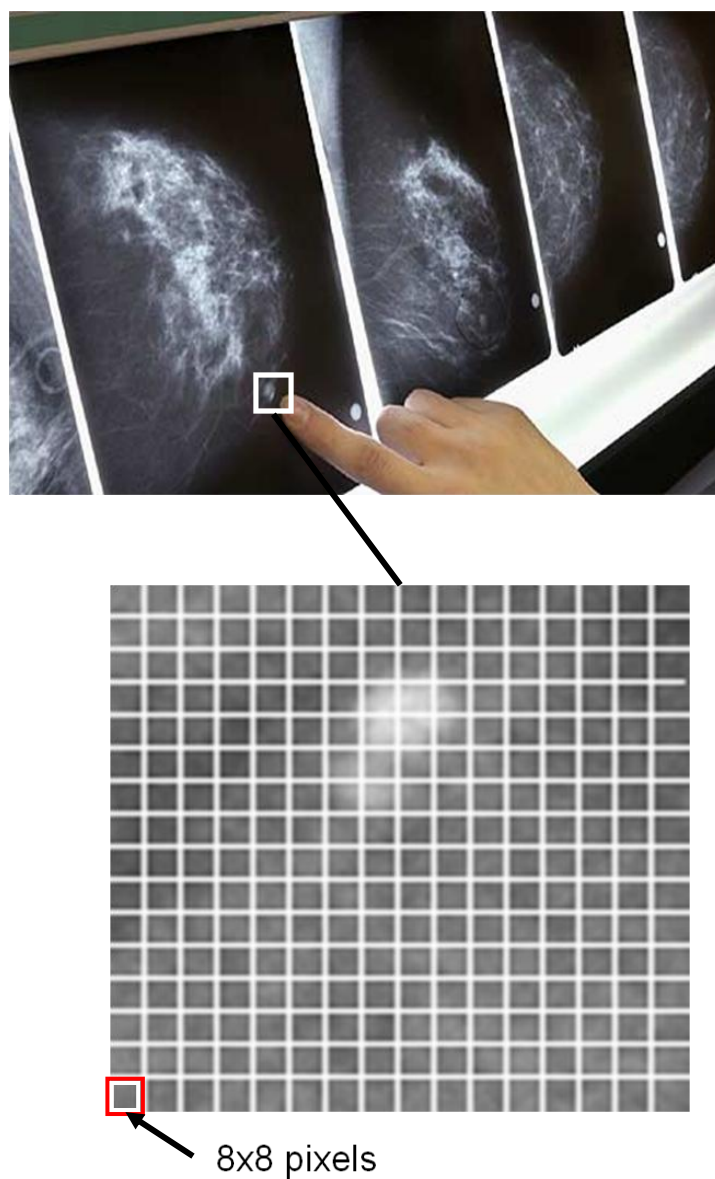
Figure 6-2 A ROI has been point out by the expert and then be devided into 256 equal non-overlapping parts

6.1.3   Additional information for dataset

As each ROI has been divided into 256 parts, each part could be seen as a small image

with 8×8 pixels with 8 bit grey scale. For each pixel, it has a range between '0' to '255'

in order to process pixel values of 8 bit depth. The detection by radiologists of microcalcification will use the grey shading and the white oral in the figure representation and mammogram ROI, as respectively have been provided by the radiologist experts. Two additional data files are one to one mappings of the original ROIs which are provided by radiologists. The values of these additions are in a smaller range than the pixel values in ROIs. These additions will be used as reference materials in order to achieve a better training system; it is also used on evaluating the performance of the algorithm in the following chapter.

6.1.3.1 White Pixel Value (WPV)

For ROIs only with benign microcalcifications, one of the corresponding additional files is on the pixel level by using '0' to indicate breast tissue pixel and '1' for whether this pixel contains information of microcalcifications. Therefore, there is a respective parts file which has then been generated for each region of interest, which simply provides the respective number of white (or to be precise, non-zero intensity) pixels for each section of the image. Since each part only has 64 pixels, therefore the range of white scale values is from 0 to 64. These white pixels count numbers will be called white pixel value (WPV) in work. The other one also uses '0' and '1' but in a level of 'part'. For each part, there is an indication of '0' to indicate benign tissue located on that part and the number '1' to indicate that that particular part contains microcalcifications.
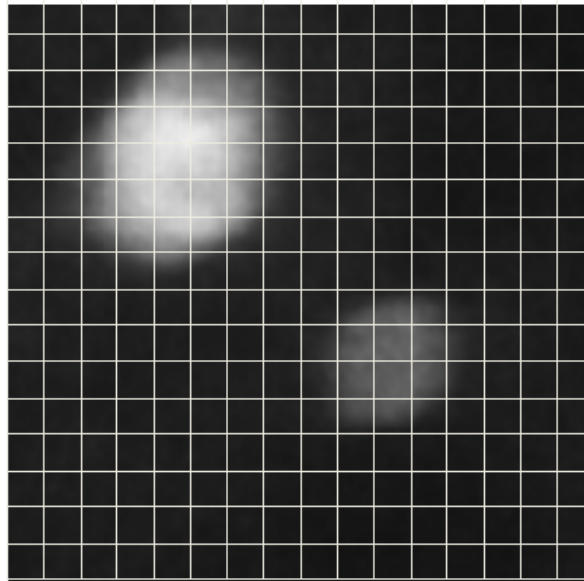
a) Original ROI with 256 equal non-overlapping parts

0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
0	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

0	0	4	30	13	22	7	0	0	0	0	0	0	0	0	0
0	5	55	64	64	64	60	9	0	0	0	0	0	0	0	0
0	29	64	64	64	64	64	49	0	0	0	0	0	0	0	0
0	44	64	64	64	64	64	64	0	0	0	0	0	0	0	0
0	55	64	64	64	64	64	59	0	0	0	0	0	0	0	0
0	54	64	64	64	64	64	32	0	0	0	0	0	0	0	0
0	30	63	64	64	64	52	4	0	0	0	0	0	0	0	0
0	0	7	32	30	15	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	19	44	53	0	0	0	0
0	0	0	0	0	0	0	0	10	64	64	64	5	0	0	0
0	0	0	0	0	0	0	0	20	64	64	50	0	0	0	0
0	0	0	0	0	0	0	0	3	32	33	7	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

b) Corresponding to ROI, '0' indicate breast tissue,'1' indicate existing microcalcifications

c) File representation of white pixels count from respective parts of mammogram ROI

Figure 6-3 An example of additional data for ROIs only with benign microcalcifications

Figure 6-3 shows one example of a ROI (Figure a) has been divided into 256 parts and each part has a relatively '0' or '1' to indicate microcalcifications in part level (Figure b) and 'white' pixel value in pixel level (Figure c).

6.1.3.2 ROIs only with malignant microcalcifications

Similarly as the ROIs with benign microcalcifications, after each image, is logically divided into 256 parts, the status of each part labelled as either being benign or malignant according to the radiologist. The number of pixels containing information of malignant microcalcifications in one part have been provided by radiologists as well. '0' is used to indicate not only benign microcalcifications but also breast tissue and '1' is used to indicate malignant microcalcifications. Therefore, the corresponding data for benign ROIs are fully '0's since there is no malignant microcalcifications.

a) An example of malignant ROI

b) The white pixel values for the corresponding ROI

Figure 6-4 An example of additional data for ROIs only with malignant microcalcifications

6.1.4   K-Fold Cross Validation

The performance of any algorithm is only as good as the data it trained on. Presenting the algorithm with a variety of images leads to the algorithms ability to deal with different problems and provide strong results. In other words, it could also be understood with the greater the number and diversity of the problems an algorithm is trained, the more stable and strong the results will be obtained. Therefore in order to obtain statistically meaningful measures of performance, computer learning algorithms should be trained and tested on large datasets. There are 31 images from 8 patients included in the training and testing together in dataset. The relatively small number of images available to train and test the CGP network has required that k-fold cross validation be used [12] . In this method, the original sample is randomly partitioned into k subsamples, each subsample is used as the validation data and the remaining k-1 subsamples are the test data, then the process repeats k times with the folds, each time it makes a different k subsample as the validation data. The advantage of k-fold cross validation is that repeated random sub-sampling leads to all samples being used for both training and validation, and each sample is used for validation exactly once. [96] The main advantage of cross validation is that all cases in the data set are used for testing. The main disadvantage is that for each different data set, different classifiers may be learned.

For detection of microcalcifications stage:

Only 18 benign microcalcifications mammogram images will be used to train and test algorithm. As described in chapter 2, benign and malignant microcalcifications have different characters, since the detection of microcalcifications stage is a pre stage to

determine whether MCGP has strength to classify microcalcifications on mammograms, therefore only using benign microcalcifications mammograms images will reduce the complexity of this algorithm and could lead to more stable and clearer results. Therefore, there are two experiments which are used to assign the images into the fold in different ways. There are 5 patients for 18 images, so in total there are 5 sub folds for each experiment, and each fold could be treated as a sub-experiment of one experiment. In the first experiment, the images of each patient are used as the testing data in 5 different folds, as shown in table 6-1. In the second experiment, a random selection of images are used from the patients as the training data, the images for the same patient could be partitioned across several folds, as shown in table 6-2. In both table 6-1 and 6-2, the 5 letters are used to represent the different patients' names, followed by a series of numbers to separate the different images which have come from the same patient. For the purpose of this investigation two experiments are conducted, each with the data assigned to the folds in a different way. In the first experiment, all images associated with each patient are used to form a separate fold; in the second experiment, images from each patient are distributed across several folds. This assignment is summarized in Tables 6-1 and 6-2 and is intended to investigate the power of the algorithm to detect microcalcifications from a particular patient. For each fold in tables 6-1 and 6-2, the remaining images are used to train the network.

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Patient A: Image 1	Patient B: Image 1	Patient C: Image 1	Patient D: Image 1	Patient E: Image 1
Patient A: Image 2	Patient B: Image 2	Patient C: Image 2	Patient D: Image 2	Patient E: Image 2
Patient A: Image 3	Patient B: Image 3	Patient C: Image 3		Patient E: Image 3
	Patient B: Image 4	Patient C: Image 4		
		Patient C: Image 5		
		Patient C: Image 6		

Table 6-2 Distributed assignment of images for k-fold cross validation

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Patient D: Image 1	Patient B: Image 1	Patient E: Image 1	Patient D: Image 2	Patient E: Image 2
Patient C: Image 2	Patient E: Image 3	Patient C: Image 6	Patient C: Image 3	Patient C: Image 4
Patient B: Image 2	Patient C: Image 1	Patient B: Image 4	Patient A: Image 3	Patient C: Image 5
Patient A: Image 2	Patient A: Image 1			Patient B: Image 3

For classification of microcalcifications stage:

There are 31 images from 8 patients which will be used in this stage to train or test the algorithm. All these images contain one or two microcalcifications that are either benign or malignant. Similar to experiment on detection of microcalcifications, there are two experiments are conducted on classification of microcalcifications, each with the data assigned to the folds in a different way. In the first experiment, which is summarized in table 6-3, patients are used to form a separate fold; in the second experiment which is

138

summarized in table 6-4, images from the same patient are distributed across several folds. In both tables, the letter 'M' represents a malignant image and letter 'B' refers to a benign image. For each fold in tables 6-3 and 6-4, the remaining images are used to train the network.

Table 6-3 Patient -centred assignment of images for k-fold cross validation

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Patient A M 1	Patient C M 1	Patient D M 1	Patient F M 1	Patient bwrcc M 1
Patient A M 2	Patient C M 2	Patient D M 2	Patient F M 2	Patient bwrcc M 2
Patient B B 1	Patient C M 3	Patient E B 1	Patient F M 3	Patient H M3
Patient B B 2	Patient C B1	Patient E B 2	Patient G B 2	Patient I B1
Patient B B 3	Patient C B2	Patient E B 3	Patient G B 2	Patient I B2
Patient B B 4	Patient C B3	Patient E B 4		Patient I B3
		Patient E B 5		
		Patient E B 6		

Table 6-4 Distributed assignment of images for k-fold cross validation

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Patient A M 1	Patient A M 2	Patient C M 3	Patient C M 2	Patient F M 2
Patient C M 1	Patient F M 1	Patient D M 3	Patient D M 2	Patient H M 2
Patient D M 1	Patient H M 1	Patient H B 3	Patient E B 3	Patient B B 4
Patient C B1	Patient E B2	Patient C B3	Patient C B 2	Patient E B 4
Patient B B 1	Patient A B1	Patient B B 3	Patient B B 2	Patient E B 5
Patient E B 4	Patient IB1	Patient E B 6	Patient G B 2	Patient I B2
		Patient IB 3		

6.2 MCGP Network Design

The premise for this Multi-chromosomes CGP (described in Chapter 4) approach is for a physically large size of mammograms. Traditionally, the current technology provides

a pixel size of 50μm^2 which is equal as a full field mammogram of 4800×6000 pixels [**97**]. Each chromosome in multi-chromosomes could be seen as an independent CGP chromosome network, the whole collection of multiple CGP chromosomes are used to evolve a single mammogram. This would be like a whole team of experts working together to solve a particular task, and this will improve the reliability and be highly effective.

Crossover or combination has not been involved within this network since it has not been found to be beneficial in CGP especially [**98**]. Therefore, it has been perceived that better evolutionary processes are obtained using mutation only.

As a result of segmentation of the parts in Section 6.1.1, each ROI image has 256 equal, non-overlapping sub-images. Having segments of equal size and shape ensures that chromosomes evaluating those parts could be identical. This provides a way of introducing multiple chromosomes. Therefore, each sub-image gets assigned its own chromosome; this means that the whole problem of evolving one large image is broken down into a range of small sub-problems. Each genotype contains 256 independent CGP chromosomes and each chromosome is encoding one sub-problem. In addition, each chromosome is a working individual at the same time; therefore, high levels of efficiency have been achieved to make the system work faster.

MCGP networks   +   Sub-images from one ROI

Networks assigned with sub-images

Figure 6-5Multi-chromosome CGP network has been assigned to individual sub-image

As each ROI part get its own CGP network assigned, the respective network is only responsble for generating the output value for one particular image part it is assigned to. Since the input values are pixel values, the output value from one CGP network is also in a range of [0,255] which can be re-arranged to compare with WPV to that part.

A ROI with equal non overlapped 256 parts

MCGP network with 256 single CGP networks

P_1^1	P_1^2	...	P_1^{15}	P_1^{16}
P_2^1	P_2^2	...	2^{15}	P_2^{16}
⋮	⋮	...	⋮	⋮
P_{15}^1	P_1^1	...	P_{15}^{15}	P_{15}^{16}
P_{16}^1	P_{16}^2	...	P_{16}^{15}	P_{16}^{16}

1/1	1/2	...	1/15	1/16
2/1	2/2	...	2/15	2/16
⋮	⋮	...	⋮	⋮
15/1	15/2	...	15/15	15/16
16/1	16/2	...	16/15	16/16

8×8 → ◯ → 1 → ☐

one part with 8×8 piexls

one CGP network

Output has a range from 0 to

$(output +1)/4$

☐

Output has a re arrangment from 0 to 64

MCGP output with 256 CGP outputs

P_1^1	P_1^2	...	P_1^{15}	P_1^{16}
P_2^1	P_2^2	...	2^{15}	P_2^{16}
⋮	⋮	...	⋮	⋮
P_{15}^1	P_1^1	...	P_{15}^{15}	P_{15}^{16}
P_{16}^1	P_{16}^2	...	P_{16}^{15}	P_{16}^{16}

Figure 6-6 The working procedure for MCGP network

6.2.1   Parameters

There are several parameters needed in order to choose and to let the algorithm give the best performance. Since the numbers of chromosomes have been decided, there was a need to decide on a size of each chromosome CGP network.

These are two aspects as the following has had to be considered:

Firstly, the network needs to be considered big enough to allow free evolution. Generally, the larger the single chromosome networks will often lead the higher fitness score; this also means there will be a large variety of possible binding abilities.

Secondly, there is a desire to design a small single chromosome network to reduce the limit of the computational demand needed for evolving this network [99].

In [95], it has been shown that the ideal combination of fitness and minimum size was found to be a network size of 32 rows and 128 columns. This affords redundancy proven to be advantaged in evolution of CGP networks by [100]. However, in this work, the rows will be chosen as 1, the number of columns will be chosen as 500 as minor. That is because; firstly, the number of rows will not affect the performance of the algorithm; secondly, since a multiple chromosome network has been chosen, the individual single chromosome network needs to be as simple as possible. This will make sure there is a better vision of the inside the CGP network. In this work, MCGP are used to evolve a single mammogram. A generalized form of the 1+ 4 evolutionary strategy has been used in which each chromosome is selected. This has the form:

(1+4) Multi-chromosome Evolutionary Strategy

1) Generate initial population of 1+4 programs randomly

2) Evaluate fitness of the population

3) Select the fittest of the population as the parent

4) Mutate the parent to generate 4 offspring to form the new population

5) Select a new parent using the following rules:

for all i chromosomes in [0, 256):

5.1) If an offspring chromosome i has a better or equal fitness that the parent chromosome i then the offspring chromosome i is as the new parent chromosome i.

5.2) Else if there are offspring chromosome i with the same fitness as the parent chromosome i, one of them is randomly selected to be the new parent chromosome i.

5.3) Else the current parent chromosome i remains as the parent chromosome i.

6) Go to step 4 unless the maximum number of generations has been reached.

The higher the number of generations, the better the fitness results will achieve. However efficiency of the network needs a binding for the number of generations. In order to lead the network to have enough room to approve the performance, with a reasonable evolution time, the number of generations will be set as fixed at as 10000. A mutation operator can alter both the function present within a grid cell and the connections between components. A good mutation rate is important because mutation

plays a crucial part in the evolution of CGP networks. There are two different areas that need to be mutated, the connections of individual nodes made with other nodes and the functions used for calculating the output of the respective node. In order to investigate the effect of the mutation rate value to resulting fitness value, it was decided to execute the algorithm using different mutation rate values. To make the results comparable all settings other than the mutation rates were kept same. The average fitness values are shown in table 6-1 with different mutation rates.

Table 6-5 Mutation rates

Mutation rate %	Average fitness
0.1	0.9324
0.25	0.9301
0.5	0.9310
1.0	0.9310
3	0.9204
5	0.9100

The results show that there are two peak values for mutation rates at 0.1% and 1%, therefore it can be said that the mutation rate should be about 1% or less to achieve a good fitness. However there are the same peak values so that it is hard to tell which value should be used. Mutation rate 1% and 0.5% have the same results. It was decided to use 1% for the algorithm in the experiment, the reason is 1% has a peak value and is the binding of the mutation rate, besides it is easy to calculate. It means that 5 genes were mutated from the parent to make each offspring (0.25% of 500*4+1 genes).

Table 6-6 The summary of the parameter values used for the evolution of the MCGP network

Parameter	Value
Number of generations	10000

Population_size	5
Percent of mutation rate	0.25
Number of runs in total	20
Number of rows in each CGP network	1
Number of rows in each CGP network	500
Levels_back	500
No parts per image	256 (16×16)
Part size	8×8
Function sets	$f_0 = x.z' + y.z$, $f_1 = x.z' + y'.z$

For each experiment, the network programming has been set to run 20 times in order to give 20 individual results for the same experiment, this is to make sure that all the experiment results are constant and stable, based on the performance of the network. One of the measures, which are used to direct address the algorithm's, is to involve the image pixel value was function set. The function set has a great effect on the speed of evolution and also on the performance of an algorithm; therefore a suitable function set is really important. A bad function set, if used, might lead more time to find a good result. The complex multiplexer was used for the function node, due to the fact it contains, not only logical AND/OR gate, but the whole algorithm works on a binary classification. The multiplexer functions represent binary IF statements and have been shown to be effective when CGP chromosomes operate on the binary data [101]. Assuming the 8 bit binary inputs from the image are defined as x,y and z, the function defines as follows:

$$f_0 = x.z' + y.z , \quad f_1 = x.z' + y'.z \qquad \qquad \textbf{Equation 6-1}$$

The symbols '.''+' and ',' indicated 8 bit bitwise logical AND/OR operations and even more complement operations.

6.3 Performance Measurement

6.3.1 Confusion Matrix

Confusion Matrix is a visualisation tool that can be used for evaluation of the performance of a classifier, in this paper, MCGP. Based on this theory, it could determine certain quality measurements for the performance, such as a receiver operating characteristic (ROC) curve.

The confusion matrix consists of two rows and two columns. The two rows represent the outcome of the prediction value or the fact or the condition determined by a gold standard. The two columns represent the outcome of the test that is to be evaluated. Both fact and test are divided into two options: a positive or a negative classification. As a result there are four different categories a test can fall into: a true positive (TP) classification , a true negative (TN) classification, a false negative (FN) classification, where a condition has not been found by the test and a false positive (FP) classification where the test has wrongly indicated a find. The full confusion matrix shown in Figure 6-6.

Figure 6-7 Confusion Matrix

Depending on the four values of TP, TN, FP, FN, it could determine other measures. The most common measures are sensibility, specificity and precision.

- The precision is also called positive predictive value (PPV) which is the chance of having cancer given a positive test result.

  PPV = TP / (TP+FP)

- Sensitivity is also called the true positive rate (TPR), or called hit rate and recall, it is the proportion of patients that tested cancer (positive) and prediction is cancer (TP) of all the patients that actually have cancer (TP+FN).

  TPR = TP / (TP+FN)

  Sensitivity can be seen as the probability the test result is cancer given that the patients have the cancer. With higher sensitivity, fewer actual cases of the disease go undetected.

- Specificity is also called true negative rate (TNR) which is the proportion of patients that tested benign and the prediction is benign (TN) of all the patients that actually have no cancer (TN+FP).

  TNR = TN / (TN+FP)

Specificity can be looked at as the probability that the test result is benign (negative) given that the patient is not sick. With higher specificity, fewer healthy people are labelled as sick.

- The false positive rate (FPR) also called false alarm rate = False positive/Total Negatives.

FPR = FP / (TN+FP)

As there are cases in which it might be possible to achieve a high TPR but a low NPR and vice-versa, it in general advisable to have at least two measures or a measure that combines different of the previously mentioned measures.

The sensitivity and specificity, as well as the performance of the classifier, can be visualized and studied using the Receiver operating characteristic (ROC) curve, also called sensitivity curve. To draw an ROC curve, only the true positive rate (TPR) and false positive rate (FPR) are needed . The TPR defines how many correct positive results occur among all positive samples available during the test. FPR, on the other hand, defines how many incorrect positive results occur among all negative samples available during the test.

A ROC space is defined by FPR and TPR as *x* and *y* axes respectively, which depicts relative trade-offs between true positive (benefits) and false positive (costs). Since TPR is equivalent to sensitivity and FPR is equal to 1- specificity, the ROC graph is sometimes called the sensitivity vs. 1-specificity plot. Each prediction result or instance of a confusion matrix represents one point in the ROC space.

Figure 6-8 shows an example of a typical ROC curve. ROC analysis is used in clinical epidemiology to quantify how accurately medical diagnostic tests can discriminate between two patient states, typically referred to as benign and cancer [**100**] . ROC analysis is used in clinical epidemiology to quantify how accurately medical diagnostic systems) can discriminate between two patient states, typically referred to as "cancer" and "benign" [**101**].
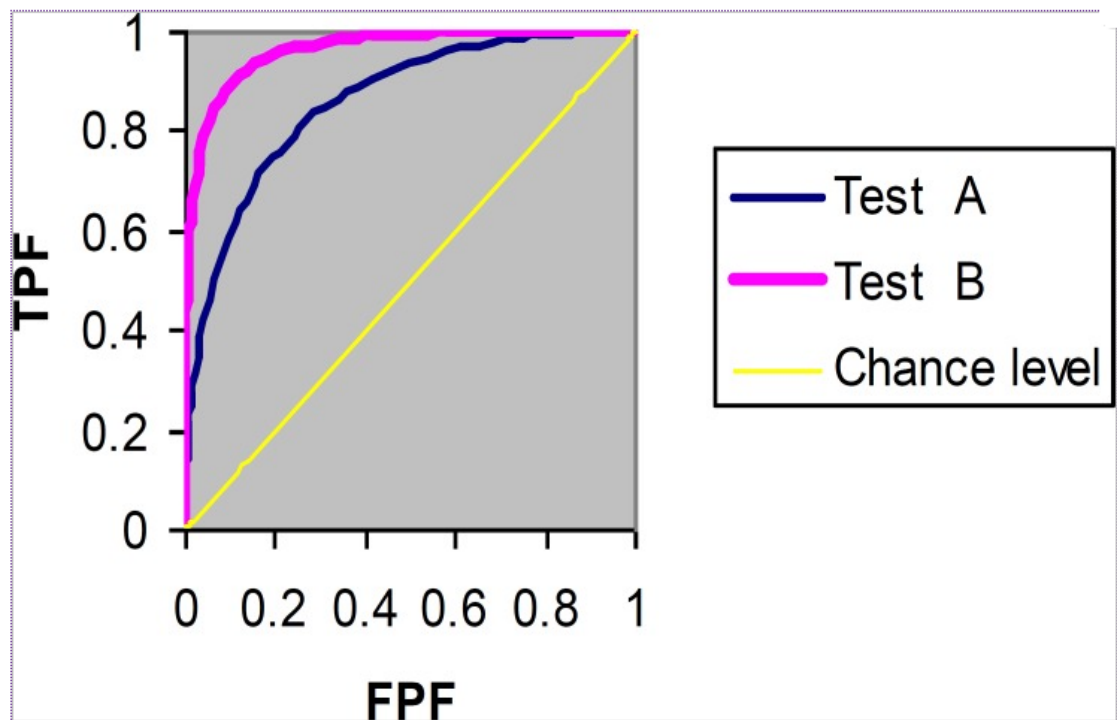
Figure 6-8 An example of ROC curve copy from [104]

In figure 6-8, Test A and B are examples of ROC curve, test B has a greater discriminate capacity than test A. Maximizing sensitivity corresponding to some large y value on the ROC curve. Maximizing specificity corresponds to a small x value on the ROC curve. Therefore a good first choice for a test cut off value is that value which corresponds to a point on the ROC curve nearest to the upper left corner of the ROC graph. This is not always true, in some screening applications it is important not to miss

detecting an abnormal therefore it is more important to maximize sensitivity minimize FN than to maximize specificity. In this case the optimal cut-off point on the ROC curve will move from the vicinity of the upper left corner over toward the upper right corner. An ROC curve lying on the diagonal line reflects the performance of a diagnostic test that is no better than chance level. The diagonal line $y = x$ represents the strategy of randomly guessing a class. For example, if a classifier randomly guesses the positive class half the time, it can be expected to get half the positives and half the negatives correct; this yield s the point (0.5, 0.5) in ROC space.

The area under the curve (AUC) summarizes the entire location of the ROC curve rather than depending on a specific operating point [103]. The AUC is an effective and combined measure of sensitivity and specificity that describes the inherent validity of diagnostic tests . [104] If the AUC area is equal to 1.0 then the ROC curve consists of two straight lines, one vertical from 0,0 to 0,1 and the next horizontal from 0,1 to 1,1. This means that the test is 100% accurate because both the sensitivity and specificity are 1.0 s there are no false positive and no false negatives. On the other side a test that cannot discriminate between normal and abnormal corresponding to an ROC curve that is the diagonal line from 0,0 to 1,1. The ROC are for this line is 0.5. ROC curve areas are typically between 0.5 and 1.0 like shown in Figure 6-8.

In this paper, the measure of the performance is called Matthews Correlation Coefficient (MCC). MCC could be calculated from a confusion matrix straight away, and used well on a binary classification. It can be used as a single value for the evaluation of test. The MCC is calculated in Equation 6-2 and in a range of [-1, 1].

MCC =1 represents a perfect prediction; MCC=0 means random prediction and -1 indicates the faulty prediction.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$

Equation 6-2

7 The Multi-Chromosome CGP on Mammograms Experiment Results

There are two stages of experiments processed in this chapter and they are the detection of microcalcifications and classification of microcalcifications. Detection of microcalcifications is the pre-stage to investigate the suitability of MCGP into mammogram applications before it is used to diagnosis mammograms. Therefore, a MCGP network is constructed to detect microcalcifications and will be used to classify microcalcifications. The Centre of Gravity has been applied to evaluate the detection system and Matthews Correlation Coefficient is used to evaluate the performance of classification of microcalcifications. There are some further experiments with interesting results which are described in this chapter.

7.1 Detection of Microcalcifications on Mammograms

The detection of microcalcifications is the development of MCGP, and specifically, the potential to train the algorithm to successfully locate microcalcifications within the mammogram. It is important to state that the motivation is not merely to detect microcalcifications.

7.1.1 Design Stage

As described above, each chromosome in MCGP is used to work on one part of a whole ROI image as using the raw pixel values. Therefore a well organised network inputs will lead a better efficiency for the MCGP performance. The network was designed to deal with a range of values from 0 to 255 in order to process pixel values of 8 bit depth. The output of the program encoded in each chromosome is automatically an integer

153

between 0 and 255. This means the ratio between inputs and outputs is different when the detection is supposed to be performed. White pixel values (WPV) are used to be the ideal value to represent the intensity of white pixels for microcalcifications. The WPV has a value between '0' and '64', and it represents how many 'microcalcification' pixels counted in one part (8×8) image. The range of values that were allowed from one chromosome CGP network is not changeable, this means the CGP network only returns a value between '0' and '255' because the value of inputs is from '0' to '255', this value needs to be assigned the meaning of how many microcalcification pixels could CGP detect in part of the image. The fitness function is giving the meaning about the potential of differences by comparing the ideal WPV to the output of CGP, and gives a value for the difference of these two for this performance. If fitness is equal to 0, then CGP could detect microcalcifications in that part successfully. Therefore, the output of one CGP chromosome needs to be rescaled by adding 1 first and then be divided by 4 to arrive at a number between 0 and 64. To calculate the fitness of a particular chromosome i (i is the serial number fir 256 parts in one ROI image), the 64 grey scale pixel values is applied from a unique section of the image as the input to the encoded program and compare the output of the program, D_i with the known number of microcalcifications pixels in the image section (WPV) (a number between 0 and 64). Then '1' is used to minus the difference of rearranged output and WPV to convert the best fitness from '0' to '1'.

The equation for calculating fitness is shown in equation7-2. The principle of the fitness function is shown in algorithm 1. The number of white pixel values of each part is M_i; the output of chrosome i is D_i. There have been two used fitness functions in the work reported. The linear fitness of a single chromosome, i, fitness is defined as follows:

$$fitness = \frac{1}{i} \sum_{i=1}^{256} \left(1 - \left|\frac{D_{i+1}}{4} - M_i\right|\right)$$ **Equation 7-1**

$$fitness = \frac{1}{i} \sum_{i=1}^{256} \left(1/\left|\frac{D_{i+1}}{4} + M_i\right|\right)$$ **Equation 7-2**

Algorithm 1 simple fitness calculation
fitness_value = 1-|number of whitepixel values of each part – (network_output+1)/4|

The fitness of the genotype is by summing all the chromosome fitness up and dividing by the number of chromosomes (256). The non-linear function penalizes the difference between the CGP output and the true microcalcification counts to a large extent than the linear fitness.

Table 7-1Assignment of patient images for k-fold cross validation and use of fitness function in experimental investigations

Experiment	k-folds assignment	Fitness function
1	Patient-centered	Liner
2	Patient-centered	Non-linear
3	Distributed	Liner
4	Distributed	Non-liner

7.1.2   Training Stage

Both linear and non-linear fitness functions have also been employed. A summary of these experiments, provided for the purpose of reference when considering results, is detailed in Table 6-5. A summary of the parameter values used for the evolution of the CGP networks for each experiment is given in Table 6-4. Fitness results for the training of the respective CGP networks are shown in Figure 6-2.
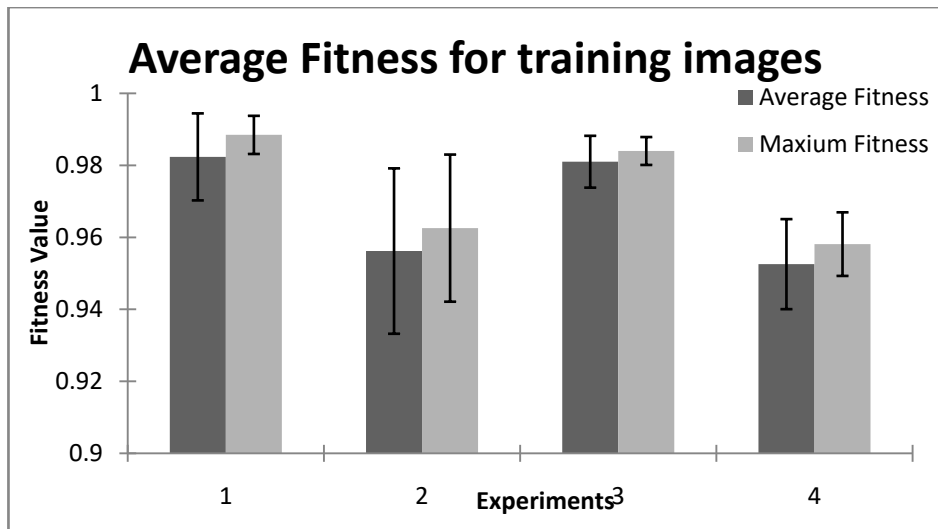
Figure 7-1 Average Fitness for Training Images

The average and best fitness values are shown in Figure 7-2. From the figure 7-2, it shows the experiments 1 and 3 with the linear function giving a little better average performance with the experiments 2 and 4 with the nonlinear function. However, the best fitness of the 4 experiments is nearly one. Also, there is a no difference in the fitness between patient-centred (Experiment 1, 2) and randomly k folds assignment (Experiment 3, 4) for the cross validation. This indicates that the CGP networks detecting microcalcification is not dependent on a patient's examples.
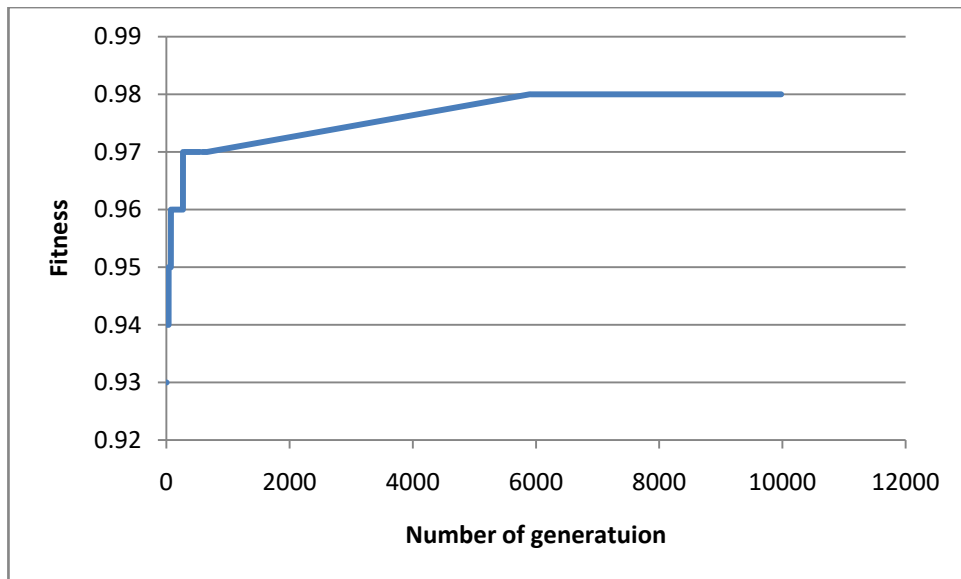
Figure 7-2Average fitness values training the algorithm on the detection of microcalcifications

Table 7-2 Training algorithm on detection of microcalcifications

Best fitness	Average fitness	Worst fitness
0.98	0.9751	0.94

Figure 7-2 shows the development of the average fitness values over the training stage of 10000 generations. The average fitness increases when generations increase and reach to peak generation at 6101 and then increases smoothly on the following generation. The maximum average fitness is 0.9817, however there is still a small difference between the best fitness and the perfect '1' fitness. One explanation for this would be the fact that each chromosome has detected that there are some microcalcifications pixels on that part. However, it cannot generate the exact numbers of microcalcification pixels as WPV in that part and therefore, an error that affects the fitness value has been caused. Although the fitness in train stage is not 'perfect', the high fitness value still indicates that the MCGP gives a good performance.

7.1.3   Testing Stage

The resulting MCGP algorithm from training stage is then tested using the images from the partitioned testing image sets shown in Table 7-1 and 7-2.
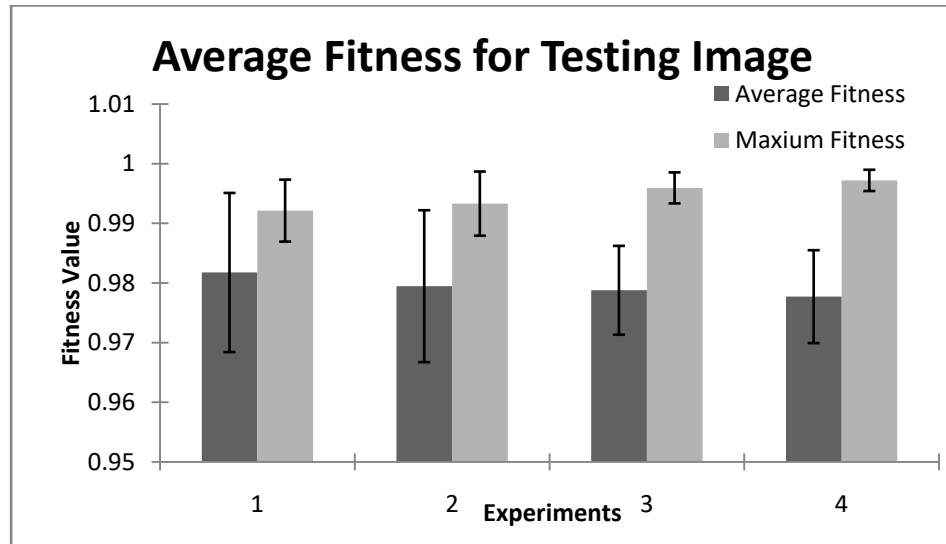
Figure 7-3 Average Fitness for Testing Images

The test result for the respective experiments is shown in Figure 7-3. It confirms the analysis in the training experiments. To give a fair comparison with the networks, the linear fitness function is applied to evaluate the data from the CGP networks which were trained using nonlinear fitness function in experiment 2 and 4 from Table 7-1. Therefore there is little tiny difference in the test performance between linear fitness and nonlinear trained network and suggests that the alternative schemes offer little benefit over each other. It should be noted that although CGP networks where trained using non-linear fitness functions in experiments 2 and 4, the linear fitness function was used to evaluate their performance on the test set of data, to enable a fair comparison with those networks evolved in experiments 1 and 3.

An example of the microcalcifications detected by an evolved network is given in Figure 7-5, which resulted from presenting the test image illustrated in Figure 7-4.

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	41	15	0
0	0	0	0	0	0	0	0	0	0	0	0	0	63	38	0
0	0	0	0	0	0	0	0	0	0	0	0	0	42	36	0
0	0	0	0	0	0	0	0	0	0	0	0	0	44	5	0
0	0	0	0	0	0	0	0	0	0	0	25	48	64	22	0
0	0	0	0	0	0	0	0	0	0	0	8	57	58	4	0
0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 7-4 An example of ROI with WPV for each parts

3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	8	8	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	3	1	0	0	0	0	0	0	0	0	0	0	41	15	0
3	0	1	0	0	0	0	0	0	0	0	0	0	63	38	0
1	1	0	0	0	0	0	0	0	0	0	0	0	42	36	0
0	0	1	0	0	0	0	0	0	0	0	0	0	44	5	0
1	0	1	0	0	0	0	0	0	0	0	25	48	64	22	0
0	0	0	0	0	0	0	0	0	0	0	8	57	58	4	0
0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0
0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 7-5 Corresponding MCGP outputs on detection of microcalcifications

Figure 7-4 shows that there is one microcalcification in a grey shadow in ROI, it is WPV with density of microcalcification pixels supplied by radiologists. Figure 6-5

shows the outputs from the MCGP network. The grey parts are the MCGP algorithms detecting the parts which contain microcalcifications. There are generally two areas of the microcalcifications in the result, one is with a small value and one is with a large value. The area contains a large output value on a similar area of microcalcification stays on the original ROI. It could tell visually that MCGP delivered a relatively good result to detect microcalcifications, however it is still processed some small non expected value. The major reason might be that, there are too many breast tissues and less microcalcification pixels, since each chromosome CGP network only deals with one part of the image.

7.1.4   Evaluation

In order to get better performance, a threshold is used to alter a higher scale to assigning the meaning of WPV to be upper level and the meaning of breast tissue to the lower level. After each chromosome CGP network produced output, it will be rescaled firstly to a WPV range from 0 to 64, then this output value will be added by a threshold. The threshold is interpreted as an indication of microcalcifications. For the outputs of MCGP, if the outputs' values are equal to or above the threshold, the outputs' values will remain same, else 0.

A set of threshold with valid values [0, 2, 4, 6, 8] are applied to the output of the multi-chromosomes CGP network for an image. Threshold=0 represents that no threshold is applied to the output of CGP. The maximum threshold value is 8 that is because the original WPV for benign microcalcifications is in part a relatively small number and the outputs from MCGP network; this ensures that for every image in 4 experiments there is at least one part in the CGP output that has a non-zero value.

An example of the outputs of MCGP network applied to the thresholds is shown in Figure 7-6.The input ROI of this example is shown on Figure 7-4. The grey parts are those with their outputs after threshold and in this particular case.

3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	8	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	3	0	0	0	0	0	0	0	0	0	0	0	0	41	15	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	63	38	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	42	36	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	44	5	0
0	0	0	0	0	0	0	0	0	0	0	0	25	48	64	22	0
0	0	0	0	0	0	0	0	0	0	0	0	8	57	58	4	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(a) MCGP outputs after threshold '2'

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	8	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	41	15	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	63	38	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	42	36	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	44	5	0
0	0	0	0	0	0	0	0	0	0	0	0	25	48	64	22	0
0	0	0	0	0	0	0	0	0	0	0	0	8	57	58	4	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(b) MCGP outputs after threshold '4'

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	41	15	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	63	38	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	42	36	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	44	0	0
0	0	0	0	0	0	0	0	0	0	0	0	25	48	64	22	0
0	0	0	0	0	0	0	0	0	0	0	0	0	57	58	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(c) MCGP outputs after threshold '6' and '8'

Figure 7-6 MCGP outputs after threshold applied for the corresponding ROI from Figure 6-5

The most common method to evaluate the detection is ROC curve, however it is not used in this experiment since there are a lot of 0s indicating the non- microcalcifications in the image, and this gives a large number of negative cases which leads to the false positive rate being too small to form a curve. For the stage of detection of microcalcifications, the locations of microcalcifications are important and it is the factor to determine the performance of the network, therefore the centre of gravity (COG) is applied. COG is a geometric property of any objects; it is the average location of the weight of an object. COG is an important indicator for the centre point of an irregular shape with a different density; it is commonly used in statistics to design the static structures. In this case, a COG is calculated to locate the centre of a microcalcifications area, if I compare the location of COG of one microcalcifications area of the image, we would say around this COG, there is an area of microcalcifications.

In this experiment, there are 18 benign images used for detecting of microcalcifications, 16 images have only one unique area of microcalcifications individually, 2 images have 2 areas of microcalcifications. Therefore, for the images that have 2 areas of microcalcifications, the image needs to be divided into 2 equally sized non-overlapped parts and each part has one unique area of microcalcifications, then calculate the COG for each microcalcifications area in one image part. For example, Figure 6-6 above shows an image with 2 areas of microcalcifications, and this image could be divided into an area with the rows between rows [1, 8] and an area with rows between [9, 16], the general form of CGP network of this image is shown on Figure 6-7.
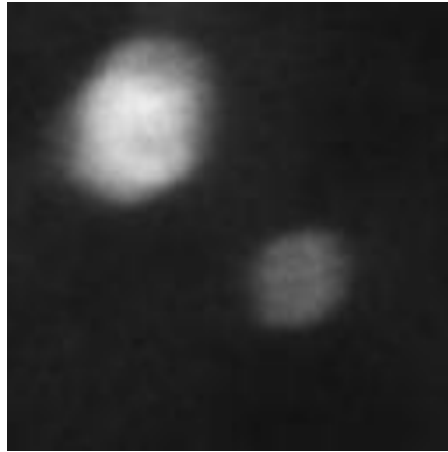
Figure 7-7 an image with 2 areas of microcalcifications

0	0	4	30	13	22	7	0	0	0	0	0	0	0	0	0
0	5	55	64	64	64	60	9	0	0	0	0	0	0	0	0
0	29	64	64	64	64	64	49	0	0	0	0	0	0	0	0
0	44	64	64	64	64	64	64	0	0	0	0	0	0	0	0
0	55	64	64	64	64	64	59	0	0	0	0	0	0	0	0
0	54	64	64	64	64	64	32	0	0	0	0	0	0	0	0
0	30	63	64	64	64	52	4	0	0	0	0	0	0	0	0
0	0	7	32	30	15	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	19	44	53	0	0	0	0	0
0	0	0	0	0	0	0	0	10	64	64	64	5	0	0	0
0	0	0	0	0	0	0	0	20	64	64	50	0	0	0	0
0	0	0	0	0	0	0	0	3	32	33	7	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 7-8 general form of CGP network of an original image with 2 areas of microcalcifications has been divided into 2 parts

The COG is used here to determine the location of a single microcalcifications distribution centre. It treats the distribution of a single microcalcification as the function only of the weight and the location the microcalcifications. The weight of each part could be seen as the white pixel number on this part. For a 2 dimensional image which has 256 equally sized non-overlapping parts which could be seen as a grid of 16 by 16, the centre of gravity COG (COG_x, COG_y) of an area of microcalcifications is defined as the average of the sum of the co-ordinates values (r_{xi}, r_{yi}). For each part i who has microcalcifications, weighted by its rescaled output for part i, since the output for each

163

part of ROI is in a range of [0,255], to make it into WPV range of [0,64], for COG of

the microcalcifications of input image, m_i is the WPV :

$$m_i = \frac{output_i + 1}{4}$$ **Equation 7-3**

$$COG_x = \frac{\sum_1^{256} r_{xi} m_i}{\sum_1^{256} m_i} \qquad COG_y = \frac{\sum_1^{256} r_{yi} m_i}{\sum_1^{256} m_i}$$ **Equation 7-4**

After the COG of the input image (COG_x, COG_y) and the COG of the output of the

CGP network (COG'_x, COG'_y) have been generated, COG differences shows as:

$$COG\ Differences = \sqrt{(COGx - COG'x)^2 + (COGy - COG'y)^2}$$

Equation 7-5

However, COG calculation has the limitations of calculating Location: (i) it could only

locate the centre of the microcalcifications area in the image without giving the exact

size of the area. To calculate the area of determining one uniform area of

microcalcifications in one image. To overcome these, binary number '0' and '1' are used

as weight. For the input image and CGP the microcalcifications further investigation is

needed. (ii) One COG is only output, '0' is indicating non-microcalcifications and '1' is

indicating microcalcifications, this could be used to replace the white pixel values on

each part. For the input image, if the white pixel value of each part is not zero, and then

replace the white pixel number to '1' to indicate the microcalcifications, others are '0'.

For the output of MCGP, if the output value is not zero, then m_i for COG become '1',

otherwise m_i stays at zero. For calculating the COG for the input ROI, if WPV is not

zero, m_i is '1', others is '0'. For the images have 2 areas of microcalcifications as Figure

6-8, the total image has been divided into two equal areas and calculate COG for

individual area then add them together to get the average.

Then the difference of COG between the MCGP network's input and output is

calculated by Equation 7-5, it is used to evaluate the ideal result for the COG difference

164

which is equal to 0. Since COG is measured under the scale of parts if the COG difference is smaller than 1, it shows the centre of gravity for both input and output are located at the same part in the image. If the COG difference is in a range of 1 to 4, it shows that the MCGP detected microcalcifications on the surrounding area of the original microcalcifications; if COG difference is larger than 4, it shows a random detection. COG difference will be used to evaluate the performance of algorithm to detect microcalcifications.
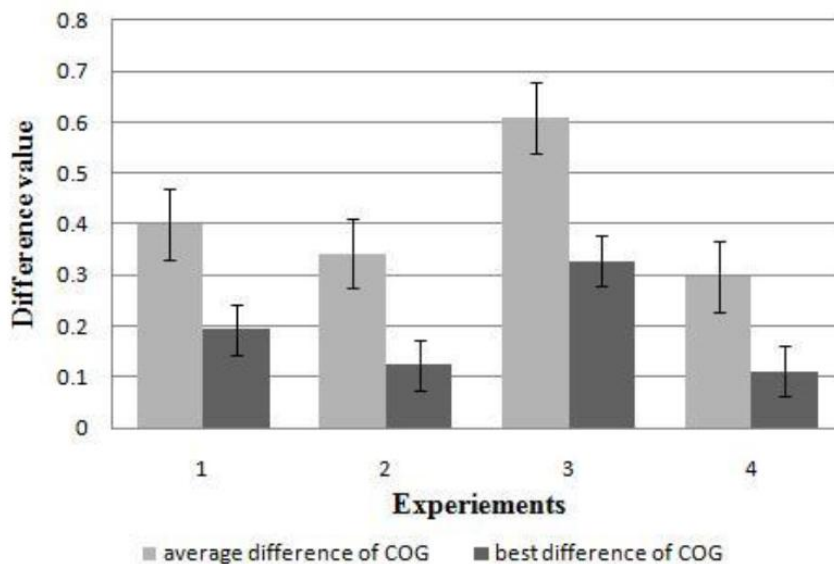
Figure 7-9 Average and best difference values for centre of gravity

The average differences of COG results of the 4 experiments (discussed in Table 6-1) are in Figure 7-9. It shows a generally small difference of the centre of gravity between the input and the output of MCGP network, but a better performance using the nonlinear fitness function than the linear function. It is apparent, however, that there is little difference in COG between patient-centered and distributed k-folds assignment for the cross-validation exercise. This shows that the MCGP networks have evolved a generic

165

microcalcifications detector that is not dependent on a particular set of patients. From the results, it could be said that the multi chromosomes of CGP network could detect the microcalcifications in mammograms successfully.
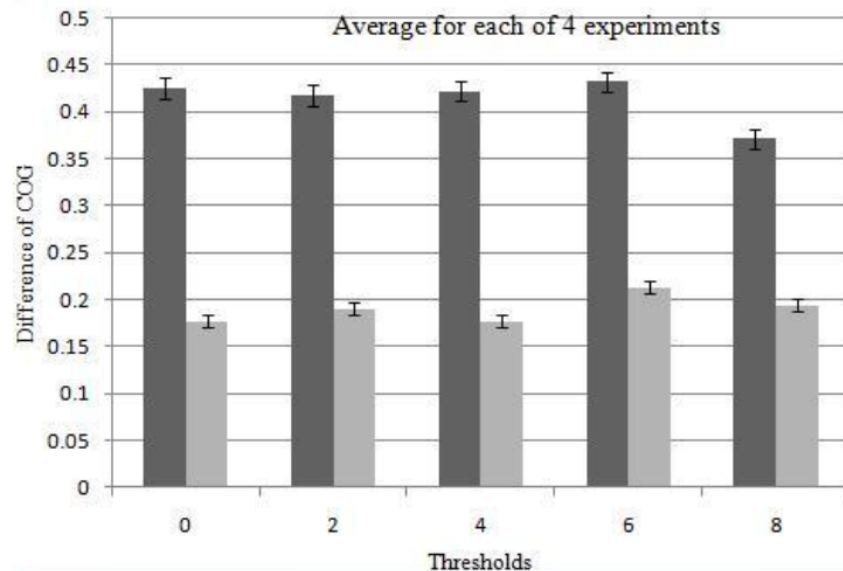
Figure 7-10 The average result of 4 experiments for different values of threshold

The average of the COG differences results for the different thresholds in Figure 6-10, shows a general low difference of COG for the thresholds [0, 2, 4, 6, 8]. The dark grey bar is the average COG differences after threshold have been applied; the light grey bar is for the best COG differences for applying threshold, therefore the weight for COG calculation is scaled to 1. Threshold 0 gives the results when calculating the COG differences which does not involve a threshold, and for any thresholds over 8. Some of the test images or the outputs of the network could not find COG differences, that is due to the output values are too small to apply to any large threshold. It is apparent; there is a little difference for the results of threshold 0 to 6 and then a better performance 8. This

can be interpreted positively; it indicates that the MCGP network outputs have small differences of COG white pixel values on each part to the relative image input.

This stage has described a novel multi-chromosomes CGP network applied to the detection of microcalcifications in mammograms, as an important stage towards using the same multiple chromosomes CGP network evolutionary algorithm for the classification of mammograms. The results presented here have demonstrated that the method is very encouraging.

7.2 Classification of microcalcifications

The MCGP network is applied by classification of microcalcifications in this stage Detection of microcalcifications could be seen as a binary classification as '0' for breast tissue and '1' for malignant microcalcifications. Besides, the aim of detection of microcalcifications, it is an important stage towards to the aim of the entire work. Therefore, the developed MCGP algorithm for detection of microcalcifications is applied to classify microcalcifications. This MCGP network will be trained to classify microcalcifications on mammograms to automatic diagnosis microcalcifications on mammograms.

A number of changes are considered, with the ranging starting from input images. The total of 31 images from the dataset are applied in the MCGP network. The training and testing images have been introduced in Chapter 5 after k fold cross validations. There are still 4 different experiments based on catalogues by patients and randomly chosen for test images, even if it has shown that there were no differences between patient-centres. (Experiment 1, 2) and randomly k folds assignment (Experiment 3, 4) for the

cross validation in stage of detection of microcalcifications. In these dataset, both malignant and benign images are used to train the system, and the training and testing are followed by Table 6-3 and 6-4.

From Chapter 2, the features of malignant microcalcifications are small and hard to locate. Only lets MCGP learning the features on malignancy microcalcifications, in order to avoiding any misdiagnosis. As mentioned in Chapter 6, after the 31 ROI images are logically segmented into 256 parts, the status of each part labelled as either being benign or malignant according to the radiologists. Both benign and malignant ROI's have WPVs for microcalcifications have different meanings as mentioned in last chapter. In the malignant image, all the microcalcifications which have been indicated by the radiologist on the images are malignant, therefore the white pixel values which are the non-zero intensity pixels for each part of the malignant. However for benign images in classification stage, the WPV values are no longer staying the same, they have been rescaled. Instead of having WPV values for benign microcalcifications, they all have '0's as the representation that there are no malignancy microcalcifications in benign images. Therefore, an example of representation of the intensity of microcalcifications of ROI will look as follows:

Left grid (benign ROI), 16×16 all zeros:

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Right grid (malignant ROI), 16×16:

0	0	4	30	13	22	7	0	0	0	0	0	0	0	0	0
0	5	55	64	64	64	60	9	0	0	0	0	0	0	0	0
0	29	64	64	64	64	64	49	0	0	0	0	0	0	0	0
0	44	64	64	64	64	64	64	0	0	0	0	0	0	0	0
0	55	64	64	64	64	64	59	0	0	0	0	0	0	0	0
0	54	64	64	64	64	64	32	0	0	0	0	0	0	0	0
0	30	63	64	64	64	52	4	0	0	0	0	0	0	0	0
0	0	7	32	30	15	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	19	44	53	0	0	0	0
0	0	0	0	0	0	0	0	10	64	64	64	5	0	0	0
0	0	0	0	0	0	0	0	20	64	64	50	0	0	0	0
0	0	0	0	0	0	0	0	3	32	33	7	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 7-11 an example of representation of non zero intensity pixel counts for benign(Left) and malignant(Right) ROIs

In figure 6-11, the benign case is on the left, it contains all zero WPV parts since there are no cancerous microcalcifications pixels. The malignancy ROI contains WPV in parts is on the right, there are two areas of malignant microcalcifications. The WPV values are relatively large numbers and means a high intensity of pixels containing malignancy microcalcification information.

The output for each single chromosome will remind as previous experiment stage, which is in a range of 0 to 255. Therefore the outputs of an entire MCGP network will be a mapping of 256 parts, in each part , the output has a range of 0 to 255. Since the input ROIs contain both benign and malignant pixels, and the standard WPV for fitness function has altered to suit on both benign and malignant images, therefore the 256 MCGP outputs in a good performance could be predicted :

- The results of 256 outputs of individual chromosomes expected   for Benign mammograms:

  1). All zeros outputs from the 256 chromosomes;

  2) Few low values of outputs from the 256 chromosomes.

- The results of 256 outputs of individual chromosomes expected for malignancy mammograms:

1). A number of large value of outputs from the 256 chromosomes;

2). A large number of non-zero outputs from the 256 chromosomes stay on or surrounding on the corresponding parts which are supposed to contain malignant microcalcifications the relative 256 parts.

This ends that there will be 3 possible results expected from the MCGP classification algorithm: the malignant image, the benign image, and the image with the likelihood of malignancy. The larger number one chromosome output is, the more numbers of malignancy microcalcifications pixels in one part of ROI has classified. If the output of each chromosome has a big value, it could be represented as a malignancy part, if it has a small or zero value, it could be represented as a benign part. Therefore, a threshold (T1) is applied to assign the meaning of benignity to the lower level and the meaning of malignancy to the higher level.

For the training stage, threshold is applied on each chromosome CGP network in order to filter the small value before fitness calculated:

```
if ( CGP output > threshold(T1))
      CGP output =  CGP output;
Else
      CGP output = 0;
```

During the process of training the algorithm to perform detection of microcalcifications, the number of outputs of MCGP network is equal to the number of inputs. Since this stage is for classification of microcalcifications, a binary classification with a value either '0' or '1' indicating a malignancy or a benignity is the

result expected. As the whole MCGP network was designed to deal with a range of values from '0' to '255' in order to process pixel values of a 8 bit depth, an additional mechanism was needed to convert this output range to a single classification integer. In order to achieve this, a fixed threshold was introduced. The MCGP network output value was then compared to the threshold (T2) to determine the classification.

After each chromosome CGP output has been rescaled, number of malignant parts is by counting the number of the parts which contain outputs of one chromosome bigger than '0' to generate an integer, threshold T2 is applied here to give a binary classification, which '0' indicate benign and '1' indicate malignant :

If number of Malignant parts$\geq$ Threshold T2

    then the input image is malignant

    Output of MCGP = 1;

else

    then the input image is Benign

    Output of MCGP network = 0;

Fitness is calculated by comparing the binary output and the diagnosed information which is supplied by a radiologist.

If   Output of MCGP network =1 && the input image is a malignant image,

    Then fitness of chromosome +1;

If   Output of MCGP network =0 && the input image is a benign image,

    Then fitness of chromosome +1;

Else

    Fitness +0;

The fitness of MCGP network $= \frac{1}{256} \sum_{i=255}^{0}$ fitness of chromosome i $\qquad$ Equation 7-6

The perfect fitness of MCGP network is 1, it means all single chromosome CGP networks have classified microcalcifications successfully in 256 parts.

7.2.1 Training stage of the Algorithm

The experiment uses parameters values based on Chapter 6 as described in table 6-6. All four experiments described by table 6-5 were carried out with the same parameters.

The experiment is based on four experiments of k fold validation. Since it is a large scale of results represented for the same experiment, the fitness results will be calculated in average and the best or the worse fitness results also showed. The training results for the algorithm on the classification of microcalcifications shows in table 7-3:

Table 7-3 Fitness results on the classification of microcalcifications

Best fitness	Average fitness	Worst fitness
0.99	0.98	0.7
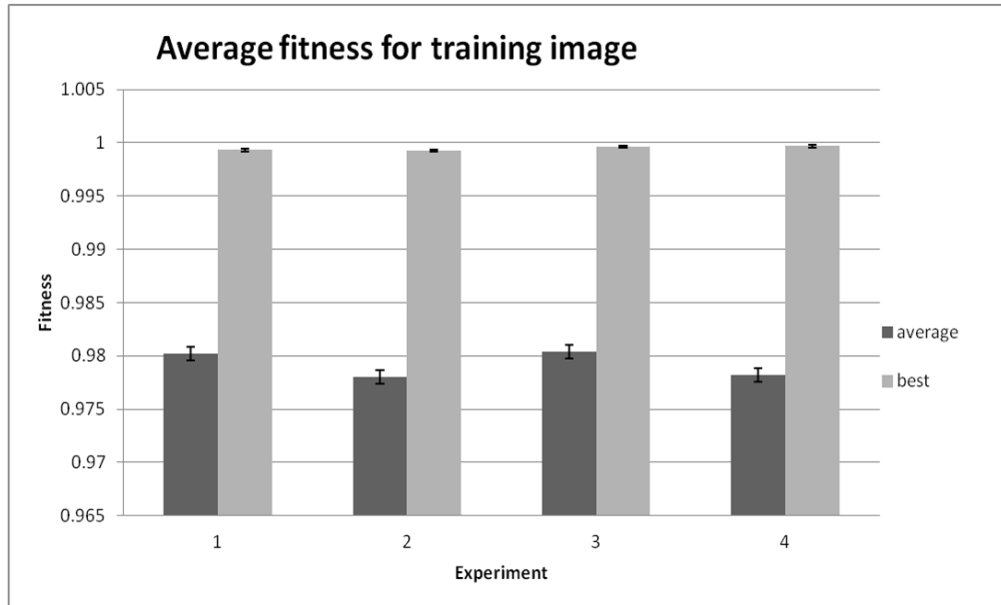
7-12 the average fitness for each experiment

As Figure 7-10 and Table 7-3 show, the MCGP network produce results showing good performance in training stage. It shows that the best fitness scores are almost 1, with the worst fitness scores being 0.7. The dark grey bar is the average fitness of all training results in experiment1, with a ±1 standard deviation. The light grey bar represents the best training fitness result with a standard deviation ±1. These mean that the MCGP network has been fully trained to achieve the best fitness score as 1. For the average fitness for all four experiments being nearly the same, there is a no difference in the fitness between patient-centred (Experiment 1, 2) and randomly k folds assignment (Experiment 3, 4) for the cross validation. This indicates that the MCGP networks classifying microcalcifications is not dependent on a patient's examples. From Table 7-3, the worst fitness is 0.7 and best fitness is 0.99. It shows the range of improvement is about 3% and it could be said that the algorithm is mainly improving the worst fitness. One explanation for this would be the fact that there are most parts of the ROIs showing benign tissues so the most algorithms in training are mainly trained with the benign

image parts. The chromosome CGP trained with these benign parts is therefore a good classifier for benign tissue. The development of average fitness against generations for all 4 experiments have been shown on Figure 7-13 and shows the ability of MCGP algorithm in learning to classify the microcalcifications.
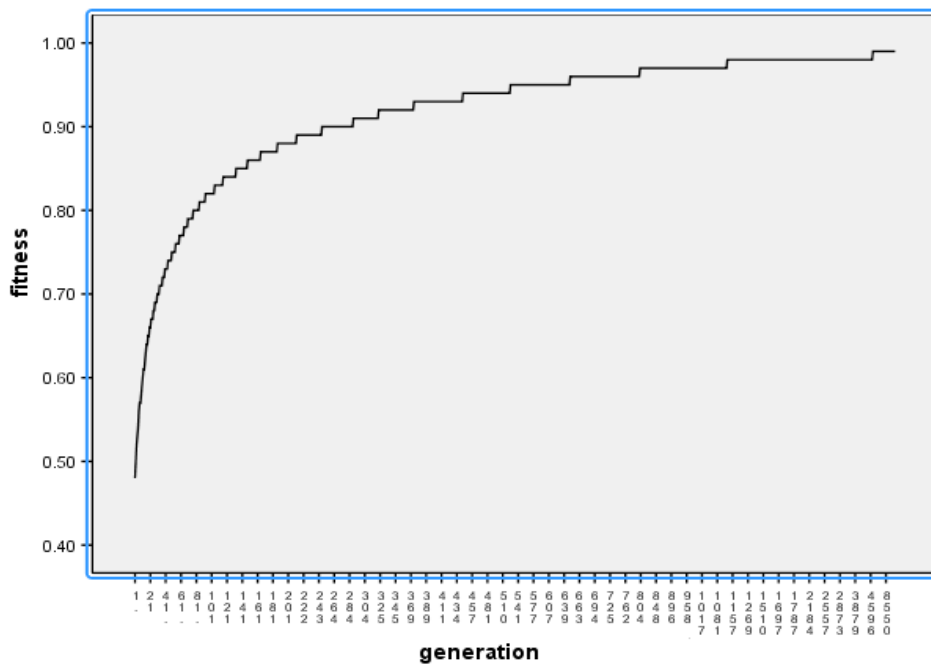
7-13 The average fitness values training the algorithm on the classification of microcalcifications

7.2.2   Testing Stage

The resulting Algorithm was then tested using images from Table 6-3 and 6-4 from the LLNL database that were not used for training the algorithm. The fitness testing results shows from Table 7-4 and the average fitness for 4 experiments (described in Charpter 7) has been shown in Figure 7-10.
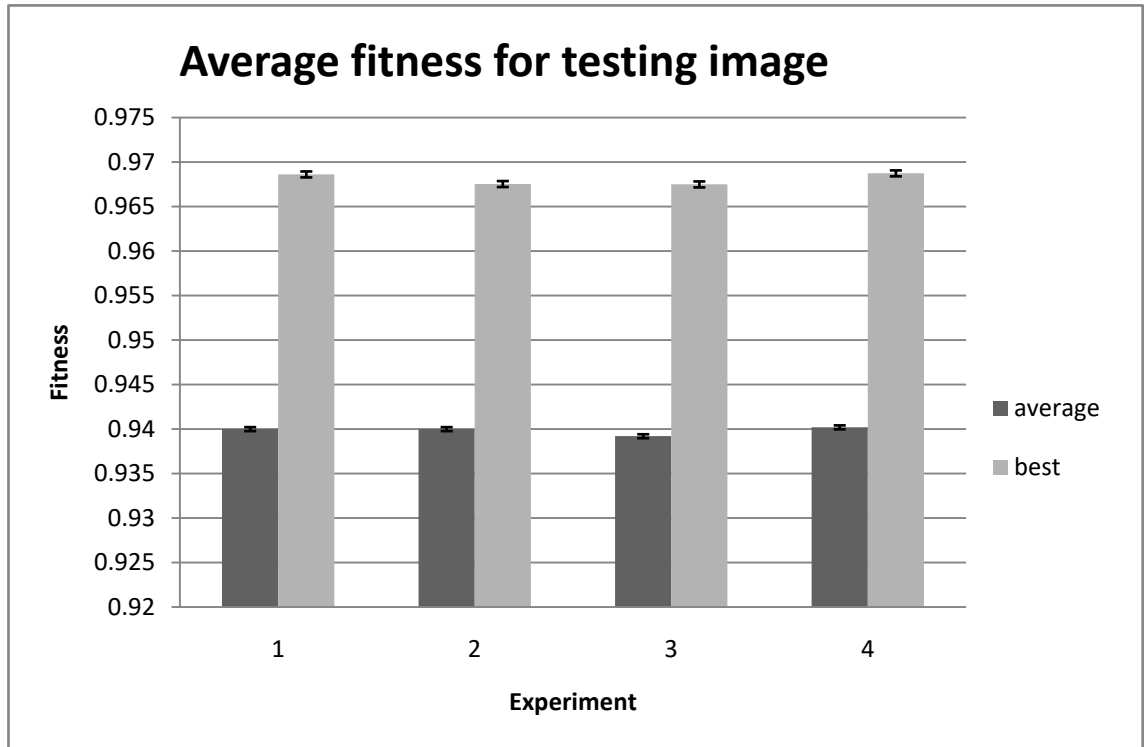
Figure 7-14Average and best fitness for 4 experiments of testing results

The testing phase shows slightly lower fitness values than the training phase, however it is still in good value, with the average fitness for these 4 experiments as above 0.935 and the best fitness value has reached 0.9683 which is really near to 1.

Table 7-4 Average results of 4 experiments of testing MCGP network

Experiments	Confusion Matrix (Average value)				Statistics				
	Bb	Bm	Mb	Mm	TP	TN	FP	FN	MCC
Experiment1	5	13	13	0	0	0.28	0.72	1	-0.75
Experiment2	5	13	13	0	0	0.28	0.72	1	-0.75
Experiment3	5	13	13	0	0	0.28	0.72	1	-0.75

Experiment14	5	13	13	0	0	0.28	0.72	1	-0.75

Table 7-4 shows the average results of a further breakdown for the best chromosomes from each MCGP network and contains both the confusion matrices as well as additional statistics measures. In confusion matrix, the first character B and M is indicating the actual classification and the second character b and m is indicating the predicted classification. These numbers link the value of TP, TN, FP, FN and values of MCC (Chapter 6) which is used to measure of performance of system are given.

Comparing the values of Table 7-4, FN rates are 1s; this means that all of the images were classified incorrectly. The values of MCCs stay as negative values indicated the faulty prediction have occurred.

In order to get the better performance of the system, a closer look of the outputs of 256 chromosome of a MCGP network for an 256 parts of testing image have been formed. This is due to the property of MCGP network which contains individual non-overlapped sub-networks. An example is shown in Figure 7-15 with one actual malignant testing image and Figure 7-16 for an actual benign images. Both of the figures contain the WPVs in 256 parts indicated by radiologist on the top of the figure, then the output of MCGP network with 256 individual chromosome outputs, the bottom is the differences between the WPVs and outputs in the corresponding part.

```
#White pixel values for image bbrccM1.tst
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0   16    2    0    0    0    0    0    0    0    0    0    0    0    0    0
    0   49   30    0    0    0    0    0    0    0    0    0    0    0    0    0
    0   11    8    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0

#CGP output values for best genotype on image bbrccM1.tst
    0    0    0    1    0    0    0    0    0    0    0    0    0    0    0    3
    0    0   32    0    3    0    0    0    0    0    1    6    1    0    1    0
    0    0    0    0    0    0    0    1    0    0    2    0    0    0    0    4
    0    0    1   32   28    0    1    1    0    0    6    0    0    0    1    0
    2   31    0    0    4    5    1    1    0    1    0    2    0    0    0    0
    0    0    3    1    0    0    2    2    0    0    2    0    0    1    0    0
    0    0    0    0    4    0    0    0    2    0    0    0    1    0    0    0
    0    0    0    0    0    0    0    1    0    0    0    0    0    0    0    2
    1    0    0    0    0    0    0    0    0    4    0    0    1    0    0    0
    0    0    8    1    3    0    0    0    0    0    0    0    0    0    0    0
    0    2    4    1    4    0    2    0    0    0    0    0    0    0    1    0
    1    2    0    0    0    3    2    0    0    0    0    0    2    1    0    0
    3    3    0    3    4    0    2    4    0    0    0    0    0    0    0    8
    0    0    4    2    0    0    4    0    1   11    1    0    0    0    1    0
    0    0    1    0    1    0    0    2    0    4    3    0    0    0    0    0
    0    0    0    1    0    3    3    0    0    1    0    8    1    1    0    6

#Difference in CGP output values from true white pixel values
#in all parts using the best genotype on image bbrccM1.tst
    0    0    0   -1    0    0    0    0    0    0    0    0    0    0    0   -3
    0    0  -32    0   -3    0    0    0    0    0   -1   -6   -1    0   -1    0
    0    0    0    0    0    0    0   -1    0    0   -2    0    0    0    0   -4
    0    0   -1  -32  -28    0   -1   -1    0    0   -6    0    0    0   -1    0
   -2  -31    0    0   -4   -5   -1   -1    0   -1    0   -2    0    0    0    0
    0    0   -3   -1    0    0   -2   -2    0    0   -2    0    0   -1    0    0
    0    0    0    0   -4    0    0    0   -2    0    0    0   -1    0    0    0
    0    0    0    0    0    0    0   -1    0    0    0    0    0    0    0   -2
   -1   16    2    0    0    0    0    0    0   -4    0    0   -1    0    0    0
    0   49   22   -1   -3    0    0    0    0    0    0    0    0    0    0    0
    0    9    4   -1   -4    0   -2    0    0    0    0    0    0    0   -1    0
   -1   -2    0    0    0   -3   -2    0    0    0    0    0   -2   -1    0    0
   -3   -3    0   -3   -4    0   -2   -4    0    0    0    0    0    0    0   -8
    0    0   -4   -2    0    0   -4    0   -1  -11   -1    0    0    0   -1    0
    0    0   -1    0   -1    0    0   -2    0   -4   -3    0    0    0    0    0
    0    0    0   -1    0   -3   -3    0    0   -1    0   -8   -1   -1    0   -6
```

Figure 7-15Testing output results of MCGP for a Malignant ROI

```
#White pixel values for image allccB3.tst
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0

#CGP output values for best genotype on image allccB3.tst
   0   0   1   0   0   0   1   0   0   0   0   8   0   0   0   7
   0   1   0   0   0   0   0   0   0   0   0   0   7   0   0  15
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   7
   0   0   0   0   1   0   0   6   1   0   8   0   0   0   0   0
   0   0   0   0   0  15   1   0   0   0   0   1   0   0   7   0
   0   0   0   0   0   0   9   0   0   0   0   0   8   0   0   1
   0   0   1   0   0   0   0   5   0   0   0   0   0   0   0   0
   0   0   0   7   0   0   0   0   0   1   0   0   0   0   0   0
   0   0   0   0   0   0   0   2  14   0   0   0   1   0   0   1
   0   8   0   0   8   0   0   0   0   0   0   0   0   0   0   0
   0   0   0  15   0  14   0   0   0   0   0   8   8   0   0   1
   0   1   0   7   0   6   1   0   0   0   0   0   0   0   0   0
   8   0   8   1   1   0   0   0   0   0   0   0   0   0   0   0
   0   0   8   0   0   0   3   1   0   1   0   8   0   0   0   0
   0   0   0   0   0   0   0   0   1   3   0   0   0   0   0   0
   0   0   0   0   1   0   0   0   0   1   0   0   0   0   0   0

#Difference in CGP output values from true white pixel values
#in all parts using the best genotype on image allccB3.tst
   0   0  -1   0   0   0  -1   0   0   0   0  -8   0   0   0  -7
   0  -1   0   0   0   0   0   0   0   0   0   0  -7   0   0 -15
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0  -7
   0   0   0   0  -1   0   0  -6  -1   0  -8   0   0   0   0   0
   0   0   0   0   0 -15  -1   0   0   0   0  -1   0   0  -7   0
   0   0   0   0   0   0  -9   0   0   0   0   0  -8   0   0  -1
   0   0  -1   0   0   0   0  -5   0   0   0   0   0   0   0   0
   0   0   0  -7   0   0   0   0   0  -1   0   0   0   0   0   0
   0   0   0   0   0   0   0  -2 -14   0   0   0  -1   0   0  -1
   0  -8   0   0  -8   0   0   0   0   0   0   0   0   0   0   0
   0   0   0 -15   0 -14   0   0   0   0   0  -8  -8   0   0  -1
   0  -1   0  -7   0  -6  -1   0   0   0   0   0   0   0   0   0
  -8   0  -8  -1  -1   0   0   0   0   0   0   0   0   0   0   0
   0   0  -8   0   0   0  -3  -1   0  -1   0  -8   0   0   0   0
   0   0   0   0   0   0   0   0  -1  -3   0   0   0   0   0   0
   0   0   0   0  -1   0   0   0   0  -1   0   0   0   0   0   0
```

Figure 7-16Test output results of MCGP for a benign ROI

Even the fitness for the testing phase is high; however it could be seen that the outputs of each chromosome CGP network produced small values of outputs for malignant image and relatively big number for benign images. These results are completely

178

opposite to the expected MCGP output results. To improve the performance of the system in testing stage, a threshold in a range of [0, 64] has to applied to assign benignity and malignancy. The value of threshold could not be so big otherwise there will be FN classification; however it could not be too small which would lead to FP classification. A middle range of [16, 32] threshold is used to apply on the assignment. Due to the outputs for the malignancy testing images, any value below 16 has inferred a misdiagnosis for the MCGP network. One example has been shown in Figure 7-18. Figure 7-19 is the outputs of MCGP for a malignancy image as input (also as Figure 7-15), it could be noticed that after a threshold valued at 32 is applied, there is no non zero outputs in MCGP output network. It shows the MCGP network give a FN when threshold = 32. When threshold = 16, almost all the output parts become 0 and the non-zero output parts are actually a benign part of the corresponding input WPVs. The same as the MCGP outputs from benign image is applied by a threshold. It shows that after a threshold=16 has been applied, there is still a few non zero parts left on the MCGP outputs. Table 7-5 indicated that the MCGP network would give a wrong diagnosis in classifying benign microcalcifications.

Table 7-5 Results for voting with threshold on the output of each chromosome network

	Confusion Matrix (Average value)				Statistics				
Experiments	Bb	Bm	Mb	Mm	TP	TN	FP	FN	MCC
Threshold 16	0	18	13	0	0	0	1	1	-1
Threshold 32	0	18	13	0	0	0	1	1	-1

```
#White pixel values for image bbrccM1.tst
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0  16   2   0   0   0   0   0   0   0   0   0   0   0   0   0
 0  49  30   0   0   0   0   0   0   0   0   0   0   0   0   0
 0  11   8   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0

#CGP output values for best genotype on image bbrccM1.tst
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0  28   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
```

Figure 7-17 The corresponding outputs of MCGP network to Figure 7-15 is applied by threshold=16

```
#White pixel values for image bbrccM1.tst
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0  16   2   0   0   0   0   0   0   0   0   0   0   0   0   0
  0  49  30   0   0   0   0   0   0   0   0   0   0   0   0   0
  0  11   8   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0

#CGP output values for best genotype on image bbrccM1.tst
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
```

Figure 7-18 The corresponding outputs of MCGP network to Figure 7-15 is applied by threshold=32

7.2.3   Further experimentation

7.2.3.1   Training and Testing with a different threshold and bias strategy

Considering the malignant microcalcifications are very small, most image parts will contain benign tissue surroundings for the classification when dividing the image into image parts. During training this leads to one chromosome CGP network that might not have had to deal with the classification of a malignant part at all. This means that some particular chromosome CGP networks will be specialised in classifying benign image properties while only few chromosome CGP networks are trained to recognise a malignant microcalcifications part. This is the problem that might arise from that and is the inability to classify malignancy during the test phase. In order to overcome this

problem, in the training stage a bias has been used to deal with the issue that there are more benign parts than malignant parts in the training images. Since the bias added on, the fitness function has altered to give the best training. The Threshold and bias algorithm is described below and applied in the training stage and the testing stage remained same.

For the training stage, threshold is applied on each chromosome CGP network in order to filter the small value before fitness calculated:

if (CGP output > threshold)
     CGP output = CGP output;
Else
    CGP output = 0;

Then bias is applied on fitness function to reduce the weight of zero WPV parts when the network is trained.

The fitness function :

If (CGP output > 0 && number_white_pixels >0)
    Fitness of chromosome = 1;
Else if (CGP output = 0 && number_white_pixels =0)
    Fitness of chromosome = 1/bias;
Else
    Fitness = 0.

There are several experiments which have been carried out to evaluate the combination value of bias and threshold. Both the threshold and the bias are given a range of [1, 64], '1' means that there is no bias and threshold applied in the system, '64' is a middle-value to test whether the improvement could lead to a better performance. After trials runs, the best average fitness occurs when bias= 30, threshold = 30, the MCGP network training fitness =0.502. In the testing stage, the performance results as follows:

Table 7-6 Results for applying bias on training system

Experiments	Confusion Matrix (Average value)				Statistics				
	Bb	Bm	Mb	Mm	TP	TN	FP	FN	MCC
Experiment1	1	17	13	0	0	0.17	0.83	1	-0.84
Experiment2	0	18	13	0	0	0	1	1	-1
Experiment3	1	17	13	0	0	0.17	0.83	1	-0.84
Experiment14	0	18	13	0	0	0	1	1	-1

From the numbers given in Figure 7-6 the algorithm with adding bias and changing fitness still cannot provide a positive performance of classifying microcalcifications. With the FN values equal to 1, it meant that the malignancy is always failed to detected, the new approach by adding bias and changing thresholds could not help to provide better performance.

The example of 256 chromosomes outputs of testing image results are shown in Figure 7-19.

```
#White pixel values for image bwrccM1.tst
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0   22    9    0    0    0    0    0    0    0    0    0    0
   0    0    0    0   17    7    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    4    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0


#CGP output values after threshold for best genotype on image bwrccM1.tst
   0    0    0    0   31    0    0    0    0    0   33    0    0    0   34    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0   55    0    0   57    0   32   50    0    0    0    0    0    0    0
   0    0    0    0    0    0    0   62    0    0    0    0    0    0    0    0
  46   35    0    0    0   32   49    0    0    0    0    0    0    0    0    0
   0    0    0   36    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0   54    0    0   64    0    0    0    0    0    0    0    0
   0    0    0    0    0   31   38    0    0    0    0    0    0    0    0   32
   0    0    0    0    0   56    0    0   60    0    0    0    0   47   40   55
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0   46   47    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0   59    0    0    0    0   59    0   32    0   43
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0   31    0    0   32    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
   0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0
```

Figure 7-19MCGP outputs with bias and thresholds

Figure 7-19 shows a MCGP outputs when bias=30 is applied on the fitness function to reduce the weight of benign microcalcifications or non microcalcification parts and threshold applied to assign the malignancy and benignity. On the output results, it shows that the each chromosome over classified malignant microcalcifications in each part of ROI, it leaded that the MCGP network over classified malignant microcalcifications. It delivered relatively smaller values for the parts which truly contain malignant microcalcifications and relatively large values for the parts which contains breast tissues or benign microcalcifications. The results from this experiment delivered a message that adding bias and threshold could not solve the indicated problem. MCGP algorithm does not have a good performance in this experiment.

7.2.3.2 Single Chromosome CGP network

The previous experiments showed that the performance of MCGP network on classification of microcalcifications is desirable. The MCGP network contains 256 individual single chromosome CGP networks, each single chromosome network is working on one equal non-overlapped part of a ROI. Therefore, for one particular single chromosome CGP sub network, may only be trained on the benign part of the training images, but is used to classify the malignant part in testing images. This could cause a problem which is, this intelligent single chromosome CGP network never ever learnt enough information. Although in the MCGP network training stage, the best genotype to represent the entire network has been recorded and used on the testing stage, this non or less training algorithm problem is still not solved. Therefore, to solve this problem, instead of putting all 256 Chromosomes CGP network for 256 parts in one image, only one chromosome CGP network is trained through all 256 parts of one image. The working procedure shows as follows:

one CGP
network

A ROI with equal nonoverlapped
256 parts

P_1^1	P_1^2	...	P_1^{15}	P_1^{16}
P_2^1	P_2^2	...	$_2^{15}$	P_2^{16}
⋮	⋮	...	⋮	⋮
P_{15}^1	P_1^1	...	P_{15}^{15}	P_{15}^{16}
P_{16}^1	P_{16}^2	...	P_{16}^{15}	P_{16}^{16}

Since the single chromosome CGP network is the subject of the MCG network, the parameters, which are used to construct single chromosomes, the CGP networks in both cases remain the same as table 6-6.

The single chromosome CGP network is applied to train each individual parts of the image. For classification stages, there are a number of parts only containing benign microcalcification information, or surrounding the breast tissue and less parts containing malignant microcalcifications. Training results shown in Table 7-7 and Figure 7-20. The fitness function will be the same as Equation 6-1 and 6-2 for both linear and non-linear fitness functions:

Table 7-7 average and best fitness for single chromosome CGP network training stage

	Average Fitness	Best Fitness	Worst Fitness
Experiment 1	0.97	0.99	0.51
Experiment 2	0.96	0.99	0.55
Experiment 3	0.97	0.99	0.34
Experiment 4	0.97	0.99	0.55

```
#White pixel values for image awrmlM1.dat
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0  17  24   5   0   0   0   0   0   0   0   0   0   0   0   0
19  64  64  58  16   0   0   0   0   0   0   0   0   0   0   0
30  64  64  64  36   3   0   0   0   0   0   0   0   0   0   0
 9  53  64  64  64  45   3   0   0   0   0   0   0   0   0   0
 1  22  60  64  64  64  46   0   0   0   0   0   0   0   0   0
 0   0   0   8  44  64  64  44   1   0   0   0   0   0   0   0
 0   0   0   0   0   2   4   0   0   0   0   0   0   0   0   0
```

```
#White pixel values for image asrccB2.dat
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
```
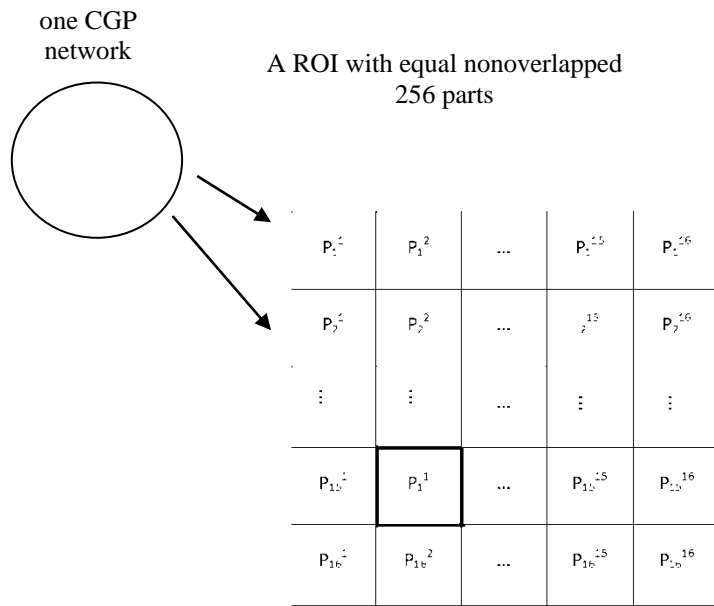
```
#CGP output values for best genotype on image awrmlM1.dat
 0   0   0   0   0   0   6   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   1   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   1
 2   0   0   0   0   0   0   0   0   0   0   0   0   0   2   0
 0   0   0  58   0   0   0   0   0   0   0   0   0   0   0   0
 1   0  51   0  63   0   0   0   0   0   0   0   0   0   0   0
 1   0   0   0  60  63   0   0   0   0   0   0   0   0   0   0
12   0   1   0   0   1   0   0   0   0   0   0   0   0   0   0
63   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
12   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
```

```
#CGP output values for best genotype on image asrccB2.dat
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   1   0   0   0   0   0   0   0   0   0   0   0   1   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   2   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   1   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   1   0   0   0   0   0   0   0
 0   0   0   0  62  60   0   0   0   0   0   0   0   0   0   0
 0   0   2  63  10   0   0   0   0   0   0   0   0   1   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0  12   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   5   0   1   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
```

Figure 7-20 Training results of single chromosome CGP network

.

Table 7-8 Confusion Matrix for the testing results of system

Confusion Matrix (Average value)				Statistics				
Bb	Bm	Mb	Mm	TP	TN	FP	FN	MCC
2	16	13	0	0	0.125	0.875	1	-0.88

From the training results, the single chromosome CGP network has been created to classify microcalcifications into benign and malignant cases. Table 7-6 shows the fitness of the training stages. These mean that the single chromosome CGP network

has been fully trained to achieve the best fitness score as 1. From Table 7-8 , it could find that the single chromosome CGP network had same problem with the MCGP network which provide better fitness values on training stages and bad performance on testing stages. Figure 7-20, the malignant outputs started to contain larger numbers of outputs valued on the benign image output. However it also the output larger numbers for benign microcalcifications that shows that the single chromosome CGP network starts to learn from the malignant microcalcification information. However it still cannot deliver a good performance.

7.2.4   Discussion

In this chapter, MCGP network has successfully detected the location and surround areas of microcalcifications by apply MCGP network in a mammogram dataset with microcalcification indicators. The detection of microcalcifications is not the entire automatic analysis of mammograms however it is one of the important part. The successful detection strongly showed the potentials of Multi chromosomes Cartesian Genetic Programming (MCGP) as a new, evolutionary algorithm on the area of automatic analysis mammograms. It also indicated that the advantages of evolutionary algorithm that they are based on learning how to do a task rather than just following the mathematics formula or sequential orders.

The limitation of self learning algorithm is it requests a reliable dataset to train the algorithm. To delivering a good performance on an algorithm, there are two important issues: 1) A suitable Algorithm which has the ability to solve the problem; 2) An reliable dataset to train the algorithm. The detection of microcalcifications by using MCGP have showed its good performance, however the performance of classification of

microcalcifications has failed by using same algorithms for detection stage and leave a problem that the algorithm has a better training results and a bad testing results, the better training results indicated that this new evolutionary algorithms are showing its the benefit on learning how to do a task and potential to automatic analysing mammograms. However the bad testing results are the evidence of a non reliable dataset cause problems on the performance of evolutionary algorithms; especially it is under the circumstances that there are different experiments have been carried out for optimizing the MCGP network in order to solve the problems. This leading the reason of non desired performance results of MCGP network on classification of microcalcifications is due to a non-reliable dataset.

Here are 4 examples of data training images, they are from this LLNL dataset and have been used to train the MCGP network for automated analysis of mammograms.

(a)   (b)   (c)   (d)
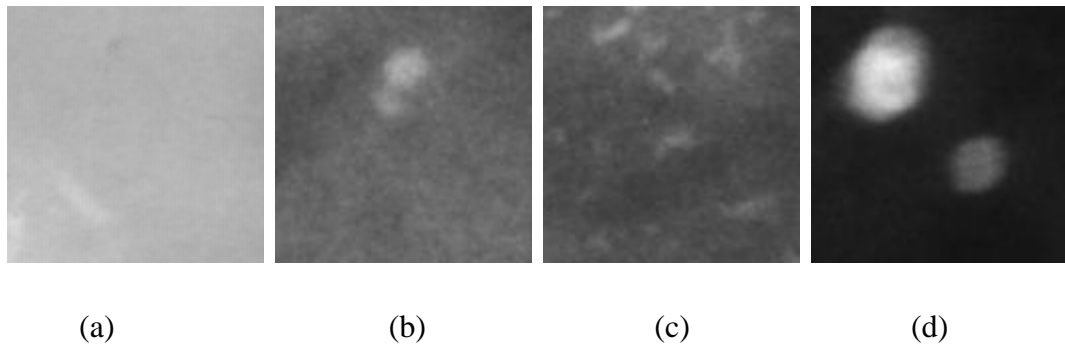
These 4 different images, (a)-(c) contains malignant microcalcifications, with image (d) being a benign image. All the 4 images have different contrast and low quality (a) images. This information could easily confuse the CGP system when it is on the training stages. The contrast of the images are not united, the differences of features between malignant and benign microcalcifications are not clear in these images, the features of

benign and malignant microcalcifications are leading confusing for MCGP network. All these issues could cause the MCGP network learn how to classify benignancy and malignancy microcalcifications not properly or miss- learning the typical features or over trained the system in training stage which leading a failure performance on testing stage. As the survey reviews dataset has been mentioned in chapter 5, a reliable, trusted dataset which the data has been selected and supplied under same circumstances is one of the important key for the successful automatic analysis on mammograms by using evolutionary algorithm. A new dataset that is overcome the existing problems has been introduced in Chapter 8.

8 Conclusion

8.1 Summary of work undertaken

8.1.1 Motivation – Diagnosis cancer in mammograms for screening

Breast cancer is the leading cause of death in women. The diagnosis of breast cancer at the earlier stage may be particularly important since it provides early treatment. The death rate of breast cancer has fallen since the breast screening program was introduced. The main method for screening patients is the mammogram - a high resolution x-ray of the breast. The hard work is the early detection of any tissues abnormality and confirmation of their cancerous natures. In addition, finding abnormalities on very early stages can also be affected by poor quality of image and other problems that might show on a mammogram. The process of identifying and evaluating signs of cancer from mammogram is a very difficult and time consuming task which requires fully trained radiologists. However, by natural and exposed to human error, there is a chance that cancer will be missed and patients being misdiagnosed. Computer aided diagnosis system has been introduced with the aim to achieve a more accurate and reliable diagnosis.

The two powerful indicators to evaluate mammograms are masses and microcalcifications. In general, masses are more difficult to classify than microcalcifications. Microcalcifications are small calcium deposits which occur as secretion from the ductal structures which have been thickened and dried. They can have the most benign cases but could also indicate malignancy. They are common on mammograms and their appearance increases with age. Over a decade, there are many

researches on the application of CAD system, and many of them evolved image processing of digital mammogram. A typical approach is pattern recognition scheme with sensing, segmentation, feature extraction, selection and classification. Evolutionary algorithm particularly genetic algorithm has previews been used with success in CAD feature selection system. Evolutionary algorithm, particularly genetic algorithm has previously been used with success in CAD feature selection system, the limitation of using evolutionary algorithm is prompted to investigate the potential benefit in the classification of microcalcifications.

8.1.2 Complexity and scale of problem Requires novel learning technology – evolutionary algorithm

To automatically analysis images especially medical image as mammograms is a complex and difficult job requiring a reliable algorithm and a suitable dataset since a good performance of an algorithm is the data it is trained on. In the real world, this is a challenge.

Since CAD system has a few problems on finding a system detection which does not necessarily have been detected by the radiologist, this might lead to the system not being trained correctly. Most CAD systems are based on classical imaging programming operations that are limited to the general understanding of how malignancy are found, as well as the limitations of CAD on practical implementation and classical image processing operations. To overcome this, evolutionary algorithms have the great advantage that they are based on learning how to do a task rather than just following a sequence of orders. EA has the great advantage that they are based on learning how to do a task rather than just following the mathematics formula or

sequential orders. EA contains programming technologies which simulate evolutionary behaviour of nature. The flexibility and learning ability of EA make it to be a good algorithm on automatic evaluating mammograms. In this work, it is decided to use Cartesian Genetic Programming for the representative EA to automatically analysis of mammograms. CGP is a special kind of Genetic Programming which is as a part of EA. Based on the aim of investigating the potential of EA on automatic analysis of Mammograms, the new representation CGP called Multi-chromosome CGP algorithm has been developed. Its structure determined that MCGP algorithm is an entire network consists of a number of small single chromosome networks. Therefore an entire mammogram is divided into a number of sub-images in order to divide one big problem into many small sub problems.

8.1.3 Suitable medical dataset required to train and evaluate algorithms

8.1.3.1 Medical datasets have common problems

Despite that EA have benefited from the traditional CAD system, there is also limitation from it. The performance of this algorithm require a good training database, more specifically, the classes represented in the training set must consist of true examples or reliable classes. However, the good mammograms dataset is a major challenge. For traditional methods, the dataset is critical and yet obtaining one is not straightforward. An obvious approach would be to obtain access to a database that has already been prepared. A common problem is the clinical protocol under which the data obtained is not similar to the requirement for the target.

With world spread adoption of computer based health system, the digitized mammogram image has increased significantly. As mammographic screening has been used on the early detection of breast cancer, a large number of public databases are or have been available in the past. However the quality, size and reliability of the data available are not under the same requirements. Also most of these databases have drawback which make them inappropriate for study.

In Mammogram dataset, most of them are digitized mammography; there are different digitizer equipments to get the measurements. However different scanner from different manufactures will digitize the analogue films in slightly different ways, therefore some X-ray film are classified by an algorithm based on which scanner was used in the digitization process.

8.1.3.2 Research has concluded that no suitable dataset currently exists

A special volume of data has been extracted from LLNL for the use in evaluating CGP that on classification of micro calcifications in Chapter 6. Additionally, according to the performance of MCGP working on the LLNL dataset, there are limitations with this dataset to evaluate the algorithm. In fact, that the LLNL dataset which has been used is relatively outdated. Based on the Implementation of this work, the non reliable dataset is the cause that the system could not provide a good performance on automatic diagnosis cancerous on mammograms. Such as there is no standard resolution for the images, the contrast differences cause the algorithm's confusion. The limited number of training data could not provide enough information to let the algorithm learn properly. Different formats and sources could generate artifacts that may interfere with training of algorithms.

8.1.4 Implementation challenging and experimental results

The multi-chromosome CGP network could not give a good performance on the provided dataset; it is not because Multi-chromosomes CGP is not working properly, the experimental results have shown that MCGP could detect microcalcifications on mammograms which leads to potentially providing a confident diagnosis. However, when the MCGP network is used on the classification stage, the extra information required to train the network from dataset could not be gained. The non-reliability of this dataset caused MCGP network confusion when it tries to learn classification.

8.1.5 Creating a new dataset

For mammography, it is difficult as the data cannot be simply being generated due to the medical situation of mammograms. However, as mentioned in section 5.2, several publically available datasets have their disadvantages. Therefore, to construct a clinically validated database of mammograms   is an important stage for the whole work, that's because the entire project requires a dataset to be absolutely stable, reliable, accurate and suitable to train and test the effectiveness of the EAs developed in diagnostic support in breast mammography. Heading towards on this aim, the acknowledgement of an expected mammogram dataset is introduced, and a new YDH dataset has been produced based on the needs, objectives and purpose of this project, information on the tools used [107] by the user interface is also available to this section.

8.1.5.1 Contents of the New Dataset

In order to construct a new dataset, it requires clinical co-operation to provide sufficient conditions and information to build up a well-served complete mammogram image data for the diagnosis of microcalcifications in mammogram.

8.1.5.2 YDH Mammography Images

This dataset contains mammograms obtained from York District Hospital, therefore, YDH has been chosen as the name of the new dataset.

The number of patients required is one of the statistical justifications to show whether the performance of the algorithm is affected by specific patients. Normally, this information comes from the clinical record. This also shows the availability of patients by showing different aspects of the records of one case from a single patient. [108] indicated that in order to obtain statistical justification for measurements of the performance of the algorithm, it should be tested on a large dataset.

The dataset contains the patient age, screening exam date, and the date when the mammogram was digitized. The patient's name or symbol name could also be shown in the dataset; this may help the radiologist for following up diagnosis or for statistics.

8.1.5.3 Type of mammogram

- The dataset needs to indicate whether the data is film screen mammograms or full digitized mammograms. If it is a film screen mammogram, the sampling rate,

number of grey levels and equation to estimate optical density from grey value for the scanner used to digitize the mammograms needs to be recorded.

- Based on chapter 2, the four standard views (medio-latra oblique and cranio caudal) from each side of the breast for one patient are indicated in the dataset.

MLO	Left
	Right
CC	Left
	Right

- What circumstances under which the mammograms have been taken.

Screening exam	screening Routine	Follow up diagnosis	Guiding biopsy

Therefore, the data could be assigned into different volumes according to the severity of the findings.

Normal	Benign without call back	Benign	Malignant

- In Mammogram dataset, most of them are digitized mammography which are digitized by different scanners, therefore a standard unit digitizer to digitize mammograms is necessary.

- To avoid any confusion made by algorithm, the data shares same contrast, pixel per inches and sample rate.

- Image format need to be united in order to present a reliable dataset for training the evolutionary algorithms especially it has been highly recommended by itself learning motivation.

Similar to the existing LLNL dataset, however still different, the image format has been suggested with the common format as follows:

The file will be a text file, and it contains 2 parts:

Header:

FILENAME: original image file name

XRES: image x resolution

YRES: image x resolution

IRES: image intensity resolution (e.g. 255 grey-levels)

MAN: digitizer manufacturer and model

Time: when to select the ROI

Marking Scheme: from 1 to 5

Body:

Pixel X coordinate and pixel Y coordinate (repeated for all pixels within area of selection

TYPE: structure type (e.g. MICROCALC or MASS)

RATING: cancer risk - HIGH, MEDIUM or LOW

CONFIDENCE: confidence of rate - HIGH, MEDIUM or LOW

RATER: rate's initials

- Marking scheme

There will be two marking schemes. One will mark the boundary of ROI as a larger marking core. Then the secondary smaller marking scheme will be used to mark the core of speculated mass or cluster of microcalcifications.

These marking schemes, which would be used by two radiologists, could result in a more confident diagnosis

There could be two fully trained radiologists to mark up the frame and ROI which contains the suspicious mass or microcalcifications. Each marking contains a descriptive value. This value is specified by mammography radiologists using the UK scoring system called RCRBG which will maintain confidence in classification.

- User Interface

Once the mammogram data in DICOM files have been supplied by YDH experts, a full trained radiologist starts to segment the Region of Interests by using Mammogram Data Acquisition System which is a user interface created by Christopher Booth, supervised by Dr. Steve Smith. This software user interface especially designed for this new data images would allow the user to convert images without changing the image quality. There are 1,237 images which are allowed to batch in this software. It would help the radiologist to zoom in the image in order to get a closer look. Once the radiologist has selected the Region of Interests, the images will be saved in one of the four conversion format-BMP, PNG, BMP+ additional information output as text file, PNG+ additional information output as text file.

- Pre-processing for image data

Once the ROI has been selected by a radiologist, the data image needs to be segmented into smaller sub-images. In order to let Multi-chromosome CGP network learn properly, the sub-image for the new dataset is no long non-overlapping. There is a sliding window operation which will be used. Figure 8-1 shows how sliding window works. The particular window used 8×8 pixel sized mask, every time this mask only slides 4 pixels length to the left or right or up or down and record the pixels into a new 8×8 sub-images. After the window slides through the whole image, there will be an 8×8 sub-images. Then each chromosome from a multi-chromosome CGP network randomly chooses one sub-image as its training inputs. This will ensure that each chromosome learns all the information from the training image to

avoid any miss training. In addition, all the sub-images contains information of microcalcification from different angles, making sure the edge of the microcalcifications will not be missed out.

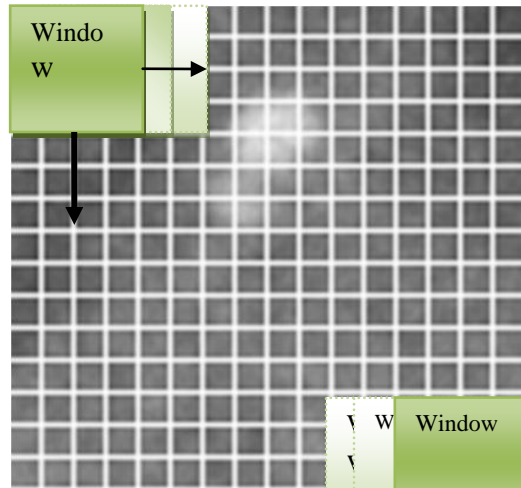

Figure 8-1Sliding window

8.1.6 An example of new image data file

The new data are generated by the contents of YDH dataset. Here is one of example about a new image dat file, dat. files are converted by the data images of dataset and they are commonly used on the input training and testing images for evolutionary algorithms , on the application of auto detected microcalcifications .

Cam739-L-MLO.bmp_lg_MarkedRegion_1_Output3
.mal 4
.i 64
.x 16
.y 16
.p 256
```
178 183 184 186 183 183 184 184 179 181 183 182 180 181 186 186 179
184 186 186 185 181 180 178 180 180 182 182 183 184 185 183 181 185
179 175 174 176 177 174 174 174 179 173 171 174 173 170 173 173 179
171 175 173 169 169 172 170 170 169 175 177 170 169 170 172 173 167
170 165 165 166 164 164 165 163 172 167 165 163 165 163 162 162 171
162 165 164 163 165 163 163 160 161 163 160 161 167 166 164 166 163
158 160 163 162 162 164 168 170 162 162 160 159 163 163 168 165 162
166 169 169 171 170 173 171 172 169 169 168 169 170 172 170 167 169
170 170 169 172 169 169 170 171 166 167 168 169 168 169 168 168 166
173 171 168 165 167 170 169 170 168 169 166 167 168 169 169 171 169
169 169 168 170 173 173 170 168 168 169 166 168 171 170 170 167 172
171 176 176 178 180 181 183 182 169 172 178 179 179 179 183 186 171
183 182 181 183 181 178 175 177 186 181 182 184 184 181 177 178 183
176 175 178 178 173 175 175 174 179 176 176 173 172 176 173 172 180
172 175 178 178 177 174 172 166 172 172 177 176 173 172 173 171 171
166 168 167 166 169 170 170 173 167 166 166 169 171 170 172 172 167
182 179 178 176 178 178 178 179 181 182 181 179 180 178 177 178 180
181 181 182 182 183 182 183 185 180 181 183 183 181 180 182 183 181
184 182 179 180 177 180 181 180 181 182 179 182 178 182 183 183 178
179 178 175 173 176 175 177 177 184 183 176 177 177 177 178 178 180
175 174 172 172 172 173 171 163 176 177 176 173 169 171 170 166 173
165 168 168 169 168 170 172 170 168 171 165 170 170 169 172 170 166
174 171 171 173 174 173 174 173 175 175 175 175 175 172 174 175 174
173 172 171 173 176 175 170 168 177 176 172 176 175 179 173 170 178
175 175 174 174 174 175 173 174 174 176 175 175 174 173 174 176 170
173 172 174 172 174 172 174 175 175 171 173 177 177 174 173 173 173
175 177 177 174 167 169 178 177 176 176 174 171 170 172 175 174 172
174 173 175 175 178 179 179 176 171 172 174 173 177 180 176 176 168
178 178 174 177 177 178 182 180 178 177 179 178 178 179 184 179 176
176 178 176 178 176 176 177 173 176 177 178 178 174 177 178 177 178
171 170 172 175 170 171 174 174 175 172 172 176 174 174 173 176 177
174 175 176 174 176 176 176 174 174 174 178 175 174 175 177 173 169
178 178 178 180 182 177 173 173 179 179 181 180 181 181 179 177 178
176 177 180 179 178 177 175 175 179 179 180 179 183 178 178 177 181
173 171 173 176 179 176 173 175 170 170 170 175 175 174 174 175 175
180 180 178 176 174 177 176 176 177 181 176 174 173 179 178 179 175
174 170 170 172 171 171 172 171 179 170 170 172 173 175 175 169 178
170 172 173 172 169 171 176 173 170 173 173 169 171 175 177 175 170
175 174 170 168 170 171 171 174 175 171 168 168 167 170 172 173 173
174 173 175 173 168 169 171 176 176 175 173 177 170 174 172 173 175
174 177 174 172 173 170 170 172 174 188 177 171 172 170 173 175 177
175 176 172 171 174 176 176 177 177 177 177 173 176 180 179 175 177
177 176 175 178 175 176 177 176 174 174 175 176 173 175 176 176 176
174 170 168 172 171 172 177 176 175 167 170 175 173 176 178 176 176
176 175 173 173 178 179 177 176 176 173 171 174 176 176 177 176 178
```

8.2 Hypothesis revisited

The work presented in this thesis addressed the following hypothesis:

"Multi-chromosome Cartesian Genetic programming is an effective evolutionary algorithm to facilitate automatic analysis of mammograms."

Based on the aim of excess investigation in the potential benefit of MCGP on a confident automatic analysis of microcalcifications on mammograms, a Multi-chromosome CGP network is developed to classify microcalcifications by using the raw pixel values. This self-learning algorithm has great advantages that they are based on learning how to do a task rather than just follow a sequence of orders. To achieve this, it relies on its training data. Two experiments have been designed towards the main aim, one is applying MCGP network on detection of microcalcifications, this stage is to investigates the suitability of whether MCGP could successfully apply on such a complex application. Both of the experiments are using LLNL dataset. From the results, it clearly showed that MCGP algorithm could deliver a confident detection of microcalcifications on mammograms; this is demonstrated on the paper [109]. On the stage of classification microcalcifications, a range of experiments have to implemented to approve that: MCGP could not perform a confident classification microcalcifications based on the non-suitable dataset This problem indicates that the appropriate datasets of equal importance to algorithms when looking for a solution to automated assessment of mammograms.

The new dataset has been provided with several standards to meet the requirement of the studies; multiple chromosomes CGP is still a promising approach since it has already showed the potential on detecting microcalcifications.

8.3 Further work

During limited time, there is still some work which could help to evaluate this automatic Automated Analysis Evolutionary Algorithms.

8.3.1 Create new datasets to known specification and hence characteristics

A standard new dataset which have been introduced particularly for this algorithm, with the aim to avoid any problems caused by the current dataset. In this dataset, data has been formatted into a new standard format. The new dat. has been generated to fit into the MCGP network for the investigation.

8.3.2 Test EAs developed

Once the new dataset has been used to train the system, the stage of test of developed by MCGP for classification of microcalcifications could be brought to the table again.

There is a new approach called sliding window mask which could be used to train the MCGP.

The fitness function has been changed since the reference number used to compare with the CGP output is a range from 0-5 to indicate the diagnosis confidence. Therefore the new fitness function is:

CGP output is in a range of 0-255, which should add 45 then divided by 6:
If the rescaled CGP output is in range of 0-50 then it rescaled into 0;
Else if the rescaled CGP output is in range of 51-100 then it rescaled into 1;
If the rescaled CGP output is in range of 101-150 then it rescaled into 2;

.

Then the number of parts which have the number 0 needs to be calculated, stored as NUM_0;
Then the number of parts which have the number 1 need to be calculated, stored as NUM_1;
Then the number of parts which have the number 2 needs to be calculated, stored as NUM_2;
Then the number of parts which have the number 3 needs to be calculated, stored as NUM_3;
Then the number of parts which have the number 4 needs to be calculated, stored as NUM_4;
Then the number of parts which have the number 5 needs to be calculated, stored as NUM_5;

Then the total fitness will be the largest number of NUM_0 to NUM_5.

The network has been constructed by the author, with the new dataset and new structured network which could be used to determine the potential benefit of MCGP on automatic analysis of microcalcifications on mammograms. The programming code is enclosed in the Appendix.

Appendix

MCGP code for Classfication of mammograms

```c
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <math.h>
#include "cgp_breast2.h"


/* validate cgp program command line and create file strings for .par
and .data files */
void validate_command_line(int argc, char* argv[], char parfile[])
{
      puts("");
      puts("*********    WELCOME TO CARTESIAN GENETIC PROGRAMMING
*********");
      puts("********* Validating command line arguments to cgp program
*********");

      if (argc!=2)
      {
            puts("INCORRECT NUMBER OF ARGUMENTS");
            puts("Type cgp <file.par> then return");
            exit(1);
      }
      if (strlen(argv[1])>(MAX_NUM_LETTERS-1))
      {
            puts("filename for parameter file is too long");
            printf("It should be less than %d
characters\n",MAX_NUM_LETTERS);
            exit(2);
      }
      strcpy(parfile,argv[1]);

}

/* creates the list of training images and reads their names into an
array
   also calculates how many training images are provided */
void get_image_names(int train)
{

      FILE *fp;

      /* create list of files to be processed */
      if (train)
      {
            system("dir *_train.dat /B > list.txt");
            fp = fopen("list.txt","r");
            if (!fp)
            {
                  puts("Error. Cannot find file list.txt");
                  exit(0);
```

```c
                }
        }
        else
        {
                system("dir *.tst /B > list-test.txt");
                fp = fopen("list-test.txt","r");
                if (!fp)
                {
                        puts("Error. Cannot find file list-test.txt");
                        exit(0);
                }
        }


        num_images =0;
        do
        {

                if (feof(fp))
                        break;

                fscanf(fp,"%s",image_names[num_images]);

                num_images++;

        }
        while (!feof(fp));

        if (num_images == 0)
        {
                puts("Empty list of images");
                exit(0);
        }

        num_images--;

}

/* creates the list of test images and reads their names into an array
   also calculates how many training images are provided
   Added so that a test score can be provided during evolution */
void test_get_image_names(void)
{

        FILE *fp;

        /* create list of files to be processed */
        system("dir *.tst /B > list-test.txt");
        fp = fopen("list-test.txt","r");
        if (!fp)
        {
                        puts("Error. Cannot find file list-test.txt");
                        exit(0);
        }

        test_num_images =0;
        do
        {

                if (feof(fp))
```

```c
                        break;

                fscanf(fp,"%s",test_image_names[test_num_images]);

                test_num_images++;

        }
        while (!feof(fp));

        if (test_num_images == 0)
        {
                puts("Empty list of images");
                exit(0);
        }

        test_num_images--;

}

/* reads the image data as a .dat file (pgm image file with extra info)
   also calculates some program globals
*/
void read_data(char datafile[], int image)
{
    int         i,j;
    char    dummy[MAX_NUM_LETTERS];
      char  imagefilename[MAX_NUM_LETTERS];
    FILE*   fp;

    fp=fopen(datafile,"r");
    if (!fp)
    {
        puts("ERROR. Missing data file");
        exit(1);
    }
    else
    {
      fscanf(fp,"%s",imagefilename);
        fscanf(fp,"%s %d",dummy,&malignancy_level);
      fscanf(fp,"%s %d",dummy,&num_inputs);
        if (num_inputs > MAX_NUM_INPUTS)
        {
                printf("\nERROR. num_inputs must be less
than %d\n",MAX_NUM_INPUTS);
                exit(0);
        }
      fscanf(fp,"%s %d",dummy,&num_xparts);
      fscanf(fp,"%s %d",dummy,&num_yparts);
      fscanf(fp,"%s %d",dummy,&num_parts_per_image[image]);
        if (num_parts_per_image[image] > MAX_NUM_PRODUCTS)
        {
                printf("\nERROR. num_parts must be less
than %d\n",MAX_NUM_PRODUCTS);
                exit(0);
        }

      for (i=0;i<num_parts_per_image[image];i++)
      {
        for(j=0;j<num_inputs;j++)
            fscanf(fp,"%u", &data_inputs[image][i][j]);
```

```c
//                   fscanf(fp,"%u", &number_white_pixels[image][i]);
            }
        fclose(fp);


    }
}


/* reads the image data as a .dat file (pgm image file with extra info)
   also calculates some program globals
*/
void test_read_data(char datafile[], int image)
{
    int           i,j;
    char    dummy[MAX_NUM_LETTERS];
      char  imagefilename[MAX_NUM_LETTERS];
    FILE*   fp;

    fp=fopen(datafile,"r");
    if (!fp)
    {
        puts("ERROR. Missing data file");
        exit(1);
    }
    else
    {
      fscanf(fp,"%s",imagefilename);
        fscanf(fp,"%s %d",dummy,&malignancy_level);
      fscanf(fp,"%s %d",dummy,&num_inputs);
        if (num_inputs > MAX_NUM_INPUTS)
        {
              printf("\nERROR. num_inputs must be less
than %d\n",MAX_NUM_INPUTS);
              exit(0);
        }
      fscanf(fp,"%s %d",dummy,&num_xparts);
      fscanf(fp,"%s %d",dummy,&num_yparts);
      fscanf(fp,"%s %d",dummy,&test_num_parts_per_image[image]);
        if (test_num_parts_per_image[image] > MAX_NUM_PRODUCTS)
        {
              printf("\nERROR. num_parts must be less
than %d\n",MAX_NUM_PRODUCTS);
              exit(0);
        }

      for (i=0;i<test_num_parts_per_image[image];i++)
      {
         for(j=0;j<num_inputs;j++)
            fscanf(fp,"%u", &test_data_inputs[image][i][j]);
            fscanf(fp,"%u", &test_number_white_pixels[image][i]);
      }
      fclose(fp);


    }
}

/* read from the parameter file all the global parameters */
void get_parameters(char parfile[])
{
      int           i;
```

```c
        char  dummy[50];
        FILE* fp;

        printf("\n********* Reading parameters defined in %s
*********\n",parfile);
        fp=fopen(parfile,"r");
        if (!fp)
        {
                printf("Missing file: %s\n",parfile);
                exit(1);
        }
        fscanf(fp,"%d %s",&population_size,dummy);
        fscanf(fp,"%lf %s",&per_cent_mutate,dummy);
        fscanf(fp,"%d %s",&num_generations,dummy);
        fscanf(fp,"%d %s",&num_runs_total,dummy);
        fscanf(fp,"%d %s",&num_rows,dummy);
        fscanf(fp,"%d %s",&num_cols,dummy);
        fscanf(fp,"%d %s",&levels_back,dummy);
        fscanf(fp,"%d %s",&progress_report,dummy);
        fscanf(fp,"%d %s",&report_interval,dummy);
        fscanf(fp,"%u %s",&global_seed,dummy);
        fscanf(fp,"%d %s",&save_best_chrom,dummy);
        fscanf(fp,"%d %s",&run_from_chrom,dummy);
        fscanf(fp,"%d %s",&malignancy_threshold,dummy);
        fscanf(fp,"%d %s",&linear_fitness,dummy);

        num_chromosomes = 1;

        /* assigned global constants */
        num_functions=0;
        num_genes_per_node=3;

        num_outputs = 1;

        for (i=0;i<MAX_NUM_FUNCTIONS;i++)
        {
                fscanf(fp,"%d %s",&number[i],&node_types[i]);
                if (number[i])
                {
                        allowed_functions[num_functions]=i;
                        num_functions++;
                        if (i>15)
                                num_genes_per_node=4;
                }
        }
        fclose(fp);

        /* get training image names */
        get_image_names(1);

        printf("\nReading image data\n");
        for (i = 0; i < num_images; i++)
        {
                read_data(image_names[i],i);
        }

        /* get testing image names and test data */
        test_get_image_names();

        printf("\nReading test image data\n");
```

210

```c
        for (i = 0; i < test_num_images; i++)
        {
                test_read_data(test_image_names[i],i);
        }

        //printf("\nnum_rows is %d\n", num_rows);
        //printf("\nnum_cols is %d\n", num_cols);
        //printf("\nnum_genes_per_node is %d\n", num_genes_per_node);

        //exit(0);

        /* assigned global constants */
        num_nodes=num_rows*num_cols;
      num_genes=num_genes_per_node*num_nodes+num_outputs;


        //printf("\nnum_images is %d\n", num_images);

        //printf("\nnum_genes is %d\n", num_genes);

        //exit(0);

    end_count=num_inputs+num_nodes;

        perfect_fitness = 1.0;


        /* initialise output (used in cgp decoding) */
      for (i=0;i<end_count+num_outputs;i++)
        output[i]=0;

        if (population_size > MAX_POPULATION_SIZE)
        {
                printf("Too large a population size
(<= %d)\n",MAX_POPULATION_SIZE);
                exit(0);
        }

        if (num_chromosomes > MAX_NUM_CHROMOSOMES)
        {
                printf("Too many chromosomes requested
(<= %d)\n",MAX_NUM_CHROMOSOMES);
                exit(0);
        }

        if (num_genes > MAX_NUM_GENES)
        {
                printf("Too many genes selected (<= %d)\n",MAX_NUM_GENES);
                exit(0);
        }

        if (num_runs_total < 1)
        {
                puts("Number of runs of EA must be at least 1");
                exit(0);
        }
        else if (num_runs_total > MAX_NUM_RUNS)
        {
                printf("Number of runs of EA must be less than %d\n",
MAX_NUM_RUNS);
```

211

```c
                exit(0);
        }

        if (num_genes < 10)
        {
                puts("Number of genes/bits must be at least 10");
                exit(0);
        }

        if ((progress_report < 0) || (progress_report > 3))
        {
                puts("Progress report parameter must be less than 4");
                exit(0);
        }

        if (levels_back > num_cols)
        {
                puts("levels_back parameter must be less than or equal to
number of columns");
                exit(0);


        }

        srand(global_seed);

        puts("********* Beginning execution ********");
}

/* write out parameter values in results file */
void write_cgp_info(char command[], char filename[MAX_NUM_LETTERS])
{
        int             i;
        FILE* fp;

        fp=fopen(filename,"w");
        fprintf(fp,"The program is         %s\n",command);
       fprintf(fp, "The training image files are. Need to add these\n");
        fprintf(fp,"population_size is     %d\n",population_size);
        fprintf(fp,"mutation rate is       %6.2lf\n",per_cent_mutate);
        fprintf(fp,"num_generations is     %d\n",num_generations);
        fprintf(fp,"num_runs is            %d\n",num_runs_total);
        fprintf(fp,"num_rows is            %d\n",num_rows);
        fprintf(fp,"num_cols is            %d\n",num_cols);
        fprintf(fp,"levels_back is         %d\n",levels_back);
        fprintf(fp,"progress report is     %d\n",progress_report);
        fprintf(fp,"report interval is     %d\n",report_interval);
        fprintf(fp,"global_seed is         %u\n",global_seed);
        fprintf(fp,"save_best_chrom is     %d\n",save_best_chrom);
        fprintf(fp,"run_from_chrom is      %d\n",run_from_chrom);
       fprintf(fp,"malignancy_threshold is   %d\n",malignancy_threshold);
        fprintf(fp,"linear_fitness is  %d\n",linear_fitness);

        for (i=0;i<MAX_NUM_FUNCTIONS;i++)
        {
                fprintf(fp,"%d %s\n",number[i],node_types[i]);
        }
        fprintf(fp,"\nHere are the Results\n");
        fclose(fp);
}
```

```c
/*  returns a random integer between 0 and range-1 */
int newrand(int range)
{
    int temp;

    temp=rand() % range;
    return(temp);
}



/* prints out all the chromosome in a geneotype both with and without
active genes */
void fprint_a_user_readable_genotype(int** genotype, char name[], int
append)
{
      int k;

      for (k = 0; k < num_chromosomes; k++)
      {

            fprint_a_user_readable_chromosome(genotype[k], name,
append);

            fprint_active_genes(genotype[k],name);
      }
}

/* prints a chromosome to a file
   when append is 1, the function appends the information to the file
   when append is 0, the function creates a new file
*/
void fprint_a_user_readable_chromosome(int* chromosome, char name[],
int append)
{
   int            i, node_label;
   int      write_bracket=1;
   FILE*    fp;

   if (append)
        fp=fopen(name,"a");
   else
        fp=fopen(name,"w");


   node_label = num_inputs-1;
   for (i=0;i<num_nodes*num_genes_per_node;i++)
   {
      if ((i+1)%num_genes_per_node == 0)
        {
            node_label++;
            fprintf(fp,"[%d]:%d)\t",chromosome[i],node_label);
            write_bracket = 1;
        }
      else
        {
            if (write_bracket)
                  fprintf(fp,"(");

          fprintf(fp,"%d,",chromosome[i]);
```

213

```c
                write_bracket = 0;
         }
      }
   fprintf(fp,"\t\t");
   for (i=0;i<num_outputs;i++)
      fprintf(fp," %d",chromosome[num_nodes*num_genes_per_node+i]);
   fprintf(fp,"\n\n");
   fclose(fp);
}


void fprint_a_genotype(int** genotype,char name[], int append)
{
      int i;

      for (i = 0; i < num_chromosomes; i++)
            fprint_a_chromosome(genotype[i],name,append);
}


/* prints a chromosome to a file
   This is in usual cgp format - not very readable
   when append is 1, the function appends the information to the file
   when append is 0, the function creates a new file
*/
void fprint_a_chromosome(int* chromosome, char name[], int append)
{
   int             i;
   FILE*    fp;

   if (append)
         fp=fopen(name,"a");
   else
         fp=fopen(name,"w");

   for (i=0;i<num_genes;i++)
   {
      if ((i+1)%num_genes_per_node == 0)
            fprintf(fp,"%d\t",chromosome[i]);
      else
         fprintf(fp,"%d ",chromosome[i]);
   }
   fprintf(fp,"\t\t");
   for (i=0;i<num_outputs;i++)
      fprintf(fp," %d",chromosome[num_nodes*num_genes_per_node+i]);
   fprintf(fp,"\n\n");
   fclose(fp);
}

/* prints a chromosome to the screen */
void print_a_chromosome(int* chromosome)
{
   int i;

   for (i=0;i<num_nodes*num_genes_per_node;i++)
   {
      if ((i+1)%num_genes_per_node == 0)
            printf(" %d\t",chromosome[i]);
      else
         printf(" %d",chromosome[i]);
   }
   printf("\t\t");
```

```c
    for (i=0;i<num_outputs;i++)
        printf(" %d",chromosome[num_nodes*num_genes_per_node+i]);
    printf("\n");
}

/* creates the active chromosome. This is the same as
   the chromosome, except all junk entries are -1
*/
void get_nodes_used(int* chromosome, int node_used[])
{
        int        i,j,index;
        int        address[MAX_NUM_GENES_PER_NODE];
        int*  active_chromosome = NULL;

        active_chromosome = create_chromosome_space();

        for (i=0;i<num_genes;i++)
            active_chromosome[i]=-1;
        for (i=num_genes-num_outputs;i<num_genes;i++)
            active_chromosome[i]=chromosome[i];

        /* first look at chromosome and identify gates not used */
        /* these are all the outputs of gates which do not appear in the
chromosome */
        for (i=0;i<num_nodes+num_inputs;i++)
            node_used[i]=0;
        /* all the nodes whose output is given by the output genes are
active */
        for (i=num_genes-num_outputs;i<num_genes;i++)
            node_used[chromosome[i]]=1;

        for (i=num_nodes+num_inputs-1;i>=num_inputs;i--)
        {
            if (node_used[i])
            {
                /* get input addresses and type of this gate */
                index=num_genes_per_node*(i-num_inputs);
                for (j=0;j<num_genes_per_node;j++)
                {
                    address[j]=chromosome[index+j];
                    active_chromosome[index+j]=address[j];
                }
                if ((address[num_genes_per_node-1]==2) ||
(address[num_genes_per_node-1]==4))
                {
                    node_used[address[0]]=1;
                }
                else if ((address[num_genes_per_node-1]==3) ||
(address[num_genes_per_node-1]==5))
                {
                    node_used[address[1]]=1;
                }
                else if ((address[num_genes_per_node-1]>=6) &&
(address[num_genes_per_node-1]<=15))
                {
                    node_used[address[0]]=1;
                    node_used[address[1]]=1;
                }
                else if (address[num_genes_per_node-1]>15)
                {
```

215

```
                                     node_used[address[0]]=1;
                                     node_used[address[1]]=1;
                                     node_used[address[2]]=1;
                         }
              }
       }

    free(active_chromosome);
}


/* prints out a chromosome showing inactive genes as -1 */
void fprint_active_genes(int* chromosome,char name[30])
{
       int            i,j,index,node_label,write_bracket = 1;
       int            num_unused_nodes=0;
       int            num_nodes_active;
       int            node_used[MAX_OUTPUT_SIZE];
       int*   active_chromosome = NULL;
       int            address[MAX_NUM_GENES_PER_NODE];
       FILE* fp;


       active_chromosome = create_chromosome_space();

       fp=fopen(name,"a");

       fprintf(fp,"Inactive genes in this chromosome are marked with
*\n");


       for (i=0;i<num_genes;i++)
            active_chromosome[i]=-1;
       for (i=num_genes-num_outputs;i<num_genes;i++)
            active_chromosome[i]=chromosome[i];

       /* first look at chromosome and identify gates not used */
       /* these are all the outputs of gates which do not appear in the
chromosome */
       for (i=0;i<num_nodes+num_inputs;i++)
            node_used[i]=0;
       /* all the nodes whose output is given by the output genes are
active */
       for (i=num_genes-num_outputs;i<num_genes;i++)
            node_used[chromosome[i]]=1;

       for (i=num_nodes+num_inputs-1;i>=num_inputs;i--)
       {
              if (node_used[i])
              {
                     /* get input addresses and type of this gate */
                     index=num_genes_per_node*(i-num_inputs);
                     for (j=0;j<num_genes_per_node;j++)
                     {
                            address[j]=chromosome[index+j];
                            active_chromosome[index+j]=address[j];
                     }
                     if ((address[num_genes_per_node-1]==2) ||
(address[num_genes_per_node-1]==4))
                     {
```

216

```c
                                node_used[address[0]]=1;
                        }
                        else if ((address[num_genes_per_node-1]==3) ||
(address[num_genes_per_node-1]==5))
                        {
                                node_used[address[1]]=1;
                        }
                        else if ((address[num_genes_per_node-1]>=6) &&
(address[num_genes_per_node-1]<=15))
                        {
                                node_used[address[0]]=1;
                                node_used[address[1]]=1;
                        }
                        else if (address[num_genes_per_node-1]>15)
                        {
                                node_used[address[0]]=1;
                                node_used[address[1]]=1;
                                node_used[address[2]]=1;
                        }
                }
        }

    node_label = num_inputs-1;
    for (i=0;i<num_nodes*num_genes_per_node;i++)
    {
            if ((i+1)%num_genes_per_node == 0)
            {
                    node_label++;
                    if (active_chromosome[i]<0)
                            fprintf(fp,"[*]:%d)\t",node_label);
                    else

fprintf(fp,"[%d]:%d)\t",active_chromosome[i],node_label);
            write_bracket = 1;
            }
            else
            {
                if (write_bracket == 1)
                            fprintf(fp,"(");

                if (active_chromosome[i]<0)
                        fprintf(fp,"*,");
                else
                        fprintf(fp,"%d,",active_chromosome[i]);

                write_bracket = 0;
            }
    }
    fprintf(fp,"\t\t");
    for (i=0;i<num_outputs;i++)

fprintf(fp," %d",active_chromosome[num_nodes*num_genes_per_node+i]);

    for (i=num_inputs;i<num_inputs+num_nodes;i++)
        if (!node_used[i])
                num_unused_nodes++;
    num_nodes_active=num_nodes-num_unused_nodes;

    fprintf(fp,"\nnumber of active gates is %d\n\n",num_nodes_active);
    fclose(fp);
```

```
        free(active_chromosome);
}

/* this routine works out how many nodes are active */
int get_num_nodes_active(int* chromosome)
{
        int    i,j,index;
    int    address[MAX_NUM_GENES_PER_NODE];
        int    num_unused_nodes=0,num_nodes_active;
        int    node_used[MAX_OUTPUT_SIZE];

        /* first look at chromosome and identify gates not used
           these are all the outputs of gates which
           do not appear in the chromosome */

        for (i=0;i<num_nodes+num_inputs;i++)
             node_used[i]=0;
        /* all the nodes whose output is given by the output genes are
active */
        for (i=num_genes-num_outputs;i<num_genes;i++)
             node_used[chromosome[i]]=1;

        for (i=num_nodes+num_inputs-1;i>=num_inputs;i--)
        {

             if (node_used[i])
             {
                    /* get input addresses and type of this gate */
                    index=num_genes_per_node*(i-num_inputs);
                    for (j=0;j<num_genes_per_node;j++)
                         address[j]=chromosome[index+j];

                    if ((address[num_genes_per_node-1]==2) ||
(address[num_genes_per_node-1]==4))
                    {
                         node_used[address[0]]=1;
                    }
                    else if ((address[num_genes_per_node-1]==3) ||
(address[num_genes_per_node-1]==5))
                    {
                         node_used[address[1]]=1;
                    }
                    else if ((address[num_genes_per_node-1]>=6) &&
(address[num_genes_per_node-1]<=15))
                    {
                         node_used[address[0]]=1;
                         node_used[address[1]]=1;
                    }
                    else if (address[num_genes_per_node-1]>15)
                    {
                         node_used[address[0]]=1;
                         node_used[address[1]]=1;
                         node_used[address[2]]=1;
                    }
             }
        }

        for (i=num_inputs;i<num_inputs+num_nodes;i++)
        if (!node_used[i])
```

```c
                num_unused_nodes++;

        num_nodes_active=num_nodes-num_unused_nodes;

        return num_nodes_active;

}

/* generate a starting population
   from a chromosome read from a file (cgp.chr)
*/
void read_from_chrom(int*** genotypes)
{
        int     i,j,k;
        FILE* fp;

    fp=fopen("cgp.gen","r");
    if (!fp)
    {
            puts("Missing file cgp.gen (contains a genotype");
        exit(1);
    }
    else
      {
        /* make starting population copies of loaded chromosome */
            for (j=0;j<population_size;j++)
        {
                if (j==0)
            {
                        k=0;
                        do
                        {
                                for (i=0;i<num_genes;i++)

        fscanf(fp,"%d",&genotypes[j][k][i]);
                                k++;
                        }
                        while (!feof(fp));

                        if ((k-1)!=num_chromosomes)
                        {
                                puts("ERROR. Number of chromosomes in
cgp.gen does not match the expected number");
                                puts("Check the number of chromosomes in
the .par file");
                                exit(0);
                        }
                }
                else
                {
                        for (k=0;k<num_chromosomes;k++)
                            for (i=0;i<num_genes;i++)

        genotypes[j][k][i]=genotypes[0][k][i];
                }
        }
            fclose(fp);
      }
}
```

```c
/* This calculates the limits that are used in the calculation
   of allowed gene values (alleles) */
void get_gene_limits(int column, int* limit_min, int* limit)
{
        int limit_max;

        limit_max=num_inputs+column*num_rows;
        if (column<levels_back)
                *limit_min=0;
        else
                *limit_min=num_inputs+(column-levels_back)*num_rows;

        *limit=limit_max-(*limit_min);
}


/* returns a random valid connection gene that
   obeys the constraints imposed by levels_back.
   Also allows program inputs to disobey levels_back */
int get_connection_gene(int limit_min, int limit)
{
        int limit_datas, rand_num;
        int gene;

        if (limit_min==0)
                gene = newrand(limit);
        else /* allows inputs to disobey levels_back */
        {
                limit_datas = limit+num_inputs;
                rand_num = newrand(limit_datas);
                if (rand_num<limit)
                        gene = rand_num+limit_min;
                else
                        gene = rand_num-limit;
        }

        return gene;
}


/* returns a random valid function gene */
int get_function_gene(void)
{
        return allowed_functions[newrand(num_functions)];
}


/* returns a random valid output gene */
int get_output_gene(void)
{
        int limit_min,limit;
        int output_gene;

        limit_min=num_inputs+(num_cols-levels_back)*num_rows;
        limit=levels_back*num_rows;

    output_gene = newrand(limit)+limit_min;

        return output_gene;
}

/* Calculates output of node in 8-bit format */
unsigned  char node_type(unsigned char in[MAX_NUM_GENES_PER_NODE],
```

```c
                       unsigned code)
{
    unsigned char result;

    if (code==0)      /* constants */
        result=0;
    else if (code==1)
        result=MAXNUM;
    else if (code==2) /* wire and inverter */
        result=in[0];
    else if (code==3)
        result=in[1];
    else if (code==4)
        result=~in[0];
    else if (code==5)
        result=~in[1];
    else if (code==6) /* two input gate functions */
        result=(in[0] & in[1]);
    else if (code==7)
        result=(in[0] & ~in[1]);
    else if (code==8)
        result=(~in[0] & in[1]);
    else if (code==9)
        result=(~in[0] & ~in[1]);
    else if (code==10)
        result=(in[0]^in[1]);
    else if (code==11)
        result=(~in[0]^in[1]);
    else if (code==12)
        result=(in[0] | in[1]);
    else if (code==13)
        result=(in[0] | ~in[1]);
    else if (code==14)
        result=(~in[0] | in[1]);
    else if (code==15)
        result=(~in[0] | ~in[1]);
    else if (code==16)  /* mux functions */
        result=((in[0] & ~in[2]) | (in[1] & in[2]));
    else if (code==17)
        result=((in[0] & ~in[2]) | (~in[1] & in[2]));
    else if (code==18)
        result=((~in[0] & ~in[2]) | (in[1] & in[2]));
    else if (code==19)
        result=((~in[0] & ~in[2]) | (~in[1] & in[2]));

    return result;
}

/* Decodes a single cgp chromosome and executes it on a part.
*/
unsigned char output_of_chromosome_on_one_part_on_one_image(int*
chromosome, int part, int image)
{
    register int i,k;
    unsigned int index,function_type;
    int count;
    /* unsigned char output_of_cgp_program; */
    unsigned char in[MAX_NUM_GENES_PER_NODE];
```

```
        /* load data_inputs into output */

        for (i=0;i<num_inputs;i++)
                output[i]=data_inputs[image][part][i];

    count=num_inputs;
    index=0;

        /* process nodes */
    for (k=0;k<num_nodes;k++)
    {
            for (i=0;i<num_genes_per_node-1;i++) /* get input data to
node */
                    in[i]=output[chromosome[index+i]];

        function_type=chromosome[index+num_genes_per_node-1];

        output[count]=node_type(in,function_type);

        count++;

        index=index+num_genes_per_node;
    }
        /* process outputs */
    for (i=0;i<num_outputs;i++)
    {
        output[count]=output[chromosome[index]];
        index++;
        count++;
    }

    /* check whether the cgp produce a number above or below a
threshold
        indicating the presence or absence of malignancy

    fit =0;
    for (i=0;i<num_outputs;i++)
    {
            output_of_cgp_program = output[end_count+i];

            if (malignant == 1)
            {

                if (output_of_cgp_program >= malignancy_threshold)
                    fit++;
            }
            else
            {
                if (output_of_cgp_program < malignancy_threshold)
                    fit++;

            }
    }
*/

    /* divide the output of cgp (which is in range 0 to 255) by 4
        to obtain valid white pixel counts per part */
    return (output[count-1]+3);
}
```

```c
/* get the output of cgp fro one part of one image and calculate
   a fitness according to how close the output is to the
   white pixel count in the image part */
unsigned char fitness_of_chromosome_on_one_part_on_one_image(int*
chromosome, int part, int image)
{
      unsigned char outvalue;
      unsigned char fit;

    outvalue =
output_of_chromosome_on_one_part_on_one_image(chromosome, part, image);

      /*
    if (linear_fitness)
          fit = 1.0 - (double)abs(number_white_pixels[image][part]-
outvalue)/64.0;
      else
          fit   = 1.0/(1.0 +
(double)abs(number_white_pixels[image][part]-outvalue));
    */
      /* NEW FITNESS METHOD */
    if ((outvalue >= 0) && (outvalue <= 43))
    {
          fit = 0;

    }
      else if ((outvalue > 43) && (outvalue <= 2*43))
    {
          fit = 1;

    }
      else if ((outvalue > 2*43) && (outvalue <= 3*43))
        {
              fit = 2;

        }
      else if ((outvalue > 3*43) && (outvalue <= 4*43))
        {
              fit = 3;

        }
      else if ((outvalue > 4*43) && (outvalue <= 5*43))
        {
              fit = 4;

        }
      else if ((outvalue >5*43) && (outvalue <= 6*43))
        {
              fit = 5;

        }

      /*
      printf("\nnumber_white_pixels[%d]
is %u\n",part,number_white_pixels[part]);
      printf("\noutvalue is %u\n",outvalue);

      getchar();
      */
```

```c
        return fit;
}




/* exceutes the genotype of an image to determine fitness
   Note in training an image is a small collection of parts
   of an image
*/
unsigned char fitness_of_genotype_on_one_image(int** genotype, int
image)
{

        int part,i;
        int fit = 0, final_fit=0;
        int number[5];
        int *pmax;


        /*
        printf("\nIn fitness_of_genotype_on_one_image:\n");
      printf("\nnum_parts_per_image[%d] is %d\n",image,
num_parts_per_image[image]);
        getchar();
      */

        for (part = 0;part < num_parts_per_image[image]; part++)
        {
            fit =
fitness_of_chromosome_on_one_part_on_one_image(genotype[0],part,
image);
            printf("\n fit for part %d is %d\n",part, fit);
            if (fit = 0)
            {
                number[0]++;
            }
            else if (fit = 1)
            {
                number[1]++;
            }
            else if (fit = 2)
            {
                number[2]++;
            }
            else if (fit = 3)
            {
                number[3]++;
            }
            else if (fit = 4)
            {
                number[4]++;
            }
            else if (fit = 5)
            {
                number[5]++;
            }
            /*
      printf("\npart is %d, fitness is %8.6lf\n",part,fit);
            getchar();
            */
```

224

```
        }
        *pmax = number[0];
        for (i=0; i<6;i++)
        {
              if (*pmax < number[i]);
              {
                  *pmax = number[i];
              }
         final_fit = i;
        }




//      return fit/(double)num_parts_per_image[image];
        return final_fit;
}


/* This runs all the chromosomes on all parts
   of a sequence of images (in .dat format)
*/
int fitness_get_best_genotype(int** genotype,int** best_genotype)
{
        int i;
        int total_fitness = 0.0;


        /*
        printf("\nIn fitness: num_images is %d\n", num_images);
        getchar();
        */

        for (i = 0 ; i < num_images; i++)
        {
               /* fitness on image is scaled between 0 and 1.0 */
//         total_fitness = total_fitness +
fitness_of_genotype_on_one_image(genotype,i);
             total_fitness = total_fitness + get_best_genotype(genotype,
best_genotype);
        }

        /* final fitness is scaled by number of images
           this means that a perfect fitness is 1.0 */
        return (int)(total_fitness/num_images);

}

/******** TESt versions of the above fitness functions */

/* Decodes a single cgp chromosome and executes it on a part.
*/
unsigned char test_output_of_chromosome_on_one_part_on_one_image(int*
chromosome, int part, int image)
{
   register int i,k;
   unsigned int index,function_type;
   int count;
   /* unsigned char output_of_cgp_program; */
```

```c
    unsigned char in[MAX_NUM_GENES_PER_NODE];


    /* load data_inputs into output */

    for (i=0;i<num_inputs;i++)
            output[i]=test_data_inputs[image][part][i];

    count=num_inputs;
    index=0;

      /* process nodes */
    for (k=0;k<num_nodes;k++)
    {
            for (i=0;i<num_genes_per_node-1;i++) /* get input data to
node */
                    in[i]=output[chromosome[index+i]];

        function_type=chromosome[index+num_genes_per_node-1];

        output[count]=node_type(in,function_type);

        count++;

        index=index+num_genes_per_node;
    }
      /* process outputs */
    for (i=0;i<num_outputs;i++)
    {
        output[count]=output[chromosome[index]];
        index++;
        count++;
    }

    /* check whether the cgp produce a number above or below a
threshold
        indicating the presence or absence of malignancy

    fit =0;
    for (i=0;i<num_outputs;i++)
    {
            output_of_cgp_program = output[end_count+i];

            if (malignant == 1)
            {

                    if (output_of_cgp_program >= malignancy_threshold)
                            fit++;
            }
            else
            {
                    if (output_of_cgp_program < malignancy_threshold)
                            fit++;

            }
    }
*/

    /* divide the output of cgp (which is in range 0 to 255) by 4
        to obtain valid white pixel counts per part */
```

```c
        return (output[count-1]+1)/4;
}

/* get the output of cgp fro one part of one image and calculate
   a fitness according to how close the output is to the
   white pixel count in the image part */
double test_fitness_of_chromosome_on_one_part_on_one_image(int*
chromosome, int part, int image)
{
        unsigned char outvalue;
        double fit = 0.0;

     outvalue =
test_output_of_chromosome_on_one_part_on_one_image(chromosome, part,
image);

     if ((outvalue == 0) && (test_number_white_pixels[image][part]==0))
             fit = 1.0;
        else if ((outvalue != 0) &&
(test_number_white_pixels[image][part]!=0))
             fit = 1.0;
        else if ((outvalue == 0) &&
(test_number_white_pixels[image][part]!=0))
           fit = -1.0;
        else if ((outvalue != 0) &&
(test_number_white_pixels[image][part] ==0))
             fit = -1.0;

        //printf("\ntest_number_white_pixels[%d]
is %u\n",part,test_number_white_pixels[part]);
        //printf("\noutvalue is %u\n",outvalue);

        //getchar();

        return fit;
}




/* exceutes the genotype of an image to determine fitness
  Note in training an image is a small collection of parts
  of an image
*/
double test_fitness_of_genotype_on_one_image(int** genotype, int image)
{

        int part;
        double fit = 0;

        /*
        printf("\nIn fitness_of_genotype_on_one_image:\n");
      printf("\nnum_parts_per_image[%d] is %d\n",image,
num_parts_per_image[image]);
        getchar();
      */

        for (part = 0;part < test_num_parts_per_image[image]; part++)
        {
```

```
            fit = fit +
test_fitness_of_chromosome_on_one_part_on_one_image(genotype[0],part,
image);

            /*
        printf("\npart is %d, fitness is %8.6lf\n",part,fit);
            getchar();
            */
    }


    return fit/(double)test_num_parts_per_image[image];
}

/* This runs all the chromosomes on all parts
   of a sequence of images (in .dat format)
*/
double test_fitness(int** genotype)
{
    int i;
    double total_fitness = 0.0;

    /*
    printf("\nIn fitness: num_images is %d\n", num_images);
    getchar();
    */

    for (i = 0 ; i < test_num_images; i++)
    {
        /* fitness on image is scaled between 0 and 1.0 */
        total_fitness = total_fitness +
test_fitness_of_genotype_on_one_image(genotype,i);
    }

    /* final fitness is scaled by number of images
       this means that a perfect fitness is 1.0 */
    return total_fitness/test_num_images;
}


/* prints out a matrix showing the output of the CGP programs for each
part.
   It shows original white pixels, cgp output and the difference
   It does this for each image.
*/
void fprint_best_genotype_output(int**  best_genotype, char name[],
char imagename[], int image)
{
    int part;
    unsigned char outvalue;
    int fitness,fit;
    int count[5]={0,0,0,0,0};

    FILE* fp;

    fp = fopen(name,"w");


    fprintf(fp,"White pixel values for image %s\n",imagename);
```

228

```c
        fprintf(fp,"malignancy level for image %s is %d\n", imagename,
malignancy_level);

//      for (part = 0; part < num_parts_per_image[image];part++)
//      {
//            fprintf(fp, "%4u",number_white_pixels[image][part]);
//      }
//      fprintf(fp,"\n");

        fprintf(fp,"\n\nCGP output values for best genotype on
image %s\n",imagename);

        for (part = 0; part < num_parts_per_image[image];part++)
        {
            outvalue =
output_of_chromosome_on_one_part_on_one_image(best_genotype[0],part,
image);
            fitness =
fitness_of_chromosome_on_one_part_on_one_image(best_genotype[0],part,
image,&count[5]);

            fprintf(fp, "%4u",fitness);
        }
        fit = fitness_of_genotype_on_one_image(best_genotype[0], image);
        fprintf(fp,"\n\nmalignancy_level for best genotype on image %s
is %d\n",imagename, fit);
        fprintf(fp,"\n");
    fprintf(fp,"");
        fclose(fp);

}

/* prints out a matrix showing the output of the CGP programs for each
part.
   It shows original white pixels, cgp output.
   It does this for each test image.
*/
void fprint_best_genotype_output_on_test_image(int**  best_genotype,
char name[], char imagename[], int image)
{
        int x, y;
        int part;
        unsigned char outvalue;

        FILE* fp;

        fp = fopen(name,"w");


        fprintf(fp,"White pixel values for image %s\n",imagename);

        part = 0;
        for (y = 0; y < num_yparts; y++)
        {
            for (x = 0; x < num_xparts; x++)
            {
                fprintf(fp,
"%4u",test_number_white_pixels[image][part]);
                part++;
            }
```

```
                fprintf(fp,"\n");
        }

        fprintf(fp,"\n\nCGP output values for best genotype on
image %s\n",imagename);

        part = 0;
        for (y = 0; y < num_yparts; y++)
        {
                for (x = 0; x < num_xparts; x++)
                {
                        outvalue =
test_output_of_chromosome_on_one_part_on_one_image(best_genotype[0],pa
rt, image);
                        part++;

                        fprintf(fp, "%4u",outvalue);
                }
                fprintf(fp,"\n");
        }

        fclose(fp);

}



/* creates initial population of chromosomes
   either having been generated from a single
   chromosome from a file or by generating
   an entire random population
*/
void initialise(int*** genotypes)
{
    int  j,k, pop_member,row,col;
      int  count;
    int  limit=0,limit_min=0;

    if (run_from_chrom)
    {
            read_from_chrom(genotypes);
    }
    else  /* generate random population */
    {
            for (pop_member = 0; pop_member < population_size;
pop_member++)
            {
                    for (k = 0; k < num_chromosomes; k++)
                    {
                            count = 0;
                            for (col=0;col<num_cols;col++)
                            {
                                    get_gene_limits(col,&limit_min,&limit);

                                    for (row=0;row<num_rows;row++)
                                    {
                                            /* connection genes */
                                            for (j=0;j<num_genes_per_node-
1;j++)
```

```c
        genotypes[pop_member][k][count+j]=get_connection_gene(limit_min,
limit);

                                        /* function gene */

        genotypes[pop_member][k][count+num_genes_per_node-
1]=get_function_gene();
                                        count=count+num_genes_per_node;
                                }
                        }
                        for (j=0;j<num_outputs;j++)

        genotypes[pop_member][k][count+j]=get_output_gene();
                } /* k loop */


        } /* pop_member loop */
    } /* generate random population */
} /* end of function */



/* calculate best population fitness and the best genotype */
double  get_best_genotype(int*** genotypes,
                          int**  best_genotype)

{
    int   i, k,j;
    double  fitness_min;
    int output;
    int     best_member;
      double fit;

    fitness_min=5.0;
    best_member=0;

    for (i = 0; i < population_size; i++)
    {
        /* printf("\npop member is %d\n",i); */
        fit =fabs((double)malignancy_level-
fitness_of_genotype_on_one_image(genotypes[i],i));


        if (fit < fitness_min)
        {
            fitness_min = fit;
            best_member = i;
        }

        if (fit == perfect_fitness) /* we solved the problem, so
stop */
            break;
    }

    /* here is the best chromosome */
    for (k = 0;k < num_chromosomes; k++)
    {
        for (i = 0; i < num_genes; i++)
            best_genotype[k][i]=genotypes[best_member][k][i];
```

```c
	}
	return fitness_min;
}




/* checks to see if the gene is not an output gene */
int is_not_output_gene(int gene)
{
	return (gene < num_genes_per_node*num_nodes);
}

/* checks to see if the gene is a function gene */
int is_function_gene(int gene, int locus)
{
	return (is_not_output_gene(gene) && (locus==(num_genes_per_node-
1)));

}

/* calculates how many mutations to do per chromosome */
int get_num_mutant(int num_genes, double per_cent_mutate)
{
	return (int)(num_genes*per_cent_mutate/100.0);
}

/* carry out one mutation on one chromosome */
void mutate(int*  chromosome)
{
	int which_gene,which_locus;
	int limit,limit_min;
	int col;

	which_gene=newrand(num_genes);
	which_locus=which_gene % num_genes_per_node;

	if (is_not_output_gene(which_gene))
	{
		if (is_function_gene(which_gene,which_locus))
		{
			if (num_functions == 1) /* redirect the mutation to
a connection */
			{
				which_locus=newrand(num_genes_per_node-1);
				which_gene=which_gene-num_genes_per_node-
1+which_locus;
			}
			chromosome[which_gene]=get_function_gene();
		}
		else /* it is a connection gene */
		{
			col = which_gene/(num_genes_per_node*num_rows);

			get_gene_limits(col,&limit_min,&limit);


	chromosome[which_gene]=get_connection_gene(limit_min,limit);
		}
	}
```

```
                else /* it is an output gene */
                {
                        chromosome[which_gene]=get_output_gene();
                }
}

/* find out how many mutations to do and mutate the chromosome */
void mutate_chromosome(int*  chromosome)
{
        int i;
        int num_mutations;

        num_mutations = get_num_mutant(num_genes,per_cent_mutate);

        for ( i = 0; i < num_mutations; i++)
        {
                mutate(chromosome);
        }

}



void mutate_genotype(int** genotype)
{
        int k;

        for (k = 0; k < num_chromosomes; k++)
                mutate_chromosome(genotype[k]);
}


/* (1+lambda evolutionary strategy where lamda = population size -1
   This defines the EA algorithm note that using mu=1 and
   lambda = population_size -1
   should give the same.
*/
void generate_new_pop_es(int*** genotypes,
                                        int**  best_genotype)
{
        int i,j,k;

    /* copy best_chromosome into last member of chromosome array */
     for (k = 0; k < num_chromosomes; k++)
     {
                for (j = 0; j < num_genes; j++)
                        genotypes[population_size-
1][k][j]=best_genotype[k][j];
        }
    /* generate new population by mutating all but last */
    for (i=0;i<population_size-1;i++)
    {
                for (k = 0; k < num_chromosomes; k++) /* copy all
chromosomes in genotype */
                {
                        for (j = 0; j < num_genes; j++) /* copy  best
genotype */
                                genotypes[i][k][j]=best_genotype[k][j];
                }
                /* mutate the genotype */
                mutate_genotype(genotypes[i]);
```

```c
    }
}


/* allocate space for a single chromosome */
int* create_chromosome_space(void)
{
      int* chromosome = NULL;

      chromosome=(int*)calloc(num_genes,sizeof(int));
    if (chromosome==NULL)
    {
            printf("ERROR.Not enough memory for a chromosome of this
length\n");
        exit(0);
    }
      return chromosome;
}

/* allocate space for a collection of chromosomes */
int** create_genotype_space(void)
{
      int k;
      int **genotype = NULL;

      genotype =(int** )calloc(num_chromosomes,sizeof(int*)); /*
create space for pointers to int pointers */
      if (genotype==NULL)
      {
       printf("ERROR.Can not allocate space for this many genotype
pointers\n");
       exit(0);
      }

      for (k = 0;k < num_chromosomes; k++) /* create array of pointers
to ints (genes) */
      {
        genotype[k] = create_chromosome_space();
      if (genotype[k]==NULL)
      {
        printf("ERROR.Not enough memory for genotypes of this
length\n");
        exit(0);
      }
   }

      return genotype;
}

/* allocate space for population of genotypes */
int*** create_genotypes_space(void)
{
      int i;
      int ***genotypes = NULL;

      genotypes = (int*** )calloc(population_size,sizeof(int**)); /*
create space for pointers to int pointers */
      if (genotypes == NULL)
      {
```

```c
        printf("ERROR.Can not allocate space for this many chromosome
pointers to pointers\n");
        exit(0);
      }

      for (i=0;i<population_size;i++) /* create array of pointers to
ints (genes) */
      {
          genotypes[i] = create_genotype_space();
          if (genotypes[i]== NULL)
          {
                      printf("ERROR.Not enough memory for a
genotype %d of this length\n",i);
                      exit(0);
          }
      }

      return genotypes;
}


/* release memory */
void free_genotypes(int*** genotypes)
{
      int i;

      for (i = 0; i < population_size; i++)
      {
          free_genotype(genotypes[i]);
      }
      free(genotypes);
}

/* release memory */
void free_genotype(int** genotype)
{
      int k;

      for (k = 0; k < num_chromosomes; k++)
          free_chromosome(genotype[k]);

      free(genotype);
}

/* release memory */
void free_chromosome(int* chromosome)
{
      free(chromosome);
}

/* Do a run of the EA */
double  EA(int* best_gen, int run,
          char prog[MAX_NUM_LETTERS], char stat[MAX_NUM_LETTERS])
{
      int         i,gen_index;
      double      best_fit=-1e10, previous_best_fit=-1e10;
      double  test_fit = 0.0;
      char  filename[MAX_NUM_LETTERS];
      int***      genotypes;
      int** best_genotype;
```

```c
        FILE* fp;

        genotypes = create_genotypes_space();
        best_genotype = create_genotype_space();

        initialise(genotypes);

        for (gen_index=1;gen_index<=num_generations;gen_index++)
        {
                if (gen_index % report_interval==0)
                        printf("\nGENERATION is %d",gen_index);

                /* find new best chromosome and its fitness */
                /* printf("\nJust before get_best_genotype\n"); */

                best_fit = get_best_genotype(genotypes,best_genotype);
                test_fit = test_fitness(best_genotype);


        /*      printf("\nJust after get_best_genotype\n"); */

                if (best_fit > previous_best_fit) /* we have an
improvement */
                {
                        if ((progress_report > 0) && (progress_report < 3))
                        {
                        fp=fopen(prog,"a");
                        printf("\nGENERATION is %u Best fitness is
now %10.8lf",gen_index,best_fit);
                        printf("\nGENERATION is %u Best test_fitness is
now %10.8lf",gen_index,test_fit);

                        fprintf(fp,"\nGENERATION is %u    Best fitness is
now %10.8lf",gen_index,best_fit);
                        fprintf(fp,"\nGENERATION is %u    Test fitness is
now %10.8lf",gen_index,test_fit);

                                fprintf(fp,"\nThe chromosome is\n");
                        fclose(fp);

                                if (progress_report == 1)

        fprint_a_user_readable_genotype(best_genotype,prog,1);

                        }

                        if (progress_report == 3)
                        {
                                fp=fopen(stat,"a");

        fprintf(fp,"%d\t%8.6lf\t%8.6lf\n",gen_index,best_fit,test_fit);
                                fclose(fp);
                        }

                        *best_gen = gen_index;
                        previous_best_fit = best_fit;
                }
                /* jump out of run if maximum fitness acheived */
                if (best_fit == perfect_fitness)
                {
```

```c
                break;
                gen_index++;
        }
        else
        {
                generate_new_pop_es(genotypes,best_genotype);
        }
    }


    fp=fopen("cgp.txt","a");
    fprintf(fp,"Run %d and gen %d acheived
fitness %10.8lf\n",run,*best_gen,best_fit);
    fprintf(fp,"Here is the chromosome\n");
    fclose(fp);

    fprint_a_user_readable_genotype(best_genotype,"cgp.txt",1);

    /*

    fprint_best_genotype_classification(best_genotype,"cgpbest-
classification.txt");
    */

    /* write output of cgp for all images to files */
    for (i = 0; i < num_images; i++)
    {
            sprintf(filename,"cgpbest-output-run%d-image%d.txt",run,
i);
            fprint_best_genotype_output(best_genotype, filename,
image_names[i], i);
    }

    /* get testing image names. These are complete dat files */
    get_image_names(0);

    /* read in image data for all testing images */
    printf("\nReading testing image data\n");
    for (i = 0; i < num_images; i++)
    {
            read_data(image_names[i],i);
    }


    /* write output of cgp for all images to files */
    for (i = 0; i < num_images; i++)
    {
            sprintf(filename,"cgpbest-output-run%d-
testimage%d.txt",run, i);
            fprint_best_genotype_output_on_test_image(best_genotype,
filename, image_names[i], i);
    }


    fprint_a_genotype(best_genotype,"cgp.gen",0);

    free_genotypes(genotypes);

    free_genotype(best_genotype);
```

```c
        return best_fit;
}


/* do mutiple runs of EA and write out results */
void run_EA(int num_runs_total)
{
        int             j;
        int             best_gen,run;
        double          worst_of_best_fit=1.0e10,best_of_best_fit=-1e10;
        double          fitness_final=0.0;
        char    prog[20],stat[20],runstring[10];
        double          fitnesses[1000], temp;
        double          av_fitness=0.0,av_best_gen=0.0,st_dev=0.0;
        FILE* best;
        FILE* fp;

        for (run=0;run<num_runs_total;run++)
        {

                sprintf(runstring,"%d",run); /* store run as characters */
                printf("\n\nRUN %d\n",run);

                if (progress_report>0)
                {
                        strcpy(prog,"cgp");
                        strcat(prog,runstring);

                        strcpy(stat,prog);
                        strcat(prog,".prg"); /* create .prg file name */
                        strcat(stat,".txt"); /* create .txt file name */

                        fp=fopen(prog,"w");  /* create empty .prg file */
                        fclose(fp);

                        fp=fopen(stat,"w");
                        fprintf(fp,"\nRUN %d\n\n",run);
                        fprintf(fp,"Gen\tfit\ttest fitness\n",run);
                        fclose(fp);
                }

                fitness_final =  EA(&best_gen,run,prog,stat);

            if (fitness_final < worst_of_best_fit)
                        worst_of_best_fit=fitness_final;

                if (fitness_final > best_of_best_fit)
                {
                        best_of_best_fit=fitness_final;

                }

                fitnesses[run]=(double)fitness_final;
                av_fitness=av_fitness+(double)fitness_final;
                if (fitness_final == num_chromosomes)
                        av_best_gen=av_best_gen+(double)best_gen;
        }

    av_fitness=av_fitness/((double)num_runs_total);
```

```c
        av_best_gen=av_best_gen/((double)num_runs_total);

        st_dev=0.0;
        for (j=0;j<num_runs_total;j++)
        {
                temp=(fitnesses[j]-av_fitness);
                temp=temp*temp;
                st_dev=st_dev+temp;
        }

        st_dev=st_dev/(double)num_runs_total;
        st_dev=sqrt(st_dev);

        best=fopen("cgp.txt","a");
        fprintf(best,"\naverage fitness  %6.4lf\n",av_fitness);
        fprintf(best,"\nstd dev          %6.4lf\n\n",st_dev);
        fprintf(best,"\nThe best solution of all runs
is  %6.2lf\n",best_of_best_fit);
        fprintf(best,"\nThe worst solution of all runs
is  %6.2lf\n",worst_of_best_fit);
        fprintf(best,"\nOf perfect solutions, the average number of
generations is  %6.4lf\n",av_best_gen);
        fclose(best);

}
```

List of Abbreviations

CGP   Cartesian Genetic Programming

MCGP   Multi-chromosomes Cartesian Genetic Programming

CAD   Computer Aided Detection

CT   Computer Tomography

CC   CranioCaudal

DDSM   Digital Database for Screening Mammography

DCIS   Dutual Carcinoma In Situ

ES   Evolutionary Strategies

EA   Evolutionary Algorithm

EC   Evolutionary Computation

EP   Evolutionary Programming

FN   False Negative

FP   False Positive

GA   Genetic Algorithm

GP   Genetic Programming

ISO   InferoSuperior Oblique

LLNL   Lawerance Livermore National Laboraty

LCIS   Lobular Carcinoma In Situ

MRI   Magnetic Resonance Imaging

MIAS   Mammography Imaging Analysis Society

MLO   MedioLateral Oblique

ML   MedioLateral

NHS   National Health Service

NPV   Negative Predictive Value

PPV   Positive Predictive Value

|ROI   Region of Interests

TN   True Negative

TNR   True Negative Rate

TP   True Positive

TPR   True Positive Rate

Reference

[1]   National statistics website. [Online]. http://www.statistics.gov.uk/cci/nugget.asp?id=575

[2]   (Accessed 10 February 2011) NHS Breast Cancer Screening Programme (2010). Annual Review 2010.   [Online].
        http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbspannualreview2010.pdf

[3]   BBC News (2011). [Online]. http://bbc.co.uk/news/uk-england-cumbria-12281456

[4]   Belfast Telegraph (2011). [Online]. http://www.belfasttelegraph.co.uk/news/localnational/northern-ireland/late-xray-hospital-shortstaffed-15073008.html

[5]   P.Hart,D.Stork,Wiley R Duda, "Pattern Classification ," *Interscience* , no. ISBN 04710566.

[6]   Health Encyclopedia. [Online]. http://www.healthscout.com/ency/68/761/main.html

[7]   Daniele Turi, Alan Williams, Katy Wolstencroft and Chris Taylor Chris Rose. (2006) DDSM: Digital Database for Screening Mammography. [Online]. http://marathon.csee.usf.edu/Mammography/Database.html

[8]   (2008) the world health organisation. [Online]. http://www.who.int/cancer/detection/breastcancer/en/

[9]   E. Siebert, et al. S.Diekmann, "Dose exposure of patients undergoing comprehensive stroke imaging by multidetector-row CT: comparison of 320-detector row and 64-detector row CT scanners," *American Journal of Neuroradiology*, pp. 1003-1009, 31(6), 2010.

[10] Patrick J. Lynch. Wikipedia. [Online]. http://en.wikipedia.org/wiki/File:Breast_anatomy_normal_scheme.png

[11] American Cancer Society, "Cancer Facts & Figures 2012.," 2012.

[12] D.B.KOPANS, *Breast Imaging*.: Lippincott William & Wilkins, Third Edition,2007.

[13] L.KAUHAVA, P.IMMONEN-R, I.PARVINEN, K.HOLLI, L.PYLKKANEN, A.KALJONEN, H.HELENIUS, P.KRONQVIST AND P.J.KLEMI, "Lower recurrence risk through mammographic screening reduces breast cancer treatment costs ," *Breast*, vol. 17(6), pp. 550-4, 2008.

[14] D. Pickuth and S.H. Heywang-Köbrunner, ""Neurosarcoidosis : Evaluation with MRI"," *Journal of Neuroradiology*, vol. Vol 27,No.3, p. p. 185, novembre 2000.

[15] Carolyn M.Kaelin,b Francesca Coltrera, *Living through breast cancer - chinese edition*. Beijing: China Renmin University Press Co.,LTD, 01-01-2008.

[16] Kelly Ford, Elizabeth Marcus, Bennett Lum, "Braest Cancer Screening, Diagnosis, and Treatment," Canada, 9 September 1999.

[17] Ramón Vilana, Jordi Bruix M.D, Concepció Bru, Carmen Ayuso, Manel Solé, Joan Rodés, "Tumor size determines the efficacy of percutaneous ethanol injection for the treatment of small hepatocellular carcinoma," *HEPATOLOGY*, vol. 16, no. 2, pp. 353-357, 1992.

[18] Virginia L. Ernster, Rachel Ballard-Barbash, William E. Barlow, Yingye Zheng, Donald L. Weaver, Gary Cutter, Bonnie C. Yankaskas, Robert Rosenberg, Patricia A. Carney, Karla Kerlikowske, Stephen H. Taplin, Nicole Urban and Berta M. Geller, "Detection of Ductal Carcinoma In Situ in Women Undergoing Screening Mammography," *Journal of The National*

Cancer Institue, vol. 94, no. 20, pp. 1546-1554, 14-08-2002.

[19] Ramsdale M L, Rust A,et al. Cooke J. Young K C, "Effect of automatic kV selection on dose and contrast for a mammographic X-ray system," *The British journal of radiology,* , vol. 70, no. 838, pp. 1036-1042, 1997.

[20] (2012) breastcancer.org. [Online]. http://www.breastcancer.org/symptoms/diagnosis/invasive.jsp

[21] Salz T, Lillie SE Brewer NT, "Systematic review: the long-term effects of false-positive mammograms," *Ann. Intern. Med*, vol. 146, no. 7, pp. 502-510, 2007.

[22] John, Jr. Conyers, David R. Obey Dr. Samuel S. Epstein, *The Politics of Cancer*., 1979.

[23] Phil B. Fontanarosa, Samuel S. Epstein, "The Politics of Cancer," *the Jounal of the American Medical Association*, vol. 284, no. 4, pp. 442-442., 2000.

[24] Anderson E, Sutt Jr E, et al. Leichti R, "Sheathing nail bending-yield strength-role in shearwall performance," in *Proceedings of the 9th World Conference on Timber Engineering*, Portland, OR, USA. 2006.

[25] J. Wittenvberg, M.Harisinghani, J.W.Chen, S.E.Jones and J.W.Patti R.Weissleder, *Primer of Diagnostic Imaging*. Mosby, 2006.

[26] NHS, "NHS Breast Screening Programme 2012 Annual Review," ISBN 978-1-84463-093-6, 2012.

[27] U.S. Preventive Services Task Force. (November, 2009) Screening for Breast Cancer. [Online]. http://www.uspreventiveservicestaskforce.org/uspstf/uspsbrca.htm

[28] U.S. food and drug adminstration. (06/02/2011) Thermographic Imaging Systems for Breast Cancer

Screening: FDA Safety Communication. [Online]. http://www.fda.gov/Safety/MedWatch/SafetyInformation/default.htm

[29] Eurostat, "Health statistics – Atlas on mortality," Belgium, 2009.

[30] Cancer Reserach Uk Primary Care Education Research Group. (2011, February) NHS Cancer Screening Programmes. [Online]. http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp.pdf

[31] (2011, January) Cancer Research UK. [Online]. http://cancerhelp.cancerresearchuk.org/type/breast-cancer/treatment/number-stages-of-breast-cancer#1

[32] American Cancer Society. (2012, August) [Online]. http://www.cancer.org/Cancer/BreastCancer/MoreInformation/BreastCancerEarlyDetection/breast-cancer-early-detection-signs-symptoms-br-ca

[33] M.D Mark Helvie, "Full Field Digital Mammography: A New Breast Cancer Screening Tool ," *Cancer news*.

[34] Sarah Vinnicombe,Valerie A. McCormack,Susan Shiel,Nick Perry,Isabel M. dos Santos Silva, "Full-Field Digital versus Screen-Film Mammography: Comparison within the UK Breast Screening Program and Systematic Review of Published Data," *Radiology*, vol. 251, no. 2, pp. 347-58, May 2009.

[35] NHSBSP Equipment Report, "COMMISSIONING AND ROUTINE TESTING OF FULL FIELD," April 2009.

[36] d Arnulf Skjennald Per Skaane, "Screen-Film Mammography versus Full-Field Digital Mammography with Soft-Copy Reading: Randomized Trial in a Population-based Screening

Program—The Oslo II Study," *Radiology*, no. 3, p. 166, 2004.

[37]   Dodd G, Costanza M, et al. Leitch A, "American Cancer Society guidelines for the early detection of breast cancer: update 1997," Canada, 1997.

[38]   S. Feig, "Increased benefit from shorter screening mammography intervals for women ages 40-49 years.," *Cross Ref Medline Web of Science*, pp. 2035-2039, 1997.

[39]   Sickles E., "Quality assurance: how to audit your own mammography practice.," 1992.

[40]   J. Lichy, and J. H. Newhouse T. M. Kolb, "Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: ananalysis of 27,825 patient evaluations," *Radiology*, vol. 225, pp. 165-75, 2002.

[41]   T. W. Wallace, and B. C. Yankaskas R. E. Bird, "Analysis of cancers missed at screening mammography," *Radiology*, vol. 184, pp. 613-617, 1992.

[42]   C.J. Vyborny, H. MacMahon, C.E. Metz, K. Doi, E.A. Sickles, H.P. Chan, "ROC studies of the effects of pixel size and unsharp-mask filtering on the detection of subtle microcalcifications, Invest.," *Radiol*, vol. 22, pp. 581-589, 1987.

[43]   Richard Eugene Woods Rafael C. González, *Digital image processing*, 3rd ed.: PEARSON, Prentice Hall, 2007.

[44]   S. Lille, K Willison V.Andolina, *Mammographic Imaging, A Practical Guide*.: J.B.LippinCott Compant=y.

[45]   R.Sankar, W.Qian, L.Zhang, "|Advances in microcalcification clusters detection in mammograms," *Computers in Biology and Medical Physics*, vol. 32, pp. 515-528, 2002.

[46] K Lam,N. Petrick,M.Helvie H.Chan, "Computerized analysis of mammgraphic microcalcifications in morphlogical and texture feature spaces," *Medical Physics*, vol. 25, no. 10, 1998.

[47] B.M. Galkin, H.D. Muir, S.A. Feig, "Evaluation of breast microcalcifications by means of optically magnified tissue specimen radiographs," *S. Brunner, B. Langfeldt (Eds.), Recent Results in Cancer Research*, vol. 105, pp. 111-123, 1987.

[48] E.A. Sickles, "Breast calcifications: mammographic evaluation," *Radiology*, vol. 160, pp. 289-296, 1986.

[49] X. Cai, X. Chen, L. Hu, X. Lou H.D. Cheng, "Computer-aided detection and classification of microcalcifications in mammograms: a survey," *Pattern Recogn*, vol. 36, pp. 2967-2991, 2003.

[50] Y. Yang, et al. I. El-Naqa, "A support vector machine approach for detection of microcalcifications," *IEEE Trans. Med. Imaging*, vol. 21, no. 12, pp. 1552-1563, 2002.

[51] Y. Yang, R.M. Nishikawa, M.N. Vernick, A. Edwards, L. Wei, "Relevance vector machine for automatic detection of clustered microcalcifications," *IEEE Trans. Med. Imaging*, vol. 24, no. 10, pp. 1552-1563, 2005.

[52] K.-Y. Li, Y.-K. Huang S.-N. Yu, "Detection of microcalcifications in digital mammograms using wavelet filter and Markov random field model," *Comput. Med. Imaging Graphics*, vol. 30, pp. 163-173, 2006.

[53] F. Rafiee-Rad, S. Pourabdollah-Nejad, H. Soltanian-Zadeh, "Comparison of multiwavelet, wavelet, Haralick, and shape features for microcalcification classification in mammograms," *Pattern Recogn*, vol. 37, pp. 1973-1986, 2004.

[54] R.M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, pp. 786-804, 1979.

[55] P,Yan,H.Zhao,and X.Zhang G.Zhang, "A computer Aided Diagnosis System in Mammography Using Artificial Neural \Networks," *BioMedical Engineering and Informatics BMEI* , pp. p823-826, volumn 2 2008.

[56] H. Fujita, J. Chen and Z.Zhang X.Zhang, "Effect of training Artificial Neural Networks on 2D image: An Example Study on Mammography," *Aritificial Intelligence and Computational Intelligence AICI*, volumn 4 2009.

[57] A.Govardhan and S.Baskaran J.S.L.Jasmine, "Microcalcification detection in digital mammograms based on wavelet analysis and neural networks," *Control, Automation, Communication and Energy Conservation* , pp. pp 1-6, 2009.

[58] G., Wang, J. Beni, "Swarm Intelligence in Cellular Robotic Systems," in *NATO Advanced Workshop on Robots and Biological Systems*, Tuscany, Italy, June 1989, pp. 26-30.

[59] Roger Alsing. (2008, December) Roger Alsing Weblog. [Online]. http://rogeralsing.com/2008/12/07/genetic-programming-evolution-of-mona-lisa/

[60] (2001, Oct.) EDGE. [Online]. http://www.edge.org/3rd_culture/mayr/mayr_print.html

[61] Charles Darwin, *The Origin of Species by Means of Natural Selection: or the Preservation of Favoured Races in the Struggle for Life.* London: John Murray, 1859.

[62] Pearson H, "Genetics: what is a gene?," *Nature*, pp. 398–401, 2006.

[63] Oxford University Press. (2013) Oxford Dictionaries. [Online]. http://www.oxforddictionaries.com/definition/english/mutation?q=mutation

[64] John R. Koza, "Genetic Programming: On the Programming of Computers by means of Natural selection," vol. 4, pp. 127-164, 1992.

[65] Thomson P. Miller J. F., "Cartesian Genetic Programming. Proceedings of the 3rd European Conference on Genetic Programming. ," *Springer LNCS 1802* , vol. 45, pp. 121-132, 2000.

[66] J. F. Miller, "An Empirical Study of the Efficiency of Learning Boolean Functions using a Cartesian Genetic Programming Approach," in *Proc. Genetic and Evolutionary Computation Conference*, 1999, pp. 1135-1142.

[67] J. F. Miller and P.Thomson, ""Cartesian Genetic Programming,"," in *Proc. European Conference on Genetic Programming, vol. 1802 of LNCS,* 2000, pp. pp. 121–132.

[68] Miller J. F. and Thomson P., "Evolving Digital Electronic Circuits for Real-Valued Function Generation using a Genetic Algorithm," in *Proceedings of the Third Annual Conference on Genetic Programming*, 1998, pp. 863-868.

[69] Dr. Julian Miller. Bio-inspired computer Lecture notes on department of Electronics. [Online]..http://www.amp.york.ac.uk/internal/ugrady/yr4/crs/bioinsp/

[70] J. F. Miller, Cavill R. J. A. Walker, "A Multi-chromosome approach to Standard and Embedded Cartesian Genetic Programming," *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO), ACM Press*, p. 903910, 2006.

[71] Julian Miller. department of Electronics/ University of York. [Online].

http://www.amp.york.ac.uk/internal/ugrady/yr4/crs/bioinsp/

[72] Walker J. A., Miller J. F., and Cavill R., "A Multi-chromosome approach to Standard and Embedded Cartesian Genetic Programming," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO),* 2006, p. pp. 903910.

[73] Völk K, Smith S. L., Miller J. F. Walker J. A., "Parallel Evolution using Multi chromosome Cartesian Genetic Programming"," *Genetic Programminmg and Evolvable Machines*, vol. Vol. 10, No. 4, pp. pp. 417-445, 2010.

[74] Smith S.L. Miller J.F., "Redundancy and Computational Efficiency in Cartesian Genetic Programming," *IEEE Transactions on Evolutionary Computation,* pp. 167-174, 2006.

[75] Miller J. F. Walker J. A., "Investigating the performance of module acquisition in Cartesian Genetic Programming. ," in *Proceedings of the 2005 conference on Genetic and Evolutionary Computation (GECCO 2005),* 2005, pp. 1649-1656.

[76] Julian. F. Miller, "Cartesian Genetic Programming," , 2011.

[77] K. Völk, S. Smith, and J. Miller J. Walker, "Parallel evolution using multi-chromosome," *Gen. Program. Evol. Mach*, vol. 10, no. 4, 2009.

[78] Merck & Co., Inc. (2009-2014) Breast Cancer. [Online]. http://www.merckmanuals.com/home/womens_health_issues/breast_disorders/breast_cancer.html

[79] Etta D. Pisano, "Current Status of Full-Field Digital Mammography," *Radiology*, vol. 214, pp. 26-28, January 2000.

[80] Felix Diekmann Ulrich Bick, *Medical Radiology, Digital Mammography*, 9783540784494th ed.

Verlag Berlin Heidelberg: Springer, 2010.

[81] Frieling JT, Hendriks JH. Karssemeijer N, "Spatial resolution in digital mammography," *PubMed.gov*, vol. 5, no. 28, pp. 413-9, 1993.

[82] F. Thibault A. Noel, "Digital detectors for mammography: the technical challenges," *EUROPEAN RADIOLOGYv*, vol. 6, no. 14, 2004.

[83] Association of European Cancer Leagues (ECL). EU Council Recommendation on Cancer Screening. [Online]. http://www.europeancancerleagues.org/cancer-in-europe/resources-on-cancer-in-europe/82-eu-council-recommendation-on-cancer-screening.html

[84] Tyne K, Naik A, et al Nelson H D, "Screening for breast cancer: an update for the US Preventive Services Task Force," *Annals of internal medicine*, pp. 151(10): 727-737, 2009.

[85] WHO Library Cataloguing-in-Publication Data, "Women and Health: today's evidence tomorrow's agenda," Switzerland, 2009. [Online]. www.who.it/

[86] Advisory Committee on Breast Cancer Screening, "Screening for breast cancer in England: past and future[J]," *Journal of Medical Screening*, vol. 13(2), pp. 59-61, 2006.

[87] Kari Tyne, Arpana Naik, Christina Bougatsos, Benjamin K. Chan,Linda Humphrey Heidi D. Nelson, "Screening for Breast Cancer: Systematic Evidence Review Update for the U. S. Preventive Services Task Force," *Annals of Internal Medicine*, vol. 10, no. 151, pp. 727-737, 17 November2009.

[88] Porter PL, Lehman C, et a Buist DS, "Factors contributing to mammography failure in women aged 40-49 years," *J Natl Cancer Inst*, vol. 19, no. 96, pp. 1432-40, 2004.

[89] Edward F. Patz Jr., "Lung Cancer Screening, Overdiagnosis Bias, and Reevaluation of the Mayo Lung Project," *Journal of the national cancer institution*, vol. 11, no. 98, pp. 724-725, 2006.

[90] American College of Radiology, "Illustrated Breast Imaging Reporting and Data System (BI-RADS)," 1998.

[91] N.T. Ridleyb, G. Rubinc, M.G. Wallisd, F.J. Gilberte,M.J. Michell A.J. Maxwella, "The Royal College of Radiologists Breast," *Clinical Radiology*, vol. 3, no. 64, pp. 624-627, 2009.

[92] K.Bowyer, D.Kopand and W.P Kegelmeyer M Heath, "The digital database for screeening mammography," *5th INternational workshop on Digital mammography*, pp. 212-218, 2001.

[93] Güld M, Thies C, et al. Lehmann T, "he irma project: A state of the art report on content-based image retrieval in medical applications," *Korea-Germany Workshop on Advanced Medical Image*, pp. 161-171, 2003.

[94] Staples J J, Hussey D H, et al. Mayr N A, "Morphometric studies in intraductal breast carcinoma using computerized image analysis[J]," *Cancer*, vol. 67(11), pp. 2805-2812, 1991.

[95] J. Miller, and S. Smith K. Völk, "Multiple network CGP for the classification of," Berlin, 2009.

[96] Prabir Burman, "Comparative Study of Ordinary Cross-Validation,v-Fold Cross-Validation and the Repeated Learning-Testing Methods," *Biometrika*, vol. 76, no. 3, pp. 503-514, 1989.

[97] James Alfred Walker, Julian F. Miller Stephen L. Smith, "Medical Applications of Cartesian Genetic Programming," in *Cartesian Genetic Programming*. York: Springer Link, 2011, pp. 309-336.

[98] J.A.WALKER, J.F.MILLER J.Clegg, "A new crossover technique for Cartesian Genetic Programming," in *GECCO '07: Proceedings of the 9th annual conference on Genetic and evolutionary computation*, ACM Press, London, 2007.

[99] J.F.Miller, *Cartesian Genetic Programming-Natural Computing Series*. York: Springer, 2011.

[100] J.F.Miller and S.Smith, "Redundancy and computational efficiency in Cartesian genetic," *IEEE Trans. Evol. Comput*, vol. 10, no. 2, pp. 167-174, 2006.

[101] Job D., Vassilev V.K. Miller J.F., "Principles in the Evolutionary Design of Digital Circuits - Part I," *Genetic Programming and Evolvable Machines*, vol. 1, no. 1, pp. 8-35, 2000.

[102] C.E. Metz, "ROC methodology in radiologic imaging," *Invest. Radiol.*, vol. 21, no. 9, pp. 720-733, 1986.

[103] L.B. Lusted, "Decision-making studies in patient management," *N. Engl. J. Med*, vol. 284, pp. 416-424, 1971.

[104] L.B. Lusted, "Decision-making studies in patient management," *N. Engl. J. Med*, vol. 284, pp. 416-424, 1971.

[105] "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, pp. 861-874, 2006.

[106] M.R. Islam1 and M.Z. Ali Sifuzzaman1, "Application of Wavelet Transform and its Advantages Compared to Fourier Transform M," *Journal of Physical Sciences*, vol. 13, no. 2, pp. 121-134, 2009.

[107] Christopher Booth, "Classification of microcalcifications," 2010.

[108] Thacker N, Clark A F. Courtney P, "Algorithmic modelling for performance evaluation[J].," *Machine Vision and Applications.*, vol. 9(5):, pp. 219-228, 1997.

[109] Julian F., Stephen L. Smith, and Yuan Zhang Miller, "Detection of microcalcifications in mammograms using multi-chromosome cartesian genetic programming. ," in *Genetic and evolutionary computation*, In Proceedings of the 12th annual conference companion.

[110] Christopher Booth, "Image analysis of mammograms towards the classification of breast cancer," Department of Electronics, University of York,.

[111] P. Courtney, N. Thacker, and A.F. Clark, "Algorithmic modeling for performance evaluation," in *Machine Vision and Applications*.: Springer, 1997, pp. pp.219-228.

[112] (2011, January) Cancer Research UK. [Online]. http://cancerhelp.cancerresearchuk.org/type/breast-cancer/treatment/number-stages-of-breast-cancer

[113] K. Bowyer,D. Kopans,R. Moore, P. Kegelmeyer Jr M. Heath, "THE DIGITAL DATABASE FOR SCREENING MAMMOGRAPHY," , Toronto,Canada, 2000.

[114] "Digital mammography," *Maidment AD*, no. 38, pp. 216-230, 2003.

[115] Mahesh M, "AAPM/RSNA physics tutorial for residents: digital mammography: an overview," *Radiographics*, no. 24, pp. 1747–1760, 2004.

[116] International Agency for Reseach on Cancer, "Cancer screening in the European Union-Report on the implementation of the Council Recommendation on cancer screening," 2008.

[117] " Breast Cancer Study," Washington, DC, 1995.

[118] B.F. Vainio H, "IARC Handbooks of Cancer Prevention: Breast Cancer Screening," vol. 7, 2002.

[119] American Cancer Society. (2011) Breast Cancer Facts & Figures 2011-2012. [Online]. http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-030975.pdf

[120] Kevin Bowyer, Daniel Kopans, Richard Moore and W. Philip Kegelmeyer Michael Heath, "The Digital Database for Screening Mammography," in *Proceedings of the Fifth International Workshop on Digital Mammography*. ISBN 1-930524-00-5: Medical Physics Publishing, 2001, pp. 212-218.

[121] IRMA (Image Retrieval in Medical Applications). [Online]. http://www.irma-project.org

[122] J.J.STAPLES,R.A.ROBINSON, J.E.VANMETRE AND D.H.HUSSEY N.A.MAYR, "Morphometric studies in intraductal breast carcinoma using computerized image analysis," *Cancer*, vol. 11, pp. 2805-12, 1991.

[123] U.Bick and F.Diekmann, "MEDICAL RADIOLOGY-Digital Mammography," Berlin, 2010.

[124] Cancer Research UK. [Online]. http://www.cancerresearchuk.org/

[125] D.PICKUTH, "klinische Radiologie systematisch," in *Uni-Med*, Bremen, 1998.

[126] MEDSCAPE. [Online]. http://www.medscape.com/

[127] Breastcancer.org. [Online]. http://www.breastcancer.org/symptoms/testing/new_research/20091203c.jsp

[128] UK Cancer Screening Programmes, "NHS breast screening ," United Kingdom, Feburary 2012.

[129] Jinfeng Li, *RUHE YINGDUI RUXIANAI.* ISBN:97887111195368: China Machine Press, 01-07-2006, p. 177.

[130] Ozekes S, Osman O, Camurcu AY, "Mammographic Mass Detection Using a Mass Template," *Korean J Radiol*, Oct-Dec 2005.

[131] J. Suckling, J. Parker, D.R. Dance, S.M. Astley, I. Hutt, C.R.M. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S.L. Kok, P. Taylor, D. Betal, J. Savage,.: Int. Work. Dig. Mammography, 1994, pp. pp.211–221.

[132] S. Smith, and E. Munday D. Hope, "Evolutionary algorithms in the classification of," *Computational Intelligence in Image and Signal Processing*, vol. in Proc, no. IEEE Symp, pp. 258-265, 2007.

[133] G. Lancaster and J, Manning D. Sutt, "Breast asymmetry and predisposition to nreast cancer," *Breast cancer research*, vol. 8, no. 2, 2006.

[134]

[135] Nielsen M Gtzsche PC, ""Screening for breast cancer with mammography".," *CochraneDatabaseSystRev*, no. doi:10.1002/14651858.CD001877.pub2.PMID 17054145., 2006.

[136] R PANCHAL and B VERMA, "A Fusion of Neural Network Based Autoassociator and Classifier for the Classification of Microcalcification Patterns," *ICONIP*, pp. 794-799, 2004.

[137] P. F., Green, R. E., & Heath, M. F. (2001). Agricultural intensification and the collapse of Europe's farmland bird populations. Donald, *Proceedings of the Royal Society of London. Series B:*

Biological Sciences, , pp. 25-29.