IMAGE UNDERSTANDING OF MOLAR PREGNANCY BASED ON ANOMALIES DETECTION

PATISON PALEE

A thesis submitted in partial fulfilment of the requirement of Staffordshire University for the degree of Doctor of Philosophy

May 2015

Abstract

Cancer occurs when normal cells grow and multiply without normal control. As the cells multiply, they form an area of abnormal cells, known as a tumour. Many tumours exhibit abnormal chromosomal segregation at cell division. These anomalies play an important role in detecting molar pregnancy cancer.

Molar pregnancy, also known as hydatidiform mole, can be categorised into partial (PHM) and complete (CHM) mole, persistent gestational trophoblastic and choriocarcinoma. Hydatidiform moles are most commonly found in women under the age of 17 or over the age of 35. Hydatidiform moles can be detected by morphological and histopathological examination. Even experienced pathologists cannot easily classify between complete and partial hydatidiform moles. However, the distinction between complete and partial hydatidiform moles is important in order to recommend the appropriate treatment method. Therefore, research into molar pregnancy image analysis and understanding is critical.

The hypothesis of this research project is that an anomaly detection approach to analyse molar pregnancy images can improve image analysis and classification of normal PHM and CHM villi. The primary aim of this research project is to develop a novel method, based on anomaly detection, to identify and classify anomalous villi in molar pregnancy stained images.

The novel method is developed to simulate expert pathologists' approach in diagnosis of anomalous villi. The knowledge and heuristics elicited from two expert pathologists are combined with the morphological domain knowledge of molar pregnancy, to develop a heuristic multi-neural network architecture designed to classify the villi into their appropriated anomalous types.

This study confirmed that a single feature cannot give enough discriminative power for villi classification. Whereas expert pathologists consider the size and shape before textural features, this thesis demonstrated that the textural feature has a higher discriminative power than size and shape.

The first heuristic-based multi-neural network, which was based on 15 elicited features, achieved an improved average accuracy of 81.2%, compared to the traditional multi-layer perceptron (80.5%); however, the recall of CHM villi class was still low (64.3%). Two further textural features, which were elicited and added to the second heuristic-based multi-neural network, have improved the average accuracy from 81.2% to 86.1% and the recall of CHM villi class from 64.3% to 73.5%. The precision of the multi-neural network

has also increased from 82.7% to 89.5% for normal villi class, from 81.3% to 84.7% for PHM villi class and from 80.8% to 86% for CHM villi class.

To support pathologists to visualise the results of the segmentation, a software tool, Hydatidiform Mole Analysis Tool (HYMAT), was developed compiling the morphological and pathological data for each villus analysis.

Acknowledgements

Table of Contents

List of Figures

List of Tables

XI

Chapter 1: Introduction

1.1. Context of the Investigation of molar pregnancy

Molar pregnancy, also known as hydatidiform mole (HM), occurs as a result of an abnormality when a sperm fertilises the egg. HM is a genetically abnormal and nonviable conception, normally associated with high risk of developing complications due to persistence of abnormal trophoblast and resulting in a miscarriage. It is an immature placenta characterised by a massive fluid accumulation within the villi. In general there is an absence of fetal blood vessels (Benirschke *et al.*, 2006). HM is classified into four distinct clinicopathologic entities: partial hydatidiform mole (PHM), complete hydatidiform mole (CHM), persistent gestational trophoblastic and choriocarcinoma. Persistent gestational trophoblastic is the disease caused by HM growing from the uterus surface to the muscle layer around the uterus surface namely myometrium. Choriocarcinoma is a rapidly growing cancer found during pregnancy. This disease is caused by the abnormal CHM tissue that continues growing. CHM are diploid androgenetic and lack normal fetal blood vessels; the villi have an abnormal budding architecture and show trophoblast proliferation. PHM are paternal triploid and have some normal villi mixed with abnormally shaped villi; the villi are irregularly shaped and identified by their only focal abnormal trophoblastic proliferation (Sebire, 2010). The morphological characteristics of CHM and PHM are different from normal placental villi (Figure 1.1). These categories of hydatidiform mole can be distinguished by means of gross morphologic and histopathological examination. Persistent trophoblastic disease is when women who have had treatment to remove a molar pregnancy still have some molar tissue left behind whereas choriocarcinoma happens when cells that were part of a normal pregnancy or a molar pregnancy become cancerous. Management of molar disease relies heavily on its early histological identification and subsequent surveillance, in order to provide early effective treatment (Sebire, 2010).

Histopathology is the microscopic study of diseased tissues. Histology studies tissues are cut into thin slices, which usually come from a surgery, biopsy or autopsy. The tissues are sectioned into very thin 2-7 micrometre sections. The slices are thinner than the average cell and they are stained with one or more pigments to enhance the contrast of the cells and to allow a visual microscopic examination. The histological slides are examined under a microscope by a pathologist who fills a report describing his/her finding and recommendations.
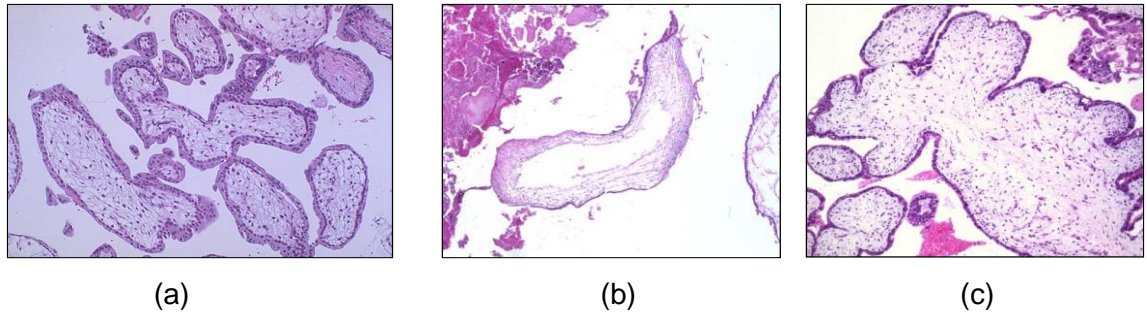
Figure 1.1. A picture of normal villi, PHM and CHM villi;

(a) normal villi, (b) PHM and (c) CHM villi.

The distinction between complete or partial hydatidiform moles is important for determining the appropriate treatment of patients. Sebire *et al.* (2003), Sumithran *et al.* (1996) and Howat *et al.* (1993) explain that the diagnosis of these moles continues to be a problem for many practicing and experienced histopathologists because in early pregnancy CHM and PHM may be difficult to distinguish morphologically from other abnormal pregnancies. Further studies by Landolsi *et al.* (2009) and Kim *et al.* (2009) confirm the challenges of histopathological diagnosis of molar pregnancy and conclude that even experienced pathologists cannot easily distinguish between Complete Hydatidiform Moles (CHM), Partial Hydatidiform Moles (PHM), and Hydropic Abortion (HA) in the early gestational period. There are several critical areas that can lead to diagnostic error, namely the diagnosis of early complete mole as partial mole, the over-diagnosis of hydatidiform mole in tubal pregnancy and the diagnosis of placental site non-villous trophoblast as placental site trophoblastic tumour or choriocarcinoma (Wells, 2007). Paul *et al.* (2010) explain that the diagnosis is difficult because morphological analysis is inadequate to mark confident diagnoses in many cases, and the histological features of complete mole at an early gestation are frequently confused with partial mole, hydropic miscarriage or non-molar chromosomal abnormalities. The characteristics of complete and partial hydatidiform moles in histopathological examination are anomalous from normal placenta cells and different from each other. The complete hydatidiform mole lacks blood vessels, and "the villi are connected to one another by their strands of connective tissues" (Benirschke *et al.*, 2006: 797), whereas partial hydatidiform moles are characterised by having some normal villi and focal trophoblastic proliferation. A visual glossary of the medical terms used in the thesis is provided in Appendix A.

In a minority of cases (15% of CHM and 0.5% of PHM), HM can develop into a persistent disease such as choriocarcinoma, a malignant form of gestational trophoblastic disease. Hence it is of critical importance to identify and distinguish hydatidiform moles

from non-molar specimens. The development of new methods that help differentiate these diagnoses in doubtful cases could be critical for treatment purposes. Hydatidiform moles are most commonly found in women under the age of 17 or over the age of 35. In the United States, the hydatidiform mole incidence is about one in 2,000 pregnancies (Smith, 2003); the incidence is about one in 1,000 in the UK. However, Gul *et al.* (1997) and Khaskheli *et al.* (2007) report higher rates of hydatidiform moles in Asia and Africa (20 cases per 1000).

1.2. Aims and objectives of this thesis

The literature review of computational cancer image analysis is concerned primarily with breast, lung, skin, cervical and prostate cancers; any research review involved with molar pregnancy tends to focus solely on the management and treatment aspects. As a result, research into molar pregnancy image analysis and understanding is still unexplored. Cancer occurs when normal cells grow and multiply without normal control. As the cells multiply, they form an area of abnormal cells, known as a tumour. Many tumours exhibit abnormal chromosomal segregation at cell division (Gisselson, 2001). These anomalies play an important role for detecting cancerous cells.

The hypothesis of this research project is that an anomaly detection approach to analyse molar pregnancy images can achieve a better image analysis and classification of molar pregnancy types than the current approaches. The focus of this thesis is the study of the two most critical hydatidiform moles, CHM and PHM; the aim is to develop a novel method that combines the theory of anomaly detection with pathologists' heuristics to identify PHM and CHM cancerous cells in molar pregnancy stained slides. The author of this thesis collaborated with two pathologists, one based at Great Ormond Hospital, London, and the other based at the Bristol University Hospital.

To achieve the aims the following objectives are carried out:
(i)   To conduct a literature review to survey current approaches to cancer detection from tissues slides and investigate criteria for validation.
(ii) To undertake a theoretical study of existing methods of anomaly detection.
(iii) To collect stained slides of molar pregnancy for analysis from open source website data.
(iv) To capture and depict pathologists' expert and strategic knowledge and the morphological features of molar pregnancy in an ontological representation.
(v)   To develop a novel method of anomaly detection of cancerous cells from stained slides of molar pregnancy based on the above ontological representation.

(vi) To apply the novel method to the data and to carry out experiments to classify the slides into their appropriate clinicopathologic category.

(vii) To validate the results using well known performance measures, namely accuracy, sensitivity and specificity factors, as well as using the knowledge of expert pathologists.

(viii) To compare the results against other current approaches.

(ix) To write a thesis and publish at least two papers.

1.3. Novel contributions

The novel contributions of this thesis are summarised as follows:

(i) A new application domain: the automated analysis of histopathology molar pregnancy tissues.

(ii) A new image understanding method that combines image processing and artificial intelligence techniques, guided by pathologists' heuristic knowledge and strategies. The proposed method is based on the anomaly detection of HM morphological features (e.g. villi, trophoblast, stroma) as identified by the experts' knowledge (Figure1.2).

(iii) A cognitive approach to image analysis: the development of an algorithm that mirrors the heuristic approach to diagnosing anomalies in villi.

1.4. Methods of investigation

Research methods can be categorised into three types: qualitative, quantitative and mixed methods. A qualitative method is related to inductive approaches, because this method is used to make sense of phenomena, to understand the reasons, motivations and opinions of people, based on empirical materials such as interviews, observations, historical texts. Ticehurst and Veal (2000) explain that there is a linkage between quantitative methods and positivism because the quantitative approach is also known as management science or operational research, linking disciplines with philosophy. They also argue that quantitative and qualitative methods are linked to positivist and interpretivist epistemologies, as shown in Figure 1.3. The research method selection is critical because the applied research method should direct/guide research purposes to the correct goal (Crotty, 1998). However, the research method is also involved with research approaches and philosophies, for example, a quantitative method is used to apply positivism and a qualitative method is used to apply interpretivism, but it can be modified depending on the particular characteristics of a research project (Karl, 2004).
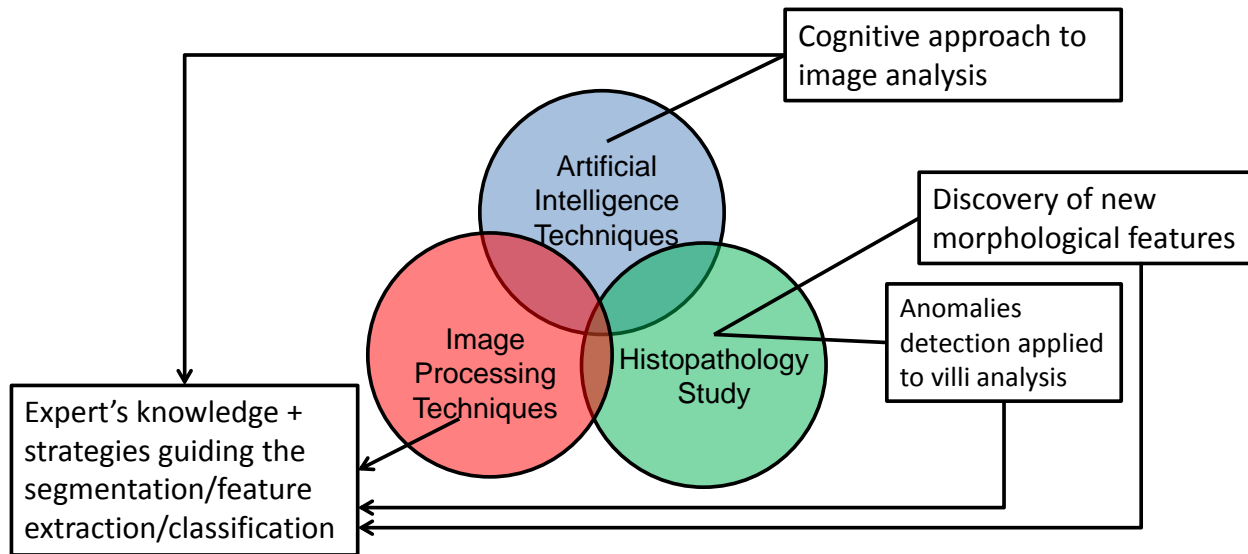
Figure 1.2. A new image understanding method diagram.

A quantitative method tends to be based on numerical measurements (e.g. graphs, surveys, and statistical data) and experimentation (Saunders *et al.*, 2009). It seeks to test hypotheses and/or seek explanations and predictions. This method is commonly used with a deductive approach based on a positivist philosophy, as shown in Figure 1.3. The deductive approach is based on using knowledge and information to perceive or produce an opinion about something, and a positivist philosophy believes in the scientific evidence (i.e. experiments and statistics) instead of ideas. Therefore, research based on a positivist philosophy usually uses a quantitative method and deductive approach to extract proved facts, predictions or trends. A mixed method is applied to analyse data by using quantitative and qualitative methods.
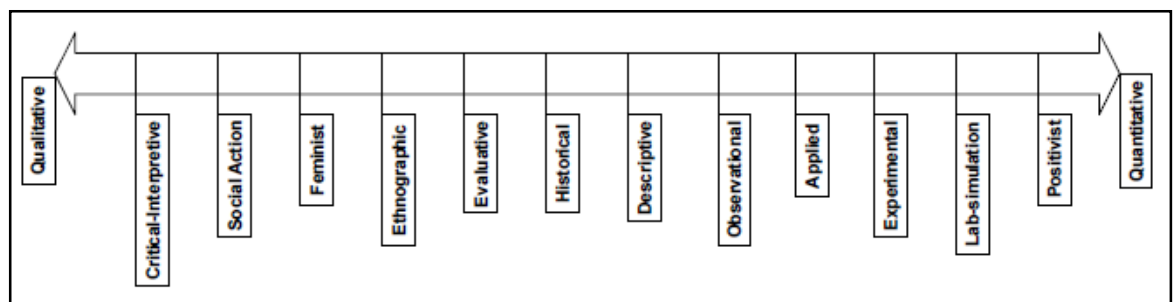
Figure 1.3. Approaches and methodologies (Ticehurst & Veal, 2000).

In this research, the quantitative method, based on the deductive approach and positivist philosophy as shown in Figure 1.4, is used to verify the hypothesis that image processing and analysis based on anomaly detection techniques and pathologists' heuristics can classify HM images into normal cells, complete and partial molar pregnancy cells. Positivism is suitable for this research because this research can validate the

classification results with expert pathologists. The basic steps used in this research are outlined below:

- A literature review to survey related work to cancer detection and relevant anomaly detection algorithms.
- A deep understanding of molar pregnancy and normal cells will be understanding in order to identify anomalies within stained slides.
- Elicitation of pathologists' expertise and strategies in identifying anomalies in molar pregnancy stained slides.
- Ontological representation of molar pregnancy morphological and clinicopathologic characteristics following elicitation of pathologists' heuristics.
- Data collection and experimentation of stained slides of molar pregnancy. Development of a novel method based on anomaly detection and molar pregnancy ontology using image processing techniques and artificial intelligence techniques.
- Implementation of the above methods to stained slides images, and further experimentation.
- Validation of final results and comparison with other related work and ground truths that are verified by expert pathologists.
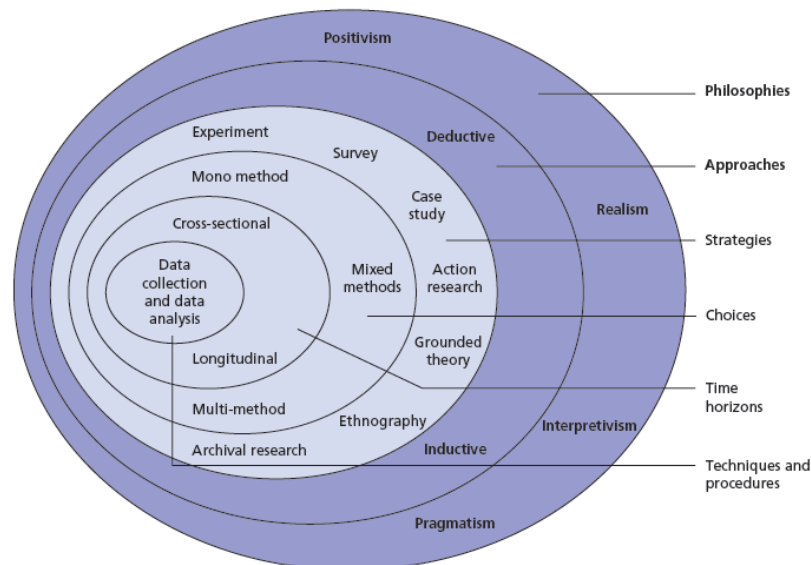
Figure 1.4. The research 'onion' (Saunders *et al.*, 2009).

1.5. Structure of the thesis

The structure of the thesis consists of seven chapters as follows:

- In Chapter 1, the introduction of the domain of study is described. The molar pregnancy is first defined, then the aims of the research and its novel contributions and method of investigation are explained.

- In Chapter 2, the literature review, image processing techniques for cancer image analysis are reviewed and categorised. The four steps of traditional image analysis methods are pre-processing, segmentation, feature extraction and classification steps. The pros and con of each technique are also discussed in this chapter.

- The review of anomaly detection is presented in Chapter 3. This chapter starts with the fundamental approaches of anomaly detection, the nature of input data and the types of anomalies. Then, anomaly detection techniques are explained and the techniques are grouped as follows: statistical anomaly detection, machine learning based anomaly detection and other approaches. Anomalies detection issues, applications of anomaly detection, and discussion are at the end of this chapter.

- Chapter 4 consists of two main parts. The first part describes the data used in this research project and explains how the pathological and morphological features that are elicited from both, the expert pathologists and medical documents, are represented ontologically. The second part introduces the novel approach of image analysis and classification guided by the ontological representation and the heuristics of the experts. It also provides a description of steps associated with the low level image processing and analysis, namely the pre-processing, segmentation and feature extraction steps.

- Chapter 5 discusses the high level processing of image analysis and classification. The early sections focus on identifying and ranking the critical features of the hydatidiform mole villi. The middle sections describe the heuristic multi-neural network approach to anomalies detection based on 15 and then 17 features villi. The experimental results of the proposed method are analysed and compared with the traditional multi-layer perceptron (MLP).

- Chapter 6 describes the Hydatidiform Moles Analysis Tool (HYMAT) developed to support low level processing tasks.

- Chapter 7 presents the conclusions of this research, its novel contributions and future work related to the study of hydatidiform mole villi.

Chapter 2: Literature Review: Image Processing and Classification Techniques

2.1. Image processing and classification techniques

This chapter reviews current image processing techniques and classification methodologies that are applied to cancer classification, detection or diagnosis of various types of images such as histological images, digital mammograms, ultrasound images and skin images.

The traditional approach used in cancer image analysis consists of four steps (Figure 2.1). The first step, i.e. the pre-processing step, is to remove unwanted objects (such as noise) and improve the quality of an image. The second step is the image segmentation step, aimed at selecting objects of interest or regions of interest from the background. The purpose of the third step is to extract noticeable features that can be used to classify the objects. The final step is used to classify or categorise the objects, using the features extracted from the previous step. The next sections describe the current methods and approaches associated with each one of the listed steps.
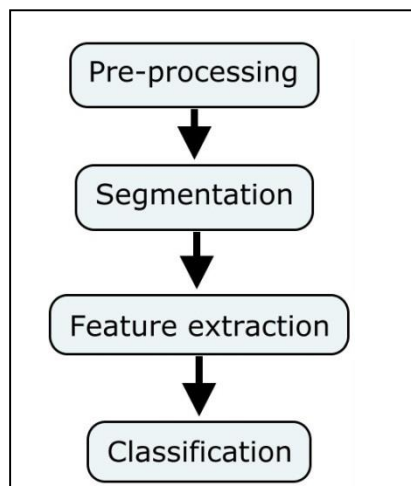
Figure 2.1. The four cancer image analysis steps

2.2. Pre-processing step

The pre-processing step is one of the important image processing tasks and it is used to improve the performance of the segmentation and feature extraction processes. It ensures that some objects that might be interesting are not removed and that useful details in the images are not eliminated. For instance, the Gaussian smoothing filter can remove noise, but this filter can also eliminate texture information in the image (Waheed

et al., 2007). Histogram equalisation improves the contrast in the ultrasound images, for example (Han *et al.,* 2007). The methods used in the pre-processing step can be grouped into noise removal and image enhancement algorithms.

The purpose of noise removal is to remove unwanted objects and background. Crisan *et al.* (2007) use thresholding to separate the bright regions from the dark regions, while Sang *et al.* (2008) apply Otsu's automatic threshold selection to select suspicious regions and to distinguish between the breast tissues and the background in mammographic images (Xin *et al.,* 2004). Otsu's threshold is also applied to distinguish between skin cancer regions and background (Dhinagar *et al.,* 2011), whereas Elizabeth *et al.* (2012b) apply the thresholding technique, based on convex edge and the centroid of a cancerous area, to select the lung area of interest from lung cancer tomography images. In their paper Xu *et al.* (2011) advocate the use of a double thresholding method to separate cancer stem cells from background and remove noise. The results show that the proposed method yields accurate segmentation results with fast execution time. Haneishi *et al.* (2001) combine thresholding with a labelling method to select lung cancer regions from x-ray CT images. The various thresholding techniques are captured in Figure 2.2.
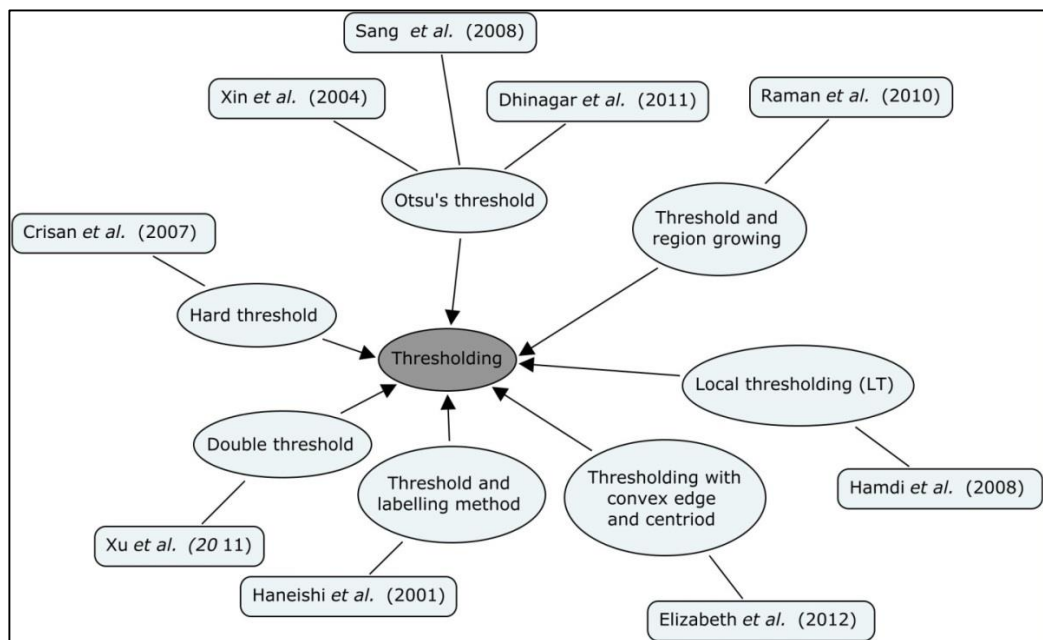
Figure 2.2. Thresholding techniques.

Other statistical methods include Gaussian filter of colour images for smoothing Fine Needle Aspiration Cytology (FNAC) images (Niwas *et al.,* 2010a), Discrete Wavelet Transforms (DWT) for eliminating the low frequency image components in a digital mammogram (Lahmiri *et al.,* 2011), background masking, based on entropy measures, for separating the background from cells (Kazmar *et al.,* 2010), and median filtering for

removing all irrelevant data in images for better classification (Xing-Li *et al.*, 2008). Salvado *et al.* (2005) also apply DWT to remove the low frequency sub-band of breast cancer digital mammograms, and then a reconstructed image is created by the high frequency sub-bands. The results show that the reconstructed image can be used to improve the contrast of digital mammograms for further diagnosis. In the paper by Lin *et al.* (2014), DWT is used to improve a breast cancer mammogram image quality. The mass signals of transformed images are enhanced before extracting features. The results show that the masses in enhanced images are easier to identify than in the original images. Markelj *et al.* (2012) review 3D and 2D data registration methods for creating an image containing more information for analysis from a cone-beam CT, CT, MR, or ultrasound image. These data registration techniques are suitable for applications where 2D and 3D images information is necessary, for example, 3D anatomical structure reconstruction.

Acceptable and widely used image enhancement methods are histogram equalisation, data normalisation, gradient enhancement, mean contrast enhancement, discrete wavelet transform and Gaussian filter. Histogram equalisation is used to improve the intensity distribution in images (Raman *et al.*, 2010), in MRI images (Naghdy *et al.*, 2010) and in ultrasound images of prostate cancer (Seok *et al.*, 2007). Data normalisation is applied to eliminate the effect of the variance of scale and to give robustness to the algorithm (Hui *et al.*, 2008). In some applications, researchers change the colour space to enhance image quality. To improve the textural and statistical features of gynaecological cancer images, Neofytou *et al.* (2008) change RGB images to the YCrCb colour system and the results indicate that the Y, Cr and Cb channels make a significant difference for statistical and textural features. Boquete *et al.* (2012) also apply YCrCb colour space to segment thermal infrared images of breast cancers. Furthermore, RGB pathological images are converted to HSV images and the H and V elements are used to extract textural features in feature extraction processes (Xiangmin *et al.*, 2008). In addition, the smoothed Fine Needle Aspiration Cytology (FNAC) image is converted into the hue, saturation, and intensity (HSI) colour system, because this yields better classification results than the RGB and CIE-Lab colour spaces (Niwas *et al.*, 2010a). Doyle *et al.* (2012) also apply the HSI colour system to RGB digital needle biopsies of prostate cancer. The advantage of the HSI colour system is that the colour information is separated from brightness. Therefore, further analysis can be done with more robust information, whereas Mouelhi *et al.* (2013a) use Fisher's linear discriminant to reduce the information of RGB and saturation values to two new features for breast cancer nuclei segmentation. The results indicate that the proposed method achieves better segmentation results than other methods. The colour spaces applied to enhance image quality for cancer image analysis are shown in Figure 2.3. In the paper by Gavrilovic *et al.* (2013) a blind method for colour

decomposition of histological images is used to deal with stain image intensity variation. The method is an extension of an ordinary linear decomposition method. Other image enhancement methods include discrete wavelet transform on digital mammogram images (Hamdi *et al.*, 2008), gradient, mean contrast, discrete wavelet transform and Gaussian filter on raw digital mammogram images (Chui-Mei *et al.*, 2008). Allwin *et al.* (2010) apply a set of morphological operators on a grey scale cyto image of cervical cancer, whilst Linguraru *et al.* (2009) and Cui *et al.* (2010) employ anisotropic diffusion to enhance the contrast of CT and microscopy images. To overcome the limitation of lung functional single-photon emission computed tomography (SPECT) images, Haneishi *et al.* (2001) create syntactic images from x-ray CT and SPECT images. The syntactic images can give the location information from x-ray CT images and the details for analysis from SPECT images.
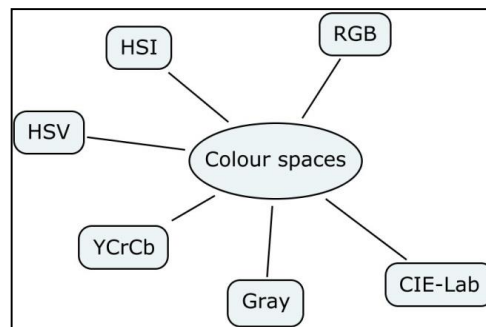
Figure 2.3. Colour spaces

2.3. Segmentation step

Segmentation is the process that aims to separate the region of interest (ROI) from the background, as this leads to better feature extraction and classification processes. Segmentation can be categorised into two major approaches: region-based and boundary-based approaches (Demir & Yener, 2005). In spite of several decades of research, segmentation remains a challenging problem. Two main challenges include over-segmentation and under-segmentation.

- **The region-based approach**

The region-based approach is based on either statistical approaches or machine learning algorithms. The simple region-based segmentation thresholding is widely used to segment the Region of Interest (ROI) in various applications.

In the paper by Naghdy *et al.* (2010) the intensity threshold is used to segment brain cancer regions from MRI images, whereas Hamdi *et al.* (2008) use Local Thresholding (LT) in digital mammograms for separating micro-calcifications from background. The pathological prostate cancer image segmentation is improved by

converting the RGB colour space image to HSV colour space, and by computing the H and V components. The area filter based on threshold values of 300 and 50 pixels is applied to the H and V components. Not only does this filter separate between lumens and artefacts in H components, but also between nucleus and artefacts in V components (Xiangmin *et al.*, 2008). Chang *et al.* (2012) apply the Multi-Reference Graph Cut (MRGC) method to deal with technical variations in sample preparation in glioblastoma multiform segmentation. A statistical based segmentation method is one of the common segmentation methods. Boquete *et al.* (2012) use the automated detection method based on Independent Component Analysis (ICA) to detect high tumour risk areas, whereas Yaguchi *et al.* (2011) apply the segmentation method based on an expectation maximization (EM) algorithm, to segment stomach cell components.

Other approaches combine two methods for improving the performance of the region based segmentation. Sang *et al.* (2008) use a region-based approach after thresholding for better segmentation results. Connected threshold region growing segmentation is used to extract features from a seed defined by a user on the threshold image, whereas Raman *et al.* (2010) perform region growing segmentation after applying the threshold to improve a low intensity and maximise a peak in the histogram of digital mammograms. Hadavi *et al.* (2014) also apply region growing based on thresholding to segment CT images of lung cancer. Kazmar *et al.* (2010) utilise Radial Symmetry Decomposition (RSD) and Blob-like Keypoint (BK) detection to segment the ROI. RSD is used as the gradient voting approach for each pixel and blob-like keypoint, with a Gaussian filter applied to select cells from the background. Mata *et al.* (2000) apply the wavelet transform to decompose the image into sub-band images and use the details responded to each scale to segment micro-calcification regions by the Gaussianity test. Karnan & Gandhi (2010) combine Markov Random Field (MRF) with a Hybrid Population based Ant Colony (HPACO) algorithm to detect the micro-calcifications from mammogram images. Rabiei *et al.* (2007) propose a primary segmentation based on a binary region mask that is defined by a user. A user helps this segmentation to eliminate as much as possible the background pixels from images by the blob number of the blob combiner. Then, the blob combiner is performed for separating between normal, benign, suspicious and malignant, to enhance Dynamic Contrast Magnetic Resonance Imaging (DCE-MRI) images. Kang *et al.* (2011) apply the k-means clustering algorithm with the information of neighbours and boundaries, to segment breast cancer regions in breast MRI images, whereas Mohapatra *et al.* (2011) apply the k-means algorithm to detect leukemia in blood cell microscopic images. In the paper by Mouelhi *et al.* (2013b), an enhanced watershed method is applied after a fuzzy active contour model to improve an automatic breast

cancer cell image segmentation method. The results show that the proposed method performs better than other current segmentation methods.

The machine learning approach to region based segmentation is adopted by Taher & Sammouda (2011) who propose Hopfield Neural Networks (HNN) and a Fuzzy c-Means (FCM) clustering algorithm to identify lung cancer on sputum colour images. HNN can segment the overlapping cytoplasm classes and is very sensitive to intensity change, whereas FCM cannot segment the overlapping cytoplasm areas. Naghdy *et al.* (2010) use k-means algorithm to separate the nuclei and cytoplasm from the background and Yue *et al.* (2010) combine k-means algorithm with Otsu's algorithm to separate the abnormal nuclei regions from all nuclei regions in cancer microscopic images. Waheed *et al.* (2007) used a watershed algorithm to separate overlapping nuclei. Rathore *et al.* (2013) apply k-means algorithm with textural features to segment colon biopsy image of colon cancer, and the results indicate that the proposed method achieves better segmentation results than a segmentation method based on circular primitive techniques, Instead of applying k-means algorithm with textural features, Vijayaraghavan *et al.* (2014) apply k-means algorithm with L*a*b* colour system to segment abnormal regions in digital mammograms of breast cancer. Acosta-Mesa *et al.* (2005) apply the Naive Bayes classifier (NB), based on a supervised learning approach, to classify the parabola features of cervical cancer of microscopic images, whereas Kekre *et al.* (2010) use the vector quantisation technique and Linde Buzo-Grey algorithm (LBG) to segment MRI images of breast cancer. In the paper by Marcomini & Schiabel (2012) a Self-Organizing Map (SOM) is applied to segment the suspicious masses boundary. The results show that the proposed algorithm still suffers from speckle noise. For skin cancer image segmentation, Amelio *et al.* (2013) apply genetic algorithms to segment skin lesions, and the proposed method achieves promising segmentation results. The segmentation techniques based on region-based approach are shown in Figure 2.4.

- **The boundary-based approach**

The boundary-based approach is the method based on finding out the border of the objects. Morphological methods are commonly used in a wide range of applications. The simple morphological segmentation approach is manual segmentation. Alolfe *et al.* (2008, 2009) apply a window of 32×32 pixels to select the group of ROI of the training and testing sets for feature extraction processes and Vani *et al.* (2010) manually segment the suspicious regions in digital mammograms as ROI.
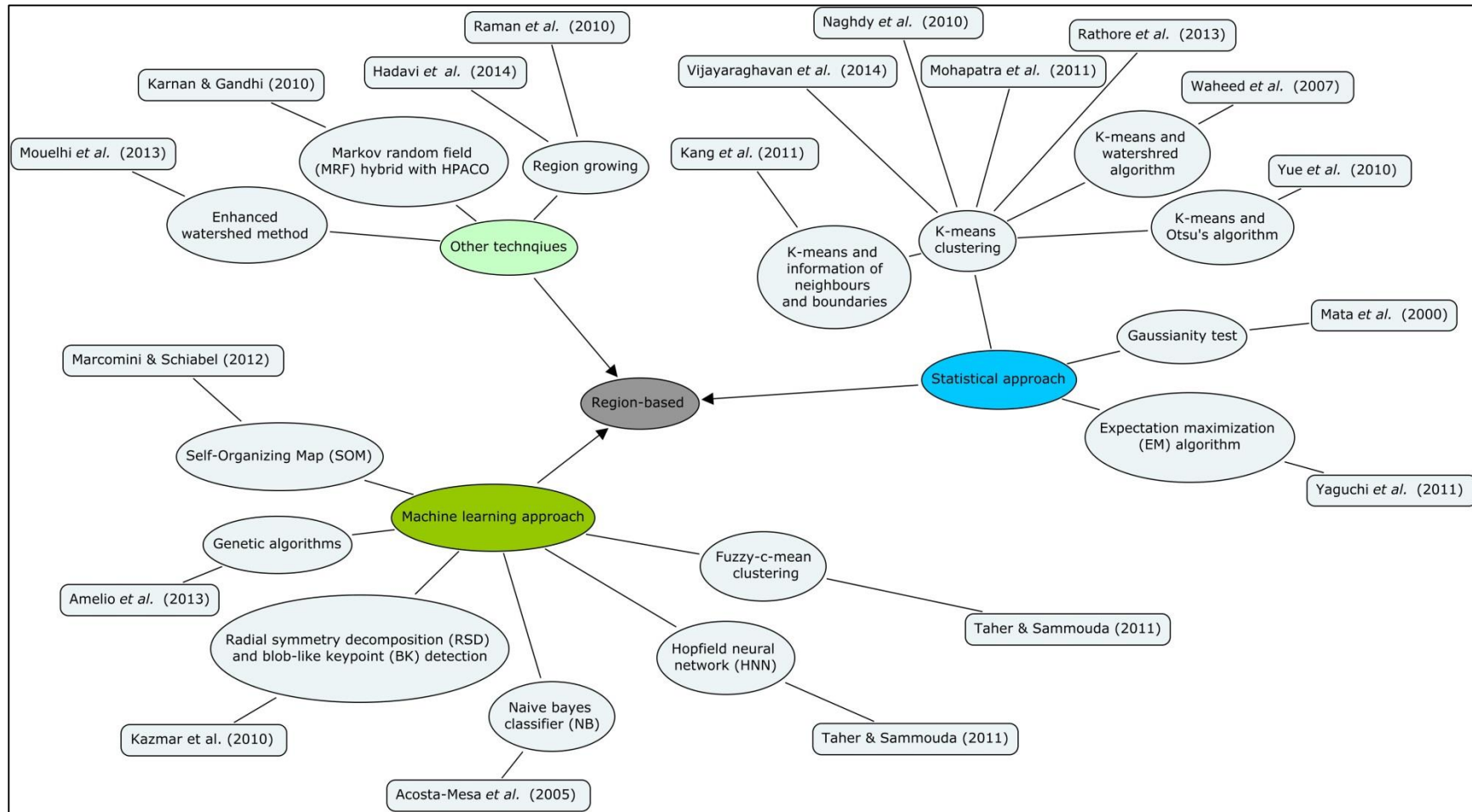
Figure 2.4. Segmentation techniques based on region-based approach.

The contour based algorithm takes advantage of the border of objects for segmentation. Crisan *et al.* (2007) apply the automatic contour trace algorithm to the digital mammograms to find out the boundary of the lesions from the images, whereas Naik *et al.* (2007) combine the contour-based approach with a level set algorithm, initialised by a user, to enhance the performance of the prostate tissue image segmentation. In a paper by Linguraru *et al.* (2009), a combination of fast marching and geodesic active contour level sets is applied to segment renal lesions of renal cancer. The snake technique is another boundary-based segmentation approach, based on internal forces and external forces in the image. Doukas *et al.* (2010) apply active contour techniques (snake) to microscopic images for cell death (apoptosis) segmentation. However, cell overlapping is one of the problems of active contour algorithms. To improve the segmentation results, Parolin *et al.* (2010) use an edge map that utilises the Wavelet Transform (WT) to guide the Gradient Vector Flow Snakes (GVF snake) to segment the lesion in dermatological images. Sakkalis *et al.* (2009) apply the magic wand and snake algorithms to segment MRI tomographic images. The results show that the segmented images can be used for further analysis. Another snake technique improvement is proposed by Chaddad *et al.* (2011). A progressive division of the image dimensions is used to improve the execution time of the snake technique, and the results indicate that the proposed method is more than 50% faster than an ordinary snake. In such research, the boundary-based greedy snake and region-based region growing algorithms are applied to select the ROI of lung cancer tomography images (Elizabeth *et al.*, 2012a). A greedy snake is used as a primary segmentation tool, then a region growing algorithm is applied to refine the segmentation results. Jayadevappa *et al.* (2011) review segmentation algorithms based on deformable models, which are boundariy-based segmentation models using internal and external forces to guide their construction. Deformable models can be categorised into parametric and geometric methods. Parametric methods consist of classic and GVF snake (Xu & Prince, 1998), whereas geometric models consist of Geometric Active Contour (GAC) model (Caselles *et al.*, 1993), (Malladi *et al.*, 1995), geodesic active contour model (Caselles *et al.*, 1997), level sets (Osher & Sethian, 1988) and variational level sets (Wang *et al.*, 2007), (Li *et al.*, 2005). The weak point of these segmentation algorithms is that the deformable models still require clear boundaries to achieve satisfactory segmentation results.

Other approaches that take advantage of the boundary and morphological features are Hough transforms on thermal infrared images, and line-scanning based on morphological features. Kuruganti *et al.* (2002) apply the Hough transform of a parabola to thermal infrared images of breast cancer for segmenting the ROI from background, whereas Qi *et al.* (2001) use the parabolic shapes of lower breast borders in thermal

infrared images that are detected by the Hough transform, in order to classify any abnormalities. Filipczuk *et al.* (2013) combine the Hough transform with a quadratic discriminant to segment nuclei of fine needle biopsies of breast cancer cytological images. The proposed techniques can remove unwanted red blood cells and overlapping nuclei and improve system robustness. Instead of using a quadratic discriminant, George *et al.* (2013) use Otsu's thresholding and FCM to improve the robustness of circular Hough transform segmentation results for breast cancer cytological image classification. Furthermore, Chang *et al.* (2009) use line-scanning, based on the morphological features of a cell in segmentation processes, to improve the quality of the cell image by using grey-level and energy methods. The overview of segmentation techniques based on the boundary-based approach is shown in Figure 2.5.
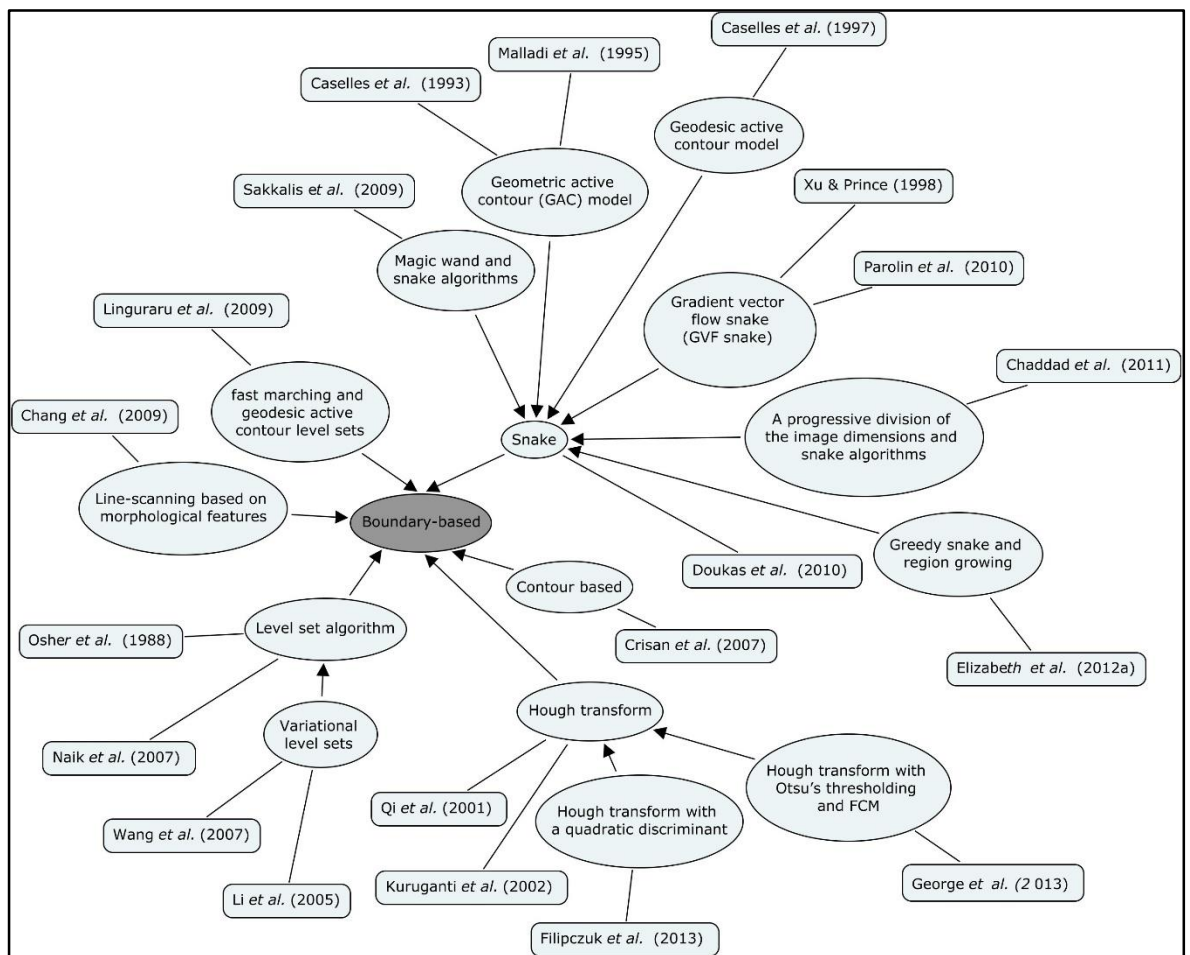
Figure 2.5. The overview of segmentation techniques based on the boundary-based approach.

2.4. Feature extraction step

Feature extraction aims at extracting the relevant features that can be applied to classify or categorise the objects of interest from each other. Types of feature extraction can be classified into statistical, textural, morphological, fractal-based and topological features, discussed in the next section. The major challenge of feature extraction is identifying the optimum number of relevant features.

- **Statistical features**

The simple statistical approaches in the feature extraction step are standard deviation, variance, mean, bias and kurtosis. For example, these five statistical measures are extracted from digital mammograms of the breast cancer diagnosis system and classification by proposed Xing-li *et al.* 2008. Allwin *et al.* (2010) use mean, standard deviation, skewness and kurtosis to identify the cervical cancer stages on cyto images. However, the results of the experiments of Kuruganti *et al.* (2002) show that mean and entropy cannot improve the accuracy of the breast cancer thermal infrared (TIR) image detection system.

To investigate critical features, Naik *et al.* (2007) apply standard deviation, variance, compactness and smoothness to prostate cancer tissue images as the features for classification processes. These features are not only used to distinguish between cancerous and normal prostate tissues, but also to discriminate between 3 and 4 Gleason grades. Neofytou *et al.* (2008) use mean, variance, median, mode, skewness, kurtosis, energy and entropy to detect textures features from hysteroscopy images of the endometrium for gynaecological cancer classification. The results demonstrate that the statistical and Grey Level Difference Statistics (GLDS) yield the highest correct classification score with a support vector machine classifier, whereas Chui-Mei *et al.* (2008) use the mean, variance, Drect Cosine Transform (DCT) coefficients and entropy as features. DCT coefficients and entropy are the most useful features of micro-calcifications due to the sensitivity of grey level changes. Vani *et al.* (2010) use fifteen features (mean grey level, variance of grey level, mean of gradients, variance of gradients, energy, inertia, entropy, homogeneity, correlation, smoothness, skewness, kurtosis, z-score, moment and range) to classify breast cancer from digital mammograms. The statistical features and application domains are shown in Table 2.1.

Table 2.1. The statistical features and application domains

Statistical features	Application domains	Authors
Standard deviation	Cervical cancer	Allwin *et al.* (2010)
	Prostate cancer tissue images	Naik *et al.* (2007)
Variance	Digital mammograms	Chui-Mei *et al.* (2008)
		Xing-li *et al.* (2008)
		Vani *et al.* (2010)
	Prostate cancer tissue images	Naik *et al.* (2007)
	Gynaecological cancer images	Neofytou *et al.* (2008)
	Lung cancer tomography images	Elizabeth *et al.* (2012a)
Mean	Digital mammograms	Chui-Mei *et al.* (2008)
		Xing-li *et al.* (2008)
		Vani *et al.* (2010)
	Cervical cancer	Allwin *et al.* (2010)
	Breast cancer (TIR) images	Kuruganti *et al.* (2002)
	Gynaecological cancer images	Neofytou *et al.* (2008)
	Lung cancer tomography images	Elizabeth *et al.* (2012a)
Median	Gynaecological cancer images	Neofytou *et al.* (2008)
Mode	Gynaecological cancer images	Neofytou *et al.* (2008)
Bias	Digital mammograms	Xing-li *et al.* (2008)
Kurtosis	Digital mammograms	Xing-li *et al.* (2008)
		Vani *et al.* (2010)
	Cervical cancer	Allwin *et al.* (2010)
	Gynaecological cancer images	Neofytou *et al.* (2008)
Skewness	Cervical cancer	Allwin *et al.* (2010)
	Gynaecological cancer images	Neofytou *et al.* (2008)
	Digital mammograms	Vani *et al.* (2010)
z-score	Digital mammograms	Vani *et al.* (2010)

The features based on the wavelet transform are used in various types of applications, such as micro-calcification classification on digital mammograms (Alolfe *et al.*, 2008), breast cancer cytological image classification (Niwas *et al.*, 2010b) and skin cancer image classification (Dhinagar *et al.*, 2011). Lahmiri *et al.* (2011) use the discrete wavelet transform with Gabor filters and uniformity and entropy measures as features to detect suspicious regions from digital mammograms.

Other statistical methods include the Average Higuchi Dimension of Radio Frequency Time series (AHDRFT), proposed as a new feature for prostate cancer detection on ultrasound images, leading to an improvement of the performance of the classification system (Moradi *et al.*, 2006). Furthermore, the Aceto-white response functions (AwRFs) is applied to extract features from colposcopy images of cervical cancer; however, these features need the specific pre-processing step to enhance images before the feature extraction step (Acosta-Mesa *et al.*, 2005). Torrent *et al.* (2010) use a dictionary database filter bank (a bank of filters, four Gaussian derivatives, a Laplacian filter, a corner detector, and two Sobel filters) extracted from digital mammograms for their automatic micro-calcification detection system.

- **Textural features**

A textural feature is the feature extracted by measuring the surface variations of the object of interest or ROI, such as smoothness, coarseness, and regularity (Demir & Yener, 2005). Various textural features are proposed for use as features in cancer classification domains, such as forty-eight textural features computed from a Spatial Grey Level Dependence (SGLD) matrix by weighted combination of the elements of the matrix (Hamdi *et al.*, 2008), textural features calculated by Grey Level Co-occurrence Matrix (GLCM) (Naghdy *et al.*, 2010), and texture edges based on Gabor filter for object structure description (Deepak *et al.*, 2012).

To enhance the system robustness and accuracy, the textural features are combined with statistical and morphological features, such as: Dual-Tree Complex Wavelets Transform (DTCWT) based decomposition method (Niwas *et al.*, 2010b), Grey Level Difference Matrix (GLDM) (Kaman *et al.*, 2010), and morphological features (e.g. line length, area fraction, quotient and Euler number) (Xiangmin *et al.*, 2008). Neofytou *et al.* (2008) apply textures features, Spatial Grey Level Dependence Matrices (SGLDM) and Grey Level Difference Statistics (GLDS), and statistical features extracted from the converted hysteroscopy images (YCrCb colour space) for the gynaecological cancer classification system. The textural features computed from each Y, Cr and Cb channels give better results than the system that only uses the GLDS and statistical features from the Y channel. Mohapatra *et al.* (2011) also combine textural features with morphological features (area, perimeter, compactness, solidity, eccentricity, elongation and form factor) to detect abnormal regions in blood microscopic images, while Elizabeth *et al.* (2012a) use eight textural features (namely, smoothness, contrast, homogeneity, dissimilarity, energy, entropy, eccentricity, correlation) and other features (such as area, major axis length, minor axis length, mean, standard deviation, orientation and proximity), for lung cancer tomography image classification. Deshpande *et al.* (2013) apply eighteen

statistical and textural features based on GLCM for breast cancer mammogram classification. Torheim *et al.* (2014) compare the GLCM extracted from Brix parameter maps with first-order features (i.e. statistical features) as used for cervical cancer MR image classification, and the experimentation results show that the GLCM features give better classification results than such first-order features. In the paper by Doyle *et al.* (2012), more than 900 statistical and textural features are used as the features of a boosted Bayesian multi-resolution (BBMR) classifier to classify digital needle biopsies of prostate cancer.

In addition, Han *et al.* (2007) use the multi-resolution autocorrelation and the brightness of tissues in ultrasound images as the features for prostate cancer detection system. The experiment results show that the multi-resolution autocorrelation can be used as an important feature of a cancer tissue. For prostate cancer classification, a set of 102 graph-based, morphological and textural features extracted from histological image is used as a set of features of the prostate cancer classification system. The experiment results show that the textural features can identify the difference in tissue patterns (Doyle *et al.*, 2007). Deepa *et al.* (2012) combine DTCWT with statistical features as a feature set of the breast cancer classification application. Waheed *et al.* (2007) use fractal dimension, ratio of area eccentricity and textural features (correlation, contrast, energy, homogeneity and entropy) extracted from pathological images as the features. These features are selected semi-automatically by a user for refining the results, and the system yields satisfactory results for renal cell carcinoma tissue classification. Kazmar *et al.* (2010) apply the fourteen Haralick textural features, extracted from nucleus, membrane, halo, structured border, float and the background, to measure the difference of six cellular objects, giving satisfactory classification results of the cancerous cells in breast cancer cell images. Yuan *et al.* (2006) apply a set of textural feature vectors as the features of skin cancer classification system, using local auto-regression and Gabor filter banks. Then, these textural features are fed to the Support Vector Machine (SVM). The experiment results show that the skin cancer classification system that used only textural features can give average accuracy of about 70%. So, the system needs further investigation to achieve better results. Ganeshan *et al.* (2012) use entropy and uniformity as features to detect oesophageal lesion regions. The results show that the proposed features can be applied to identify the lesion regions and that these features relate to biological features of oesophageal cancer. In the paper of Muthukarthigadevi *et al.* (2013), Laws' texture energy measures and Haralick's texture features are used as features to classify breast cancer regions in digital mammograms. Textural features and associated application domains are summarised in Table 2.2.

Table 2.2. Textural features and application domains.

Textural features	Application domains	Authors
Spatial Grey Level Dependence (SGLD) matrix	Digital mammograms Gynaecological cancer images	Hamdi *et al.* (2008) Neofytou *et al.* (2008)
Grey Level Co-occurrence Matrix (GLCM)	MRI images of brain cancer Cervical cancer MR images Breast cancer mammograms	Naghdy *et al.* (2010) Torheim *et al.* (2014) Deshpande *et al.* (2013)
Grey Level Difference Matrix (GLDM)	Gynaecological cancer images Pathological images of prostate cancer Digital mammograms	Neofytou *et al.* (2008) Xiangmin *et al.* (2008) Kaman *et al.* (2010)
Gabor filter	Skin cancer images Brain, lung, colon, breast and retina	Yuan *et al.* (2006) Deepak *et al.* (2012)
Dual-tree complex wavelets transform (DTCWT)	Cytological images of breast cancer Digital mammogram images	Niwas *et al.* (2010b) Deepa *et al.* (2012)
Smoothness	Lung cancer tomography images	Elizabeth *et al.* (2012a)
Contrast	Renal cell carcinoma tissue images Lung cancer tomography images	Waheed *et al.* (2007) Elizabeth *et al.* (2012a)
Homogeneity	Renal cell carcinoma tissue images Lung cancer tomography images	Waheed *et al.* (2007) Elizabeth *et al.* (2012a)
Dissimilarity	Lung cancer tomography images	Elizabeth *et al.* (2012a)
Energy	Renal cell carcinoma tissue images Lung cancer tomography images	Waheed *et al.* (2007) Elizabeth *et al.* (2012a)
Entropy	Renal cell carcinoma tissue images Lung cancer tomography images Oesophageal cancer images	Waheed *et al.* (2007) Elizabeth *et al.* (2012a) Ganeshan *et al.* (2012)
Correlation	Renal cell carcinoma tissue images Lung cancer tomography images	Waheed *et al.* (2007) Elizabeth *et al.* (2012a)
Uniformity	Oesophageal cancer images	Ganeshan *et al.* (2012)
Multi-resolution autocorrelation	Prostate cancer ultrasound images	Han *et al.* (2007)
Haralick textural features	Breast cancer cell images	Kazmar *et al.* (2010) Muthukarthigadevi *et al.* (2013)
Laws' texture energy measures	Breast cancer mammograms	Muthukarthigadevi *et al.* (2013)

- **Morphological features**

A morphological feature is the feature based on the size and shape characteristics of the object of interest or ROI. The size can be described by radius, area, perimeter of the object and the shape can be described by the compactness, roundness, smoothness, length of the major and minor axis, symmetry, concavity as well as perimeter (Demir & Yener, 2005). The seven boundary features, such as area of overlap, distance, perimeter, compactness and smoothness, are extracted from the prostate cancer tissue, in the study

by Naik *et al.* (2007). In addition, Naghdy *et al.* (2007) use nuclei, to cytoplasm (N/C) ratio, diameter of nuclei, shape factor and roundness of nuclei as the features of cervical histology image classification, and the nuclei to cytoplasm ratio gives satisfactory classification results. Seok *et al.* (2007) apply clinical knowledge based features (i.e., the cancerous region location and shape) for their prostate cancer detection system as these features can decrease the false positive rate of the detection system. Qi *et al.* (2001) use the Hough transform characteristic to classify the breast cancer on thermal infrared images, whereas Li *et al.* (2013) apply enhance roughness index (ERI) as a feature for breast cancer classification.

In some research projects, morphological features are combined with other features such as statistical and textural features. For example, the combinations of morphological and multi-phase intensity features are used as features for the cysts and type of renal cancer classification system (Linguraru *et al.*, 2009). Furthermore, morphological features (e.g. line length, area fraction, quotient and Euler number) and textural features extracted from the nuclei and cell components, are used to classify prostate pathological images. Line length is calculated by the ratio between perimeter of the pathological objects and area of the whole binary image. Area fraction is computed by the ratio between area of the pathological objects and area of the whole binary image (Xu *et al.*, 2008). The morphological feature, namely ratio of area eccentricity, is combined with textural features and fractal dimension for the renal cell carcinoma tissue classification task. The experiment results show that using the ratio of area eccentricity feature alone cannot give satisfactory results for the renal cell carcinoma tissue classification: therefore, the combined features are necessary for this difficult task (Waheed *et al.*, 2007). Filipczuk *et al.* (2013) also extract twenty-five features, combined with statistical, morphological and textural features for fine needle biopsies cytological images of breast cancer classification. The experimentation results emphasise the combined feature sets, which play an important role for complex cancer image classification applications. The morphological features and application domains are summarised in Table 2.3.

- **Fractal-based features**

The concept of a fractal is normally associated with geometrical objects satisfying two criteria, namely self-similarity and fractional dimensionality. The widely used and accepted methods for measuring the fractal dimension are the box-counting, perimeter-stepping and pixel dilation methods (Demir & Yener, 2005). Waheed *et al.* (2007) apply the spectral analysis method for measuring the fractal dimension on pathological images and the results indicate that the fractal dimension can improve classification performance. Crisan

et al. (2007) use the box-counting algorithm to measure fractal dimension on digital mammograms, because the fractal dimension is higher in the cancerous area than in the benign area. To classify breast cancer ultrasound images, the fractal dimension features, extracted by the fractal Brownian motion model, are used as features for breast lesion classification on ultrasound images (Kai *et al.*, 2007).

Table 2.3. The morphological features and application domains.

Morphological features	Application domains	Authors
Area	Prostate cancer tissue images Fine needle biopsies cytological images	Naik *et al.* (2007) Filipczuk *et al.* (2013)
Perimeter	Prostate cancer tissue images Fine needle biopsies cytological images	Naik *et al.* (2007) Filipczuk *et al.* (2013)
Compactness	Prostate cancer tissue images	Naik *et al.* (2007)
Smoothness	Prostate cancer tissue images	Naik *et al.* (2007)
Shape factor	Cervical histology image	Naghdy *et al.* (2007)
Roundness	Cervical histology image	Naghdy *et al.* (2007)
Nuclei to cytoplasm (N/C) ratio	Cervical histology image	Naghdy *et al.* (2007)
Enhance roughness index (ERI)	Digital mammograms	Li *et al.* (2013)
Line length	Contrast-enhanced CT of renal cancer Pathological images	Linguraru *et al.* (2009) Xu *et al.* (2008)
Area fraction	Contrast-enhanced CT of renal cancer Pathological images	Linguraru *et al.* (2009) Xu *et al.* (2008)
Quotient	Contrast-enhanced CT of renal cancer Pathological images	Linguraru *et al.* (2009) Xu *et al.* (2008)
Euler number	Contrast-enhanced CT of renal cancer Pathological images	Linguraru *et al.* (2009) Xu *et al.* (2008)
Ratio of area eccentricity	Renal cell carcinoma tissue images	Waheed *et al.* (2007)

- **Topological features**

Topological features are based on the spatial distribution measuring of objects. The Voronoi diagrams and Delaunay triangulations can be applied to measure and record the relationship between neighbouring objects. Delaunay triangulations focus on the number of cells that are connected together, and on the average length of the object connections. The characteristics of the tree connection of the Delaunay triangulation can be used as topological features (Demir & Yener, 2005). For example, the Voronoi Diagram, Delaunay triangulation, minimum spanning tree (MST) and co-adjacency matrix are used to describe the spatial distribution of the nuclei (Doyle *et al.*, 2007).

2.5. Classification step

Classification is the process of categorising or classifying the objects of interest and can be divided into statistical approaches and machine learning approaches. The statistical classification approaches include Bayesian decision rule, k-nearest neighbour classifier, k-means classification and Gaussian Mixture Model (GMM). The machine learning approaches are classification approaches based on artificial intelligence approaches such as a Naive Bayes classifier (NB), Gentleboost algorithm, Fuzzy classifier, Artificial neural network (ANN), Support Vector Machine (SVM), Linear Discriminant Analysis (LDA) classifier, Hopfield neural network (HNN), Fuzzy c-Means (FCM) clustering algorithm, Case based reasoning (CBR) classification algorithm, self-organising map neural network, Kernel Fisher discriminant (KFD), Relevance vector machine (RVM) and Extreme learning machine (ELM) algorithm.

- **Statistical classification approaches**

The classification techniques based on statistical approaches are widely used in cancer classification in various types of images. Waheed *et al.* (2007) use multi-class Bayesian decision rule to classify between benign and malignant renal cell carcinoma on the pathological images. Qi *et al.* (2001) apply the k-means classifier to classify breast cancer regions from thermal infrared images, while Hamdi *et al.* (2008) apply Fisher's linear discriminant to categorise the normal and abnormal microcalcifications on digital mammograms. In addition, the Quadratic Discriminant Analysis (QDA) method is used to classify the features of colon cancer in the histological images, and the results indicate that QDA gives better results than Fisher's score (Esgiar *et al.*, 2007). Meng *et al.* (2008) apply a classifier based on similarity measuring (city block) to compute similarity between the data values of ovarian cancer samples. This method is selected because the data values extracted from samples are binary values. Zheng *et al.* (2007) develop a breast cancer classification application by combining k-means classification with multilayer perceptron network, based on the back-propagation algorithm to train and classify the data set of benign and malignant regions in ultrasound images. Niwas *et al.* (2010b) apply a k-nearest neighbour classifier to classify between the cancerous and non-cancerous cells on cytological images. Petroudi *et al.* (2013) also use a k-nearest neighbour classifier for breast cancer classification. Li *et al.* (2012a) use a classification method based on GMM to classify breast tumour mammographic images. Statistical classification techniques are summarised in Figure 2.6. The main challenge is the assumption that the samples or features are independent. However, for some applications, the samples or features depend on each other (Demir & Yener, 2005). To find the best classifier for

breast cancer diagnosis, Filipczuk *et al.* (2013) compare the classification performance between the statistical based k-nearest neighbour classifier with the machine learning based classifiers: naive Bayes, decision tree and support vector machine. The results show that the machine learning based classifiers perform better than the k-nearest neighbour classifier.
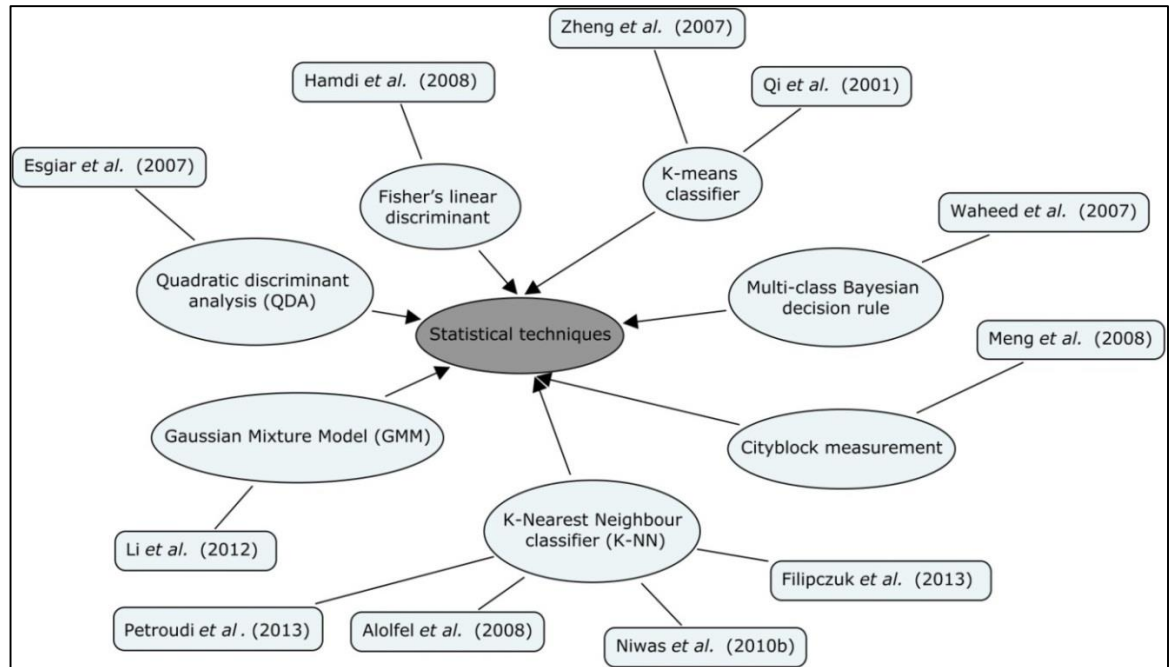
Figure 2.6. Statistical classification techniques.

- **Machine learning classification approaches**

Machine learning classification approaches take advantage of learning from data to classify the classes of objects from each other (Demir & Yener, 2005). The support vector machine (SVM) is a common supervised machine learning approach applied to cancer applications. It is used by Lahmiri *et al.* (2011) to classify cancerous breast cells, by Allwin *et al.* (2010) in their study of cervical cancer and by Han *et al.* (2007) and Naik *et al.* (2007) in their analysis of prostate cancer tissues. Xu *et al.* (2008) and Doyle *et al.* (2007, 2008) have also applied SVM to classify histology image slides, while Yuan *et al.* (2006) have used SVM based on texture only, with acceptable results, though they recommend adding more features to improve the classification. Deepa *et al.* (2012) and Kim *et al.* (2013) use SVM to distinguish between benign and malignant regions in digital mammogram images, whereas Torheim *et al.* (2014) apply SVM to classify MR images of cervical cancer.

To solve the unclassifiable region problems in SVM and to improve the performance of the SVM classifier, Xing-li *et al.* (2008) propose Fuzzy Support Vector Machines (FSVM) to classify the different noise levels of digital mammograms of breast cancer. The experiments yield better results and decrease the computational cost.

Furthermore, Alolfe *et al.* (2008) use four classifiers, such as a SVM, k-nearest neighbour, neural network and fuzzy classifiers, to classify micro-calcification on digital mammography. Due to the small number of samples, the classification results are not satisfactory. To improve the classification, Alolfe *et al.* (2009) apply a new classification method, combining a SVM classifier with Linear Discriminant Analysis (LDA) to classify the breast cancer in digital mammograms, and the results show that the proposed classifier gives greater accuracy than SVM, LDA, and Fuzzy c-Means (FCM) classifiers, whereas Kounelakis *et al.* (2012) apply SVM and Relief-F to classify statistical and biological features of brain gliomas.

As well as SVM, artificial neural networks (ANNs) are widely used for cancer applications. They are used by Niwas *et al.* (2010a) to classify images of breast tissue samples and by Karnan & Gandhi (2010) in their study of digital mammograms. Moradi *et al.* (2006) have applied multi-layer perceptron neural networks, based on the back propagation algorithm, to classify cancer and non-cancer regions, while Mini *et al.* (2011) apply Probabilistic neural network (PNN) to classify breast cancer in digital mammograms.

To improve the classification of ANNs, Neofytou *et al.* (2008) combine PNN with SVM, while Zheng *et al.* (2007) apply a multilayer perceptron network and the k-means classification method to classify ultrasound images of tumours. Naghdy *et al.* (2010) also apply ANN and neuro-fuzzy classifiers to a real time brain cancer classification application. The results demonstrate that the proposed classifiers give precisie classification results. Chui-Mei *et al.* (2008) combine a Self-Organising Map (SOM) neural network with a fuzzy criterion classifier, to classify microcalcification on digital mammograms, whereas Elizabeth *et al.* (2012a) apply a Radial Basis Function Neural Network (RBFNN) to classify lung cancer tomography images. The experiment's results show that the proposed method yields 94.44% accuracy.

A decision tree classifier has also been used in various applications. In the paper by Peng *et al.* (2010a), C4.5 decision tree is applied to classify the cervical nuclei on microscopic images. The results demonstrate that the proposed method yields a promising accuracy (97.8%). To improve the classification of cervical nuclei on microscopic images, Peng *et al.* (2010b) use a C4.5 decision tree as classifier and F-score as a feature selection algorithm. The results demonstrate that the computational cost can be reduced by F-score.

Other classifiers have also been applied to cancer classification applications. Acosta-Mesa *et al.* (2005) use the Naive Bayes classifier (NB) to classify cervical cancer from colposcopic images. Raman *et al.*, (2010) propose a performance based Cased Based Reasoning (CBR) classification algorithm to classify breast cancer in digital mammograms. Vani *et al.* (2010) apply the Extreme Learning Machine (ELM) algorithm to

classify the masses in digital mammograms and the results demonstrate that the ELM algorithm is not affected by local minima, whereas Yaguchi *et al.* (2011) apply Linear Discriminant Analysis (LDA) to classify biopsy images of stomach cancer. Torrent *et al.* (2010) propose the Gentleboost algorithm, based on the concept that the sum of weak classifiers can deliver a strong classifier, and the proposed algorithm improve the robustness of the system. Wei *et al.* (2005) compare the performance of classifiers such as SVM, Kernel Fisher Discriminant (KFD), Relevance Vector Machine (RVM) and committer machines based on Adaptive Boosting algorithm. The results show that SVM, KFD and RVM yield the best performance for microcalcification classification in digital mammograms. George *et al.* (2013) compare four machine-learning based classifiers, namely, Multilayer Perceptron (MLP), PNN, Learning Vector Quantisation and SVM for breast cancer cytological image classification applications. The results show that the two best classifiers of this image type are PNN and SVM. In the paper of Lin *et al.* (2014), Particle Swarm Optimisation (PSO) is applied to classify anomalous areas in digital mammograms of breast cancer. The classification techniques based on machine learning approaches and the application domains are summarised in Table 2.4.

2.6. Conclusion

The review of the literature related to the computational image analysis of cancer demonstrates a much research activity, though limited to the study of breast, lung, skin, cervical and prostate cancers. The literature covering molar pregnancy is confined to the management and care of molar pregnancy (Gul *et al.*, 1997; Khaskheli *et al.*, 2007; Sebire, 2010). Common techniques applied in particular to image processing and segmentation include either statistical methods or artificial intelligence and machine learning approaches. In the study carried out by Yu *et al.* (2010), the C4.5 decision tree yields the most promising accuracy (97.8%) in cervical classification of microscopic images. Xiaojing *et al.* (2006), who apply a set of textural feature vectors and SVM to classify skin cancer with 70% accuracy, suggest further investigation to achieve better results. This study applies the traditional approach to analyse the cancer images consisting of pre-processing, segmentation, feature extraction and classification steps. Segmentation is based on fuzzy c-means clustering and HSV colour space as advocated by Niwas *et al.* (2010a) and Doyle *et al.* (2012) to extract the regions of interest of the cancer stained slides. The literature review also shows that knowledge and expert heuristics are still unexplored and have not been included in these computational approaches. Such expertise has the potential of improving the performance of the proposed neural network configuration aimed at classifying cancerous cells into their appropriate types. Consequently, this project puts great emphasis on eliciting expert

knowledge to guide the system in detecting abnormal cells. It is believed that pathologists' intuitive and tacit knowledge is crucial to guide the novel neural network configuration and to enhance the accuracy of the classification system.

Table 2.4. Classification techniques based on machine learning approaches.

Classification techniques	Application domain	Authors
Support vector machine (SVM)	Cancerous breast cells	Lahmiri *et al.* (2011) and George *et al.* (2013)
	Cervical cancer	Allwin *et al.* (2010) and Torheim *et al.* (2014)
	Prostate cancer tissues	Han *et al.* (2007) and Naik *et al.* (2007)
	Histology image slides	Doyle *et al.* (2007, 2008) and Xu *et al.* (2008)
	Skin cancer images	Yuan *et al.* (2006)
	Digital mammogram images	Deepa *et al.* (2012) and Kim *et al.* (2013)
SVM classifier with linear discriminant analysis (LDA)	Digital mammograms	Alolfe *et al.* (2009)
SVM and Relief-F	Brain gliomas	Kounelakis *et al.* (2012)
Fuzzy support vector machines (FSVM)	Digital mammograms	Xing-li *et al.* (2008)
Fuzzy classifiers	Digital mammography	Alolfe *et al.* (2008)
Neural network	Digital mammography	Alolfe *et al.* (2008)
	Images of breast tissue samples	Niwas *et al.* (2010a)
	Digital mammograms	Karnan & Gandhi (2010)
Multi-layer perceptron neural networks	Prostate cancer images	Moradi *et al.* (2006)
	Cancerous breast cells	George et al. (2013)
Probabilistic neural network (PNN)	Digital mammograms	Mini *et al.* (2011)
	Cancerous breast cells	George et al. (2013)
PNN with SVM	Gynecological cancer	Neofytou *et al.* (2008)
Multilayer perceptron network and k-means	Ultrasound images of tumours	Zheng *et al.* (2007)
ANN and neuro-fuzzy classifiers	Brain cancer	Naghdy *et al.* (2010)
Radial basis function neural network (RBFNN)	Lung cancer tomography images	Elizabeth et al. (2012a)
Self-organising map (SOM) neural network with a fuzzy criterion classifier	Digital mammograms	Chui-Mei *et al.* (2008)
C4.5 decision tree	Cervical microscopic images	Peng *et al.* (2010a)
C4.5 decision tree as classifier and F-score	Cervical microscopic images	Peng *et al.* (2010b)
Naive Bayes classifier (NB)	Cervical cancer images	Acosta-Mesa *et al.* (2005)
Cased based reasoning (CBR)	Digital mammograms	Raman *et al.*, (2010)
Extreme learning machine (ELM)	Digital mammograms	Vani *et al.* (2010)
Gentleboost algorithm	Skin cancer images	Torrent *et al.* (2010)
Learning vector quantisation	Cancerous breast cells	George et al. (2013)
Particle swarm optimisation (PSO)	Digital mammograms	Lin *et al.* (2014)

Chapter 3: Literature Review: Anomaly Detection

3.1. Introduction

The philosophical approach adopted by the study is based on the fact that cancerous cells are best understood by analysing anomalies detected in cancer slides. As Genest (2001) explains, irregular clinicopathological features are an important for scientific study and are utilised by pathologists for supporting their diagnostic classification of problematic cancerous cells. This has led to the investigation of anomaly detection theory and approaches. Anomaly detection aims to detect data patterns that are irregular, or at demonstrating unusual or abnormal behaviour, for example, the detection of unusual traffic patterns in computer networks (Kumar, 2005), or uncommon credit card transactions (Aleskerov *et al.*, 1997). These irregular data patterns are called anomalies, outliers, discordant observations, contaminants, surprises or peculiarities, depending on the context of investigation. While most research projects tend to discard these irregularities, this study considers them as important and critical to the diagnostic classification of cancerous cells. The aims of this chapter are to review current anomaly detection approaches in order to select the most appropriate methods for cancer cell classification.

Anomalies are data patterns that exhibit abnormal behaviour when compared against defined normal behaviour. For example, in a simple spatial data set, as shown in Figure 3.1, a region (Normal) is considered as a normal region. The points, V, W, X, Y and Z, which are far away from a normal region, are considered as anomalies.

Noise removal and noise accommodation are related but distinct from anomaly detection (Rousseeuw & Leroy, 1987; Teng *et al.*, 1990). They are proposed for removing unwanted noise in a data set. Noise is a data pattern of no interest to the study being carried out, and tends to obstruct access to the interesting data. Noise removal is a process that aims at removing the contaminated data before data analysis techniques are applied. Noise accommodation improves statistical model estimation for protecting the system from anomalous observations (Huber, 1974). Anomaly detection and novelty detection are also related. However, novelty detection aims at finding out previously undetected data patterns (Markou & Singh, 2003a; Markou & Singh, 2003b; Saunders & Gero, 2000), so the difference between novel patterns and anomalies is that novel patterns are not integrated into the model of the system until relevant data patterns are

detected. Also anomaly detection focuses on the abnormal data patterns at the beginning of the analysis.
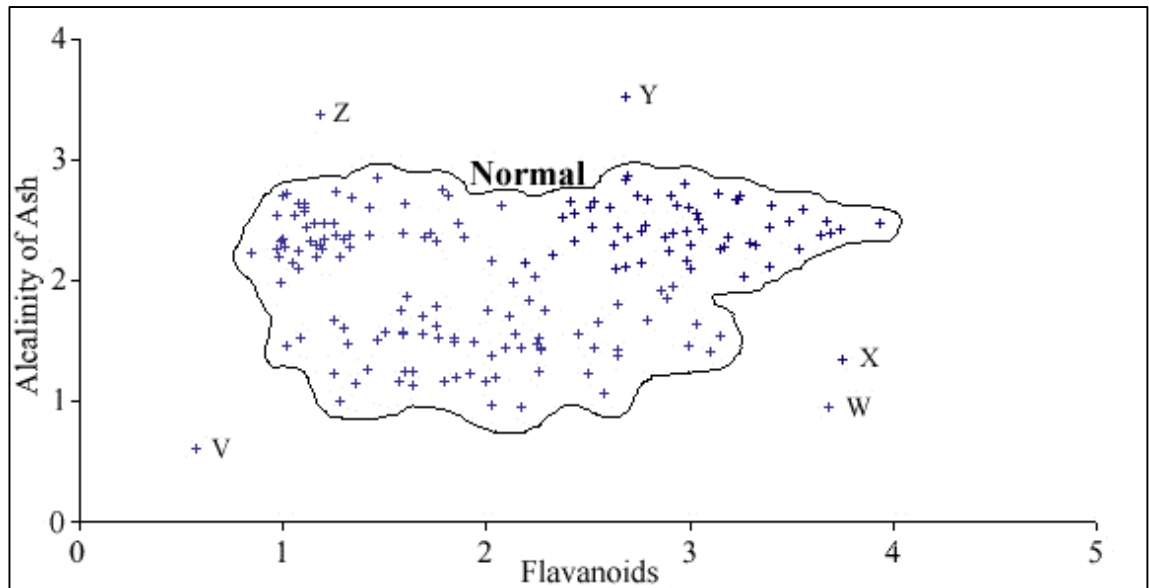
Figure 3.1. An example of anomaly data in two-dimensional data distribution (Blake & Merz, 1998).

3.2. Challenges

Anomalies are data patterns that do not exhibit expected behaviour. An anomaly detection approach defines a normal behaviour region and any data pattern that does not lie in a normal region is considered anomalous. However, this simple anomaly detection approach becomes complex and challenging because of the following factors (Chandola *et al.*, 2009):

(i) It is difficult to define a normal region that can cover all possible normal behaviour data patterns. The imprecise nature of the boundary between normal and abnormal behaviour makes the normal region hard to define. Observations close to the boundary may be considered as normal behaviour.

(ii) Anomalies are also difficult to detect in the detection of criminal or malicious behaviour where this is disguised as normal behaviour. So, in this case, defining normal behaviour is more difficult because malicious actions by criminals can change over time.

(iii) The definition of an anomaly is different depending on the domain. For instance, a small deviation in data value in the stock market domain may be considered as normal, whereas in the medical diagnosis domain it may be considered as an anomaly. So, detection techniques that can operate and yield satisfactory results in one domain may not be applicable in other domains.

(iv) The availability of data for anomaly detection systems for training, testing and validation is another critical challenge.

(v) Noise can be confused with actual anomalies and can make the detection task more difficult. So, noise removal techniques that can remove noise and retain the anomalies must be cautiously applied.

Because of these challenges, anomaly detection is a difficult problem to resolve. Most anomaly detection techniques are applied to solve specific problems in specific domains. So, when the factors are changed the anomaly detection techniques need to be adapted to suit appropriate contexts. Consequently, researchers have applied a variety of anomaly detection techniques from different perspectives, some are statistically based and others use machine learning, data mining, spectral theory, and information theory, all applied in a range of domains from industrial intrusion detection, financial fraud detection, to medical diagnostic detection.

3.3. Fundamental approaches of anomaly detection

This section describes the relevant factors that need to be taken into consideration when developing appropriate anomaly detection systems. The relevant factors consist of the nature of input data, data labels, types of anomalies and the output of anomaly detection.

3.3.1. Nature of input data

The nature of input data is an important issue in anomaly detection techniques. Input can include sets of collection data such as points, objects, records, patterns, vectors, events, samples, cases, entities and observations (Tan *et al.*, 2005). A set of attributes can be used to describe each set of collection data. These attributes can be classified into three types: binary, categorical and continuous.

Sets of data can be involved with others; for example, sequence data are sets of data that are ordered by linear function, such as data recorded in time-series, sequences of genome data and sequences of protein data. In addition, spatial data are sets of data related to neighbours in a spatial domain, such as traffic data and ecological data. Graph data includes sets of data connected by a graph relationship.

3.3.3. Data availability

Anomaly detection can deal with unusual data patterns in real-world problems, and new types of data: normal and anomalous classes need, to be defined before further analysis can take place (Dunning & Friedman, 2014).

Data labels are applied to mark a data sample whether it is in a normal or anomalous class. The labelling process is expensive and normally undertaken by an expert. In addition, this process takes time and several attempts to get labelled data for training processes. In general, in anomaly labelling tasks, a set of anomalous data is more difficult to label than a set of normal data because (i) accessing all possible types of anomalous behaviour is a challenge, (ii) anomalous data patterns are often dynamic in nature, and (iii) new anomalies may arise for which no training data is labelled and available (Chandola, 2009). In addition, when new anomalous data this needs to be labelled and added to a set of anomalous data. Anomaly detection can be categorised into three modes based on the available labels of samples, described below.

- **Supervised anomaly detection**

The anomaly detection techniques that train the data set in a supervised mode are called supervised anomaly detection techniques. These techniques assume that a training data set of normal classes and anomalous classes are both available. Therefore, the system is trained by both normal and anomalous classes. Two major issues need to be considered when supervised anomaly detection is applied. First, the unbalanced in class distribution makes the training process more difficult; the population in the anomalous class is usually smaller than that for the normal class. This issue also occurs in data mining and machine learning application domains. Second, the accuracy and representative labels in an anomalous class are difficult to obtain.

- **Semi-supervised anomaly detection**

The anomaly detection techniques that train the data set in a semi-supervised mode are called semi-supervised anomaly detection techniques. These techniques assume that, in a training data set, only a normal class is labelled and an anomalous class is not. Therefore, the system is only trained with a normal class. These techniques can be applied to a wider range of applications than supervised anomaly detection techniques.

- **Unsupervised anomaly detection**

The anomaly detection techniques that do not train the data set are called unsupervised anomaly detection techniques. These techniques assume the values of normal data are far away, in term of geometrical distance, from the values of anomalous data, and distant enough to classify by the classification techniques. Thus, the system can operate without a training process. However, these techniques may have a high false alarm rate when the boundaries of normal and anomalous classes cannot be clearly distinguished.

3.3.2. Types of anomalies

Anomalies can be categorised into three categories: point anomalies, contextual anomalies and collective anomalies.

- **Point Anomalies**

A point anomaly is a data sample with different individual characteristics from the rest of samples. For example, points V, W, X, Y and Z in Figure 3.1 are considered anomalous, due to their position and distance with respect to the region labelled "Normal". In financial transactions the exceptionally high amount spent on a recurring item can be considered as a point anomaly.

- **Contextual Anomalies**

A contextual anomaly is a data sample considered to be anomalous in a specific context. This type of anomaly is also referred to as conditional (Song *et al.*, 2007). Two types of attributes are identified: contextual and behavioural attributes. Contextual attributes are attributes defined for the neighbourhood or the context of a sample of data; for instance, the longitude and latitude of a location or the position in x-axis and y-axis are the context attributes in spatial domains. Behavioural attributes are those determined by non-contextual characteristics. For example, in a dataset of the average world rainfall in a given year, a behavioural attribute is the amount of rainfall at a specific location, whereas the longitude and latitude of a rainfall location are its contextual attributes.

- **Collective Anomalies**

A collective anomaly is a collection of data that is different from the entire data set. Some data samples are not considered an anomaly as individual data values, but when collected together, may be considered as a collective anomaly. For example, in the human electrocardiogram output shown in Figure 3.2 (Dunning & Friedman, 2014), the samples of data between 1206 and 1210 in the x-axis are collective anomalies because they remain at the same at low value for long time.
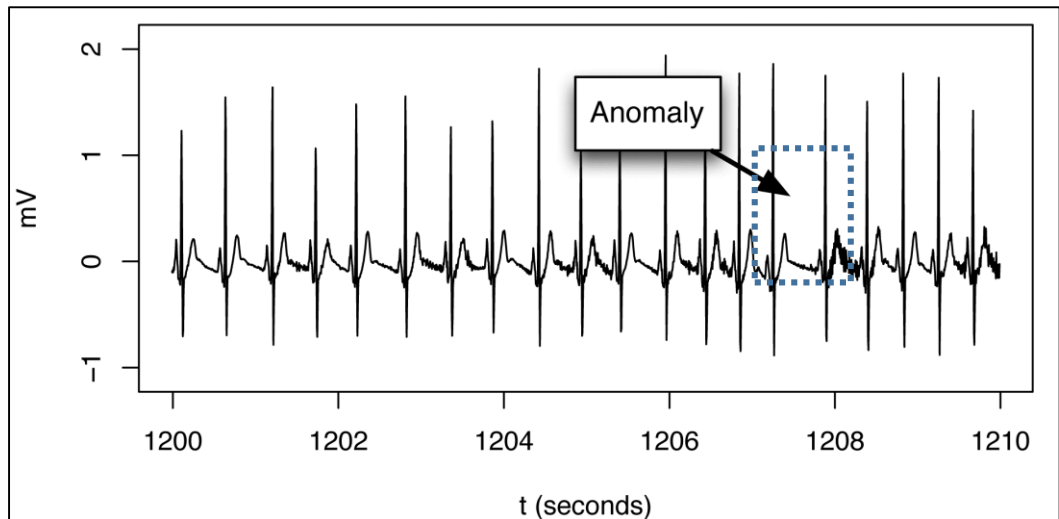
Figure 3.2. Collective anomaly samples in a human electrocardiogram output (Dunning & Friedman, 2014; 13).

3.3.4. Output of anomaly detection

An important issue in anomaly detection is how to report the results. The traditional outputs, which are given by anomaly detection techniques, can be categorised into types: scores and labels.

- **Scores**

Scoring techniques are applied to assign the anomaly score to each data sample. The anomaly score depends on the degree to which a data sample is considered an anomaly. Users or the system can select a certain number of the more significant anomalies to analyse, or can select the interesting anomalies by using a threshold.

- **Labels**

Labelling are techniques that assign a label to an instance as normal or anomalous. The user can use a threshold to select the interesting anomalies when scoring techniques are applied. On the other hand, label techniques do not allow the user to control directly the outputs, so the user needs to control the labelling process indirectly by changing the parameters in each technique.

3.4. Anomaly detection techniques

In this section, various techniques applied to detect anomalies are described. The techniques are based on statistical, machine learning, information theoretic and spectral approaches.

3.4.1. Statistical anomaly detection techniques

Statistical techniques are the primary means whereby anomalies are detected (Hodge & Austin, 2004). Statistical techniques assume that in a given stochastic model anomalies appear in its regions with low probability, whereas normal data instances appear with high probability. Parametric and nonparametric techniques are both applied to fit a statistical detection model. These two techniques are discussed in the next two sections.

3.4.1.1. Parametric techniques

The evaluation time of parametric techniques is fast when the techniques are used to test a new data instance, and these techniques are appropriate for large data sets. Parametric techniques are based on the assumption that normal data can be generated by a parametric distribution calculated by a probability density function $f(x, \theta)$ where θ is estimated from the given data and x denotes an observation. The anomaly score of each observation x can be computed by the inverse of the probability function. Sometimes, another statistical hypothesis test based on discordancy test might be used (Barnett & Lewis, 1994). If the statistical test rejects the null hypothesis, then the observation x is declared anomalous. Parametric techniques can be classified into three sub-groups as follows:

- Gaussian model

Some techniques are based on the assumption that the data is generated from a Gaussian distribution. Maximum likelihood estimates (MLE) are used to calculate the distance between the estimated mean and a data instance, and used as the anomaly score for that instance. Shewhart (1931) proposes a basic outlier detection technique for his quality control domain application. This technique defines the data instances that are more than 3σ from the mean μ, as outliers. Another simple statistical technique is the box plot rule technique, which is applied to detect both univariate and multivariate anomalies in medical domain applications (Laurikkala *et al.*, 2000; Horn *et al.*, 2001; Solberg & Lahti, 2005), and in the data of turbine rotors (Guttormsson *et al.*, 1999). A box plot depicts graphically the data using summary attributes in quartiles.

Grubbs' test (known as the maximum normed residual test) is based on the assumption that data is generated by a Gaussian distribution, and this test is applied to detect anomalies in a univariate data set (Grubbs, 1969; Stefansky, 1972; Anscombe & Guttman, 1960). The standard deviation and mean are calculated and a Z score (i.e. a standard score) for each test data instance is declared an anomaly if

$$Z > \frac{N-1}{\sqrt{N}} \sqrt{\frac{t^2_{\alpha/(2N),N-2}}{N-2+t^2_{\alpha/(2N),N-2}}}$$

where N is the size of the data and $t^2_{\alpha/(2N),N-2}$ is defined as a threshold for separating the anomalous and normal data instances. Laurikkala *et al.* (2000) apply the Grubb's test to multivariate data and reduce multivariate data samples to univariate scalars using the Mahalanobis distance, which is a measure of the distance between a point and a distribution. This distance is calculated by the following equation:

$$y^2 = (x - \bar{x})' S^{-1} (x - \bar{x})$$

Where x and $\bar{x}$ are a test data instance and the mean of the sample, respectively, and S is the variance-covariance matrix of the sample. Other approaches using variants of Grubbs' test are applied to deal with multivariate data sets (Aggarwal & Yu, 2001; 2008; Laurikkala *et al.*, 2000). Shekhar *et al.* (2001) apply the test to handle graph-structured data and Sarawagi *et al.* (1998) propose online analytical processing (OLAP) data cubes. Surace & Worden (1998) and Surace *et al.* (1997) apply the t-test to detect damages in the structure of beams.

- Regression model

Regression model, based anomaly detection techniques, has been extensively explored in time series data (Abraham & Chuang, 1989; Abraham & Box, 1979; Fox 1972). The traditional regression model involves two steps. First, the data is used to fit a regression model. Second, the residual of each data instance, which cannot be fitted by a regression model, is used to determine the anomaly score.

Anomalies in the training data can affect the regression model and give inaccurate results. To solve this problem, Rousseeuw & Lerory (1987) proposed a robust regression to deal with anomalies when fitting a regression model, whereas Bianco *et al.* (2001) and Chen *et al.* (2005) proposed Autoregressive Integrated Moving Average (ARIMA) models, based on a similar approach to robust regression. Tsay *et al.* (2000) improved ARIMA models using a variant to operate in multivariate data, and to develop ARMA further Galeano *et al.* (2004) transformed the multivariate time series to a univariate time series by linearly combining the multivariate time series components.

- Hybrid approach

These techniques are based on modelling the data using a mixture of parametric statistical distributions. These techniques can be categorised into two groups. The first group uses the mixture of parametric distributions to model both normal and anomalous

data instances. It is concerned with classifying the data instance into normal or anomalous distributions. Abraham & Box (1979) propose the technique using the assumption that normal data is generated from a Gaussian distribution $\left(N(0, \sigma^2)\right)$ while the anomalies, calculated from a Gaussian distribution, have larger variance $\left(N(0, k^2\sigma^2)\right)$ than normal data with the same mean. The same labelling techniques, using the Grubbs' test on both normal and anomalous data, are used by Lauer (2001), Eskin, (2000) Abraham & Box (1979), Box & Tiao (1968) and Agarwal (2005). The second group of techniques generates a mixture model of the normal instances. So, the data instances that cannot be classified into the trained models are defined as anomalies. Agarwal (2006) uses Gaussian mixture models to create a model of the normal data instances. This second group has been applied by Hickinbotham & Austin (2000a) and Hollier & Austin (2002) to detect airframe strain data, by Spence *et al.* (2001) and Tarassenko (1995) to detect anomalies in mammographic images, to detect network intrusion (Yamanishi & Takeuchi, 2001; Yamanishi *et al.*, 2004) and in biomedical signal data (Roberts & Tarassenko, 1994; Roberts, 1999; 2002). Byers & Raftery (1998) use a mixture of Poisson distributions to create the model of normal data instances for anomaly detection.

3.4.1.2. Nonparametric techniques

Nonparametric techniques detect anomalies based on nonparametric statistical models and on assumptions regarding the data, for example smoothness or density, to create the model. Nonparametric techniques can be categorised into two sub-groups as follows:

- Histogram

Histogram based techniques use histograms to represent data instances, based on frequency or counting. Histogram based techniques are commonly used to detect fraud (Fawcett & Provost, 1999) and intrusion detection (Eskin, 2000; Eskin *et al.*, 2001; Denning, 1987). The simple histogram based anomaly detection technique for univariate data can be categorised into two steps. First, a histogram is built on the different values of the training data features. Second, the test instance is tested by the built histogram. The test instance is declared an anomaly if the instance does not fall into any bin of the built histogram. An anomaly score for each instance is computed by a variant of the frequency (height) of the bin in which it falls. The key to the performance of a histogram based on anomaly detection technique is the size of the bin. If the size of the bins is too small the technique will give a high false alarm rate, and if the size of the bins is too large, the technique will give a high negative rate. For this reason, the optimal size of the bins is a challenge in this histogram technique.

This technique is also applied to multivariate data to detect anomalies in various applications, such as in system call intrusion data (Endler, 1998), network intrusion detection (Ho *et al.*, 1999; Yamanishi & ichi Takeuchi, 2001; Yamanishi *et al.*, 2004), structure damage detection (Manson, 2002; Manson *et al.*, 2001; 2000), web-based attack detection (Kruegel & Vigna, 2003; Krügel *et al.*, 2002), text data detection (Allan *et al.*, 1998), fraud detection (Fawcett & Provost, 1999). Packet Header Anomaly detection (PHAD) and Application Layer Anomaly Detection (ALAD) are a variant of the simple histogram technique used in network intrusion detection (Mahoney & Chan, 2002) and online wireless sensor network anomaly detection (Xie *et al.*, 2012).

- **Kernel function**

Kernel functions based on anomaly detection techniques are related to a non-parametric technique and apply probability density estimation, called Parzen windows estimation (Parzen, 1962). The difference between a kernel function-based technique and Parzen windows estimation is the way density is estimated. A semi-supervised statistical technique, based on kernel functions is introduced to find anomalies (Desforges *et al.*, 1998). A test data instance located in the low probability area is classified as an anomaly. Chow & Yeung (2002) use this technique in network intrusion detection applications, while Bishop (1994) applies it for novelty detection in oil flow data instances. Tarassenko (1995) uses a similar technique for a mammographic image analysis application.

3.4.2. Machine learning based anomaly detection techniques

Machine learning is a means of learning a model (classifier) from a set of labelled data (training sets). Then, the trained classifier is then used to classify the unknown data sets into the classes of the learned model. In the training step, these techniques learn a model using the available labelled data, and then, a set of test data is classified into normal or anomalous class by the classifier in the testing step. The assumption of machine learning based anomaly detection techniques is that a classifier can be trained by features from a feature space to classify between normal and anomalous classes. The learning steps of machine learning based anomaly detection techniques are illustrated in Figure 3.3.

- **Neural networks**

Stefano *et al.* (2000) and Odin & Addison (2000) divide a neural networks-based anomaly detection technique into two steps. First, a set of normal data is fed to a neural network that learns the difference between normal classes. Second, the trained neural network is applied to classify the test input. If the classification system accepts the input,

the input data is categorised into normal classes and if the network rejects the input data then it is categorised into an anomalous class.
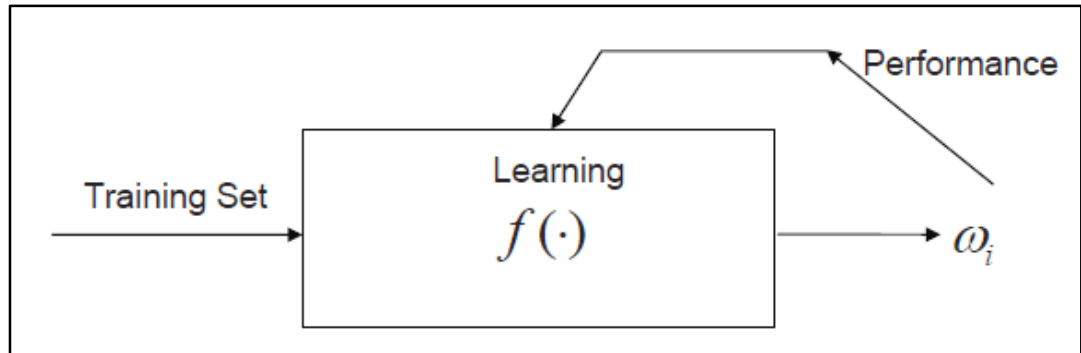
Figure 3.3. Machine learning based anomaly detection techniques (Thottan *et al.*, 2010).

Hawkins *et al.* (2002) and Williams *et al.* (2002) apply Replicator Neural Networks to detect anomalies. A multi-layer feed forward neural network is generated and used to train the data using three hidden layers. Multi layered percreptrons (MLPs) are used for anomaly detection in various application domains, such as satellite imagery (Augusteijn & Folkert, 2002), digit recognition (Cun *et al.*, 1990), Cushing's syndrome data (Sykacek, 1997), intrusion detection (Ghosh *et al.*, 1999; 1998), fraud detection in mobile phone networks (Barson *et al.*, 1996), jet engine vibration data (Nairac *et al.*, 1997) and airframe strain data (Hickinbotham & Austin, 2000b). The reliable comparisons of neural networks such as MLPs, Gaussian MLP (GMLP) and radial basis function (RBF) networks are studied in the paper of Vasconcelos *et al.* (1995; 1994). The studies show that RBF is the most reliable among three networks for two dimensional synthetic data. In addition, Martinez (1998) proposes neural competitive learning trees to identify novelty in a driving performance overtime data set, and the proposed method yields better performance than other methods (standard competitive learning (standard CL) and frequency sensitive competitive learning (FSCL)).

Auto-associative networks are also applied to detect anomalies in diverse research areas, for example, novelty detection in dynamic behaviour (Aeyels, 1991), syntactic data (Byungho & Sungzoon, 1999; Ko & Jacyna, 2000; Song *et al.*, 2001), helicopter gearbox fault detection (Japkowicz *et al.*, 1995), standard Reuters data (Manevitz & Yousef, 2000), motor failure detection (Petsche *et al.*, 1996), synthetic data of hard disk drives (Sohn *et al.*, 2001), turbine generator rotors (Streifel *et al.*, 1996), network hub CPU usage data (Thompson *et al.,* 2002), damage of structure (Worden, 1997) and motor vibration data (Diaz & Hollmen, 2002). Instead of using auto-associative networks, Moya *et al.* (1993) apply an adaptive resonance theory based technique to identify anomalies in satellite

images, whereas Dasgupta & Nino (2000) use this technique for a network intrusion detection application. Caudell & Newman (1993) also apply the adaptive resonance theory based technique to detect anomalies in time series and databases. Albrecht *et al.* (2000) use a technique based on radial basis function for anomaly detection in a speech recognition application, while Bishop (1994) applies this technique to identify anomalies in oil flow data instances. Radial basis function-based techniques are also applied to detect anomalies in various applications, such as military aircraft (Brotherton & Johnson, 2001), a diesel engine cooling system (Li *et al.*, 2002), jet engine vibration data (Nairac *et al.*, 1999; 1997), a credit card fraud detection application (Ghosh & Reilly, 1994), and fault detection in the automotive industry (Jakubek & Strasser, 2002).

Anomaly detection techniques based on Hopfield networks are used in various applications. For instance, Jagota (1991) uses Hopfield networks to detect anomalies in large memory collections, whereas Crook & Hayes (2001) and Crook *et al.* (2002) apply the networks to detect novelty in robot behaviour. Addison *et al.* (1999) study the effect of feature extraction to novelty detection by Hopfield networks, while Murray (2001) applies the networks to embedded systems.

Oscillatory networks are other networks used to detect anomalies. Ho & Rouat (1997; 1998) apply the oscillatory networks to detect a novelty in digit number datasets and the results indicate that the proposed method is better than Hopfield and backpropagation networks. Kojima & Ito (1999) also apply the oscillatory networks to identify novelty in character patterns. Borisyuk *et al.* (2000) propose sparse distributed memory neural networks based on oscillatory networks. The proposed method is tested with syntactic data and the results show that the networks can detect novelty in the storage of input signals. Martinelli & Perfetti (1994) apply a cellular neural network (CNN) based on oscillatory networks to identify novelty in two-dimensional syntactic data. The results show that the proposed method is robust to noise, but further experiments need to be carried out to verify the method.

- **Bayesian networks**

Bayesian networks are used to evaluate a set of normal and anomalous classes. The probabilities estimated by a Bayesian network indicate whether the test data belongs to a normal or anomalous class. This technique is applied to detect network intrusion (Barbara *et al.*, 2001; Sebyala *et al.*, 2002; Valdes & Skinner, 2000; Mingming, 2000; Bronstein *et al.*, 2001), anomaly in text data (Baker *et al.* 1999) and outbreaks (Wong *et al.*, 2002; 2003). Mascaro *et al.* (2014) apply static and dynamic Bayesian network models to detect anomalies in vessel tracking applications. The overall performance of combined models is better than using a static or dynamic model alone. Dereszynski & Dietterich

(2007) also apply a dynamic Bayesian network model for anomaly detection in remote sensors.

- **Support vector machines**

Ratsch *et al.* (2002) apply SVM based on a one-class learning technique to learn the training boundary of data sets. The common radial basis function (RBF) is used as the Kernel of SVM. In the testing phase, the input data sets are fed to the classification system. If the system accepts the input, the input data is classified into the normal class and if the network rejects the input data, it is categorised as the anomalous class. SVM based on anomaly detection techniques are applied to detect audio signal data (Davy & Gosil, 2002), to intrusion detection systems (Eskin *et al.,* 2002). King *et al.* (2002) proposed a novelty detection method based on SVM in power generation plants, whereas Catania *et al.* (2012) apply an autonomous labelling approach to label well-known attacks to improve the performance of a novelty detection system based on SVM. Görnitz *et al.* (2013) integrate an automatic filter to SVM to enhance the performance of a network intrusion detection system, whereas Song *et al.* (2013) modify a filtering process to improve SVM performance of an intrusion detection application. Song *et al.* (2002) proposed a robust support vector machines (RSVM) to identify the anomalies in training data sets. Instead of applying RSVM to improve training sets, Hu *et al.* (2003) implement the RSVM to detect anomalies in intrusion detection systems, and the results show that the RSVM achieves a better accuracy and lower false alarms than a nearest neighbour based method.

- **Rule-based approach**

The steps of a basic rule-based technique can be separated into two stages. First, a set of normal data is used to train a rule learning algorithm, namely decision trees and/or RIPPER with associated confidence values (Paredes-Oliva *et al.*, 2012). Second, the classification system uses these rules to classify the test input. The anomaly score is based on the inverse of the confidence associated with the best rule. Agrawal & Srikant (1995) use association rule mining for anomaly detection classification with an unsupervised mode to create the rules from customer transaction data, whilst Tan *et al.* (2005) apply a support threshold to improve the robustness of the generated rules. Association rule mining techniques are used by Mahoney & Chan (2002; 2003) for intrusion detection, Lee *et al.* (2000) apply the technique in system call intrusion detection, whereas Brause *et al.* (1999) apply them to detect credit card fraud and Yairi *et al.* (2001) for fraud detection in spacecraft housekeeping data.

3.4.3. Other approaches

3.4.3.1. Nearest neighbour-based anomaly detection techniques

The nearest neighbour based anomaly detection techniques assume that anomalies occur far from their nearest neighbours, whereas normal data group in neighbourhood. These techniques need distance or similarity measurement tools to measure distance or similarity between two data sets. Euclidean distance is typical for continuous data. Nevertheless, for other types of data, Boriah *et al.* (2008) and Chandola *et al.* (2008) use a simple matching coefficient and more sophisticated distance measures for categorical data. For multivariate data, Tan *et al.* (2005) compute distance and similarity for each data attribute and combine distance or similarity together. Nearest neighbour based anomaly detection techniques can be categorised into two major groups:

- **Techniques that calculate the anomaly score using the distance between a data set and its kth nearest neighbour.**

Byers & Raftery (1998) use this technique for land mines detection on satellite ground images, and Guttormsson *et al.* (1999) use this for anomaly detection in DC field windings of large synchronous turbine generators. To improve the performance of this basic technique, various approaches are proposed. The anomaly score of a set of data is calculated by the sum of the distances with its k-nearest neighbours (Eskin *et al.*, 2002; Angiulli & Pizzuti, 2002; Zhang & Wang, 2006), and is used in peer group analysis in credit card fraud detection (Bolton & Hand, 1999), whereas Knorr *et al.* (2000) propose a different way of computing the anomaly score by counting the number of nearest neighbours that are at a defined distance. Om & Kundu (2012) combine k-means, k-nearest neighbour and Naive Bayes to reduce the false alarm rate of intrusion detection systems. Xie *et al.* (2013) apply k-nearest neighbours based on hypergrid intuition to detect anomalies in wireless sensor networks (WSNs) and which solves a lazy learning issue of the traditional k-nearest neighbours for WSNs anomaly detection.

- **Techniques that calculate the anomaly score using the relative density of each data set.**

The data instance located in a low-density neighbourhood is classified into an anomalous class, and the data instance located in a high-density neighbourhood is classified into a normal class. The disadvantage of this technique is that it operates unsuccessfully in regions with varying density. In Figure 3.4, the data instance p_1 is detected as an anomaly but the instance p_2 will not be detected, because the low density of the cluster C_1 makes the distance between the instances in the cluster C_1 greater than the distance between p_2 and the instances in the cluster C_2. Nevertheless, the instance p_2 might be considered as an anomaly because the distance is greater than the instances in

both the C_1 and C_2 clusters. To solve the low density problem, Breunig *et al.* (1999; 2000) propose a local outlier factor (LOF) that can compute the ratio of the average local density of the k-nearest neighbours of an instance and the data instance local density. With LOF the number k of k-nearest neighbours is divided by the size of the defined hypersphere, whereas COF (connectivity based outlier factor) estimates densities by the shortest path of neighbours (Tang *et al.*, 2002). COF can capture some regions better than LOF, for example, the straight line, as shown in Figure 3.5. Hautamaki *et al.* (2004) propose outlier detection using in-degree number (ODIN), a simple version of LOF and the anomaly score of the instance can be computed by the inverse of ODIN. Papadimitriou *et al.* (2002) introduce Multi-Granularity Deviation Factor (MDEF) that uses the variance of LOF as the density value of the data instance and the inverse of the standard deviation is defined as the anomaly score.
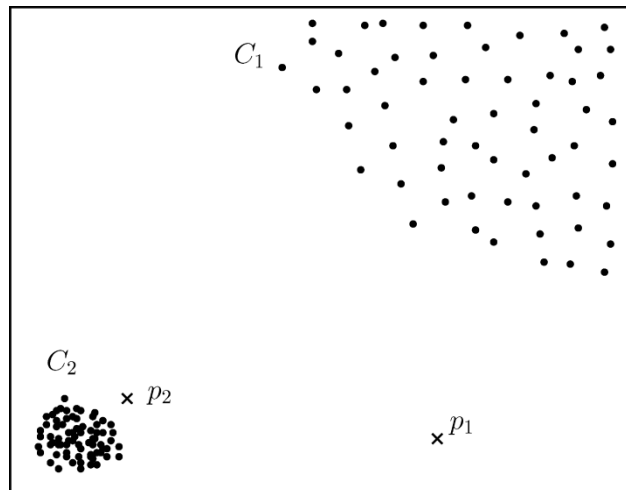
Figure 3.4. Advantages of local density-based techniques over global density-based techniques (Chandola *et al.*, 2009; 15:25).
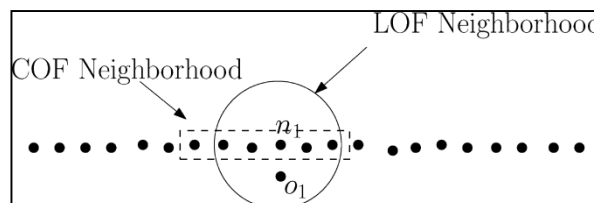
Figure 3.5. Difference between the neighbourhoods computed by LOF and COF (Chandola *et al.*, 2009; 15:25).

3.4.3.2. Clustering

The major difference between clustering-based and nearest neighbour-based techniques is that nearest neighbour-based techniques evaluate each data instance relative to its local neighbour data instances, whereas clustering-based techniques validate each data instance relative to the cluster that the instance belongs to.

Clustering is applied to categorise data instances that have similar characteristics or behaviour into groups or clusters. Basu *et al.* (2004) extend one technique from an unsupervised technique to semi supervised clustering. Clustering based anomaly detection techniques can be categorised into three groups.

The first group of clustering based techniques depends on the assumption that anomalies cannot be categorised into any cluster, while normal data instances can be categorised into a cluster. Ester *et al.* (1996) propose an algorithm based on this assumption and this algorithm is called DBSCAN. Guha *et al.* (2000) propose the ROCK algorithm and Ertoz *et al.* (2003) propose SNN clustering. Yu *et al.* (2002) develop the FindOut algorithm extended from the WaveCluster algorithm (Sheikholeslami *et al.*, 1998). The primary objective of these techniques is to find clusters, and not to detect anomalies.

The second group of clustering based techniques depends on the assumption that anomalies are located far from the closest cluster centroid, whereas normal data instances are located close to the closest cluster centroid. The techniques relying on this assumption can be divided into two steps. First, a clustering algorithm is applied to categorise the data. Second, the distance to the closest centroid of each data instance is computed and used as the anomaly score. Anomaly detection techniques based on this assumption are widely used. Smith *et al.* (2002) apply k-means clustering, self-organising maps (SOM) and expectation maximisation (EM) to categorise training data, and used the clusters to categorise the test data. Kohonen (1997) uses SOM based on anomaly detection in semi-supervised mode for the classification system. In addition, Labib & Venuri (2002), Smith *et al.* (2002) and Ramadas *et al.* (2003) apply the SOM technique to intrusion detection application domains, whereas Harris (1993), Ypma & Duin (1998) and Emamian *et al.* (2000) use the SOM technique for fraud detection applications. Furthermore, to deal with sequence data, various techniques are proposed (Blender *et al.*, 1997; Bejerano & Yona, 2001; Vinueza & Grudic, 2004; Budalakoti *et al.*, 2006).

However, the techniques in the second group are unable to detect the anomalies that form clusters by themselves. So to solve this problem, the third group of clustering based anomaly techniques has been introduced. This group depends on the assumption that anomalies lie in sparse clusters while normal data instances lie in large and dense clusters. Therefore, clusters that have a smaller size or less density than the defined threshold are anomalous. He *et al.* (2003) define an anomaly score, called Cluster-Based

Local Outlier Factor (CBLOF) using the FindCBLOF technique. This score measures the size of the cluster as well as the distance of a data instance from the centroid of its cluster. To improve the performance of the previous techniques, Eskin *et al.* (2002) and Portnoy *et al.* (2001) apply fixed width clustering, called a linear time approximation algorithm. They define the width of clusters by a user-specified parameter, whereas Mahoney *et al.* (2003) define the width of clusters from the data. The k-d trees based anomaly detection techniques are used to segment the linear time data in astronomical datasets (Chaudhary *et al.*, 2002) and the modified k-d trees technique using the partition hyperplane method (namely special cut) is applied to separate data into clusters.

3.4.3.3. Information theoretic approach

The information content of a data set is analysed by information theoretic techniques, which use different measures, for example, Kolmogorov complexity, entropy and relative entropy. Kolmogorov complexity refers to the shortest object descriptive code and can be used to describe an object, whereas entropy is an uncertainty measure of a random variable, and relative entropy is a distance measure of two probability distributions. Information theoretic anomaly detection techniques are based on the assumption that the irregularity information content of the data set is influenced by anomalous data in the data set. These techniques compute the complexity C of a data set D. It finds the minimal subset of instances, I, that maximise the following equation:

$$C(D) - C(D - I)$$

The data instances in the subset are declared to be anomalies. Various measuring methodologies of the complexity of a data set (C) are proposed Li & Vitanyi (1993) introduced Kolmogorov complexity, which was used by Arning *et al.* (1996) to measure the size of the regular expression, and by Keogh *et al.* (2004) to measure the size of the compressed data. Other measuring theories, such as entropy and relative uncertainty, have also been applied (Lee & Xiang, 2001; He *et al.*, 2005; 2006; Ando, 2007). He *et al.* (2006) apply a Local Search Algorithm (LSA) to estimate a subset of linear data and use entropy to measure the complexity, whereas Ando (2007) applies the same technique to measure the complexity of information bottleneck data

Lin *et al.* (2005), Chakrabarti *et al.* (1998) and Arning *et al.* (1996) apply the information theoretic technique to sequential data, whereas Noble & Cook (2003) apply this technique to graph data. Additionally, Lin & Brown (2006) apply spatial data. The key challenge of information theoretic based anomaly detection techniques is the optimal size of the substructure that can give satisfactory results.

3.4.3.4. Spectral anomaly approach

Spectral anomaly detection techniques estimate the data by using a combination of attributes that can take the volume of unsteadiness in data. These techniques are based on the assumption that normal data instances appear differently from anomalies when data are embedded to a lower dimensional subspace but anomalous data instances show different results.

Principal component analysis (PCA) is a typical Spectral anomaly detection technique. Robust PCA (Huber, 1974) is applied to calculate the principal components from the normal training data covariance matrix (Shyu *et al.*, 2003). An anomaly score can be computed by using their distance from the principal components.

Spectral anomaly detection based on robust PCA is used for the network intrusion detection application domains (Shyu *et al.,* 2003; Lakhina *et al.*, (2005); Thottan & Ji, 2003), whereas Fujimaki *et al.* (2005) apply this technique to detect anomalies in spacecraft components. Lee *et al.* (2013) propose an over-sampling principal component analysis (osPCA) algorithm to deal with two-dimensional synthetic data and real-world data sets, and the results show that osPCA achieves better accuracy and efficiency than a traditional PCA. Jolliffe (2002) apply PCA to project data into a subspace, whist Parra *et al.* (1996) use the projection of data instances that are the results of PCA to distinguish between normal and anomalous data. Dutta *et al.*, (2007) also apply this method for anomaly detection in astronomy catalogues.

Sun *et al.* (2007) implement an anomaly detection technique using Compact Matrix Decomposition (CMD) to identify anomalies in a sequence of graphs. The proposed method uses the estimated errors as criteria for anomaly detection. A spectral technique is also applied to detect anomalous data in a time series of graphs (Ide & Kashima, 2004).

3.5. Anomalies Detection Issues

Based on the above review of current anomaly detection methods, it is important to identify their strengths as well as their limitations before deployment.

One of the key aspects of anomaly detection techniques that it is necessary to consider is the computational complexity. Machine learning and statistical techniques require expensive training times, but in the testing stage, the techniques often operate rapidly. These two types of anomaly detection techniques can be usually trained in an offline mode and the test data instances can be tested in a real-time. On the other hand, nearest neighbour-based, information theoretic and spectral techniques do not require a training process but these techniques have expensive testing times.

To detect anomalies in complex data sets, nearest neighbour and clustering approaches cannot cope with a high dimensionality, because their distance measures cannot distinguish normal from anomalous instances. Spectral techniques perform well because they can map high dimensionality data to lower dimension. Machine learning techniques require labels for normal and anomaly instances, which are difficult to get.

Generally, anomaly detection techniques are based on the assumption that anomalies rarely occur in data sets. However, this assumption is not always true for some applications, such as worm detection in computer networks, because the worm traffic anomalies are more frequent than the normal traffic. So, unsupervised techniques are difficult to operate in this case, whereas supervised and semi-supervised techniques perform well in detecting bulk anomalies (Sun *et al.*, 2007; Soule *et al.*, 2005).

3.6. Applications of anomaly detection

Anomaly detection is applied to many applications such as intrusion detection, fraud detection, industrial damage detection, medical and public health anomaly detection applications. This section focuses primarily on medical image processing applications.

3.6.1. Medical and public health anomaly detection

In medical and public health application domains, anomaly detection techniques operate with the patient records, such as abnormal patient conditions, errors of instruments and errors of recording, focusing on data related to age of patients, blood types and weights. Wong *et al.* (2003) use a contextual anomaly detection method based on a Bayesian network for disease outbreak detection, whereas Lin *et al.* (2005) apply collective anomaly detection techniques to detect anomalies in Electrocardiograms (ECG) and Electroencephalograms (EEG). On the other hand, Li *et al.* (2012b) use k-nearest neighbours to detect anomalies in ECG and to visualise and analyse the ECG signals. Salem *et al.* (2013) apply the anomaly detection method based on Haar wavelet decomposition and Hampel filters to identify anomalies in online medical wireless sensor data, achieving high accuracy and reducing false alarm rates. Neural networks are applied to explore the detection of possible breast lesions in medical digital imaging (Ferrero *et al.*, 2006), to detect signs of acute myocardial infarction (AMI) in ECGs. Burke *et al.* (1995) conclude that the prediction accuracy of neural networks was more accurate than statistical methods in predicting 5-year survival of 25 cases used in their study. Hauskrech (2007) investigate probabilistic models such as Bayesian networks to detect unusual patient-management decisions. Their aim is to develop computational tools using collected patient data to detect unusual patient-management patterns that can eventually

alert clinicians to unusual treatment choices. Babbar & Chawla (2010) propose Bayesian networks as a technique to capture real outliers and represent causal knowledge of two medical domains: hepatitis and breast cancer. Churilov *et al.* (2005) describe a clustering method to extract risk-grouping rules for prostate cancer patients making use of the patient's age, tumour stage, Gleason score, and PSA level. Goldstein & Dengel (2012) present an unsupervised histogram-based anomaly detection algorithm to model univariate feature densities using histograms with a fixed or dynamic bin width. These histograms are used to compute an anomaly score for each data instance from the breast cancer data set. They conclude that the Histogram-based Outlier Score (HBOS) is up to 5 times faster than clustering based algorithms and up to 7 times faster than nearest-neighbour based methods.

3.6.2. Image processing related applications

Anomaly detection techniques are applied in various types of static images, for example, satellite imagery (Augusteijn & Folkert, 2002; Byers & Rafter, 1998; Moya *et al.*, 1993; Torr & Murray, 1993; Theiler & Cai, 2003), spectroscopy (Chen *et al.*, 2005; Davy & Godsill, 2002; Hazel, 2000), functional magnetic resonance imaging (Scarth *et al.*, 1995), bullet hole image classification (Song *et al.*, 2002), digit recognition (Cun *et al.*, 1990) and mammographic image analysis (Spence *et al.*, 2001; Tarassenko, 1995). Furthermore, to detect anomalies in hyperspectral images, various techniques are proposed. Khazai *et al.* (2011) apply a adaptive support vector method based on a global Support Vector Data Description (SVDD) to detect anomalies in hyperspectral images and propose a fast-adaptive support vector method to improve the performance of the previous anomaly detection technique (Khazai *et al.*, 2013), whereas Messinger & Albano (2011) propose an anomaly detection technique based on a graph theoretic approach. This technique has a comparable performance with the traditional anomaly detection based on Mahalanobis distance, namely the RX algorithm. Liangliang *et al.* (2010) improve the kernel RX algorithm by applying an anomaly detection technique based on background endmember extraction. This technique can improve the performance of anomaly detection based on the kernel RX algorithm. Ma *et al.* (2010) use local tangent space alignment (LTSA) to reduce the dimension of hyperspectral image data sets. Mousazadeh & Cohen (2010) apply an anomaly detection technique based on non-causal autoregressive-autoregressive conditional heteroskedasticity, which is a statistical measure used to calculate the different variance of random variables from others, (AR-ARCH) to detect anomalies in sonar images. Liu & Zheng (2011) propose the technique based on discrete cosine (DC) parameters extracted from JPEG images. This technique gives better performance than a YUV (Y is the brightness component and U and V are colour

components) image data processing method. Anderson *et al.* (2012) apply a combination of anomaly techniques based on size-contrast filters and mean shift clustering, Gaussian mixture model and maximally stable extremal regions algorithms to detect explosive hazards on a road. The results show that the proposed techniques can detect explosive hazard on a road as well as human visual inspection.

Another task of anomaly detection in image processing is video surveillance (Diehl & Hamshire, 2002; Singh & Markou, 2004; Pokrajac *et al.*, 2007; Saligrama & Chen, 2012). Li *et al.* (2014) use an anomaly detection method based on a mixture of dynamic texture models for crowd scenes detection. The results confirm that the proposed method is better than a traditional anomaly detection method. Instead of using anomaly detection for video surveillance applications, Khazai *et al.* (2013) apply a single-feature based anomaly detector (SFAD) to detect anomalies in sub-pixels of hyperspectral images and the proposed method achieves promising detection results. The types of anomalies in this application domain are point and contextual anomalies. The major challenge in this domain is the large amount of data input, and anomaly detection techniques need to operate online with video data.

3.6.3. Other domains

Anomaly detection techniques are applied to various application domains listed below.

- I**ntrusion detection**

Intrusion detection refers to the detection of dangerous activity or malicious behaviour that can harm a computer system (Phoha, 2002). Anomaly detection techniques are appropriate for intrusion detection applications because the intrusion behaviour is different from the ordinary behaviour of the system. The challenge of anomaly detection in these applications is the dynamic changes in data introduced by intruders, which make detection very challenging.

Jeong *et al.* (2010) apply a hierarchical approach to detect anomalies (e.g. scanning attacks, Distributed Denial of Service (DDos) and Worm) in computer network systems, whereas Wang *et al.* (2010) use a fast anomaly detection technique based on the estimation of Hurst parameter for network traffic anomaly detection.

Zhao *et al.* (2010) apply the Bayesian inference of statistical methods to improve the performance of a machine learning anomaly detection technique in network intrusion detection, whereas Om & Kundu (2012) apply an intrusion detection system integrating k-means, k-nearest neighbour and Naive Bayes classifiers to detect anomalies in KDD-99 data set, to reduce the false alarm rate. Shanbhag *et al.* (2010) evaluate six algorithms in

network traffic: Holt-Winter forecasting, cumulative sum, wavelet analysis, k-means, support vector machine (SVM) and one-class neighbour machine (OCNM). They conclude that a combination of algorithms provides a better accuracy than one single algorithm, and accurate anomaly detection is achieved by applying a diversity of algorithms. Paredes-Oliva *et al.* (2012) use frequent item-set mining (FIM) and C5.0 algorithms to recognise network traffic anomalies. On the other hand, Choraś *et al.* (2012) apply an anomaly detection framework based on Matching Pursuit (MP) algorithm to detect anomalies in network traffic.

In network intrusion, Catania *et al.* (2012) use an autonomous labelling approach to improve a novelty detection system based on SVM, whereas Görnitz *et al.* (2013) propose a semi-supervised anomaly detection method based on SVM. Tartakovsky *et al.* (2013) apply an online anomaly detection based on Shiryaev–Roberts's procedure to detect intrusion activities. In the paper of Khorchani *et al.* (2012), a modal logic, namely visibility logic (VL), is applied to detect anomalies in firewall rules. To improve anomaly detection based on hidden Markov models (HMMs), Khreich *et al.* (2012) propose an intrusion detection system based on a receiver operating characteristic (ROC), and efficiently adapt ensembles of HMMs (EoHMMs).

- **Fraud detection**

Fraud detection involves criminal activity detection in commercial and financial organisations. This type of detection is divided into credit card fraud detection, mobile phone fraud detection, insurance claim fraud detection and insider trading detection. Noto *et al.* (2012) apply feature regression and classification (FRaC) to study mobile phone fraud detection problems by monitoring their usage activity. Statistical profiling using histograms is applied by Donoho (2004) and Aggarwal (2005) and an information theoretic approach is adopted by Arning *et al.* (1996). Gaber *et al.* (2013) implement a user's behaviour model for fraud detection in mobile transaction data. The primitive results show that the proposed model can be used as a prototype for fraud detection, whereas Wei *et al.* (2013) propose online banking fraud detection method based on data mining, namely ContrastMiner. The results indicate that the proposed method achieve better detection results than traditional fraud detection methods. Instead of using a data mining based technique, Anderka *et al.* (2014) apply sequence-based anomaly detection to detect anomalies in ATM data.

- **Anomaly detection in textual data**

Anomaly detection plays an important role in detecting new topics or new articles in text data. A clustering based anomaly detection technique is one of the techniques

applied to detect anomalies in text data, for example, anomalies in text reports (Srivastava & Zane-Ulman, 2005) and anomalies in reports of aerospace problems (Srivastava, 2006). Instead of using a clustering technique, Manevitz & Yousef (2000) apply neural networks to retrieve information in the standard Reuters data set, and the results show that the proposed method can be used to detect anomalies in text data. Manevitz & Yousef (2002) also test one-class support vector machines for information retrieval in the standard Reuters data set. The experimentation results confirm that neural networks are more robust than SVM for anomaly detection in text data.

- **Sensor networks**

Anomaly detection techniques are applied in a network of sensors, to detect fault in a sensor. Failures of sensor events are unusual, so the anomaly detection techniques are appropriate tools to identify sensor failures events. The challenge of this application is the need to operate with online data from a network of sensors.

Various techniques are applied to detect anomalies in sensor networks. Bayesian networks are used to identify outliers in wireless sensor networks (Janakiram *et al.*, 2006) and to detect anomalies in data streams of remote sensors (Dereszynski & Dietterich, 2007). Instead of using Bayesian networks-based techniques, Curiac & Volosencu (2012) propose a detection technique combining a set of classifiers based on average based classifier, autoregressive linear predictor, neural networks, neural network autoregressive predictor and adaptive neuro-fuzzy inference system (ANFIS), to detect anomalies in wireless sensor networks. Moshtaghi *et al.* (2014) use the anomaly detection method based on an adaptive model of clustering ellipsoids, to detect attacks and fault sensors in the IBRL dataset (Intel Berkeley research lab). The proposed method achieves better detection results than using a traditional method solely based on clustering ellipsoids.

3.7. Summary

In this chapter, a review of types of anomalies, current techniques and associated challenges have been discussed. A description of some important domain applications has also been presented, and issues have been discussed.

This research advocates the adoption of anomaly detection techniques to the analysis of stained slides of molar pregnancy; this approach is novel for molar pregnancy cancer. The literature review reveals that most image processing approaches have focused on MRI, digital mammograms, ultrasound images or histological images while this research is focused on stained slide images that represent a new type of images. The primary focus of current approaches is either on improving segmentation or classification of images. The focus of this research is to identify anomalies in stained slides, combining

artificial intelligence techniques with experts' strategic knowledge. The anomaly detection approach has been useful in non-medical domains and some medical domains such as disease outbreak detection (Wong *et al.*, 2003), anomaly detection in online medical wireless sensors (Salem *et al.*, 2013), anomalies in Electrocardiograms (ECG) (Lin *et al.*, 2005; Li *et al.*, 2012), and Electroencephalograms (EEG) (Lin *et al.*, 2005). This has inspired us to extend this approach to the novel medical domain of molar pregnancy. As seen from the literature review and anomaly detection applications, this approach has not yet been applied to molar pregnancy cancer and to this type of medical images.

The elicitation of expert pathologists and histopathologists' knowledge is aimed at guiding the anomaly detection approach in identifying the types of anomalies, discussed in Section 3.3.2, which can be related to the morphological characteristics of partial and complete hydatidiform mole villi. These types of anomalies are represented ontologically and guide the segmentation and classification stages in identifying anomalies at each image analysis step, to allow the handling of any uncertainties and fuzziness in the images.

Chapter 4: A Heuristic Based Approach to Anomalies Detection of Hydatidiform Mole Villi (Low Level Processing)

4.1. Introduction

This chapter describes a novel approach to analyse the villi of hydatidiform mole (HM) stained sides. This approach focuses on detecting anomalies of HM by analysing computationally the morphological features (e.g. red blood cells, trophoblast, and stroma) of its villi. The analysis of villi and detection of their anomalies is based on the elicitation of pathologists' heuristic knowledge and strategies, which are embedded in the image analysis and anomalies detection. This approach is based on a cognitive understanding of image analysis, to mirror the pathologists' heuristic diagnosis of anomalies in villi.

This chapter reviews the important contribution of expert knowledge to medical image analysis and explains how the pathologists' tacit knowledge and strategies are elicited and integrated in the developed heuristic approach aimed at detecting anomalies in HM slides. The analysis and anomalies detection is divided into two main phrases: low level processing, which is discussed in this chapter, and high level processing, described in the next chapter.

4.2. Review of experts' knowledge medical image analysis

The literature shows that experts' knowledge plays an important role in medical image analysis. For example, Hamarneh *et al.* (2009) apply expert segmented binary masks of the cardiac muscle of each of the images to construct the shape histograms and the histograms, are used to guide the watershed algorithm in the segmentation step, whereas Martin *et al.* (2010) use MRI images that are manually segmented by experts to construct atlas-based-segmentation-methods of a deformable model. Kang *et al.* (2013) apply the Fuzzy c–Means (FCM) based clustering algorithm and qualitative medical knowledge on geometric properties of different tissues for the tissue classification system. McKenna *et al.* (2012) use anonymous knowledge works trained by manually segmented premalignant colorectal polyps image, to improve the interpretation of computer-aided detections for CT colonography. Pitiot *et al.* (2004) apply a series of rules derived from analysing template's dynamics, based on experts' experience, to develop deformable templates for classifying MRI of brain images, whereas Perner (2002) uses expert

knowledge to develop a decision tree algorithm to diagnose Hep-2 cell-images. McInerney *et al.* (2002) develop a deformable model based on expert strategies to analyse MRI brain images. The organisms of the deformable model are created by expert analysis routines.

The current contributions of expert domain knowledge tend to be limited to diagnosis, segmentation or evaluation purposes. This approach captures not only their expertise but also, more importantly, their heuristics and strategies in dealing with uncertainties and fuzziness. The overall aim is to support histopathologists in detecting anomalies and classifying hydatidiform moles into their appropriate types. The proposed approach is implemented in two phases: the low-level processing phase consists of segmentation and feature extraction steps for processing HM slides obtained from the pathologists, and the high-level processing focuses on classification of slides. The pathologists' heuristic knowledge and strategies for identifying anomalies in the morphological features of the villi are applied at both levels and represented ontologically. The following sections discuss the steps taken to process the HM slides at the low level processing: the high level processing is explained in the next chapter. The two levels are illustrated in Figure 4.1.

4.3. Anomaly detection in molar pregnancy villi

One of the many challenges of this research is that, in early pregnancy, CHM and PHM cannot be easily distinguished from other anomalous pregnancy diseases (Sebire *et al.*, 2003). The distinction between CHM and PHM is important for determining the appropriate counselling and treatment of patients. The diagnosis of these moles continues to be a problem for many practicing and experienced histopathologists because, in early pregnancy, CHM and PHM are difficult to distinguish morphologically from other abnormal pregnancy products (Sebire *et al.*, 2003; Sumithran *et al.*, 1996; Howat *et al.*, 1993). Paul *et al.* (2010) explain that in many cases morphological analysis is inadequate for making a confident diagnosis and that the histological features of complete moles at an early gestation are frequently confused with partial moles, hydropic miscarriage or non-molar chromosomal abnormalities.

First phrase: low level processing

Histogram equalisation

Pre-processing

FCM on HSV colour space

Segmentation

Anomaly Detection

Feature Extraction

Morphology, statistics and texture

Second phrase: high level processing

Classification

Anomaly Detection

Normal placental villi

PHM villi

CHM villi

Experts' knowledge and strategies: ontological representation

- Molar_pregnancy
- Complete_mole
  - Villi_CM
    - No_normal_placental_tissue
    - Red_blood_cells_CM
      - 'Few_or_no_red_blood_cells_(RBC)'
    - Stroma_CM
      - Enlarged
      - Irregular
      - Karyorrhexis
    - Trophoblast_CM
      - Oval/irregular
    - Vessel_CM
      - Collapsed_and_inconspicuous
      - Fetal_absent
- Nomal_placenta
  - Villi_NP
    - Red_blood_cells_NP
      - 'Nucleated_(RBC1)'
    - Stroma_NP
      - 'Contain_Hofbauer_cells_(rounded_shape)'
      - 'Reticular_(like_a_net)'
      - Without_stroma_fibrosis
    - Trophoblast_NP
      - Cytotrophoblast
      - Syncytiotrophoblast
    - Vessel_NP
- Other_abnormalities
- Partial_mole
  - Villi_PM
    - Red_blood_cells_PM
      - Nucleus
        - 'Nucleated_fetal_(RBC2)'
    - Stroma_PM
      - Hydrops
        - Mild
      - Shape
        - Scalloping/dentate
    - Trophoblast_PM
      - Excess
        - Focal
        - Less_marked
      - Irregular_proliferation
    - Vessel_PM
      - Collapsed
      - Fetal_present
      - Inconspicuous

Figure 4.1. The system architecture.

The analysis faces other challenges, some of which are related to the limited material available for analysis, a problem shared by the expert histopathologists themselves, and others to the angle of sectioning these villus samples, which is entirely random. Also some villi may have a perfect cross-section, whereas others are sectioned either longitudinally or obliquely. This may in part explain the variation in size and shape of villi within samples. One particular concern expressed by the expert pathologists is that PHM stained slides can contain PHM and normal placenta villi, which emphasises the complexity of the detection task. To address some of these issues, it was agreed that the analysis of each villus is best carried out with respect to other villi in the image.

Given the above challenges, the detection of anomaly types is based on the study and detection of irregular features of villi. Four main types of anomalies are identified. For example, a point anomaly is used to analyse abnormal percentage of red blood cells (RBC), based on the knowledge that normal villi contain a higher proportion of RBC than CHM and PHM. Contextual anomalies help analyse the relative percentage of stroma region inside a given villus, as a high percentage of the stroma region is usually found in normal placental and PHM villi and a low percentage is found in CHM villi. Morphological anomalies relate to villi shape and size. An example of a contextual anomaly is the size of the stroma region in normal villi (Figure 4.2 (a)), which is larger than the stroma regions found in PHM (Figure 4.2(b)) and CHM villi (Figure 4.2(c)). Density anomalies refer to villi that display markedly reduced vessel density or show significantly more staining. These anomalies, which are elicited from experts and from medical documents, are described below in Section 4.7.
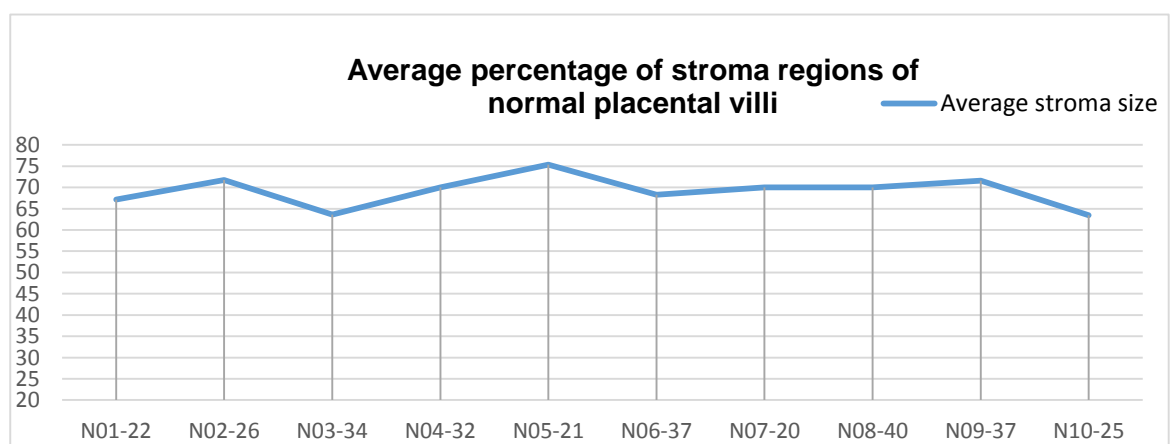
Figure 4.2. (a) Average percentage of stroma regions of normal placental villi.

56
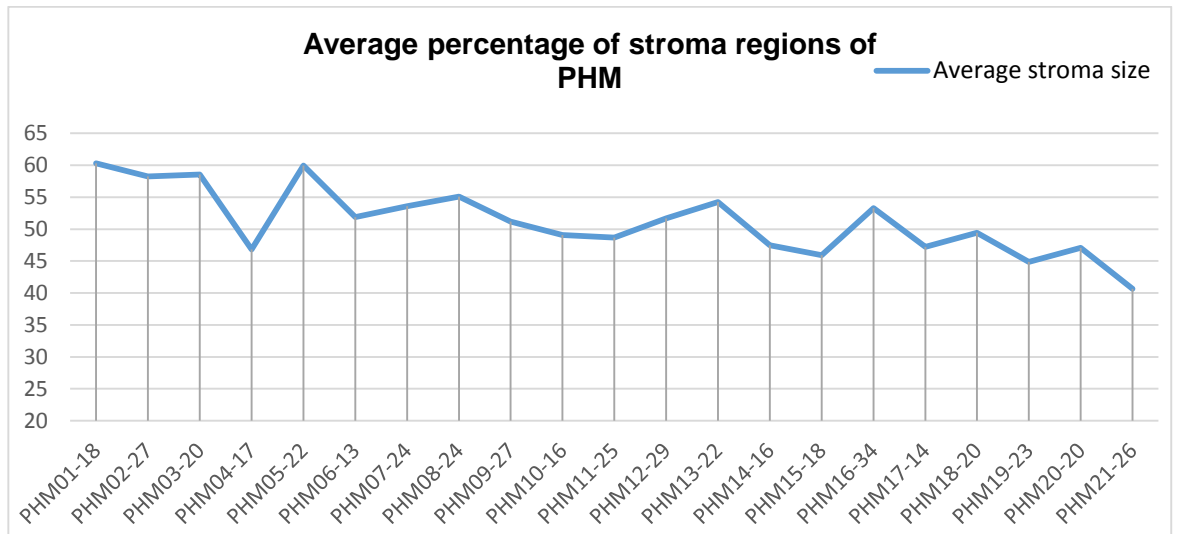
Figure 4.2. (b) Average percentage of stroma regions of PHM.
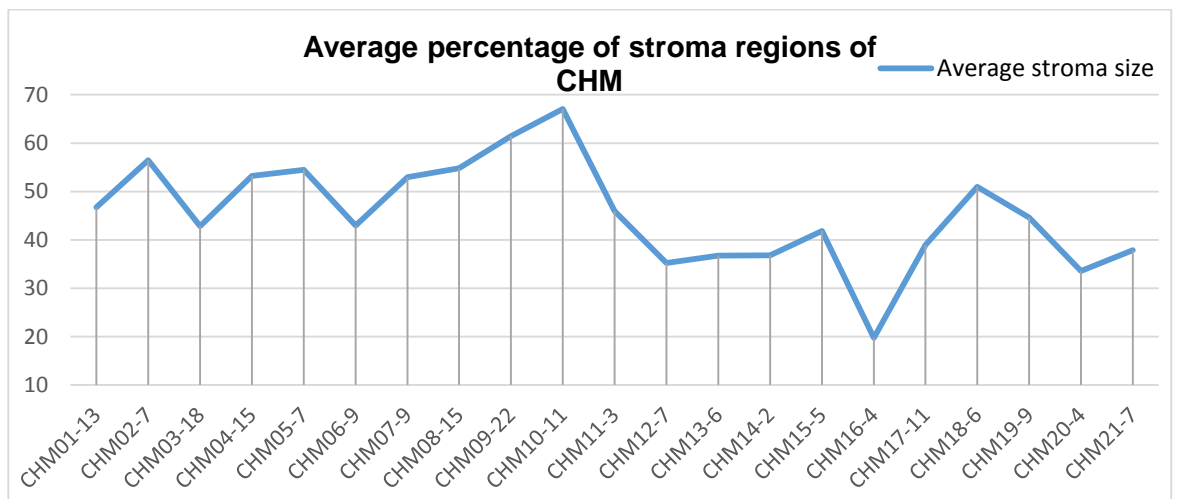
Figure 4.2. (c) Average percentage of stroma regions of CHM.

4.4. Knowledge elicitation of HM anomalies

Knowledge elicitation is the expert knowledge acquisition process that can be used to capture the knowledge not only from experts but also from many sources (Shadbolt & Smart, 2015). This process is used to develop and to deploy knowledge-based systems. High-quality knowledge acquisition to create a robust and useful system is an expensive activity and time-consuming. The knowledge elicitation technique can be categorised into interviews, protocol analysis, critical decision method, concept sorting, repertory grids, laddered grids, limited information task and concept mapping and process mapping.

An interview technique aims at extracting the elicit information about how a specific task is executed or how to make an accurate decision, whereas protocol analysis

(PA) aims at developing essential structure, rules and processes from the protocols. A critical decision method (CDM) combines interviewing and protocol analysis (Zsambok & Klein, 1997) while concept sorting uses experts' knowledge to sort the priority of a fixed set of concepts. The repertory grids technique aims at revealing a conceptual map of a domain created by the concept sorting technique (Shaw & Gaines, 1987), whereas laddered grids are designed to create a hierarchy graph representing the relationship between domain and problem-solving elements. A limited information task technique aims at extracting the experts' strategies by using the questions that experts ask the elicitor to narrow down the information scope required by the elicitor (Hoffman, 1987). Concept mapping and process mapping are based on diagramming techniques focusing on developing a 2-dimensional network of labelled grids and nodes (Milton, 2012).

The limitations of knowledge elicitation are categorised as follows:

(i) Knowledge elicitation techniques are different, and no official techniques can be a universal and effective tool (Hart, 1985).

(ii) The problems of the interview technique are that experts might repeat themselves when they try to explain their knowledge. So the elicitor needs to focus on the important issues during the interview.

(iii) The protocol analysis is effective when recorded case histories are available.

The difficulty in extracting tacit knowledge from experts is a well-known challenge. One of the difficulties related to the fact that experts' knowledge resides in their head and often it is stored subconsciously. The conversion of their tacit knowledge and heuristics to explicit knowledge is the most challenging process, as experts may find it difficult to communicate or may not be consciously aware of the knowledge (Tagger, 2005). This was evident in this research project as knowledge elicitation from the two expert pathologists was carried out over seven meetings taking 2-3 hours each. In the early stages of meetings, two expert pathologists provided guidance to understand fundamental knowledge of molar pregnancy, and essential strategies to build the molar pregnancy ontological representation. The following two meetings discussed the critical morphological features of normal placental, PHM and CHM villi used by the experts to distinguish between these three types of villi. This was followed up by three further meetings to discuss the developed ontological representation of hydatidiform moles and the results obtained from the segmentation and classification tasks. In addition, frequent phone calls, skype meetings and emails were used to clarify unclear issues, to elicit

further knowledge, to request additional villi samples, and to discuss the need for supplementary experimentation results. The phone calls and skype meetings varied between 30 - 45 minutes. The elicitation of the knowledge and strategies of pathologists evolved during these sessions and allowed them to identify additional hidden strategies. These meetings have helped gaining an understanding of how to analyse and interpret the villi in the stained slide images. In addition to the experts' knowledge the researcher extracted the domain knowledge of molar pregnancy from the medical documents; these two types of knowledge were captured in an ontological representation (Palee *et al.*, 2013).

4.5. Ontological representation of anomalies in villi

Gruber defines ontology as `the specification of conceptualisations, used to help programs and humans share knowledge' (Gruber, 1993). The conceptualisation expresses knowledge about the world in terms of entities (things, relationships and constraints). The specification is the representation of this conceptualisation in a concrete form, encoded in a knowledge representation language. As yet, there are no standardised methodologies for building ontologies, so a pragmatic and task-oriented approach to building an ontology was taken. The ontological representation of the morphological features of HM and of their anomalies has been developed with the help of the two expert pathologists. It consists of concepts representing the entities of the HM domain and their critical regions: the vessel, trophoblast, RBC and stroma regions (Figure 4.3). Each region is further described in terms of its specific/anomalous characteristics and their relationship to a given region.
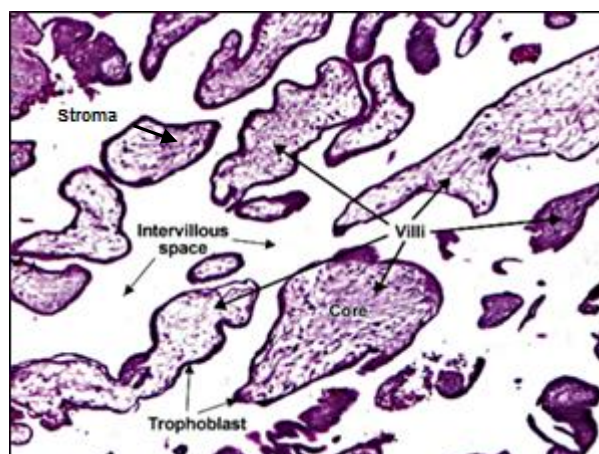

Figure 4.3. Villus' morphological features.
(http://www.pathologyatlas.ro/uterine-curretage-biopsy.php)

In summary, villi are ramifications of placenta with fetal vessels that are the "business end" of the placenta, covered with trophoblast. Stroma is the connective tissue core of an organ. Trophoblast is the epithelium that covers the placenta (Benirschke *et al.*, 2006). The morphological features of normal placental villi are a regular trophoblast, non-stroma fibrosis, a reticular (like a net) stroma structure, defined red cells and rounded villi. The PHM villi are marked by: an excess of focal trophoblast, irregular trophoblast proliferation, stroma hydrops and mild stroma, scalloping and dentate stroma, collapsed vessel, fetus present in a vessel, and presence of some red blood cells. The CHM villi are detected by their oval or irregular trophoblast, enlarged and irregular stroma, stroma karyorrhexis, collapsed and inconspicuous vessel, absence of fetus, few or no red blood cells and irregular villi shape. The CHM trophoblast appearance is usually thicker than PHM and normal placental villi, due to the non-polar proliferation. The stroma's texture, its irregular shape and size are also significant features in CHM, PHM and normal placental villi (Seckl *et al.*, 2010). The top level of ontological representation of hydatidiform moles is shown in Figure 4.4; further description of the ontology is given in Appendix B.

4.6. HM Data

Tissues, which are obtained via spontaneous or surgical uterine evacuation, are processed, embedded and then sectioned into thin 2-5 micrometers sections and then stained, usually with haematoxylin and eosin, to enhance the contrast of the nucleus and cytoplasm within cells for microscopic examination. These stained slide images are examined by the pathologists to classify the type of hydatidiform mole.

The analysis is carried out on the stained slide images captured and pre-classified and pre-labelled as normal, PHM, CHM. by the pathologists who were based at the University Hospital Bristol and London Great Ormond Street Hospital, using a microscope at 40-times-magnification and 20-times magnification. These two levels of magnification are important: although the 40-times-magnification can capture more villi details such as trophoblast thickness and stroma structure, it cannot capture large villi because of the narrow field of view as shown in Figure 4.5a. To capture complete villi the analysis is also applied to the 20-times magnification stained slide images (Figure 4.5b). A total of 986 villi were analysed, consisting of 215 pre-classified complete hydatidiform moles (CHM), 467

partial hydatidiform moles (PHM) and 304 normal villi images, extracted from 52 stained slide images (10 normal, 21 PHM and 21 CHM villi slide images). An example of each type of HM is shown in Figure 4.6. The villi in each stained slide, which was provided by the pathologists, are assigned a new identification number. For example, Nxx-yy denotes normal placental villi, PHMxx-yy denotes partial hydatidiform moles and CHMxx-yy denotes complete hydatidiform moles. So, N03-21 denotes normal placental villi no. 21 of a villi slide image no. 3. Further details of pre-labelled villi images used in feature extraction steps are discussed in Section 4.10.
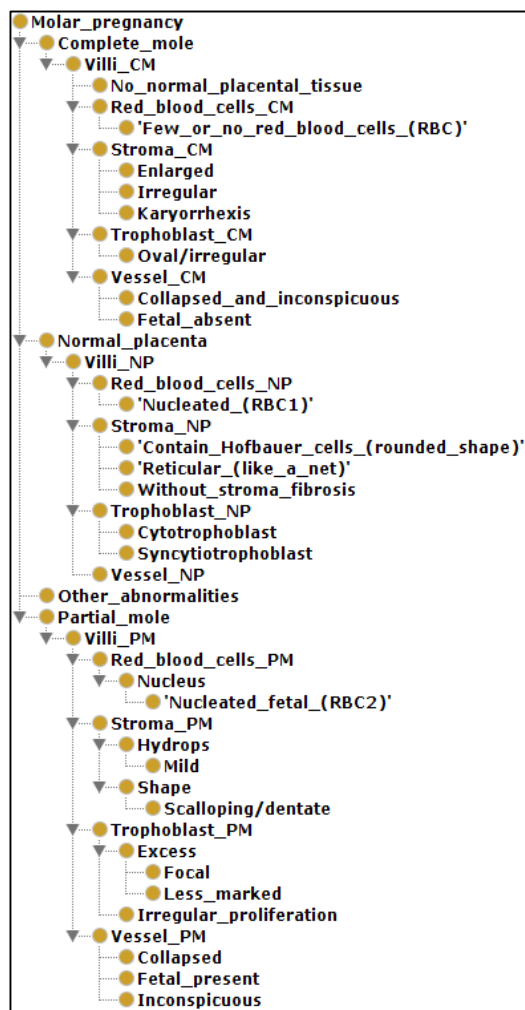
Figure 4.4. A top level of ontological representation of hydatidiform moles.
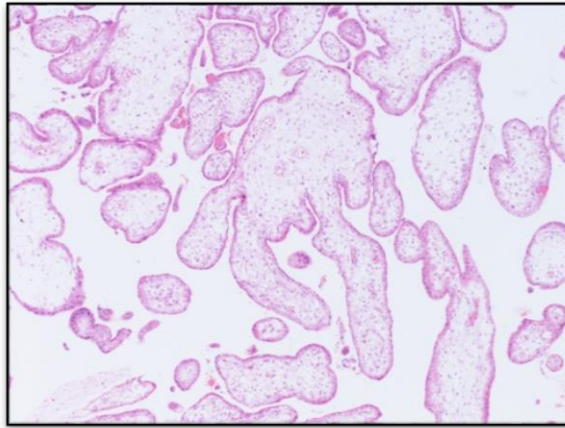
Figure 4.5. (a) A 40-times-magnification HM
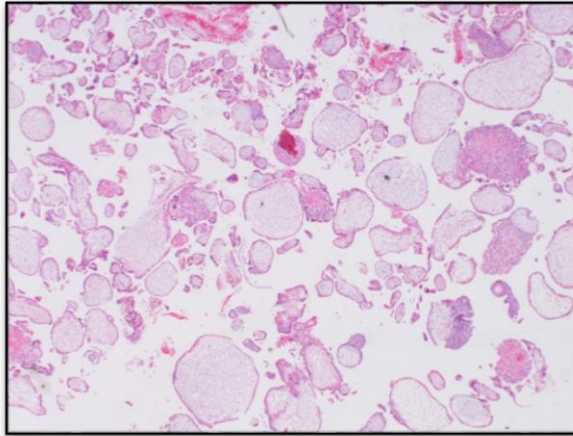
Figure 4.5. (b) A 20-times-magnification HM
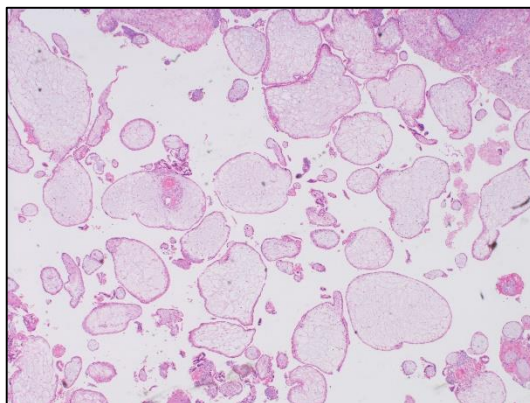
Figure 4.6 (a) A 20-times-magnification
normal placenta stained slide image.

Stroma hydrops and mild
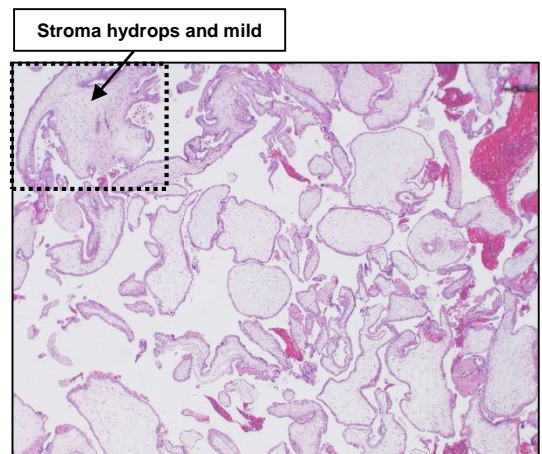
Figure 4.6. (b) A 20-times-
magnification PHM stained slide image
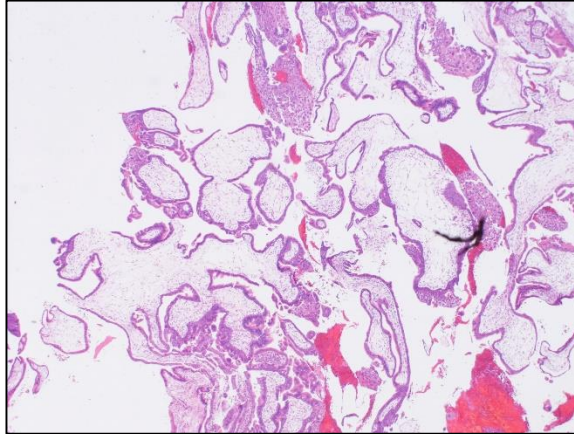
Figure 4.6. (c) A20-times-magnification CHM stained slide image

4.7. Image analysis guided by the ontological representation

The ontological representation of HM morphological features and their anomalies, described in Section 4.5 and Figure 4.4, is integrated at the segmentation and feature extraction steps, which are the low level phase of the image processing analysis. This phase combines HSV colour space with Fuzzy c-Means (FCM) clustering algorithm to analyse the stained HM slides against a set of four critical morphological features and associated anomalies (Table 4.1):

1) Red blood cell (RBC) feature: density and distribution within given villi:
    a) The percentage of red blood cells inside stroma regions is calculated by computing the percentage of segmented red blood cells located inside stroma boundaries.

2) Villi's shape and size:
    a) Villi pixels are measured by counting all pixels segmented as villi regions by the segmentation algorithms.
    b) The major axis is the longest axis of villi boundaries. As shown in Figure 4.7 the major axis is the diameter of an object.

c) The minor axis is the shortest axis of villi boundaries. The minor axis is the distance between villi's boundaries located along the line that is perpendicular to the major axis as shown in Figure 4.7.

d) The elongation ratio is the ratio between major and minor axis.

e) The difference area between villi's bounding box and villi areas is measured by subtracting the villi areas from villi bounding box. This feature is applied to capture the irregularity of villi shape.

f) The notion of four quadrants (e.g. A, B, C and D in Figure 4.7), which is proposed by the two experts, is applied when the shape of PHM and CHM villi is non-rounded and irregular.

g) The number of villi boundary corner points detected by the corner detection algorithms.

h) The ratio between the number of villi boundary corner points and all pixels belonging to villi perimeter is computed. This feature indicates scale robustness.
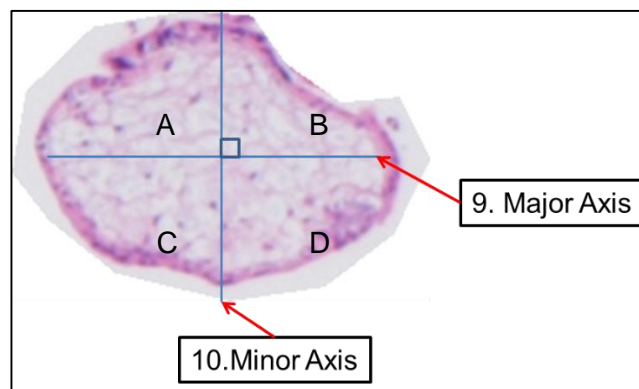
Figure 4.7. Major and minor axes.

3) Stroma's morphological and textural characteristics:

a) The percentage of stroma regions inside the villi is measured by calculating the proportion of segmented stroma areas located inside villi boundaries and villi regions.

b) The percentage of edge pixels inside stroma regions is calculated by computing the percentage of edge pixels detected by the canny edge detector located inside stroma boundaries.

c) The variance of greyscales of stroma regions is applied to measure the texture of stroma regions. This feature is calculated by the variance of greyscale of segmented stroma areas.

4) Trophoblast thickness and proliferation:

a) The percentage of trophoblast inside the villi is calculated by computing the proportion of segmented trophoblast regions located inside villi boundaries and villi regions.

b) The trophoblast skeleton per trophoblast perimeter ratio is computed by the ratio between all skeletonized pixels of trophoblast areas and all pixels belonging to trophoblast perimeter.

c) Trophoblast analysis is based on the experts' knowledge and strategies that indicate that the shape of PHM and CHM trophoblast is irregular. Therefore, this feature is computed by the different trophoblast areas within the four quadrants defined by the centroid of villi.

Based on experts' heuristics, this study came across a few major difficulties. The first issue relates to the RBC's density and distribution within a given villus, which is a distinguishing feature that can be used to separate normal placental villi from CHM and PHM. However, the RBC feature depends on the time period taken between staining and examination, as well as on gestation age. If pathologists leave a tested stained slide for a certain period of time before examining it under a microscope, the RBC inside the villi die and the dead red blood cells are difficult to identify, whereas if the examination follows staining immediately, the live red blood cells can be observed by microscope. The second issue relates to the pregnancy period. The red blood cells inside villi are fewer in the early pregnancy stage but they increase significantly in later stages. However, HM slides of later pregnancy stages are unavailable because of pregnancy termination. The third issue involves the shape and size of villi, which can be a significant discriminatory feature. The size of normal placental villi is usually smaller than CHM and PHM villi. Additionally, the number of villi boundary corner points of PHM and CHM villi is usually higher than in normal placental villi, because the shape of normal villi is rounded, whereas the shape of PHM and CHM villi boundary is irregular, consisting of corners and curves (Figure 4.6). These size and shape features can be used to distinguish between the three categories of

villi, but abnormal cases in each category make the analysis more difficult. For example, if the size of an abnormal case of a normal placental villus is larger than the average size of normal placental villi, the villi might be classified as CHM or PHM due to their irregular size, although they are a normal placental villus. Therefore, this last descriptive feature is necessary but not sufficient on its own. The last issue relates to the fact that some villi in some stained slide images are incomplete and it is difficult to extrapolate their boundaries.

Table 4.1. Elicited morphological features and associated anomalies.

Morphological features given by expert pathologists	The proposed features	Types of anomalies
RBC factor: density and distribution within a given villi	Percentage of red blood cells inside stroma regions	Point anomaly
Villi: shape and size	Villi size	Morphological anomaly
	Number of villi boundary corner points	Morphological anomaly
	Different area between villi's bounding box and villi area	Morphological anomaly
	Major axis	Morphological anomaly
	Minor axis	Morphological anomaly
	Elongation ratio	Morphological anomaly
	Ratio between number of villi boundary corner points and all pixels belonging to villi perimeter	Morphological anomaly
	The notion of four quadrants	Morphological anomaly
Stroma's morphological characteristics	Percentage of stroma regions inside villi	Contextual anomaly
Stroma's textural characteristics	Percentage of edge inside stroma regions	Contextual anomaly
	Variance of grey scale of stroma regions	Point anomaly
Trophoblast thickness and proliferation	Percentage of trophoblast inside villi	Contextual anomaly
	Trophoblast skeleton per trophoblast perimeter ratio	Density anomaly
	Trophoblast analysis	Morphological anomaly

4.8. A heuristic approach to anomaly detection in image segmentation

An important process of an image analysis system is image segmentation, and this process can impact on later image processing steps such as feature extraction and classification. However, current image segmentation algorithms still have limitations for complex structures of various shapes and fuzzy boundaries. Thus, to address these problems the segmentation approach in this study is guided by the ontological representation of villi and their associated anomalies.

Based on the study by Niwas *et al.* (2010a) and Doyle *et al.* (2012), the segmentation algorithm applies fuzzy c-means clustering on HSV colour space to extract the four regions of interest (ROI): villi, red blood cells, trophoblast and stroma regions. The HSV or HSI colour space is used to describe the colours in terms of human interpretation, and is suitable for image processing techniques (Gonzalez & Woods, 2002). HSV colour space contains three colour channels: hue, saturation and value. The hue is a circular colour model containing the range from 0 to 360 degree as shown in Figure 4.8. The saturation value describes the density of the colour and the range is from 0 to 1, whereas the value channel measures the brightness of the colour and the range is from 0 to 1. *H*, *S* and *V* are mathematically defined below (Su *et al.,* 2011), where *R*, *G* and *B* relate to the intensity of red, green and blue channels of RGB colour space.

$$H = \cos^{-1} \frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\sqrt{(R-G)^2 + (R+B)(G-B)}} \tag{1}$$

$$S = 1 - \frac{3}{R+G+B}(\min(R,G,B)) \tag{2}$$

$$V = \frac{1}{3}(R+G+B) \tag{3}$$

To deal with the limitations of Euclidean distance measurement different techniques are applied. For the hue channels, circular distance measurement techniques (Lu *et al.*, 2010) based on cosine and sine functions are applied to measure the circular value in a hue channel (0 – 360 degree) as shown in Figure 4.9 (Lu *et al.*, 2010).

For a grey pixel (*S<0.1* or *V<0.15*), only the value channel is considered and the distance measurement is only computed by the Euclidean distance:

67

$$d(x_k, v_i) = |V_k - V_i| \tag{4}$$

where V_i and V_k are the intensity of the value channel in pixels i and k.

For a colour pixel, the circular distance measurement is applied to transform hue, saturation and value to C_1, C_2 and C_3. These parameters are calculated by the following equations:

$$\begin{aligned} C_1 &= S \times \cos(H) \\ C_2 &= S \times \sin(H) \\ C_3 &= Value \end{aligned} \tag{5}$$

Then, the new parameters are used as features of segmentation or classification methods.
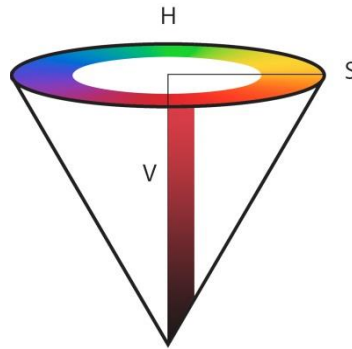
Figure 4.8. HSV colour space (Su *et al.,* 2011).
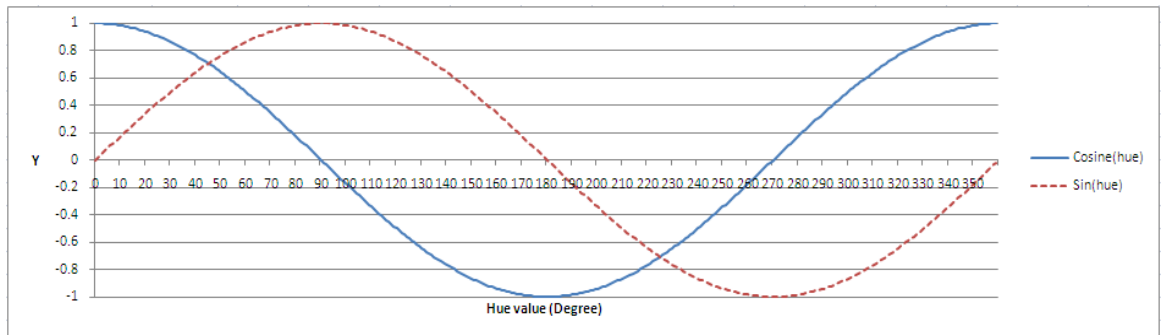
Figure 4.9. The cosine and sine functions of hue value in HSV colour space.

The Fuzzy c-Means clustering (FCM) technique is a popular tool to categorise data points. The k-means clustering technique is based on a discrete membership value, whereas FCM technique uses a variable membership value instead of using a discrete value (Dhawan, 2003). The technique aims at minimising the objective function and updates a membership value. The objective function $J_m(U, v)$ is defined as below:

$$J_m(U,v) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} d_{ij}^{2} = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} \left\| x_j - v_i \right\|^2 \qquad (6)$$

$$0 \le u_{ij} \le 1 \text{ for all } i, j$$

$$\sum_{i=1}^{c} u_{ij} = 1 \text{ for all } j \text{ and } 0 < \sum_{j=1}^{n} u_{ij} < n \text{ for all } i \qquad (7)$$

Note that n is the number of data points, c is the number of clusters, v_i is the centre vector for cluster i, u_{ij}^{m} is the degree of membership for the jth data point x_j in cluster i, m is weighting exponent and d_{ij}^{2} is squared distance. The computing and updating steps of fuzzy c-means clustering are the same as the k-means clustering steps, but the fuzzy c-means clustering technique uses equation (6) as the objective function and (7) as the updating function.

- **RBC segmentation**

The circular distance measurement is used to deal with the limitations of the distance measurement on HSV colour space. The FCM clustering technique is used to classify red blood cell regions from the background and other cell components. If some villi contain undesirable artefacts such as Hofbauer cells, purple rounded cells, (Benirschke *et al.*, 2006) in the segmented red blood cells, FCM is applied iteratively using two different redness measurements (*H* and *S*) to remove these artefacts from those RBCs. If the average *H* value of the segmented cells is less than 0.1 or the average S value is equal or greater than 0.2, then the segmented cells belong to RBC, otherwise they are removed. This is also used to solve the over-segmentation problems by removing the red blood cell regions located outside stroma or cell's boundary regions, as shown in Figure 4.10.

- **Trophoblast and stroma segmentation**

FCM is also applied to segment trophoblast regions from stroma regions and their background, based on saturate channel of HSV colour space. The ROI can be classified into three groups: background, trophoblast and stroma regions. Because the saturation channel is sensitive to a bright pixel, the brighter regions classified by FCM are labelled as stroma regions, whereas the dark regions are classified as trophoblast regions. Any other purple regions, which overlap with stroma regions, are removed. Because the Euclidian distance is limited to measuring the range of the red colour in hue channel between 0 and 20 degrees, the circular distance measurement based on sine and cosine functions is

69

selected. This algorithm improves the accuracy of trophoblast segmentation but cannot segment the red blood cell regions – which are outliers- located near trophoblast regions, as shown in Figure 4.11.

- **Villus segmentation**

To segment the placenta's villus boundary regions, histogram equalisation is applied first, to improve the quality of the villus image by adjusting image intensities and enhancing contrast; this helps separate villus boundary regions from their background. Then, FCM based on the circular algorithm is applied to segment the ROI (placenta villus boundary regions) from the background, as shown in Figure 4.12.

The stroma region located inside the villus boundary is correctly segmented by the proposed algorithm: the bright regions located inside the villous region are segmented by the algorithm as shown in Figure 4.13. This algorithm cannot select unconnected regions of stroma which are rare cases as shown in Figure 4.14. Furthermore it can only locate the largest region located inside the villus boundary and cannot deal with multiple stroma regions within a given villus. Over- and under-segmentation problems are discussed in Section 4.9.
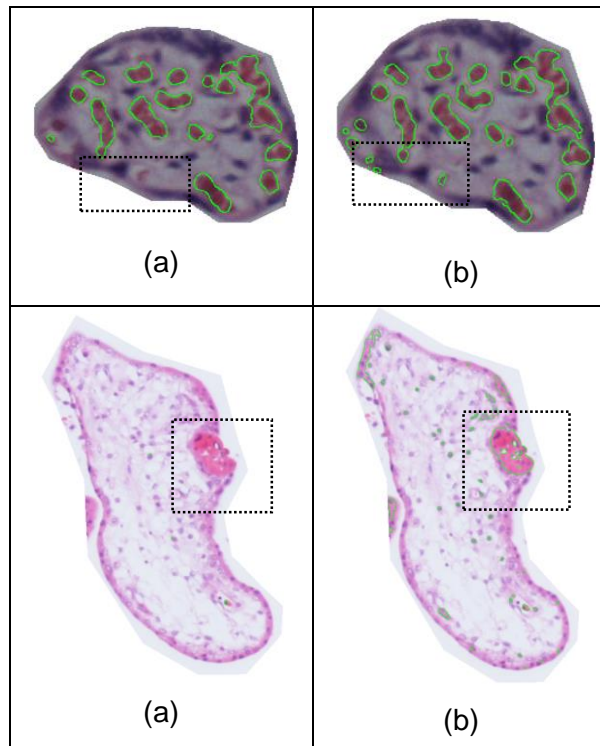

Figure 4.10. (a) Segmentation based on Euclidean distance measurement technique, (b) Segmentation based on the circular measurement technique.
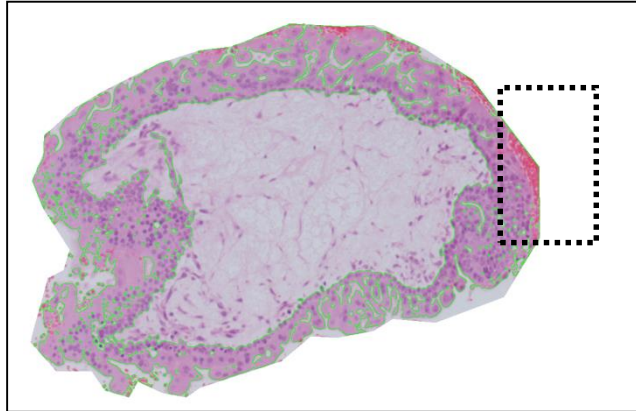
Figure 4.11. Outliers (i.e. red region) located between the villus' boundary and trophoblast region.
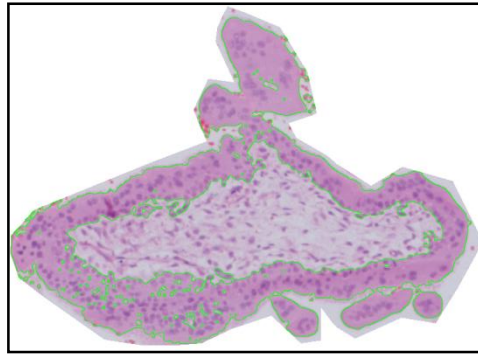
Figure 4.12. Segmentation of the trophoblast and villus boundary regions.

Figure 4.13. Stroma regions segmented by the proposed algorithm.
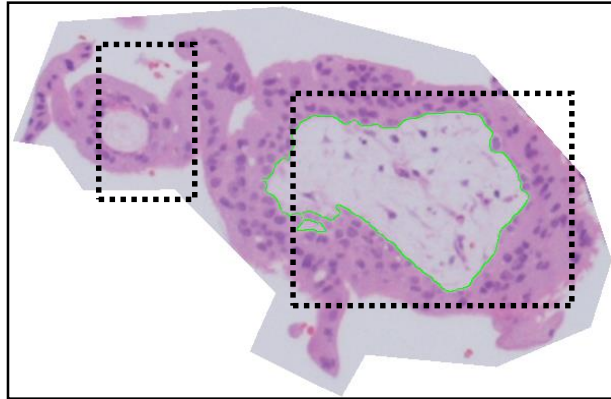
Figure 4.14. Example of multiple stroma regions.

4.9. Discussion of segmentation results

The proposed FCM on HSV colour space approach has been applied to segment 986 villi images, 304 normal placental villi, 467 PHM villi and 215 CHM villi. The segmentation results suffer from over- and under-segmentation problems, as shown in Table 4.2. Over-segmentation produces a region larger than the actual ROI (regions of interest), whereas under-segmentation ignores critical regions. Before subjecting the villi to the segmentation algorithm each villus is manually extracted from the slide. The algorithm correctly segmented 773 villi out of 986 (78.4%); the best segmentation was achieved with PHM villi (83.5%) followed by normal villi (82.2%) and the lowest score referred to CHM villi (62%); 185 villi were incorrectly segmented because of over-segmentation problems and 26 villi were under-segmented. Over- and under-segmentation problems are caused by artefacts such as incomplete villi, villous tissues and red blood cells in intervillous space, as shown in Figure 4.15a, and Figure 4.15b, and Appendix C.

To address the problem of over-segmentation, a careful manual extraction of each villus was carried out to help reducing the artefacts of the selected images and improving the segmentation results (Figure 4.15c and Figure 4.15e). This involved the user to draw a line closer to the villus' boundaries and avoid undesirable artefacts attached to the villus (Figure 4.15d). The number of correctly segmented villi increased from 78.4% to 94.5%, and the correctly segmented CHM villi increased from 61.9% to 85.6%. Though over segmented normal placental villi were resolved, 8 villi were still under-segmented (Table 4.3). In the future, additional textural features may be further elicited and added to the algorithm to address the under-segmentation problems.

The segmentation method has a few limitations. It assumes that a villus contains one stroma region, and therefore cannot cope with multiple-stroma regions. Additionally, the boundaries of some villi are missing and the boundaries between the trophoblast and red blood cells are fuzzy. Finally, the outliers that are located outside the villus boundaries remain unresolved.
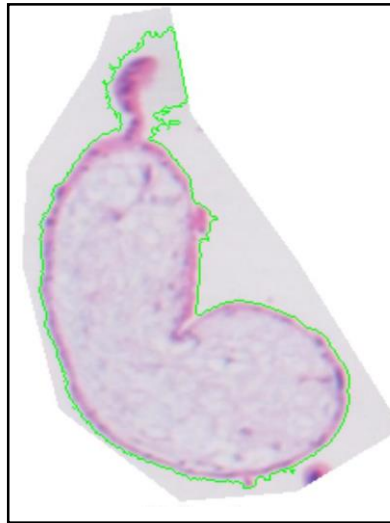
N05-19

Figure 4.15. (a) Over-segmentation case
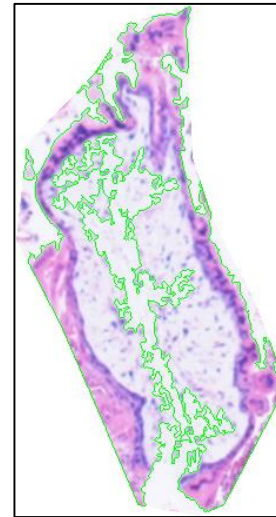
CHM19-3

Figure 4.15. (b) Under-segmentation case.
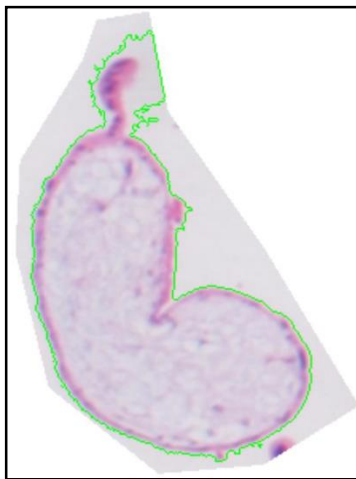
N05-19

Figure 4.15. (c) Over-segmentation problem
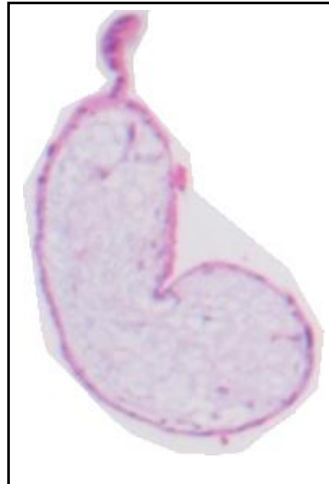
N05-19

Figure 4.15. (d) The manual extraction of villi.
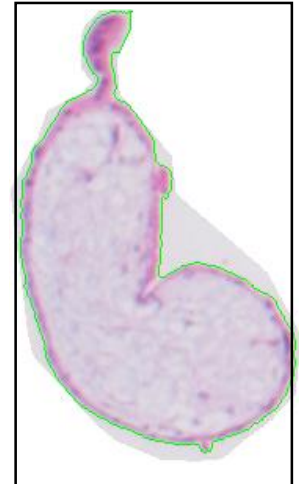
N05-19

Figure 4.15. (e) Improved segmentation result.

Table 4.2. Segmentation results based on FCM and HSV colour space.

Type of Images	Total number of villi samples	Number of correct segmented villi	Number of incorrect segmented villi	Percentage of correct segmented villi	Over-segmented villi	Under-segmented villi
Normal	304	250	54	82.2%	46	6
PHM	467	390	77	83.5%	70	7
CHM	215	133	82	61.9%	69	13
Total	986	773	213		185	26

Table 4.3. Improved segmentation results based on FCM and HSV colour space

Type of Images	Total number of samples	Number of correct segmented villi	Number of incorrect segmented villi	Percentage of correct segmented villi	Over-segmented villi	Under-segmented villi
Normal	304	296	8	97.4%	0	8
PHM	467	452	15	96.8%	8	7
CHM	215	184	31	85.6%	18	13
Total	986	932	54		26	28

4.10. Feature extraction step

As the data consists of villi whose boundaries are ill-demarcated and contain interruptions, this step and the subsequent steps are carried out solely complete and intact placental villi. In this study Nxx denotes the normal placental images, consisting of 294 normal placental villi, whereas the PHM (455 complete villi) and CHM images (190 complete villi) are denoted by PHMxx and CHMxx respectively, as shown in Table 4.4.

This study aims at extracting the remarkable morphological features and categorise them into four regions (RBC, villi, stroma and trophoblast). Then, the anomalies of PHM and CHM are analysed and are compared with the features of normal placental villi.

The feature extraction, also guided by the ontological representation of HM, aims to extract the anomalies associated with the four villi components listed in Table 4.1 and illustrated in Figure 4.16. This step applies a set of statistical measures on the villi pixels to analyse morphological and texture characteristics of the villi components for normal, PHM and CHM slides, described below.

- **Villi**

A number of features are analysed:

First, the villi size (in terms of pixels) and the number of villi boundary corner points are used as the morphological features of villi. With the exception of slide images N01-22, N05-21 and N07-20, the size of normal villi is relatively constant (Figure 4.17a), whereas the size of PHM villi shows some variety but is still confined to a small boundary range (Figure 4.17b) and the size of CHM villi identifies a number of high peaks in Figure 4.17c. The graph in Figure 4.18 reveals a clear boundary between CHM villi and normal and PHM villi; this can be a useful distinctive feature to separate CHM villi from the other two types. Further graphs are included in Appendix D.

Table 4.4. Hydatidiform Moles dataset.

Image No.	Total No. of villi	No. of intact villi	Image No.	Total No. of villi	No. of intact villi	Image No.	Total No. of villi	No. of intact villi
N01	25	22	PHM01	18	18	CHM01	13	13
N02	28	26	PHM02	28	27	CHM02	7	7
N03	34	34	PHM03	21	20	CHM03	21	18
N04	32	32	PHM04	17	17	CHM04	18	15
N05	21	21	PHM05	23	22	CHM05	9	7
N06	38	37	PHM06	13	13	CHM06	9	9
N07	20	20	PHM07	24	24	CHM07	9	9
N08	40	40	PHM08	24	24	CHM08	15	15
N09	37	37	PHM09	27	27	CHM09	22	22
N10	29	25	PHM10	17	16	CHM10	11	11
Total	304	294	PHM11	26	25	CHM11	4	3
			PHM12	32	29	CHM12	8	7
			PHM13	22	22	CHM13	7	6
			PHM14	16	16	CHM14	2	2
			PHM15	18	18	CHM15	7	5
			PHM16	34	34	CHM16	5	4
			PHM17	14	14	CHM17	13	11
			PHM18	20	20	CHM18	11	6
			PHM19	25	23	CHM19	12	9
			PHM20	20	20	CHM20	5	4
			PHM21	28	26	CHM21	7	7
			Total	467	455	Total	215	190
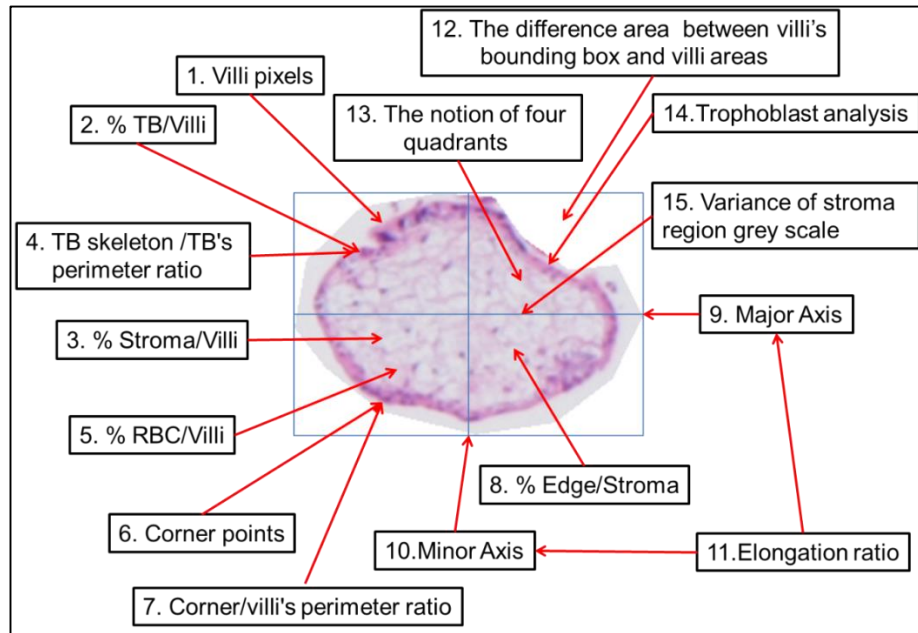
Figure 4.16. Villi's features.
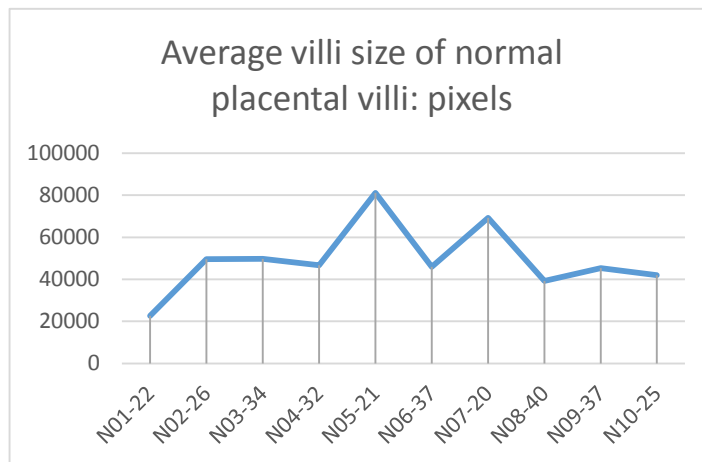
Figure 4.17. (a) Average villi size of normal placental villi.

(e.g. N01 denotes the slide number and 22 the number of villi in N01)
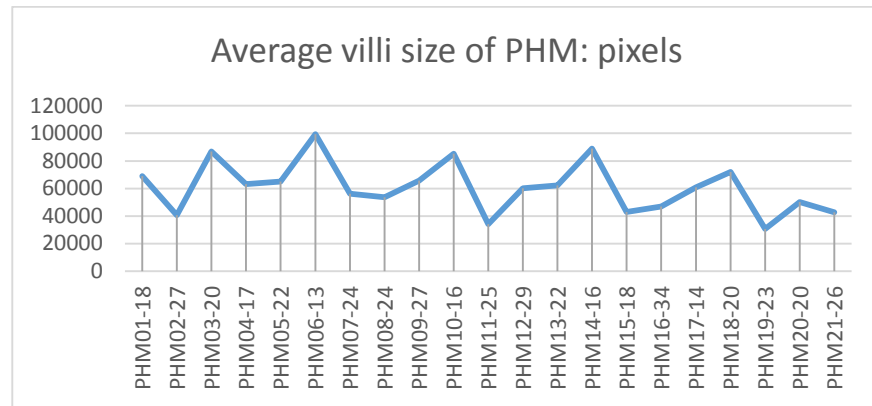
Figure 4.17. (b) Average villi size of PHM.
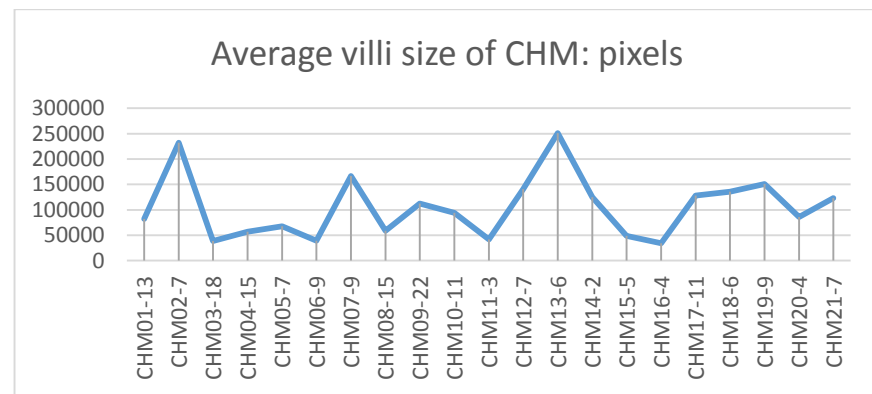
Figure 4.17. (c) Average villi size of CHM.

Second, the villi boundary corner points are used to measure the shape of villi. Normal villi have a rounded shape, whereas PHM and CHM villi are irregular and have a large number of villi boundary corners (Figure 4.19). The ratio between number of villi boundary corner points and all pixels belonging to villi perimeter of PHM and CHM is higher than for normal villi, outlining their irregular shape (Figure 4.20).

Third, the major axis, minor axis and elongation ratio of villi are used to measure the degree of shape irregularity. The values of the major axis, minor axis and elongation ratio of PHM and CHM villi tend to be higher than for normal placental villi, as shown in Figure 4.21 and Figure 4.22. However, all of these three features are not stable. Therefore, a further fourth anomalous characteristic measure is required to deal with uncommonly shaped villi in normal placental, PHM and CHM slides. Consequently, the different area between the villi's bounding box and villi area, and the notion of four quadrants are analysed to check if, compared to normal placentas, their values, in PHM and CHM samples are higher, due to their irregular and uneven shapes, or smaller, due to their rounded and smoothed shapes, as shown in Figures 4.23 and 4.24.
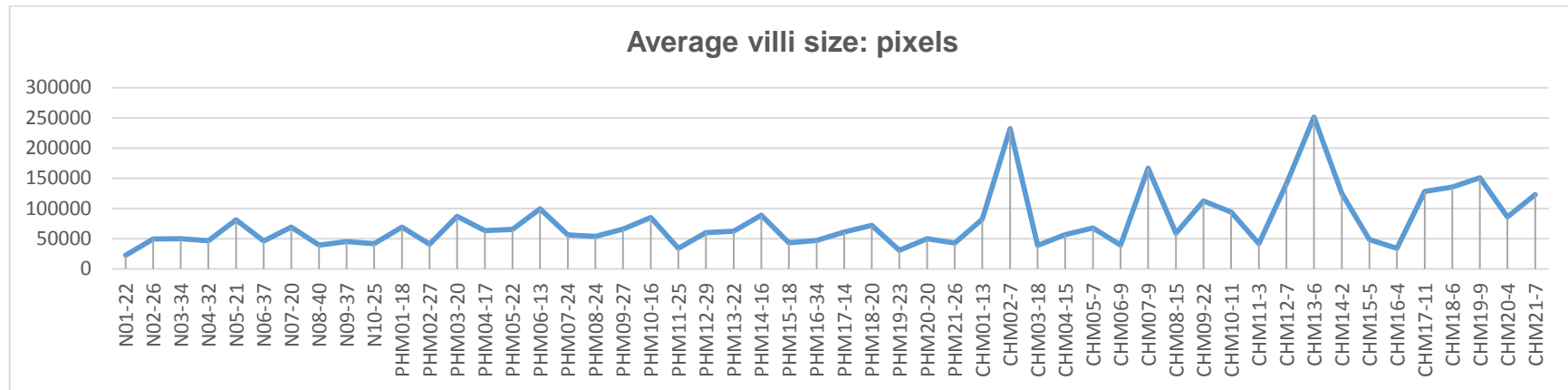
Figure 4.18. Average villi size.
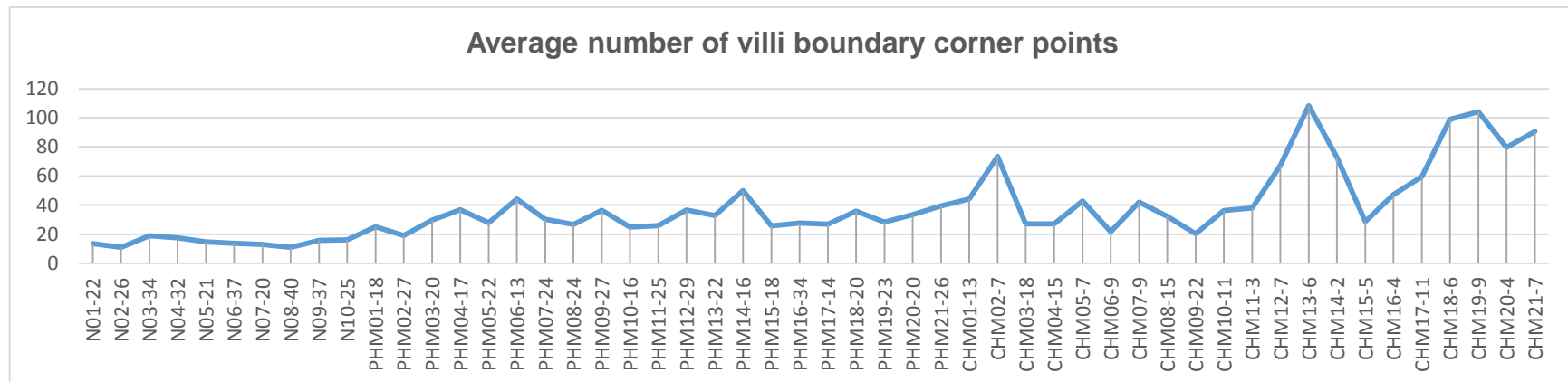
Figure 4.19. Average number of villi boundary corner points
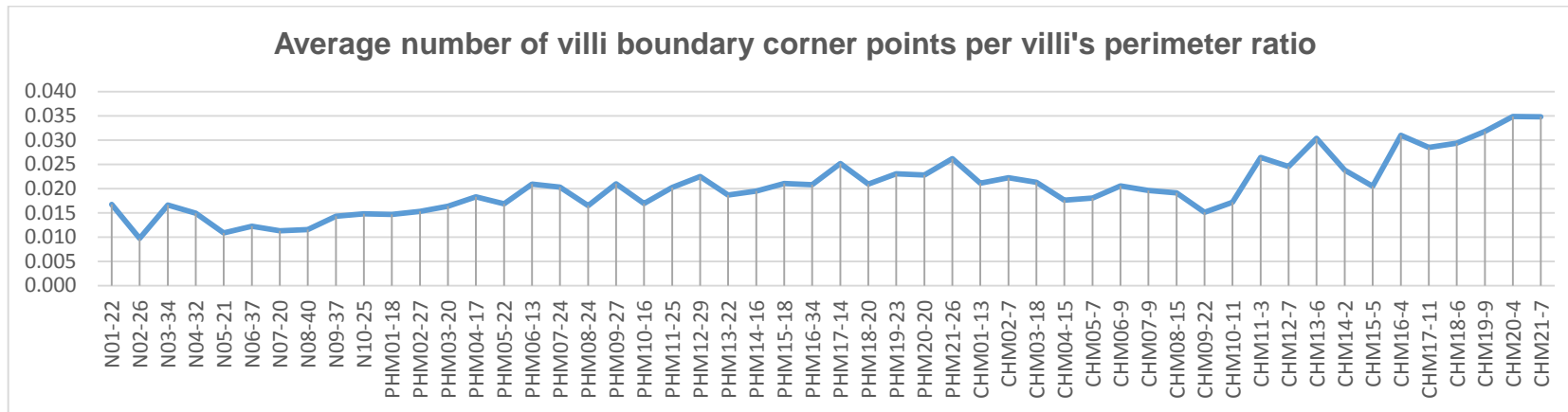
Figure 4.20. Ratio between number of villi boundary corner points and all pixels belonging to villi perimeter.
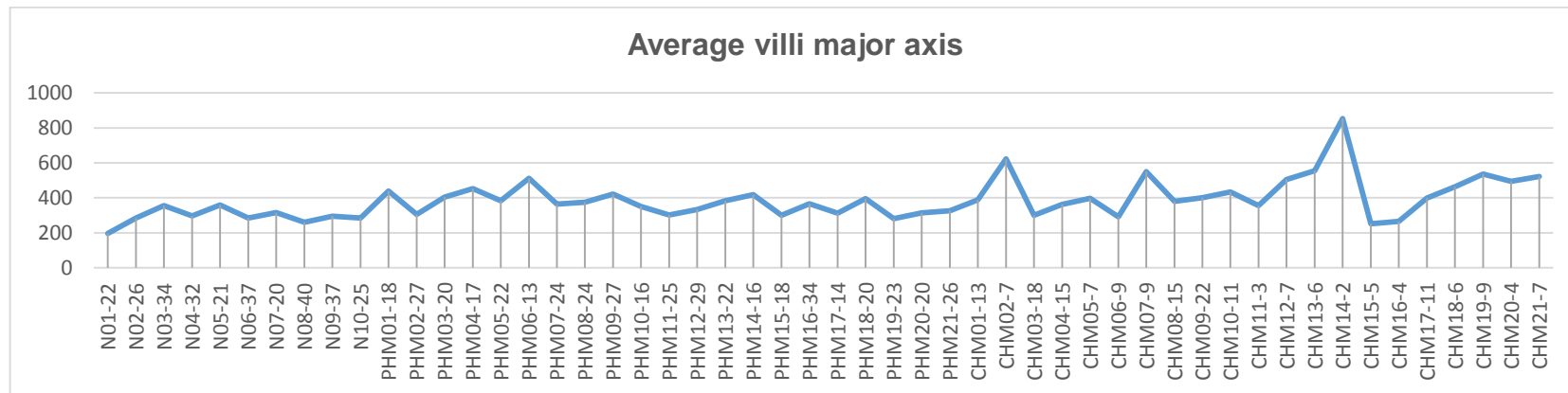
Figure 4.21. (a) Major axis.
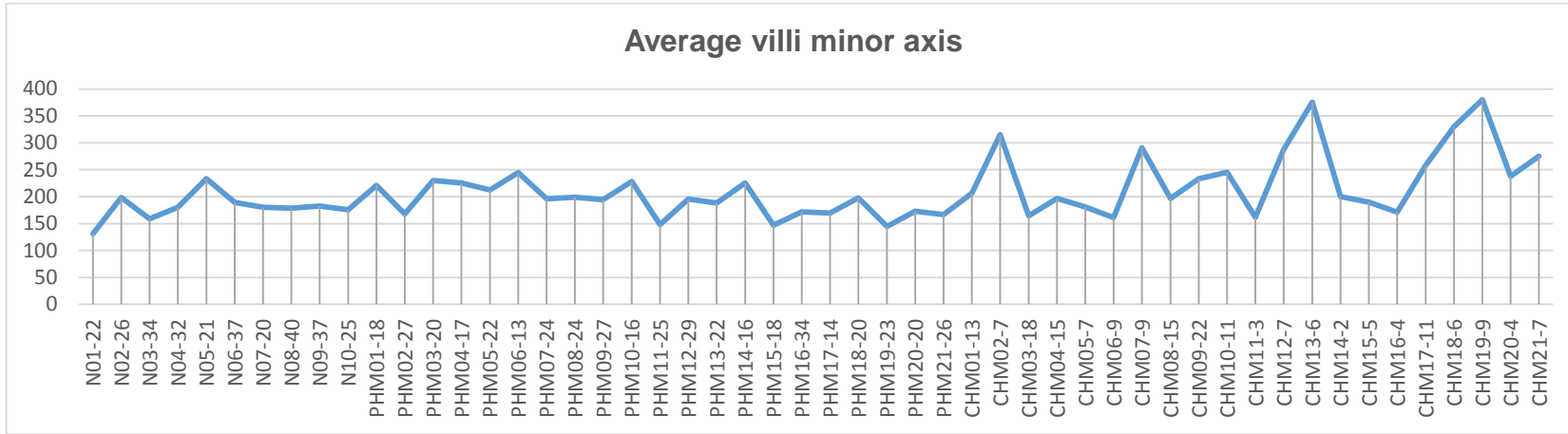
Figure 4.21. (b) Minor axis.
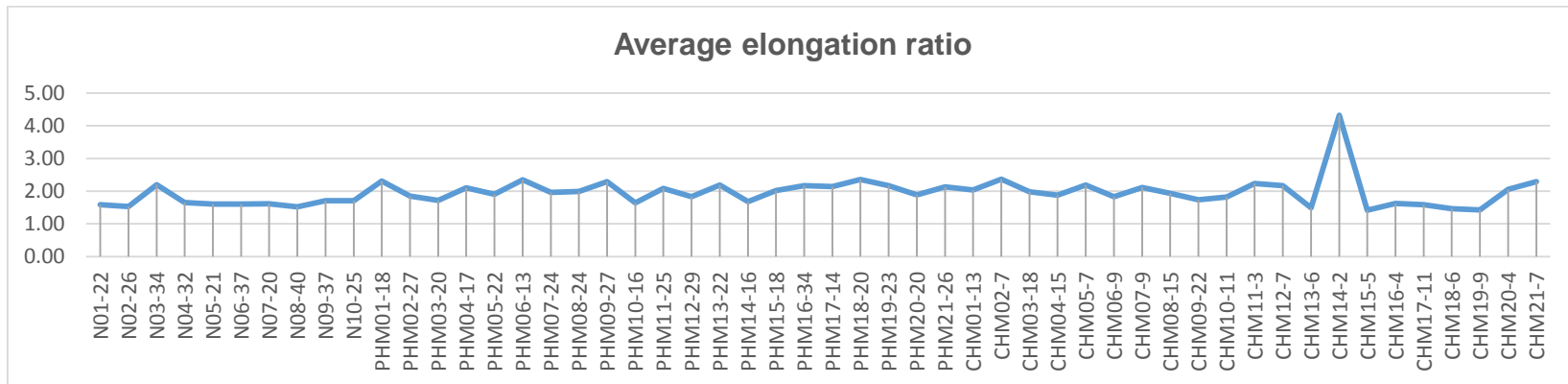
Figure 4.22. Elongation ratio.
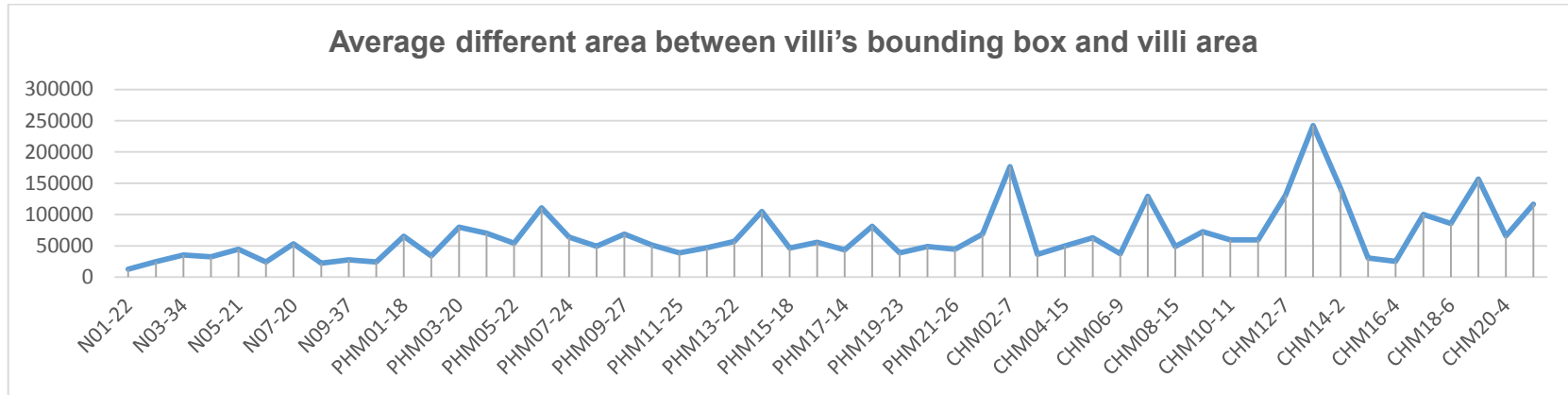
Figure 4.23. Different area between villi's bounding box and villi area.
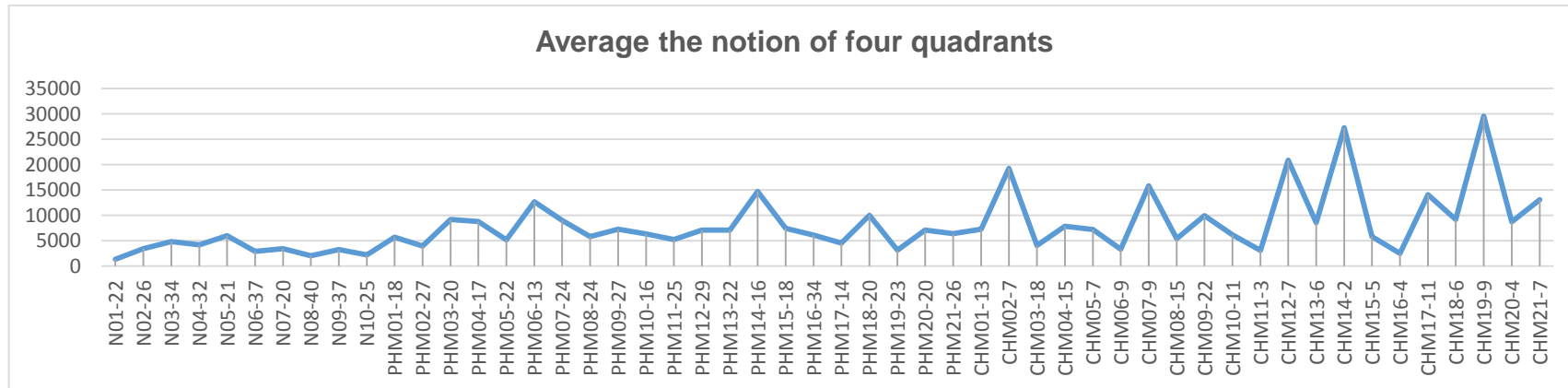
Figure 4.24. The notion of four quadrants.

- **Red blood cells**

These cells, which are located inside the stroma region, are an important feature used to distinguish normal placental villi from PHM and CHM villi. The density of RBCs indicates the percentage of segmented red blood cells located inside stroma boundaries in terms of pixels. The analysis is based on the assumption that normal villi contain a higher proportion of RBCs than CHM and PHM, so the density of RBC values in the three villi types (Figure 4.25) needs to be considered in the classification step.

- **Stroma's morphological characteristics**

These are important contextual anomalies expressed in terms of three measures. The first measure focuses on the percentage of stroma regions inside villi and it is high in normal placental and PHM and low in CHM villi, because of their trophoblast proliferation. If the percentage of stroma regions inside the villi in a given slide is low (e.g. N03 and PHM21) or high (e.g. CHM08, CHM09, CHM10 and CHM18) these villi can be labelled as anomalous (Figure 4.29).

The second measure, which computes the percentage of edge pixels inside stroma regions, is used to capture textural characteristics of stroma surfaces. High values of this measure are commonly found in normal placental and CHM stroma regions, and low values in PHM stroma regions. Therefore, the high values of the percentage of edge inside stroma regions of PHM samples can be considered as anomalous (Figure 4.26).

The third measure analyses the variance of greyscale of stroma regions. This measure is referred to as a point anomaly. The variance of greyscales of stroma regions is commonly high in CHM samples and low in PHM and normal placental villi samples (Figure 4.27). This is because the stroma structure of normal placental villi is reticular whereas it is scalloped dentate and mild in PHM and enlarged, irregular and karyorrhectic in CHM.

- **Trophoblast thickness and proliferation**

The percentage of trophoblast inside villi, and trophoblast skeleton per trophoblast perimeter ratio are examples of contextual and density anomalies; they are used to measure trophoblast thickness and proliferation. The percentage of trophoblast inside normal and PHM villi should be low, but the villi in N03, N10 and PHM04 slides reveal some anomalies because of their high percentage of trophoblast inside their villi. However, the percentage of trophoblast inside CHM villi are expected to be high, but the

villi in CHM10 and CHM19 contain a low percentage of trophoblast inside their villi (Figure 4.28).

The density anomaly is captured by the trophoblast skeleton per trophoblast perimeter ratio, which is low in normal and PHM villi but high in CHM villi. However, Figure 4.29 shows exceptional cases, such as the slides N03, N10, PHM04, PHM09, CHM08, CHM09 and CHM10.

The analysis of trophoblast is an example of a morphological anomaly and it is particularly relevant to PHM and CHM trophoblasts, which show strong irregularities in CHM, some irregularities in PHM and mild changes in normal placental villi (Figure 4.30).

4.11. Summary

This chapter focused on describing the novel heuristic approach to detecting anomalies in villi, based on the two low level image processing steps: segmentation and feature extraction. The tacit knowledge elicited from experts in detecting anomalies is ontologically represented and integrated into the segmentation and feature extraction algorithms. The knowledge and strategies evolved during this study, as the findings were not conclusive and allowed the experts to identify additional hidden strategies. The segmentation method based on FCM applied to HSV colour space has achieved high accuracy in normal placental (82.2%) and PHM villi (83.5%) but low accuracy for CHM villi (61.9%), caused by over- and under-segmentation problems. This was partially caused by the fact that the extraction of villi was carried out with a large margin (see Figure 4.15c). By reducing that margin (see Figure 4.15e) the segmentation was improved and the total number of correctly segmented villi increased from 78.4% to 94.5% and from 61.9% to 85.6% in CHM.

The low level experimental study has indicated that no single feature has enough discriminating power to deal with fuzzy and complex villi shapes and sizes. It has also revealed inconsistencies of villi characteristics within each component and within each type of villi. There are some basic common properties but also a set of exceptions that emphasises the complexity of the task on hand. Further meetings with the experts have led to the elicitation of additional tacit features which have been added and integrated into the next high level of processing (i.e. the classification step); these are discussed in the next chapter.
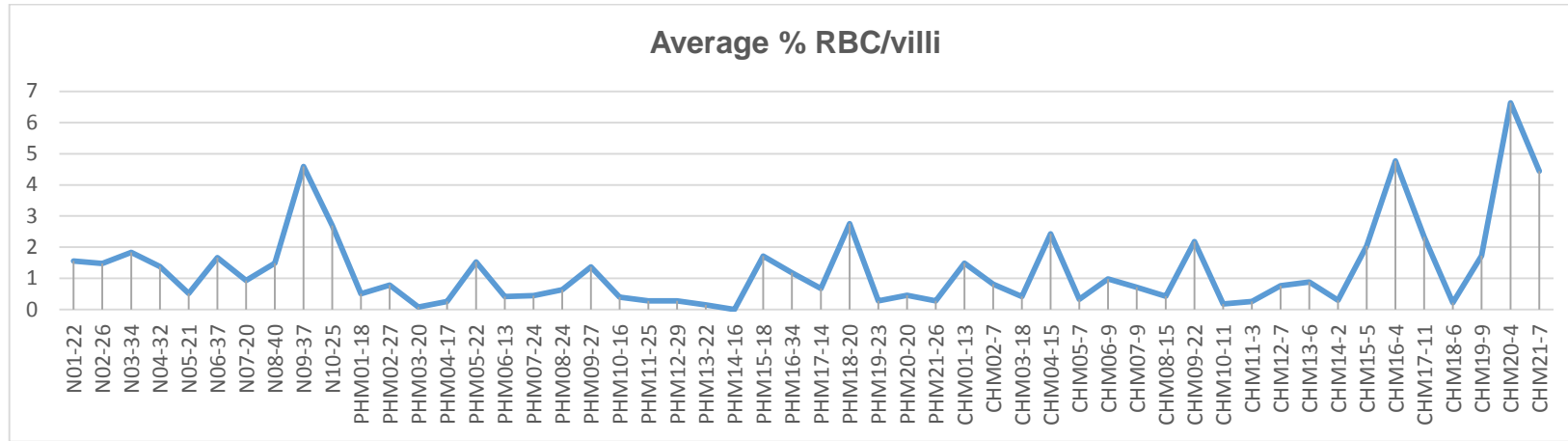
Figure 4.25. Density of RBC per villi.
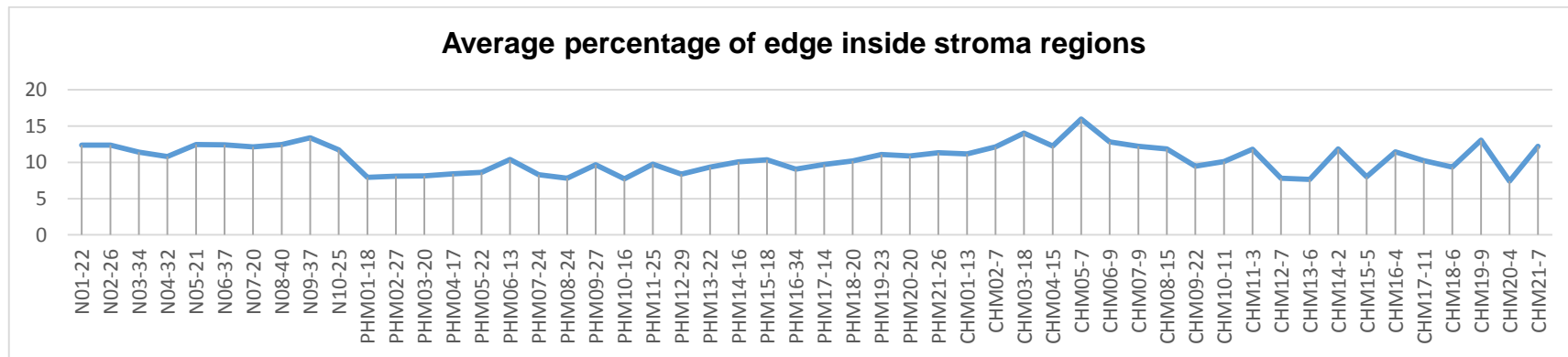
Figure 4.26. Percentage of edge inside stroma regions.
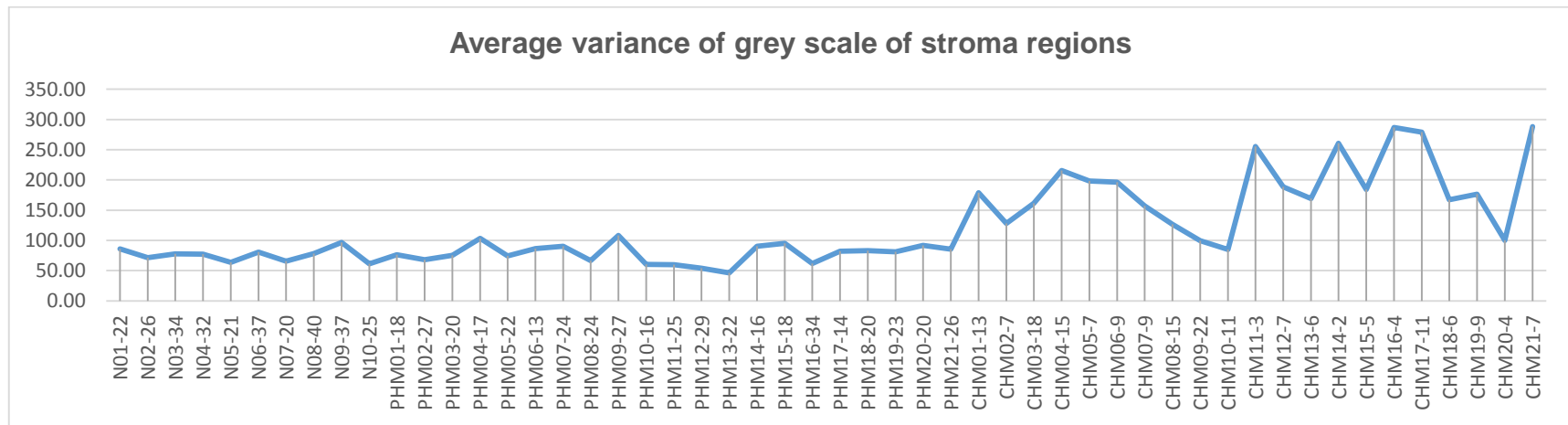
Figure 4.27. Variance of grey scale of stroma regions.
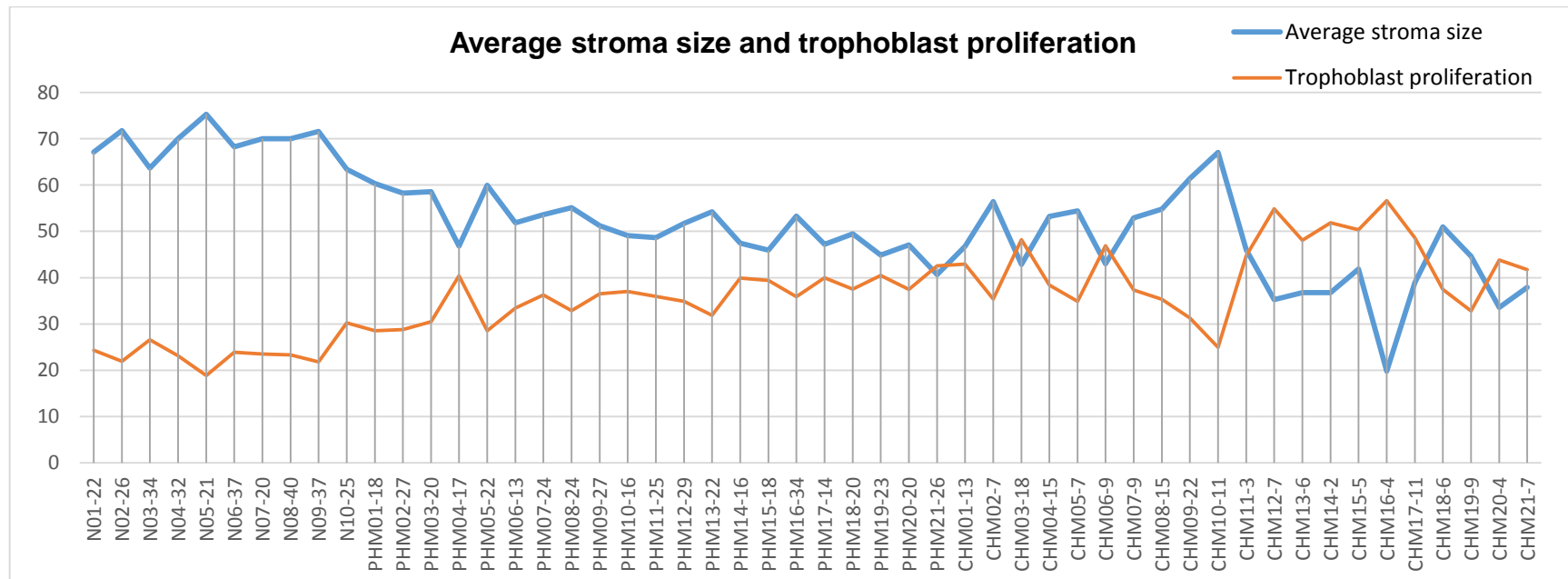
Figure 4.28. Average stroma size and trophoblast proliferation
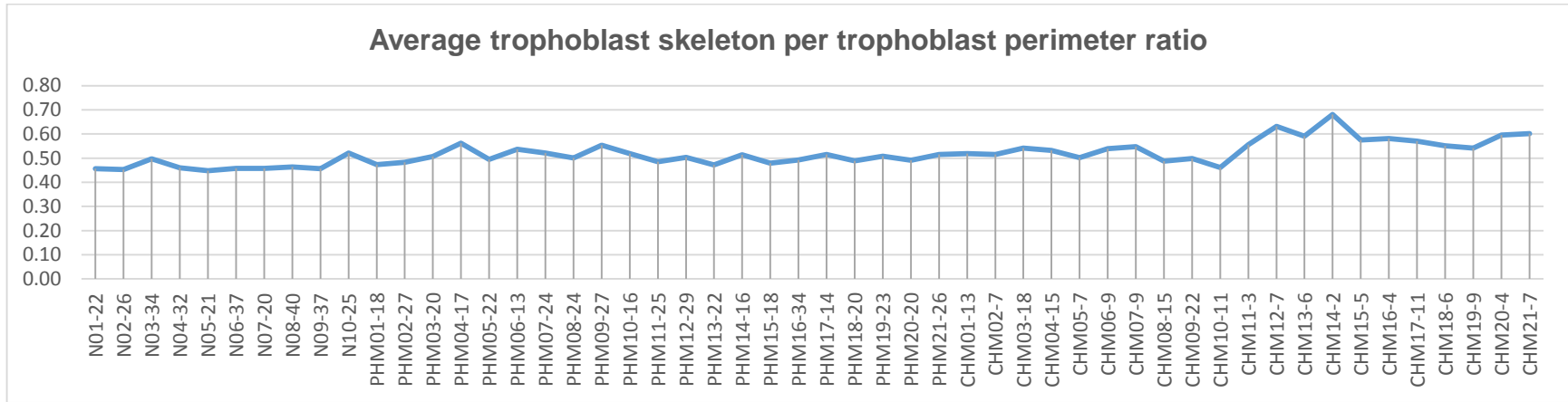
Figure 4.29. Average trophoblast skeleton per trophoblast perimeter ratio.
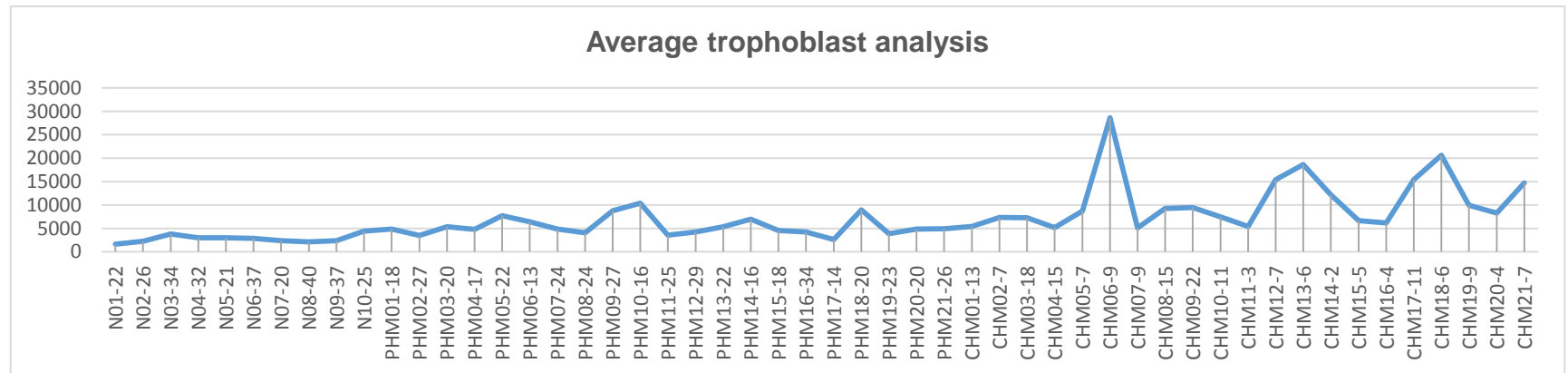
Figure 4.30. Trophoblast analysis.

Chapter 5: A Heuristic Neural Network Approach to Anomalies Detection of Hydatidiform Mole Villi

5.1. Introduction

This chapter focuses on the second phase of the novel approach, which relates to the high level processing of the villi stained sides. The aim of this chapter is twofold: to identify and rank the most critical features for investigating their discriminative power, and to classify the anomalies of the villi using neural networks that are designed to simulate experts' cognitive processes.

Whilst the low level processing step (i.e. segmentation step) is based on 986 villi and has included complete and incomplete villi, the high level processing phase is based on intact and complete villi and includes 939 villi (i.e. 294 normal, 455 PHM and 190 CHM villi) selected from a total of 52 stained slide images (10 normal, 21 PHM and 21 CHM). The first part of this chapter describes the four sets of experiments carried out to select the most discriminative features, using principal component analysis. The ranking of these features is determined using five criteria, described in Section 5.3.2. These features are then used to train a multi-neural network, consisting of three sub-networks, to classify the villi into their appropriate anomalous types. The second part describes the tool that was developed to support these low level and high level analyses of the villi. The findings are summarised in the conclusion.

5.2. The novel classification approach

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. An ANN is a computational machine–learning approach based on a biological neuron system of a human brain. It is composed of a set of interconnected neurones designed to solve specific problems. An artificial neuron consists of input and activation functions. The input function is the summation of input signals multiplied by weights, and the activation function is the function that is applied to determine the neuron output. If the value of the input function is greater than or equal to the activation level of the activation function, then an output signal equal to the activation value is sent through the neuron output. Typical activation functions are step, sign, linear, and sigmoid functions (Negnevitsky, 2005). The

diagram of a neuron is shown in Figure 5.1. An ANN is configured through a learning process for a specific application, such as pattern recognition or data classification.
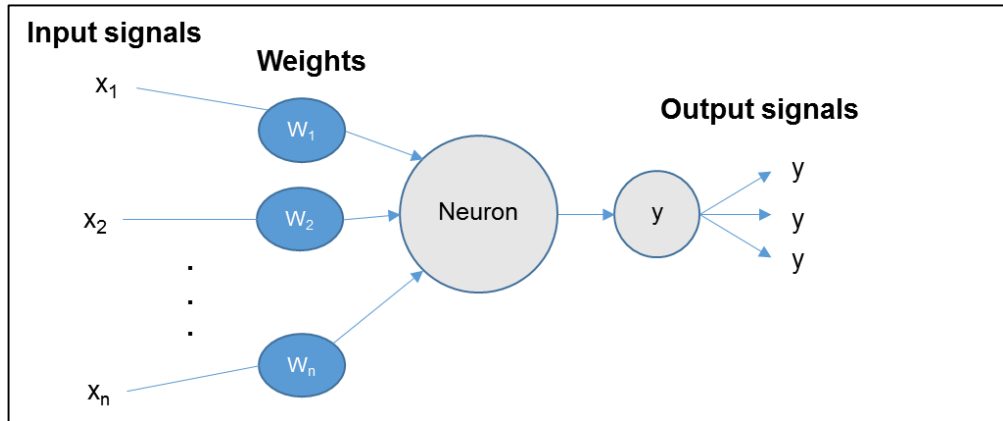
Figure 5.1. Diagram of a neuron (Negnevitsky, 2005)

This study focuses on an artificial neural network, namely a multilayer perceptron. A multilayer perceptron is a neural network containing an input layer, one or more hidden layers and an output layer. The input layer comprises input nodes connected to sources or to features, and the outputs of these input nodes are linked to the hidden layer containing hidden nodes. The outputs of the hidden nodes are used as inputs to output nodes in the output layer (Russell & Norvig, 2010), as shown in Figure 5.2. Generally, input nodes receive signals from input sources and transfer the received signals to all the hidden nodes in the hidden layer. Then, the hidden nodes in the hidden layer process the signals and apply the weights and bias. Next, the output layer determines the output signals (Negnevitsky, 2005). The activation function is the function that the hidden nodes use to decide their outputs. In this study, the sigmoid function is chosen as hidden nodes activation function, in order to achieve practical and continuous outputs.

The choice of the number of hidden layers and of the number of nodes for each layer is one of the most critical problems in constructing the ANN. Each added hidden layer could increase the computational cost (Negnevitsky, 2005). The Kolmogorov's theorem states that a single hidden layer is enough for a multilayer perceptron to solve any linear and non-linear problem and the number of hidden nodes is between the number of input nodes and output nodes (Blum, 1992). Therefore, the number of hidden layers and hidden nodes need to be as small as possible, to enable the network to generalise whilst at the same time encapsulate the requisite complexity (Karsoliya, 2012).
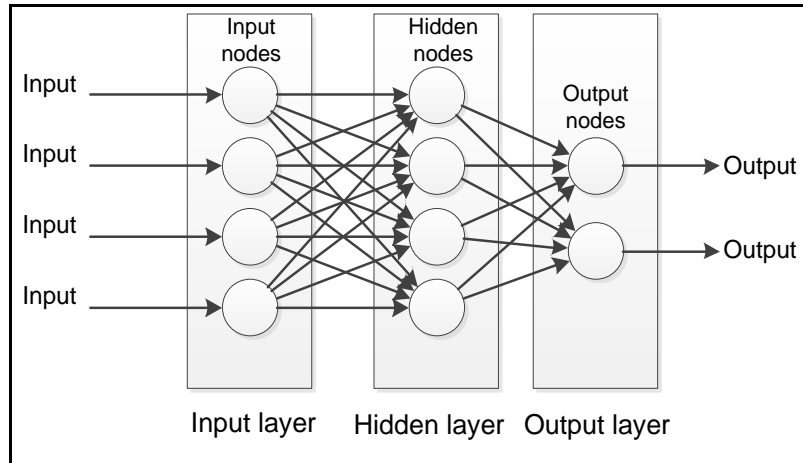
Figure 5.2. Three layers and nodes of MLP

This thesis has developed a novel neural network configuration approach that simulates the heuristic approach of the experts for detecting anomalies in villi. The following sections describe the development of this novel configuration approach based on the elicitation of experts' heuristics. Much of their medical expertise lies in laid-down experience, gathered over a number of years, and becomes so "routinised" that they find it difficult to articulate it, or it is taken for granted. Experts may not always agree in their approach and may differ among themselves. In developing the neural network configuration approach a task analysis approach is adopted to capture experts' heuristics, and each task is divided into subtasks that are further explored in terms of a set of experiments, in order to manage the complexity of the task.

The development of the novel approach was initially based on the set of 15 features, described in the previous chapter, to train the MLP and classify the villi. As the villi samples used in the segmentation step contained incomplete villi, it was agreed with the experts that the classification should be based only on complete and intact villi, a practice adopted by the experts themselves. Consequently, the classification is based on 939 villi out of the original 986, consisting of 294 normal, 455 PHM and 190 CHM intact and complete villi.

5.3. Experimental study 1

The first task adopted by the experts and by this experiment is to focus on the analysis of normal and CHM villi slides, to distinguish clear boundaries between these two types of villi. The 15 features extracted from these normal and CHM villi are fed into a feed forward and back propagation MLP. Based on Kolmogorov theorem, this MLP uses one hidden layer with hidden nodes ranging from one to 15, as shown in Figure 5.3.
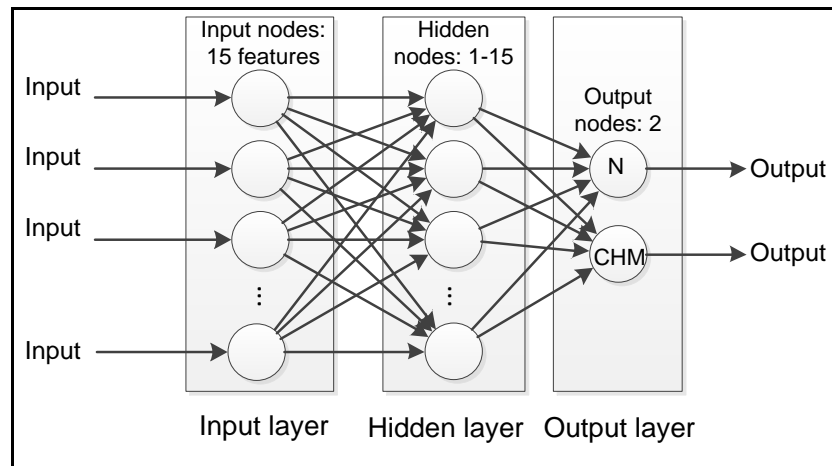
Figure 5.3. MLP for normal and CHM villi classification.

To validate the classifier, in this experiment a 10-fold cross-validation and a repeated 10-fold cross validation are conducted. Refaeilzadeh et al. (2009) explain that recent theoretical and experimental studies have shown that the k-fold cross-validation is the better informed estimation method out of the five popular strategies for computing the accuracy of a classifier. Table 5.1 summarises the limitations and benefits of each strategy.

The number k can be any number, but the most typical number is ten, which is selected in this experiment. Cross validation is carried out in terms of two major steps. First, the data is divided into ten portions. Second, ten iterations of training and validation are executed. In each iteration, one portion is sequentially taken out and used as a validation set, while the other nine portions are used as a training set, as illustrated in Figure 5.4. The samples in each portion need to be stratified to ensure that each portion represents the approximately equal samples of each class. The repeated 10-fold cross-validation method is the method that repeats the 10-fold cross-validation N times by shuffling sampling sets and then computes the average and standard deviation of their accuracies. In this experiment, the number N is set as five, to validate the accuracy and reliability of MLP. A confusion matrix is also applied to measure the results in terms of accuracy, recall, false positive rate, true negative rate, false negative rate and precision (Dhawan, 2003). It is also used to compare the predicted results against its actual outputs (Figure 5.5). Accuracy is the portion of correctly predicted samples out of the total number of analysis samples, whereas precision is the percentage of predicted samples that are relevant, and recall is the percentage of relevant samples that are predicted.

Table 5.1. Validation method comparisons (Refaeilzadeh *et al.*, 2009).

Validation method	Pros	Cons
Resubstitution Validation	Simple	Over-fitting
Hold-out Validation	Independent training and test	Reduced data for training and testing; Large variance
k-fold cross validation	Accurate performance estimation	Small samples of performance estimation; Overlapped training data; Elevated Type I error for comparison; Underestimated performance variance or overestimated degree of freedom for comparison
Leave-One-Out cross-validation	Unbiased performance estimation	Very large variance
Repeated *k*-fold cross-validation	Large number of performance estimates	Overlapped training and test data between each round; Underestimated performance variance or overestimated degree of freedom for comparison
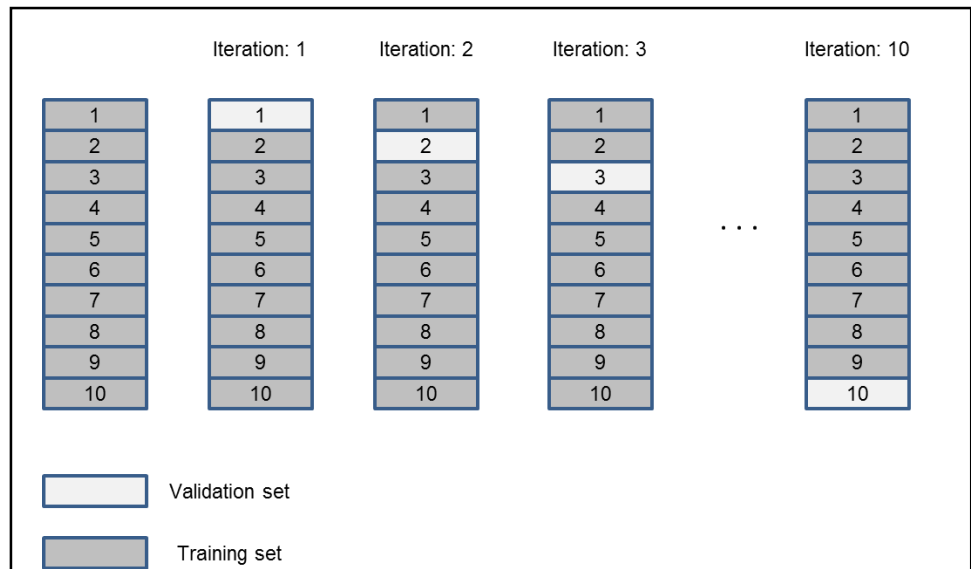
Figure 5.4. 10-fold cross validation.

		Actual	
		Positive	Negative
Predicted	Positive	True positive	False positive
	Negative	False negative	True negative

Figure 5.5. A confusion matrix (Dhawan, 2003).

The results of the classifier, based on five sampling sets, indicate that the lowest average accuracy of validation sets is 89.3%, based on nine hidden nodes, and the highest value is 91.0%, based on 14 hidden nodes (Table 5.2). However, the standard deviation of the classifier with 14 hidden nodes is higher (1.3) than other hidden node settings, namely 1, 6, 11, 12 and 13 hidden nodes.

Additionally, the graph in Figure 5.6 of the average accuracy of validation sets of five sampling sets of 10-fold cross-validation shows that the reliability of MLP with 14 hidden nodes achieves above 90%, which is higher than for the other ANNs settings. The standard deviation of the 14 hidden nodes MLP is 1.3, which is higher than for the MLP with 1, 6, 11, 12 and 13 hidden nodes. The one hidden layer MLP with six hidden nodes yields high average accuracy and low standard deviation. Therefore, the optimal MLP setting achieving high average accuracy and low standard deviation is one hidden layer MLP with six hidden nodes (90.3% average accuracy and 0.5 standard deviation), as shown in Table 5.3. To improve the classification results, principal component analysis is applied in the next section to deal with dimensional complexity.
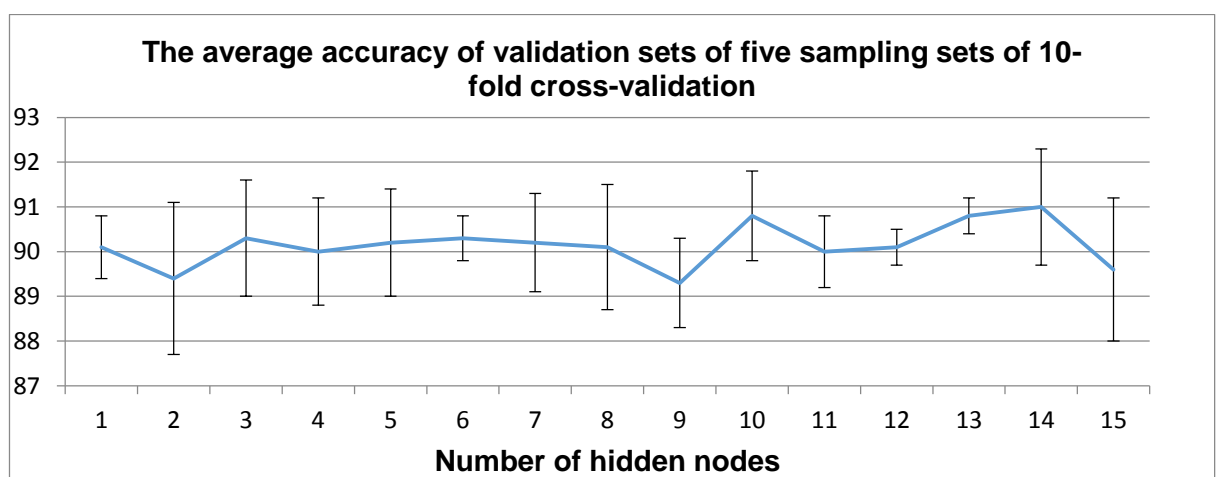
Figure 5.6. The average accuracy of validation sets of five sampling sets of 10-fold cross-validation.

Table 5.2. Results of five sampling sets of 10-fold cross-validation.

Number of hidden nodes	Sampling set: 1		Sampling set: 2		Sampling set: 3	
	Average of % Train accuracy	Average of % Test accuracy	Average of % Train accuracy	Average of % Test accuracy	Average of % Train accuracy	Average of % Test accuracy
1	92.8	90.3	92.9	90.3	92.2	90.9
2	92.5	90.3	93.2	90.9	89.3	86.5
3	92.9	90.9	92.7	89.7	92.8	89.7
4	92.4	89.1	92.1	88.6	93.2	90.3
5	92.8	90.7	93.5	89.9	92.3	91.5
6	94.1	90.1	92.4	89.7	93.0	90.1
7	93.9	90.5	93.0	90.9	92.6	89.5
8	93.4	90.3	94.3	89.7	94.1	91.3
9	94.0	89.3	94.0	89.4	94.3	87.8
10	94.4	90.1	94.9	91.9	94.1	91.5
11	94.3	88.8	94.2	90.9	93.9	90.1
12	94.5	90.3	93.9	89.5	94.6	90.5
13	94.8	91.3	94.6	90.9	94.2	90.5
14	94.7	91.9	94.4	91.7	94.6	92.2
15	93.9	88.4	95.2	90.7	94.3	87.8

Number of hidden nodes	Sampling set: 4		Sampling set: 5	
	Average of % Train accuracy	Average of % Test accuracy	Average of % Train accuracy	Average of % Test accuracy
1	92.8	89.0	92.7	90.1
2	92.8	89.5	92.6	89.9
3	92.6	88.8	92.6	92.2
4	93.4	90.1	93.2	91.7
5	93.2	88.4	92.7	90.5
6	92.4	90.3	93.4	91.1
7	94.3	91.3	93.8	88.7
8	94.2	88.0	93.6	91.3
9	93.9	89.3	93.4	90.7
10	94.2	89.5	94.9	90.9
11	94.2	89.5	94.7	90.5
12	94.5	90.1	94.3	90.3
13	94.9	90.3	94.3	90.9
14	94.8	89.3	94.4	90.1
15	95.0	89.2	94.7	91.7

Table 5.3. Average accuracy and standard deviation of five sampling sets of 10-fold cross-validation.

Number of hidden nodes	Average accuracy		Standard deviation	
	Training set	Validation set	Training set	Validation set
1	92.7	90.1	0.3	0.7
2	92.1	89.4	1.6	1.7
3	92.7	90.3	0.1	1.3
4	92.9	90.0	0.6	1.2
5	92.9	90.2	0.5	1.2
6	93.0	90.3	0.7	0.5
7	93.5	90.2	0.7	1.1
8	93.9	90.1	0.4	1.4
9	93.9	89.3	0.3	1.0
10	94.5	90.8	0.4	1.0
11	94.3	90.0	0.3	0.8
12	94.4	90.1	0.3	0.4
13	94.6	90.8	0.3	0.4
14	94.6	91.0	0.2	1.3
15	94.6	89.6	0.6	1.6

5.3.1. Principal component analysis

The first subtask is to reduce the dimension of the set of correlated features while retaining the essential information. To this end, the principal component analysis (PCA) is applied, aimed at transforming orthogonally a set of correlated variables into a new set of uncorrelated components. The principal components (PCs) are found by computing the eigenvectors and eigenvalues of the data covariance matrix. For example, the correlated two dimensional data in Figure 5.7a is difficult to classify, so PCA can maximise the variance. The plotted data of two PCs in Figure 5.7b show less correlated data than the original data (Jolliffe, 2002).

To reduce the dimension of a data set, PCs selection is necessary. The retained number of PCs is supposed to represent the information of the original data as much as possible. A threshold method is common for the PCs selection. The primary threshold method for PC selection is the cumulative percentage of variation. This method is based on the cumulative percentage of information on data that PCs can contribute (Jolliffe, 2002). The typical threshold is between 70% and 90%, but the threshold can be higher or lower than the conventional threshold, depending on the type of application and the nature of its data.
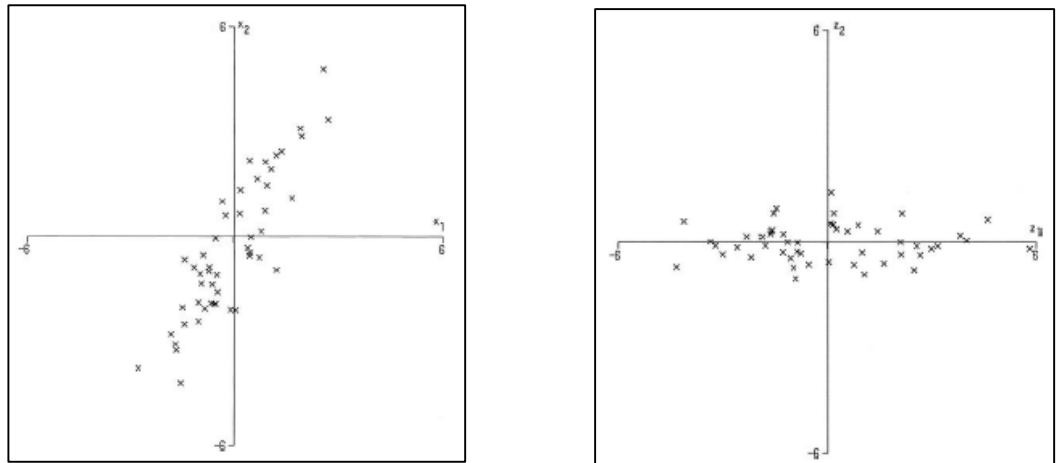
Figure 5.7. (a) Two dimension data (Jolliffe, 2002). Figure 5.7. (b) Two PCs (Jolliffe, 2002).

In this experiment PCA is applied to the projected data features (PCs), and the cumulative percentages of data contained in each of these projected features are shown in Table 5.4. We can choose the threshold in term of the percentage of the information on data that we want to feed to the classifier. For example, if we select six principal components, we will have six dimensions of features containing 91.38% of data information. The cumulative percentages of data given by ten principal components provide less than 1% of data information, and the addition of 10-15 components can only increase the data information by 2.49%. In this experiment, nine principal components are selected as they contain 97.51% of data information. Although, nine components are optimal in this experiment, 15 components have been further tested to investigate the discriminating power of 2.49% of data information.

Once the optimal nine principal components are selected, they are trained by MLP with 10-fold cross-validation. The experiments are run with five shuffling sampling sets; the average accuracy and standard deviation are reported in Tables 5.5 and 5.6. The results indicate that the lowest average accuracy is 89.2%, based on four and five hidden nodes, and the highest average accuracy is 90.5%, based on one hidden layer and nine hidden nodes, with a standard deviation of 0.6. Based on the rule of thumb of MLP, which recommends that the number of hidden nodes should be as small as possible to keep the model generalisation, this study uses a MLP with six hidden nodes and one hidden layer, as it achieves 90% average accuracy and a low standard deviation of 0.7.

Table 5.4. The cumulative percentage of data of each principal component

Principal components	Cumulative percentage
1	40.38
2	63.96
3	74.12
4	81.50
5	87.10
6	91.38
7	94.62
8	96.19
9	97.51
10	98.34
11	98.93
12	99.46
13	99.77
14	99.95
15	100.00

Table 5.5. Average accuracy and standard deviation of five sampling sets of 10-fold cross-validation of nine principal components

Number of hidden nodes	Average accuracy		Standard deviation	
	Training set	Validation set	Training set	Validation set
1	91.3	89.5	0.1	1.1
2	91.5	89.9	0.2	1.4
3	91.6	89.6	0.1	0.5
4	91.3	89.2	0.2	0.4
5	92.4	89.2	0.4	1.2
6	92.5	90.0	0.2	0.7
7	92.5	89.4	0.4	0.6
8	92.9	89.4	0.2	1.1
9	93.4	90.5	0.2	0.6

Table 5.6. Average accuracy and standard deviation of five sampling sets of 10-fold cross-validation of 15 principal components

Number of hidden nodes	Average accuracy		Standard deviation	
	Training set	Validation set	Training set	Validation set
1	92.9	90.7	0.3	0.6
2	92.7	89.4	0.4	0.8
3	93.3	90.5	0.3	0.7
4	92.7	89.5	0.7	0.8
5	93.4	89.3	0.4	1.0
6	93.7	90.1	0.2	0.9
7	93.4	90.2	0.4	0.8
8	94.2	90.2	0.2	0.8
9	94.2	90.3	0.5	1.9
10	94.1	91.1	0.6	1.3
11	94.7	90.6	0.4	0.7
12	94.8	91.0	0.3	1.2
13	94.7	91.1	0.4	0.9
14	94.5	89.9	0.5	1.1
15	94.8	90.6	0.4	1.4

5.3.2. Feature ranking

The second subtask is to investigate the ranking of these 15 features to understand the relationship between the ranked features, based on their information content and pathologists' knowledge and strategies.

Feature selection techniques can be categorised into three groups: wrappers, embedded methods and filters (Guyon & Elisseeff, 2003). Wrappers use a black-box machine learning approach to score variables according to their predictive power, whereas embedded methods perform variable selection in the process of training.

Instead of using a machine-learning approach, the filter methods select variables based on the characteristics of the variables. To measure their discriminative power this study focused on the filter method and applied five evaluation criteria: (i) t-test, (ii) entropy, (iii) minimum attainable classification error (Chernoff bound), (iv) area between the empirical receiver operating characteristic (ROC) curves and the random classifier slope, and (v) absolute value of the standardised u-statistic of a two-sample unpaired Wilcoxon test (Mann-Whitney). A linear classifier is applied to measure the discriminative power of each feature.

- **T-test results**

The results of the feature ranking method based on t-test show that the most important feature is the variance of grey scale of stroma regions, achieving accuracy of 85.3% (Table 5.7). The second group of ranked features consists of the percentage of trophoblast per villi, percentage of stroma per villi, corner per perimeter ratio and corner points, achieving from 73.4 to 75.8% accuracy respectively. The least significant feature is the percentage of red blood cells per villi (53.3% accuracy).

The feature ranking method based on t-test is applied to rank features by using the distance calculated from the mean of each class. Although this method cannot guarantee the discriminating power ranking of the listed features, it gives a ranking of the features that relate well to the classification results of a linear classifier, as shown in Figure 5.8.

Table 5.7. The features ranked by t-test.

Ranking	Features	Accuracy %
1	Variance of grey scale of stroma regions	85.3
2	% TB/villi	75.8
3	% Stroma/villi	74.6
4	Corner/villi's perimeter ratio	74.2
5	TB skeleton /TB's perimeter ratio	74.0
6	Corner points	73.4
7	Trophoblast analysis	70.9
8	Major axis	62.4
9	Villi's bounding box and villi areas	68.2
10	The notion of four quadrants	65.1
11	Villi pixels	61.6
12	Minor axis	55.0
13	Elongation ratio	61.2
14	% Edge/stroma	56.6
15	% RBC/villi	53.3

- **Entropy**

Entropy is implemented to measure the uncertainty of information content in data. The results of the feature ranking method based on entropy show that the top three essential features are: trophoblast analysis, the number of villi boundary corner points, and the variance of grey scale of stroma regions (Table 5.8). The less significant features include the percentage of edge inside stroma regions, the percentage of red blood cells and elongation ratio. The ranked features indicate the development of the discriminative power of the features, but some of the listed features are still not associated with the results of the linear classifier, such as the variance of grey scale of stroma regions, the

notion of four quadrants, the different area between villi's bounding box and villi areas, villi pixels, major axis and the percentage of red blood cells inside stroma regions (Figure 5.9).
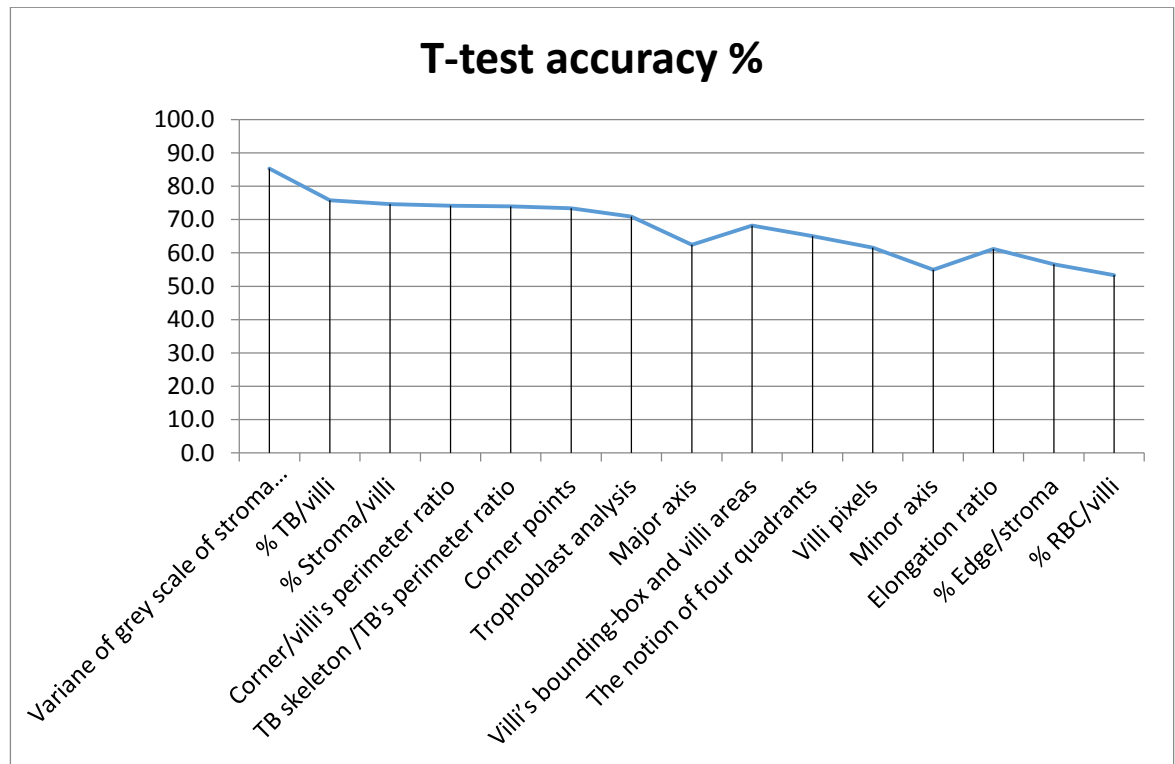
Figure 5.8. The accuracy of features ranked by t-test.

• **ROC curves criterion**

The results of the feature ranking method based on the area between the empirical receiver operating characteristic (ROC) curves and the random classifier slope, show that the features ranked by this criterion are related to the classification results of a linear classifier. The top six features are: the variance of grey scale of stroma regions, the percentage of trophoblast per villi, the percentage of stroma per villi, the ratio between number of villi boundary corner points and all pixels belonging to villi perimeter, the number of villi boundary corner points and the trophoblast skeleton per trophoblast perimeter ratio (Table 5.9). The ranked features show that the most important feature is the variance of grey scale of stroma regions, followed by the second group consisting of the percentage of trophoblast per villi, the percentage of stroma per villi, the ratio between number of villi boundary corner points and all pixels belonging to villi perimeter, the number of villi boundary corner points, and the trophoblast skeleton per trophoblast perimeter ratio (Figure 5.10). The least significant features include the minor axis, the percentage of edge pixels inside stroma regions and villi pixels. This criterion can be used to order the features using their discriminating aspect.

Table 5.8. The features ranked by entropy.

Ranking	Features	Accuracy %
1	Trophoblast analysis	70.9
2	Corner points	73.3
3	Variance of grey scale of stroma regions	85.3
4	The notion of four quadrants	65.1
5	Villi's bounding box and villi areas	68.2
6	Villi pixels	61.6
7	% TB/villi	75.8
8	% Stroma/villi	74.6
9	TB skeleton /TB's perimeter ratio	74.0
10	Corner/villi's perimeter ratio	74.2
11	Minor axis	55.0
12	Major axis	62.4
13	Elongation ratio	61.2
14	% RBC/villi	53.3
15	% Edge/stroma	56.6
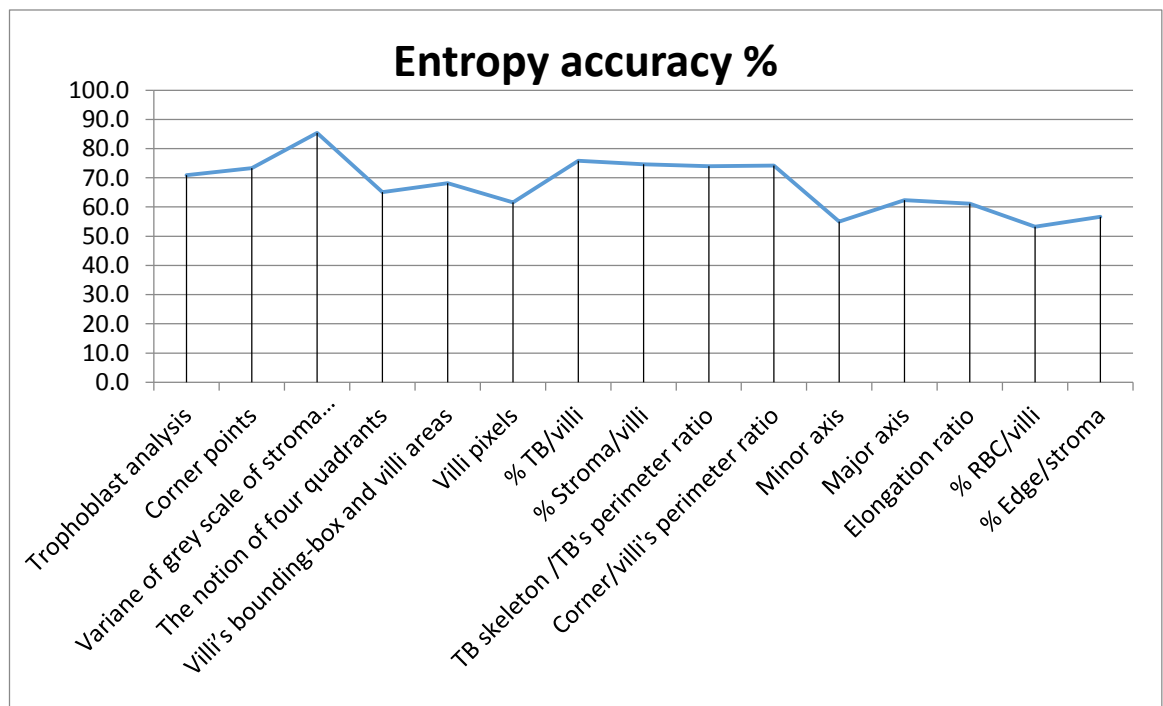
Figure 5.9. The accuracy of features ranked by entropy.

Table 5.9. The features ranked by area between ROC and the random classifier slope.

Ranking	Features	Accuracy %
1	Variance of grey scale of stroma regions	85.3
2	% TB/villi	75.8
3	% Stroma/villi	74.6
4	Corner/villi's perimeter ratio	74.2
5	Corner points	73.3
6	TB skeleton /TB's perimeter ratio	74.0
7	Trophoblast analysis	70.9
8	Villi's bounding box and villi areas	68.2
9	The notion of four quadrants	65.1
10	% RBC/villi	53.3
11	Major axis	62.4
12	Elongation ratio	61.2
13	Villi pixels	61.6
14	% Edge/stroma	56.6
15	Minor axis	55.0
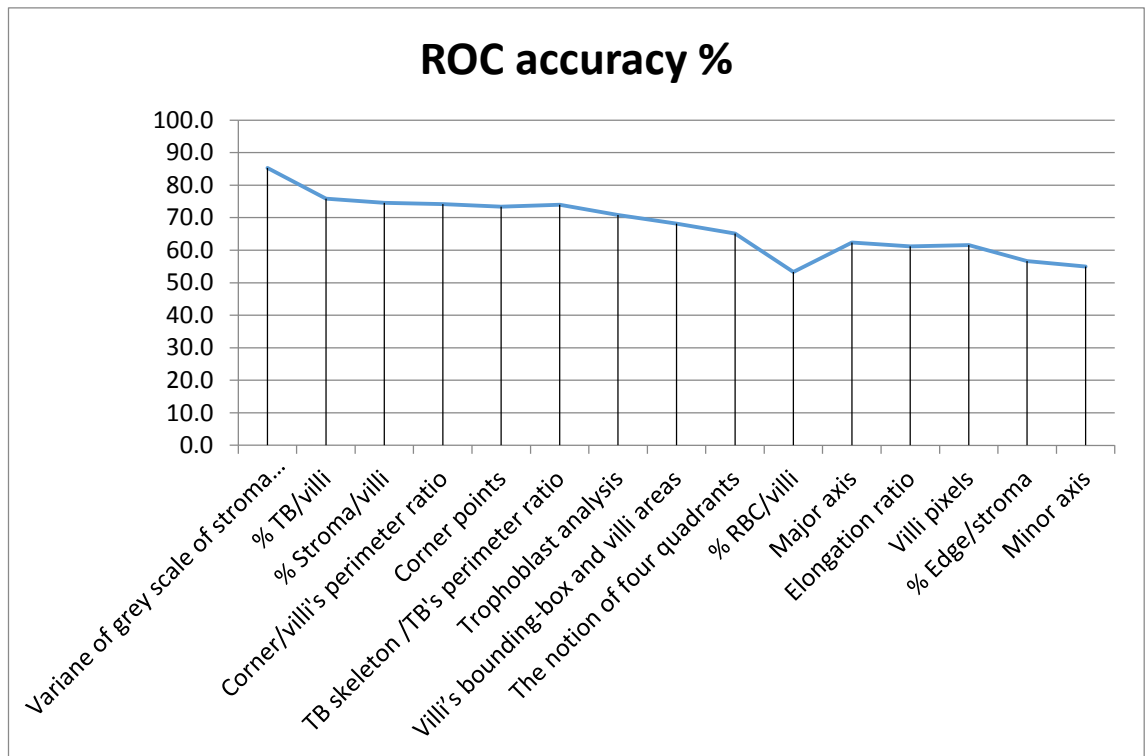
Figure 5.10. The accuracy of features ranked by the area between ROC and the random classifier slope.

- **Chernoff bound criterion**

Minimum attainable classification error (Chernoff bound) is used to measure the distance that the expected value of a random variable can diverge from the mean of data. The important features, listed by this criterion, are: trophoblast analysis, number of villi boundary corner points, and variance of grey scale of stroma regions. The less important features are: percentage of edge pixels inside stroma regions, elongation ratio, and percentage of red blood cells (Table 5.10). The ranked features show the trend of the discriminative power of the features, but some of the listed features are still not related to the classification results of the linear classifier, for example, the notion of four quadrants, villi pixels, major axis and the percentage of red blood cells inside stroma regions (Figure 5.11).

Table 5.10. The features ranked by minimum attainable classification error (Chernoff bound).

Ranking	Features	Accuracy %
1	Trophoblast analysis	70.9
2	Corner points	73.3
3	Variance of grey scale of stroma regions	85.3
4	The notion of four quadrants	65.1
5	Villi pixels	61.6
6	Villi's bounding box and villi areas	68.2
7	% TB/villi	75.8
8	% Stroma/villi	74.6
9	TB skeleton /TB's perimeter ratio	74.0
10	Minor axis	55.0
11	Corner/villi's perimeter ratio	74.2
12	Major axis	62.4
13	% RBC/villi	53.3
14	Elongation ratio	61.2
15	% Edge/stroma	56.6
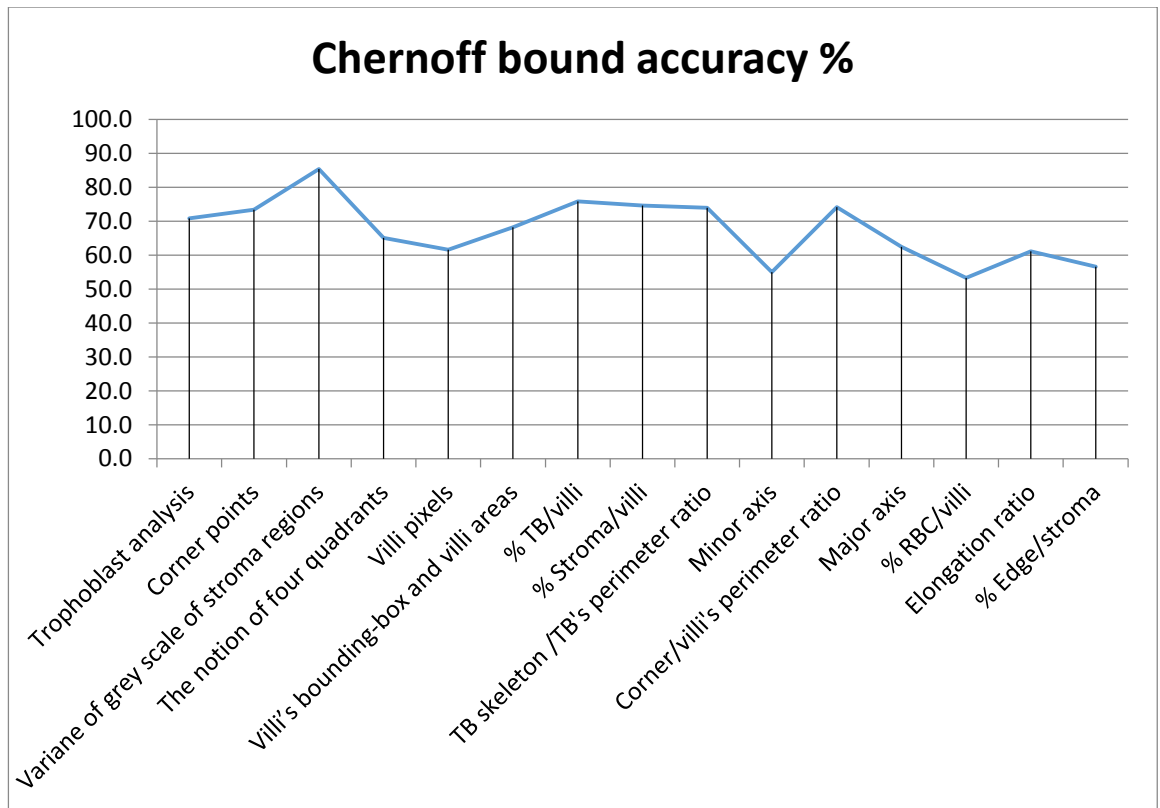
Figure 5.11. The accuracy of features ranked by the minimum attainable classification error (Chernoff bound).

- **Wilcoxon test**

The absolute value of the standardised u-statistic of a two-sample unpaired Wilcoxon test (Mann-Whitney) is a non-parametric statistical test applied to measure the difference between data distributions in two different conditions. The top three important features selected by this criterion are: percentage of stroma per villi, percentage of red blood cells inside stroma regions, and percentage of edge inside stroma regions, whereas the least significant feature listed is the number of villi boundary corner points (Table 5.11). The results show that the features ranked by this criterion are unrelated to the classification results of the linear classifier, as shown in Figure 5.12.

Table 5.11. The features ranked by absolute value of the standardised u-statistic of a two-sample unpaired Wilcoxon test (Mann-Whitney).

Ranking	Features	Accuracy %
1	% Stroma/villi	74.6
2	% RBC/villi	53.3
3	% Edge/stroma	56.6
4	Minor axis	55.0
5	Villi pixels	61.6
6	Elongation ratio	61.2
7	Major axis	62.4
8	The notion of four quadrants	65.1
9	Variance of grey scale of stroma regions	85.3
10	Villi's bounding box and villi areas	68.2
11	Trophoblast analysis	70.9
12	% TB/villi	75.8
13	TB skeleton /TB's perimeter ratio	74.0
14	Corner/villi's perimeter ratio	74.2
15	Corner points	73.3
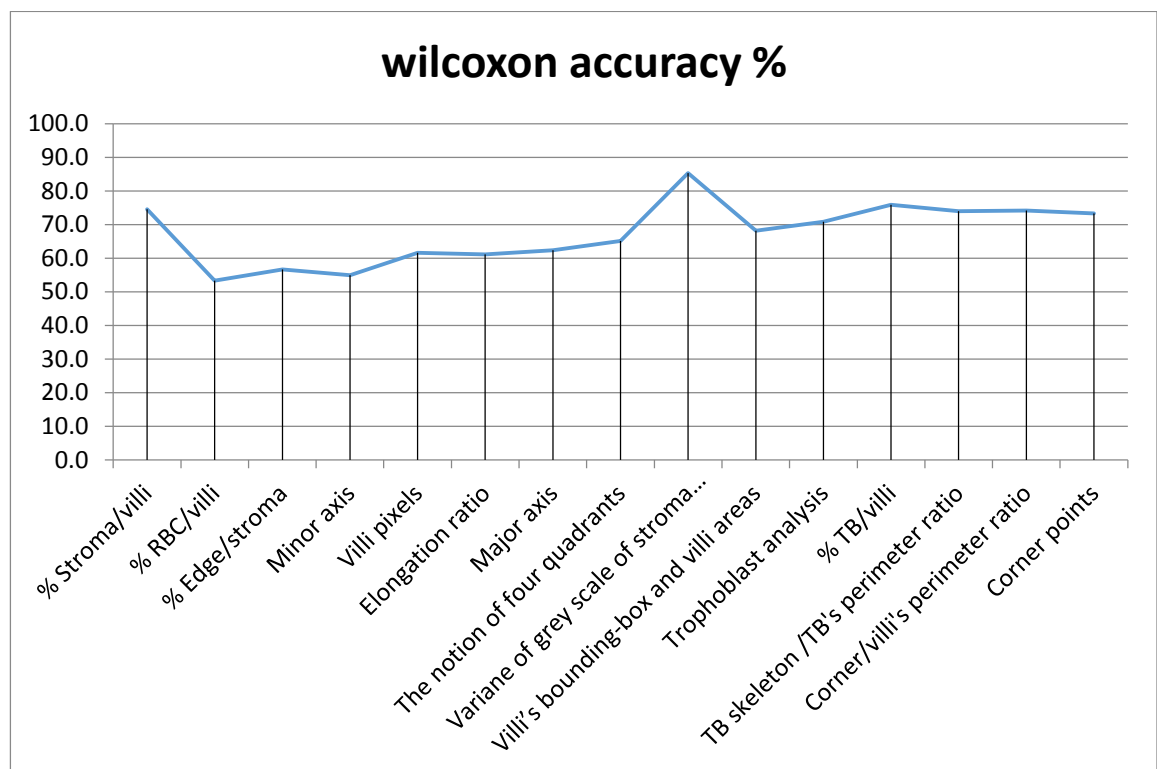
Figure 5.12. The accuracy of features ranked by absolute value of the standardised u-statistic of a two-sample unpaired Wilcoxon test (Mann-Whitney).

Based on the above evaluation criteria, the three most important features are identified as: variance of grey scale of stroma regions, percentage of trophoblast per villi, and percentage of stroma per villi (Table 5.12). The t-test and ROC criteria have ranked the variance of grey scale of stroma regions as the most important feature, with 85.3% accuracy. They have also listed the percentage of trophoblast per villi as the second criterion, with 75.8% accuracy. The percentage of stroma per villi appeared to be the third criterion identified by t-test and ROC, with 74.6% accuracy. However, it is ranked as the first criteria by Wilcoxon test.

The least important feature by t-test, entropy, Chernoff bound and ROC criteria is the percentage of red blood cells inside stroma regions, although Wilcoxon test ranked it second, similarly the percentage of edge inside stroma regions.

This feature ranking experiment shows that the textural feature (e.g. variance of grey scale of stroma regions) is more important than the size and shape of villi, whereas the experts claim to consider the size and shape first, before texture. As the experts explain, the distinction between PHM and CHM villi is the most difficult task. The ranked features can be good indicators to help distinguish PHM from normal and CHM villi.

5.4. Experimental study 2

Having identified the distinguishing features between normal and CHM villi, the experts set out to include PHM villi in the comparative analysis of anomalous villi. The aim of the second task and experiment 2 is to classify villi images into these three categories: normal, PHM and CHM villi. Experiment 2 uses a traditional multilayer perceptron consisting of the same 15 input nodes as experiment 1, but with output nodes set to three classes. The number of hidden layers is again set to one layer, and the number of hidden nodes is between one and 15 (Figure 5.13). In this experiment, repeated 10-fold cross-validation is also applied to evaluate the results and validate the accuracy and reliability of MLP. The accuracy, precision and recall are also applied to measure the MLP performance.

Table 5.12. The features ranked by the five criteria.

Ranking	T-test	Entropy	ROC	Chernoff bound	Wilcoxon test
1	Variance of grey scale of stroma regions	Trophoblast analysis	Variance of grey scale of stroma regions	Trophoblast analysis	% Stroma/villi
2	% TB/villi	Corner points	% TB/villi	Corner points	% RBC/villi
3	% Stroma/villi	Variance of grey scale of stroma regions	% Stroma/villi	Variance of grey scale of stroma regions	% Edge/stroma
4	Corner/villi's perimeter ratio	The notion of four quadrants	Corner/villi's perimeter ratio	The notion of four quadrants	Minor axis
5	TB skeleton /TB's perimeter ratio	Villi's bounding box and villi areas	Corner points	Villi pixels	Villi pixels
6	Corner points	Villi pixels	TB skeleton /TB's perimeter ratio	Villi's bounding box and villi areas	Elongation ratio
7	Trophoblast analysis	% TB/villi	Trophoblast analysis	% TB/villi	Major axis
8	Major axis	% Stroma/villi	Villi's bounding box and villi areas	% Stroma/villi	The notion of four quadrants
9	Villi's bounding box and villi areas	TB skeleton /TB's perimeter ratio	The notion of four quadrants	TB skeleton /TB's perimeter ratio	Variance of grey scale of stroma regions
10	The notion of four quadrants	Corner/villi's perimeter ratio	% RBC/villi	Minor axis	Villi's bounding box and villi areas
11	Villi pixels	Minor axis	Major axis	Corner/villi's perimeter ratio	Trophoblast analysis
12	Minor axis	Major axis	Elongation ratio	Major axis	% TB/villi
13	Elongation ratio	Elongation ratio	Villi pixels	% RBC/villi	TB skeleton /TB's perimeter ratio
14	% Edge/stroma	% RBC/villi	% Edge/stroma	Elongation ratio	Corner/villi's perimeter ratio
15	% RBC/villi	% Edge/stroma	Minor axis	% Edge/stroma	Corner points
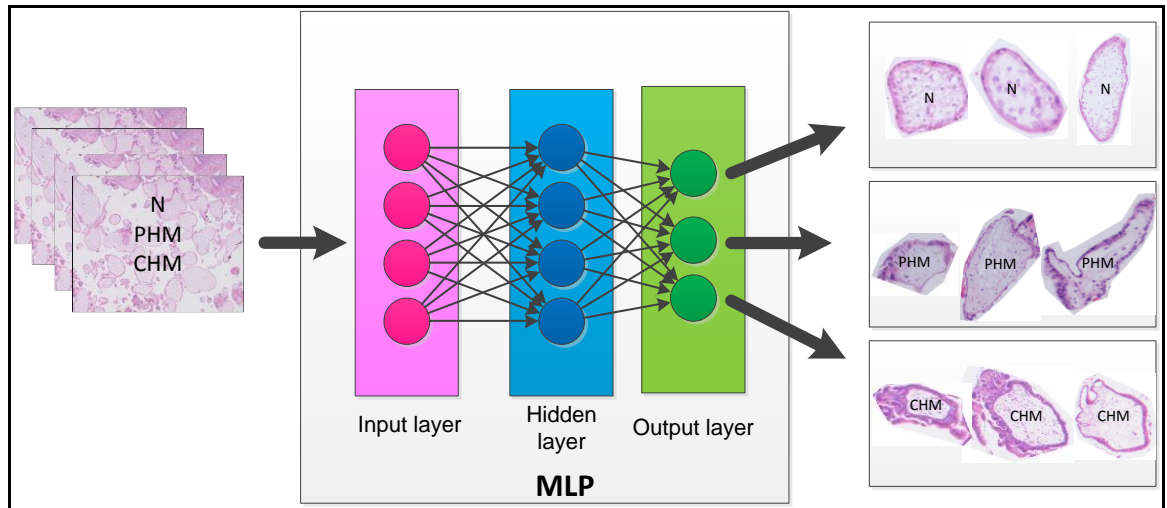
Figure 5.13. The MLP diagram of normal, PHM and CHM villi images classification.

The MLP classification results of ten sampling sets of 10-fold cross-validation indicate that the number of hidden nodes achieving the best average accuracy (80.5%) and standard deviation (0.5) of validation sets is five hidden nodes. Although, 12 hidden nodes MLP yields 81.1% average accuracy, the average standard deviation of validation sets is higher (0.9) than for five hidden nodes, as shown in Table 5.13. The graph of the average accuracy of validation sets indicates that the performance of MLP in term of efficiency is steady from 2 to 15 hidden nodes, as illustrated in Figure 5.14. Furthermore, the one hidden layer MLP with five hidden nodes achieves 81.3% precision for the normal villi class, 81.2% for the PHM villi class and 80.1% for the CHM villi class. The MLP also achieves high recall for the normal (83.2%) and PHM (86.2%) villi classes but the MLP yields low recall for the CHM villi class (62.8%) as shown in Table 5.14.
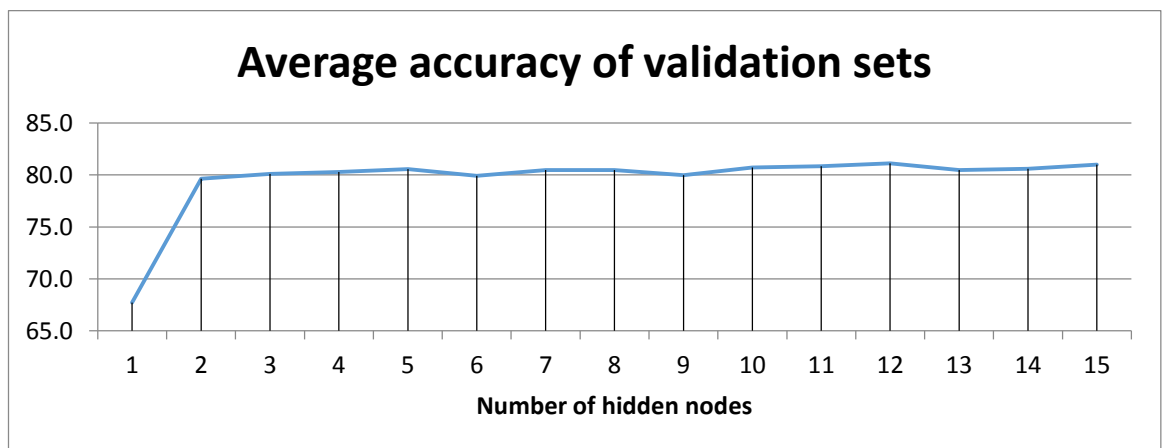

Figure 5.14. Average accuracy of validation sets.

Table 5.13. MLP classification results of ten sampling sets of 10-fold cross-validation.

Number of hidden nodes	Average accuracy		Standard deviation	
	Training set	Validation set	Training set	Validation set
1	69.1	67.7	1.3	1.6
2	82.2	79.6	0.7	0.9
3	82.8	80.1	0.4	0.6
4	83.4	80.3	0.3	0.8
5	83.7	80.5	0.3	0.5
6	83.9	79.9	0.6	0.7
7	84.3	80.5	0.4	1.1
8	84.7	80.4	0.5	0.6
9	84.8	80.0	0.4	0.7
10	85.1	80.7	0.5	0.7
11	85.0	80.8	0.6	0.9
12	85.3	81.1	0.6	0.9
13	85.7	80.5	0.5	1.1
14	85.9	80.6	0.6	0.5
15	85.8	81.0	0.6	0.9

Table 5.14. The precision and recall of MLP.

	Normal	PHM	CHM
Precision	81.3	81.2	80.1
Recall	83.2	86.2	62.8

These results confirm the difficulty of villi classification when the PHM villi are added to the analysis. The expert pathologists claim that the early stage of PHM and CHM are difficult to distinguish by using solely morphological features. To improve the accuracy and recall of MLP, meetings with experts helped improve the configuration of the neural network discussed in the next section.

5.5. Experimental study 3

The aim of this experiment is to capture the heuristic approach of the expert pathologists in improving the design of traditional MLP. The pathologists' approach to the detection of anomalies begins by first separating normal villi from non-normal villi, then they focus on the non-normal villi and separate PHM from CHM villi. This approach is reflected in the novel multi-neural network configuration.

This novel multi-neural network configuration consists of three sub-networks, namely NN1, NN2 and NN3 (Figure 5.15). NN1 is applied to classify normal from non-

normal villi. Then, the villi classified as non-normal villi are trained by NN2 to distinguish between PHM and CHM villi. The villi classified as normal villi are further trained by NN3 to classify normal villi from PHM villi. The primary settings of NN1, NN2 and NN3 consist of 15 input nodes, one hidden layer, hidden nodes varying from one to 15, and two output nodes.
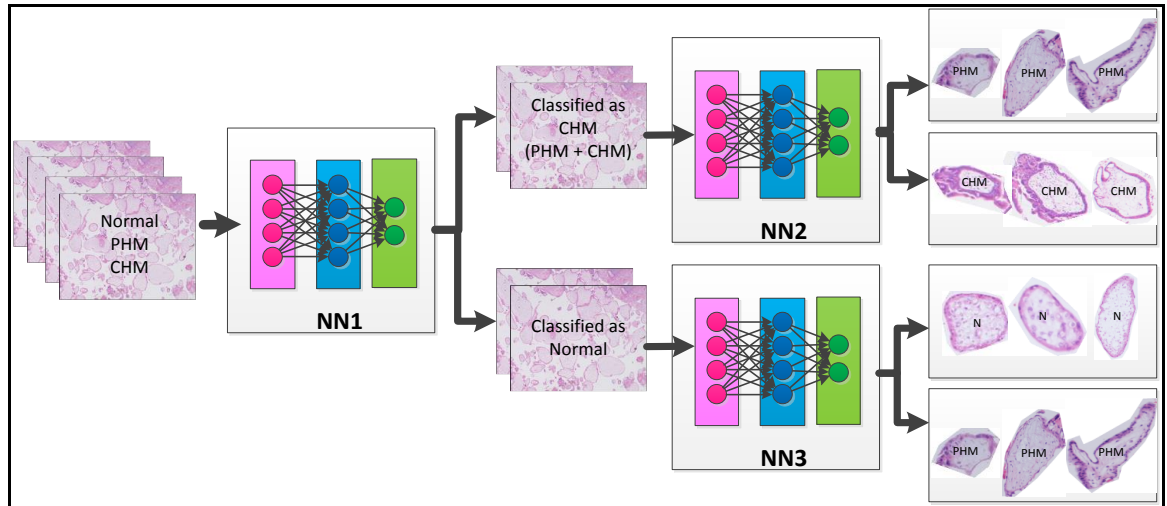
Figure 5.15. Multi-neural network architecture.

The results of the three sub-neural networks indicate that the best number of hidden nodes is six, with 93.1% average accuracy for NN1, three with 89.0% average accuracy for NN2, and five with 91.4% average accuracy for NN3.

The multi-neural network classification results based on the repeated 10-fold cross-validation show that the overall average accuracy is slightly higher (81.2%) than the traditional MLP (80.5%) with standard deviation of 1, as shown in Table 5.15.

The multi-neural network configuration has improved the classification performance in term of average accuracy, precision and recall, achieving 81.2% average accuracy (Table 5.16), and achieved high precision for all three classes: normal (82.7%), PHM (81.3%) and CHM (80.8%) villi. The recall has also increased for normal (83.0%), PHM (87.1%) and CHM (64.3%). However, the improvement in average accuracy is still limited to a small gain compared to MLP results.

Table 5.15. Average accuracy of multi-neural network approach.

Sampling set	Average accuracy of validation set			
	NN1 = 6	NN2 = 3	NN3 = 5	Overall accuracy
1	93.3	88.5	92.2	80.4
2	93.2	89.1	91.4	81.4
3	93.2	89.4	90.6	80.4
4	93.5	89.0	90.9	81.3
5	93.6	88.8	91.4	80.2
6	92.8	89.2	91.2	81.1
7	92.4	88.9	91.2	80.5
8	93.3	88.1	92.1	83.2
9	93.3	89.2	90.8	82.4
10	92.6	89.7	92.1	81.4
Average	93.1	89.0	91.4	81.2
Standard deviation	0.4	0.5	0.6	1.0

Table 5.16. The precision and recall of MLP vs. Multi-neural network (MNN).

ANNs	Precision			Recall			Average accuracy
	Normal	PHM	CHM	Normal	PHM	CHM	
MLP	81.3	81.2	80.1	83.2	86.2	62.8	80.5
MNN	82.7	81.3	80.8	83.0	87.1	64.3	81.2

5.6. Experimental study 4

The previous limited improvement led to elicitation of two additional features, namely dark regions inside stroma, and dark regions inside trophoblast. Dark regions inside stroma and dark regions inside trophoblast represent the texture of stroma and trophoblast regions respectively. The dark regions inside stroma are high in normal placental and low in PHM and CHM villi samples (Figure 5.16). In normal placental villi the structure of the stroma is a net-like structure, whereas it is scalloping, dentate and mild in PHM and enlarged, irregular and karyorrhectic in CHM. The dark regions inside trophoblast are also used to capture the textural features of trophoblast thickness and proliferation of villi. The values of dark regions inside trophoblast are high in CHM and low in PHM and normal placental villi samples (Figure 5.17). This is due to the trophoblast regions in CHM villi being thicker and more proliferated than in normal placental and PHM villi.

These two additional features are added to the 15 features and used to train both, the MLP and the MNN based on the same settings as previous experiments, except that the number of input nodes is configured to 17 and the number of hidden nodes ranges from one to 17.

The MLP results indicate that the highest average accuracy (85.1%) and lowest standard deviation (0.6) are achieved with five hidden nodes (Table 5.17). Although, the MLPs with 12, 14, 16 and 17 hidden nodes achieve slightly better accuracies (Figure 5.18) their standard deviation is higher than for the five hidden nodes. The MLP with nine hidden nodes yields 85.2% average accuracy and 0.5 standard deviations, which are slightly lower than the MLP with five hidden nodes. However, to keep the number of hidden nodes as small as possible in order to maintain the classifier generalisation, it is concluded that the MLP with five hidden nodes is the best setting for the 17 features classification. The average accuracy of the MLP is significantly improved from 80.5% to 85.1% by adding these two textural features.

Table 5.17. MLP classification results of ten sampling sets of 10-fold cross-validation.

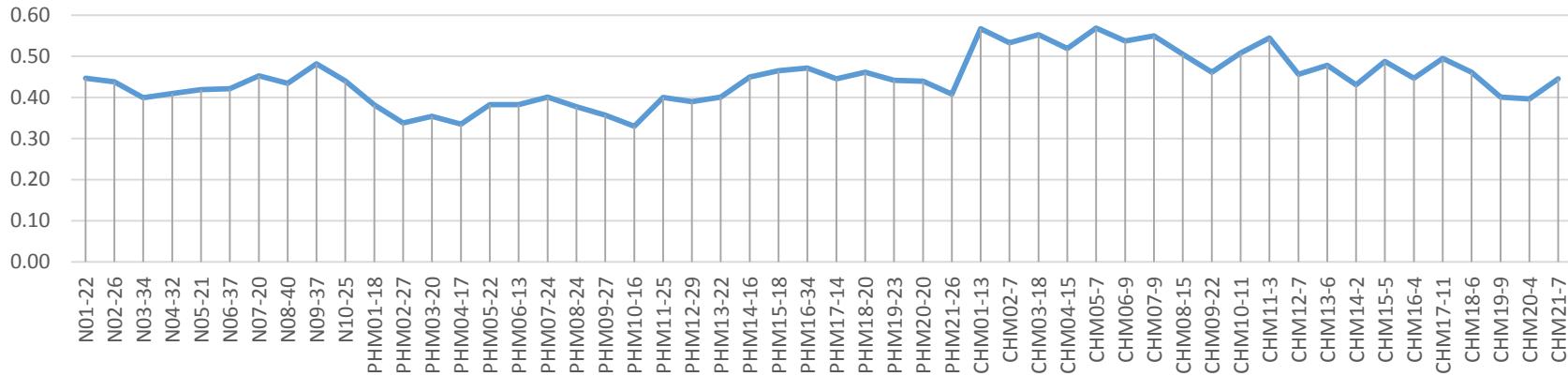| Number of hidden nodes	Average accuracy		Standard deviation	
	Training set	Validation set	Training set	Validation set
1	70.4	69.5	1.6	1.8
2	85.4	82.9	2.6	2.7
3	87.1	84.3	1.3	1.2
4	87.8	84.3	0.3	0.6
5	88.1	85.1	0.4	0.6
6	88.3	84.4	0.3	0.6
7	88.6	84.2	0.2	0.6
8	88.7	84.9	0.4	0.9
9	88.9	85.2	0.6	0.5
10	89.3	85.1	0.6	0.6
11	89.9	84.7	0.6	0.7
12	89.8	85.3	0.4	0.8
13	89.6	84.7	0.4	0.7
14	90.2	85.2	0.5	0.6
15	90.2	85.0	0.5	0.9
16	90.4	85.6	0.4	0.7
17	90.7	85.6	0.3	0.7
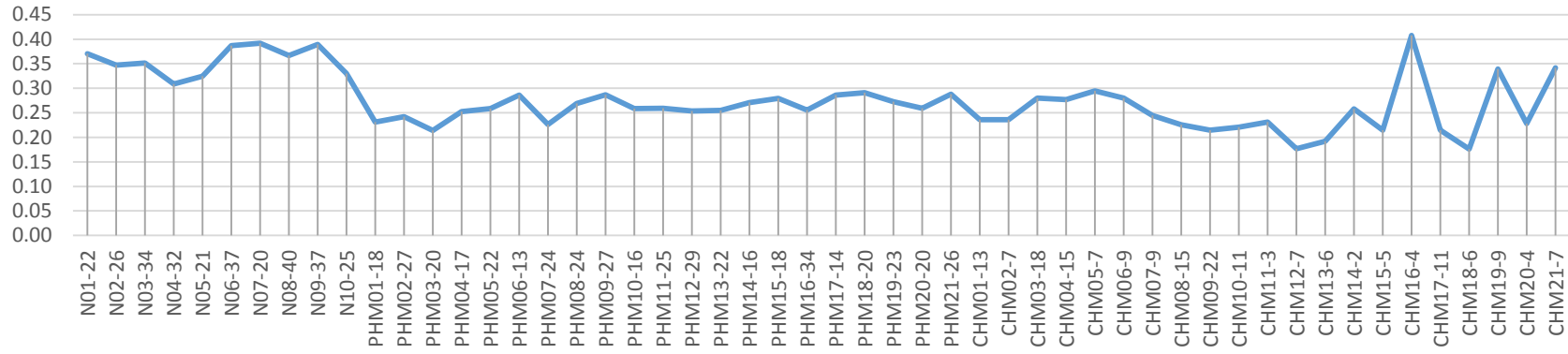
Figure 5.16. Dark regions inside the stroma.

Figure 5.17. Dark regions inside the trophoblast.
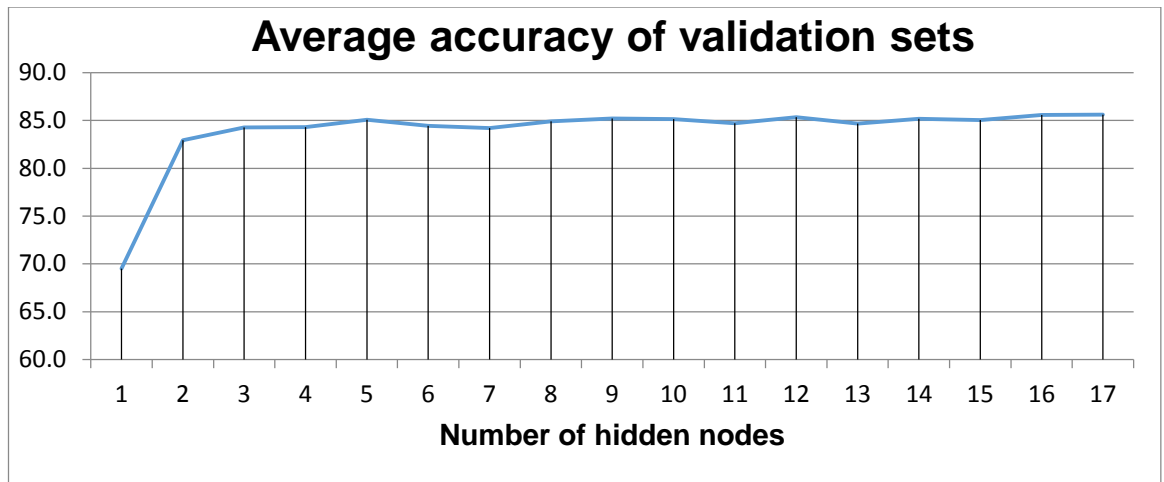
Figure 5.18. Average accuracy of validation sets of MLP with 17 features.

The results of the multi-neural networks show that the best numbers of hidden nodes for NN1, NN2 and NN3 hidden nodes are five, with 96.7% average accuracy, eight, with 92.8% average accuracy and 14, with 94.9% average accuracy respectively. The repeated 10-fold cross-validation indicates that the average accuracy is slightly better (86.1%) than MLP (85.1%) with a standard deviation of 0.6 (Table 5.18).

Table 5.18. Average accuracy of multi-neural network approach with 17 features.

Sampling set	Average accuracy of validation set			
	NN1 = 5	NN2 = 8	NN3 = 14	Overall accuracy
1	96.9	92.4	94.9	85.0
2	96.5	92.7	95.2	86.7
3	96.5	93.0	94.7	85.5
4	96.6	92.7	94.1	86.8
5	96.9	93.2	94.9	85.9
6	97.1	93.5	94.9	86.1
7	96.7	92.8	94.6	86.4
8	96.0	92.8	95.4	85.9
9	96.6	92.6	95.0	85.7
10	96.8	92.5	95.4	87.0
Average	96.7	92.8	94.9	86.1
Standard deviation	0.3	0.3	0.4	0.6

In summary, these additional two features have improved the average accuracies of MLP and MNN, and MNN has reached an average accuracy higher than MLP (i.e. 86.1% vs. 85.1%). Furthermore, the precision of the multi-neural network has increased from 82.7% to 89.5% for normal villi class, from 81.3% to 84.7% for PHM villi class and from 80.8% to 86% for CHM villi class. Similarly, the recall has improved from 83.0% to 87.7% for the normal villi class, from 87.1% to 90.3% for the PHM villi class and from 64.3% to 73.5% for the CHM villi class.

Table 5.19. Precision and recall of 15 and 17 features classified by MLP and MNN.

15 FS ANNs	Precision			Recall			Average accuracy
	Normal	PHM	CHM	Normal	PHM	CHM	
MLP	81.3	81.2	80.1	83.2	86.2	62.8	80.5
MNN	82.7	81.3	80.8	83.0	87.1	64.3	81.2

17 FS ANNs	Precision			Recall			Average accuracy
	Normal	PHM	CHM	Normal	PHM	CHM	
MLP	86.2	84.6	86.5	87.5	88.0	74.3	85.1
MNN	89.5	84.7	86	87.7	90.3	73.5	86.1

5.7. Post-processing study

To further improve the results, a majority voting technique is applied to the above findings. This is a practice followed by expert pathologists for finalising their diagnosis by taking a holistic view. For example, if a hydatidiform mole slide image contains ten villi and eight are identified as normal placental villi, then the two remaining villi are classified as normal, even though they contain a few CHM villi characteristics. The results of majority voting indicate that the average portion of villi classified as normal (91.4%) in normal slides is higher than in PHM (8.2%) and CHM (0.5%) (Table 5.20). The average portion of villi in PHM slides classified as PHM (91.5%) is also higher than the portions classified as normal (4.5%) and as CHM (4.0%). The average portion of CHM villi in CHM slide is 81.6%, which is lower than normal villi (91.4%) in normal slides and PHM villi (91.5%) in PHM slides. However, the average portion of CHM (81.6%) is higher than normal (2.8%) and PHM villi (15.6%) in CHM slides.

The maximum and minimum values of the portions of the three classes (normal, PHM and CHM villi) are shown in the range of variation of classified villi. The maximum and minimum values of the portions in normal slides are 97.5% and 76.5% for normal villi, 23.5% and 2.5% for PHM villi and 4.5% and 0% for CHM villi. The gap between maximum and minimum values indicates that the distribution of classified villi is low and the low standard deviation of each class also confirms the low distribution behaviour of normal

slide classification (Table 5.20). The trends of PHM slides are the same as in normal slides, whereas the ranges between minimum and maximum values of classified villi portions represent the high distribution of classification results, for example, 0 – 18.2% for normal villi, 0 – 42.9% for PHM villi and 54.5 – 100% for CHM villi. The standard deviations of classified normal, PHM and CHM villi are 5.9, 15.9 and 17.5, which emphasises the high variation of the classification results of CHM slides. Although, the classification results of CHM slides show high variation, the minimum of CHM villi portion (54.5%) is still more than 50%. Therefore, majority voting can be used to reduce the misclassifications in the whole slide images (Figure 5.19).

5.8. Conclusion

This chapter describes the second phase of the novel approach, which relates to the high level processing of the villi stained slides.

The first task is to apply principal component analysis (PCA) and feature ranking to deal with the complexity of dimensions and to rank the 15 features. The variance of grey scale of stroma regions is ranked as the most critical feature, with 85.3% accuracy, by the t-test and ROC criteria, and the percentage of trophoblast per villi is selected as the second criteria, with 75.8% accuracy, followed by the percentage of stroma per villi with 74.6% accuracy. The least important features are the percentage of red blood cells inside stroma regions and the percentage of edge inside stroma regions. This feature ranking experiment indicates that a textural feature, namely the variance of grey scale of stroma regions, is more significant than the morphological features (size and shape) of villi, while the experts consider the size and shape before texture as the most important features. The results of PCA and feature ranking also confirm that no single feature can give enough discriminative power for villi classification.

The second task aims at classifying the villi into normal placental, PHM and CHM villi, based on the 15 features defines in Section 4.7. The traditional MLP approach has provided low recall, whereas the heuristic MLP approach has improved the precision of all three classes and the recall of PHM and CHM villi classes, as well as the overall average accuracy. The multi-neural network configuration has helped to show which features are more important. While the experts claim that the important features are the size and shape of villi, the experimental studies have confirmed that texture features are more important than the size and shape of villi. Although the results of the multi-neural network, which are based on 15 anomalous features, show an improvement in average accuracy (81.2%), the recall of CHM villi class is still low (64.3%). The two additional textural features, namely dark regions inside the stroma and dark regions inside the trophoblast,

have increased the average accuracy of multi-neural network configuration (experiment 4) from 81.2% to 86.1%, and the recall of CHM villi class from 64.3% to 73.5%.

To support this study, HYMAT is developed to assist pathologists by visualising the results of the segmentation and provide them with a tool to further analyse the villi and explore further properties. This is described further in the next chapter.

Table 5.20. The majority voting results of normal, PHM and CHM slides classified by the multi-neural network.

Normal slides

	Normal	PHM	CHM
Max %	97.5	23.5	4.5
Min %	76.5	2.5	0.0
Average %	91.4	8.2	0.5
SD	6.4	6.5	1.4

PHM slides

	Normal	PHM	CHM
Max %	16.7	100.0	15.4
Min %	0.0	77.8	0.0
Average %	4.5	91.5	4.0
SD	4.9	7.0	4.3

CHM slides

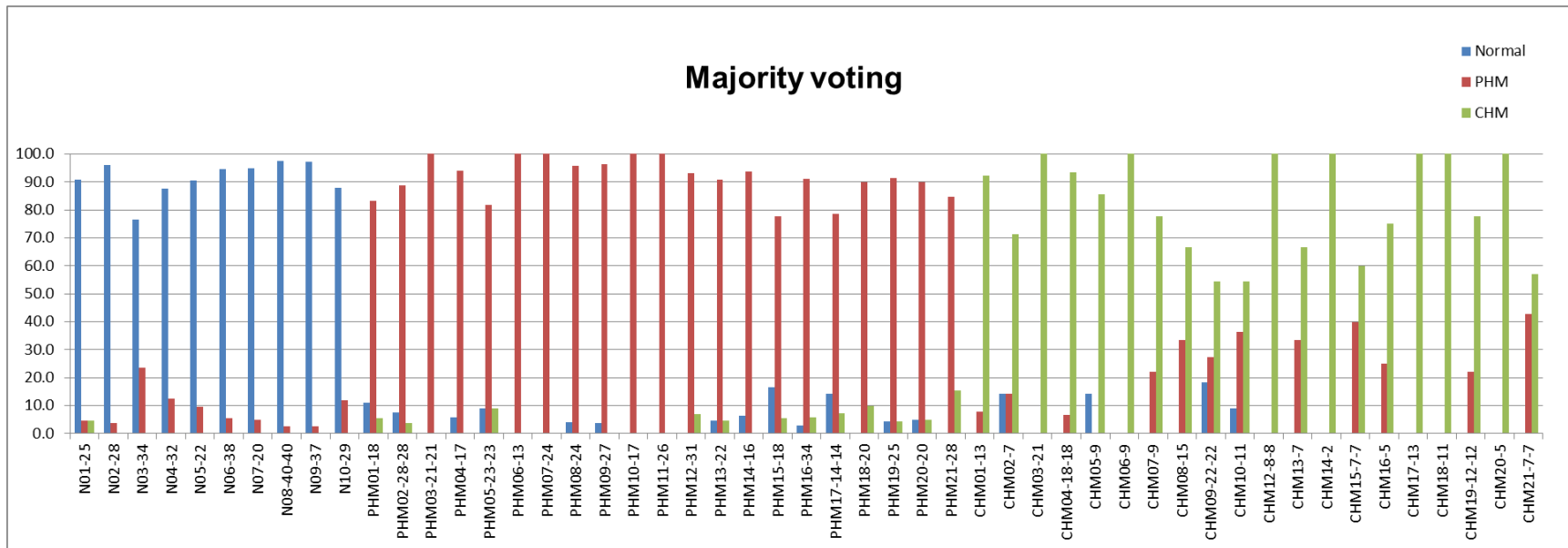|	Normal	PHM	CHM
Max %	18.2	42.9	100.0
Min %	0.0	0.0	54.5
Average %	2.8	15.6	81.6
SD	5.9	15.9	17.5

Figure 5.19. The majority voting results of normal, PHM and CHM slides.

Chapter 6: Hydatidiform Mole Analysis Tool (HYMAT)

6.1. Introduction

One of the main goals of this research is to support pathologists in distinguishing complete hydatidiform moles (CHM) from partial hydatidiform moles (PHM). The distinction between these two types is challenging, yet important, because CHM has a greater malignant potential than PHM and each mole produces a different prognosis and patient management. Although these moles can be identified using clinical, ultrasonographic, gross morphological, histological and genetic criteria, the final diagnosis must be confirmed by pathologists (Vassilakos *et al.*, 1977; Petignat *et al.*, 2003). Many experienced pathologists confirmed the difficulties in distinguishing these moles. Their histological examination aims at analysing trophoblast proliferation, villus contour and its scalloping shape, and the presence of red blood cells in fetal vessels (Sebire *et al.*, 2003). Most misclassifications relate to the absence of strict morphological criteria that can help differentiate the PHM from CHM and normal placenta, because of the significant overlap of features. Some typical cases are not easily detected by focusing solely on morphological criteria. Misclassifications can also occur if no fresh tissue is available, and if abundant tissue of maternal origin is present (Bell *et al.*, 1999). The heuristic approach described in chapter 4 has provided a series of strategies, which have been transformed into a set of software tools, designed to support the low level processing of the stained villi slides. These tools, referred to herewith as Hydatidiform Moles Analysis Tool (HYMAT), focus on the pre-processing, segmentation, feature extraction and analysis tasks, and are described further in the next sections.

6.2. Hydatidiform Moles Analysis Tool (HYMAT)

The system architecture of HYMAT is based on the heuristic approach of expert pathologists and consists of four main steps: pre-processing, segmentation, feature extraction and analysis steps (Figure 6.1). The pre-processing step aims at enhancing the image quality and removing noise or unwanted regions. The histogram equalisation algorithm is applied to support segmentation and enhance the contrast of foreground and background in the pre-processing step (Figure 6.2). The output of this algorithm is illustrated in Figure 6.3.
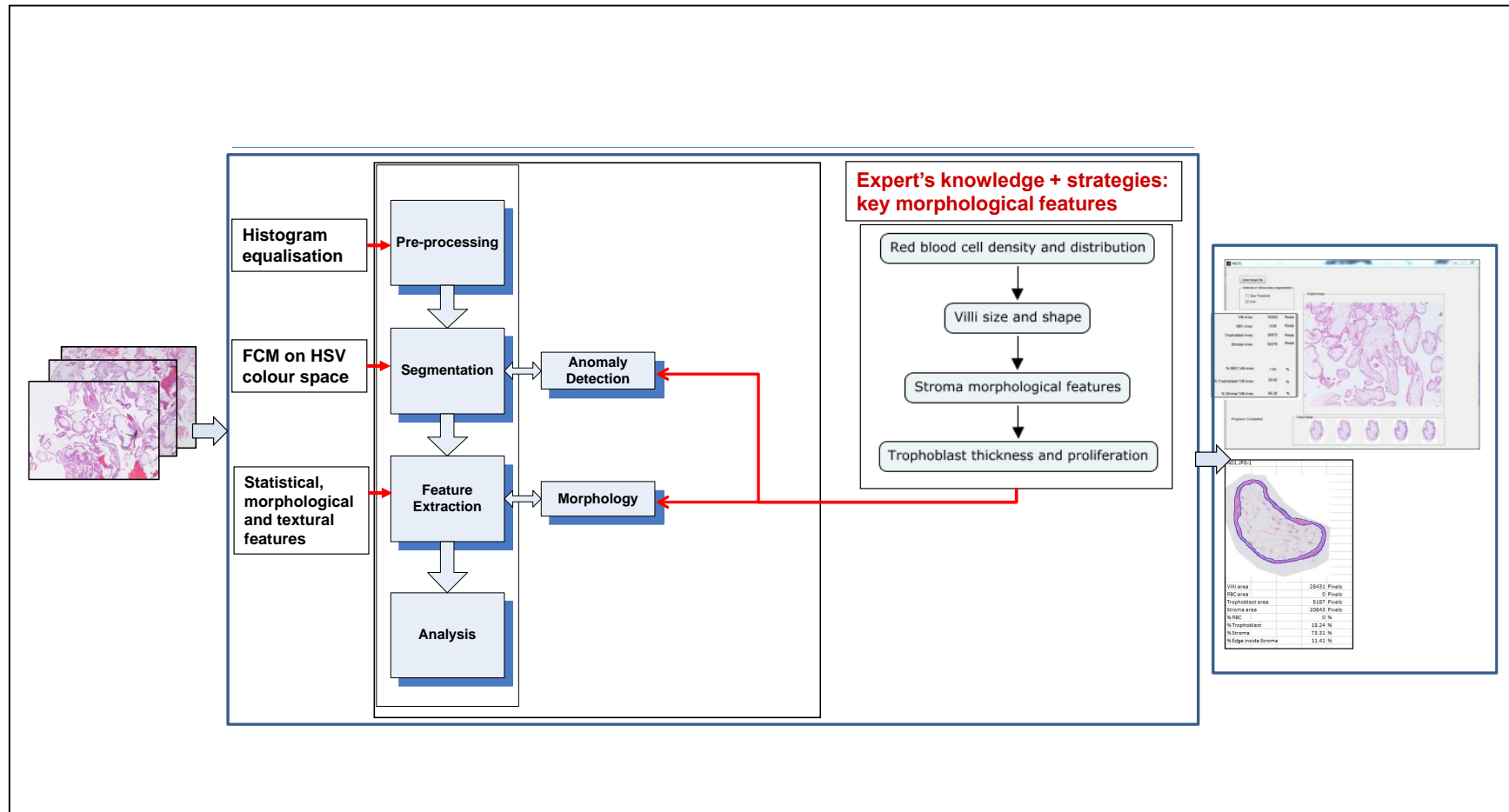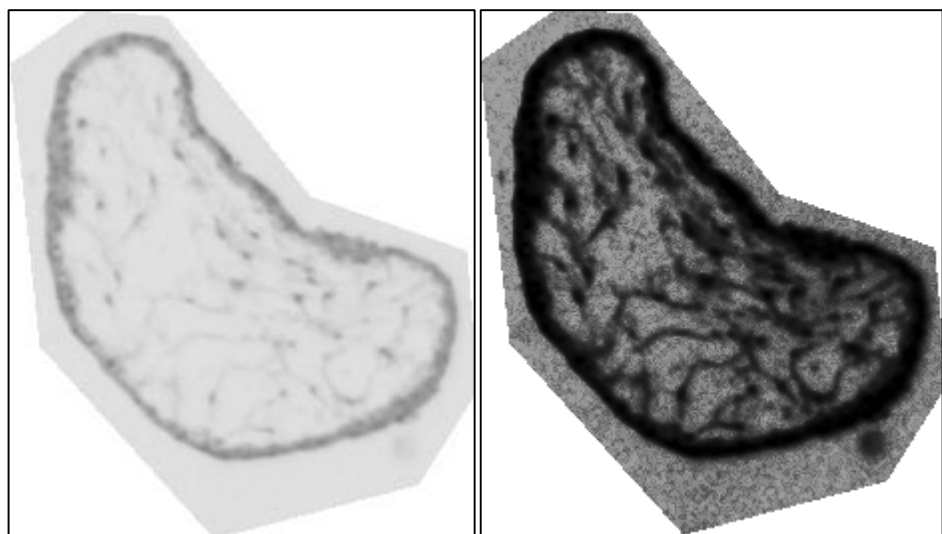
Figure 6.1. The system architecture.

```
// Create a cumulative histogram
for ( i=0; i < SIZE_OF_ HISTOGRAM; ++i )
{
    sum += histogram[i];
    sum_Histogram[i] = sum;
}
// transform image using sum_Histogram
for ( i = 0; i < pixel_Count; ++i )
{
Output_Image[i] = sum_Histogram[image[i]]*MAX_INTENSITY/ pixel_Count;
}
```

Figure 6.2. Histogram equalisation algorithm.

(a) An original villus image   (b) a processed image

Figure 6.3. A histogram equalisation result.

Segmentation is the process of dividing an image into segments that share similar spectral, spatial, and texture characteristics. In this step, the Fuzzy c-Means (FCM) clustering technique, based on HSV colour space, is implemented to segment villi boundaries, red blood cells (RBC), trophoblast and stroma regions. The most common algorithms applied are boundary–based segmentation techniques, active contour techniques, gradient vector flow snakes (GVF snake) and Hough transforms, which can

121

achieve promising segmentation results when they are applied to segment clear boundary objects. However, as our villi's boundaries are unclear, fuzzy and broken, the boundary-based techniques cannot yield satisfactory results. In addition, active contour algorithms cannot cope with overlapping cells (Doukas *et al.*, 2010). To deal with unclear and fuzzy boundaries of villi components, the literature review showed that the FCM technique is the most suitable segmentation approach for the histopathological analysis of placental villi. The RGB (i.e Red Green Blue) villi images are converted to HSV (i.e. Hue Saturation Value) images because HSV colour space can provide more discriminating power in terms of colour and this colour space is similar to a human visual system. The outputs of this step are given in Figures 6.4 and 6.5, showing the four segmented regions, namely villi boundary, RBC, stroma and trophoblast regions.

The third step focuses on extracting meaningful parts of the stained slide, such as characteristics related to the morphological and pathological villi features, as elicited from the experts. This step consists of algorithms responsible for encoding the image contents in a descriptive way. The algorithms combined HSV colour space with FCM associated anomalies, as described in Table 4.1.

The fourth step aims at analysing and interpreting the results achieved from the previous steps and producing a descriptive summary for each villus stored in a Microsoft Excel sheet (Figure 6.6), as requested by the experts. The analysis results include raw morphological and analysed feature groups. The raw morphological feature group contains villi, RBC, trophoblast and stroma areas in terms of pixel, whereas the analysed feature group contains the percentage of red blood cells inside stroma regions, the percentage of stroma regions inside villi, the percentage of trophoblast inside villi and the trophoblast skeleton per trophoblast perimeter ratio. These two groups are also displayed next to the original image window and the thumbnail images resulting from the segmentation step are displayed below the original image window as shown in Figure 6.4.
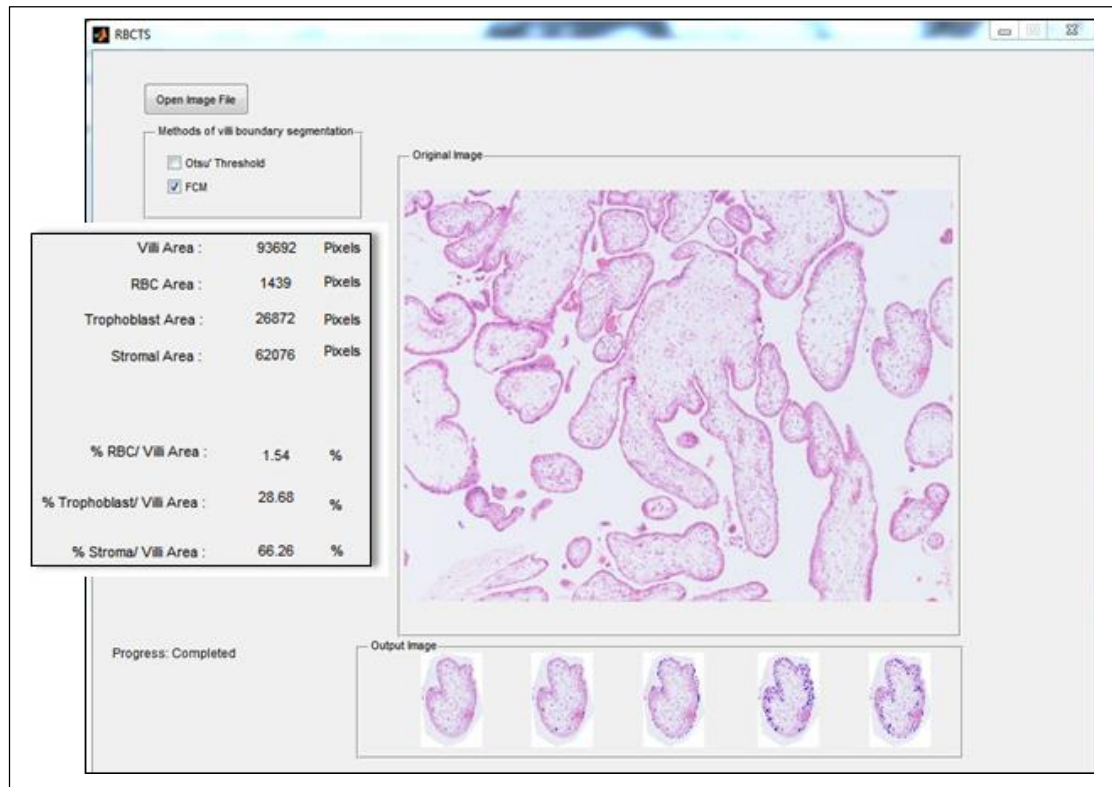
Figure 6.4. HYMAT GUI.

6.3. Implementation of HYMAT

HYMAT is developed using MATLAB, which is a high-level language and interactive tool that allows modelling, simulation and visualisation of data using automation capabilities and built-in functions. HYMAT uses features available from the MATLAB Image Processing Toolbox, namely a set of standard functions associated with image segmentation and image enhancement. In addition to these MATLAB features, algorithms related to FCM, fuzzy clustering, and circular distance measurement are developed within MATLAB to support the segmentation and feature extraction steps.

MATLAB also provides a graphical user interface (GUI) based on the point-and-click control approach, which is used to design the interface of HYMAT as illustrated in Figure 6.4. The outputs of the segmentation and feature extraction steps are stored into an Excel sheet that provides the segmented villus and its statistical, morphological and textual features (Figures 6.5 and 6.6). HYMAT is developed in 32 and 64 bit versions, to be compatible with the pathologists' computers.
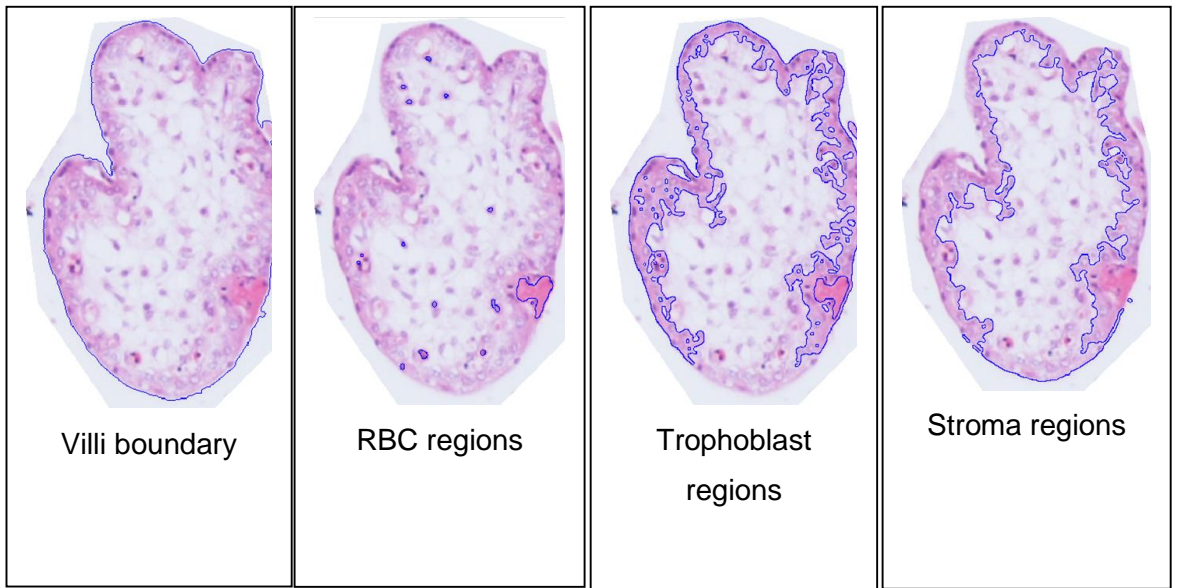
123

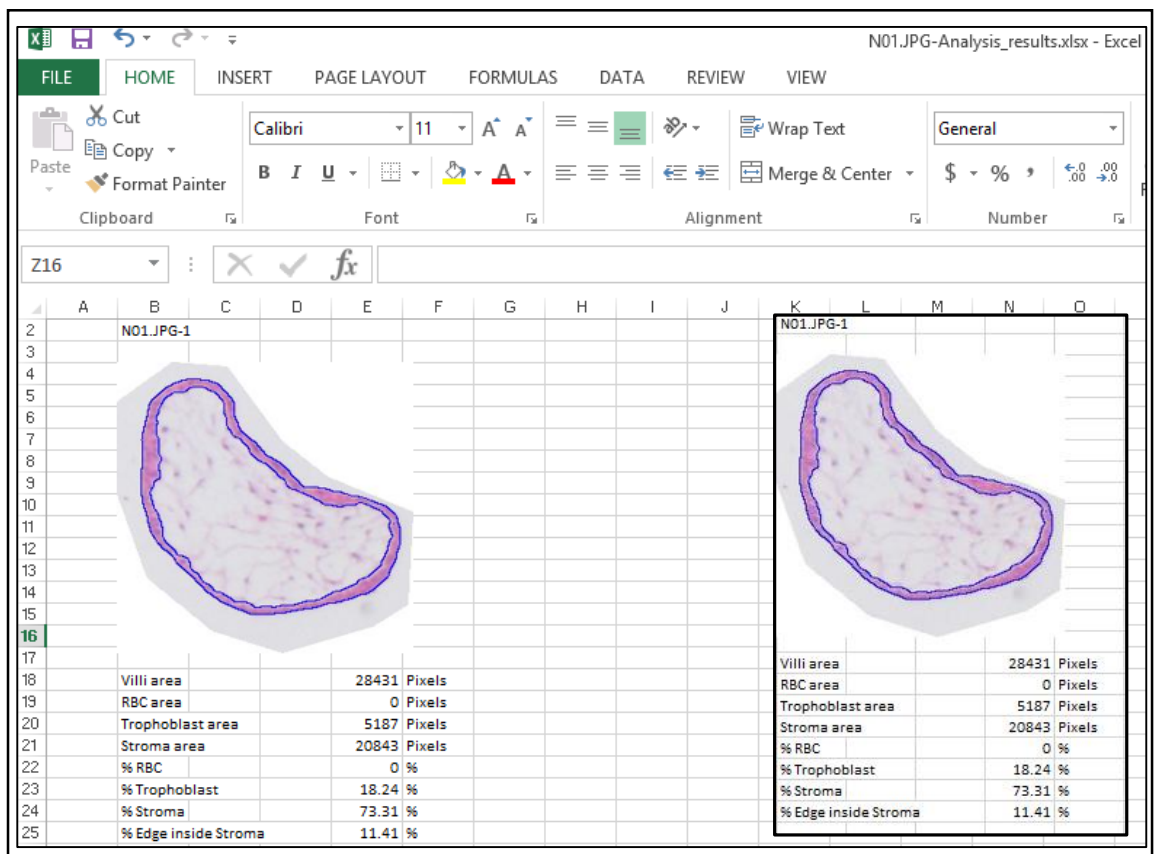| Villi boundary | RBC regions | Trophoblast regions | Stroma regions |

Figure 6.5. Segmentation results of HYMAT.

Figure 6.6.Analysis results of villus N01-2 stored in Microsoft Excel.

6.4. Discussion and conclusion

HYMAT is software developed as part of this research to support solely the low level processing of the heuristic approach discussed in the previous chapters. This tool is also a great asset to the pathologists, who are interested in using it to support their histopathological analysis. It is hoped that the pathologists can use the tool to acquire more stained slides and expand the analyses of villi for future knowledge discovery applications.

The current tool, however, suffers from a few limitations. Currently, the extraction of each single villi from the stained slide is carried out manually by the user: once the villi is extracted HYMAT can pre-process, segment, feature extract and analyse the villi computationally. It is intended to develop the tool in the future so that the system can extract each single villus without need of a user. Though the developed segmentation techniques of HYMAT can handle fuzzy boundary regions, they suffer from over- and under-segmentation problems. Further work is required to (i) extend the current segmentation algorithms so that not only intact villi but also villi with broken/missing boundaries can be processed by HYMAT, and (ii) to integrate the high level processing into HYMAT. The approach used in HYMAT can also be applied to other types of cancerous cells images such as breast, prostate and brain cancerous cells.

Chapter 7: Conclusions and Future work

7.1. Introduction

Molar pregnancy, also known as hydatidiform mole (HM), is an unsuccessful pregnancy resulting from the over-expression of paternal genes and the proliferation of tissue that should develop into the placenta during pregnancy. The classification between complete or partial hydatidiform moles is important for choosing an appropriate treatment method. Kim *et al.* (2009) describe the challenges of the histopathological diagnosis of molar pregnancy and state that even experienced pathologists cannot easily distinguish between Complete Hydatidiform Moles (CHM) and Partial Hydatidiform Moles (PHM). To address this challenge, a novel image understanding, based on anomaly detection guided by experts' heuristic knowledge and strategies, is developed to assist pathologists in HM classification tasks.

The hypothesis of this research project is that an anomaly detection approach for analysing molar pregnancy images can achieve a better image analysis and classification of molar pregnancy types than the current approaches. In this research, three types of villi, normal placental, PHM and CHM, are analysed.

In this study, experts' heuristic knowledge and strategies play an important role, and the definition of their heuristics is an essential aspect of the development of this approach. The definition of a heuristic is: anything that gives a direction guidance to solve a problem. However, heuristics cannot guarantee the final solution (Koen, 1985).

A cognitive approach is a tool for developing an algorithm that mirrors heuristic approach to analyse anomalies in villi. The limitations of the cognitive approach need to be considered before the approach is applied to create the algorithm. One of the important limitations is that experts solve problems mentally, and it is difficult for them to communicate their experience and approaches (Hinds & Pfeffer, 2003). Furthermore, the steps of problem solving, as explained by experts, are more conceptual and abstract representations than the steps described by novices. This can make the expert knowledge difficult to capture.

The developed novel approach is based on capturing the normal and anomalous features associated with the three types of HM (i.e. normal, partial and complete HM) from the experts and the medical documents. These tacit and explicit types of knowledge are used to guide the analysis and classification of villi. The segmentation of the villi image

has primarily been based on Fuzzy c-Means clustering (FCM) and HSV colour space, whereas feature extraction has applied statistical measures to study the morphological and textural features of villi. The classification of the villi was carried out using a multi-neural network architecture that simulated the experts' heuristics.

The heuristic multi-neural network (MNN) configuration is applied to classify the villi into three classes: normal placental, PHM and CHM, based on the 15 features explained in Chapter 4. The results indicate that the multi-neural network improves the classification results, as compared with traditional multi-layer perceptron (MLP) in terms of the overall average accuracy (80.5% to 81.2%), the precision of the three classes and the recall of PHM and CHM villi classes. However, the recall of the CHM villi class is still low (64.3%) Two textural features, namely dark regions inside the stroma and dark regions inside the trophoblast are added to improve the classification results, and the results indicate that the 17 features increase the average accuracy of MNN from 81.2% to 86.1% and the recall of the CHM villi class from 64.3% to 73.5%.

7.2. Research contributions

The novel contributions of this research project are summarised as follows:

(i) A new application domain: this research has focused on the study of the histopathology of molar pregnancy images. The literature review of computational cancer image analysis tends to focus on breast, lung, skin, cervical, and prostate cancers. The current research into molar pregnancy focused solely on the management and treatment aspects. To the best of our knowledge, until now no analysis and classification of these aspects have been carried out. Therefore, research into molar pregnancy image analysis and understanding is still widely open.

(ii) A new image understanding method is developed, combining image processing and artificial intelligence techniques with guidance by experts' knowledge and heuristics. The proposed method focuses on the anomaly detection of hydatidiform mole features. This approach is able to classify the anomalies into four categories, namely: point, contextual, morphological and density anomalies. To our knowledge, this approach has not been adopted in the HM medical diagnosis papers.

127

(iii) A novel multi-neural network architecture that simulates experts' heuristics. This approach is an attempt at modelling computationally the way expert pathologists detect anomalies in villi.

(iv) A cognitive approach to image analysis: the development of an algorithm that mirrors heuristic approach to diagnosing anomalies in villi. What distinguishes this work from others involving experts' knowledge is as follows:

- Whilst current research extracts the morphological knowledge of the expert to analyse their images, this research focuses solely on the anomaly detection of the villi and the critical features in terms of density, size, shape and proliferation.

- Our research has identified new essential features, such as:
  - the number of villi boundary corner points used to describe the shape of villi,
  - the trophoblast skeleton per trophoblast perimeter ratio applied to measure trophoblast thickness and proliferation,
  - the dark regions inside the stroma and the dark regions inside the trophoblast, used to capture stroma and trophoblast surface characteristics.

(v) The development of a tool to support the extraction of critical features, which is used to analyse anomalies. This tool is referred to herewith as HYMAT (hydatidiform mole analysis tool). Though HYMAT is a minor research contribution, it has been found to be a very useful tool and has been adopted by the two expert pathologists involved in this project.

7.3. Limitations

Although this research has achieved significant improvements, there are a number of limitations that have influenced its development. These are listed below.

(i) The villi used in this study belong to the early trimester and the distinction between PHM and CHM is usually unclear; even experienced pathologists cannot easily distinguish between them.

(ii) The study is based on 986 villi; however, the analysis could only be carried out on 939 villi that are intact. It was difficult to obtain a larger set of villi, due to the fact

that HM is a rare disease and the availability and access to the villi samples is extremely difficult.

(iii) The complexity of the pathologists' task in articulating their strategies and heuristics. The meetings with the experts demonstrated the difficulty in obtaining consensus, as they have their personal approach to the detection of anomalous villi.

(iv) The proposed approach can segment the four villi regions defined in Chapter 4 and can deal with fuzzy boundary regions. However, the segmentation results still suffer from over- and under-segmentation problems. This requires further research to improve the segmentation.

(v) The current approach can segment the placental image containing complete villi but fails to segment a villus containing multi-stroma regions. The developed method is limited to a villus with one stroma region.

(vi) The current version of HYMAT works in the semi-automatic mode that requires a user to operate the tool. This task is time-consuming, and the user has to learn how to run the tool. Further improvements are to be conducted for future work.

7.4. Conclusions and future work

This research has proposed a new image understanding, based on image processing and artificial intelligence techniques guided by experts' heuristic knowledge and strategies, to analyse three types of villi, consisting of normal placental, PHM and CHM. The anomaly detection of HM features is applied to analyse villi images.

Initially, the pathologists relied on 15 features to detect anomalous villi. This study has identified two additional features that could be used to improve the detection of anomalies. It has also identified the most critical features, which are now adopted by the pathologists. The final analysis results show that a textural feature, namely the variance of grey scale of stroma regions, is more important than the morphological features (size and shape) of villi, whereas the experts believed previously that they had focused on the size and shape before texture. Furthermore, the analysis has confirmed that there is no single criterion that can reliably distinguish the villi.

In the future, more villi samples should be added to the current data set to improve the segmentation phase and classification results of the proposed method. The current segmentation is primarily based on colour intensity. It is believed that texture could be

used in addition to colour intensity to reduce over and under-segmentation. This research has identified textural features to be important criteria in classifying the villi. As this research demonstrated that the textural feature has a higher discriminative power than size and shape, future work would benefit from the experts exploring further textural features. The developed heuristic neural network configuration could be enhanced by integrating fuzzy systems into the hidden layers to address the issue of fuzziness of villi characteristics. The current research is based on normal placental, PHM and CHM villi; other types of abnormal villi, such as persistent gestational trophoblastic and choriocarcinoma, could be investigated to support pathologists in their complex and challenging classification of HM. Further development of the HYMAT tool should be carried out to include the classification of HM and to improve the GUI interface.

References

Abraham, B., & Box, G. E. P. (1979) Bayesian analysis of some outlier problems in time series. *Biometrika 66*, vol. 2, pp. 229–236.

Abraham, B., & Chuang, A. (1989) Outlier detection and time series modeling. *Technometrics 31*, vol. 2, pp. 241–248.

Acosta-Mesa, H. G., Zitova, B., Rios-Figueroa, H. V., Cruz-Ramirez, N., Marin-Hernandez, A., & Hernandez-Jimenez, R. (2005) Cervical cancer detection using colposcopic images: A temporal approach. *Computer Science, 2005. ENC 2005. Sixth Mexican International Conference,* pp. 158-164.

Addison, J., Wermter, S., & Macintyre, J. (1999) Effectiveness of feature extraction in neural network architectures for novelty detection. In *Proceedings of the 9th International Conference on Artificial Neural Networks*, vol. 2, pp. 976–981.

Aeyels, D. (1991) On the dynamic behaviour of the novelty detector and the novelty filter. In *Analysis of Controlled Dynamical Systems: Progress in Systems and Control Theory*.

Agarwal, D. (2005) An empirical Bayes approach to detect anomalies in dynamic multidimensional arrays. In *Proceedings of the 5th IEEE International Conference on Data Mining. IEEE Computer Society*, pp. 26-33.

Agarwal, D. (2006) Detecting anomalies in cross-classified streams: A Bayesian approach. *Journal of Knowledge and Information Systems*, vol. 11, no. 1, pp. 29-44.

Aggarwal, C. & Yu, P. (2001) Outlier detection for high dimensional data. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pp. 37-46.

Aggarwal, C. (2005) On abnormality detection in spuriously populated data streams. In *Proceedings of the 5th SIAM Data Min. Conference*, pp. 80–91.

Aggarwal, C. C. & Yu, P. S. (2008) Outlier detection with uncertain data. In *Proceedings of the International Conference on Data Mining (SDM)*, pp. 483-493.

Agrawal, R., & Srikant, R. (1995) Mining sequential patterns. In *Proceedings of the 11th International Conference on Data Engineering*, IEEE Computer Society, pp. 3–14.

Albrecht, S., Busch, J., Kloppenburg, M., Metze, F., & Tavan, P. (2000) Generalized radial basis function networks for classification and novelty detection: Self-organization of optional Bayesian decision, *Neural Netw*, vol. 13, pp. 1075–1093.

Aleskerov, E., Freisleben, B., & Rao, B. (1997) Cardwatch: A neural network based database mining system for credit card fraud detection. In *Proceedings of the IEEE Conference on Computational Intelligence for Financial Engineering*, pp. 220–226.

Allan, J., Carbonell, J., Doddington, G., Yamron, J., & Yang, Y. (1998) Topic detection and tracking pilot study. In *Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop*, pp. 194–218.

Allwin, S., Kenny, S. P. K., & Manian, V. (2010) Classification of stages of maligancies using textron signatures of a cervical cyto image. *Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference,* pp. 1-4.

Alolfe, M. A., Mohamed, W. A., Youssef, A. M., Mohamed, A. S., & Kadah, Y. M. (2009) Computer aided diagnosis in digital mammography using combined support vector machine and linear discriminant analyasis classification. *Image Processing (ICIP), 2009 16th IEEE International Conference,* pp. 2609-2612.

Alolfe, M. A., Youssef, A. M., Kadah, Y. M., & Mohamed, A. S. (2008) Computer-aided diagnostic system based on wavelet analysis for microcalcification detection in digital mammograms. *Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International,* pp. 1-5.

Amelio, A., & Pizzuti, C. (2013) Skin lesion image segmentation using a color genetic algorithm. In *Proceeding of the fifteenth annual conference companion on Genetic and evolutionary computation conference companion*, pp. 1471-1478.

Anderka, M., Klerx, T., Priesterjahn, S., & Büning, H. K. (2014) Automatic ATM Fraud Detection as a Sequence-based Anomaly Detection Problem. In *Proceedings of the 3rd International Conference on Pattern Recognition Applications and Methods* (ICPRAM'14).

Anderson, D. T., Stone, K. E., Keller, J. M., & Spain, C. J. (2012) Combination of anomaly algorithms and image features for explosive hazard detection in forward looking infrared imagery. *Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 5, no. 1, pp. 313–323.

Ando, S. (2007) Clustering needles in a haystack: An information theoretic analysis of minority and outlier detection. In *Proceedings of the 7th International Conference on Data Mining*, pp. 13–22.

Angiulli, F. & Pizzuti, C. (2002) Fast outlier detection in high dimensional spaces. In *Proceedings of the 6th European Conference on Principles of Data Mining and Knowledge Discovery*. Springer-Verlag, pp. 15–26.

Anscombe, F. J. & Guttman, I. (1960). Rejection of outliers. *Technometrics*, vol. 2, no. 2, pp. 123-147.

Arning, A., Agrawal, R., & Raghavan, P. (1996) A linear method for deviation detection in large databases. In *Proceedings of the 2nd International Conference of Knowledge Discovery and Data Mining*, pp. 164–169.

Atallah, M. J., Szpankowski, W., & Gwadera, R. (2004) Detection of significant sets of episodes in event sequences. In *Fourth IEEE International Conference on Data Mining*, pp. 3-10.

Augusteijn, M., & Folkert, B. (2002) Neural network classification and novelty detection. *Int. J. Rem. Sens*, vol. 23, no. 14, pp. 2891–2902.

Babbar, S. and Chawla, S., (2010) On Bayesian Network and Outlier Detection. In *Proceedings of the 16th International Conference on Management of Data*, p. 125.

Baker, D., Hofmann, T., Mccallum, A., & Yang, Y. (1999) A hierarchical probabilistic model for novelty detection in text. In *Proceedings of the International Conference on Machine Learning*.

Barbara, D., Couto, J., Jajodia, S., & Wu, N. (2001) Detecting novel network intrusions using Bayes estimators. In *Proceedings of the 1st SIAM International Conference on Data Mining*.

Barnett, V., & Lewis, T. (1994) *Outliers in statistical data*, New York: Wiley.

Barson, P., Davey, N., Field, S. D. H., Frank, R. J., & Mcaskie, G. (1996) The detection of fraud in mobile phone networks. *Neural Netw. World*, vol. 6, pp. 4.

Basu, S., & Meckesheimer, M. (2007) Automatic outlier detection for time series: an application to sensor data. *Journal of Knowledge and Information Systems*, vol. 11, no. 2, pp. 137-154.

Basu, S., Bilenko, M., & Mooney, R. J. (2004) A probabilistic framework for semi-supervised clustering. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM Press, pp. 59–68.

Bejerano, G. & Yona, G. (2001) Variations on probabilistic suffix trees: statistical modeling and prediction of protein families. *Bioinformatics*, vol. 17, no. 1, pp. 23–43.

Bell, K. A., Van Deerlin, V., Addya, K., Clevenger, C. V., Van Deerlin, P. G., & Leonard, D. G. (1999) Molecular genetic testing from paraffin-embedded tissue distinguishes nonmolar hydropic abortion from hydatidiform mole. *Molecular Diagnosis*, vol. 4, no. 1, pp. 11-19.

Benirschke, K., Kaufmann, P., & Baergen, R., (2006) Pathology of the human placenta fifth edition. New York: Springer.

Bianco, A. M., Garcia Ben, M., Martinez, E. J., & Yohai, V. J. (2001) Outlier detection in regression models with arima errors using robust estimates. *Journal of Forecasting*, vol. 20, no. 8, pp. 565-579.

Bishop, C. (1994) Novelty detection and neural network validation. In *Proceedings of the IEEE Conference on Vision, Image and Signal Processing*, vol. 141, pp. 217–222.

Blake, C., & Merz, C. J. (1998) {UCI} Repository of machine learning databases. *http://www.ics.uci.edu/mlearn/MLRepository.html, University of California, Irvine, Dept. of Information and Computer Sciences*.

Blender, R., Fraedrich, K., & Lunkeit, F. (1997) Identification of cyclone-track regimes in the north atlantic. *Quart. J. Royal Meteor. Soc.*, vol. 123, no. 539, pp. 727–741.

Blum, A. (1992) *Neural networks in C++*. New York: Wiley.

Bolton, R. & Hand, D. (1999) Unsupervised profiling methods for fraud detection. In *Proceedings of the Conference on Credit Scoring and Credit Control VII*, pp. 235-255.

Boquete, L., Ortega, S., Miguel-Jiménez, J. M., Rodríguez-Ascariz, J. M., & Blanco, R. (2012) Automated detection of breast cancer in thermal infrared images, based on independent component analysis. *Journal of Medical Systems*, vol. 36, no. 1, pp. 103-111.

Boriah, S., Chandola, V., & Kumar, V. (2008) Similarity measures for categorical data: A comparative evaluation. In *Proceedings of the 8th SIAM International Conference on Data Mining*, pp. 243–254.

Borisyuk, R., Denham, M., Hoppensteadt, F., Kazanovich, Y., & Vinogradova, O. (2000) An oscillatory neural network model of sparse distributed memory and novelty detection. *Biosystems*, vol. 58, pp. 265–272.

Bouckaert, R. R. (2003) Choosing between two learning algorithms based on calibrated tests. In *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, pp. 51-58.

Box, G. E., & Tiao, G. C. (1968) A Bayesian approach to some outlier problems. *Biometrika*, vol. 55, no. 1, pp. 119-129.

Branch, J., Szymanski, B., Giannella, C., Wolff, R., & Kargupta, H. (2006) In-network outlier detection in wireless sensor networks. In *Proceedings of the 26th IEEE International Conference on Distributed Computing Systems*, p. 51.

Brause, R., Langsdorf, T., & Hepp, M. (1999) Neural data mining for credit card fraud detection. In *Proceedings of the IEEE International Conference on Tools with Artificial Intelligence*, pp. 103–106.

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (1999) Optics-of: Identifying local outliers. In *Proceedings of the 3rd European Conference on Principles of Data Mining and Knowledge Discovery*. Springer-Verlag, pp. 262–270.

Breunig,M. M.,Kriegel, H.-P.,Ng, R. T., & Sander, J. (2000) LOF: Identifying density-based local outliers. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*. ACM Press, pp. 93–104.

Brockett, P. L., Xia, X., & Derrig, R. A. (1998) Using Kohonen's self-organizing feature map to uncover automobile bodily injury claims fraud. *Journal of Risk and Insurance*, vol. 65, no. 2, pp. 245-274.

Bronstein, A., Das, J., Duro, M., Friedrich, R., Kleyner, G., Mueller, M., Singhal, S., & Cohen, I. (2001) Self-aware services: Using bayesian networks for detecting anomalies in internet-based services. In *IEEE/IFIP International Symposium on Integrated Network Management Proceedings*, pp. 623-638.

Brotherton, T., & Johnson, T. (2001) Anomaly detection for advanced military aircraft using neural networks. In *Proceedings of IEEE Aerospace Conference*, vol. 6, pp. 3113-3123.

Brotherton, T., Johnson, T., & Chadderdon, G. (1998) Classification and novelty detection using linear models and a class dependent-elliptical basis function neural network. In *Proceedings of the IJCNN Conference*, vol. 2, pp. 876-879.

Budalakoti, S., Srivastava, A., Akella, R., & Turkov, E. (2006) Anomaly detection in large sets of highdimensional symbol sequences. Tech. rep. NASA TM-2006-214553, NASA Ames Research Center.

Burke, H. B., Goodman, P. H., Rosen, D. B., Henson, D. E., Weinstein, J. N., Harrell, F. E., ... & Bostwick, D. G. (1997) Artificial neural networks improve the accuracy of cancer survival prediction. *Cancer*, vol. 79, no. 4, pp. 857-862.

Byers, S. D., & Raftery, A. E. (1998) Nearest neighbour clutter removal for estimating features in spatial point processes. *J. Amer. Statis. Assoc*, vol. 93, pp. 577–584.

Byungho, H., & Sungzoon, C. (1999) Characteristics of autoassociative MLP as a novelty detector. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, Vol. 5, pp. 3086–3091.

Calderara, S., Heinemann, U., Prati, A., Cucchiara, R., & Tishby, N. (2011) Detecting anomalies in people's trajectories using spectral graph analysis. *Journal of Computer Vision and Image Understanding*, vol. 115, no. 8, pp. 1099-1111.

Caselles, V., Catté, F., Coll, T., & Dibos, F. (1993) A geometric model for active contours in image processing. *Numerische mathematik*, vol. 66, no. 1, pp. 1-31.

Caselles, V., Kimmel, R., & Sapiro, G. (1997) Geodesic active contours. *International journal of computer vision*, vol. 22, no. 1, pp. 61-79.

Catania, C. A., Bromberg, F., & Garino, C. G. (2012) An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection. *Expert Systems with Applications*, vol. 39, no. 2, pp. 1822-1829.

Caudell, T., & Newman, D. (1993) An adaptive resonance architecture to define normality and detect novelties in time series and databases. In *Proceedings of the IEEE World Congress on Neural Networks*, pp. 166–176.

Chaddad, A., Tanougast, C., Dandache, A., Al Houseini, A., & Bouridane, A. (2011) Improving of colon cancer cells detection based on Haralick's features on segmented histopathological images. In *IEEE International Conference on Computer Applications and Industrial Electronics (ICCAIE)*, pp. 87-90.

Chakrabarti, S., Sarawagi, S., & Dom, B. (1998) Mining surprising patterns using temporal description length. In *Proceedings of the 24rd International Conference on Very Large Data Bases.Morgan Kaufmann*, vol. 98, pp. 606-617.

Chandola, V., Banerjee, A., & Kumar, V. (2009) Anomaly detection: A survey. *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, pp. 1-58.

Chandola, V., Boriah, S., & Kumar, V. (2008) Understanding categorical similarity measures for outlier detection. *Tech. rep. 08-008, University of Minnesota.*

Chang H., Han J., Borowsky A., Loss L., Gray JW., Spellman PT., &, Parvin, B. (2013) Invariant Delineation of Nuclear Architecture in Glioblastoma Multiforme for Clinical and Molecular Association. *IEEE Transaction on Medical Imaging*, vol. 32, no. 4, pp. 670-682.

Chang, C. W., Lin, M. Y., Harn, H. J., Harn, Y. C., Chen, C. H., Tsai, K. H., & Hwang, C. H. (2009) Automatic segmentation of abnormal cell nuclei from microscopic image analysis for cervical cancer screening. In *IEEE International Conference on Nano/Molecular Medicine and Engineering (NANOMED),* pp. 77-80.

Chatzigiannakis, V., Papavassiliou, S., Grammatikou, M., & Maglaris, B. (2006) Hierarchical anomaly detection in distributed large-scale sensor networks. In *Proceedings of the 11th IEEE Symposium on Computers and Communications (ISCC)*, IEEE Computer Society, pp. 761–767.

Chaudhary, A., Szalay, A. S., & Moore, A. W. (2002) Very fast outlier detection in large multidimensional data sets. In *Proceedings of the ACM SIGMOD Workshop in Research Issues in Data Mining and Knowledge Discovery (DMKD)*.

Chawla, S., & Sun, P. (2006) SLOM: a new measure for local spatial outliers. *Journal of Knowledge and Information Systems*, vol. 9, no. 4, pp. 412-429.

Chen, D., Shao, X., Hu, B., & Su, Q. (2005) Simultaneous wavelength selection and outlier detection in multivariate regression of near-infrared spectra. *Anal. Sci*, Vol. 21, no. 2, pp. 161–167.

Choraś, M., Saganowski, Ł., Renk, R., & Hołubowicz, W. (2012) Statistical and signal-based network traffic recognition for anomaly detection. *Journal of Expert Systems*, vol. 29, no. 3, pp. 232-245.

Chow, C. & Yeung, D.-Y. (2002) Parzen-window network intrusion detectors. In *Proceedings of the 16th International Conference on Pattern Recognition*, *IEEE Computer Society*, vol. 4, pp. 385–388.

Chui-Mei, T., Tai-Lang, J., & Chi-Wen, H. (2008) Self organizing map neural network with fuzzy screening for micro-calcifications detection on mammograms. *Soft Computing in Industrial Applications, 2008. SMCia '08. IEEE Conference,* pp. 421-425.

Churilov, L., Bagirov, A., Schwartz, D., Smith, K., & Dally, M. (2005) Data mining with combined use of optimization techniques and self-organizing maps for improving risk grouping rules: application to prostate cancer patients. *Journal of Management Information Systems*, vol. 21, no. 4, pp. 85-100.

Clarke, L., P., Qian, W., & Li, L. (1998) Computer-Assisted Method and Apparatus for Analysis of x-Ray Images using Wavelet Transforms. U.S. Pat. 5,799,100.

Crisan, D. A., Dobrbscu, R., & Planinsic, P. (2007) Mammographic lesions discrimination based on fractal dimension as an indicator. *Systems, Signals and Image Processing,*

2007 and 6th EURASIP Conference Focused on Speech and Image Processing, Multimedia Communications and Services. 14th International Workshop, pp. 74-77.

Crook, P. A., Marsland, S., Hayes, G., & Nehmzow, U. (2002) A tale of two filters: Online novelty detection. In *Proceedings of the International Conference on Robotics and Automation*, pp. 3894–3899.

Crook, P., & Hayes, G. (2001) A robot implementation of a biologically inspired method for novelty detection. In *Proceedings of the Towards Intelligent Mobile Robots Conference.*

Crotty, M. (1998) The foundation of social research: meaning and perspective in the research process. *London*, Sage publications.

Cui, Y., Jin, J. S., Park, M., Luo, S., Xu, M., Peng, Y., & Santos, L. D. (2010) Computer aided abnormality detection for microscopy images of cervical tissue. In *International Conference on Complex Medical Engineering (CME)*, pp. 63-68.

Cun, Y. L., Boser, B., Denker, J. S., Howard, R. E., Hubbard, W., Jackel, L. D., & Henderson, D. (1990) Handwritten digit recognition with a back-propagation network. In *Advances in Neural Information Processing Systems*, pp. 396–404.

Curiac, D. I., & Volosencu, C. (2012) Ensemble based sensing anomaly detection in wireless sensor networks. *Journal of Expert Systems with Applications*, vol. 39, no. 10, pp. 9087-9096.

Dasgupta, D., & Nino, F. (2000) A comparison of negative and positive selection algorithms in novel pattern detection. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, vol. 1, pp. 125–130.

Davy, M., & Godsill, S. (2002) Detection of abrupt spectral changes using support vector machines, an application to audio signal segmentation. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing.*

Deepa, S., & Bharathi, V. S. (2012) An efficient digital mammogram image classification using DTCWT and SVM. In *Proceedings of the Second International Conference on Computational Science, Engineering and Information Technology,* pp. 288-293.

Deepak, K. S., Rai, H. G., Syed, S., & Krishna, P. R. (2012) Texture edge statistics for efficient retrieval of biomedical images. In *Proceedings of the 5th ACM COMPUTE Conference: Intelligent & scalable system technologies*, pp. 1-6.

Demir, C., & Yener, B. (2005). Automated cancer diagnosis based onhistopathological images: A systematic survey, *Rensselaer Polytechnic Institute, Department of Computer Science*.

Denning, D. E. (1987) An intrusion detection model. *IEEE Trans. Softw. Eng.*, vol. 13, no. 2, pp. 222–232.

Dereszynski, E. & T. Dietterich (2007) Probabilistic models for anomaly detection in remote sensor data streams. In *23rd Conference on Uncertainty in Artificial Intelligence (UAI-2007)*. pp. 75–82.

Desforges, M., Jacob, P., & Cooper, J. (1998) Applications of probability density estimation to the detection of abnormal conditions in engineering. In *Proceedings of the Institute of the Mechanical Engineers*, vol. 212, pp. 687-703.

Deshpande, D. S., Rajurkar, A. M., & Manthalkar, R. M. (2013) Medical image analysis an attempt for mammogram classification using texture based association rule mining. In *IEEE Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, pp. 1-5.

Dhawan, A. P. (2003). *Medical image analysis*. London: Wiley-IEEE Press.

Dhinagar, N. J., & Celenk, M. (2011) Non-invasive detection and classification of skin cancer from visual and cross-sectional images. In *Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*, pp. 1-7.

Diaz, I., & Hollmen, J. (2002) Residual generation and visualization for understanding novel process conditions. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, *IEEE*, pp. 2070–2075.

Diehl, C. & Hampshire, J. (2002) Real-time object classification and novelty detection for collaborative video surveillance. In *Proceedings of the IEEE International Joint Conference on Neural Networks, v*ol. 3, pp. 2620-2625.

Donoho, S. (2004) Early detection of insider trading in option markets. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 420–429.

Dorronsoro, J. R., Ginel, F., Sanchez, C., & Cruz, C. S. (1997) Neural fraud detection in credit card operations. *IEEE Trans. Neural Netw*, vol. 8, no. 4, pp. 827–834.

Doukas, C. N., & Maglogiannis, I. (2007) Automated cell apoptosis characterization using active contours. In *Proceedings of 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* pp. 812-815.

Doyle, S., Agner, S., Madabhushi, A., Feldman, M., & Tomaszewski, J. (2008) Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features. In *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 496-499.

Doyle, S., Feldman, M., Tomaszewski, J., & Madabhushi, A. (2012) A boosted bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. In *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 5, pp. 1205-1218.

Doyle, S., Hwang, M., Shah, K., Madabhushi, A., Feldman, M., & Tomaszeweski, J. (2007) Automated Grading Of Prostate Cancer Using Architectural And Textural Image Features. *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium,* pp. 1284-1287.

Du, W., Fang, L., & Peng, N. (2006) Lad: Localization anomaly detection for wireless sensor networks. *J. Paral. Distrib. Comput*, vol. 66, no. 7, pp. 874–886.

Dunning, T., & Friedman, E. (2014) *Practical Machine Learning: A New Look at Anomaly Detection*. California, USA: O'Reilly Media, Inc.

Dutta, H., Giannella, C., Borne, K., & Kargupta, H. (2007) Distributed top-k outlier detection in astronomy catalogs using the DEMAC system. In *Proceedings of the 7th SIAM International Conference on Data Mining*, pp. pp. 473-478.

Elizabeth, D. S., Nehemiah, H. K., Raj, C. R., & Kannan, A. (2012a) Computer-aided diagnosis of lung cancer based on analysis of the significant slice of chest computed tomography image. *IET image processing*, vol. 6, no. 6, pp. 697-705.

Elizabeth, D. S., Nehemiah, H. K., Raj, C., & Kannan, A. (2012b) A novel segmentation approach for improving diagnostic accuracy of CAD systems for detecting lung cancer from chest computed tomography images. *Journal of Data and Information Quality (JDIQ)*, vol. 3, no. 2, pp. 1-16.

Emamian, V., Kaveh, M., & Tewfik, A. (2000) Robust clustering of acoustic emission signals using the Kohonen network. In *Proceedings of the IEEE International Conference of Acoustics, Speech and Signal Processing*, vol. 6, pp. 3891–3894.

Endler, D. (1998) Intrusion detection: Applying machine learning to solaris audit data. In *Proceedings of the 14th Annual Computer Security Applications Conference. IEEE Computer Society*, pp. 268-279.

Ertoz, L., Steinbach, M., & Kumar, V. (2003) Finding topics in collections of documents: A shared nearest neighbour approach. In *Clustering and Information Retrieval*, pp. 83–104.

Esgiar, A. N., & Chakravorty, P. K. (2007) Fractal based classification of colon cancer tissue images. *Signal Processing and its Applications, 2007. ISSPA 2007. 9th International Symposium,* pp. 1-4.

Eskin, E. (2000) Anomaly detection over noisy data using learned probability distributions. In *Proceedings of the 17th International Conference on Machine Learning*, pp. 255–262.

Eskin, E., Arnold, A., Prerau, M., Portnoy, L., & Stolfo, S. (2002) Ageometric framework for unsupervised anomaly detection. In *Proceedings of the Conference on Applications of Data Mining in Computer Security*, pp. 78–100.

Eskin, E., Lee,W., & Stolfo, S. (2001) Modeling system call for intrusion detection using dynamic window sizes. In *Proceedings of DARPA Information Survivability Conference and Exposition (DISCEX)*.

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, pp. 226–231.

Fawcett, T., & Provost, F. (1999) Activity monitoring: noticing interesting changes in behavior. In *Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 53–62.

Ferrero, G., Britos, P., García-Martínez, R., (2006) in IFIP International Federation for Information Processing, Volume 218, Professional Practice in Artificial Intelligence, eds. J. Debenham, (Boston: Springer), pp. 1-10.

Filipczuk, P., Fevens, T., Krzyzak, A., & Monczak, R. (2013) Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. In *IEEE Transactions on Medical Imaging*, vol. 32, no. 12, pp. 2169-2178.

Forrest, S., Warrender, C., & Pearlmutter, B. (1999) Detecting intrusions using system calls: Alternate data models. In *Proceedings of the IEEE ISRSP*, pp. 133–145.

Fox, A. J. (1972) Outliers in time series. *J. Royal Statis. Soc. Series B*, vol. 34, no. 3, pp. 350–363.

Fujimaki, R., Yairi, T., & Machida, K. (2005) An approach to spacecraft anomaly detection problem using kernel feature space. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pp. 401-410.

Gaber, C., Hemery, B., Achemlal, M., Pasquet, M., & Urien, P. (2013) Synthetic logs generator for fraud detection in mobile transfer services. In *IEEE International Conference on Collaboration Technologies and Systems (CTS)*, pp. 174-179).

Galeano, P., Pea, D., & Tsay, R. S. (2004) Outlier detection in multivariate time series via projection pursuit. *Statistics and econometrics working articles ws044211*, Departamento de Estad¨ıstica y Econometr¨ıca, Universidad Carlos III.

Ganeshan, B., Skogen, K., Pressney, I., Coutroubis, D., & Miles, K. (2012). Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: Preliminary evidence of an association with tumour metabolism, stage, and survival. *Clinical Radiology,* vol. 67, no. 2, pp. 157-164.

Gavrilovic, M., Azar, J., Lindblad, J., Wahlby, C., Bengtsson, E., Busch, C., & Carlbom, I. (2013) Blind Color Decomposition of Histological Images. *IEEE Transactions on Medical Imaging*, vo. 32, no. 6, pp. 983 – 994.

Genest, D. R. (2001) Partial hydatidiform mole: clinicopathological features, differential diagnosis, ploidy and molecular studies, and gold standards for diagnosis. *International Journal of Gynecologic Pathology*, vol. 20, no. 4, pp. 315-322.

George, Y. M., Zayed, H. H., Roushdy, M. I., & Elbagoury, B. M. (2014) Remote Computer-Aided Breast Cancer Detection and Diagnosis System Based on Cytological Images. *IEEE Systems Journal*, vol. 8, no. 3, pp. 949-964.

Ghosh, A. K., Schwartzbard, A., & Schatz, M. (1999) Learning program behavior profiles for intrusion detection. In *Proceedings of the 1st USENIX Workshop on Intrusion Detection and Network Monitoring*, pp. 51–62.

Ghosh, A. K.,Wanken, J., & Charron, F. (1998) Detecting anomalous and unknown intrusions against programs. In *Proceedings of the 14th Annual Computer Security Applications Conference*, pp. 259-267.

Ghosh, S., & Reilly, D. L. (1994) Credit card fraud detection with a neural-network. In *Proceedings of the 27th Annual Hawaii International Conference on System Science*, vol. 3, pp. 621-630.

Gisselson, D. (2001) Chromosomal instability in Cancer: Causes and Consequences. Atlas of Genetics and Cytogenetics in Ontology and Haematology, May 2001.

Goldberger, A. L., Amaral, L. A. N., Glass, L.,Hausdorff, J. M., Ivanov, P.C.,Mark, R. G.,Mietus, J. E.,Moody, G. B., Peng, C.-K., & Stanley, H. E. (2000) Physiobank, physiotoolkit, and physionet *components of a new research resource for complex physiologic signals*, ,*Circulation*, vol. 101, no. 23, pp. e215-e220.

Goldstein, M., & Dengel, A. (2012). Histogram-based outlier score (hbos): A fast unsupervised anomaly detection algorithm. In *Proceedings of the 35th German Conference on Artificial Intelligence*, pp. 59-63.

Gonzalez, R. C., & Woods, R. E. (2002) *Digital image processing* (2nd ed). New Jersey, USA: Prentice Hall.

Görnitz, N., Kloft, M. M., Rieck, K., & Brefeld, U. (2013) Toward supervised anomaly detection. *Journal of Artificial Intelligence Research*, vol. 46, pp. 235-262

Grubbs, F. (1969) Procedures for detecting outlying observations in samples. *Technometrics*, vol. 11, no. 1, pp. 1–21.

Gruber, T.R. (1993) Towards Principles for the Design of Ontologies Used for Knowledge Sharing. In *Roberto Poli Nicola Guarino, editor, International Workshop on Formal Ontology, Padova, Italy, 1993, Technical report KSL-93-04, Knowledge Systems Laboratory, Stanford University.*

Guha, S., Rastogi, R., & Shim, K. (2000) ROCK: A robust clustering algorithm for categorical attributes. *Inform. Syst.*, vol. 25, no. 5, pp. 345–366.

Gul, T., Yilmaztürk, A., & Erden, A. C., (1997) A review of trophoblastic diseases at the medical school of Dicle University. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, vol. 74, pp. 37–40.

Guttormsson, S. E., Marks, R. J., El-Sharkawi, M. A., & Kerszenbaum, I. (1999) Elliptical novelty grouping for on-line short-turn detection of excited running rotors. *IEEE Transactions on Energy Conversion*, vol. 14, no. 1, pp. 16-22.

Guyon, I., & Elisseeff, A. (2003) An introduction to variable and feature selection. *The Journal of Machine Learning Research*, vol. 3, pp. 1157-1182.

Gwadera, R., Atallah, M. J., & Szpankowski, W. (2005a) Markov models for identification of significant episodes. In Proceedings of the 5th SIAM International Conference on Data Mining, pp. 404-414.

Gwadera, R., Atallah, M. J., & Szpankowski, W. (2005b) Reliable detection of episodes in event sequences. *Journal of Knowledge and Information Systems*, vo. 7, no. 4, pp. 415-437.

Hadavi, N., Nordin, M., & Shojaeipour, A. (2014) Lung cancer diagnosis using CT-scan images based on cellular learning automata. In *IEEE International Conference on Computer and Information Sciences (ICCOINS)*, pp. 1-5.

Hamarneh, G., & Li, X. (2009) Watershed segmentation using prior shape and appearance knowledge. *Image and Vision Computing*, vol. 27, no. 1, pp. 59-68.

Hamdi, N., Auhmani, K., & Hassani, M. M. (2008) Design of a high-accuracy classifier based on fisher discriminant analysis: Application to computer-aided diagnosis of microcalcifications. *Computational Sciences and its Applications, 2008. ICCSA '08. International Conference,* pp. 267-273.

Han, S.M., Lee, H.J. & Choi, J.Y. (2007) Prostate Cancer Detection using Texture and Clinical Features in Ultrasound Image. In *Information Acquisition, 2007. ICIA '07. International Conference on*. pp. 547-552.

Haneishi, H., Ue, H., Takita, N., Toyama, H., Miyamoto, T., Yamamoto, N., & Mori, Y. (2001) Lung image segmentation and registration for quantitative image analysis. In *IEEE Nuclear Science Symposium Conference Record*, vol. 3, pp. 1390-1393.

Haraldsson, H., Edenbrandt, L., & Ohlsson, M. (2004). Detecting acute myocardial infarction in the 12-lead ECG using Hermite expansions and neural networks. *Artificial Intelligence in Medicine*, vol. 32, no. 2, pp. 127-136.

Harris, T. (1993) Neural network in machine health monitoring. *Professional Engineering*.

Hart, A. (1985). Knowledge elicitation: issues and methods. *Computer-aided design*, vol. 17, no. 9, pp. 455-462.

Hauskrecht, M., Valko, M., Kveton, B., Visweswaran, S., & Cooper, G. F. (2007) Evidence-based Anomaly Detection in Clinical Domains. In *Proceedings of AMIA Annual Symposium*, pp. 319–323.

Hautamaki, V., Karkkainen, I., & Franti, P. (2004) Outlier detection using k-nearest neighbour graph. In *Proceedings of the 17th International Conference on Pattern Recognition*, vol. 3, pp. 430–433.

Hawkins, S., He, H., Williams, G. J., & Baxter, R. A. (2002) Outlier detection using replicator neural networks. In *Proceedings of the 4th International Conference on Data Warehousing and Knowledge Discovery*, pp. 170–180.

144

Hazel, G. G. (2000) Multivariate Gaussian MRF for multi-spectral scene segmentation and anomaly detection. *GeoRS*, vol. 38, no. 3, pp. 1199–1211.

He, Z., Deng, S., Xu, X., & Huang, J. Z. (2006) A fast greedy algorithm for outlier mining. In *Proceedings of the 10th Pacific-Asia Conference on Knowledge and Data Discovery*, pp. 567–576.

He, Z., Huang, J. Z., Xu, X., & Deng, S. (2004) Mining class outliers: Concepts, algorithms and applications. In *Advances in Web-Age Information Management*, pp. 589-599.

He, Z., Xu, X., & Deng, S. (2003) Discovering cluster-based local outliers. *Pattern Recognition Letters*, vol. 24, no. 9, pp. 1641-1650.

He, Z., Xu, X., & Deng, S. (2005) An optimization model for outlier detection in categorical data. In *Proceedings of the International Conference on Intelligent Computing*. Lecture Notes in Computer Science, vol. 3644, pp 400-409.

Hickinbotham, S. J., & Austin, J. (2000a) Novelty detection in airframe strain data. In *Proceedings of the 15th International Conference on Pattern Recognition*, vol. 2, pp. 536–539.

Hickinbotham, S. J., & Austin, J. (2000b) Novelty detection in airframe strain data. In *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, vol. 6, pp. 24–27.

Hinds, P. J., & Pfeffer, J. (2003) Why organizations don't "know what they know": Cognitive and motivational factors affecting the transfer of expertise. *Sharing expertise: Beyond knowledge management*, pp. 3-26.

Ho, L. L., Macey, C. J., & Hiller, R. (1999) A distributed and reliable platform for adaptive anomaly detection in ip networks. In *Active Technologies for Network and Service Management*, pp. 33-46.

Ho, T. V., & Rouat, J. (1997) A novelty detector using a network of integrate and fire neurons. Lecture Notes in Computer Science, vol. 1327, pp. 103–108.

Ho, T. V., & Rouat, J. (1998) Novelty detection based on relaxation time of a network of integrate-and-fire neurons. In *Proceedings of the 2nd IEEE World Congress on Computational Intelligence*, pp. 1524–1529.

Hodge, V. J., & Austin, J. (2004) A survey of outlier detection methodologies. *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85-126.

Hoffman, R. R. (1987) The problem of extracting the knowledge of experts from the perspective of experimental psychology. *AI Magazine*, vol. 8, no. 2, pp. 53-66.

Hollier, G., & Austin, J. (2002) Novelty detection for strain-gauge degradation using maximally correlated components. In *Proceedings of the European Symposium on Artificial Neural Networks*, pp 257–262.

Horn, P. S., Feng, L., Li, Y., & Pesce, A. J. (2001) Effect of outliers and nonhealthy individuals on reference interval estimation. *Clinical Chem*, vol. 47, no. 12, pp. 2137–2145.

Howat, A. J., Beck, S., Fox, H., Harris, S. C., Hill, A. S., Nicholson, C. M., & Williams, R. A. (1993) Can histopathologists reliably diagnose molar pregnancy? *J Clin Pathol*, vol. 46, no. 7, pp. 599–602.

Hu, W., Liao, Y., & Vemuri, V. R. (2003) Robust anomaly detection using support vector machines. In *Proceedings of the International Conference on Machine Learning*. Morgan Kaufmann Publishers Inc., pp. 282–289.

Huber, P. J. (1974). *Robust statistics*. Wiley, New York.

Hui, M., Wenxue, H., Jialin, S., & Liqiang, W. (2008) Feature extraction and analysis of ovarian cancer proteomic mass spectra. *Bioinformatics and Biomedical Engineering, 2008. ICBBE 2008. the 2nd International Conference,* pp. 668-671.

Ide, T. & Kashima, H. (2004) Eigenspace-based anomaly detection in computer systems. In *Proceedings of the 10th ACMSIGKDD International Conference on Knowledge Discovery and Data Mining*, pp.440–449.

Ilgun, K., Kemmerer, R. A., & Porras, P. A. (1995) State transition analysis: A rule-based intrusion detection approach. *IEEE Transaction on Software Engineering*, vol. 21, no. 3, pp. 181–199.

Jagota, A. (1991) Novelty detection on a very large number of memories stored in a hopfield-style network. In *Proceedings of the International Joint Conference on Neural Networks*, vol. 2, p. 905.

Jakubek, S., & Strasser, T. (2002) Fault-diagnosis using neural networks with ellipsoidal basis functions. In *Proceedings of the American Control Conference*, vol. 5, pp. 3846–3851.

Janakiram, D., Reddy, V., & Kumar, A. (2006) Outlier detection in wireless sensor networks using Bayesian belief networks. In *Proceedings of the 1st International Conference on Communication System Software and Middleware*, pp. 1–6.

Japkowicz, N., Myers, C., & Gluck, M. A. (1995) A novelty detection approach to classification. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 518–523.

Jayadevappa, D., Srinivas Kumar, S., & Murty, D. S. (2011) Medical Image Segmentation Algorithms using Deformable Models: A Review. *IETE Technical Review*, vol. 28, no. 3, pp. 248-255.

Jeong, C. Y., Chang, B. H., & Na, J. C. (2010) A hierarchical approach to traffic anomaly detection using image processing technique. In *IEEE Sixth International Conference on Networked Computing and Advanced Information Management (NCM)*, pp. 592-594.

Jin, X., & Chow, T. W. (2013) Anomaly detection of cooling fan and fault classification of induction motor using Mahalanobis–Taguchi system. *International Journal of Expert Systems with Applications*, vol. 40, no. 15, pp. 5787-5795.

Jolliffe, I. (2002) *Principal component analysis*. New York, John Wiley & Sons, Ltd.

Kadota, K., Tominaga, D., Akiyama, Y., & Takahashi, K. (2003) Detecting outlying samples in micro-array data: A critical assessment of the effect of outliers on sample classification. *Chem-Bio Informatics*, vol. 3, no. 1, pp. 30–45.

Kai, Z., Tian-fu, W., Jiang-li L., & De-yu, L. (2007) Recognition of breast ultrasound images using A hybrid method. *Complex Medical Engineering, 2007. CME 2007. IEEE/ICME International Conference,* pp. 640-643.

Kang, D., Shin, S. Y., Sung, C. O., Kim, J. Y., Pack, J. K., & Choi, H. D. (2011) An improved method of breast MRI segmentation with Simplified K-means clustered images. In *Proceedings of the 2011 ACM Symposium on Research in Applied Computation*, pp. 226-231.

Kang, H., Pinti, A., & Taleb–Ahmed, A. (2013) Automatic tissue classification by integrating medical expert anatomic ontologies. *International Journal of Advanced Operations Management*, vol. 5, no. 1, pp. 3-13.

Karl, K. (2004) A researcher's dilemma - philosophical and methodological pluralism. *Electronic journal of business research methods*, vol. 2, no. 2, pp. 119-128.

Karnan, M., & Gandhi, K. R. (2010) Diagnose breast cancer through mammograms, using image processing techniques and optimization techniques. *Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference,* pp. 1-4.

Karsoliya, S. (2012) Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. *International Journal of Engineering Trends and Technology*, vol. 3, no. 6, pp. 713-717.

Kazmar, T., Smid, M., Fuchs, M., Luber, B., & Mattes, J. (2010) Learning cellular texture features in microscopic cancer cell images for automated cell-detection. *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE,* pp. 49-52.

Kekre, H. B., Sarode, M. T. K., & Gharge, M. S. M. (2009) Detection and Demarcation of Tumor using Vector Quantization in MRI images. *International Journal of Engineering Science and Technology*, vol.1, no.2, pp. 59-66.

Keogh, E., Lonardi, S., & Ratanamahatana, C. A. (2004) Towards parameter-free data mining. In *Proceedings of the 10th ACMSIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM Press, pp. 206–215.

Khaskheli, M., Khushk, I. A., Baloch, S., & Shah, H. (2007) Gestational trophoblastic disease: experience at a tertiary care hospital of Sindh. *J Coll Physicians Surg Pak*, vol. 17, pp. 81–3.

Khazai, S., Safari, A., Mojaradi, B., & Homayouni, S. (2011) A fast-adaptive support vector method for full-pixel anomaly detection in hyperspectral images. *Geoscience and Remote Sensing Symposium (IGARSS),* pp. 1763–1766.

Khazai, S., Safari, A., Mojaradi, B., & Homayouni, S. (2013) An approach for subpixel anomaly detection in hyperspectral images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 769-778.

Khorchani, B., Hallé, S., & Villemaire, R. (2012) Firewall anomaly detection with a model checker for visibility logic. In *IEEE Conference on Network Operations and Management Symposium (NOMS),* pp. 466-469.

Khreich, W., Granger, E., Miri, A., & Sabourin, R. (2012) Adaptive ROC-based ensembles of HMMs applied to anomaly detection. *Pattern Recognition*, vol. 45, no. 1, pp. 208-230.

Kim, H. I., Shin, S., Wang, W., & Jeon, S. I. (2013) SVM-based Harris corner detection for breast mammogram image normal/abnormal classification. In *Proceedings of the 2013 Research in Adaptive and Convergent Systems*, pp. 187-191.

Kim, K. R., Park, B. H., Hong, Y. O., Kwon, H. C., & Robboy, S. J. (2009) The villous stromal constituents of complete hydatidiform mole differ histologically in very early

pregnancy from the normally developing placenta. *Am J Surg Pathol*, vol. 33, pp. 176–85.

King, S., King, D., P., Anuzis, K. A., Tarassenko, L., Hayton, P., & Utete, S. (2002) The use of novelty detection techniques for monitoring high-integrity plant. In *Proceedings of the International Conference on Control Applications*, vol. 1, pp. 221–226.

Knorr, E. M., Ng, R. T., & Tucakov, V. (2000) Distance-based outliers: algorithms and applications. *The VLDB Journal—The International Journal on Very Large Data Bases*, vol. 8, no. 3-4, pp. 237-253.

Ko, H., & Jacyna, G. (2000) Dynamical behavior of autoassociative memory performing novelty filtering. In *IEEE Trans. Neural Netw*, vol. 11, pp. 1152–1161.

Koen, B. V. (1985) *Definition of the Engineering Method*. Washington DC, USA, ASEE Publications.

Kohonen, T., Ed. (1997) *Self-Organizing Maps*. New York, USA, Springer-Verlag.

Kojima, K., & Ito, K. (1999) Autonomous learning of novel patterns by utilizing chaotic dynamics. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, vol. 1, pp. 284–289.

Kou, Y., Lu, C. T., & Chen, D. (2006) Spatial Weighted Outlier Detection. In *Proceedings of the SIAM Conference on Data Mining*, pp. 614–618.

Kounelakis, M., Zervakis, M., Giakos, G., Postma, G., Buydens, L., & Kotsiakis, X. (2012) On the relevance of glycolysis process on brain gliomas. *Biomedical and Health Informatics*, vol. 17, no. 1, pp. 128 – 135.

Kruegel, C., & Vigna, G. (2003) Anomaly detection of web-based attacks. In *Proceedings of the 10th ACM conference on Computer and communications security*, pp. 251-261.

Krügel, C., Toth, T., & Kirda, E. (2002) Service specific anomaly detection for network intrusion detection. In *Proceedings of the 2002 ACM symposium on Applied computing*, pp. 201-208.

Kumar, V. (2005) Parallel and distributed computing for cybersecurity. *IEEE Distributed Systems Online*. vol. 6, no. 10.

Kuruganti, P. T., & Hairong Qi. (2002) Asymmetry analysis in breast cancer detection using thermal infrared images. *Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint,* vol. 2, pp. 1155-1156.

Labib, K. & Vemuri, R. (2002) NSOM: A real-time network-based intrusion detection using self-organizing maps. *Networks and Security,* pp. 1-6.

Lahmiri, S., & Boukadoum, M. (2011) Hybrid discrete wavelet transform and gabor filter banks processing for mammogram features extraction. *New Circuits and Systems Conference (NEWCAS), 2011 IEEE 9th International,* pp. 53-56.

Lakhina, A., Crovella, M., & Diot, C. (2005) Mining anomalies using traffic feature distributions. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, pp. 217–228.

Landolsi, H., Missaoui, N., Yacoubi, M. T., Trabelsi, A., Rammeh-Rommani, S., Hidar, S., Gribaa, M., Hmissa, S., & Mokni, M. (2009) Assessment of the role of histopathology and DNA image analysis in the diagnosis of molar and non-molar abortion: a study of 89 cases in the center of Tunisia. *Pathol Res Pract*, vol. 205, pp.789–96.

Lauer, M. (2001) A mixture approach to novelty detection using training data with outliers. In *Proceedings of the 12th European Conference on Machine Learning*, pp. 300–311.

Laurikkala, J., Juhola, M., & Kentala, E. (2000) Informal identification of outliers in medical data. In *Proceedings of the 5th International Workshop on Intelligent Data Analysis in Medicine and Pharmacology*, pp. 20–24.

Lee, W., & Xiang, D. (2001) Information-theoretic measures for anomaly detection. In *Proceedings of the IEEE Symposium on Security and Privacy*, pp. 130–143.

Lee, W., Stolfo, S. J., & Mok, K. W. (2000) Adaptive intrusion detection: A data mining approach. *Artificial Intelligence Review*, vol. 14, no. 6, pp. 533-567.

Lee, Y. J., Yeh, Y. R., & Wang, Y. C. F. (2013) Anomaly detection via online oversampling principal component analysis. In *IEEE Transactions on Knowledgeon Knowledge and Data Engineering,* vol. 25, no. 7, pp. 1460-1470.

Li, C., Xu, C., Gui, C., & Fox, M. D. (2005) Level set evolution without re-initialization: a new variational formulation. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 430-436.

Li, M., & Vitanyi, P.M.B. (1993) *An Introduction to Kolmogorov Complexity and Its Applications*, Springer-Verlag.

Li, W., Mahadevan, V., & Vasconcelos, N. (2014) Anomaly detection and localization in crowded scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no.1, pp. 18-32.

Li, Y., Pont, M. J., & Jones, N. B. (2002) improving the performance of radial basis function classifiers in condition monitoring and fault diagnosis applications where unknown faults may occur. *Patt. Recog. Lett.*, vol. 23, no. 5, pp. 569–577.

Li, Z., Shin, S., Jeon, S. I., Son, S. H., & Pack, J. K. (2012a) A new histogram-based breast cancer image classifier using Gaussian mixture model. *In Proceedings of the 2012 ACM Research in Applied Computation Symposium*, pp. 143-147.

Li, Z., Wang, W., Shin, S., & Choi, H. D. (2013). Enhanced roughness index for breast cancer benign/malignant measurement using Gaussian mixture model. In *Proceedings of the 2013 Research in Adaptive and Convergent Systems*, pp. 177-181.

Li, Z., Xu, W., Huang, A., & Sarrafzadeh, M. (2012b) Dimensionality reduction for anomaly detection in electrocardiography: A manifold approach. In *IEEE Ninth International Conference on Wearable and Implantable Body Sensor Networks (BSN),* pp. 161-165.

Liangliang, W., Zhiyong, L., Jixiang, S., & Shilin, Z. (2010) Anomaly detection algorithm for hyperspectral images based on background endmember extraction and kernel RX algorithm. *Computer Application and System Modeling (ICCASM),* vol. 13, pp. V13-288-V13-291.

Lin, J., Keogh, E., Fu, A., & Herle, H. V. (2005) Approximations to magic: Finding unusual medical time series. In *Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems*, pp. 329–334.

Lin, S., & Brown, D. E. (2006) An outlier-based data association method for linking criminal incidents. *Journal of Decision Support Systems*, vol. 41, no. 3, pp. 604-615.

Lin, W. C., Hsu, S. C., & Cheng, A. C. (2014) Mass Detection in Digital Mammograms System Based on PSO Algorithm. In *IEEE International Symposium on Computer, Consumer and Control (IS3C)*, pp. 662-668.

Linguraru, M. G., Shijun Wang, Shah, F., Gautam, R., Peterson, J., & Linehan, W. M. (2009) Computer-aided renal cancer quantification and classification from contrast-enhanced CT via histograms of curvature-related features. *Engineering in Medicine and Biology Society*, pp. 6679-6682.

Liu, G., & Zheng, Z. (2011) Anomaly target detection algorithm based on JPEG images. In *IEEE International Conference on Multimedia Technology (ICMT)*, pp. 2952-2955.

Lu, C. T., Chen, D., & Kou, Y. (2003) Algorithms for spatial outlier detection. In *Third IEEE International Conference on Data Mining*, pp. 597-600.

Lu, Y., Gao, W., & Liu, J. (2010). Color matching for colored fiber blends based on the fuzzy c-mean cluster in HSV color space. In *Proceedings of Fuzzy Systems and Knowledge Discovery (FSKD)*, vol. 1, pp. 452-455.

Ma, J. & Perkins, S. (2003) Time-series novelty detection using one-class support vector machines. In *Proceedings of the International Joint Conference on Neural Networks*, vol. 3, pp. 1741–1745.

Ma, L., Crawford, M. M., & Tian, J. (2010) Anomaly detection for hyperspectral images using local tangent space alignment. In *IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 824-827.

Macdonald, J. W. & Ghosh, D. (2007) Copa–cancer outlier profile analysis. *Bioinformatics*, vol. 22, no. 23, pp. 2950–2951.

Mahoney, M. V., & Chan, P. K. (2002) Learning nonstationary models of normal network traffic for detecting novel attacks. In *Proceedings of the 8th ACMSIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 376–385.

Mahoney, M. V., & Chan, P. K. (2003) Learning rules for anomaly detection of hostile network traffic. In *Proceedings of the 3rd IEEE International Conference on Data Mining*, pp. 601–604.

Malladi, R., Sethian, J., & Vemuri, B. C. (1995) Shape modeling with front propagation: A level set approach. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158-175.

Manevitz, L. M., & Yousef, M. (2000) Learning from positive data for document classification using neural networks. In *Proceedings of the 2nd Bar-Ilan Workshop on Knowledge Discovery and Learning*.

Manevitz, L. M., & Yousef, M. (2002) One-class SVMS for document classification. *J. Mach. Learn. Res.*, vol. 2, pp. 139–154.

Manson, G. (2002) Identifying damage sensitive, environment insensitive features for damage detection. In *Proceedings of IES Conference*, pp. 187-197.

Manson, G., Pierce, G., & Worden, K. (2001) On the long-term stability of normal conditions for damage detection in a composite panel. In *Proceedings of the 4th International Conference on Damage Assessment of Structures*, pp. 359-369.

Manson, G., Pierce, S. G., Worden, K., Monnier, T., Guy, P., & Atherton, K. (2000) Long-term stability of normal condition data for novelty detection. In *Proceedings of the Conference on Smart Structures and Integrated Systems*, pp. 323–334.

Marcomini, K. D., & Schiabel, H. (2012) Nodules segmentation in breast ultrasound using the artificial neural network self-organizing map. In *Proceedings of the World Congress on Engineering*, vol. 2.

Markelj, P., Tomaževič, D., Likar, B., & Pernuš, F. (2012) A review of 3D/2D registration methods for image-guided interventions. *Medical image analysis*, vol. 16, no. 3, pp. 642-661.

Markou, M., & Singh, S. (2003a) Novelty detection: a review—part 1: statistical approaches. *Journal of Signal processing*, vol. 83, no. 12, pp. 2481-2497.

Markou, M., & Singh, S. (2003b) Novelty detection: a review—part 2: neural network based approaches. *Journal of Signal processing*, vol. 83, no. 12, pp. 2499-2521.

Marsland, S., Nehmzow, U., & Shapiro, J. (2000a) Novelty detection for robot neotaxis. In *Proceedings of the 2nd International Symposium on Neural Compuatation*, pp. 554–559.

Marsland, S., Nehmzow, U., & Shapiro, J. (2000b) A real-time novelty detector for a mobile robot. In *Proceedings of the EUREL Conference on Advanced Robotics Systems*.

Marsland, S., Nehmzow, U., & Shapiro, J. 1999. A model of habituation applied to mobile robots. In *Proceedings of Towards Intelligent Mobile Robots Conference. Department of Computer Science,Manchester University*, Technical report UMCS-99-3-1.

Martin, S., Troccaz, J., & Daanen, V. (2010) Automated segmentation of the prostate in 3D MR images using a probabilistic atlas and a spatially constrained deformable model. *Medical physics*, vol. 37, pp. 1579.

Martinelli, G., & Perfetti, R. (1994) Generalized cellular neural network for novelty detection. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications,* vol. 41, no. 2, pp. 187-190.

Martinez, D. (1998) Neural tree density estimation for novelty detection. IEEE Transaction on. Neural Networks, vol. 9, no. 2, pp. 330–338.

Mascaro, S., Nicholso, A. E., & Korb, K. B. (2014) Anomaly detection in vessel tracks using Bayesian networks. *International Journal of Approximate Reasoning*, vol. 55, no. 1, pp. 84-98.

Mata, R., Nava, E., & Sendra, F. (2000) Microcalcifications detection using multiresolution methods. *Pattern Recognition, 2000. Proceedings. 15th International Conference,* vol.4., pp. 344-347.

McInerney, T., Hamarneh, G., Shenton, M., & Terzopoulos, D. (2002). Deformable organisms for automatic medical image analysis. *Medical Image Analysis*, vol. 6, no. 3, pp. 251-266.

McKenna, M. T., Wang, S., Nguyen, T. B., Burns, J. E., Petrick, N., & Summers, R. M. (2012) Strategies for improved interpretation of computer-aided detections for CT colonography utilizing distributed human intelligence. *Medical Image Analysis*, vol. 16, no. 6, pp. 1280-1292.

Meng, H., Hong, W., Song, J., & Wang, L. (2008, May). Feature extraction and analysis of ovarian cancer proteomic mass spectra. In *The 2nd International Conference on Bioinformatics and Biomedical Engineering, ICBBE 2008.* pp. 668-671.

Messinger, D. W., & Albano, J. (2011) A graph theoretic approach to anomaly detection in hyperspectral imagery. In *IEEE 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1-4.

Milton, N. (2012) Acquiring Knowledge from Subject Matter Experts. *In J. Kantola & W. Karwowski (Eds.), Knowledge Service Engineering Handbook. CRC Press, Boca Raton, Florida, USA.*

Mingming, N. Y. (2000) Probabilistic networks with undirected links for anomaly detection. In *Proceedings of the IEEE Systems, Man, and Cybernetics Information Assurance and Security Workshop*, pp. 175–179.

Mini, M. G. (2011) Neural network based classification of digitized mammograms. In *Proceedings of the Second Kuwait Conference on e-Services and e-Systems*, pp. 1-5.

Mohapatra, S., Patra, D., & Satpathy, S. (2011) Automated leukemia detection in blood microscopic images using statistical texture analysis. In *Proceedings of the 2011 International Conference on Communication, Computing & Security*, pp. 184-187.

Moradi, M., Abolmaesumi, P., Isotalo, P. A., Siemens, D. R., Sauerbrei, E. E., & Mousavi, P. (2006) Detection of prostate cancer from RF ultrasound echo signals using fractal analysis. *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE,* pp. 2400-2403.

Moshtaghi, M., Leckie, C., Karunasekera, S., & Rajasegarar, S. (2014) An adaptive elliptical anomaly detection model for wireless sensor networks. *International Journal of Computer Networks*, vol. 64, pp. 195-207.

Mouelhi, A., Sayadi, M., & Fnaiech, F. (2013a) A supervised segmentation scheme based on multilayer neural network and color active contour model for breast cancer nuclei detection. In *International Conference on Electrical Engineering and Software Applications (ICEESA)*, pp. 1-6.

Mouelhi, A., Sayadi, M., & Fnaiech, F. (2013b) Hybrid segmentation of breast cancer cell images using a new fuzzy active contour model and an enhanced watershed method. In *IEEE International Conference on Control, Decision and Information Technologies (CoDIT)*, pp. 382-387.

Mousazadeh, S., & Cohen, I. (2010) Anomaly detection in sonar images based on wavelet domain noncausal AR-ARCH random field modeling. In *Proceedings of 26th IEEE Convention of Electrical and Electronics Engineers in Israel*, pp. 17-20.

Moya, M., Koch, M., & Hostetler, L. (1993) One-class classifier networks for target recognition applications. In *Proceedings of the World Congress on Neural Networks, International Neural Network Society*, pp. 797–801.

Murray, A. F. (2001) Novelty detection using products of simple experts: A potential architecture for embedded systems. *Neural Netw.*, vol. 14, no. 9, pp. 1257–1264.

Muthukarthigadevi, R., & Anand, S. (2013) Detection of architectural distortion in mammogram image using wavelet transform. In *IEEE International Conference on Information Communication and Embedded Systems (ICICES)*, pp. 638-643.

Naghdy, Naghdy, G., Ross, M., Todd, C., & Norachmawati, E. (2010) Classification cervical cancer using histology images. *Computer Engineering and Applications (ICCEA), 2010 Second International Conference,* vol.1, pp. 515-519.

Naik, S., Madabhushi, A., Tomaszeweski, J., & Feldman, M. D. (2007) A quantitative exploration of efficacy of gland morphology in prostate cancer grading. *Bioengineering Conference, 2007. NEBC '07. IEEE 33rd Annual Northeast,* pp. 58-59.

Nairac, A., Corbett-Clark, T., Ripley, R., Townsend, N., & Tarassenko, L. (1997) Choosing an appropriate model for novelty detection. In *Proceedings of the 5th IEEE International Conference on Artificial Neural Networks*, pp. 227–232.

Nairac, A., Townsend, N., Carr, R., King, S., Cowley, P., & Tarassenko, L. (1999) A system for the analysis of jet engine vibration data. *Integ. Comput.-Aided Eng.*, vol. 6, no. 1, pp. 53–56.

Negnevitsky, M. (2005) *Artificial intelligence: a guide to intelligent systems*. Essex, Pearson Education.

Neofytou, M. S., Tanos, V., Pattichis, M. S., Kyriacou, E. C., Pattichis, C. S., & Schizas, C. N. (2008) Color multiscale texture classification of hysteroscopy images of the endometrium. *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE,* pp. 1226-1229.

Ng, R. T. & Han, J. (1994) Efficient and effective clustering methods for spatial data mining. In *Proceedings of the 20th International Conference on Very Large Data Bases*, pp. 144-155.

Nishiya, K., Hasegawa, J., & Koike, T. (1982) Dynamic state estimation including anomaly detection and identification for power systems. *Generation, Transmission and Distribution,* vol. 129, no. 5, pp. 192-198.

Niwas, S. I., Palanisamy, P., & Sujathan, K. (2010a) Complex wavelet as nucleus descriptors for automated cancer cytology classifier system using ANN. *Computational Intelligence and Computing Research (ICCIC), 2010 IEEE International Conference,* pp. 1-5.

Niwas, S. I., Palanisamy, P., & Sujathan, K. (2010b) Wavelet based feature extraction method for breast cancer cytology images. *Industrial Electronics & Applications (ISIEA), 2010 IEEE Symposium,* pp. 686-690.

Noble, C. C., & Cook, D. J. (2003) Graph-based anomaly detection. In *Proceedings of the 9th ACMSIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 631–636.

Noto, K., Brodley, C., & Slonim, D. (2012) FRaC: a feature-modeling approach for semi-supervised and unsupervised anomaly detection. *Data mining and knowledge discovery*, vol. 25, no. 1, pp. 109-133.

Odin, T., & Addison, D. (2000) Novelty detection using neural network technology. In *Proceedings of the COMADEN Conference,* pp. 731-743.

Om, H., & Kundu, A. (2012) A hybrid system for reducing the false alarm rate of anomaly intrusion detection system. In *1st International Conference on Recent Advances in Information Technology (RAIT),* pp. 131–136.

Osher, S., & Sethian, J. A. (1988) Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *Journal of computational physics*, vol. 79, no. 1, pp. 12-49.

Palee, P., Sharp, B., Noriega, L., Sebire, N. J. & Platt, C. (2013) Image Analysis of Histological Features in Molar Pregnancies, *Expert Systems with Applications*, vol. 40, no. 17, pp. 7151-7158.

Papadimitriou, S., Kitagawa, H., Gibbons, P. B., & Faloutsos, C. (2002) Loci: Fast outlier detection using the local correlation integral: *Tech. rep. IRP-TR-02-09*, Intel Research Laboratory.

Paredes-Oliva, I., Castell-Uroz, I., Barlet-Ros, P., Dimitropoulos, X., & Sole-Pareta, J. (2012) Practical anomaly detection based on classifying frequent traffic patterns. In *IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 49-54.

Parolin, A., Herzer, E., & Jung, C. R. (2010) Semi-automated diagnosis of melanoma through the analysis of dermatological images. *Graphics, Patterns and Images (SIBGRAPI), 2010 23rd SIBGRAPI Conference,* pp. 71-78.

Parra, L., Deco, G., & Miesbach, S. (1996) Statistical independence and novelty detection with information preserving nonlinear maps. *Neural Comput.*, vol. 8, no. 2, pp. 260–269.

Parzen, E. (1962) On the estimation of a probability density function and mode. *Annals Math. Stat.*, vol. 33, pp. 1065–1076.

Paul, M., Goodman, S., Felix, J., Lewis, R., Hawkins, M., & Drey, E. (2010) Early molar pregnancy: experience in a large abortion service. *Contraception*, vol. 81, pp. 150–6.

Peng, Y., Park, M., Xu, M., Luo, S., Jin, J.S., Cui, Y. & Wong, W.S.F., (2010a). Detection of nuclei clusters from cervical cancer microscopic imagery using C4.5. In *2010 2nd International Conference on Computer Engineering and Technology (ICCET)*. pp. 593--597.

Peng, Y., Park, M., Xu, M., Luo, S., Jin, J.S., Cui, Y. & Wong, W.S.F. (2010b). Clustering nuclei using machine learning techniques. In *International Conference on Complex Medical Engineering (CME), 2010 IEEE/ICME*. vol. 3, pp. 52-57.

Perner, P. (2002) Image mining: issues, framework, a generic tool and its application to medical-image diagnosis. *Engineering Applications of Artificial Intelligence*, vol. 15, no. 2, pp. 205-216.

Petignat, P., Billieux, M. H., Blouin, J. L., Dahoun, S., & Vassilakos, P. (2003) Is genetic analysis useful in the routine management of hydatidiform mole? *Human Reproduction*, vol. 18, no. 2, pp. 243-249.

Petroudi, S., Constantinou, I., Tziakouri, C., Pattichis, M., & Pattichis, C. (2013) Investigation of AM-FM methods for mammographic breast density classification. In *IEEE 13th International Conference on Bioinformatics and Bioengineering (BIBE)*, pp. 1-4.

Petsche, T., Marcantonio, A., Darken, C., Hanson, S., Kuhn, G., & Santoso, I. (1996) A neural network autoassociator for induction motor failure prediction. In *Proceedings of the Conference on Advances in Neural Information Processing*, vol. 8, pp. 924–930.

Phoha, V. V. (2002) *Internet Security Dictionary*. New York, USA, Springer-Verlag.

Phuong, T. V., Hung, L. X., Cho, S. J., Lee, Y., & Lee, S. (2006) An anomaly detection algorithm for detecting attacks in wireless sensor networks. *Intel. Secur. Inform.*, vol. 3975, pp. 735–736.

Pitiot, A., Delingette, H., Thompson, P. M., & Ayache, N. (2004) Expert knowledge-guided segmentation system for brain MRI. *NeuroImage*, vol. 23, pp. S85-S96.

Pokrajac, D., Lazarevic, A., & Latecki, L. J. (2007) Incremental local outlier detection for data streams. In *Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining*, pp. 504-515.

Portnoy, L., Eskin, E., & Stolfo, S. (2001) Intrusion detection with unlabeled data using clustering. In *Proceedings of the ACM Workshop on Data Mining Applied to SecurityI*, pp. 5-8.

Qi, H. & Head, J.F. (2001) Asymmetry analysis using automatic segmentation and classification for breast cancer detection in thermograms. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*. pp. 2866-2869 vol.3.

Quinn, J. A., & Sugiyama, M. (2014) A least-squares approach to anomaly detection in static and sequential data. *Pattern Recognition Letters*, vol. 40, pp. 36-40.

Rabiei, H., Mahloojifar, A., & Farmer, M. E. (2007) PROVIDING CONTEXT FOR TUMOR RECOGNITION USING THE WRAPPER FRAMEWORK. *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium,* pp. 1252-1255.

Ramadas, M., Ostermann, S., & Tjaden, B. C. (2003) Detecting anomalous network traffic with selforganizing maps. In *Proceedings of the Conference on Recent Advances in Intrusion Detection*, pp. 36–54.

Raman, V., Sumari, P., & Raj, G. D. P. (2010) Performance based CBR mass detection in mammograms. *Communication Control and Computing Technologies (ICCCCT), 2010 IEEE International Conference,* pp. 565-568.

Rathore, S., Hussain, M., & Khan, A. (2013) A novel approach for colon biopsy image segmentation. In *IEEE International Conference on Complex Medical Engineering (CME)*, pp. 134-139.

Ratsch, G., Mika, S., Scholkopf, B., & Muller, K. (2002) Constructing boosting algorithms from SVMs: an application to one-class classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1184-1199.

Refaeilzadeh, P., Tang, L., & Liu, H. (2009) Cross-validation. In *Encyclopedia of database systems*, pp. 532-538.

Roberts, S. (1999) Novelty detection using extreme value statistics. In *Proceedings of the IEEE Vision, Image and Signal Processing Conference*, vol. 146, pp. 124-129.

Roberts, S. (2002) Extreme value statistics for novelty detection in biomedical signal processing. In *Proceedings of the 1st International Conference on Advances in Medical Signal and Information Processing*, pp.166-172.

Roberts, S., & Tarassenko, L. (1994) A probabilistic resource allocating network for novelty detection. *Neural Computation*, vol. 6, no. 2, pp. 270-284.

Rousseeuw, P. J., & Leroy, A. M. (1987) *Robust Regression and Outlier Detection*. New York, USA, John Wiley & Sons, Inc.

Ruotolo, R., & Surace, C. (1997) A statistical approach to damage detection through vibration monitoring. In *Proceedings of the 5th Pan-American Congress of Applied Mechanics*, pp. 314-317.

Russell, S., & Norvig, P. (2010) *Artificial intelligence: a modern approach*, 3rd ed. (International edition), Upper Saddle River: Prentice-Hall.

Sakkalis, V., Marias, K., Roniotis, A., & Skounakis, E. (2009) Translating Cancer Research into Clinical Practice: A Framework for Analyzing and Modeling Cancer from Imaging Data. In *IEEE International Conference on Intelligent Systems Design and Applications*, pp. 347-350.

Salem, O., Liu, Y., & Mehaoua, A. (2013) A lightweight anomaly detection framework for medical wireless sensor networks. In *IEEE Conference on Wireless Communications and Networking (WCNC)*, pp. 4358-4363.

Saligrama, V., & Chen, Z. (2012) Video anomaly detection based on local statistical aggregates. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 2112-2119.

Salvado, J., & Roque, B. (2005) Detection of calcifications in digital mammograms using wavelet analysis and contrast enhancement. In *IEEE International Workshop on Intelligent Signal Processing*, pp. 200-205.

Sang, H. L., Jong, H. K., Jeong, S. P., Jung, M. C., Sang, J. P., Yun, S., & Jung (2008) Computerized segmentation and classification of breast lesions using perfusion volume fractions in dynamic contrast-enhanced MRI. *BioMedical Engineering and Informatics, 2008. BMEI 2008. International Conference,* vol. 2, pp. 58-62.

Sarawagi, S., Agrawal, R., & Megiddo, N. (1998) Discovery-driven exploration of OLAP data cubes. In *Proceedings of the 6th International Conference on Extending Database Technology*, pp.168-182.

Saunders, M., Lewis, P., & Thornhill, A. (2009) *Research methods for business students* fifth edition. Essex, UK, Financial Times/ Prentice Hall.

Saunders, R., & Gero, J. (2000) The importance of being emergent. In *Proceedings of the Conference on Artificial Intelligence in Design*.

Sbarufatti, C., Manes, A., & Giglio, M. (2013) Performance optimization of a diagnostic system based upon a simulated strain field for fatigue damage characterization. *International Journal of Mechanical Systems and Signal Processing*, vol. 40, no. 2, pp. 667-690.

Scarth, G., Mcintyre, M.,Wowk, B., & Somorjai, R. (1995) Detection of novelty in functional images using fuzzy clustering. In *Proceedings of the 3rd Meeting of the International Society for Magnetic Resonance in Medicine*, p. 238.

Sebire, N. J., (2010) Histopathological diagnosis of hydatidiform mole: contemporary features and clinical implications.Fetal. *Pediatr Pathol*, vol. 29, pp. 1–16.

Sebire, N. J., Fisher, R. A., & Rees, H. C. (2003) Histopathological Diagnosis of Partial and Complete Hydatidiform Mole in the First Trimester of Pregnancy. *Pediatric and Developmental Pathology*, vol. 6, no. 1, pp. 69-77.

Sebyala, A. A., Olukemi, T., & Sacks, L. (2002) Active platform security through intrusion detection using naive Bayesian network for anomaly detection. In *Proceedings of the London Communications Symposium*.

Seckl, M.J., Sebire, N.J. & Berkowitz, R.S. (2010) Gestational trophoblastic disease, *The Lancet*, vol. 376, pp. 717-729.

Seok, M. H., Hak, J. L., & Jin, Y. C. (2007) Prostate cancer detection using texture and clinical features in ultrasound image. *Information Acquisition, 2007. ICIA '07. International Conference,* pp. 547-552.

Shadbolt, N. R., & Smart, P. R. (2015). Knowledge Elicitation : Methods, Tools and Techniques. *Evaluation of Human Work*, 4th ed, pp. 1-43.

Shanbhag, S., Gu, Y., & Wolf, T. (2010) A Taxonomy and Comparative Evaluation of Algorithms for Parallel Anomaly Detection. *In IEEE Proceedings of 19th International Conference on Computer Communications and Networks (ICCCN)*, pp. 1-8.

Shaw, M. L. G., & Gaines, B. R. (1987) An Interactive Knowledge Elicitation Technique using Personal Construct Technology. In *A. L. Kidd (Ed.), Knowledge Acquisition for Expert Systems: A Practical Handbook*. New York, USA, Plenum Press.

Sheikholeslami, G., Chatterjee, S., & Zhang, A. (1998) Wavecluster: A multi-resolution clustering approach for very large spatial databases. In *Proceedings of the 24rd International Conference on Very Large Databases*, pp. 428–439.

Shekhar, S., Lu, C.-T., & Zhang, P. (2001) Detecting graph-based spatial outliers: Algorithms and applications (a summary of results). In *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 371–376.

Shewhart, W. A. (1931) *Economic Control of Quality of Manufactured Product*, Milwaukee, USA, ASQ Quality Press.

Shyu, M.-L., Chen, S.-C., Sarinnapakorn, K., & Chang, L. (2003) A novel anomaly detection scheme-based on principal component classifier. In *Proceedings of the 3rd IEEE International Conference on Data Mining*, pp. 353-365.

Singh, S. & Markou, M. (2004) An approach to novelty detection applied to the classification of image regions. *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 4, pp. 396–407.

Smith, Kim, I., Hancock, B. W., Newlands, E. S., Berkowitz, R. S., & Cole, L. A. (2003) *Gestational Trophoblastic Diseases, 2ⁿᵈ edition*, Sheffield: International Society for the Study of Trophoblastic Diseases.

Smith, R., Bivens, A., Embrechts, M., Palagiri, C., & Szymanski, B. (2002) Clustering approaches for anomaly-based intrusion detection. In *Proceedings of the Intelligent Engineering Systems through Artificial Neural Networks*, pp. 579–584.

Smyth, P. (1994) Markov monitoring with unknown states. *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 9, pp. 1600-1612.

Snyder, D. (2001) Online intrusion detection using sequences of system calls. M.S. thesis, Department of Computer Science, Florida State University.

Sohn, H., Worden, K., & Farrar, C. (2001) Novelty detection under changing environmental conditions. In *Proceedings of the 8th Annual SPIE International Symposium on Smart Structures and Materials*, pp. 108-118.

Solberg, H. E., & Lahti, A. (2005) Detection of outliers in reference distributions: Performance of Horn's algorithm. *Clinical Chem.*, vol. 51, no. 12, pp. 2326–2332.

Song, J., Takakura, H., Okabe, Y., & Nakao, K. (2013) Toward a more practical unsupervised anomaly detection system. *Journal of Information Sciences*, vol. 231, pp. 4-14.

Song, Q., Hu, W., & Xie, W. (2002) Robust support vector machine with bullet hole image classification. *IEEE Trans. Syst. Man Cyber.—Part C: Applications and Reviews*, vol. 32, no. 4, pp. 440–448.

Song, S., Shin, D., & Yoon, E. (2001) Analysis of novelty detection properties of auto-associators. In *Proceedings of the Conference on Condition Monitoring and Diagnostic Engineering Management*, pp. 577–584.

Song, X., Wu, M., Jermaine, C., & Ranka, S. (2007) Conditional anomaly detection. *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 5, pp. 631–645.

Song, Y., Wen, Z., Lin, C. Y., & Davis, R. (2013) One-class conditional random fields for sequential anomaly detection. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, pp. 1685-1691.

Soule, A., Salamatian, K., & Taft, N. (2005) Combining filtering and statistical methods for anomaly detection. In *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*, pp. 31-31.

Spence, C., Parra, L., & Sajda, P. (2001) Detection, synthesis and compression in mammographic image analysis with a hierarchical image probability model. In *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, *IEEE Computer Society*, vol. 3, pp. 3–10.

Srivastava, A. (2006) Enabling the discovery of recurring anomalies in aerospace problem reports using high-dimensional clustering techniques. In *Proceedings of the IEEE Aerospace Conference*, pp. 17–34.

Srivastava, A., & Zane-Ulman, B. (2005) Discovering recurring anomalies in text reports regarding complex space systems. In *Proceedings of the IEEE Aerospace Conference*, pp. 3853–3862.

Stefano, C., Sansone, C., & Vento, M. (2000) To reject or not to reject: that is the question: An answer in the case of neural classifiers. *IEEE Trans. Syst. Manag. Cyber.*, vol. 30, no. 1, pp. 84–94.

Stefansky, W. (1972) Rejecting outliers in factorial designs. *Technometrics*, vol. 14, no. 2, pp. 469-479.

Streifel, R., Maks, R., & El-Sharkawi, M. (1996) Detection of shorted-turns in the field of turbinegenerator rotors using novelty detectors–development and field tests. *IEEE Trans. Energy Conv.*, vol. 11, no. 2, pp. 312–317.

Su, C., Chiu, H., & Hsieh, T. (2011). An efficient image retrieval based on HSV color space. In *Proceedings of International Conference Electrical and Control Engineering (ICECE)*, pp. 5746-5749.

Subramaniam, S., Palpanas, T., Papadopoulos, D., Kalogeraki, V., & Gunopulos, D. (2006) Online outlier detection in sensor data using non-parametric models. In *Proceedings of the 32nd International Conference on Very Large Data Bases (VLDB*, pp. 187–198.

Sumithran, E., Cheah, P.-L., Susil, J. B., & Looi, L.-M. (1996) Problems in the histological assessment of hydatidiform moles: A study on consensus diagnosis and ploidy status by fluorescent in situ hybridisation. *Pathology*, vol. 28, no. 4, pp. 311–315.

Sun, J., Qu, H., Chakrabarti, D., & Faloutsos, C. (2005) Neighborhood formation and anomaly detection in bipartite graphs. In *Proceedings of the 5th IEEE International Conference on Data Mining*. pp. 418–425.

Sun, J., Xie, Y., Zhang, H., & Faloutsos, C. (2007) Less is more: Compact matrix representation of large sparse graphs. In *Proceedings of the 7th SIAM International Conference on Data Mining*, pp. 6–22.

Sun, P. & Chawla, S. (2004) On local spatial outliers. In *Proceedings of the 4th IEEE International Conference on Data Mining*, pp. 209–216.

Sun, P., Chawla, S., & Arunasalam, B. (2006) Mining for outliers in sequential databases. In *Proceedings of the SIAM International Conference on Data Mining*.

Surace, C. & Worden, K. (1998) A novelty detection method to diagnose damage in structures: An application to an offshore platform. In *Proceedings of the 8th International Conference of Off-Shore and Polar Engineering*, vol. 4, pp. 64–70.

Surace, C., Worden, K., & Tomlinson, G. (1997) A novelty detection approach to diagnose damage in a cracked beam. In *Proceedings of the SPIE*, vol. 3089, pp. 947–953.

Sykacek, P. (1997) Equivalent error bars for neural network classifiers trained by Bayesian inference. In *Proceedings of the European Symposium on Artificial Neural Networks*, pp. 121–126.

Tagger, B. (2005) An enquiry into the extraction of tacit knowledge. Retrieved April, 30, 2006.

Taher, F., & Sammouda, R. (2011) Lung cancer detection by using artificial neural network and fuzzy clustering methods. *GCC Conference and Exhibition (GCC), 2011 IEEE,* pp. 295-298.

Tan, P.-N., Steinbach, M., & Kumar, V. (2005) *Introduction to Data Mining*. Addison-Wesley.

Tang, J., Chen, Z., Chee Fu, A. W., & W.Cheung, D. (2002) Enhancing effectiveness of outlier detections for low density patterns. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 535–548.

Tarassenko, L. (1995) Novelty detection for the identification of masses in mammograms. In *Proceedings of the 4th IEEE International Conference on Artificial Neural Networks*, vol. 4, pp. 442–447.

Tartakovsky, A. G., Polunchenko, A. S., & Sokolov, G. (2013) Efficient computer network anomaly detection by changepoint detection methods. *IEEE Journal of Selected Topics in Signal Processing,* vol. 7, no. 1, pp. 4-11.

Teng, H., Chen, K., & Lu, S. (1990) Adaptive real-time anomaly detection using inductively generated sequential patterns. In *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, pp. 278–284.

Theiler, J. P., & Cai, D. M. (2003) Resampling approach for anomaly detection in multispectral images. In *Proceedings of the International Society for Optics and Photonics*, pp. 230-240.

Thompson, B., Ii, R. M., Choi, J., El-Sharkawi, M., Huang, M., & Bunje, C. (2002) Implicit learning in autoencoder novelty assessment. In *Proceedings of the International Joint Conference on Neural Networks*, pp. 2878–2883.

Thottan, M., & Ji, C. (2003) Anomaly detection in IP networks. *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2191-2204.

Thottan, M., Liu, G., & Ji, C. (2010). Anomaly detection approaches for communication networks. *Algorithms for Next Generation* Networks, Springer: London.

Tibshirani, R., & Hastie, T. (2007) Outlier sums for differential gene expression analysis. *Biostatistics*, vol. 8, no. 1, pp. 2-8.

Ticehurst, G.W., & Veal, A.J. (2000) *Business research methods: a managerial approach*. Sydney, Australia, Longman, Pearson education Pty limited.

Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M., Mehra, R., Sun, X. W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., Lee, C., Montie, J. E., Shah, R., Pienta, K. J., Rubin, M., & Chinnaiyan, A. M. (2005) Recurrent fusion of tmprss2 and ets transcription factor genes in prostate cancer. *Science*, vol. 310, no. 5748, pp. 603–611.

Torheim, T., Malinen, E., Kvaal, K., Lyng, H., Indahl, U. G., Andersen, E. K., & Futsaether, C. M. (2014) Classification of Dynamic Contrast Enhanced MR Images of Cervical Cancers Using Texture Analysis and Support Vector Machines. *In IEEE Transactions on Medical Imaging*, vol.33, no.8, pp.1648-1656.

Torr, P. & Murray, D. (1993) Outlier detection and motion segmentation. In *Proceedings of the SPIE*. Sensor Fusion VI, S. Schenker, Ed., vol. 2059, pp. 432–443.

Torrent, A., Oliver, A., Llado, X., Marti, R., & Freixenet, J. (2010) A supervised micro-calcification detection approach in digitised mammograms. *Image Processing (ICIP), 2010 17th IEEE International Conference,* pp. 4345-4348.

Tsay, R. S., Pea, D., & Pankratz, A. E. (2000) Outliers in multi-variate time series. *Biometrika*, vol. 87, no. 4, pp. 789–804.

Upadhyaya, B. R., Mathai, G., & Green, J. D. (1988). Data clustering and prediction for fault detection and diagnostics. *American Control Conference,* pp. 650-651.

Valdes, A., & Skinner, K. (2000) Adaptive, model-based monitoring for cyber attack detection. In *Proceedings of the 3rd International Workshop on Recent Advances in Intrusion Detection*, pp. 80–92.

Vani, G., Savitha, R., & Sundararajan, N. (2010) Classification of abnormalities in digitized mammograms using extreme learning machine. *Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference,* pp. 2114-2117.

Vasconcelos, G. C., Fairhurst, M. C., & Bisset, D. L. (1995) Investigating feed-forward neural networks with respect to the rejection of spurious patterns. *Patt. Recog. Lett.*, vol. 16, no. 2, pp. 207–212.

Vasconcelos, G., Fairhurst, M., & Bisset, D. (1994) Recognizing novelty in classification tasks. In *Proceedings of the Neural Information Processing SystemsWorkshop on Novelty Detection and Adaptive Systems Monitoring*.

Vassilakos, P., Riotton, G., & Kajii, T. (1977) Hydatidiform mole: two entities. A morphologic and cytogenetic study with some clinical consideration. *American journal of obstetrics and gynecology*, vol. 127, no. 2, pp. 167-170.

Vijayaraghavan, R., Eswari, C., & Raajan, N. R. (2014) Analysis of Ductal carcinoma using K-means clustering. In *IEEE International Conference on Electronics and Communication Systems (ICECS)*, pp. 1-4.

Vilalta, R., & Ma, S. (2002). Predicting rare events in temporal domains. In *Proceedings of the IEEE International Conference on Data Mining*, pp. 474-481.

Vinueza, A. & Grudic, G. (2004) Unsupervised outlier detection and semi-supervised learning. Tech. rep. CU-CS-976-04, University of Colorado at Boulder.

Waheed, S., Moffitt, R. A., Chaudry, Q., Young, A. N., & Wang, M. D. (2007) Computer aided histopathological classification of cancer subtypes. *Bioinformatics and Bioengineering, 2007. BIBE 2007. Proceedings of the 7th IEEE International Conference,* pp. 503-508.

Wang, S. Y., Lim, K. M., Khoo, B. C., & Wang, M. Y. (2007) An extended level set method for shape and topology optimization. *Journal of Computational Physics*, vol. 221, no. 1, pp. 395-421.
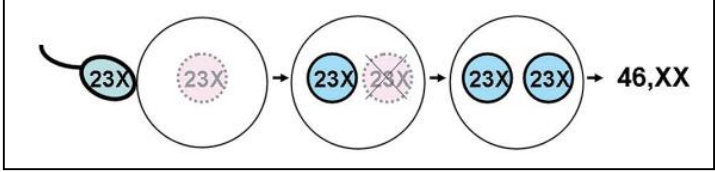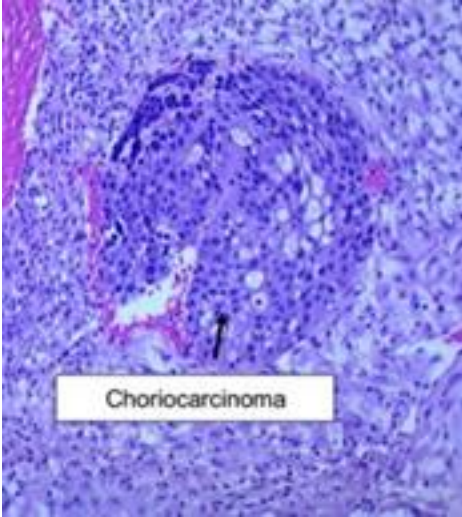
Wang, T., Hamann, A., Spittlehouse, D. L., & Murdock, T. Q. (2012) ClimateWNA-high-resolution spatial climate data for western North America. *Journal of Applied Meteorology and Climatology*, vol. 51, no. 1, pp. 16-29.

Wang, X., Pang, L., Pei, Q., & Li, X. (2010) A scheme for fast network traffic anomaly detection. In *IEEE International Conference on Computer Application and System Modeling (ICCASM)*, vol. 1, pp. 592-596.

Wei, L., Yang, Y., Nishikawa, R. M., & Jiang, Y. (2005). A study on several machine-learning methods for classification of malignant and benign clustered microcalcifications. *Medical Imaging, IEEE Transactions on*, vol. 24, pp. 371-380.

Wei, W., Li, J., Cao, L., Ou, Y., & Chen, J. (2013) Effective detection of sophisticated online banking fraud on extremely imbalanced data. *International Journal of World Wide Web*, vol. 16, no. 4, pp. 449-475.

Weiss, G. M., & Hirsh, H. (1998). Learning to Predict Rare Events in Event Sequences. In *Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining*, pp. 359-363.

Wells, M. (2007) The pathology of gestational trophoblastic disease: recent advances. *Pathology*, vol. 39, pp. 88–96.

Williams, G., Baxter, R., He, H., Hawkins, S., & Gu, L. (2002) A comparative study of RNN for outlier detection in data mining. In *Proceedings of the IEEE International Conference on Data Mining*, pp. 709-712.

Wong, W.-K., Moore, A., Cooper, G., & Wagner, M. (2002) Rule-based anomaly pattern detection for detecting disease outbreaks. In *Proceedings of the 18th National Conference on Artificial Intelligence*, [Online], Available: http://www.cs.cmu.edu/*sim*awm/antiterror.

Wong, W.-K., Moore, A., Cooper, G., & Wagner, M. (2003) Bayesian network anomaly pattern detection for disease outbreaks. In *Proceedings of the 20th International Conference on Machine Learning*, pp. 808–815.

Worden, K. (1997) Structural fault detection using a novelty measure. *J. Sound Vibr.*, vol. 201, no. 1, pp. 85–101.

Xiangmin, X., Xiaojie, X., Yusheng, H., & Zhuocai, W. (2008) On the classification of prostate pathological images based on gleason score. *Intelligent Information Technology Application Workshops, 2008. IITAW '08. International Symposium,* pp. 605-608.
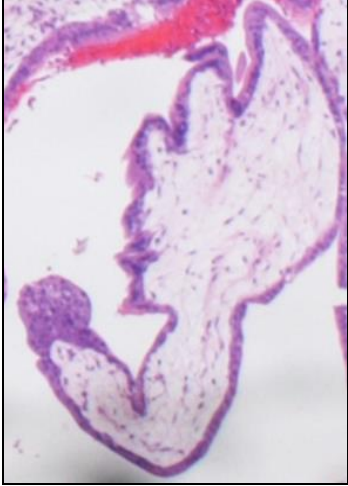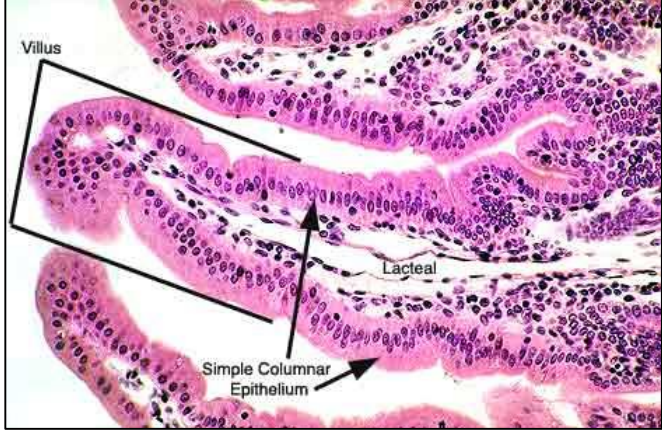
Xiaojing, Y., Zhengyu, Y., Zouridakis, G., & Mullani, N. (2006) SVM-based texture classification and application to early melanoma detection. *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE,* pp. 4775-4778.

Xie, M., Hu, J., & Tian, B. (2012) Histogram-based online anomaly detection in hierarchical wireless sensor networks. In *IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom),* pp. 751-759.

Xie, M., Hu, J., Han, S., & Chen, H. H. (2013) Scalable hypergrid k-NN-Based online anomaly detection in wireless sensor networks. In *IEEE Transactions on Parallel and Distributed Systems,* vol. 24, no. 8, pp. 1661-1670.

Xin, Y., & Pengcheng, S. (2004) Microcalcification detection based on localized texture comparison. *Image Processing, 2004. ICIP '04. 2004 International Conference,* Vol. 5, pp. 2953-2956.

Xing-li, B., & Xu, Q. (2008) Medical image classification based on fuzzy support vector machines. *Intelligent Computation Technology and Automation (ICICTA), 2008 International Conference,* vol. 2, pp. 145-149.

Xu, C., & Prince, J. L. (1998) Snakes, shapes, and gradient vector flow. In *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359-369.

Xu, J. W., Pham, T. D., & Zhou, X. (2011) A double thresholding method for cancer stem cell detection. In *IEEE 7th International Symposium on Image and Signal Processing and Analysis (ISPA)*, pp. 695-699.

Xu, X., Xing, X., Huang, Y., & Wang, Z. (2008) On the Classification of Prostate Pathological Images Based on Gleason Score. In *Intelligent Information Technology Application Workshops, 2008. IITAW '08. International Symposium on*. pp. 605-608.

Yaguchi, A., Kobayashi, T., Watanabe, K., Iwata, K., Hosaka, T., & Otsu, N. (2011) Cancer detection from biopsy images using probabilistic and discriminative features. *IEEE International Conference on Image Processing (ICIP)*, pp. 1609-1612.

Yairi, T., Kato, Y., & Hori, K. (2001) Fault detection by mining association rules from housekeeping data. In *Proceedings of the International Symposium on Artificial Intelligence, Robotics and Automation in Space*, pp. 18-21.

Yamanishi, K., & Takeuchi, J. I. (2001) Discovering outlier filtering rules from unlabeled data: combining a supervised learner with an unsupervised learner. In *Proceedings*
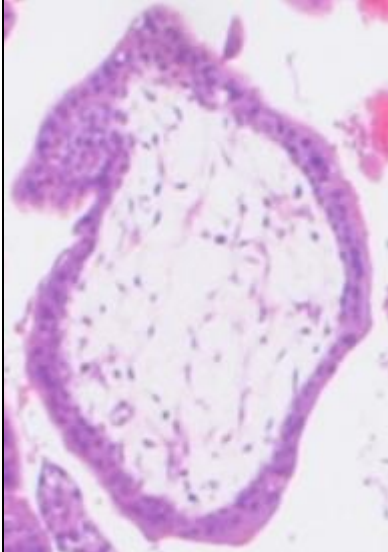
of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 389-394.

Yamanishi, K., Takeuchi, J. I., Williams, G., & Milne, P. (2004) On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms. *Data Mining and Knowledge Discovery*, vol. 8, no. 3, pp. 275-300.

Ypma, A. & Duin, R. (1998) Novelty detection using self-organizing maps. In *Progress in Connectionist Based Information Systems*, vol. 2, pp. 1322–1325.

Yu, D., Sheikholeslami, G., & Zhang, A. (2002) Findout: Finding outliers in very large datasets. *Journal of Knowledge and Information Systems*, vol. 4, no. 4, pp. 387–412.

Yu, P., Park, M., Min X., Suhuai L., Jin, J. S., & Yue, C., (2010) Detection of nuclei clusters from cervical cancer microscopic imagery using C4.5. *Computer Engineering and Technology (ICCET), 2010,* vol. 3, pp. 593-597.

Yuan, X., Yang, Z., Zouridakis, G., & Mullani, N. (2006) SVM-based Texture Classification and Application to Early Melanoma Detection. In *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE.* pp. 4775-4778.

Yue, C., Jin, J. S., Park, M., Suhuai, L., Min, X., & Yu, P. (2010) Computer aided abnormality detection for microscopy images of cervical tissue. *Complex Medical Engineering (CME), 2010 IEEE/ICME International Conference,* pp. 63-68.

Zhang, J. & Wang, H. (2006) Detecting outlying subspaces for high-dimensional data: The new task, algorithms, and performance. *Knowl. Inform. Syst.*, vol. 10, no. 3, pp. 333–355.

Zhang, K., Shi, S., Gao, H., & Li, J. (2007) Unsupervised outlier detection in sensor networks using aggregation tree. In *Advanced Data Mining and Applications*, vol. 4632, pp. 158–169.

Zhang, M., Raghunathan, A., & Jha, N. K. (2013) MedMon: Securing medical devices through wireless monitoring and anomaly detection. *In IEEE Transactions on Biomedical Circuits and Systems*, vol. 7, no.6, pp. 871-881.

Zhao, D., Xu, Q., & Feng, Z. (2010) Analysis and Design for Intrusion Detection System Based on Data Mining. *In IEEE Second International Workshop on Education Technology and Computer Science (ETCS)*, vol. 2, pp. 339-342.
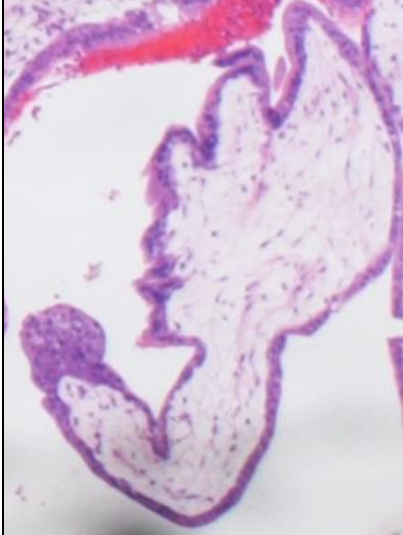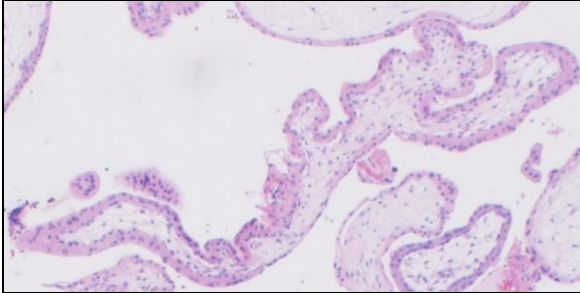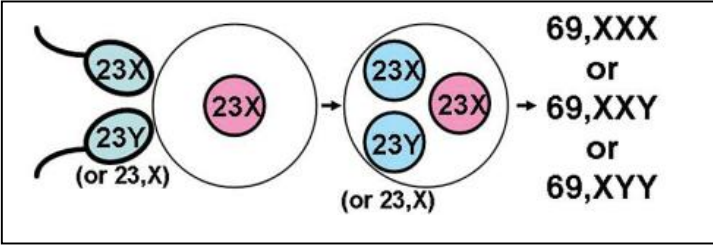
Zheng, K., Wang, T. F., Lin, J. L., & Li, D. Y. (2007) Recognition of breast ultrasound images using a hybrid method. *In Complex Medical Engineering, 2007. CME 2007. IEEE/ICME* pp. 640-643.

Zsambok, C. E., & Klein, G. A. (Eds.) (1997) *Naturalistic Decision Making*. New Jersey, USA, Lawrence Erlbaum Associates.
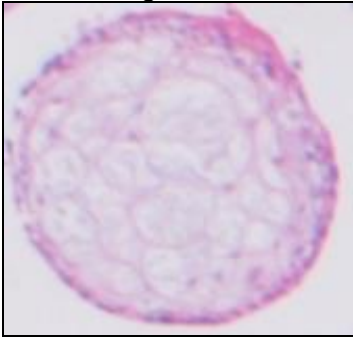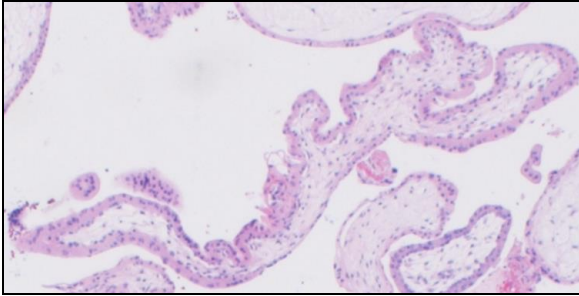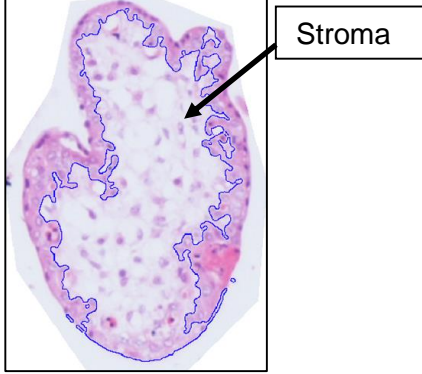
Appendix A. Glossary
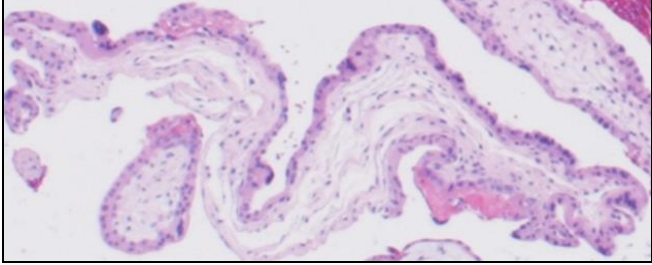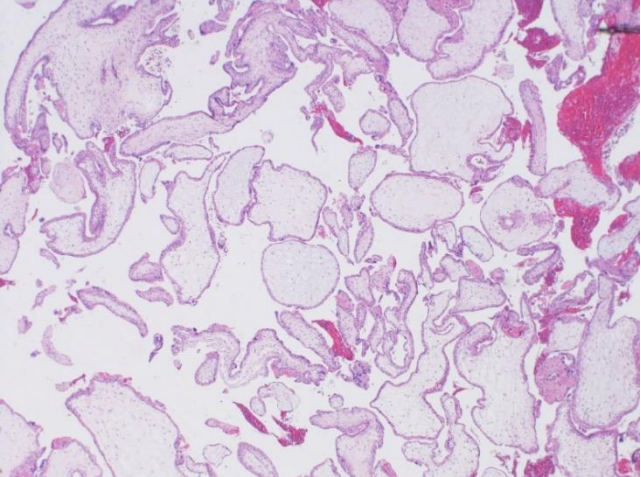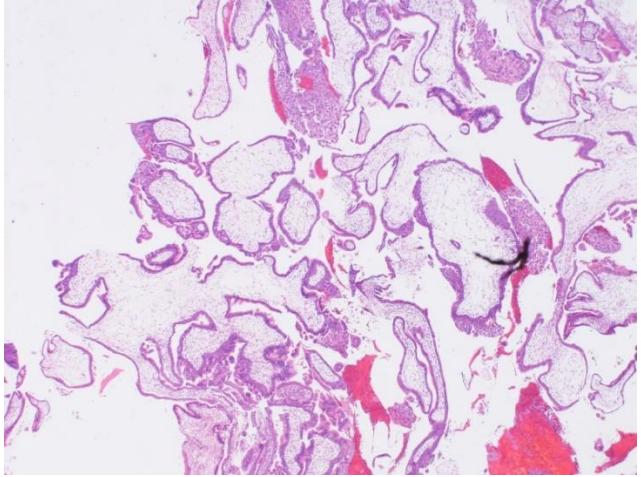
Terms	Description
Androgenetic diploid	The state of a cell comprising of two sets of chromosomes, derived from the father and from the mother respectively; in humans, the diploid number is 46. Androgen is a generic term referring to usually a hormone (e.g. androsterone, testosterone) which stimulates activity of the accessory male sex organs, Courtesy of *Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition*. (2003). Retrieved July 28 2015 from http://medical-dictionary.thefreedictionary.com
Choriocarcinoma	A malignant tumour of trophoblasts Courtesy of Retrieved July 28 2015 from Geneva Foundation for Medical Education and Research "http://www.jmedicalcasereports.com/content/figures/1752-1947-4-379-1-l.jpg"

Terms	Description
Complete hydatidiform moles (CHM)	It is an abnormal placenta but no fetus (Benirschke *et al.*, 2006). (Source: HYMAT tool)
Epithelium	Epithelium, the cellular covering of internal and external surfaces of the body, consists of cells joined by small amounts of cementing substances. Courtesy of Retrieved July 28 2015 from University of New England "http://faculty.une.edu/com/abell/histo/villusweb.jpg"

Terms	Description
Focal abnormal trophoblastic proliferation	 (Source: HYMAT tool)
Histopathological examination (or histopathology)	The study of the microscopic anatomical changes in diseased tissue (Benirschke *et al.*, 2006).
Hydatidiform (mole)	A rare placental mass of growth resulting from the proliferation of the trophoblast and the degeneration of the chorionic villi, indicating an abnormal pregnancy (Benirschke *et al.*, 2006).
Karyorrhexis (n.) Karyorrhectic (adj.)	The fragmentation of nucleus in a cell. Courtesy of Retrieved July 28 2015 from University of Pittsburgh School of Medicine "http://path.upmc.edu/cases/case127/images/micro8.jpg"
Morphology	The form and structure of a particular organism, tissue, or cell.

Terms	Description
Non-polar proliferation	Nonpolar atoms and molecules do not interact freely with water. This is relevant in cell biology because water is a large component of living organisms. (Source: HYMAT tool)
Partial hydatidiform moles (PHM)	It is an abnormal placenta and some fetal development (Benirschke *et al.*, 2006). (Source: HYMAT tool)
Paternal triploid	A chromosomal abnormality affected by the presence of an entire extra chromosomal set. (Van den Veyver & Al-Hussaini, 2006)

Terms	Description
Placenta	An organ that connects between the developing fetus and the uterine wall to exchange nutrient, waste, and gas between mother and fetus. *Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition.* (2003). Retrieved July 28 2015 from http://medical-dictionary.thefreedictionary.com/villus
Regions of Interests (ROI)	It is a portion of an image used to filter or perform some operations on.
Reticular	Resembling a net. (Source: HYMAT tool)
Scalloping dentate and mild (in PHM)	A series of indentations on a normally smooth margin of a structure (Source: HYMAT tool)
Stroma	Connective tissue core of an organ (Benirschke *et al.*, 2006). (Source: HYMAT tool)

Terms	Description
Stroma hydrops and mild	Swelling from excessive accumulation of watery fluid in stroma of villi. (Source: HYMAT tool)
Trimester images of PHM and CHM	PHM and CHM images are only available at first trimester. With the increasing use of ultrasonography to assess complicated pregnancies these moles are normally evacuated before 12 weeks' gestation (Seckl *et al.,* 2010). PHM at first trimester (Source: HYMAT tool). CHM at first trimester (Source: HYMAT tool).

Terms	Description
Trophoblasts	Trophoblasts are specialised cells of the placenta with an important role in embryo implantation and interaction with the maternal uterus (Benirschke *et al.*, 2006). (Source: HYMAT tool)
Villus (pl. vílli)	A small vascular projection from the surface, especially of a mucous membrane Villi are covered with epithelium that diffuses and transports fluids and nutrients. *Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition.* (2003). Retrieved July 28 2015 from http://medical-dictionary.thefreedictionary.com/villus (Source: HYMAT tool)

References:

Benirschke, K., Kaufmann, P., & Baergen, R., (2006) Pathology of the human placenta fifth edition. *New York: Springer.*

Brown, M. J. (1992) Miller-Keane encyclopedia & dictionary of medicine, nursing, & allied health. *WB Saunders Company: Philadelphia.*

McGraw-Hill. (2002) Concise Dictionary of Modern Medicine. New York: McGraw-Hill Companies.

Seckl, M. J., Sebire, N. J., & Berkowitz, R. S. (2010) Gestational trophoblastic disease. The Lancet, vol. 376, no. 9742, pp. 717-729.

Van den Veyver, I. B., & Al-Hussaini, T. K. (2006) Biparental hydatidiform moles: a maternal effect mutation affecting imprinting in the offspring. *Human reproduction update*, vol. 12, no. 3, pp. 233-242.

Appendix B. Ontological Representation of Hydatidiform Moles
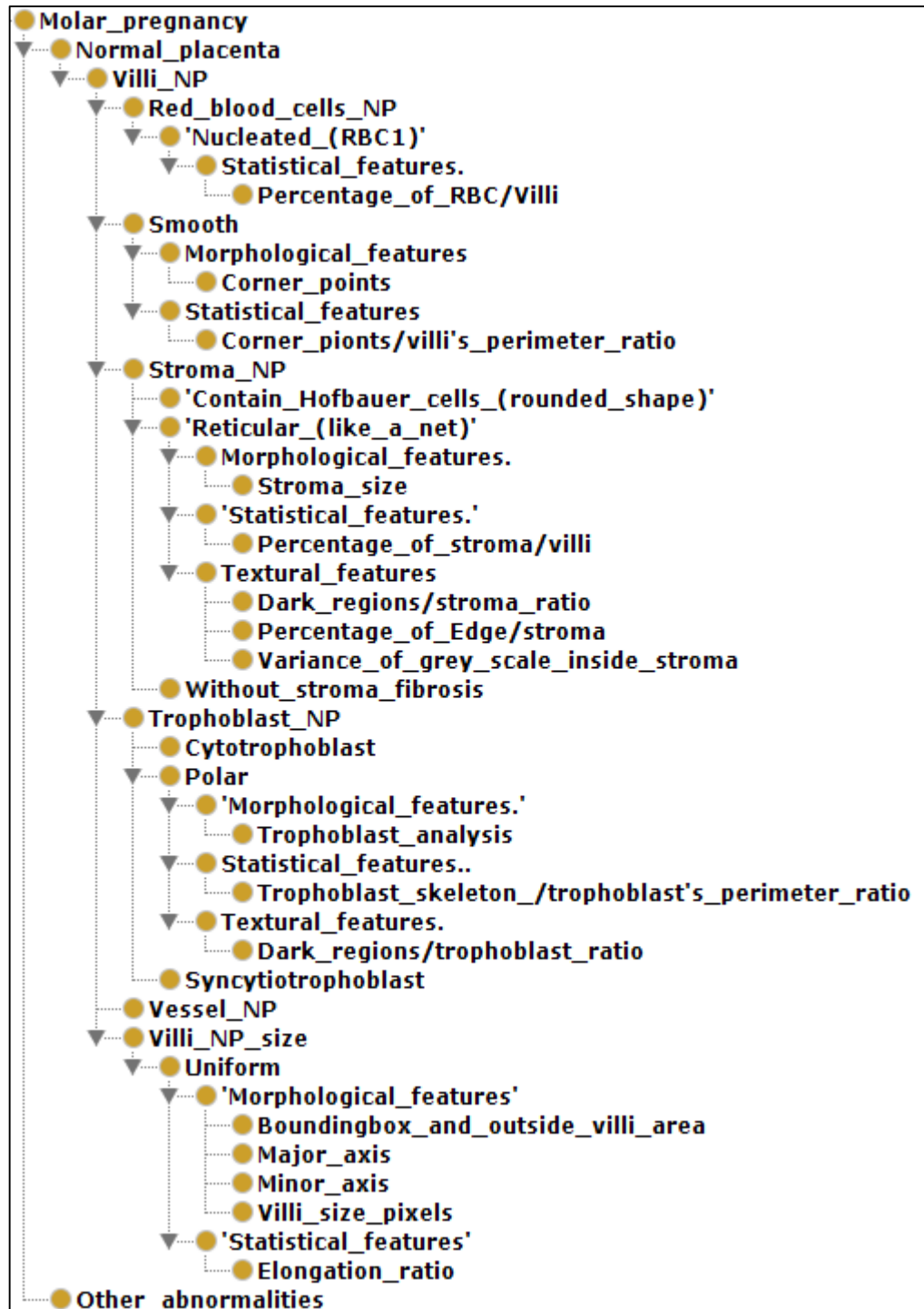
Figure B.1. Ontological representation of normal placental villi.

Figure B.2. Ontological representation of partial hydatidiform moles.

```
●Molar_pregnancy
▼─●Complete_mole
  ▼───●Villi_CM
    │───●No_normal_placental_tissue
    ▼───●Red_blood_cells_CM
      ▼───●'Few_or_no_red_blood_cells_(RBC)'
        ▼───●Statistical_features.
          │───●Percentage_of_red_blood_cells_inside_stroma_regions
    ▼───●Stroma_CM
      │───●Big_central_cistern
      ▼───●Enlarged
        ▼───●'Statistical_features'
          │───●Percentage_of_stroma/villi
      ▼───●Irregular
        ▼───●'Morphological_features'
          │───●Stroma_size
      ▼───●Karyorrhexis
        ▼───●Textural_features
          │───●Dark_regions/stroma_ratio
          │───●Percentage_of_edge_inside_stroma_regions
          │───●Variance_of_grey_scale_of_stroma_regions
    ▼───●Trophoblast_CM
      ▼───●Hyperplasia
        ▼───●'Morphological_features.'
          │───●Trophoblast_analysis
        ▼───●'Statistical_features.'
          │───●Trophoblast_skeleton/trophoblast's_perimeter_ratio
        ▼───●Textural_features.
          │───●Deark_regions/trophoblast_ratio
      ▼───●Oval/irregular
        ▼───●Morphological_features.
          │───●Trophoblast_size
      │───●Two_layers
    ▼───●Vessel_CM
      │───●Collapsed_and_inconspicuous
      │───●Fetal_absent
    ▼───●Villi_CM_size
      ▼───●Varied
        ▼───●Morphological_features
          │───●Corner_points
          │───●Different_area_between_villi's_bounding_box_and_villi_area
          │───●Major_axis
          │───●Minor_axis
          │───●The_notion_of_four_quadrants
          │───●Villi_size
        ▼───●Statistical_features
          │───●Corner_points/villi's_perimeteer_ratio
          │───●Elongation_ratio
```

Figure B.3. Ontological representation of complete hydatidiform moles.

Appendix C. Over- and Under-Segmentation Examples

Notation: N denotes normal placental villi, PHM denotes partial hydatidiform moles and CHM denotes complete hydatidiform moles.

Example: N03-21 denotes normal placental villi slide no. 3 with villi no. 21.

N03-21

N03-25

N05-19

N06-1
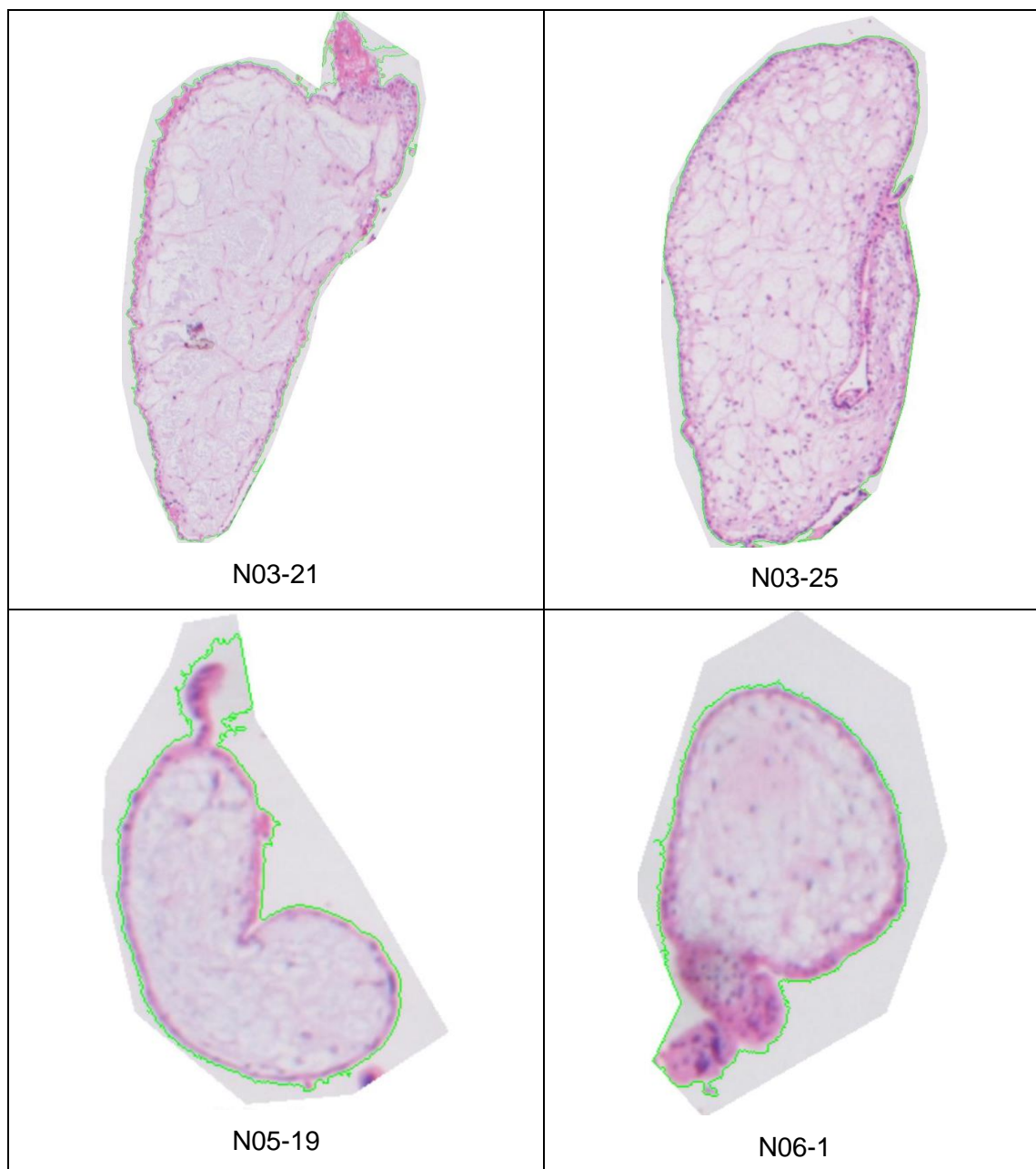
Figure C.1. Over-segmentation examples of normal placental villi

PHM01-5

PHM01-15

PHM02-8

PHM03-9
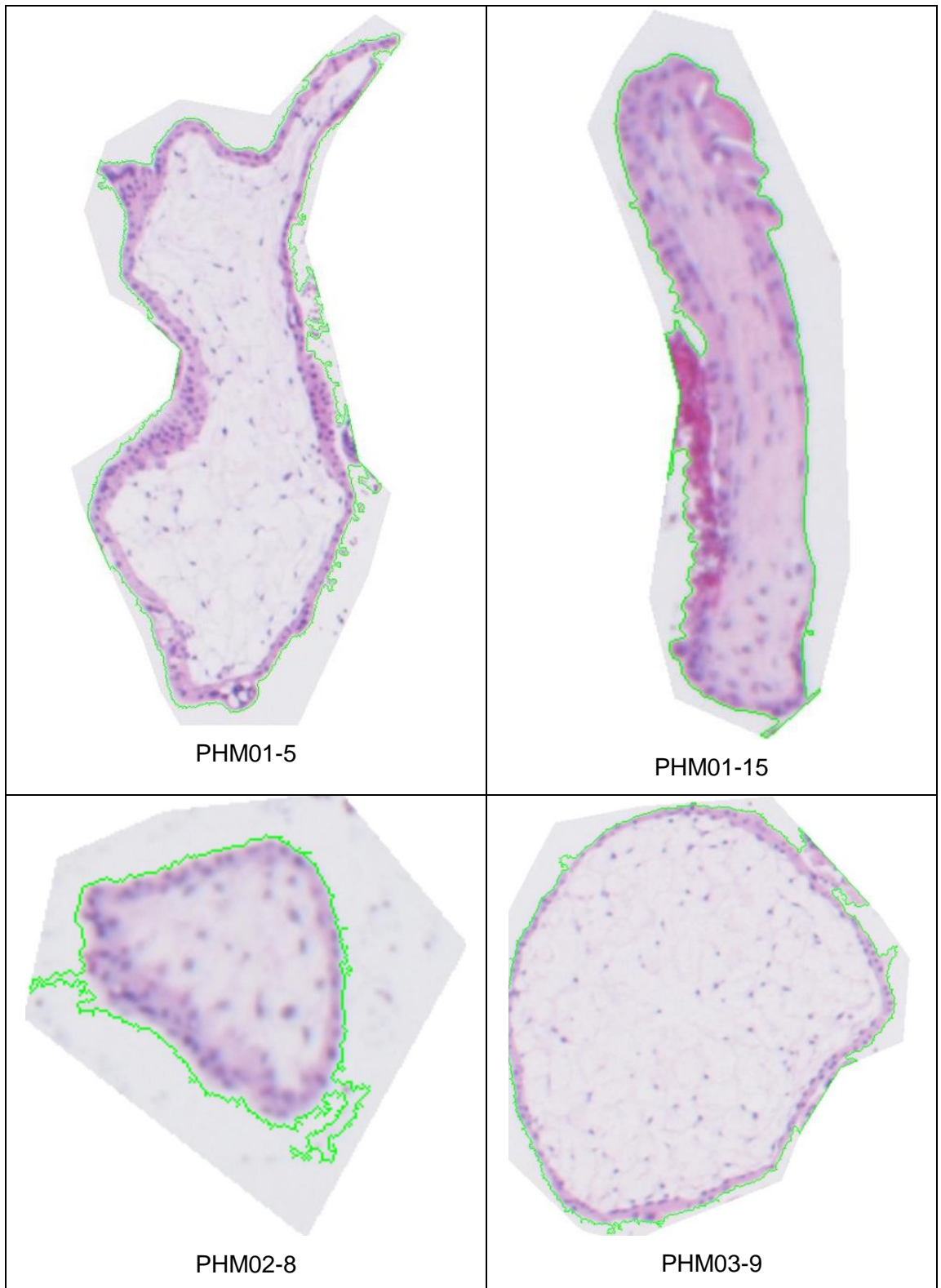
Figure C.2. Over-segmentation examples of PHM villi

CHM01-1

CHM03-20

CHM04-2

CHM04-8
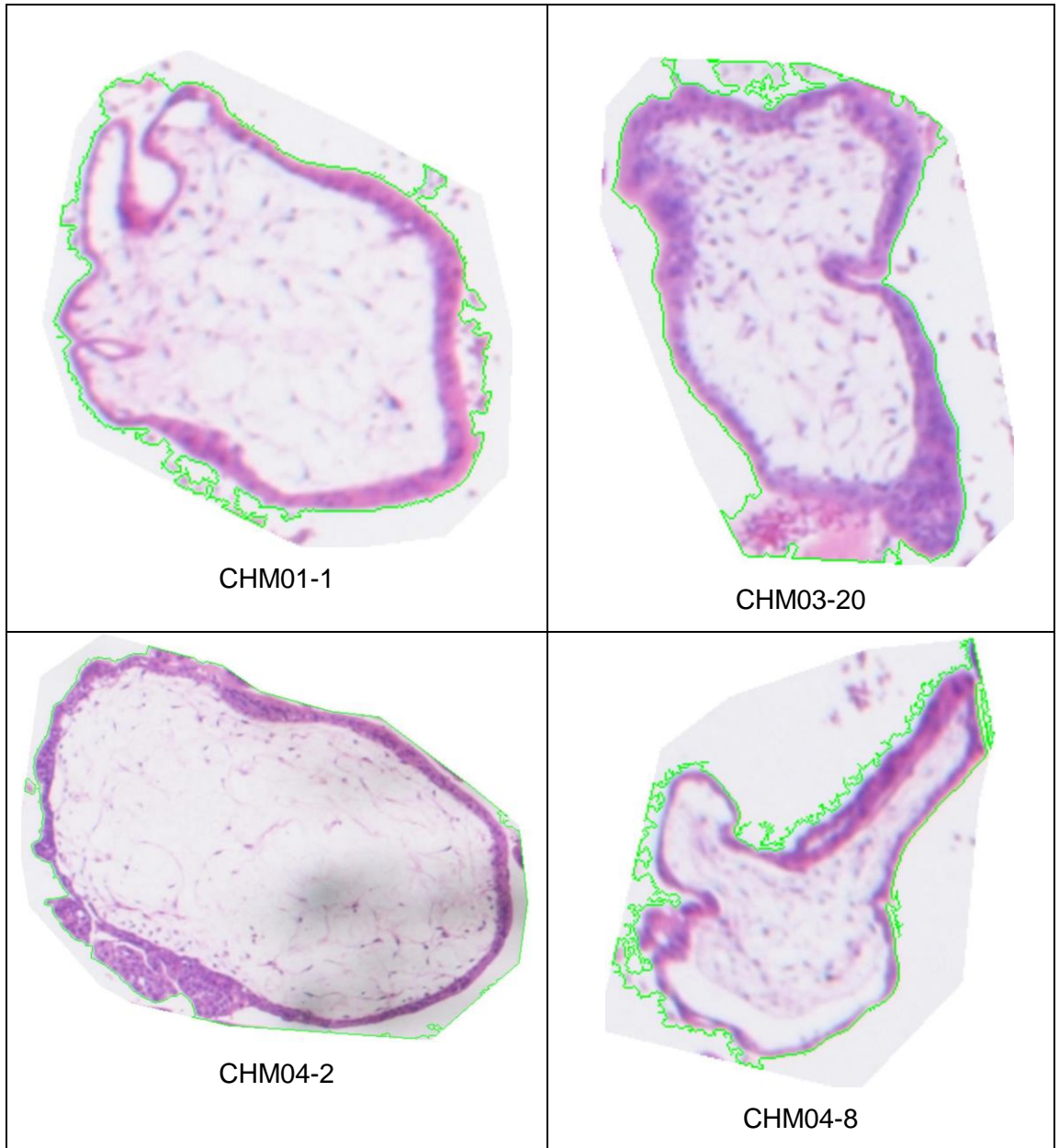
Figure C.3. Over-segmentation examples of CHM villi

N01-4

N01-17

N02-20

N10-13
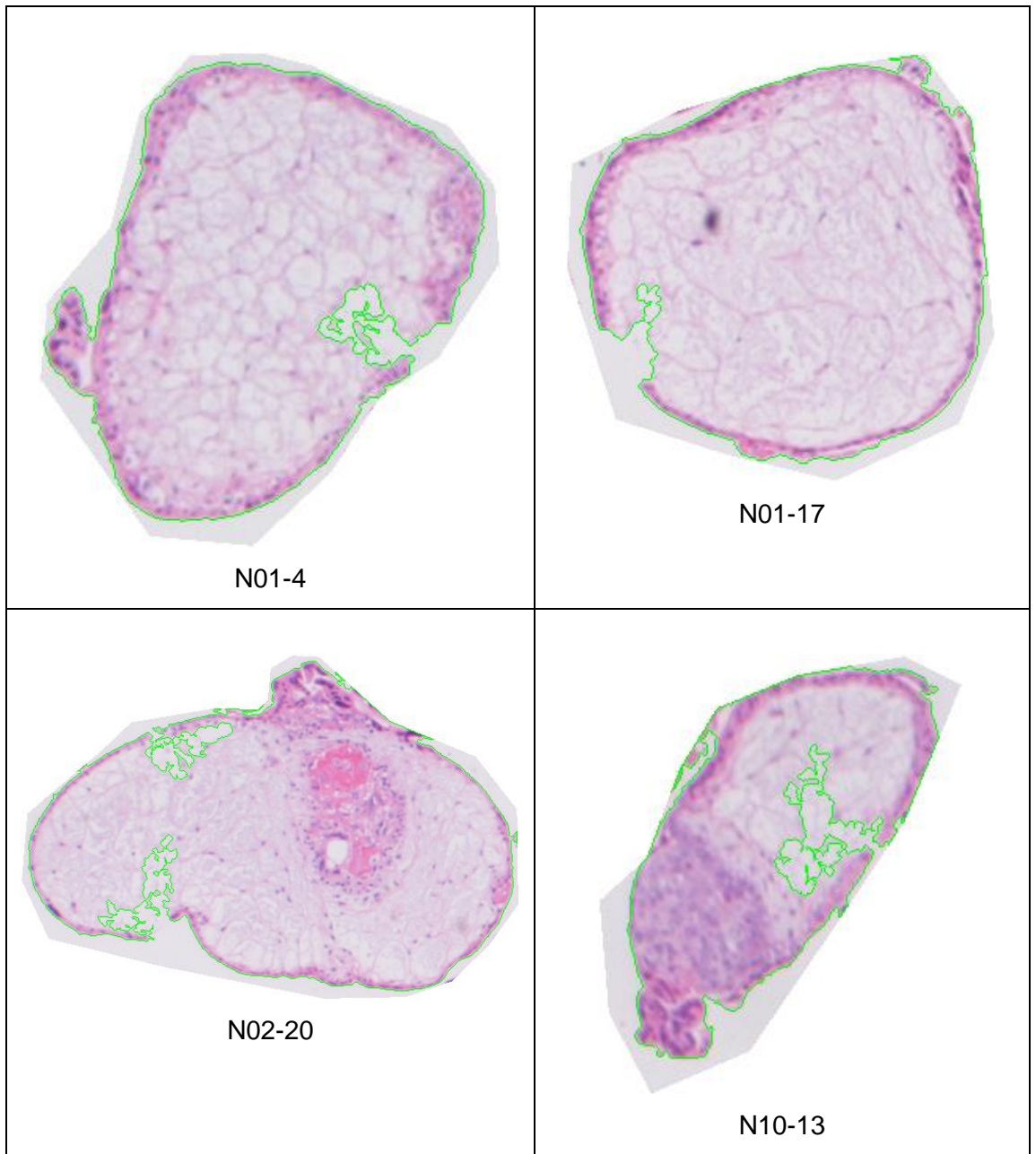
Figure C.4. Under-segmentation examples of normal placental villi

PHM10-11

PHM11-10

PHM12-4

PHM19-9
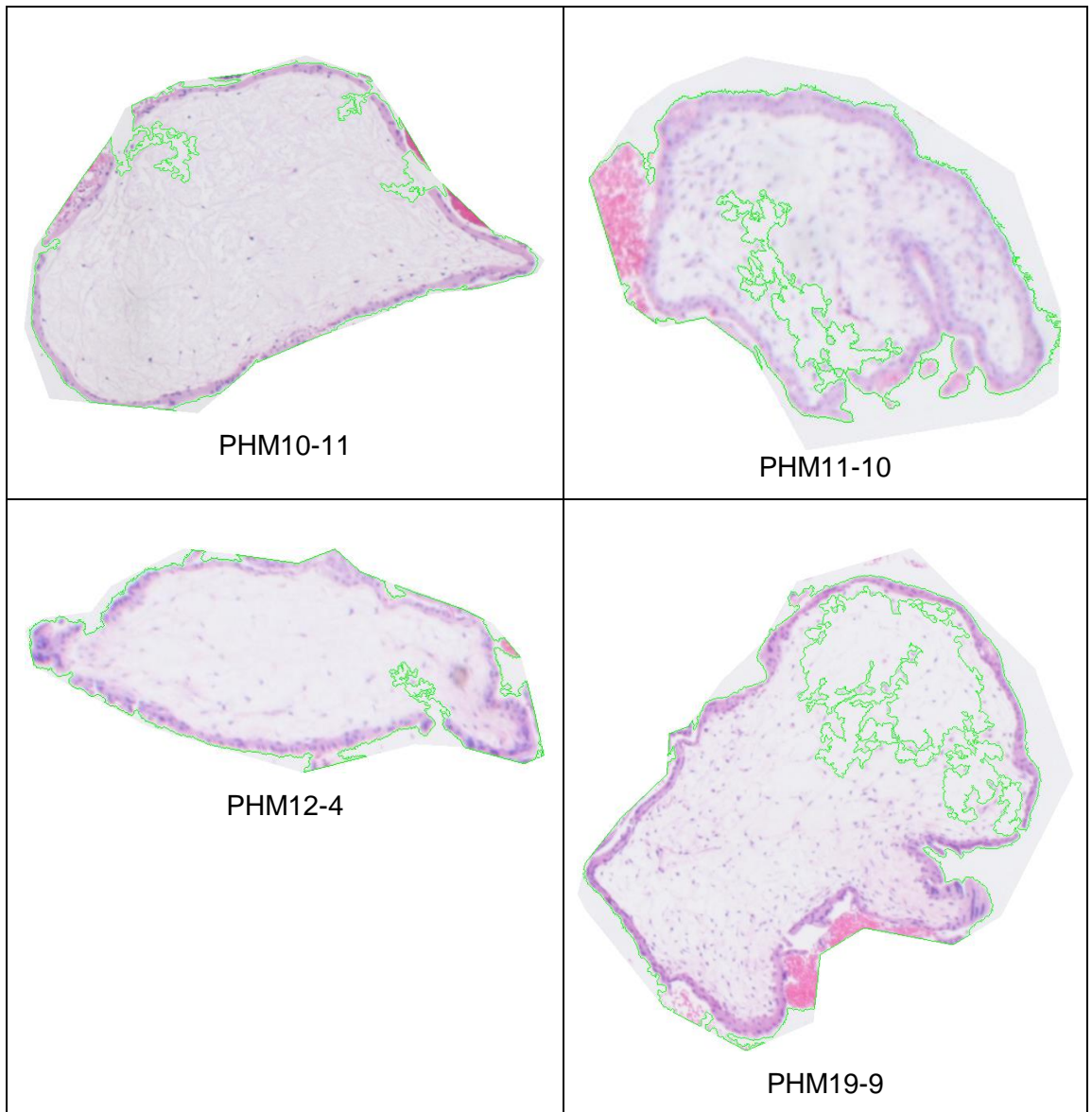
Figure C.5. Under-segmentation examples of PHM villi

CHM03-6

CHM04-12

CHM12-6

CHM19-3
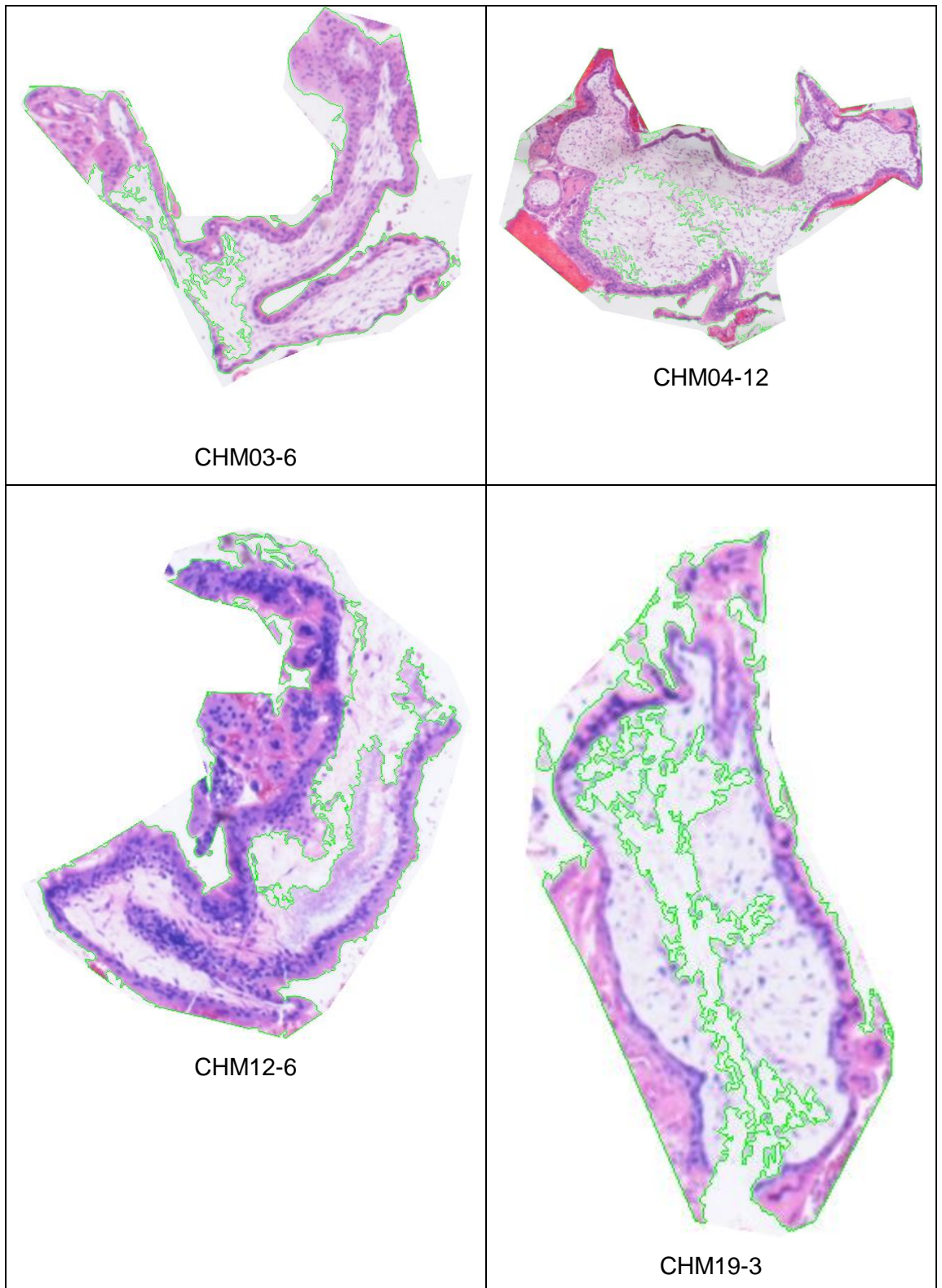
Figure C.6. Under-segmentation examples of CHM villi

Appendix D. Figures of the 15 Features
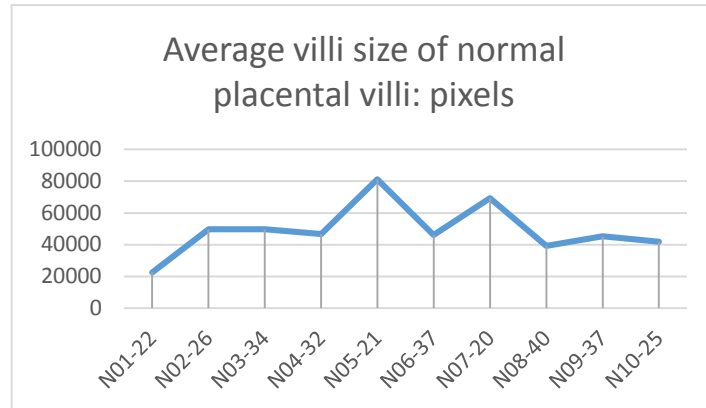
(a) Average villi size of normal placental villi.
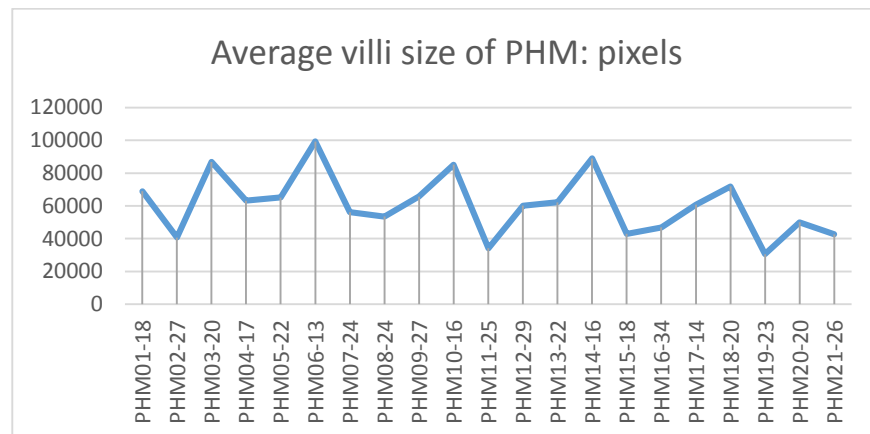
(b) Average villi size of PHM.
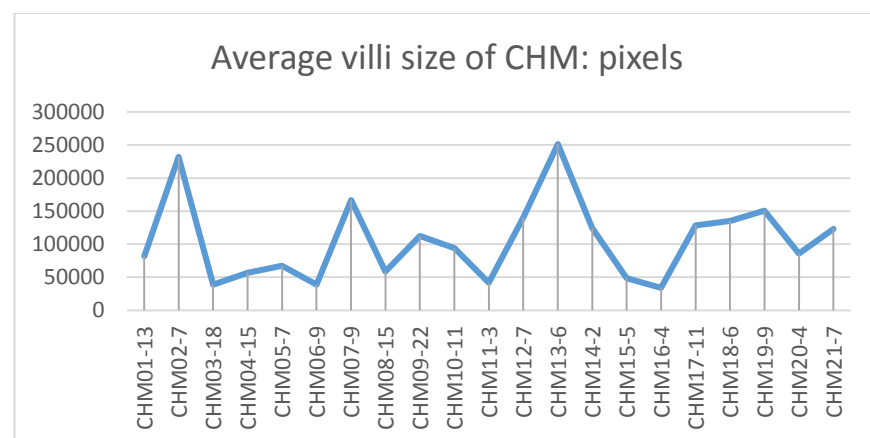
(c) Average villi size of CHM.

Figure D.1. Average villi size of molar pregnancy images.
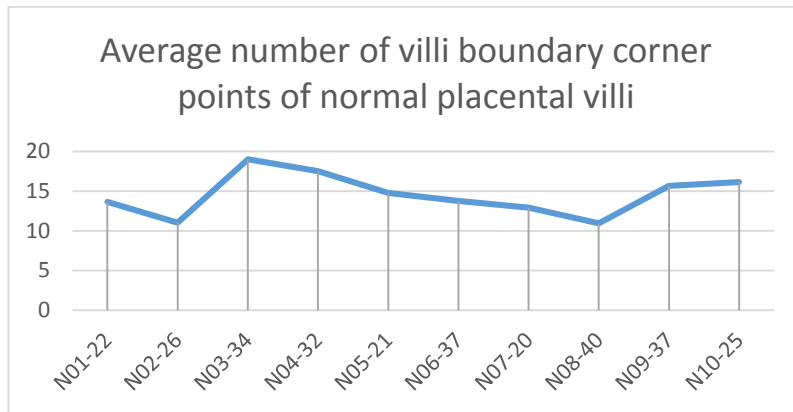
(a) Average number of villi boundary corner points of normal placental villi.
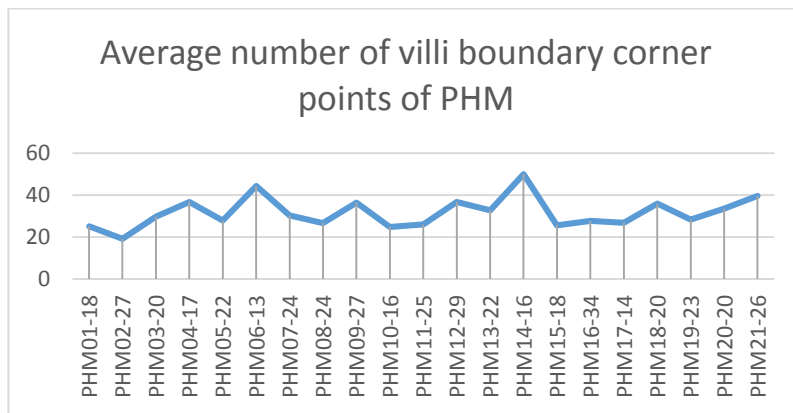
(b) Average number of villi boundary corner points of PHM.
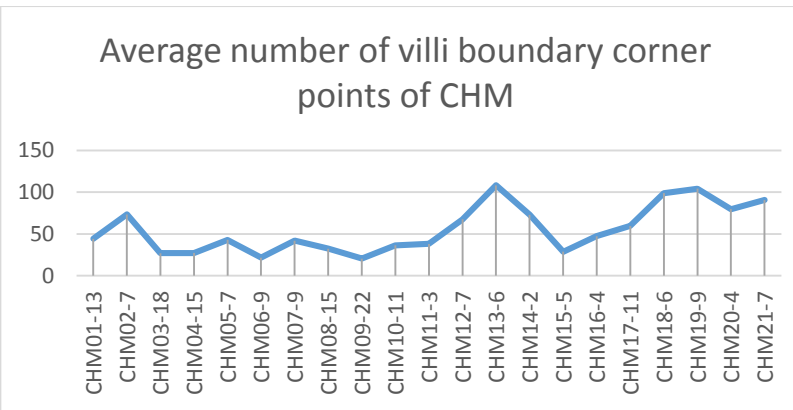
(c) Average number of villi boundary corner points of CHM.

Figure D.2. Average number of villi boundary corner points of molar pregnancy images.
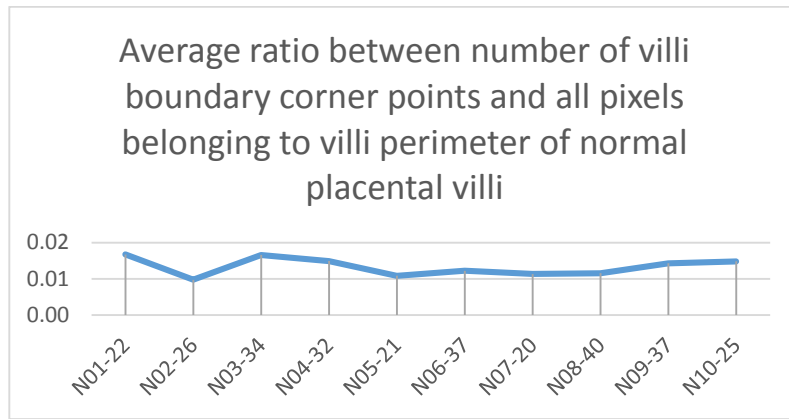
Average ratio between number of villi boundary corner points and all pixels belonging to villi perimeter of normal placental villi

(a) Ratio between number of villi boundary corner points and all pixels belonging to villi perimeter of normal placental villi.

Average ratio between number of villi boundary corner points and all pixels belonging to villi perimeter of PHM

(b) Ratio between number of villi boundary corner points and all pixels belonging to villi perimeter of PHM.
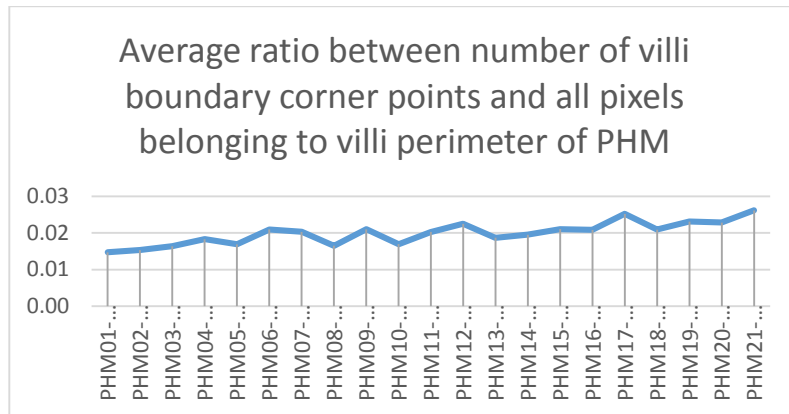
Average ratio between number of villi boundary corner points and all pixels belonging to villi perimeter of CHM
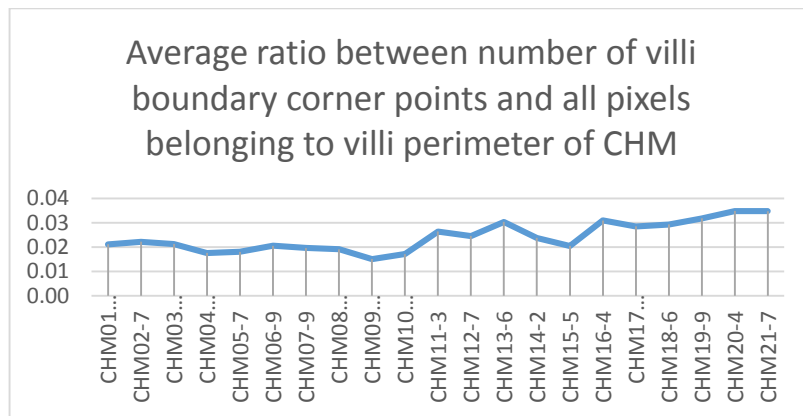
(c) Ratio between number of villi boundary corner points and all pixels belonging to villi perimeter of CHM.

Figure D.3. Ratio between number of villi boundary corner points and all pixels belonging to villi perimeter of molar pregnancy images.
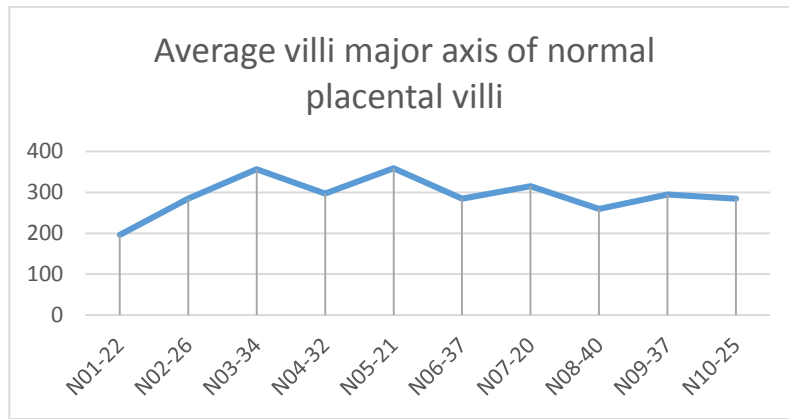
(a) Major axis of normal placental villi.
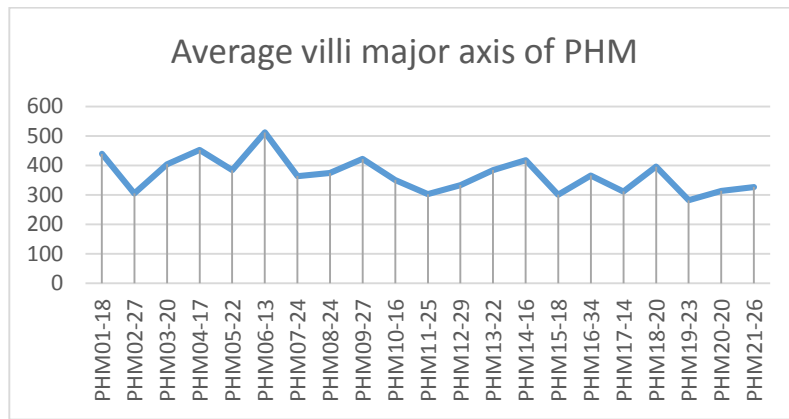
(b) Major axis of PHM.
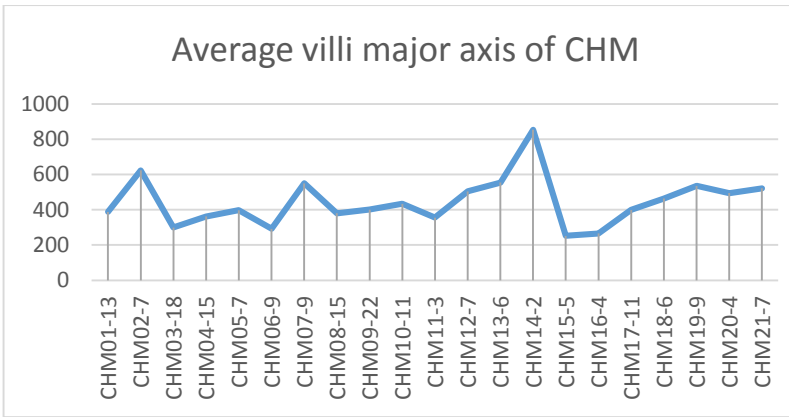
(c) Major axis of CHM.

Figure D.4. Major axis of molar pregnancy images.
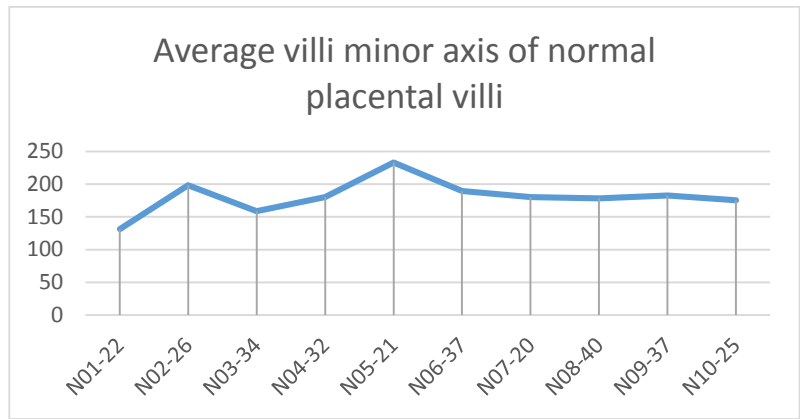
(a) Minor axis of normal placental villi.
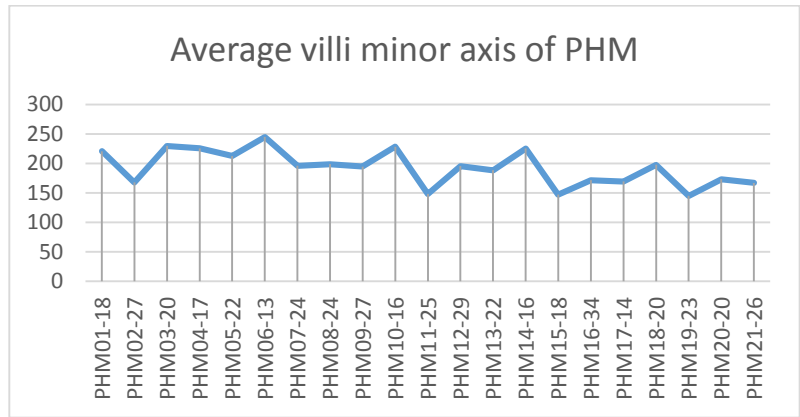
(b) Minor axis of PHM.
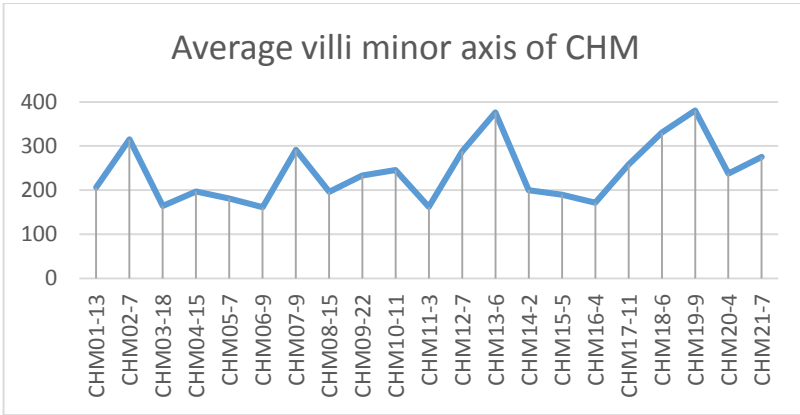
(c) Minor axis of CHM.

Figure D.5. Minor axis of molar pregnancy images.
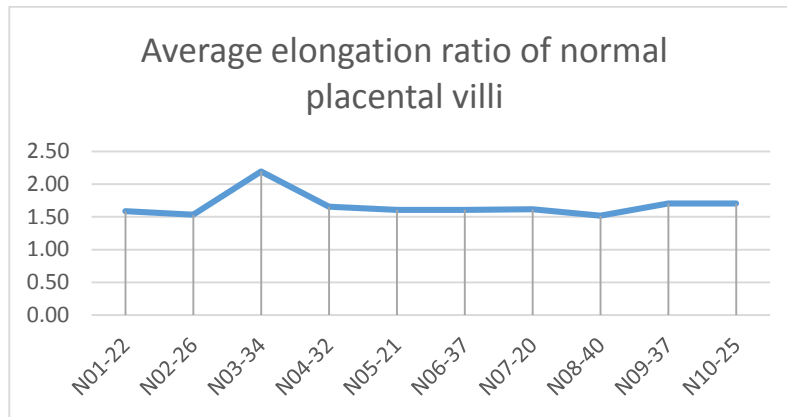
(a) Elongation ratio of normal placental villi.
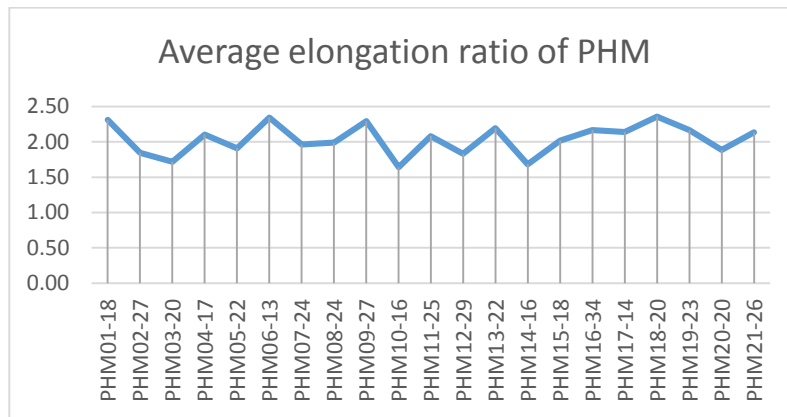
(b) Elongation ratio of PHM.
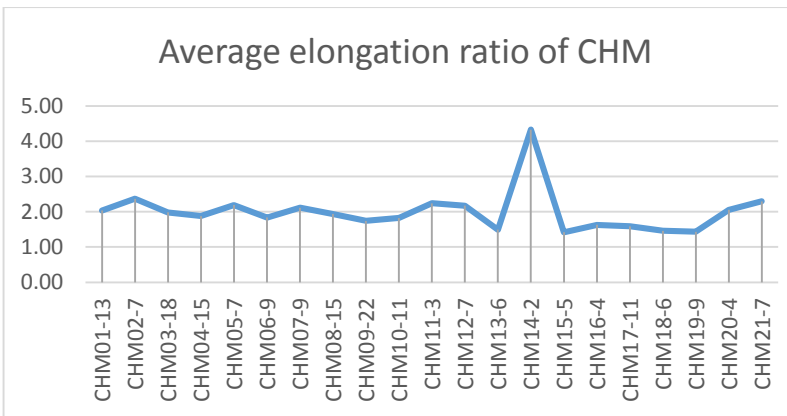
(c) Elongation ratio of CHM.

Figure D.6. Elongation ratio of molar pregnancy images.
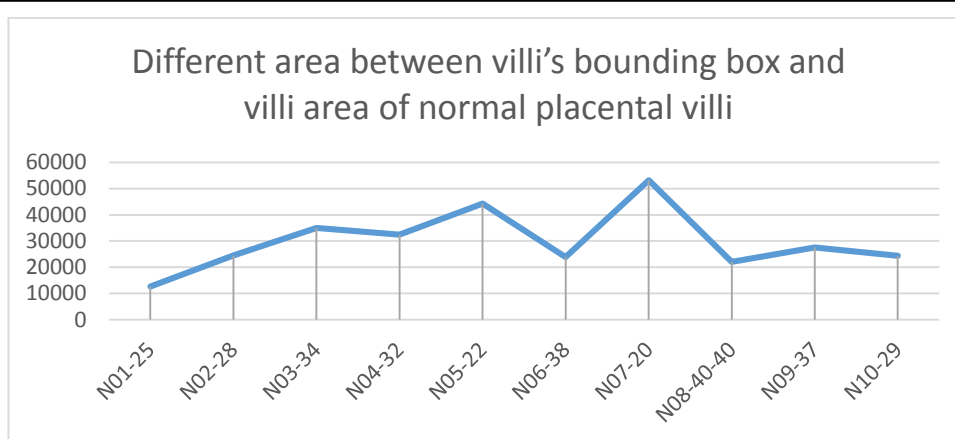
(a) Different area between villi's bounding box and villi area of normal placental villi.
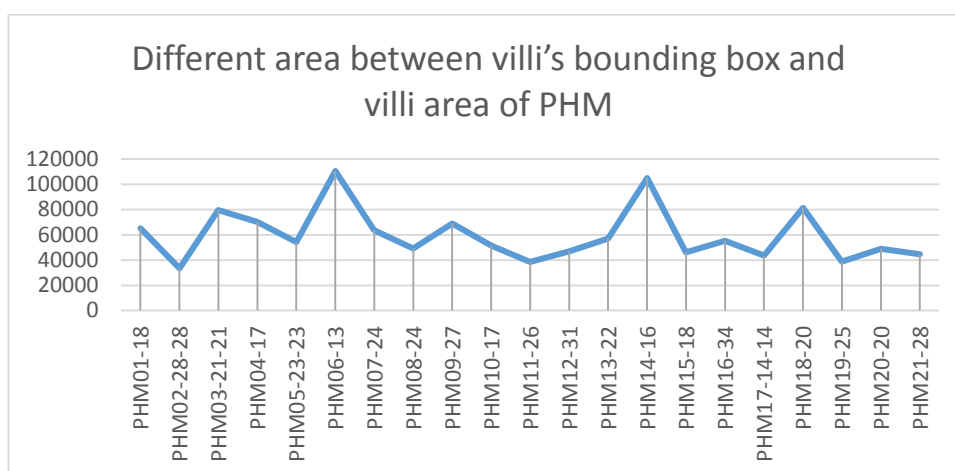
(b) Different area between villi's bounding box and villi area of PHM.
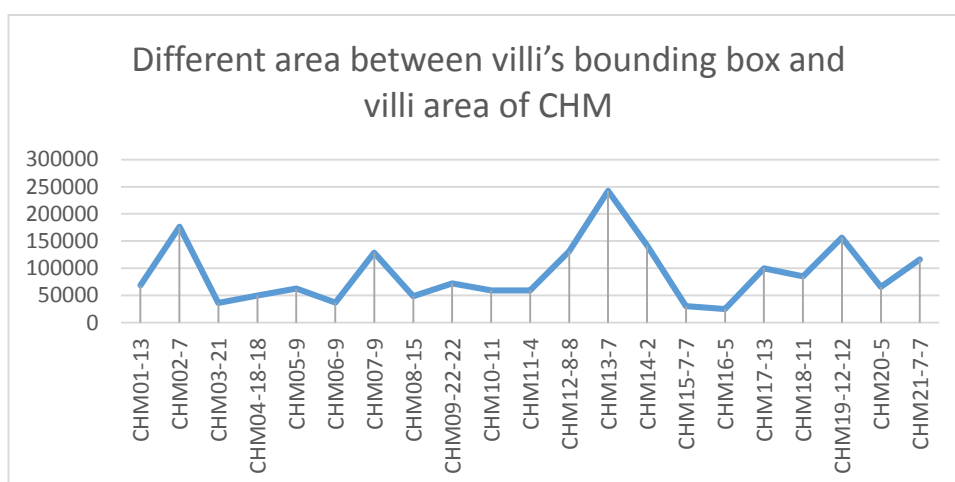
(c) Different area between villi's bounding box and villi area of CHM.

Figure D.7. Different area between villi's bounding box and villi area of molar pregnancy images.
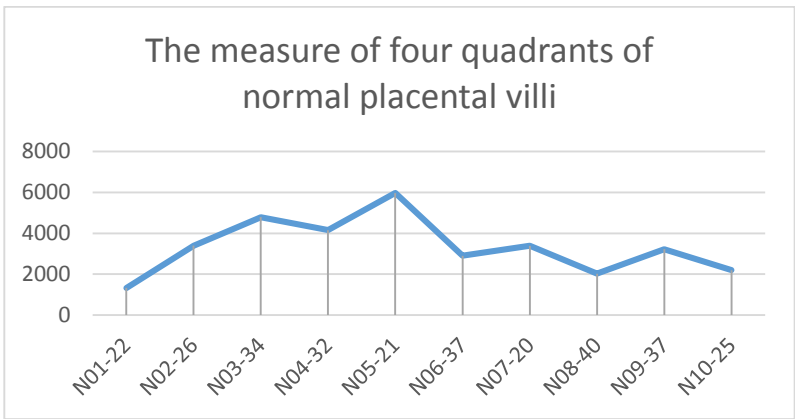
(a) The measure of four quadrants of normal placental villi.
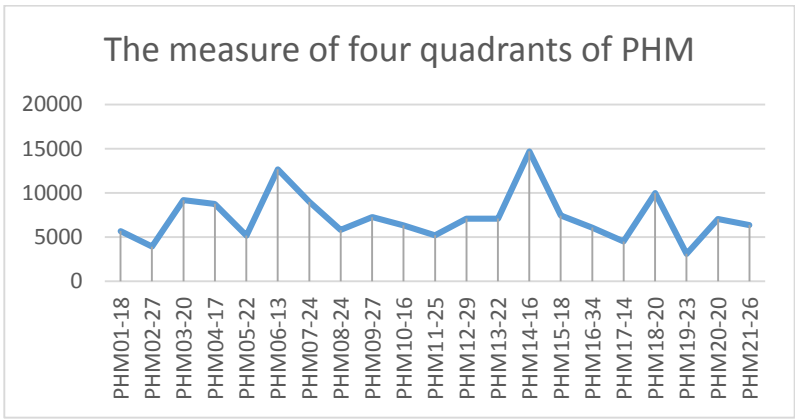
(b) The measure of four quadrants of PHM.
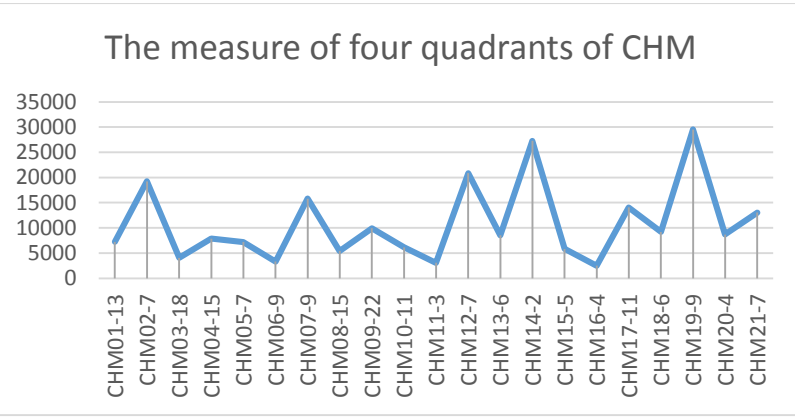
(c) The measure of four quadrants of CHM.

Figure D.8. The measure of four quadrants of molar pregnancy images.
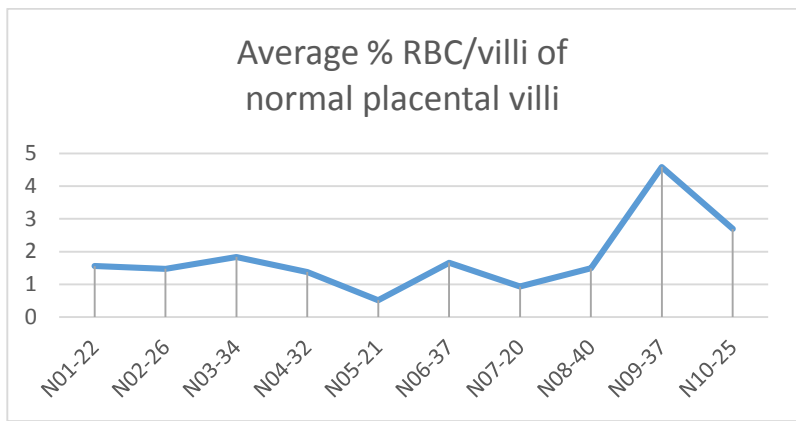
(a) Density of RBC per villi of normal placental villi.
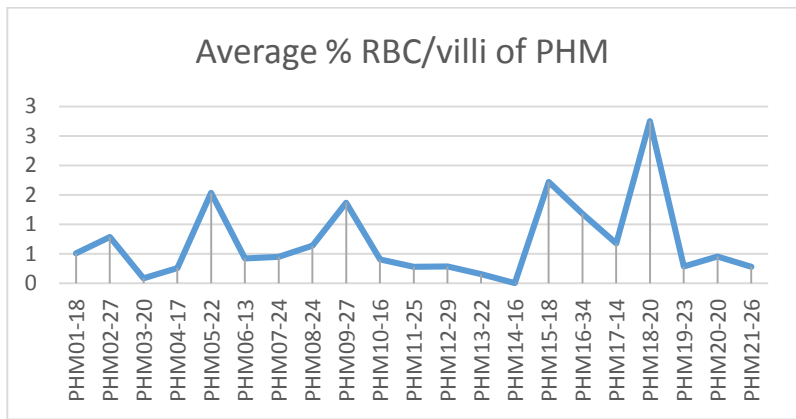
(b) Density of RBC per villi of PHM.
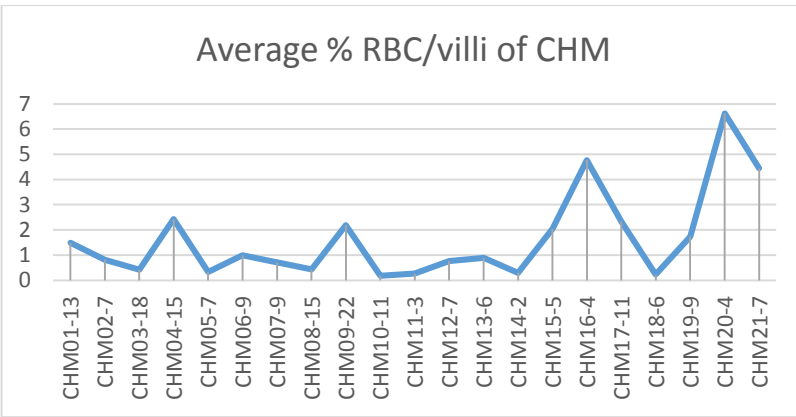
(c) Density of RBC per villi of CHM.

Figure D.9. Density of RBC per villi of molar pregnancy images.
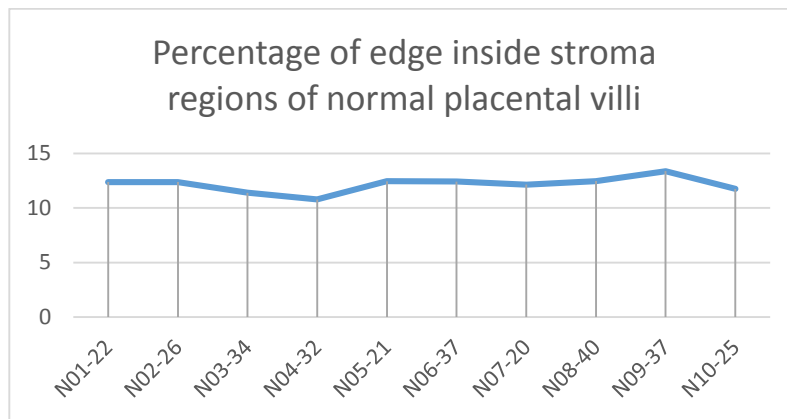
(a) Percentage of edge inside stroma regions of normal placental villi.
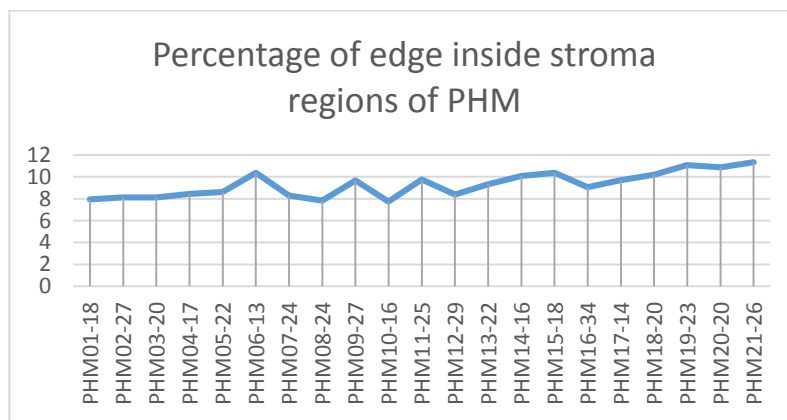
(b) Percentage of edge inside stroma regions of PHM.
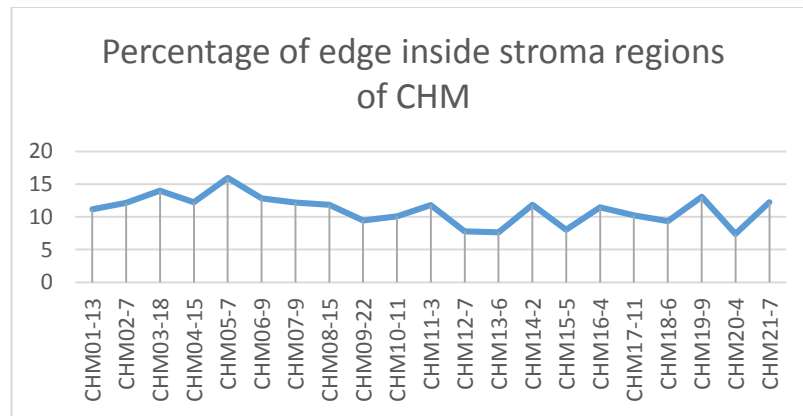
(c) Percentage of edge inside stroma regions of CHM.

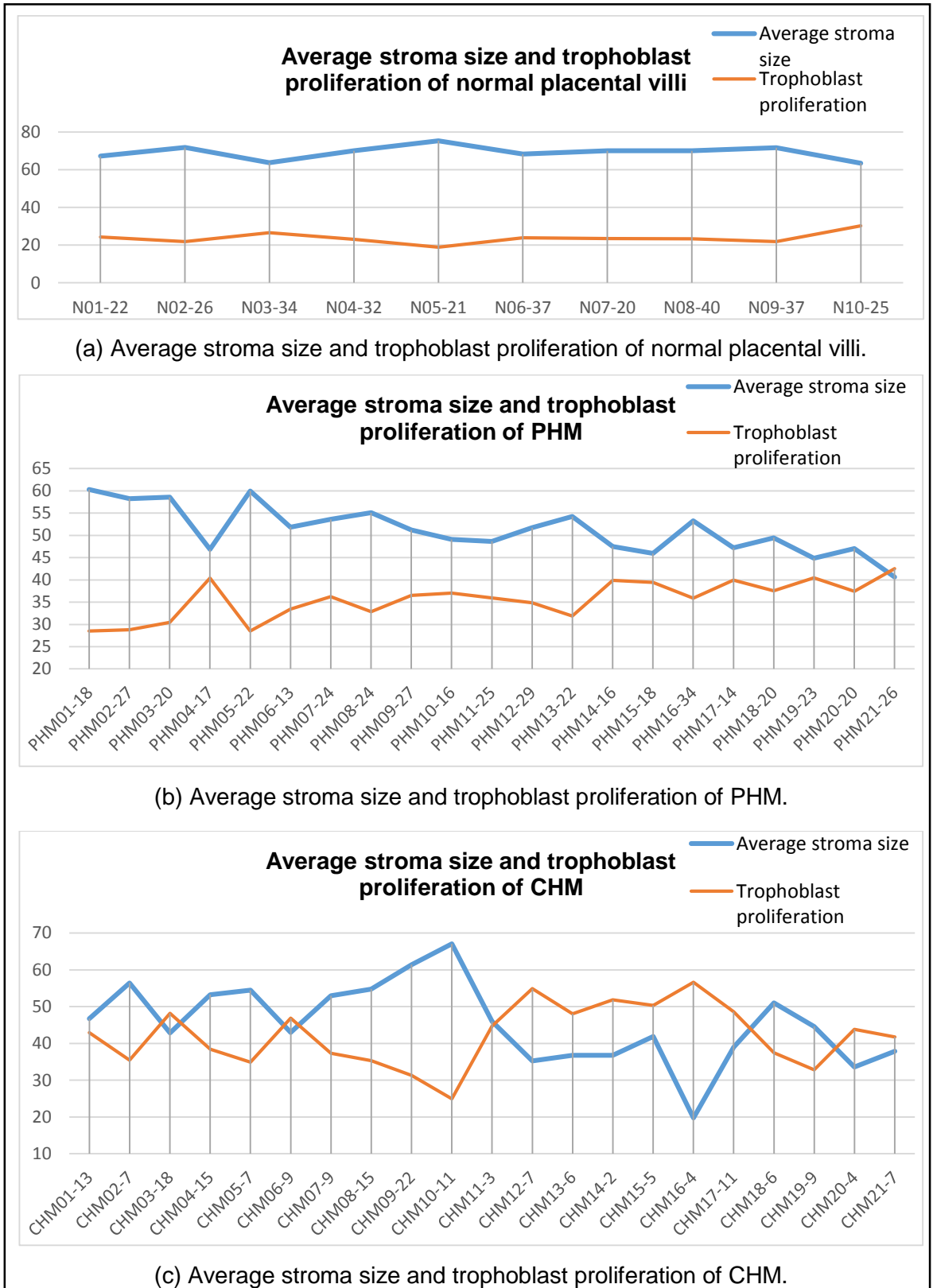Figure D.10. Percentage of edge inside stroma regions of molar pregnancy images.
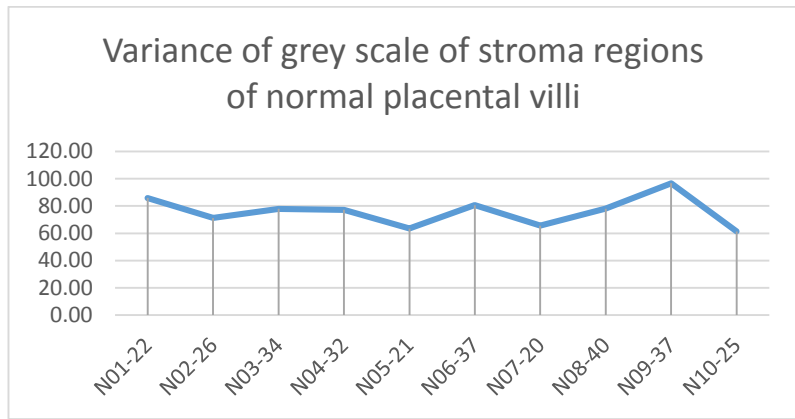
(a) Variance of grey scale of stroma regions of normal placental villi.
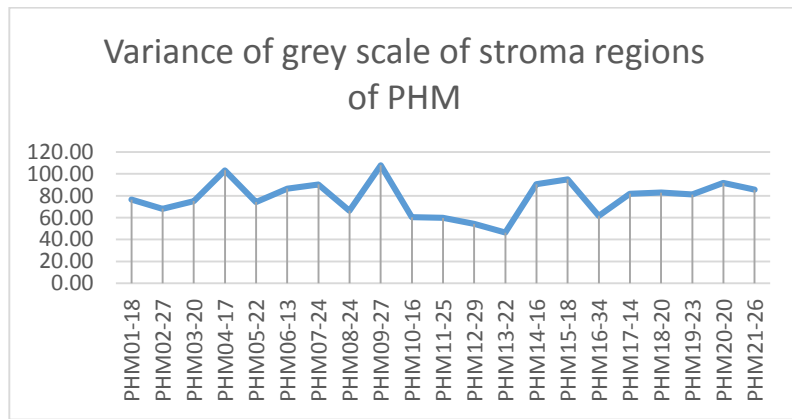
(b) Variance of grey scale of stroma regions of PHM.
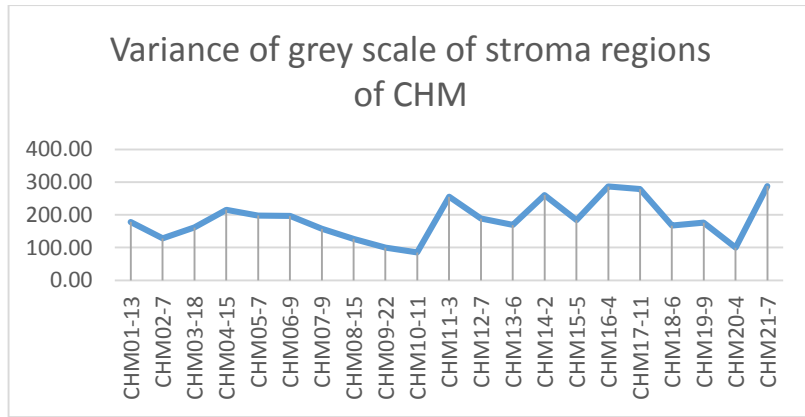
(c) Variance of grey scale of stroma regions of CHM.

Figure D.11. Percentage of edge inside stroma regions of molar pregnancy images.

(a) Average stroma size and trophoblast proliferation of normal placental villi.

(b) Average stroma size and trophoblast proliferation of PHM.

(c) Average stroma size and trophoblast proliferation of CHM.

Figure D.12. Average stroma size and trophoblast proliferation of molar pregnancy images.
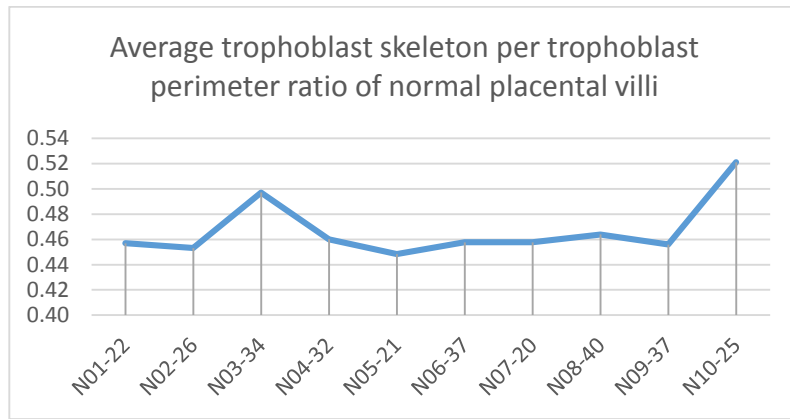
(a) Average trophoblast skeleton per trophoblast perimeter ratio of normal placental villi.
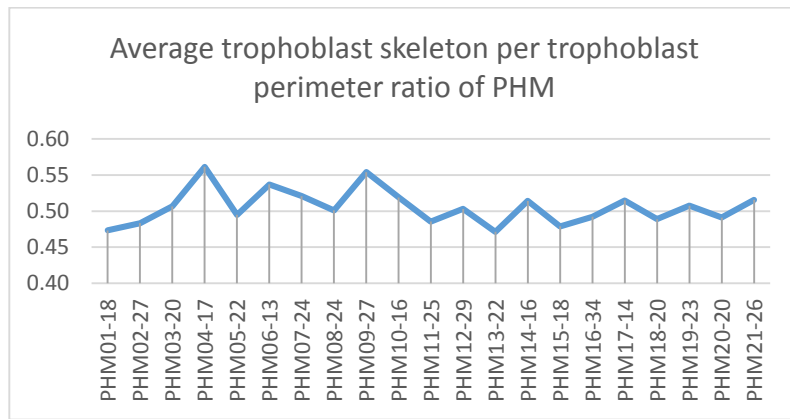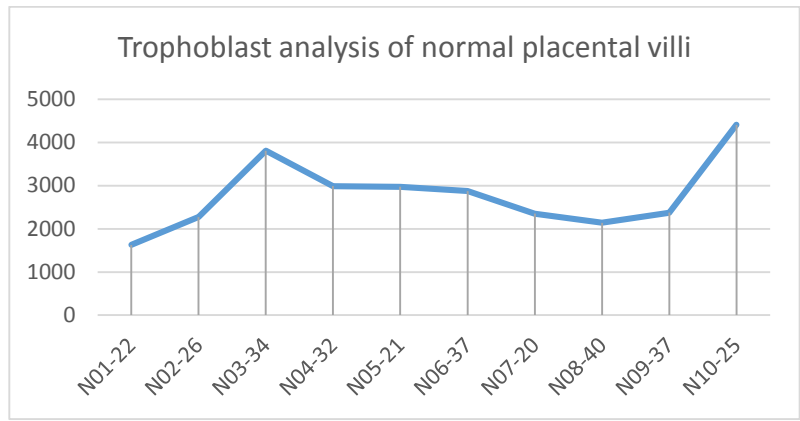

(b) Average trophoblast skeleton per trophoblast perimeter ratio of PHM.
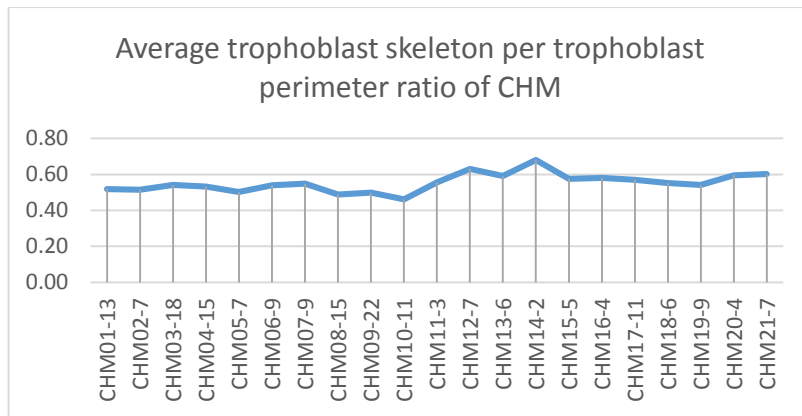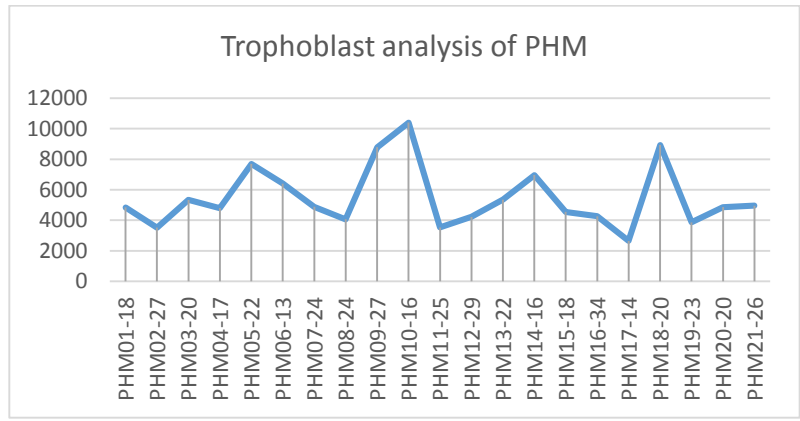

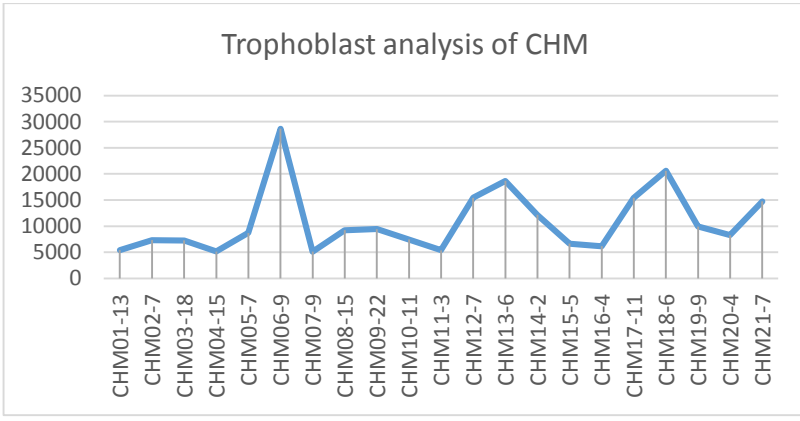(c) Average trophoblast skeleton per trophoblast perimeter ratio of CHM.

Figure D.13 Average trophoblast skeleton per trophoblast perimeter ratio of molar pregnancy images.

(a) Trophoblast analysis of normal placental villi.

(b) Trophoblast analysis of PHM.

(c) Trophoblast analysis of CHM.

Figure D.14 Trophoblast analysis of molar pregnancy images.

Appendix E. Publications

- Palee, P., Sharp, B., Noriega, L., Sebire, N. J., & Platt, C. (2013). Image analysis of histological features in molar pregnancies. *Expert Systems with Applications*, vol. 40, no. 17, pp. 7151-7158.

- Palee, P., Sharp, B., Noriega, L. & Pascal S. (2013) Hydatidiform Mole Analysis Tool based on Anomaly Detection, *International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*.