UNIVERSITY OF
LINCOLN

Automated Characterisation and Classification of Liver Lesions From CT Scans

Hussein Alahmer

Doctor of Philosophy

2018

School of Computer Science
University of Lincoln

Automated Characterisation and Classification of Liver Lesions From CT Scans

Hussein Alahmer

A thesis submitted in partial fulfilment of the requirements of
the university of Lincoln for the degree of Doctor of Philosophy

February 2018

Abstract

Cancer is a general term for a wide range of diseases that can affect any part of the body due to the rapid creation of abnormal cells that grow outside their normal boundaries. Liver cancer is one of the common diseases that cause the death of more than 600,000 each year. Early detection is important to diagnose and reduce the incidence of death. Examination of liver lesions is performed with various medical imaging modalities such as Ultrasound (US), Computer tomography (CT), and Magnetic resonance imaging (MRI). The improvements in medical imaging and image processing techniques have significantly enhanced the interpretation of medical images. Computer-Aided Diagnosis (CAD) systems based on these techniques play a vital role in the early detection of liver disease and hence reduce liver cancer death rate. Moreover, CAD systems can help physician, as a second opinion, in characterising lesions and making the diagnostic decision. Thus, CAD systems have become an important research area. Particularly, these systems can provide diagnostic assistance to doctors to improve overall diagnostic accuracy.

The traditional methods to characterise liver lesions and differentiate normal liver tissues from abnormal ones are largely dependent on the radiologists experience. Thus, CAD systems based on the image processing and artificial intelligence techniques gained a lot of attention, since they could provide constructive diagnosis suggestions to clinicians for decision making. The liver lesions are characterised through two ways: (1) Using a content-based image retrieval (CBIR) approach to assist the radiologist in liver lesions characterisation. (2) Calculating the high-level features that describe/characterise the liver lesion in a way that is interpreted by humans, particularly Radiologists/Clinicians, based on the hand-crafted/engineered computational features (low-level features) and learning process. However, the research gap is related to the high-level understanding and interpretation of the medical image contents from the low-level pixel analysis, based on mathematical processing and artificial intelligence methods. In our work, the research gap is bridged if a relation of image contents to medical meaning in analogy to radiologist understanding is established.

This thesis explores an automated system for the classification and characterisation of liver lesions in CT scans. Firstly, the liver is segmented automatically by using anatomic medical knowledge, histogram-based adaptive threshold and morphological operations. The lesions and vessels are then extracted from the segmented liver by

applying AFCM and Gaussian mixture model through a region growing process respectively. Secondly, the proposed framework categorises the high-level features into two groups; the first group is the high-level features that are extracted from the image contents such as (Lesion location, Lesion focality, Calcified, Scar, ...); the second group is the high-level features that are inferred from the low-level features through machine learning process to characterise the lesion such as (Lesion density, Lesion rim, Lesion composition, Lesion shape,...). The novel Multiple ROIs selection approach is proposed, in which regions are derived from generating abnormality level map based on intensity difference and the proximity distance for each voxel with respect to the normal liver tissue. Then, the association between low-level, high-level features and the appropriate ROI are derived by assigning the ability of each ROI to represents a set of lesion characteristics. Finally, a novel feature vector is built, based on high-level features, and fed into SVM for lesion classification. In contrast with most existing research, which uses low-level features only, the use of high-level features and characterisation helps in interpreting and explaining the diagnostic decision.

The methods are evaluated on a dataset containing 174 CT scans. The experimental results demonstrated that the efficacy of the proposed framework in the successful characterisation and classification of the liver lesions in CT scans. The achieved average accuracy was 95.56% for liver lesion characterisation. While the lesion's classification accuracy was 97.1% for the entire dataset.

The proposed framework is developed to provide a more robust and efficient lesion characterisation framework through comprehensions of the low-level features to generate semantic features. The use of high-level features (characterisation) helps in better interpretation of CT liver images. In addition, the difference-of-features using multiple ROIs were developed for robust capturing of lesion characteristics in a reliable way. This is in contrast to the current research trend of extracting the features from the lesion only and not paying much attention to the relation between lesion and surrounding area. The design of the liver lesion characterisation framework is based on the prior knowledge of the medical background to get a better and clear understanding of the liver lesion characteristics in medical CT images.

Declaration

I declare that the work presented in this thesis is my own work, except where otherwise mentioned and referenced in the text.

Acknowledgements

First and foremost, I would like to offer my sincere gratitude to my supervisors Dr Amr Ahmed and Dr Xujiong Ye for their insightful guidance, invaluable patience and supported me throughout my years at University of Lincoln. Without their encouragement, advice and motivation, this thesis would not have been completed. It has been my great honour to carry out the PhD research under their supervision. Sincere thank also go to my second supervisor Dr Tryphon Lambrou for his kind advice.

In addition to these, I need also to thank the King Hussein Medical Centre and all of its staff members for their help and support in my research. I am grateful to Anas Amayreh and Dr Walaa Arabiat for their kindness and support in collecting the dataset. Moreover, I would like to thank the ImageClef organiser for making the dataset available for the experiments. I would extend my thanks to the current and former members of the School of Computer Science administration team, for their support and kindness. I would also like to thank my colleagues in the School of Computer Science, especially to Saddam, Ashiqur Rahman and Casmir for their companionship in stressful times, and kindness during my PhD journey.

Last but not least, my love and gratitude also have to go to my Mum, Dad, Brothers and Sisters for almost unbelievable support. The words are not enough to express my gratitude for all your encouragement, patience and wisdom all these years that have guided me to come this far.

Ruba, my eternal love and soulmate, you have always been there with me in rain or shine throughout this journey, there is nothing I could have done without your love, patient, support and encouragement. Esam, my lovely son, thanks for giving me happiness and filling my life with hope.

And the one above all of us, I praise God, the Almighty for answering my prayers for granting me the capability to complete my thesis. Thank you so much, O Allah.

Contents

Acronyms

2D	Two-Dimension
3D	Three-Dimension
AC	Autocorrelation Coefficients
ACC	Accuracy
AFCM	Alternative Fuzzy C-Means
AFP	Alpha-Fetoprotein
ANFIS	Adaptive Neuro-Fuzzy Inference System
AUC	Area Under Curve
BOF	Bag-of-visual-Feature
BoVW	Bag-of-Visual-Word
CAD	Computer-Aided Diagnosis
CBCR	Content Base Case Retrieval
CBIR	Content-based image retrieval
CCL	Connected Component Labeling
CES	Clinical Experience Sharing
CLAHE	Contrast Limited Adaptive Histogram Equalization
CLEF	Cross Language Evaluation Forum
COM	Co-occurrence Matrices
CT	Computer Tomography
CVHE	Constrained Variable Histogram Equalization
DoF	Difference-of-Features

DSC	Dice Similarity Coefficient
DWT	Discrete Wavelet Transform
DWT	Discrete Wavelet Transform
FCM	Fuzzy C-Means
FDM	Fractal Dimension Measurements
FM	Fractal Model
FN	False Negative
FNH	Focal Nodular Hyperplasia
FOS	First Order Statistics
FP	False Positive
FPVF	False Positive Volume Fraction
GA	Genetic Algorithm
GLCM	Gray-Level Co-Occurrence Matrix
GLDM	Gray Level Difference Matrices
GLH	Gray Level Histogram
GLRLM	Gray Level Run Length Matrix
GM	Gradient Matrices
HCC	Hepatocellular Carcinoma
HOG	Histogram of Oriented Gradients
HU	Hounsfield Unit
HV	Hepatic Vein
IVC	Inferior Vena Cava
JSM	Jaccard Similarity Metric
KNN	K-Nearest Neighbour
LASSO	Least Absolute Shrinkage and Selection Operator
LBP	Local Binary Pattern
LDA	Linear Discriminant Analysis

LR	Logistic Regression
LTE	Law's Texture Energy
MI	Mutual Information
ML	Machine Learning
MRI	Magnetic Resonance Imaging
NB	Naive Bayes
NMR	Nuclear Magnetic Resonance
NN	Neural Network
NPV	Negative Predictive Value
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PPV	Positive Predictive Value
PV	Portal Vein
RAM	Random Access Memory
Ref.	Reference
RF	Radio Frequency
RLM	Run Length Matrices
ROC	Receiver Operating Characteristic
ROI	Region of Interest
RVD	Relative Volume Difference
Seg.	Segmentation
SGLDM	Spatial Gray Level Dependence Matrices
SN	Sensitivity
SNR	Signal-to-Noise-Ratio
SP	Specificity
SPECT	Single-Photon Emission Computerised Tomography
SVM	Support Vector Machine

TEM	Laws' Texture Energy Measures
TFCM	Texture Feature Coding Method
TN	True Negative
TP	True Positive
TPVF	True Positive Volume Fraction
US	Ultrasonography
VOE	Volumetric Overlap Error
WNN	Weighted Nearest-Neighbour

List of Figures

List of Tables

Chapter 1

Introduction

The liver is an important organ that performs vital functions such as detoxification of hormones, drugs, filters the blood from waste products and production of proteins required for blood clotting (Kumar et al., 2011). Therefore, liver diseases have to be considered seriously. However, diseases can occur without warning or symptoms. Cancer is a leading cause of death and liver cancer is the second leading cause of cancer death worldwide (Ferlay et al., 2015). The earliest possible detection of such a disease becomes critical to successful treatment. Therefore, Early detection will help to reduce the cancer death.

Medical imaging plays an important role in the diagnosis process through providing visual information of the body organs. There are various imaging modalities such as Computed tomography (CT) scan, Ultrasound, Positron Emission Tomography (PET), and Magnetic Resonance Imaging (MRI) used to diagnose liver lesions. The CT scan is often preferred for diagnosing liver diseases, as it is considered to be of higher accuracy and cheaper than MRI (Arakeri et al., 2011; Dankerl et al., 2013). However, doctors rely on their experience for the final diagnosis (Krupinski, 2011). The growing disparity of imaging protocols and producing a huge volume of image data daily is a challenge to image interpretation, even for experienced doctors (Andriole et al., 2011). As a consequence, errors and inconsistencies in interpretation of medical images now represent the weakest side of clinical imaging (Robinson, 1997; Berlin, 2007; Sabih et al., 2011).

Computer-aided classification and characterisation for interpreting images have been considered one of the major research subjects. Particularly, these systems can provide diagnostic assistance to clinicians for the improvement of diagnosis and increase the accuracy (Kumar et al., 2012). Hence, the success of medical image interpretation depends on two main processes: (1) visual perception by determining important visual patterns and (2) potential linking of image content, clinical context, and diagnostic decision which needs a deep understanding of clinical aspects of diseases (Tourassi et al., 2013). The majority of unsuccessful interpretation of medical images is due to misinterpretation of perceptual perception (Berlin, 1996; Taylor, 2017). This reflects

upon the importance of working on medical images interpretation for better characterising liver lesions that could support radiologists to describe lesions. Nevertheless, the automated lesion characterisation is a challenging task, due to the semantic annotation problem of mapping between high-level features (semantic annotation), low-level features and selecting the appropriate of the region of interest (ROI). This problem requires imposing new methods that could effectively address these challenges and improve the interpretation of medical images.

This thesis presents an efficient framework for characterising and classifying liver lesions based on CT images. For a better capturing of the lesion characteristics, the framework is proposing multiple ROIs (inside, border and surrounding lesion) to calculate the high-level features by considering the ability of each ROI that represents a set of lesion characteristics, which tackles the challenge of medical image interpretation. This is in contrast with traditional work that relied only on the lesion area itself, as a single ROI, to predict the characterisation without considering the relation between the lesion, margin and surrounding area. In addition, this work enhances lesion classification accuracy by using high-level features to classify the respective lesions and interpret the diagnosis decision. In contrast, the current traditional researches use low-level features only (black box) to classify liver lesions, which can not interpret the diagnosis decision.

1.1 Motivation

Liver cancer is one of the major death factors in the world. Global Cancer Statistics (Jemal et al., 2011) reported that, liver cancer was the fifth most commonly diagnosed and the second-leading cause of cancer death for the men. While in women, it is the seventh most frequently diagnosed and the sixth most common cause of cancer death, as presented in Figure 1.1. Moreover, incidence statistic's rate of liver cancer was increasing across many parts of the world where most patients who are diagnosed with liver cancer die within six months of diagnosis. Early detection will help to reduce the cancer death and becomes critical to successful treatment.

Using computer-aided systems for medical diagnosis field and treatment procedures is a rapidly growing research area at present. Computer-based interpretation and analysis of medical images help the radiologists to diagnose disease faster and more accurate. The CAD systems play an important role in the diagnosis processes related to liver lesions by providing doctors with fully automated or semi-automated computing tools to assist them in diagnosis and treatment tasks.

The ultimate motivation of this work is to characterise the liver lesions, by calculating the high-level features automatically. The high-level features are then used to enhance the classification accuracy. The use of high-level features in classification task helps in interpreting and explaining the diagnosis decision. Currently, there is

a significant amount of research devoted to segmentation of the liver, lesion detection and classification, while the liver lesion characterisation is still relatively new and challenging area. In addition, Most of the current classification-based techniques have proposed and/or optimised the hand-crafted features. The traditional classification systems give the diagnosis without decision interpretation, due to the low-level features lack the human perceptions and therefore cannot be directly interpreted by the humans. Hence, the importance of this work is to characterise the liver lesion automatically and using high-level features to classify the lesion with the advantage of interpreting the diagnosis decision. The characterisation approach gives more information about the lesion rather than simply classifying lesion into benign or malignant.

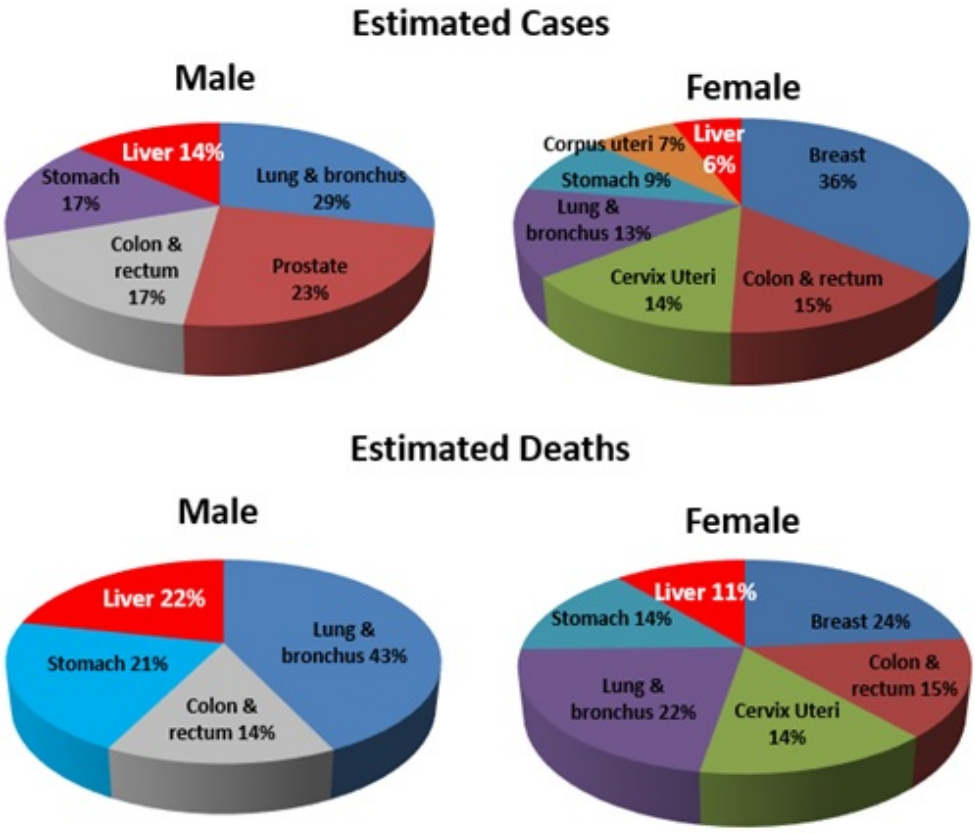


Figure 1.1: Estimated New Cancer Cases and Deaths Worldwide by Sex.

1.2 Aims and Objectives

More specifically, the main aim of this thesis is to address the liver lesion characterisation (mainly) based on CT images, ultimately improve the performance of the lesion classification. For characterisation of liver lesions in CT image, high-level features are needed to be calculated. To achieve this aim a novel characterisation framework is proposed. Furthermore, a novel feature vector is built, based on high-level features to improve the lesion classification performance, which provides the understanding explanation for classification through the lesion characterisation.

The objectives that helped to build the framework and fulfil this aim are, as follows:

- To investigate different liver lesion characterisation frameworks that characterise the lesion from CT images.
- To investigate different liver lesion classification frameworks that classify liver lesions into benign and malignant where the classification task is used as a case study to demonstrate how the characterisation could be used and useful in a realistic task (classification/diagnose).
- To develop a better ROI selection approach that captures all the lesion characteristics such as lesion morphology, brightness and patterns of contrast enhancement.
- To overcome the CT image challenge issues such as the variation of intensity and texture ranges between study cases due to the imaging devices settings such as images resolution and spacing.
- To develop an efficient and more accurate framework for characterising liver lesions by combining medical prior-knowledge and Multiple ROIs with Difference-of-features (DoF) technique to infer high-level features, and simulate radiological observations for liver characterisation.

1.3 Challenges

This section discusses the main challenges and issues that were faced and addressed in this thesis, which need to be overcome, to fulfil the work goals, as follows:

- The targeted liver lesion characterisation, as a high-level image understanding, is challenging. However, this thesis proposed a automatic lesion characterisation framework to link the image content to visual and logical/semantic features through mapping between quantitative imaging features, semantic features and ROI selection.
- The second major challenge in lesion characterisation/ classification is to improve the accuracy. Thus, selecting the appropriate region of interest to extract the lesion features is important to better characterisation/ classification performance. However, proposed Multiple ROIs area contribute different information about the lesion, and particularly surrounding area (outside), because the Malignant lesion affects the surrounding area differently compared to, the Benign lesion (Heiken, 2007). Utilising the features from inside, border, and outside lesion area supports in better differentiation between benign and malignant lesion.

- The third major challenge is the variation of the CT images that were collected from different institutes or machines which might use different settings in image acquisitions. However, the proposed Difference-of-Features (DoF) helps to overcome the challenge issues such as the variation of intensity and texture ranges between study cases due to the imaging devices settings such as images resolution and spacing.

1.4 Contributions

The primary contribution of this thesis is an effective liver lesion characterisation/ classification framework based on multiple ROIs and DoF that predicts radiological observations related to the characteristics of the lesion. To achieve this, the technical contributions are presented below:

- Characterising liver lesions with a large number of high-level features compared to the related work. The high-level features are categorised into two groups; the first group is the visual features from the image contents; the second group is visual and logical/semantic features inferred from low-level features.
- Proposing Multiple ROIs selection approach through generating abnormality level map based on the intensity difference and the proximity distance for each voxel with respect to the normal liver tissue. As well as, capturing the lesion characteristics through multiple ROIs/ DoF to calculate the high-level features by assigning the ability of each ROI to represent a set of lesion characteristics such as assigning the surrounding lesion ROI to characterise lesion rim.
- Mapping between low-level, high-level features and ROI selection to build a robust characterisation framework for interpretable images to infer semantic features and making a diagnostic decision with a human-interpretable explanation.

The proposed framework also has another benefit to enhance lesion classification performance through utilising the high-level features to build a novel feature vector, the use of lesion characterisation helps in interpreting and explaining the classification and is more intuitive to clinicians.

1.5 Thesis Overview

The rest of this thesis is organised as follows:

- Chapter 2 provides an overview of the medical background. This chapter gives the necessary information about liver and liver anatomy. It also presents the diagnostic imaging and liver diseases types. Furthermore, presents the datasets

and also elaborates all the metrics that were used to evaluate and validate the proposed framework.

- Chapter 3 discusses review of key related literature work. The previous studies are divided into two main categories: (1) liver lesion classification and (2) liver lesion characterisation.

- Chapter 4 presents a general overview of liver lesion characterisation pre-processing, with details about liver segmentation, lesion detection and vessels extraction proposed method.

- Chapter 5 presents the proposed framework for liver lesion characterisation/classification, with details of its components, the methods and the techniques that were adopted.

- Chapter 6 presents and discusses the experimental setup, extensive and comparative evaluation, as well as benchmarking of the proposed framework against related work. In addition, the statistical analysis is adopted to confirm the significance of the proposed framework.

- Chapter 7 Concludes the thesis work, and presents some potential future work.

1.6 List of Publications

Published paper:

- Hussein Alahmer, and Amr Ahmed. "Computer-aided classification of liver lesions using contrasting features difference." In: ICMISC 2015 : 17th International Conference on Medical Image and Signal Computing, International Science Index, Biomedical and Biological Engineering , 2(11), pp.398-405, 2015.

- Hussein Alahmer, and Amr Ahmed. "Hierarchical classification of liver tumor from CT images based on difference-of-features (DOF)." International Conference of Signal and Image Engineering, InProceedings of the World Congress on Engineering. pp.490-495, 2016. **Best Student Paper Award WCE2016.**

- Hussein Alahmer, Amr Ahmed. "Computer-aided Classification of Liver Lesions from CT Images Based on Multiple ROI." In Procedia Computer Science, Volume 90, pp.80-86, ISSN 1877-0509, 2016.

Under Review Work:

- Hussein Alahmer, Amr Ahmed, and Xujiong Ye. "Automatic Liver Lesion Characterisation: From Low-level Features to High-level Features in CT Scans". journal of IEEE Transactions on Biomedical Engineering.

- Hussein Alahmer, Amr Ahmed, and Xujiong Ye. "Liver Lesion Classification through characterisation and comparison with low-level features in CT". journal of Expert Systems With Applications.
- Hussein Alahmer, Amr Ahmed, and Xujiong Ye. "Liver Lesion Classification through Combination of Multiple ROIs and Difference-of-Features". journal of Computerized Medical Imaging and Graphics.

Chapter 2

Research Background

This chapter presents the medical background related to the liver with an emphasis on current facts. Section 2.1 displays general information about liver anatomies such as location, internal structure and functions. Section 2.2 provides a short descriptions about diagnostic imaging with a brief discussion about the imaging modalities used in our work, while liver diseases types are presented in Section 2.3. Section 2.4 presents the datasets and introduces the evaluation metrics that will be used to evaluate the work. Finally, the chapter is concluded in Section 2.5.

2.1 Liver anatomy

The liver is the largest gland in the body. Weighting in the adult men 1.4 to 1.5 kg and 1.2 to 1.4 kg in the adult women (Wolf, 1990), it sits mainly on the upper right of the abdomen and rests just under the diaphragm, and to the right of the stomach, intestine, spleen and overlying the gallbladder (Ger, 1989; Stringer, 2014), as shown in Figure 2.1.

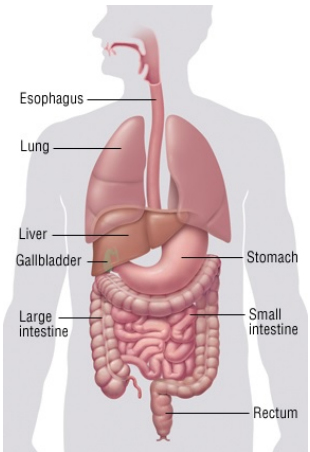


Figure 2.1: Liver location in the body.(School, 2014)

The liver performs hundreds of vital life functions (Mitra and Metcalf, 2012) such as detoxification of hormones, drugs, filter the blood from waste products, production

of proteins required for blood clotting, metabolism of carbohydrates, synthesis, breakdown and storage (as glycogen); production of chemicals (bile) that aid in digestion.

The liver anatomy can be described using two different aspects; anatomically from anterior view, the liver is divided into two portions, a right and a left lobe, as shown in Figure 2.2 (a). This division, however, is not used surgically because it is not showing the internal features of vessels and biliary ducts, which are very important in liver surgery. Functionally based on Couinaud classification, the liver is divided into eight functionally independent segments, as shown in Figure 2.2 (b). Each segment has its own vascular inflow, outflow and biliary drainage (Bismuth, 2013).

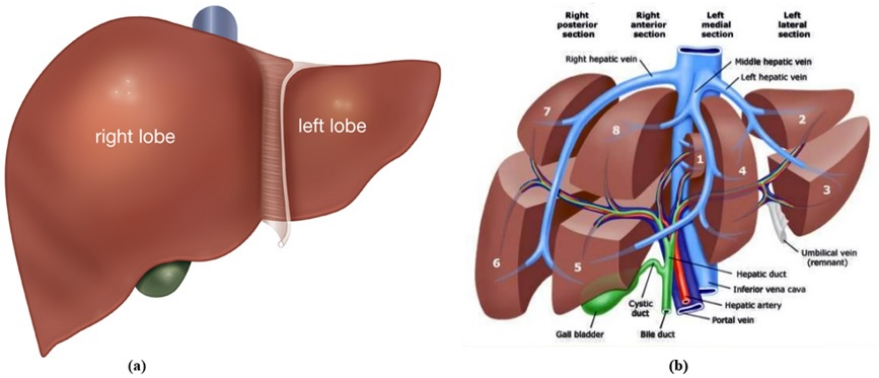


Figure 2.2: Liver; (a) liver anterior view; (b) segmental anatomy according to Couinaud (Farid et al., 2013).

The liver contains four vessel structures: portal vein, hepatic vein, hepatic arteries and bile duct. The liver is connected to two main blood supply. 80% of the blood rich in nutrients is derived from the portal vein and the remaining percentage of blood through hepatic artery, rich in oxygen (Kan and Madoff, 2008). The right hepatic vein divides the right lobe into anterior and posterior segments. The middle hepatic vein divides the liver into right and left lobes. Left hepatic vein divides the left lobe into medial and lateral part. Portal vein divides the liver into upper and lower segment. Hereby, the arrangement of these vessels used in dividing liver into eight functionally independent segments. (Kang et al., 2014a)

2.2 Medical imaging modalities

Liver imaging is important for lesion detection and characterisation, especially monitoring the patients with a history of disease to evaluate treatment response. The liver is considered as a site of metastasis spread from lung, stomach, colon and pancreas (Sahani and Kalva, 2004a). The liver pathology examination is performed through various medical imaging modalities such as US, MRI, PET, CT, etc. The imaging modalities can be generally categorised into two groups: ionising radiation group which use either X-rays or gamma rays and non-ionising radiation that use of sound waves or radio

wave combined with magnetic fields, as illustrated in Figure 2.3. Depending on the specific application requirements, one or more of these imaging methods can be used. A brief review of the most commonly used imaging modalities for liver lesions diagnosis, which can be detected and characterised by these methods.

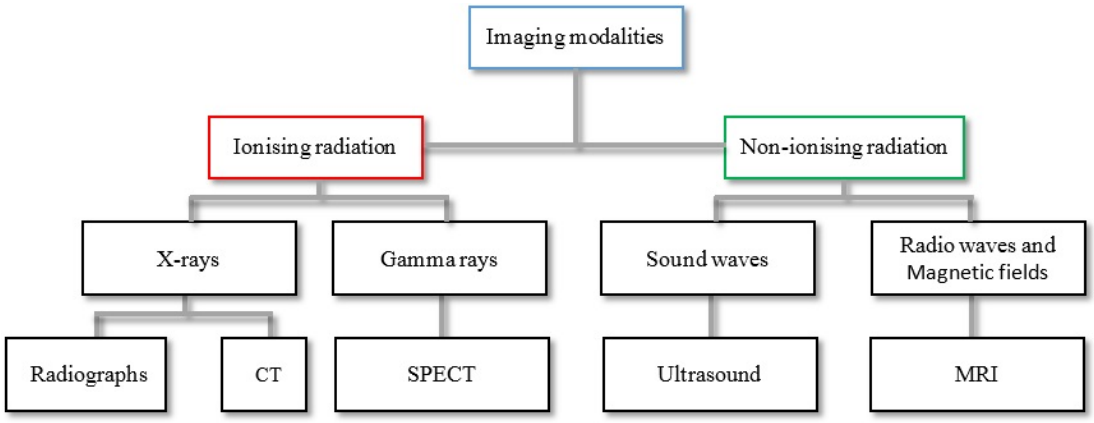


Figure 2.3: Medical imaging modalities, categorised according to ionising and non-ionising radiation, with further specific categorisation for each category.

2.2.1 Ultrasound Imagining

The US device consist of three parts: (1) console comprising a computer, controls, keyboard, disk storage and printer. (2) Video display monitor. (3) The transducer, which is a small hand-held device attached to the US device through a cord, that used for scanning, as displayed Figure 2.4. The US uses high-frequency sound waves upper the human hearing limit ($> 9 - 18$ MHz) (Afonso, 2015) to create a live image for internal organs.

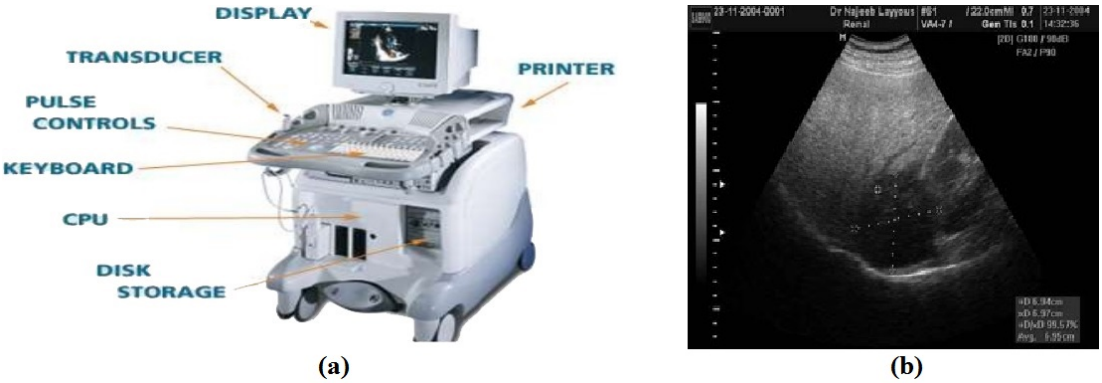


Figure 2.4: Ultrasound. (a) Ultrasound device. (b) Sample of ultrasound image.

The US imaging is a painless, safe, inexpensive procedure and easily available. Moreover, US imaging is very sensitive in differentiating between a cystic lesion and a solid lesion (Toriyabe et al., 1997) and assessing the relation between lesions and hepatic vessels(Schmidt et al., 2000; Catheline et al., 2000). Despite of that, the CT and MRI are considered as a higher sensitive for detecting and characterising focal and

solid lesion than US (Glover et al., 2002; Toriyabe et al., 1997; Nam et al., 2011), where the US sensitivity for liver metastases detection ranges between 40% and 70%(Paulson, 2001). In contrary, the US suffers from low specificity, and inability to detect small lesions that less than 1 cm (Dhar et al., 2016).

2.2.2 MRI imaging

MRI is a 3D imaging technique, which is used to generate images of the internal organs in the body for diagnosis. MRI produce high-quality images by utilising the science of nuclear magnetic resonance (NMR)(Pichler et al., 2008). The patient is placed on a sliding table that enters into a tunnel with a strong magnetic field around the organs to be diagnosed. The signal-to-noise-ratio (SNR) and quality of the image are determined by the strength of the magnetic field where the current systems utilised either 1.5T or 3T field strength (Han and Talavage, 2011), due to higher field strengths provide a greater SNR and theoretically improve spectral resolution; however, potential limiting factors include the increased field inhomogeneity leading to increased metabolite line widths.

The human body is mostly water. Water molecules (H2O) contain hydrogen nuclei (protons), which become aligned in a magnetic field. The hydrogen atoms are excited with radio frequency (RF) electromagnetic pulse, which makes the hydrogen alignment appear in a heterogeneous manner. When excitation by radio frequency is finished, all the hydrogen atoms come back into homogeneous alignment again while the released of RF reflects the characteristic of proton's chemical environment (Keevil, 2001). The generated radio frequency signal from hydrogen atoms is received by coil to produce images in frequency, or *k*-space. The inverse Fourier transform of the signal (*k*-space) constructs a 2D or 3D image (Tsao et al., 2003).

A contrast agent such as gadolinium-based is used with MRI scan to enhancing the detail of the fabric and making vascular tissue more visible through changing the magnetic properties of the blood as shown in Figure 2.5. MRI scan is, in general, a safe modality without exposing the human body to ionising radiation (Schenck, 2000). Moreover, MRI provides a high soft tissue differentiation (Aisen et al., 1986). Despite of that, MRI is not always preferable especially when another imaging modality could produce the same information. in spite of that, MRI images usually affected by various artefacts due to its sensitivity to the noise, making it very challenging to generate MRI images. Furthermore, MRI equipment is costly and has high maintenance costs, time-consuming, and claustrophobia-exacerbating (van Herk and Kooy, 1994; Wintermark et al., 2007).

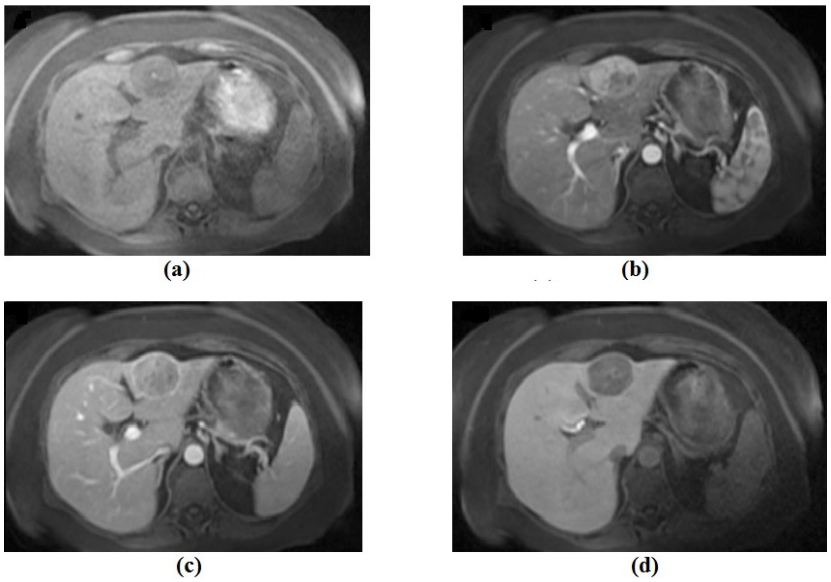


Figure 2.5: Example of MRI image Axial fat-suppressed T1-weighted with different phases. (a) unenhanced phase. (b) artial phase. (c) portal phase. (d) late phase at 20-min.

2.2.3 CT scan imaging

CT is one of most widely used 3D imaging technique, which can be used to produce horizontal or axial tomographic images through the combination of X-ray and computer technology. It shows detailed image of any part of the human body including fat, muscles, internal organs and bones (Aisen et al., 1986). To perform the CT scan, the patient lies on a table slides through the centre of a X-ray tube, which contains rotating X-ray beam and detectors. The X-ray beams circle around the human body, emits the X-ray that passes through the specific part of the body from different angle to generate different views for the same organ. The attenuation of the X-ray beams measure by the detector, then the computer interprets the data that is received from the CT (Hu, 1999; Hsieh, 2003a). In CT image, the proportion of radiation attenuation absorbed by the tissues represents the density of the organs that is called Hounsfield Unit (HU) (Razi et al., 2014). The corresponding HU value is calculated by Equation 2.1.

$$HU_x = 1000 \times \frac{\mu_x - \mu_{water}}{\mu_{air} - \mu_{water}} \tag{2.1}$$

Where μ_{water} and μ_{air} are respectively the linear attenuation coefficients of water and air . In a voxel x with average linear attenuation coefficient μ presented as μ_x. The intensities range of CT image in HU from -1000 to +1000 (-1024 to +1024 or +3072, depending on the particular vendor's encoding). The different organs presented as different voxels in the image which led to different HU values. The HU of tissue intensity is based on two values: air has a minimum HU value as -1000 HU and water as 0 HU. (Hsieh, 2003b)

Usually contrast agents such as iodine are taken with CT scan imaging by injec-

tion before the image acquisition. The benefit of contrast enhanced image is to find the pathology by enhancing the contrast between a tumour and the normal tissue surrounding structure. After injecting the contrast, images are acquired after varied time lengths in order to visualise the dynamic of the blood flow within tissues. These multiple images enhance the detection and characterisation of their tumours.

In clinical practice, the images are divided into four phases based on the time duration from taking the injection (Roy et al., 2014a). The first phase is the non-enhanced image. The second is arterial phase after 25-40 seconds from succeeding the injection. The third is portal phase after about 60-75 seconds. The final is delayed phase take place after 3-5 minutes taking the contrast agent. Figure 2.6 depicts sample CT images in four phases.

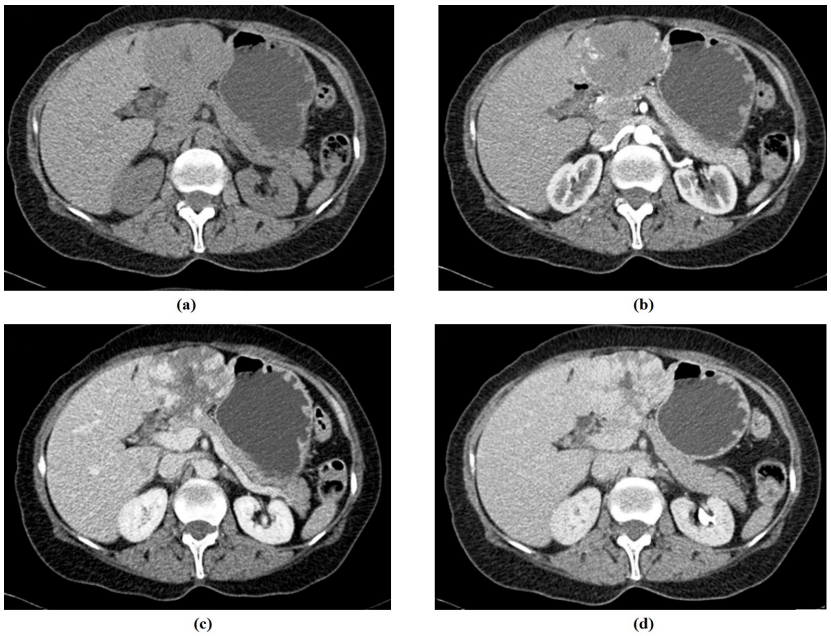


Figure 2.6: Example of CT scan image in four phases. (a) CT image before contrast injection. (b) arterial phase. (c) portal phase. (d) delay phase

2.2.4 CT, MRI, and US comparison

CT scan, MRI and US are commonly used in oncology patients where each have various advantages and limitations.The CT scan can offer high resolution images with short acquisition time which make it preferable in emergency situation and trauma cases. Moreover, it provides a good contrast between bone, tissue and vessels at the same time. However, MRI is more sensitive and provides much more soft tissue details than CT. Further to this, the ability of MRI to change the image contrast by changing in magnetic field and radio wave to highlight different types of tissue. On the other side, US is safe, readily available and cheap imaging compared to CT and MRI. However, US relies on the skill and expertise of the radiologist and equipment quality. Table 2.1 depicts the comparison between the three imaging modalities of CT, MRI and US.

Points	CT scan	MRI	US
Cost (Wintermark et al., 2007)	Cheaper than MRI	More expensive	Cheapest one
Time required (van Herk and Kooy, 1994; Jung et al., 2009)	Quick compared to MRI and US (Less than 10 minutes with less than 30 second for actual scan).	Required a long time (15 minutes up to 2 hours)	More than CT and less than MRI
Imaging techniques (Pichler et al., 2008)	X-ray	Strong magnetic field and radio frequency (RF) electromagnetic pulse	High frequency sound waves
Radiation (Pichler et al., 2008)	Ionising radiation	Non-ionising radiation	Non-ionising radiation
Risk of contrast Agent (Schenck, 2000; Feinstein et al., 2010)	Rare allergic reaction but more common compared to MRI. Risk in cases of renal insufficiency and dehydration.	Very rare allergic. The risk of reaction in patients have a history of kidney or liver disorders.	Safe, with a low incidence of side effects compared to CT and MRI.
Comfort (Wippold, 2007; Peris et al., 2010)	Rarely causes claustrophobia	* Anxiety, especially patients have claustrophobia. * Discomfort because the patients must to stay on a hard table for a long time.	Comfort
Limitations (Schenck, 2000; Cohen and Afdhal, 2010)	Not fit for overweight patients (more than 200 kg)	* Not fit for overweight patients (more than 160 kg). * Not suitable for Patients with Cardiac Pacemakers or having tattoos and metal implants.	* Disrupted by air or gas. * Operator dependency.
Preferred for (Aisen et al., 1986; Wintermark et al., 2007; Cohen and Afdhal, 2010)	Bone injuries, cancer detection, biopsy, lung and chest and emergency.	Soft tissue evaluation, spinal cord injury, brain tumours and tendon injury.	Check the organs in the abdominal cavity, examine the blood flow, and examine foetuses
Details of soft tissues (Aisen et al., 1986; Afonso, 2015)	Showing bone, tissues and vessels at same time	Providing more details than CT	More noises
Details of bone (Aisen et al., 1986; Cohen and Afdhal, 2010)	Providing much details than MRI	Less details	Usually not used for bony structures

Table 2.1: Comparison between CT scan, MRI and Ultrasound.

2.3 Liver lesion types

A tumour is an abnormal growth of cells or tissues. There are different types of liver tumours, as illustrated in Figure 2.7. Some tumours are benign (non-cancerous) such as Cyst, Hemangioma, while others are malignant (cancerous) such as Hepatocellular carcinoma (HCC), metastasis. There are several criteria used to diagnose liver lesion. First, blood tests: several blood tests can be done to check the liver problem such as test levels of alpha-fetoprotein (AFP). High levels of this protein could be a sign of liver cancer. Second, medical imaging such as CT, MRI are used mainly for diagnosis. These imaging techniques allow doctors to see inside the body without surgery. Imaging techniques frequently employ contrast agents to improve image resolution and enhance pathology detection. However, liver biopsy is considered the gold-standard diagnostic method to identify the liver cancer. The biopsy is performed by inserting a thin needle through the skin to obtain a small sample of liver tissue. (Rockey et al., 2009a)

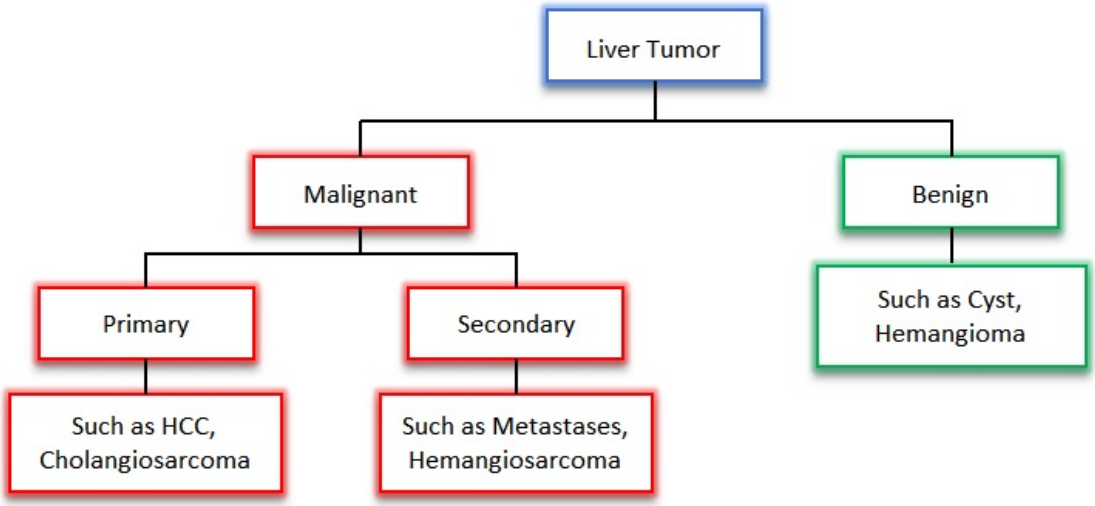


Figure 2.7: Liver lesion types.

Liver biopsy is used when liver diagnosis is not clear. However, biopsy is usually not recommended due to the risk of bleeding from the site of needle entry point or having some cancer cells follow the pathway of the needle and spread outside of the liver. Typically, a liver biopsy is performed by an expert radiologist using CT scan or US to help guide the biopsy needle.

2.3.1 Benign tumours

There are various types of benign tumours in the liver as presented in Table 2.2 (Fergusson, 2012a; Assy et al., 2009a; Suh et al., 2011; Grazioli et al., 2010), all of which are abnormal growths cells or structures. However, benign tumours do not spread to other organs. Benign tumours are quite common and usually do not produce symptoms and, mostly need no treatment. However, treatment of benign lesions depends on

the size, location, and symptoms of the lesions. Generally, some benign lesion may produce symptoms such as abdominal pain or have a huge size may effect on liver function required a treatment which often requires surgical intervention to remove the tumour.

Hemangiomas

Hepatic hemangioma lesion is the most common non-cancerous, solid tumour. Usually seen in women in the right hepatic lobe with no specific age range but rarely occur under adulthood (Glinkova et al., 2004). Pathologically, hemangioma consists of many tangles of blood vessels appear as blood-filled spaces, may contains clot, calcification and scar (Dilsiz et al., 2009). Overall, Hemangioma may be found isolated or multiple with diameter less than 5 cm to giant lesion (10-20 cm)(Maruyama et al., 2013; Erdogan et al., 2007). The small lesion not causing symptoms or risk developing cancer or requiring treatment but a large lesion is more likely to cause symptoms (abdominal pain) and may be needed to eradication the lesion.

Hepatic Adenoma

Hepatic adenoma is one of the rare solitary benign lesion and mostly identified in young women which is correlated with increased estrogen levels especially oral contraceptives or taking higher potency hormones such as anabolic steroids or glycogen storage disease (Dhingra and Fiel, 2014; Furlan et al., 2008). The hepatic adenoma affects architecture of normal liver cells and shows the arrangement of infected cells as sheet and cords. This type of lesions causes pain in the upper right of the abdomen. Typically, the lesion diameter is 5-10 cm and surrounded by a capsule from infected cells (Grazioli et al., 2001). The treatment, in this case, is necessary by surgically removed, due to the risk of malignant transformation or bleeding (Socas et al., 2005).

Focal Nodular Hyperplasia (FNH)

The FNH lesion is second most prevalent benign after hepatic hemangioma and mostly occurred in young women similar to hepatic adenoma, due to taking oral contraceptives (Navarro et al., 2014; Venturi et al., 2007). It usually causes pain in the upper right of the abdomen but rarely grows or bleeds and has no risk of malignant transformation. The lesion diameter less than 5 cm and often located close to the liver surface (Halankar et al., 2012). The FNH contains central scar which is differentiated from hepatic adenoma. The lesion is diagnosed confidently by imaging modalities such as CT scan and MRI, but it is often subtle on US (Grazioli et al., 2012).

Hepatic cysts

The hepatic cyst is one of the most common benign lesion. This lesion is sacs filled with fluids that occur in the liver with generally less than 3 cm diameter and could be isolated or multiple lesion (Bakoyiannis et al., 2013; Miliadis et al., 2010). Typically,

it is asymptomatic, no affect liver function and has no risk malignant potential (Schiff, 2009).

Benign lesion	Outline
Hemangiomas	• The most common benign solid lesions of the liver. • common in the right lobe of the liver. • diameter less than 5 cm to giant lesion (10-20 cm) • Rare spontaneous rupture (bleeding) (Glinkova et al., 2004; Vilgrain et al., 2000)
Hepatic adenomas	• Benign solid neoplasms of the liver • Most commonly seen in young women. • Typically solitary, although multiple adenomas also can occur. (Socas et al., 2005)
Focal nodular hyperplasia (FNH)	• FNH is the second most common benign tumour of the liver. • More common in women (similar to adenomas) • Usually do not rupture spontaneously. • has central scar (Grazioli et al., 2010)
Hepatic cysts	• Fluid-filled structures of the liver (Bakoyiannis et al., 2013). • Different types of hepatic cysts include (Zhang et al., 2009b) : Simple liver cysts ;Biliary cysts ;Parasitic cysts ;Cystadenomas

Table 2.2: Types of benign tumour.

2.3.2 Malignant tumours

Liver cancer is a type of malignant tumour. There are two types of malignant (Reynolds, 2001). First, primary liver cancer, where tumours have originated in the liver. Second, metastatic secondary liver cancer, where tumours spread from cancer sites elsewhere

in the body such as lung, colon, and breast. As listed in Table 2.3 (Moreira, 2015; Garrean et al., 2008; Robinson, 2000; Al-Salem, 2014; Reynolds, 2001)

Malignant tumour in early stage often does not cause symptoms until it has reached an advanced stage. So many cases are diagnosed fairly late. Liver cancer treatment is determined by tumour characteristics such as number, size, and location in the liver; tumour effect on liver functions or cause cirrhosis and/or tumour has spread outside the liver.

Malignant lesion	Outline
Hepatocellular carcinoma (HCC)	• Most common type of primary liver cancer. • 75% people with cancer have HCC disease. • HCC can start as a single tumour or as multiple tumours through the liver. • Multiple sites are more common than single tumour. (Davis et al., 2008; Zviniene, 2012)
Cholangiocarcinoma	• Start in the bile ducts of the liver. • 10-20% people have liver cancer diagnosed as Cholangiocarcinoma. (Iwatsuki et al., 1998)
Hemangiosarcoma	• Rare type of liver cancer; Begins in the blood vessels of the liver; • Grows quickly. • Diagnosed fairly late. (Rademaker et al., 2000a)
Hepatoblastoma	• Very rare type of liver cancer. • Usually seen in children under 4 year's age. (Schnater et al., 2003; Faraj et al., 2008)

Table 2.3: Types of malignant tumour.

Hepatocellular carcinoma (HCC)

Hepatocellular carcinoma (HCC), also called malignant hepatoma is a primary common liver cancer. Most cases of HCC are due to either a viral hepatitis infection or cirrhosis (alcoholism being the most common cause of hepatic cirrhosis) (Davis et al., 2008). HCC usually common in men more than women especially as old ages and

could be seen focal, multifocal or diffuse. US, CT and MRI are used to diagnose HCC but the CT and MRI is better than US in detection. However, in some cases, the biopsy is required for diagnostic accuracy.

Cholangiocarcinoma

Cholangiosarcoma is a second common primary malignant tumour of bile duct that accounts approximately 10-20% of all primary liver cancer (Reddy and Faust, 2005; Khan et al., 2005). The Cholangiocarcinoma is characterised on CT and MRI as irregular shape with peripheral enhancement in portal phase (Blechacz and Gores, 2008).

Hemangiosarcoma

Hemangiosarcoma is a rare metastases malignant liver lesion that usually occurs with no symptoms in elderly men (Rademaker et al., 2000b). The Hemangiosarcoma originates in spleen and often spreads to the liver and lung. The most important of hemangiosarcoma characteristic is an aggressive lesion, grows quickly and effects on blood vessels (Shoemaker et al., 2016).

Hepatoblastoma

Hepatoblastoma is another type of malignancy tumour of fetal hepatocytes, which is the most frequent in children and infants (Schnater et al., 2003). Usually, the right lobe is affected more than left lobe with no symptoms (Davenport et al., 2012). The treatment is made by Surgical tumour removal or liver transplantation (Ang et al., 2007).

The next section will introduce our dataset, experimental setup and evaluation matrices that used in this thesis.

2.4 Dataset and Evaluation Measurements

In this section, we will first present the datasets used in this thesis. Furthermore, we introduces the evaluation measurements that will be used to evaluate the proposed framework for liver segmentation, lesion detection, vessels extraction and also liver lesion classification/characterisation.

2.4.1 Datasets

In order to provide a rigorous evaluation and benchmarking of the proposed framework, the following two datasets (patients anonymised) were used during experiments:

1. Dataset I, obtained from ImageClef (Marvasti et al., 2015).

2. Dataset II, obtained from King Hussein Medical Center, Amman, Jordan.

The overall dataset comprises of two datasets, collected from two different institutions with a total number of 174 CT scan images and divided into 80 cases malignant (34 case HCC and 46 case Metastases) and 94 cases benign (56 case Cysts and 38 case Haemangiomas), as presented in Figure 2.8.

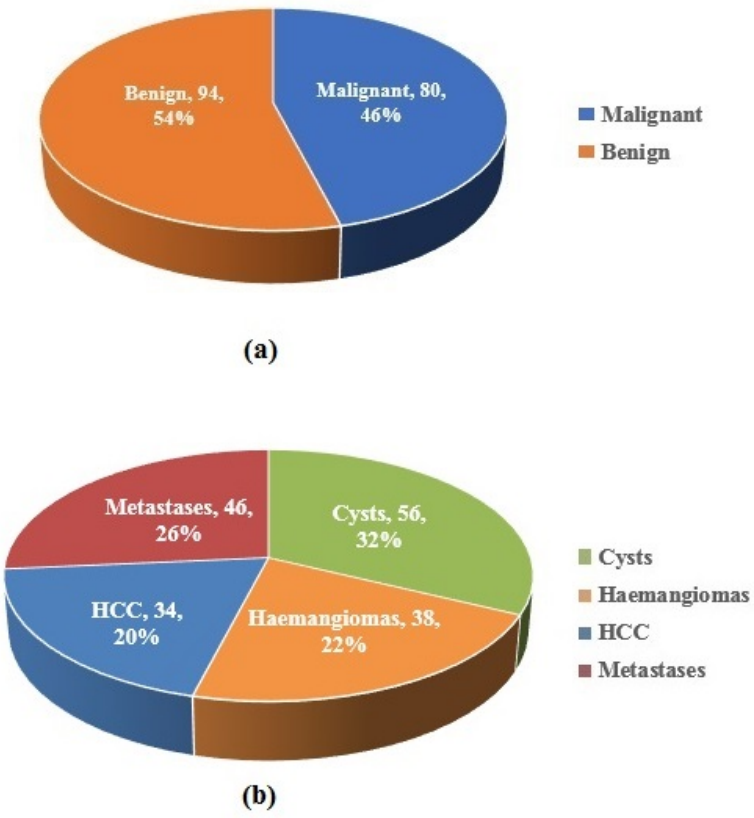


Figure 2.8: The overview of the dataset size; (a) Split dataset based on the lesion category (Malignant/ Benign); (b) Split dataset based on the lesion types (Cysts, Haemangiomas, HCC, Metastases).

2.4.1.1 Dataset I

The dataset I is given in ImageClef 2014 for liver CT annotation task (Marvasti et al., 2015). a 50 3D abdominal CT scans acquired from 46 patient and divided into 29 cases malignant and 21 cases benign with total number of lesions 137. Among the 50 CT scans, 98 lesion is malignant, while 39 lesion is benign. The dataset I contains four common types of liver lesions; two types are benign (14 case Cysts and 7 cases Haemangiomas) and two types are malignant (13 case HCC and 16 case Metastases) , as illustrated in Figure 2.9.

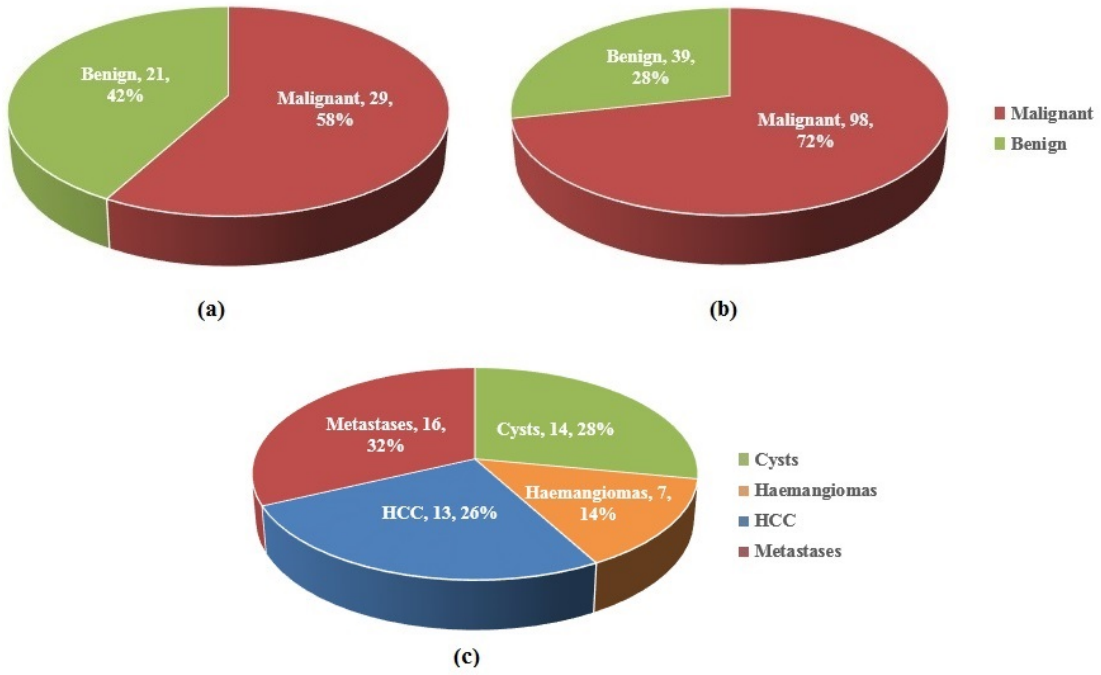


Figure 2.9: The overview of the dataset I; (a) Split dataset based on the lesion category (Malignant/ Benign); (b) Split dataset based on the number of the lesions for each of benign and malignant; (c) Split dataset based on the lesion types (Cysts, Haemangiomas, HCC, Metastases).

The CT images had varied resolutions (x: 190-308, y: 213-387) voxels in plane and contain between 41 and 588 slices depending on the field-of-view and the slice thickness. The slice thickness between 0.5 and 5 mm and the slice spacing varies from 0.674 to 1.007 mm.

The reference segmentation are available for the liver and lesions by ImageClef. In addition, the liver lesion characterisation are provided by a radiologist on each scan. The RadLex ontology (Mejino Jr et al., 2008) was used to characterise of the high-level features (semantic features) for the lesion annotation.

2.4.1.2 Dataset II

The dataset II is obtained from King Hussein Medical Centre, Amman, Jordan and the patient details is anonymised. The total number of lesions 328 obtained from 124 CT scan, acquired from 115 patient. Among the 328 lesions, 117 were benign while 211 were malignant. In overall, the 124 CT case divided into 51 case malignant and 73 case benign. The dataset II contains four common types of liver lesions; two types are benign (42 case Cysts and 31 case Haemangiomas) and two types are malignant (21 case HCC and 30 case Metastases) , as illustrated in Figure 2.10.

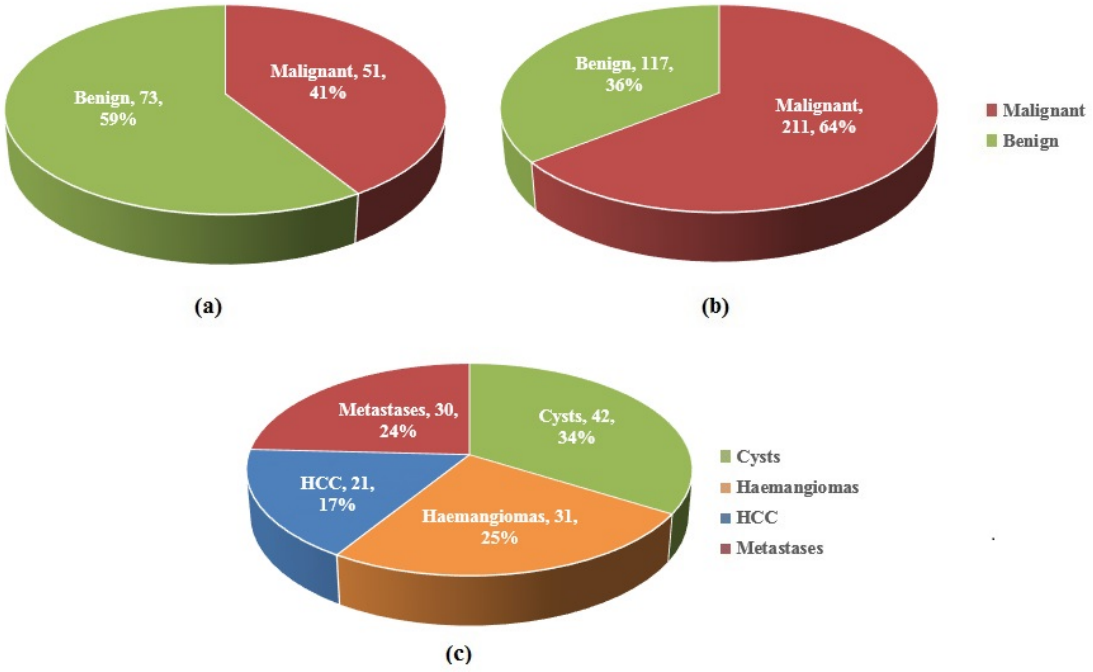


Figure 2.10: The overview of the dataset II; (a) Split dataset based on the lesion category (Malignant/ Benign); (b) Split dataset based on the number of the lesions for each of benign and malignant; (c) Split dataset based on the lesion types (Cysts, Haemangiomas, HCC, Metastases).

The dataset was composed of 76 men and 39 women who ranged in age from 24 to 86 years with mean age of 53.7 years. All scans were acquired between 2009 and 2016. The CT scan images have a resolution of 512x512 voxels in plane and contain between 159 and 482 slices depending on the field-of-view and the slice thickness. The slice thickness between 2 and 5 mm and the slice spacing varies from 0.65625 to 1 mm and the voxel size range from 0.6836 to 0.8789 mm.

All of the CT scans were acquired in triple phase. On each CT scan, the liver lesion detection and characterisation was provided by three experts radiologist, with more than 15 years experience based on RadLex ontology.

2.4.2 Evaluation Measures

This section will present in details all the evaluation measurements that used during the evaluation the proposed framework in both side: liver segmentation, lesion detection, vessels extraction and liver lesion classification/characterisation. The next part of the evaluation measurements section presents the different evaluation methods which are used to assess the accuracy of the segmentation.

2.4.2.1 Segmentation Evaluation

This section presents an evaluation of the liver segmentation, lesion detection and vessels extraction for the proposed framework. To measure the accuracy of the proposed framework in segmentation, all the the evaluation methods that mentioned in this section are used.

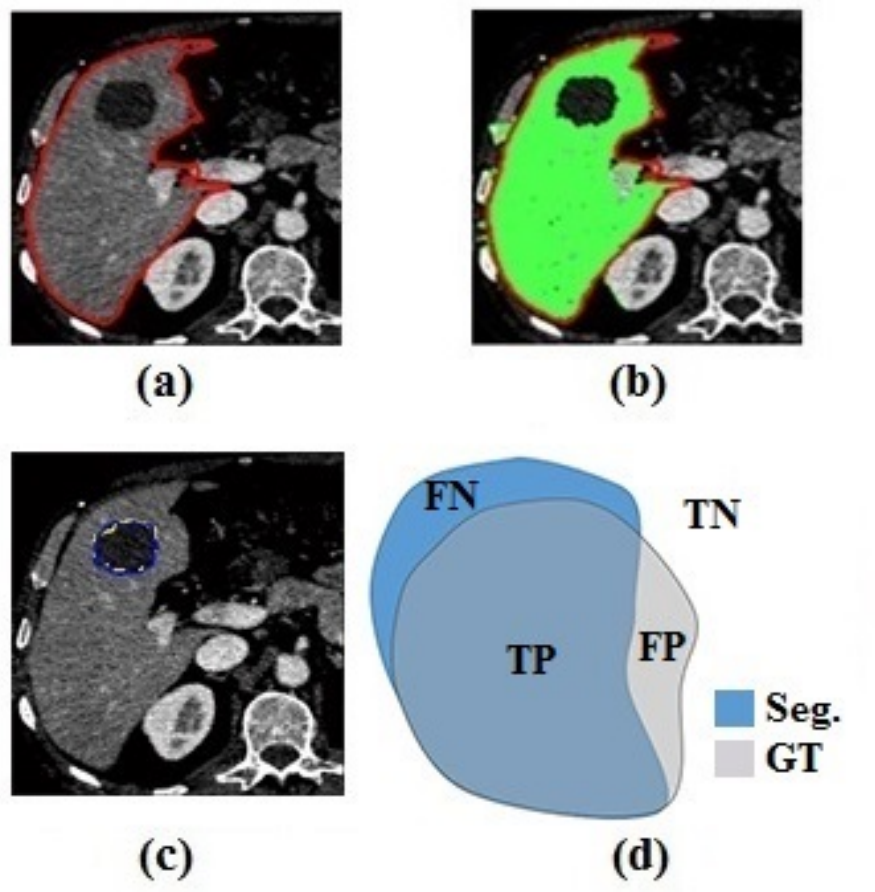


Figure 2.11: The evaluation of liver/lesion segmentation. (a) Ground truth of Liver segmentation by radiologist. (b) Overlap liver segmentation proposed system (Seg.) and ground truth (GT). (c) Yellow line is a ground truth of the lesion drawn by expert and blue line is the segmented lesion by the proposed system . (d) Set matching indicated are the True Negative (TN), False Positive (FP), False Negative (FN), and True Positive (TP) areas.

As shown in Figure 2.11, we define (X) as a set of all pixels in the image. The ground truth $T \in X$ as the set of pixels that were labelled as liver by the radiologist. Similarly, we defined $S \in X$ as the set of pixels that were labelled as liver by the proposed system.

A true positive set is defined as $TP = T \cap S$, the set of pixels common to (T) and (S). True negative is define as $TN = \bar{T} \cap \bar{S}$, the set of pixels that were labelled as non-liver in both sets. Similarly, the false positive set is $FP = \bar{T} \cap S$ and the false negative set is $FN = T \cap \bar{S}$.

Moreover, this section investigates the different evaluation measurement methods that are used to evaluate the system performance. To measure the accuracy of segmentation, let (S) is a segmentation result from a segmentation ground truth (T) that is segmented manually by an expert radiologist. where each of (S) and (T) is a binary mask image. There are many methods to assess the segmentation accuracy which reflects the accuracy of under-segmentation ($|S \cap T| \leq |T|$) or over-segmentation ($|S \cap T| \leq |S|$) or both. These measures are presented as follows:

- Volumetric Overlap Error (*VOE%*):

  The lower absolute values on the (VOE) indicate better segmentation results. The $(VOE = 0\%)$ for perfect segmentation result $(S = T)$. On the other hand, the $(VOE = 100\%)$ for no intersection between segmentation and ground truth $(S \cap T = \phi)$. The (VOE) between two sets of segmentations (S) and (T) is given as a percentage and calculated as Equation 2.2:

$$VOE(S,T) = \left(1 - \frac{|S \cap T|}{|S \cup T|}\right) \times 100\% \tag{2.2}$$

- Relative Volume Difference (*RVD%*)

  The perfect match obtained when $RDV = 0\%$ and larger than zero otherwise. In case of empty segmentation $(S = \phi)$ the $RDV = -100\%$. However, the perfect segmentation value of 0% can also be obtained when the volume of over-segmentation $(|S \backslash (S \cap T)|)$ is the same value under-segmentation $(|T \backslash (S \cap T)|)$. The RVD Equation 2.3 is shown as follow:

$$RVD(S,T) = \left(\frac{|S| - |T|}{|T|}\right) \times 100\% \tag{2.3}$$

- True Positive Volume Fraction (*TPVF%*)

  This method evaluates only the under-segmentation, where $TPVF = 0\%$ when there is no intersection between (S) and (T) and $TPVF = 100\%$ when is no under-segmentation $T \subseteq S$. *TPVF* equation is presented in Equation 2.4.

$$TPVF(S,T) = \left(\frac{|S \cap T|}{|T|}\right) \times 100\% \tag{2.4}$$

- False Positive Volume Fraction (*FPVF%*)

  This method evaluates only the over-segmentation. The FPVF denotes the amount of the pixels from the CT image (I) falsely identified in segmentation (S). However, neither $TPVF$ nor $FPVF$ can evaluate segmentation accuracy alone, both metrics $TPVF$ and $FPVF$ shall be taken always into consideration. The *FPVF* is defined in Equation 2.5.

$$FPVF(S,T) = \left(\frac{|S - (S \cap T)|}{|S|}\right) \times 100\% \tag{2.5}$$

- Jaccard Similarity Metric (*JSM%*)

  The perfect segmentation (100%) obtained when the segmented area is equal the ground truth $(S = T)$. In case of no intersect between segmentation and ground truth $S \cap T = \phi$ the JSM value equal zero. The JSM equation is shown in Equation 2.6.

$$JSM(S,T) = \left(\frac{|S \cap T|}{|S \cup T|}\right) \times 100\% \tag{2.6}$$

- Dice Similarity Coefficient (*DSC%*)

  The DSC is equal to 100% if the segmentation is accurate without false positive $(S = R)$, and DSC is equal zero if the intersect between segmentation and ground truth is null $S \cap T = \phi$. The DSC is calculated via the following Equation 2.7

$$DSC(S,T) = \left(\frac{2 \times |S \cap T|}{|S| + |T|}\right) \times 100\% \tag{2.7}$$

2.4.2.2 Classification/Characterisation Evaluation

In order to evaluate the performance and validate the proposed framework for liver lesion classification/characterisation, a number of different metrics are adopted during the evaluation and benchmarking. Each metric focuses on a different performance aspect. Moreover, we will adopt the standard metrics that are commonly used in medical system evaluation, to facilitate comparison with the current relevant work. The most used metrics of validation are Sensitivity, Specificity, Positive predictive value, Negative predictive value and accuracy (Šimundić, 2008)

Confusion Matrix

Several metrics such as sensitivity, specificity and accuracy that used to measure the system performance can be calculated from the confusion matrix. In this section, the confusion matrix will be discussed first.

Confusion matrix consists of two columns and two rows that represent the number of false positive (FP), false negative (FN), true positive (TP) and true negative (TN), as illustrated in Figure 2.12.

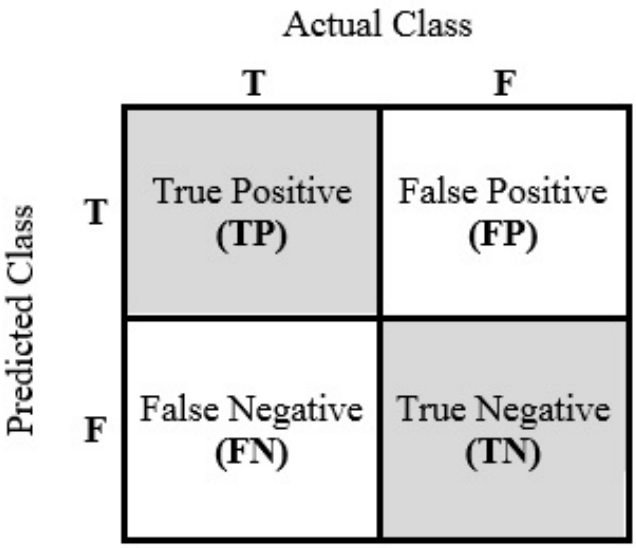


Figure 2.12: Confusion matrix.

Sensitivity (SN)

Sensitivity is calculated as the number of true positive prediction over the number of actual positive cases, as defined in Equation 2.8. However, the sensitivity deals only with the positive cases.

$$Sensitivity(SN) = \frac{TP}{TP + FN} \tag{2.8}$$

Specificity (SP)

Specificity is computed as the number of true negative prediction over a number actual negative cases, as depicted in Equation 2.9. However, the specificity concerns only on the negative cases.

$$Specificity(SP) = \frac{TN}{TN + FP} \tag{2.9}$$

Accuracy (Acc)

Accuracy is the most common metric method which used to evaluate the system performance. The high accuracy means the better system performance. The advantage of

this metric lies in its simplicity and defined by the Equation 2.10.

$$Accuracy(Acc) = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.10}$$

Positive Predictive Value (PPV)

PPV is the proportion of positive predict results that are actually true. It effectively estimates an overall true positive probability for the system and computed by the Equation 2.11.

$$PositivePredictiveValue(PPV) = \frac{TP}{TP + FP} \tag{2.11}$$

Negative Predictive Value (NPV)

NPV is defined as the proportion of the cases with a negative results who are correctly predicted. The NPV is calculated by Equation 2.12.

$$NegativePredictiveValue(NPV) = \frac{TN}{TN + FN} \tag{2.12}$$

Receiver Operating characteristic (ROC) curve

ROC curve is a graphical plot that visually describe the performance of the system. The curve is constructed by plotting the rate of false positive against the true positive rate (Flach et al., 2003; Fawcett, 2006). The *x* axis represents the false positive rate (FPR), and can be computed as $1 - \text{specificity}$. It shows the percentage of positive samples that are incorrectly classified. The *y* axis comprises the true positive rate (TRP), also known as sensitivity. It represents the percentage of positive samples that are correctly identified.

2.5 Conclusion

This chapter started with an overview of the liver anatomy with presenting three different modalities that used in the liver diagnosis and a brief discussion about liver diseases types. Based on that, the hepatic imaging in oncologic patients is undertaken to detect and characterise liver tumours and differentiate between tumour types, especially either benign or malignant. Moreover, to monitoring the lesion development and response to the treatment. As presented previously each imaging modality has advantages and limitations compared to each other. However, selecting the most appropriate imaging technique to detecting and characterising liver lesions rely on the kind of information that required to the clinical decision. The US usually is used in initial examination for the patient who has no past medical history or allergy to contrast agent. Although the ultrasound has several advantages, such as being safe, inexpensive and easily available. However, The US is not an adequate examination for FNH

diagnosis or patient has HCC with cirrhosis. CT scan is widely used for their high sensitivity and specificity to detecting cancers and ability to imaging the entire body. Moreover, CT is faster, more comfort, cheaper than MRI. Unlike CT, which use radiation, MRI offers an attractive alternative to ionising radiation based CT with highly detailed images of soft tissues. However, MRI is expensive, less available, noisy and need longer acquisition time compared to CT.

In overall, The appropriate imaging modalities should be selected on a case-by-case basis to assessed the lesion, depending on several factors such as lesion type, availability of imaging facilities, imaging suitability, and cost issues.

Chapter 3

Literature Review

This chapter presents a review of key related literature work. Various computer-aided diagnosis (CAD) approaches using CT images have been proposed to study liver lesions. These approaches can be divided into two main categories: (1) Liver lesion classification, (2) Liver lesion characterisation. Following this order, the literature will be categorically presented. Hence, Section 3.1 and Section 3.2 reviews related work in image pre-processing and image segmentation techniques respectively. Section 3.3 reviews related work on lesion classification, while lesion characterisation work is presented in Section 3.4. Section 3.5 links the thesis with the literature review to contribute this research field. Finally, the chapter is concluded in Section 3.6.

3.1 Image Pre-processing

CT is one of the important imaging modality that helps to locate the pathological variation in the human body such as cancer. CT provides high contrast images of body tissues or organs with good spatial resolution, and the information provided by CT varies from other imaging examinations such as US and MRI. The flexibility of CT not only provides diagnostic information but is also useful in guiding many clinical procedures like surgeries, interventions and radiation therapies (Kaur and Juneja, 2018a). The quality of CT image plays a vital role in the excellence of medical investigations that can easily be affected by the presence of artifacts during acquisition procedure. The artifacts present in image effects both the diagnostic and automated computerised analysis tasks, such as image segmentation and lesion detection. Thus, a clean CT image is essential to extract the precise medical diagnosis information. However, CT images usually contain noise that can affect the medical diagnosis tasks, mainly in the low contrast images. Hence, a medical image contaminated with noise may not be able to provide correct diagnosis results which can be harmful to the health of patients.

In this thesis, the CT images acquired from different patients subjected to variable radiation doses with a random amount of noise have been considered, which were acquired by different CT scanner system that supports different iterative reconstruction

(IR) algorithms, such as SAFIRE and ADMIRE IR algorithm. However, CT images are usually degraded with the Gaussian or Poisson noise which has a negative impact on the quality of the CT images (Gravel et al., 2004; Sanches et al., 2008). Therefore, de-noising techniques are of great interest in CT imaging, and it is an important preprocessing step to reduce anatomical variability.

3.1.1 Spatial and temporal filter

Spatial filtering involves eliminating noise from the CT image by convolving it with a smoothing function in the spatial domain. However, this technique blurs the sharp edges of the image by reducing its variance, and the amount of blurring depends on the convolution function used for filtering which is same as reducing large frequency values from the image. In CT imaging, convolution filtering is susceptible as data acquired in the frequency domain is multiplied by a smoothing function which lessens large spatial frequency values. In this smoothing, there is a trade-off between the reduction in noise and spatial resolution of the image. The choice of the temporal filter must be made by considering suitable relation to the sampling interval for avoiding the false artifacts. A temporal filter with a limited frequency response reduces the signal at the boundaries of the image, whereas adequate frequency response produces additional noise in the image through aliasing.

The frequency domain filters and local smoothing method reduces noise, but they do not preserve details, texture and fine structure of the image and removes them as they mimic functional aspects of noise. Initially, the effort to reduce the noise based on linear filtering was made by (Lee, 1980). (Tsagaan et al., 2001) and (Liu et al., 2016) utilised 3D Gaussian smoothing filter for removing artifacts in CT image. However, the Gaussian filter have the disadvantage of blurring the edges while reducing the noise. In the same manner, (Lin et al., 2006) and (Campadelli et al., 2009b) used median filter for de-noising CT images which preserve the edges while removing noise.

3.1.2 Anisotropic diffusion filter

(Perona and Malik, 1990) proposed an edge detection and multi-scale smoothing scheme named anisotropic diffusion filter (ADF). ADF addresses the limitation of spatial filtering and considerably improves the image quality by maintaining edges of the object as it effectively removes noise in identical regions and ensures edge sharpening. This filter describes the problem in the form of a heat equation which relies on second-order partial differential equation (PDE) in an anisotropic medium, and it can attain good trade-off among noise removal and edge preservation. Here, image smoothening is framed as a diffusion process that can be stopped or repressed at edges by choosing the strengths of the local gradient in various orientations.

(Linguraru et al., 2011) and (Glisson et al., 2011) employed ADF filter for noise

reduction in CT images, which usually removes small image details by transforming image statistics because of its edge enhancement effect and generates staircase effect in the image. (Zhang et al., 2009a) used ADF filter on brain CT images and concluded that it gives better de-noising results as compared to other techniques with a high computation time.

3.1.3 Total variation de-noising

(Rudin et al., 1992) proposed another de-noising technique based on Rudin, Osher, and Fatemi (ROF) model, known as a total variation (TV) which is capable of smoothing homogeneous areas of the image. It is iterative constrained based optimisation algorithm that performs de-noising by minimising the TV norm with the constraints considered from the image noise statistics. The TV de-noising method uses regularisation expression given by Equation 3.1:

$$\min_{t \in Y} \frac{\|h-k\|^2}{2\gamma} + M\left(k\right) \tag{3.1}$$

Where h represents the noisy image and k represents the desired image of size $n \times n$, Y is the Euclidean space $O^{n \times n}$, and $\gamma > 0$ is Lagrange multiplies and, Euclidean norm is represented by $\|.\|$. The function $M\left(k\right)$ is discrete total variation of k, as depicted in Equation 3.2.

$$M\left(k\right) = \sum_{j \leq n} \left| \left(\nabla h\right)_{i,j} \right| \tag{3.2}$$

However, many improvements have been proposed in the total variation algorithm, but still, this method suffers from staircase effects and eliminates important details from the image. This technique works on the principle that images with extreme and probably false detail possess high total variation, which means it will have a high absolute gradient. Then dropping the total variation of the image removes the undesired details by keeping the essential details such as edges or texture of the image. The advantage of using this technique over other de-noising techniques is that it is effective in preserving the edges and performs de-noising even on images with low signal to noise ratio.

3.2 Image Segmentation

The aim of image segmentation is to divide an image into a set of areas that may correspond in medical imaging with the type of tissue, structure or function (Despotović et al., 2015). The segmentation process is useful to classify the pixels to different

anatomical areas, such as tissues, bones, muscles and blood vessels. in addition, it is utilised to define the pathological areas in the organs (Memon et al., 2006). The segmentation techniques are divided into three groups (manual, semi-automated and fully-automated) based on the degree of user interaction. Liver, lesions and vessels segmentation from CT is a prerequisite for treatment planning and CAD systems. Analysis of vasculature structure from CT images has received interest due to possible contributions towards a variety of medical applications. Accurate annotation of the blood vasculature structure can result in improved diagnosis and more accurate liver surgical and resection planning (Kirbas and Quek, 2003)

The segmentation techniques are applied to different types of imaging modalities. In the literature, these algorithms are commonly used in CT image. The segmentation algorithms can be categorised according to several criteria, such as user interaction, imaging modality, algorithm properties, ect. Regarding the method properties, the algorithms can be classified in grey level-based and contour-based techniques. The next subsections will present in detail the various common segmentation techniques.

3.2.1 Grey Level-based

Grey level is the most intuitive and most obvious image features (Chen et al., 2013). The threshold-based, clustering-based, and region growing approaches belong to this group of techniques (Soler et al., 2001a; Ruskó et al., 2009a; Chen et al., 2009). The gray level-based methods are easy to apply and have a lower computational cost (Ning et al., 2010). These methods have some drawbacks such as less robust to noise and gradient changes; some of these algorithms are semi-automatic and need user's interaction (Zheng et al., 2018).

3.2.1.1 Region Growing

The region growing (RG) method is considered as one of the simplest techniques for image segmentation. The RG technique is partitioning of an image into similar regions of connected pixels by incrementally adding new points to the region. This approach starts with a single or a group pixels that are called initial seed points. The seed point can be selected manually by the user or automatically by a computer algorithm. Starting from the seed point and based on some pre-defined criteria, such as intensity, variance, color or texture etc., they operate by appending new pixels seeded region as long as they meet the pre-defined criteria (Susomboon et al., 2007; Ruskó et al., 2009a; Gloger et al., 2010). However, This method involves two major challenges: selecting the appropriate seed point in the region of interest and defining the criteria such as threshold value for stopping the growth process. The process is iterated on until no more pixels can be appended to the region. The main advantage of RG method is that it is able to divide regions that have similar characteristics and generate a connected

area (Rogowska, 2000).

(Rusko et al., 2007a; Kumar et al., 2013b) proposed a region growing approach for automatic liver segmentation where the estimated liver intensity range based on histogram analysis is used to threshold the algorithm. The multiphase CT image used later by (Ruskó et al., 2009b) to improve the segmentation accuracy. An automatic liver lesion segmentation approach based on region growing and watershed algorithm proposed by (Anter et al., 2013). (Lu et al., 2014) proposed a region growing algorithm to segment the liver from CT image. The Quasi-Monte Carlo approach is utilised to select the optical seed points. However, the liver contour gotten by this method was not smooth. On the other hand, a region growing approach based on intensity similarity and the spatial proximity between voxels are used in vessels extraction task (Selle et al., 2002a; Shang et al., 2008; Jiang et al., 2013a; QIAN, 2017). (Jiang et al., 2013b) proposed a region growing based on spectrum information for vessels segmentation task.

The fundamental drawback of RG algorithm is the noisy data can be a problem for RG approach as it may result in holes or over segmentation (Pham et al., 2000). In addition, it can be computationally expensive (Adams and Bischof, 1994).

3.2.1.2 Thresholding Based

The segmentation methods which based on threshold are often used to make a rough segmentation to determine the region of interest or seed points as pre-processing. The thresholding approach can be used to create a binary mask by replacing each pixel in an image with a black pixel if the image intensity value is less than the specified threshold value, or a white pixel if the image intensity is greater than the threshold. Thresholding methods use histogram properties to classify pixels, thus these algorithms are optimised for a determined type of images with similar histograms properties (Soler et al., 2001a). The threshold can be chosen manually by the user or automatically by a computer algorithm. However, the threshold selection is the major challenging in this approach. (Soler et al., 2001a) proposed the first threshold method for liver lesion segmentation. (Seo and Park, 2005) proposed adaptive multi-modal threshold to find the range of gray-level values of the liver structure, then an optimal threshold for lesion segmentation was proposed by (Park et al., 2005). (Ciecholewski and Ogiela, 2007) proposed the use of histogram equalisation for simplifying the choice of a threshold value of liver lesions. (Rusko et al., 2007b; Huang et al., 2011; Kumar et al., 2013b) have also proposed liver lesion segmentation approaches where the initial lesion boundaries are segmented using histogram analysis and thresholding followed by post-processing steps for a refined segmentation. The advantage of this approach is a simple and easy implementation and low computational cost.

3.2.1.3 Clustering Based

The main idea of the clustering based method is that in n-dimensional feature space, the distance between samples is shorter if they belong to the same class and the similarity of samples from same class is higher. The clustering approaches separate the pixels of an image into a specific number of clusters based on gray-level similarity. These algorithms are iterative, starting with some initial clusters and continuing to modify the clusters until they change. Fuzzy c-means incorporates the fuzzy concept to k-means, in which each data point can belong to more than one cluster. The data points have sets of weights that indicate the degree of their belonging to the clusters. FCM is a popular unsupervised technique in the field of image processing, especially for liver lesion segmentation (Hong et al., 2001). Firstly, a set of tissue classes of the liver and lesion should be predetermined. Then, a membership value is assigned to each voxel related to the tissue classes, based on the features such as intensity, texture, etc. The method performance depends on the initialisation and selecting accurate cluster centres. However, there are two main issues in clustering based approaches: one is how to estimate the similarity of samples, the other is how to determine the threshold of similarity. (Massoptier and Casciaro, 2008) proposed a method using k-means clustering technique to segment the tumor inside the liver. (Häme, 2008) proposed fuzzy c-means clustering with spatial smoothing followed by deformable models for a refined segmentation. (Gunasundari and Ananthi, 2012; Kumar et al., 2013a) proposed the use FCM for liver lesion segmentation. (Moghbel et al., 2016) propose an automatic liver lesion segmentation based on a hybrid approach integrating cuckoo optimisation and fuzzy c-means method with random walkers algorithm. (Foruzan and Chen, 2016) proposed a combination of the SVM, watershed and scattered data approximation approaches to initially segment a lesion. The segmentation results are refined using sigmoid edge model. The other method combined FCM with graph cut approach (Wu et al., 2017) to improve the lesion segmentation results. The lesion was extracted using confidence connected region growing method to reduce computational cost. Then, initial foreground/background regions were labeled automatically, and a kernelised FCM with spatial information was incorporated in graph cuts approach to increase segmentation accuracy. (Ahmadi et al., 2016) proposed liver vessels extraction approach based on FCM in conjunction with genetic algorithms. The advantages of clustering based algorithms are: they are fully automatic and they can handle multiple tasks of segmentation. In addition, they have a relatively low computational cost (Das and Sabut, 2016).

3.2.2 Contour-based

Contour-based approaches generally achieve good results in segmentation process but a more complex interaction, initialisation, and/or training process is required in or-

der to obtain the desired results. However, these characteristics can be inappropriate in a clinical environment (López-Mir et al., 2015). Probabilistic atlases, level-sets, deformable models and statistical shape models are approaches used in segmentation which are based on contour properties (Heimann et al., 2007a, 2009; Casciaro et al., 2012a; Wang et al., 2013a; Yang et al., 2014).

3.2.2.1 Probabilistic Atlases

Probabilistic atlases (PA) utilise both the prior shape and the spatial location information to achieve a refined segmentation. Moreover, relations between adjacent structures can be effectively formulated using multi-class atlases. The PA is constructed by a manual segmentation from a large number of anatomical images. These images are registered into a standard space via affine transformations. The procedure for producing PA starts with averaging the images and corresponding segmentations where the PA is employed into a Bayesian frame. Afterwards, the probability of each pixel/voxel to belong to the specific organs is computed. lastly, a simple thresholding or an iterative conditional mode approach is used to extract the wanted part based on the posterior probability (Park et al., 2003a; Shimizu, 2006).

(Zhou et al., 2006) proposed a probabilistic model consisting of location and density probabilities. In this approach, spatial relations between liver, bone structure and diaphragm are used for defining a standard anatomical structure surrounding the liver with this anatomical structure being utilised as a reference with all cases being deformed to it. Density probabilities of liver intensities are assumed as Gaussian and calculated utilising the regions under PA. After that, both probabilities are used for segmenting the liver. (Slagmolen et al., 2007) suggested a PA and corresponding intensity atlas using affine registration based on maximisation of mutual information (MI) followed by a non-rigid registration using B-splines via MI and surface distance between the reference segmentation and the floating image segmentation as a similarity metric. However, as in this PA is not directly build on the image to be segmented, an additional step is required to register both intensity atlas and PA onto the image of interest and then final segmentation is obtained by thresholding the PA. (Li et al., 2010) proposed a probabilistic liver atlas combined with a rib cage atlas. In this method, problems of mapping derived from the variability in the liver shape are avoided by achieving a more accurate mapping of probabilistic atlas onto the input CT volume. The main drawback of this AP approach is that the PA requires a lot of training data to be collected and manually segmented (Park et al., 2003b). In addition, it has high computational costs (Ji et al., 2013).

3.2.2.2 Level Sets

Level set method (LSM) relying on boundary tracking by dynamic variations to extract the object (Osher and Fedkiw, 2006). LSM segments the object by asking the user to

draw a rough contour outside/inside the target, and then the contour will shrink/extend. When the contour corresponds to the target boundary, the shrinking/growing procedure will be stopped. The direction of the shrinking/extending and finding the end point of the procedure is controlled by the speed function (Leventon et al., 2000). This approach requires training (Wang et al., 2013b) or an initial iteration to form the initial curve (Yang et al., 2014) or drawing the initial boundary (Peng et al., 2014). However, some efforts have focused on automating the initial boundaries but the computational cost has increased significantly (Casciaro et al., 2012b) or decreased the algorithm's accuracy (Ciecholewski, 2014). LSM has gained attention for liver, lesion and vessels segmentation with a number of proposed methods. (Dawant et al., 2007) segmented the liver based on 2D level sets with a dynamic speed function. A 3D level set method based on a medium level of intervention for the liver segmentation in CT slices suggested by (Wimmer et al., 2007). (Platero et al., 2011) proposed a variation of level set where shape priors are incorporated into edge-based and region-based models. (Yang et al., 2014) proposed a semi-automatic approach to segment the liver from CT image. The liver is initially extracted by a level set algorithm with multiple seed points selected by the user, and followed by a threshold-based method to refine the initial segmentation. The automatic liver segmentation method utilising level set proposed by (Wang et al., 2016). The initial liver segmentation is obtained using a probabilistic atlas method with a maximum a posteriori classification. The segmentation is then refined using a shape and intensity prior based level set approach. On the opposite, a level set approach had been used for lesion segmentation. (Smeets et al., 2008) proposed a level set approach based on statistical pixel classification with supervised learning for the lesion segmentation. A multi-resolution 3D level set method coupled with adaptive curvature approach for the classification of the pixels into lesion and background proposed by (Jimenez-Carretero et al., 2011). Furthermore, the level set is amongst the most popular vessel segmentation approaches (Toledo et al., 2000). (Selle et al., 2002b) proposed pixel clustering approach for liver vessel segmentation and then refined by level sets approach. (Jin et al., 2013; Hibet-Allah et al., 2016) proposed Hessian-based multi-scale filtering and a level set approach for vascular trees extraction. (Lu et al., 2017) proposed a liver vessel segmentation based on the variational level set approach, which uses non-local robust statistics to suppress the influence of noise in the images. (Selvalakshmi et al., 2017) proposed a new fuzzy Bernstein polynomial level set algorithm for segmentation of liver lesions and hepatic vein where the initial level set is evolved directly from the initial liver segmentation by spatial fuzzy clustering approach. However, The main drawbacks of the level set approaches are the time-consuming, computationally expensive and the space complexity.

3.2.2.3 Deformable Models and Statistical Shape Models

Deformable models are curves or surfaces defined within an image domain that can move under the influence of internal forces, which are defined within the curve or surface itself, and external forces, which are computed from the image data. The internal forces are designed to keep the model smooth during deformation. The external forces are defined to move the model toward an object boundary or other desired features within an image. On the other side, a statistical shape model (SSM) is built by given a set of examples of a shape. SSMs utilise prior information by providing global shape constraints based on a training set. The main drawback is that SSM requires a large training dataset. Furthermore, it often lacks the flexibility to adapt accurately to a structure with high variations in shape. Thus, many of the proposed SSMs are followed by a free-form deformation step utilising deformable models (Heimann et al., 2006; Kainmüller et al., 2007; Zhang et al., 2010a; Wang et al., 2015a).

(Saddi et al., 2007) suggested an automatic liver SSM with deformable models based on non-rigid template matching for a refined segmentation. A hybrid method utilising evolutionary algorithm,SSM and a deformable mesh proposed by (Heimann et al., 2007b). Initialisation of SSMs is done by a multi-resolution algorithm using an evolutionary approach. A deformable mesh is fitted to extract a liver based on internal forces and external forces. The internal forces describe the deviation of the mesh from the underlying SSM, while the external forces model the fit to the image data. (Kainmüller et al., 2007) proposed an automatic liver segmentation by matching the SSM to the CT data and enhance the initial results by a deformable mesh. The SSM consists of around 7000 landmarks and is built by 112 liver shapes. (Seghers et al., 2007a) proposed to employ local information for solving the flexibility problem of SSM. Unlike the global shape models, in this implementation, a statistical model for each edge of the mesh is utilised for capturing mean and covariance of the edge vector. (Erdt and Kirschner, 2010) proposed an SSM approach combining learned local constraints with constraints directly obtained from current curvature for coping with the drawback of SSM in regions with high curvature. (Li et al., 2015) proposed an automatic liver segmentation framework based on shape constraints and deformable graph cut in CT Images. Firstly, the CT image is smoothed using curvature anisotropic diffusion filtering to construct a statistical shape model based on the principal component analysis. Secondly, the mean shape model is moved using thresholding and Euclidean distance transformation to obtain a coarse position in a CT image, and then the initial mesh is locally and iteratively deformed to the coarse boundary, which is constrained to stay close to a subspace of shapes describing the anatomical variability. Finally, the deformable graph cut was proposed to refine the liver surface detection. (Saito et al., 2017) proposed an automatic liver segmentation method using SSM where SSM-guided expectation-maximisation (EM) algorithm without using spatial standardisation for better handling of pathological livers. The segmentation is then

refined using a graph cut based approach.

3.2.2.4 Graph Cut

Graph Cut (GC) is an alternative to boundary based segmentation methods proposed by (Boykov et al., 2001). The main idea of graph cuts is to represent the image to an undirected weighted graph. Each node represents each pixel of the image. Each edge connected a pair of adjacent pixels. The weight of edge indicates the similarly of gray level, colour or texture between each pair. The GC segments the image into background and the object by finding the minimum cost function between all possible cuts of the graph. The best cut is to make the similarity in a sub-graph maximum and the similarity between sub-graphs minimum. However, the user is required to provide the seeds representing the background and the object to be segmented. (Beichel12 et al., 2007) suggested graph cut approach for initial liver segmentation. Afterwards, requiring the user to fix arbitrary segmentation errors. (Massoptier and Casciaro, 2007) proposed a GC approach for liver segmentation initialised by adaptive thresholding. (Stawiaski et al., 2008) proposed an interactive liver lesion segmentation using graph cut algorithm for an initial segmentation and watersheds method for a refine segmentation. (Shimizu et al., 2010) proposed a combination of statistical atlas-based algorithm and graph cuts method. Prior shape is estimated by using the PA proposed in (Park et al., 2003a) and then an implicit SSM is fitted as proposed in (Leventon et al., 2000). (Pamulapati et al., 2011) introduced a 4D graph-based to segment liver vessels and lesions based on multiphase CT image to model the differential enhancement of the liver structures and Hessian-based shape likelihoods to avoid the common pitfalls of graph cuts with under segmentation and intensity heterogeneity. (Chen et al., 2012) proposed a strategic combination of the active appearance model (AAM), live wire (LW), and graph cut for abdominal organ segmentation. The AAM combined with LW was used for initial object recognition and delineation. The shape constraint generated from the initial recognised object was integrated into the graph cuts cost computation to improve segmentation accuracy. (Wu et al., 2016) proposed an automatic liver segmentation approach using a histogram-based adaptive thresholding approach and morphological operations. Then, the initial segmentation was enhanced by using supervoxel-based graph cuts algorithm. A graph cuts and border marching based method for liver segmentation proposed by (Liao et al., 2017). The pixel-wise and patch-wise features is employed to enhance the original data and highlight the liver region. However, a major drawback of this technique is their reliance on initialisation as well as the need for a dense neighbourhood system in order to avoid geometric artefacts and jagged boundaries (Daněk et al., 2012).Furthermore, it has a higher computational complexity.

In summary: image segmentation methods, including liver, lesions and vessels segmentation, can be grouped according to several criteria, such as user interaction,

algorithm properties, etc. The literature review divided these methods into grey level-based and contour-based approaches. Contour-based approaches generally achieve better results. However, a hard user interaction, initialisation, and/or training process is required, in order to obtain the desired results. Additionally, the computational cost increases considerably in these methods. These reasons can be inappropriate in a clinical environment, making their application limited (Campadelli et al., 2009a; López-Mir et al., 2013). On the other hand, grey level-based methods can obtain good results with a reasonable computational cost and the low training required. Thus, they could be optimal for clinical environments (Ruskó et al., 2009a; Gloger et al., 2010). However, pre- and/or post-processing steps are required in order to increase the segmentation accuracy by reducing the image noise and adjacent organs connections (Gloger et al., 2010).

3.3 Liver Lesion Classification

There are several systems that were proposed by researchers to classify liver tumor according to the tissue types (Normal, Benign, and Malignant). Computer aided liver lesion classification can be divided into three types based on liver segmentation and lesion detection. First, manual liver segmentation and lesion detection where drawn by experienced radiologist. Second, semi-automated liver segmentation and lesion detection. Finally, fully-automated liver segmentation and lesion detection.

Manual liver segmentation and lesion detection; these techniques completely depend on expert radiologist to mark lesions, while focusing on extracting the appropriate features to feed classifiers to identify lesion types (Duda et al., 2006; Huang et al., 2006; Dramiński et al., 2007; Ganeshan et al., 2009; Wang et al., 2009; Ye et al., 2009; Duda et al., 2013; Yang et al., 2013a; Rao et al., 2014; Doron et al., 2014a). The advantage of this approach is avoiding the error rate in segmentation and lesion detection. However, some of the lesions are invisible and unreachable by the human perception (Mir et al., 1995). Furthermore, segmentation accuracy depends only on the skills and experience of the user. In addition, it is a very time consuming and a tedious process.

Semi-automated classification systems; this type of systems needs the radiologist to intervene in segmentation and detection of lesions (Kretowski, 2002; Stoitsis et al., 2006; Mougiakakou et al., 2007a), where it could be done in two possible ways: (1) The radiologist outlines the region of interest with the mouse clicks then the automatic segmentation algorithm is applied. (2) The automatic segmentation is followed by manual checking and editing of the segment boundaries. However, the accuracy of this method depends on the experience of radiologist to enter the parameter for a correct segmentation.

Fully automated systems (Mala and Sadasivam, 2010; Chi et al., 2013b; Kumar et al., 2013a); the main advantage of this method is that no user interaction is needed,

which saves the time and can provide more useful information and helps the radiologist in diagnosis of diseases. However, low contrast between organs and the high shape variability of the liver makes automatic segmentation a hard task.

Recent research efforts have been contributing considerably to enhance the performance of lesion classification (Kumar et al., 2012). Generally, the classification system is performed in three systematic sequential stages, which are (1) Liver lesion extraction to define the region of interest (ROI), (2) Features extraction to compute low-level features from segmented ROI (3) Classification process based on extracted features.

The majority of liver lesion classification methods, utilises different low-level features calculated directly from single ROI (segmented lesion) to design robust CAD systems. Feature extraction refers to identifying a set of distinguishing and sufficient features from medical image such as CT scan image for lesion diagnosis. The extracted features capture key informations of the image such as intensity, texture, and shape properties.

Generally, all proposed CAD systems shared in two main stages for lesion classification, which are feature extraction and classification (Kang et al., 2014b; Roy et al., 2014b). Features extraction is an important phase in the CAD system to describe the lesion with its essential characteristic (Rockey et al., 2009b; Fergusson, 2012b). Basically, there is a large diverse set of features to be used. Those come under three categories; Bag-of-Visual features, texture feature and combined (texture,intensity and shape) features, as fully illustrated in Figure 3.1.

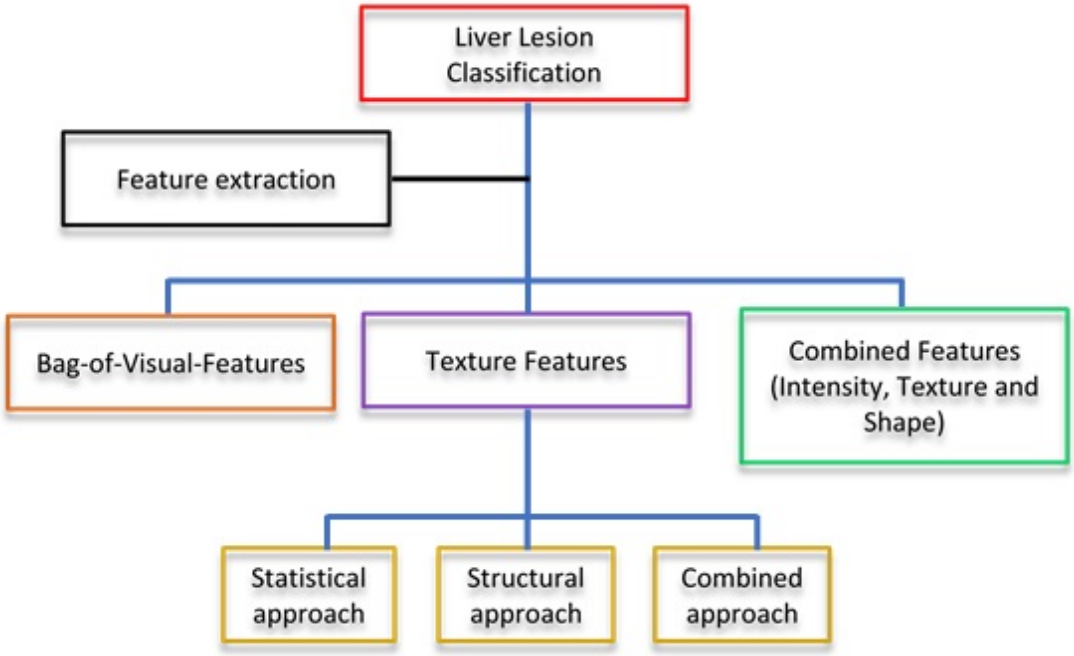


Figure 3.1: Literature on liver lesion classification, categorised according to work on low-level features type.

3.3.1 Bag-of-Visual Features

This section presents the most relevant key work regarding lesion classification based on ROI selection, according to the categorisation shown in Figure 3.1. Hence, Bag-of-Visual-Features will be introduced in this section.

The Bag-of-Visual-Words (BoVW) can be used for image classification, through treating image features as words. The ROI is abstracted by several local patches to generate a numerical vector called feature descriptors (Winn et al., 2005). The feature descriptors-represented patches are converted to codebook (Van Gemert et al., 2010) which is considered as a representative of several similar patches in the feature space.

Image patch representation and bag-of-visual-words (BoVW) were used in lesion classification by analysing multiple region of interest, (Safdari et al., 2013) proposed classification and lesion detection system, where a visual word histogram was used to build a dictionary through using local descriptors and representing a region in the image. The accuracy result reported is more than 95% on dataset size of 73 CT images. (Wang et al., 2015b) proposed CAD system to classify normal livers and livers with lesions. The dataset comprised 151 cases (76 normal and 75 abnormal). The patches are extracted densely from a given ROIs with 400 to 500 patches. The histogram of oriented gradients (HOG) and intensity are extracted as the features of patches inside ROI. The coding dictionary is created from feature clustering in the training set. Each patch feature is coded with sparse constraint to generate a coding scheme from both training and testing images. Bag of visual features (BOF) is used to represent the ROI. The SVM is adopted for classification with accuracy of 96.15%.

Regarding to recent developed BoVW, (Diamant et al., 2016) proposed single dictionary BoVW for the automatic lesion classification. Two datasets were used to evaluate the system performance with total number of 194 CT images. The visual word histograms are generated based on creating two separate dictionaries from two ROIs for interior and lesion margin region where all the patches inside ROIs are clustered by using k-means algorithm. The two histograms from interior and margin lesion are concatenated to build a new feature vector that represents the given lesion. The proposed method shows result with 93% of accuracy by using nonlinear SVM with histogram intersection kernel. The best accuracy was obtained using the combination of two datasets parameter: (1) patch size 7x7 pixels, dictionary size of 160 words and word size of 14 for the first dataset and (2) patch size 9x9 pixels, dictionary size of 200 words and word size of 12 for the second dataset.

An approach similar to the preceding ones was used by (Diamant et al., 2015). In their work, the analysis of BoVW was performed on single (portal phase) and multi-phase CT scans. The given ROI is divided into uniform patches (patch size between 5 and 13 pixels), from which visual words are computed. The histogram of visual word is utilised to generate feature vectors for each image. A mutual information (MI)

criterion was used to improve BoVW model by selecting the most relevant words from a generated dictionary. For classification task, the SVM with histogram intersection kernel was used. The experiments on 85 CT images, used optimal parameter of BoVW-MI (patch size of 11, visual of word size of 10 principal component analysis (PCA) coefficient and dictionary size 260 words), gave the weighted average classification sensitivity and specificity of about 82.4% and 92.7% for multi-phase CT scan and 70.6% and 86.9% based on single phase. This work was later tested by (Diamant et al., 2017) on three different tasks: chest x-ray pathology identification, liver lesion classification with dataset size 118 portal phase CT images and lesion classification in breast mammograms, by considering different parameter of BoVW for each task. Table 3.1 represents a different classification system based on bag-of-visual features.

In summary: BoVW and image patches are a crucial approach to any classification system. The BoVW were originally proposed for text document analysis, and it was further adapted for image analysis (Bosch et al., 2007). The existing methods tried to capture all the characteristics of lesion through dividing the ROI into patches (Diamant et al., 2016). However, the widely used BoVW approach with image patches for lesion classification basically (1) uses k-means for coded vector calculation to generate sparse dictionary learning (codebook learning) (Jurie and Triggs, 2005), which approximates any local descriptor using one learned visual word only and leads to large reconstruction error of local descriptors (Wang et al., 2010, 2017); (2) Through literature, there exists a limited work based on extracted patches that the accuracy of the existing methods mainly depends on the number and size of the patches in addition to the dictionary size , which may work only with specific conditions such as specific dataset and specific machine settings. As a consequence, the performance is varied significantly under ROI selection (Singh et al., 2014) and different acquisition conditions (Bharti et al., 2017) such as different operators and settings. In proposed work, the multiple ROIs (internal, border, and surrounding area) fused with the difference-of-features between the internal lesion and surrounding area employed as a new feature vector to well-represent the lesion characteristic, and the various behaviours between benign and malignant.

Author	Year	Dataset	Patch size	Number of dictionary word	Features	Accuracy
(Safdari et al., 2013)	2013	Cyst (25); Metastases (24); Haemangiomas (24)	9x9	250	Visual word histogram	95.89%
(Wang et al., 2015b)	2015	Normal (76); Abnormal (75)	16x16; 24x24; 32x32	500	HOG; Intensity	96.15%
(Diamant et al., 2015)	2015	Haemangiomas (27); Focal Nodular Hyperplasia (16); HCC (29); Cholangiocarcinoma (13)	11x11	260	Visual word histogram	Portal phase (Sensitivity 70.6%; Specificity 86.9%); Multi-phase (Sensitivity 82.4%; Specificity 92.7%)
(Diamant et al., 2016)	2016	Cyst (61); Haemangiomas (53); Metastases (80)	7x7	160	Visual word histogram	93%
(Diamant et al., 2017)	2017	Cyst (22); Haemangiomas (32); HCC (29); Metastases (35)	11x11	750	Visual word histogram	sensitivity 83.1%; specificity 93.6%

Table 3.1: Liver lesion classification literature work based on bag-of-visual features.

3.3.2 Texture Features

Texture is a powerful discriminating visual feature which has been widely used in pattern recognition and computer vision for identifying visual patterns with properties of homogeneity that cannot result from the presence of only a single colour or intensity. Most of biomedical images acquired and represented in grayscale are often highly textured, and consequently, examination of biomedical images usually requires interpretation of organ/tissue/lesion appearance, i.e., the local intensity variations, based on different texture properties such as smoothness, coarseness, regularity, and homogeneity. Since texture acquires such distinguished importance, it is becoming one of the most commonly used characteristics in biomedical image classification and retrieval. Hence, texture features gained more attention in liver tumor classification, such as: Co-occurrence Matrices (COM) (Haralick et al., 1973a; Bankman, 2008; Conners and Harlow, 1980); Gradient Matrices (GM) (Lerski et al., 1993); Gray Level Histogram (GLH); Gray-Level Co-Occurrence Matrix (GLCM); Texture Feature Coding Method (TFCM) (Horng et al., 2002); Gray Level Difference Matrices (GLDM) (Weszka et al., 1976); Spatial Gray Level Dependence Matrices (SGLDM); First Order Statistics (FOS); Run Length Matrices (RLM) (Galloway, 1975; Albregtsen et al., 2000); Gray Level Run Length Matrix (GLRLM); Wavelet coefficient statistics; Discrete Wavelet Transform (DWT) (Mallat, 1989); Autocorrelation Coefficients (AC) (Gonzales and Woods, 2007); Laws Texture Energy (LTE) (Laws, 1980); Fractal Model (FM) (Mandelbrot and Pignoni, 1983; Chen et al., 1989, 1998; Li et al., 2009; Sankar and Thomas, 2010).

The intensity and shape features were used mostly for liver segmentation and lesion detection (Seghers et al., 2007b; Mitrea et al., 2009). But the majority of research was developed based on texture features as lesion descriptors (Duda et al., 2004), which can reveal subtle characteristics of the lesion and robustness. Furthermore, the computerised methods for texture characterisation being able to overpass the limits of the subjective human eye (Mitrea et al., 2009). Textural features usually play only a secondary role in feature description that deals with coloured images. However, in medicine, textural features gain an extra importance, because (a) gray-level features alone may not have enough discriminatory power and (b) some diseases can affect the organs in such a manner that CT images of those organs show texture changes with less intensity changes. Moreover, texture features have been utilised in several ways: (1) Standalone, (2) Through merging with other features (e.g. intensity and shape). The rest of the literature can be categorised in two groups based on feature extraction, which are texture features, and combination features (intensity, shape, and texture), as illustrated in Figure 3.1.

Furthermore, texture features were used in different CT image phase such as non-contrast CT image, single phase-enhanced CT image after contrast agent such as portal phase, and multiphase CT image (non-enhanced, arterial phase, portal phase, and de-

layed phase). It will be followed by extensive literature review. This section will be further categorised into three main approaches: (1) Statistical approach, (2) Structural approach and (3) Combined Statistical and Structural approach.

3.3.2.1 Statistical approach

Statistical texture analysis approaches (Eltoukhy et al., 2012) are more extensively used for texture feature extraction. They recognise the texture as a quantitative measure of the arrangement of intensities in an area. One of the earliest attempts to utilise texture in classification was proposed in 1995 by (Mir et al., 1995), which works direct on lesion without segmentation. This study explored whether the texture features could be providing information to differentiate between malignant and normal liver tissue that are unreachable to the human perception through find the most appropriate texture features. COM, RLM, GLDM texture features were used in this work (Mir et al., 1995). The dataset consists of three groups of liver CT tissues were classified: normal liver, abnormal tissue (malignant) which divided into malignancy with clearly visible, and malignance was not visible. Entropy, local homogeneity (COM), and gray level distribution (RLM) texture features were found the most useful features to classify invisible malignancy tissue with a confidence level above 99%. However, the result represented using confidence level not accuracy.

While (Ganeshan et al., 2009) proposed a system to determine whether texture analysis of CT images in liver region apparently healthy were changed through the presence of a malignant tumor in patients with colorectal cancer. The dataset consists of three groups of tissues. Firstly, healthy tissue with no malignancy. Secondly, malignancy tissue but no liver involvement. Finally, liver metastases tissue (malignant). The tumor region was manually constructed as region of interest. Statistical parameter of texture (mean gray-level intensity, Entropy, and Uniformity) were extracted from both unfiltered images as well as filtered images (highlighting fine, medium and coarse texture). However, the experiments revealed that texture features derived from unfiltered images for three groups were not significantly different, while statistically different when used filtered images.

Semi-automatic classification system was presented by (Stoitsis et al., 2006). Image pre-processing was used to enhance image quality and define tumor as ROI. The proposed system was able to classify four types of liver tissues: Normal, cyst, hemangioma, and HCC. Five texture features (FOS, COM, GLDM, LTE, and FM) were extracted for each tumor. The most useful features were found using a feature selection, based on Genetic Algorithms. Classification was carried out by Neural Network. The best classification accuracy was equal to 90.63%. Later, (Mougiakakou et al., 2007a) extended this work by considering a different features set (FOS, SGLDM, GLDM, TEM, and FDM).

Another classification system was proposed by (Wang et al., 2009). The system classified three types of hepatic tissue (Normal, HCC, and Hemangioma). The ROIs of tumor were defined by experienced radiologists. For each ROI, four texture features (FOS, SGLDM, GLRLM, and GLDM) were extracted to fed SVM classifier. The classifier was used two strategies to construct multiclass SVMs: one-against-all (Liu and Zheng, 2005), one-against-one (Hsu and Lin, 2002). The performance of classification was assessed through 5-fold cross validation. The best accuracy has been observed by the multiclass SVM using one-against-one method which was 97.78%. However, the proposed system used a fixed size ROI (32x32), while the lesion has variable size larger or smaller than selected ROI.

Yet another fully automated classification system was presented by (Mala and Sadasivam, 2010). The proposed system was able to detect liver and classify two types of tumors (fatty and cirrhosis liver). Wavelet based statistical texture features is extracted and fed three types of Neural Network (Probabilistic Neural Network, Linear Vector Quantisation, and Neural Network and Back Propagation Neural Network). The system evaluation used was 10-fold cross validation. The experiment showed higher accuracy of 96%. However, the proposed system work on the noise-free dataset. Moreover, the accurate of liver segmentation and boundary are very important and critical in classification accuracy.

(Duda et al., 2004) proposed a classification system based on triphase liver CT image. For each phase, texture features (gray-level histogram, COM, RLM, and Laws) were extracted from manually traced ROI. These features were combined in one feature vector to classify liver tissue into normal or one of two main primary tumors (HCC, cholangiocarcinoma). Decision Dipolar Tree classifier was used in classification. The proposed system applied was 10-fold cross validation method and repeated 5 times to evaluate system accuracy. The experiments showed that the higher accuracy gained when used features from triphase (non-enhanced, arterial phase, portal phase) together rather than used features from each phase separately. The accuracy recorded for each phase was 95.5%, 93.9%, and 93.9% respectively, while triphase together was 99.7%. However, another work later by the same team (Duda et al., 2006) confirmed that texture parameters derived from three subsequent acquisition moments (triphase) improves the classification accuracy.

Another approach based on four phases CT image was proposed by (Chi et al., 2013b). The importance of this study is to improve radiologist's accuracy in tumor diagnosis. The proposed system used a hybrid generative-discriminative (Chi et al., 2013a) and nonrigid B-spline registration (Yushkevich et al., 2006) methods to localise tumor on multiphase image. Multi-phase density and co-occurrence matrix features were extracted for each phase and combined in one feature vector. The precision-recall curve and the Bull's Eye Percentage Score (BEP) (Manjunath et al., 2002) were used to evaluate the system performance. The experimental results showed improvement

in BEP score when used multi-phase image to obtain 78%, while the score recorded 63%-65% using a single phase image. However, the proposed system can be used in content-based image retrieval for focal liver lesions by providing visually similar lesions.

Another tumor classification based on multi-phase CT image (pre-contrasted phase, arterial phase, portal phase and delayed phase) was presented by (Ye et al., 2009). The dataset consists from four types of liver tissue (Normal, cyst, haemangioma, and HCC); each lesion was drawn by experienced radiologist. Temporal features (relative signal intensity, intensity change tendency, and signal enhancement ratio), FOS, SGLM features were extracted from each phase to be fed into SVM classifier. K-fold cross validation was used to evaluate system accuracy. The best classification accuracy was achieved when mixed features (Temporal features, FOS, and SGLM) were 95.5%, 97.2%, and 96.4% for normal-abnormal, cyst-other disease, and carcinoma-haemangioma sub problems respectively. However, Boundary of each lesion was drawn by experienced radiologist. Within the identified lesion, a 16x16 pixels square was extracted as ROI for extracting features and classification. Which means the proposed system is not capable of handling small lesions. Furthermore, the temporal features are quite limited as the different features are computed over the mean value of the pixels (heterogeneous lesions might be hard to distinguish in this case).

Another system, developed by (Krishan and Mittal, 2015), focused on enhancing CT images using two different algorithms: contrast limited adaptive histogram equalisation (CLAHE) and constrained variable histogram equalisation (CVHE). Where CVHE and CLAHE were used to enhance CT image and lesion respectively. Twenty texture features based on spatial gray level dependence matrices are extracted from a lesion cut size of 25x20 pixels. The classification (normal and abnormal) was performed with the SVM for 97% accuracy. However, the dataset is limited where it contains only malignant and normal liver. Moreover, the proposed system used a fixed cut size ROI (25x20), while the lesion has variable size larger or smaller than selected ROI. In (Obayya et al., 2016), the adaptive neuro-fuzzy inference system (ANFIS) was used to classify liver lesion as benign or malignant. The liver is extracted automatically by applying threshold and boundary extraction algorithm on CT image. Then FCM approach is used for lesion segmentation. GLCM and discrete wavelet transform (DWT) features are extracted to train ANFIS classifier separately. The proposed system was tested on 100 images evenly split between benign and malignant. The author reported that the DWT is more effective than GLCM with an accuracy of 96% and 90% respectively. However, The DWT feature has has some limitations in capturing relevant information such as a lack of shift invariance, which means that small shifts in the image can cause a significant variations in values of wavelet coefficients at different scales (Kingsbury, 2001). Table 3.2 represents a different classification system based on statistical texture features.

Author	Year	Dataset	Features	Accuracy
(Mir et al., 1995)	1995	Normal (20); Abnormal clear visible malignant (20); Abnormal invisible malignant (20)	COM; RLM; GLDM	invisible malignant 99%
(Stoitsis et al., 2006)	2006	Normal (76); Cyst (19); Hemangioma (28); HCC (24)	FOS; COM; GLDM; LTE; FM;	90.63%
(Duda et al., 2006)	2006	Normal (150); HCC (150); Cholangiocrcinoma (150)	FOS; RLM; COM; LTE	non-enhanced, 95.5%; arterial phase, 93.9%; portal phase, 93.9%
(Mougiakakou et al., 2007a)	2007	Normal (76); Cyst (19); Hemangioma (28); HCC (24)	FOS; COM; GLDM; LTE; FM	84.96%
(Ganeshan et al., 2009)	2009	Healthy (15); Malignant not in liver (9); Metastases (8);	FOS; COM	-
(Ye et al., 2009)	2009	Normal (64); HCC (26); Cyst (14); Haemangioma (27);	COM; FOS	95.5%
(Wang et al., 2009)	2009	Normal (30); HHC (30); Hemangioma (30)	FOS; SGLDM; GLRLM; GLDM	97.78%
(Mala and Sadasivam, 2010)	2010	Fatty liver (100); Cirrhotic liver (100)	BWT and FOS; BWT and COM	96%
(Chi et al., 2013b)	2013	HCC (16); Metastases (10); Hemangioma (16); Cysts (15); Liver abscess (7); FNH (5)	COM; FOS	-
(Krishan and Mittal, 2015)	2015	Normal (20); Malignant (20)	SGLDM	97%
(Obayya et al., 2016)	2016	Malignant (50); Benign (50)	GLCM; DWT	Using GLCM 90%; Using DWT 96%

Table 3.2: Liver lesion classification literature work based on statistical texture features.

3.3.2.2 Structural approach

In structural texture analysis approaches, (Kitasaka et al., 2003) described a texture as the composition of well-defined texture elements such as regularly spaced parallel lines. The properties and placement rules of the texture elements define the image texture.

The lesion classification system was proposed by (Huang et al., 2006). The system has been adapted in order to differentiate between the two sets of liver tumors: HCC (malignant) and hemangiomas (benign). The tumor region was manually selected and extracted as a circular sub-image from CT image. The equalisation autocorrelation coefficients between neighbouring pixels within the image were used as features to identify liver lesion. Support vector machine (SVM) (Vapnik, 2013) was used in classification. The evaluation of proposed system used k-fold cross validation (Duda et al., 2012). Accuracy of classification was equal to 81.7%.

(Ramamoorthy et al., 2015) examined the feature extraction methods for a better classification accuracy through enhancing texture recognition. The rotation-invariant texture features are extracted by two methods: (1) using individual Gabor filter and (2) combining the multi-scale property of Gabor filters and rotation-invariant property of Local Binary Pattern (MGRLBP). The system was tested on dataset of 56 images of four different classes (30 images are split equally between normal and hemangioma, 26 images are divided into fatty liver and cyst). The SVM classifier was used in classification task. The author reported that the proposed MGRLBP performed better than individual feature extraction. However, the size of the dataset used for evaluation is small. Furthermore, the dataset consisted only from normal liver tissues and benign lesions.

3.3.2.3 Combined Statistical and Structural approach

The texture can be represented by structure or statistical approaches. In this section, the combination between two approaches will be presented. Table 3.3 shows a comparison between different classification system based on combined (statistical and structural) texture features.

One of the earliest semi-automatic classification systems was proposed by Kretowski (Kretowski, 2002). The objective of the work is classifying liver tissue to normal or metastases after taken the contrast agent (arterial phase, and portal phase). Four type of texture features (FOS, GM, COM, and RLM) were extracted and fed Dipolar Decision Trees (Bobrowski and Kretowski, 2000) classifier. The arterial phase showed the highest accuracy. The tumor classification system based on three typical acquisition moments (without contrast, arterial, and portal phase of contrast agent) was proposed by (Duda et al., 2013). ROI is drawn on each phase of image at the same size and of the same anatomical position. Several types of texture features (FOS, COM, RLM,

GLDM, GM, TFCM, AC, and LTE) were used to differentiate between four types of liver tissue (Normal, cholangiocarcinoma, cirrhosis, and HCC). To select the most robust features, a simplified Monte Carlo (Dramiński et al., 2007) was applied as a feature selection method. An Adaptive Boosting algorithm (Quinlan, 2014) with a C4.5 tree (Freund and Schapire, 1995) was used in classification. The experiments result shown that a small set of features (consist from 12 features) gained higher accuracy exceeding 90%, while all set of features (183 features) recorded an accuracy rate of 88.94%. However, the features are unstable and dependent on the lesion characteristic on each phase. Where the lesion may appear in phases and disappear in another.

Another work was presented by (Quatrehomme et al., 2012). The proposed system work on four phases liver CT image (non-enhanced, arterial phase, portal phase, and delayed phase) to classify five types of liver tumors; three types are benign (cysts, adenomas, haemangiomas); and two types are malignant (HCC, and metastases). Four different types of texture features (gray level histogram, unser histograms statistics (Unser, 1986), LTE, and Gaussian Markov Random Fields (Cross and Jain, 1983)) were extracted and then labelled by a SVM classifier. A Leave-One-Out cross validation was used to evaluate the system performance. The experiments showed that a significant improvement was achieved using multiphase CT texture analysis, in particular on haemangiomas tumors. However, the proposed system is unable to classify HCC lesion since it was able to classify only 4 cases out of 13 cases correctly.

The most recently fully automated classification system proposed by (Kumar et al., 2013a), specialised in differentiation between HCC (malignant) and hemangioma (benign). From each ROI, four texture features set (gray level, co-occurrence of gray level, wavelet coefficient statistics, and contourlet coefficient statistic) were extracted. A probabilistic neural network classifier was used in tumor classification. The highest accuracy achieved is 96.7%, which had been obtained with contourlet coefficient co-occurrence features. However, the proposed system can be extended for other types of liver diseases but the performance measures and accuracy mainly depend on the number of samples used.

Author	Year	Dataset	Features	Accuracy
(Kretowski, 2002)	2002	Normal (192); Metastases (405)	FOS; GM; COM; RLM	92.3%
(Quatrehomme et al., 2012)	2012	HCC (13); Metastases (38); Cyst (25); Adenomas (10); Haemangiomas (9)	LTE; UHS; MRE; FOS;	-
(Duda et al., 2013)	2013	Normal (573); Cholangiocarcinoma (222); Cirrhosis (433); HCC (319);	RLM; FOS; GLDM; COM; TFCM; LTE; GM; AC	90%
(Kumar et al., 2013a)	2013	Malignant (150); Benign (150)	Gray level; GLCM; Wavelet coefficient statistics; Contourlet coefficient statistic	96.7%

Table 3.3: Liver lesion classification literature work based on combined (statistical and structural) texture features.

3.3.3 Combined Feature

In contrast with the previous techniques of using texture only to classify liver lesions, a mix of texture with other features (intensity, shape), were used in the literature as well. Table 3.4 displays a different classification system based on combined features (Intensity, texture and shape features).

In 2005, (Lee et al., 2005) proposed a system to classify three different types of lesions, namely: Cyst, Hemangioma, and HCC. Lesions were drawn by experienced radiologist from non-contrast CT image. The combination features (mean gray level, entropy, local variance) were extracted to fed Back-Propagation Cerebellar Model Articulation Controller (BP-CMAC) neural network classifier. The accuracy of proposed system recorded 87%. However, the proposed system based on extracted features directly from given ROI not provide robust and accurate performance.

Later, the approach was extended by utilising shape descriptor in addition to GLCM features (Lee et al., 2007). A sequential forward selection algorithm was used to reduce feature space and adopted 4-layer pyramid scheme in classification. First layer in classifier distinguished between normal and abnormal liver tissue and cyst from abnormal liver tissue in the second layer. The third layer used to identify Hemangioma and last

layer recognised HCC from undefined liver tissues. Three different types of classifier were used, namely: SVM, MILP, and RBF neural network. The accuracy of classification was 89.5%, 82.1%, and 86.7% respectively. However, the proposed system is confused with hepatoma and cavernous hemangioma diseases and cause false positive results. Recently, the same approach was improved by utilising four features from the ROI, namely: edge, roundness, contrast, and internal texture (Lee et al., 2014). The extracted feature sets were fed to SVM classifier to classify the lesions with accuracy 93.7%.

Another approach based on portal phase of contract enhancement CT image was proposed by (Yang et al., 2013b). The proposed system was able to detect and classify three types of tumors, namely: hyperdense, hypodense, and heterogeneous. The combination features (Fast discrete curvelet transform, Biorthogonal wavelet, Histogram, and intensity superpixel) were extracted to fed Naive Bayes Nearest Neighbor (NBNN). The accuracy for classification has achieved 93%. However, the important advantage of nonparametric method when compared with other methods, that required training for detection of liver lesions and classification and need adjust the algorithm parameters carefully, which makes it flexible and easy to implement. On the other hand, Image dataset and the local descriptors are considered as main role in the detection performance and only depend on single-phase CT slices to detect lesion.

Additional study in liver lesion classification was provided by (Doron et al., 2014b). The combination of texture features (GLCM, LBP, Gabor, GLBP) and intensity feature (gray level intensity) are obtained from a given lesion. For classification module, SVM and KNN classifier were used to distinguish between four types of liver tissues, namely: Cyst, Hemangioma, Metastases, and Healthy tissue. The best result of 97% accuracy was obtained with combination of Gabor, LBP and Intensity features using SVM classifier. However, the main disadvantage of LBP feature is that the spatial relations among LBPs are often eliminated within the LPB histogram generation process, because they are selected in a single histogram and results to loss of global image information (Guo et al., 2010). While, computation complexity is considered the main main disadvantage of Gabor wavelet, due to producing a large number of redundant features at different scales (Baaziz et al., 2010).

The semi-automatic lesion classification system based on multiphase CT images was proposed by (Chang et al., 2017). Three types of features were extracted from the lesions in each phase, including texture, shape and kinetic curve. The GLCM texture feature was calculated 3D texture data of the lesion. The 3D shape features were obtained using compactness, elliptic model and margin to describe the lesion shape. A kinetic curve was created from each phase of CT image to represent differences in density of the lesion between each phase. The most useful features were found using a feature selection, based on Backward elimination. The extracted feature sets were fed to binary logistic regression classifier to classify the lesions with leave-one-out

cross validation. A total of 71 cases including 49 benign and 22 cases of malignant were used to evaluate classification performance. The highest accuracy of 81.69% was achieved through combining all of the features. However, the majority of the dataset cases were benign lesions, where only 30% of the cases are malignant.

Author	Year	Dataset	Features	Accuracy
(Lee et al., 2005)	2005	Cyst (55); Hepatoma (33); Hemangioma (33)	Mean gray level; Entropy; Local variance	87%
(Lee et al., 2007)	2007	Cyst (76); HCC (30); Haemangiomas (40)	GLCM; Shape descriptor	89.5%
(Yang et al., 2013b)	2013	Hyperdense; Hypodense; Heterogeneous	Fast discrete curvelet transform; Biorthogonal Wavelet; Histogram; Intensity superpixel	93%
(Lee et al., 2014)	2014	Cyst (76); HCC (30); Haemangiomas (40)	Edge; Roundness; Contrast; Internal texture	93.7%
(Doron et al., 2014b)	2014	Cyst (43); Haemangiomas (24); Metastases (25); Healthy tissue (20)	GLCM; LBP; Gabor; GLBP; Gray level intensity	97%
(Chang et al., 2017)	2017	Malignant (22); Benign (49)	3D texture (GLCM); 3D shape (Compactness, Elliptic model and Margin); Kinetic curve	82.69%

Table 3.4: Liver lesion classification literature work based on combined features (intensity, texture and shape features).

In summary: According to the previous literature about liver lesion classification based on texture and combination features extracted from lesion ROI, the majority of researcher are focused on feature extraction, and usually using the absolute value of features that extracted from lesion area. Moreover, the statistical texture features gained more attention comparing to other types of texture features or intensity and shape feature (Duda et al., 2004). However, the characteristics of malignant lesions

differ from benign lesions in terms of shape, boundary and effect on surrounding liver tissues (Nicolau et al., 2006; Assy et al., 2009b; Murakami and Tsurusaki, 2014a). Furthermore, the classification through black box low-level features is meaningless for radiologist because it does not provide the understandable information behind the classification decision. There exists a limited work that benefited from all the characteristics of the lesion such as border, shape and surrounding area. In addition, interpreting the classification results through high-level features. Hence, our proposed system utilises all the lesion characteristics (internal, border and surrounding area) with the combined between intensity, texture and shape features to enhance the system accuracy. In addition, proposing multiple ROIs to calculate the high-level features by considering the ability of each ROI that represents a set of characteristics, and then using high-level features to classify the lesion. In contrast with most existing research, which use low-level features only, the use of high-level features and characterisation helps in interpreting and explaining the classification and is more intuitive to clinicians.

3.4 Liver Lesion Characterisation

Recently, Liver lesion characterisation has become one of the major research topics in the field of medical imaging and diagnostic radiology. This research not only concerns in tumor classification, but also characterisation the tumor through describe it (size, location, boundaries).

CaReRa (Case Retrieval in Radiology) (Reddy and Faust, 2006) is a prototype Content Base Case Retrieval (CBCR) implementation of the Clinical Experience Sharing (CES) concept, which focuses on liver cases. The objective of CaReRa is search and retrieves past cases relevant to the query case. Radiologists describe their observation of CT image in free-text radiology reports. Unfortunately, radiology reports written in natural language are often vague, incomplete and error-prone, which makes them challenge for automatic processing (Oberkampf et al., 2015). Moreover, it becomes difficult to retrieve valuable information from reports because of their unstructured nature (Iroju and Olaleke, 2015). Hence, it is beneficial to represent the observations in a structured manner. To represent imaging observations of the liver domain, The Ontology of Liver for Radiologists (ONLIRA) has been developed based on the radiologists' manual annotation process (Seghers et al., 2007c). It enables description of liver CT in a structured way with an emphasis on its properties and the relations between liver, veins and liver lesions (Marvasti et al., 2017).

ImageClef (Caputo et al., 2014) was part of the Cross Language Evaluation Forum (CLEF) 2014. ImageClef has four main tasks; Liver CT annotation is one of these tasks. The objective of this task was to answer on multiple choice questions to annotate liver based on an analysis of the image features to generate a structured report (Caputo et al., 2014). The dataset provided for this task contained 50 CT images for training

and 10 CT images for testing, segmented liver and lesion. All the participants (Kumar et al., 2014; Spanier and Joskowicz, 2014; Nedjar et al., 2015) were asked to answer questions that generated from ONLIRA (Kokciyan et al., 2014).

This section will present the most relevant key works regarding liver lesion characterisation following the categorisation depicted in Figure 3.2. The presented literature in this section covers work that only focus on semantic features (high-level features) to characterised liver lesion. This is in contrast with the previous section that included lesion classification that operated on low-level features. The literature of this section will be categorised based on the approaches that used to characterise lesion. The techniques that used machine learning to characterise lesion are presented in Section 3.4.1, lesion characterisation based on Case-based Retrieval is presented in Section 3.4.2.

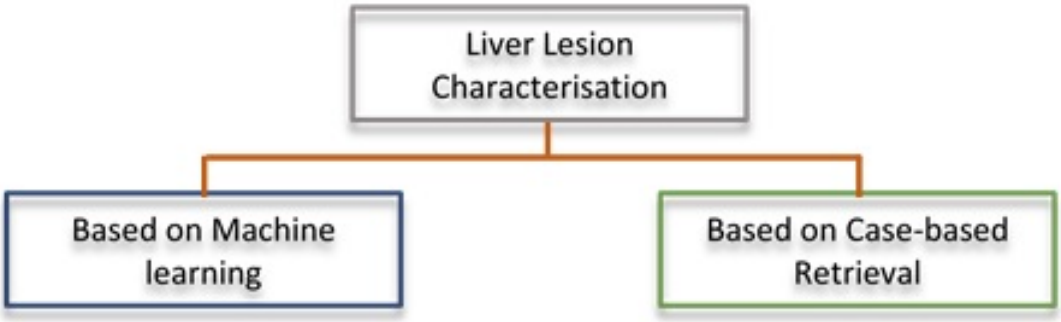


Figure 3.2: Literature on liver lesion characterisation.

3.4.1 Based on Machine Learning

Machine learning approaches are widely used in medical imaging classification (Pourghassem and Ghassemian, 2008; Setia et al., 2008) and annotation (Tommasi et al., 2008; Wennerberg et al., 2011). The extracted low-level features are fed as input to machine learning to relate them to high-level features. (Gimenez et al., 2012) proposed a liver lesion characterisation system to predict radiological observations from CT images in the portal phase. Intensity, texture and shape features were extracted from the 79 pre-identified lesions. The least absolute shrinkage and selection operator (LASSO) was used in lesion characterisation. A Leave-One-Out cross validation was used to evaluate the system performance. This work was later extended by (Agarwal Vibhu, 2013) by implementing a stack of two auto-encoders consisting of the 21x21 randomly sampled patches. The neural network and 10-cross validation were used to evaluate the system performance. The system accuracy is not significant to the results obtained in the study proposed by (Gimenez et al., 2012) (this was validated using a 2 sample t test that gave a p-value of 0.74, rejecting the null hypothesis). However, the proposed methods are only characterised binary semantic features that can be presented by positive or negative observation, for instance, whether the lesion is heterogeneous or not.

Regarding the utilisation of semantic features, (Depeursinge et al., 2014) proposed a framework to predict semantic terms of the liver lesions using linear combination of texture features based on Riesz wavelets and SVM. The lesion ROI was divided into 12x12 patches to analyse the lesion structure. The feature vector was built by applying the Steerable Riesz wavelets for each patch. The distances between all visual semantic terms are calculated to establish a non-hierarchical computationally-derived ontology of visual semantic terms containing inter-term synonymy and complementarity. The computationally-derived ontology is complementary to the RadLex ontology because it allows connecting semantic concepts with their actual appearance in CT images. For example, heterogeneous and homogeneous feature are very close to each other in RadLex because they both describe the uniformity of lesion enhancement, but they are opposed to each other in the computationally-derived ontology since they are visually antonymous in terms of texture characterisation. The system evaluation used 74 liver cases annotated with 18 visual semantic features set. Leave-one-patient-out cross-validation resulted in an average area under the ROC curve of 0.853 for predicting the presence of each semantic feature. However, the proposed system is limited to characterise some of the important semantic features such as overall lesion shape and discrimination between the margin of the lesion and lesion shape.

According to the ImageClef task, (Kumar et al., 2014) designed multi-class classification scheme based on SVM to annotate liver lesions. Each semantic feature (question) has multiple answers (labels) such as lesion shape (regular, irregular, round, oval,...), for each answer has a classifier that is trained to separate it from other answers. Two stages of SVM were used to annotate each semantic feature where the second stage is activated only when the first stage assigned multiple labels. The first stage is 1-vs-all SVM and the second stage is 1-vs-1 SVM classifier. The best overall accuracy of 91% was obtained by the RBF kernel with scaling factor equal 1. (Spanier and Joskowicz, 2014) tried four different types of machine learning algorithm (linear discriminant analysis (LDA), logistic regression (LR), K-nearest neighbors (KNN), and SVM) to link between extracted low-level features and the high-level features based on training dataset. Leave-one-out cross validation was used to build a training model for every semantic feature and classifier. The lesion location (lobe and segment) was estimated from the image itself by measuring the centre of the lesion and the liver. The lesion characterised in the right lobe if the lesion centre is on the right part of the liver. Almost the same performance was obtained by using any previous classifier, with accuracy of 91%. However, the location of the lesion was characterised in a primitive way with high probability of error where the localisation of lesions depends on knowledge of liver anatomy (Sibulesky, 2013; Majno et al., 2014).

Another recent ImageClef task participant, (Nedjar et al., 2015) proposed a classification approach to characterise liver lesion. The characterisation process divided into two phases. The first stage is pre-processing to segment the actual lesion from

given lesion box. Thereafter, features vector is generated through extracted texture and shape features from the segmented lesion. The second stage is classification process. Each semantic feature is classified by using random forest classifier. The overall characterisation accuracy was 82.5%

Later, (Kurtz et al., 2015) proposed a new framework that enables retrieving similar images based on semantic features and user feedback. The given lesions are automatically annotated with semantic terms. Each ROI is divided into 12x12 patches extracted from internal and margin of the lesion. A gray-level intensity histogram and multi-scale Riesz wavelets were extracted from each patch to build the feature vector. The SVM was used to annotate lesions. The predicted semantic features were used to enhance the performance of retrieving similar images. The system evaluation used 72 images from 42 patients annotated with 18 visual semantic features set. The area under the receiving operator characteristic (AUC) curve was used to measure the annotation system performance, 0.76 for all 18 semantic features. Table 3.5 represents a different characterisation system based on machine leaning.

Author	Year	Dataset	ML	Validation	Accuracy
(Gimenez et al., 2012)	2012	79 case	LASSO	Leave-one-out	81.6%
(Agarwal Vibhu, 2013)	2013	79 case	Neural Network	10-cross validation	89.5%
(Depeursinge et al., 2014)	2014	74 case	SVM	Leave-one-patient-out	ROC curve 0.853
(Kumar et al., 2014)	2014	50 case	SVM	10 cases	91%
(Spanier and Joskowicz, 2014)	2014	50 case	LDA; LR; KNN; SVM	10 cases	91%
(Nedjar et al., 2015)	2015	50 case	Random forest	10 cases	82.5%
(Kurtz et al., 2015)	2015	72 case	SVM	Leave-one-patient-out	AUC curve 0.76%

Table 3.5: Liver lesion classification literature work based on machine leaning.

3.4.2 Based on Case-Based Similarity

Content-based image retrieval (CBIR) is a computer vision technique that gives a way for searching relevant images in large databases (Müller et al., 2004). The low-level features (texture, intensity and shape) of the lesion are utilised to select which cases are similar to the query case (Smeulders et al., 2000). Another method for liver characterisation task by using content-based image retrieval technique was introduced in (Kumar et al., 2014). The most similar training images were used to select the answers for query un-annotated image. The Euclidean distance was used to calculate the simi-

larity between image feature vectors. Then, a weighted voting scheme was applied to select more similar cases to the test image to answering questions. This work was later extended by (Kumar et al., 2016). The weighted nearest-neighbour (WNN) search with sequential feature selection used to find the most similar training images to select the labels for the test images. However, the higher frequency of the labels in the dataset would have a greater chance of being selected as the best answer. (Nedjar et al., 2015) proposed another method based on retrieval-based approach to characterise liver lesion, by using a specific signature of the liver. The CT image was normalised into a rectangular block with the size of 200x190 and encode it by applying 1D Log-Gabor wavelets. The output of the Gabor filter was divided into a small block of size 5x5. For each block, the dominant angular direction was quantised into four levels by using the Daugman algorithm. After that, the Hamming distance (HD) was used as a similarity metric to retrieve the top five similar images to the query image. The majority voting between retrieved images has been used in characterisation. The experiment showed a slightly improved in accuracy compared to the classification approach; with overall accuracy of 83.6%. However, the dataset is small compared to the number of annotations. Moreover, there are some of the annotations such as the lesion location cannot be estimated correctly through CBIR.

In summary: according to previous literature about liver lesion characterisation, the majority of works have been done through machine learning technique. Through literature, the existing works predicted the semantic description of the lesion but did not characterise the effects of the lesion on the anatomical structures as well as the relation between the lesion and the surrounding liver area. Furthermore, lack in characterised some of the high-level features that rely on the case itself such as the proximity of the lesion to the hepatic vasculature or characterised lesion in a primitive way with the high probability of error such as lesion location. In addition, some of works used unbalanced dataset such as ImageClef dataset (Kumar et al., 2014; Spanier and Joskowicz, 2014; Nedjar et al., 2015; Kumar et al., 2016), where it has limited training data compared to the number of semantic features and more than 70% of the high-level features not applicable or it has the same semantic features for all the dataset. Hence, the proposed lesion characterisation system will be divided the semantic features into two groups: (1) high-level features that extracted from the case itself such as lesion location. (2) high-level features that predicted through training data. Moreover, proposed multiple ROIs (internal, border and surrounding lesion) to captured all the lesion characteristic. In addition, the balanced dataset will be used in training task that have enough training data compared to the high-level features.

3.5 Connection the Thesis with Previous Studies

The previous works have shown different techniques used to classify/ characterise the liver lesions from CT images. The segmentation process is considered the first step for semi-automated and fully automated CAD systems. However, all CAD systems share two main stages: the feature extraction and the classification/ characterisation stage.

In this thesis, particular attention was paid to the efficiency of the proposed algorithms as well as their accuracy. Regarding the segmentation process efficiency, the grey level-based approach was adopted for this task. This due to this approach has the advantages of a low computational cost, no training requirements and no user interaction (Ruskó et al., 2009a; Gloger et al., 2010). The prior medical knowledge, preprocessing and post-processing were used to enhance the algorithm accuracy. Thus, the reasonable balance between accuracy, robustness and computational cost in framework design can offer a suitable solution for clinical use.

This chapter has also given an overview of different CAD systems for liver lesion classification and characterisation. The majority of the proposed works used the texture features for lesion diagnosis, as the combination features achieved promising results in this task. In this thesis, the statistical features such as GLCM will be employed, due to its considered the most popular method to drive the spatial statistical texture features, providing information about the spatial arrangement and intensities distribution in the image, and also outperforming other techniques such as wavelet features (Bayram et al., 2011). Furthermore, the intensity and shape features will be also used. This is because the malignant lesions differ from benign lesions not only in surface texture but also in shape, boundary and intensity (Nicolau et al., 2006; Assy et al., 2009b; Murakami and Tsurusaki, 2014a).

However, the liver lesion classification and characterisation accuracy is usually affected by detecting lesion appearance in CT image. These characteristics are observed differently according to the region of interest selection approaches. Lesion characterisation based on CT image methods, using existing ROI selection approaches, have a limit to represent all the lesion characteristics such as the relation between liver and lesion. Thus, the performance of the CAD systems using the current ROI selection methods was variable according to the feature extraction techniques. Hence, to overcome this limitation and obtain a better and more stable framework performance than the current methods, we proposed a multiple ROIs selection approach to well-represent the lesion characteristics.

It is crucial to emphasise that, to the best of my knowledge, all previous studies and others in literature used hand-designed features (such as texture, intensity, etc.) which are fed into a classifier (such as SVM, RF, etc.), in an attempt to classify liver lesions. However, the main limitation for using hand-designed features in lesion classification is that the diagnosis decision cannot be explained in human-level understanding, mak-

ing it less reliable for physicians. Hence, to overcome this drawback, the proposed lesion classification framework will benefit from the lesion characterisation. This will be done through utilising the high-level features to classify liver lesions. The use of high-level features provides a human-interpretable explanation of the lesion diagnostic decision to better-trusted diagnosis.

3.6 Conclusion

This chapter started with an overview of the various CAD systems based on CT image has been presented. Then, a categorised review of related literature on classification and characterisation was discussed. The review highlighted each one's approaches and limitations and how the proposed system will address each of them. The disease diagnosed in most cases were malignant such as HCC, metastases, benign tumors such as cyst, hemangioma, which distinguished from healthy liver tissue. Texture features were used widely in classification tumor system such as COM, RLM, FOS, FM, LTE and GLCM as mentioned in the literature. Different datasets (tumors types, tumors size, dataset size, image resolution) were used in different classifiers. Furthermore, different evaluation approaches were used to calculate the system performance. As a result, it is difficult to infer the best features and classifier method which can be used in classification/ characterisation. However, some texture features have shown to be reliable in different CT image phases. For instance, GLCM and COM texture feature has been utilised successfully to enhance the classification accuracy for different CT image phases (non-enhanced and enhanced images). While FM feature was frequently considered just for one phase of CT image (non-enhanced). In addition, FOS and RLM features were most often used for enhanced CT image (after administration of contrast agent). Texture features gained more attention, while Intensity and shape features used in segmentation and lesion detection. Combination features (texture, intensity, and shape) gained promising results compared to texture features.

As overall conclusion, most of literature work was diagnosed liver lesion by extracting the features from the lesion only and not paying much attention to the relation between lesion and surrounding area. The selected ROI in existing works have a limit for representing all liver lesion characteristic. Thus, the performance of these systems was variable according to used feature extraction approaches. Moreover, there are some of the semantic features such as the lesion location cannot be estimated correctly through CBIR or machine learning. The proposed system will utilise the lesion and surrounding area to capture all the characteristic of the lesion where multiple ROIs area are selected by considering the ability that each ROI represents the kind of characteristics and exploiting relationships between low-level and high-level features.

Chapter 4

Liver Image Analysis in CT

This chapter shows the overall structure of our computer aided detection framework for liver segmentation, lesion detection and vessels extraction. The aim of this thesis is to develop an automated CAD system for liver lesion with the main focusing on lesion characterisation/classification.

Hence, the liver lesion characterisation and classification process will be introduced in the next chapter in details. This chapter is organised as follows: Section 4.1 will introduce the automatic liver segmentation framework. Section 4.2 will present the automatic lesion detection. Section 4.3 will discuss the framework for the main vessels extraction. Section 4.4 will discuss the obtained segmentation results. Finally, the chapter is concluded in Section 4.5.

4.1 Liver segmentation

Liver segmentation from abdominal images is an important step in many diagnostic and surgical procedures. The manual liver segmentation process is subjective and very time consuming (Suzuki et al., 2013), because a radiologist has to extract the liver on many CT slices (Suzuki, 2011). In addition, the accurate segmentation depends on his/her experience (Arakeri et al., 2011). Figure 4.1 depicts the block diagram of an automatic liver segmentation process using CT images.

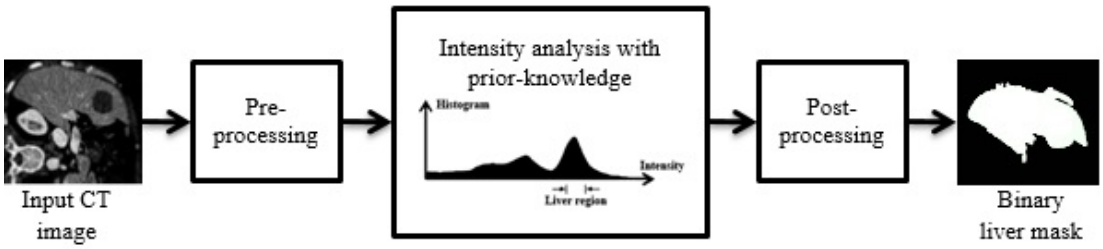


Figure 4.1: Detailed block diagram of Our framework for liver segmentation.

The our framework has focused mainly on three steps for liver segmentation:

- Pre-processing step is used to enhance the CT image quality and remove the noise to get more surety and ease in segmenting the liver. Noises in the CT are the main cause of false-positive and false-negative segmentation results, which may lead to relatively high error rates of segmentation result.
- Intensity analysis with prior-knowledge to extract liver area automatically.
- post-processing step is used to enhance the segmentation result.

4.1.1 Pre-processing

CT is one of the important imaging modality that provides a high contrast of body tissue or organs with good spatial resolution (Zhu et al., 2012; Kaur and Juneja, 2018b). The quality of CT image is important in medical diagnosis that can be affected by the presence of artifacts during acquisition procedure. The most commonly affected noises in medical CT image are impulse noise "Salt and Pepper", Gaussian and Poisson noise (Javed et al., 2016). The noise in CT image affects both the diagnosis process and automated computerised analysis tasks, such as segmentation, three-dimensional image reconstruction, and visualisation. Therefore, reducing noise in the CT image is considered an important step, because it helps to improve the performance of the image analysis such as segmentation (Bhadauria and Dewal, 2012).

The main goal of applying the pre-processing step is to enhance the image quality, smoothness and reduce the noise that occurred by defects of CT scan device. In addition, it emphasises beneficial image features and quality for better segmentation accuracy and speed. Image filtering is a major pre-processing method used for many purposes including smoothing, sharpening and contrast stretching. In practice, as well as reducing noise, it is important to preserve the edges of the image where the edges provide important information about the visual appearance of the image. Because of this, anisotropic diffusion filter (ADF) and the median filter have been adopted for this task.

Perona and Malik proposed an edge detection and multiscale smoothing algorithm called anisotropic diffusion filter (ADF) (Perona and Malik, 1990). ADE addresses the limitation of spatial filtering and improves image quality by preserving the edges of the object as it eliminates noise in similar areas and ensures edge sharpening. However, this algorithm usually removes small image detail and change image statistics because of its edge enhancement effect. This filter describes the problem in the form of a heat equation based on the second-order partial differential equation (PDE) in an anisotropic medium, and it can achieve a good trade-off between noise removal and edge preservation (Li et al., 2013). Here, image smoothing is framed as a diffusion process that can be stopped or repressed at edges by selecting the strengths of the local gradient in various orientations. In this algorithm, boundaries are preserved by con-

volving the image I in the direction orthogonal to image gradient. The ADF process is given by Equation 4.1.

$$\frac{\partial I(y,i)}{\partial i} = div\left(c\left(y,i\right)\nabla I\left(y,i\right)\right) \tag{4.1}$$

Where $\nabla I\left(y,i\right)$ gives the value of image gradient at voxel y and iteration i, $\frac{\partial I(y,i)}{\partial i}$ is the partial derivative of $I\left(y,i\right)$ and the edge-stopping function $c\left(y,i\right)$ is illustrated in Equation 4.2.

$$c\left(y,i\right) = g\left\|\nabla I\left(y,i\right)\right\| = e^{-\left\|\nabla I(y,i)\right\|/R^2} \tag{4.2}$$

Where, R is the diffusion parameter and c is the flux function that controls the rate of diffusion.

On the other side, Median filter is a nonlinear approach used to remove noise from images. The median filter is demonstrably effective at removing noise whilst keeping edges for a given, fixed window size (Chang and Chu, 2012). Furthermore, Median filter is particularly effective at removing "salt and pepper" type noise. Because of this, the median filter is widely used in filtering the image by moving through the image pixel by pixel, replacing each value with the median value of neighbouring pixels. The median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the middle (median) pixel value.

In this work, the noise present in the CT image is removed by employing a 3x3 median filter, as depicted in Equation 4.3. The median filter was selected due to its retains the edge information within the input image where Gaussian and Mean filters tend to blur the edges in the image (Moghbel et al., 2016). This is because the Median filter does not create new unrealistic pixel values in the case of the filtering window laying over an edge.

$$f(x,y) = median_{(3,3)}[I(x,y)] \tag{4.3}$$

Where *f(x,y)* shows the filtered image and *I(x,y)* shows the original image before applying median filter. Figure 4.2 shows the original CT image and the output images achieved after de-noising algorithm ADF and median filer respectively.

The visual interpretation of the images is used to differentiate between noisy and denoised image content. However, to further validate the results at the quantitative level, some objective criteria were used such as Peak signal to noise ratio (PSNR), and they do not rely on the visual appearance of the images. This parameter computes PSNR between two images as a quality measurement between the original and filtered image.

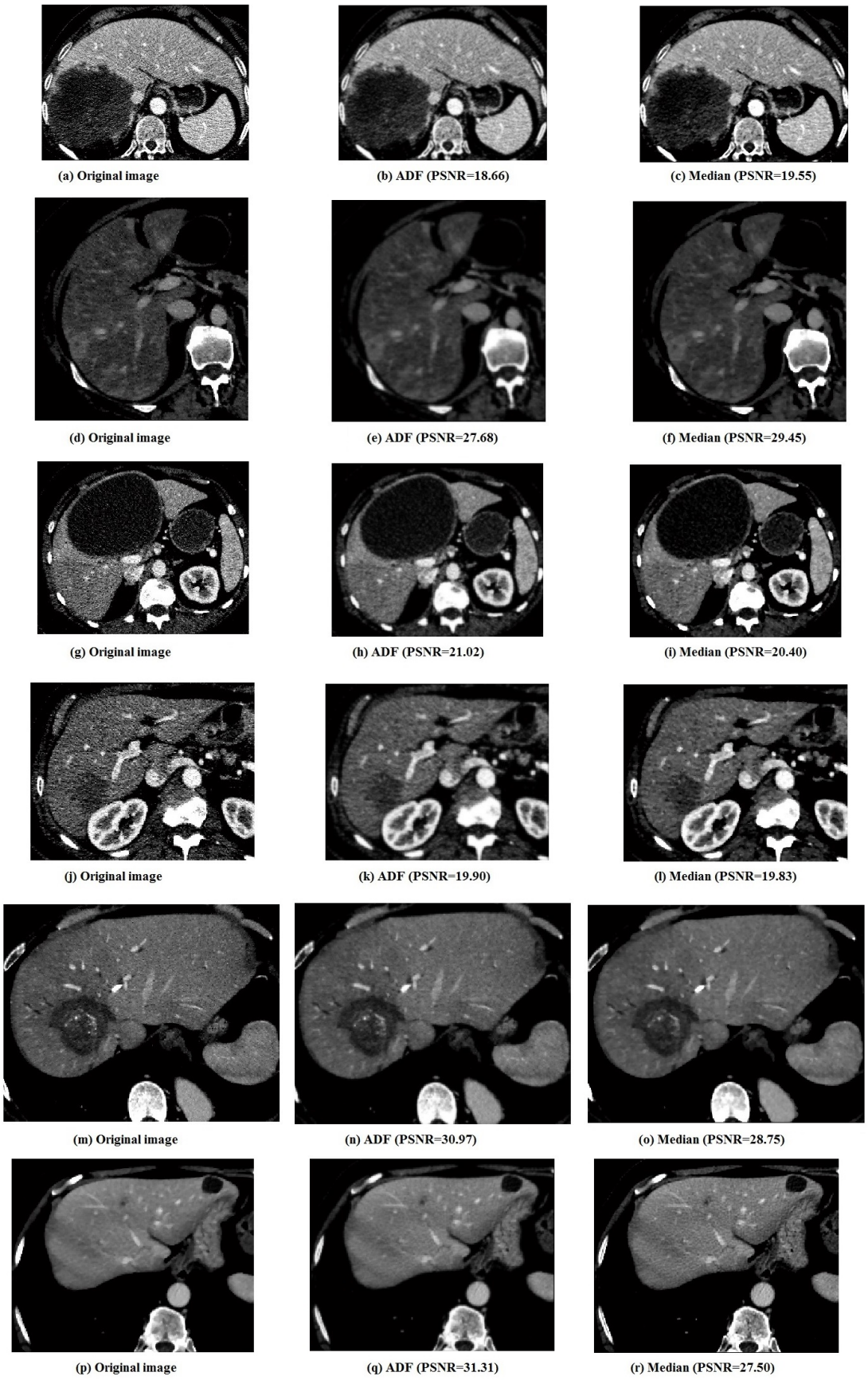


Figure 4.2: Pre-processing step to remove the CT image noise. (a, d, g, j, m, p) the original input CT image. (b, e, h, k, n, q) the output CT image after ADF. (c, f, i, l, o, r) the output CT image after median filter.

The Mean Square Error (MSE) and the PSNR are the two error metrics used to compare noise reduction quality of the image. The MSE represents the cumulative squared error between the denoised and the original image, whereas PSNR represents a measure of the peak error. The lower the value of MSE, the lower the error. In contrast, the higher PSNR value means a better noise reduction algorithm. Equation 4.4 and Equation 4.5 depicts the respective formulas to compute MSE and PSNR respectively.

$$MSE = \frac{1}{mn} \sum_{y=1}^{m} \sum_{x=1}^{n} \left(I(x,y) - I'(x,y) \right)^2 \tag{4.4}$$

$$PSNR = 10 \log_{10} \left(\frac{(I_{\max})^2}{MSE} \right) \tag{4.5}$$

Where I is the original with image dimensions m and n. I' represents the image after noise reduction process. $I_{\max}$ is the maximum pixel value for the image.

The performance of both filters (median filter and ADF) is evaluated through PSNR metric over the 50 cases. The Figure 4.3 depicts the graphical results of PSNR for ADF and median filter approach, which demonstrates that ADF has better results in some cases as compared to the median filter algorithm. However, the median filter is better than ADF algorithm in the low quality of CT images. This is due to the CT image with the low quality has black and bright pixels, causing noise in the image where the dark and bright pixels are reduced better by using the median filter.

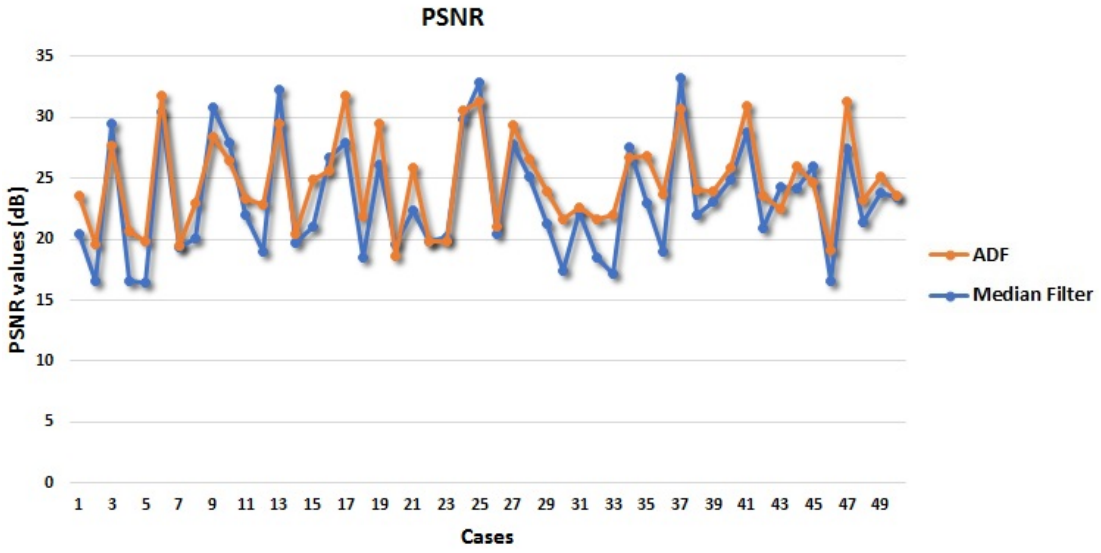


Figure 4.3: Pre-processing step to remove the CT image noise. (a, d, g, j, m, p) the original input CT image. (b, e, h, k, n, q) the output CT image after ADF. (c, f, i, l, o, r) the output CT image after median filter.

In this thesis, regarding the experiments, the median filter provided a good results

compared to ADF method. In addition, the median filter computational cost is significantly lower than the ADF algorithm. Hence, the median filter was adopted to fulfill this task.

4.1.2 Liver Segmentation

First step in semi/fully automated diagnostic system is liver segmentation then lesion detection. The variety of sizes and shapes of the liver from a patient to another, and similar intensity of liver with other organs, makes the liver segmentation a task difficult (Soler et al., 2001b; Seo et al., 2005). However, manual segmentation from clinical setting perspective is tedious and excessively time consuming (Militzer et al., 2010).

After enhancing the input CT image by reducing the noise and smoothing the image, the next step of the framework is segmenting the liver from CT scan. The liver segmentation is done through CT intensity analysis and prior-knowledge, which is defined as assumptions based on medical knowledge. The framework is based on the following medical assumptions:

- The liver is the largest internal organ in the human body (Bandiera et al., 2015).
- The liver is located in the front of the abdominal cavity in the right quadrant (Ger, 1989; Stringer, 2014).
- The contrast agent typically makes the liver appear brighter than the other abdominal organs. Where the liver absorbs more contrast agent more than other organs (Sahani and Kalva, 2004a).
- The liver parenchyma in CT image is almost homogeneous(Han et al., 2015).

For the liver segmentation, the framework used a contrast-enhanced CT images that are acquired in the portal phase. where the contrast agent makes the liver appear brighter than surrounding abdominal organs. In addition, liver is derived 80% of the blood from the portal vein (Kan and Madoff, 2008), which can be exploited to separate the liver from other organs. The histogram of the input CT scan is analysed. According to (Rusko et al., 2007a; Corson et al., 2011) the intensity of the liver in a portal phase is in the range of [-50, 250] HU. Figure 4.4 shows the HU intensity ranges for the abdominal organs in contrast-enhanced CT scan.

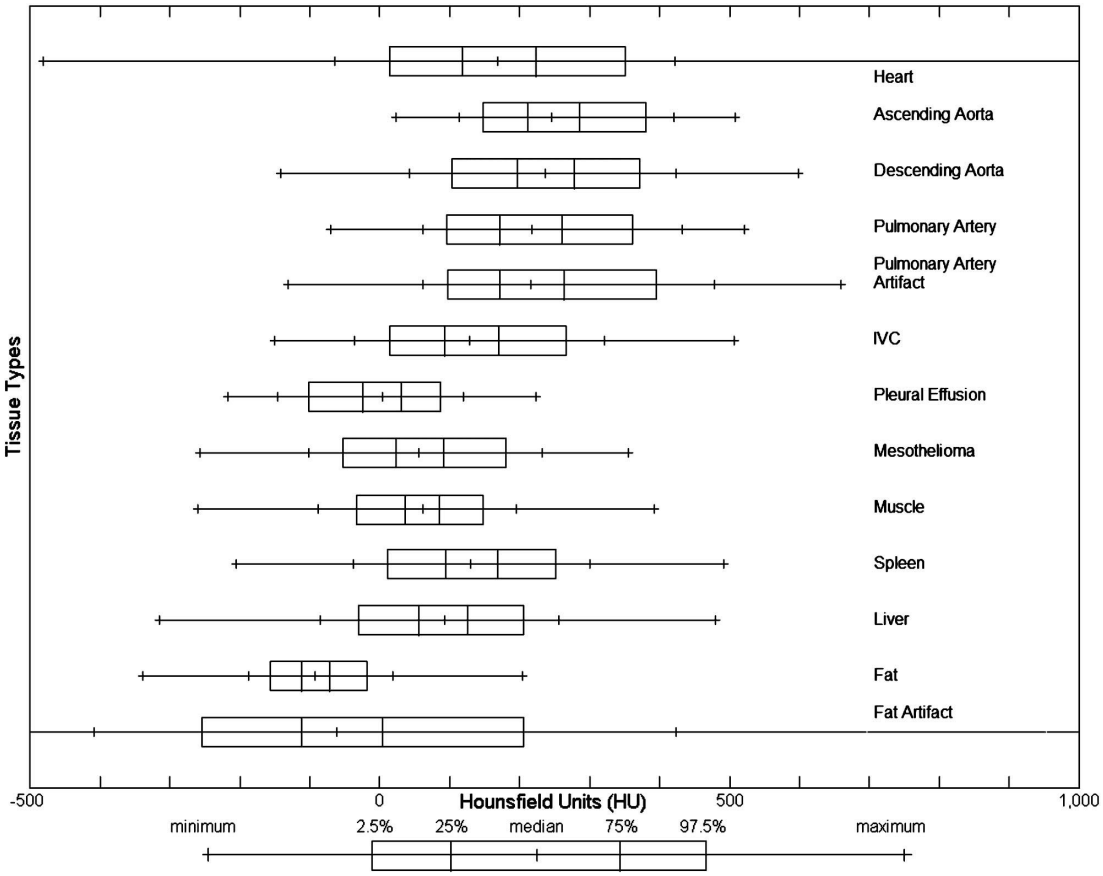


Figure 4.4: The distributions of HU values for different abdominal organs in contrast-enhanced CT scan (Corson et al., 2011).

The [-50, 250] HU intensity range was applied on CT image to excludes air, fat and bones that having voxels intensity out of this range, which makes the localisation of the liver more robust. According to our assumption that liver is located in the right quadrant of the abdomen, the histogram of the right side with the range of [-50, 250] HU was calculated and smoothed to determine the mean liver voxels intensity, as depicted in Figure 4.5.

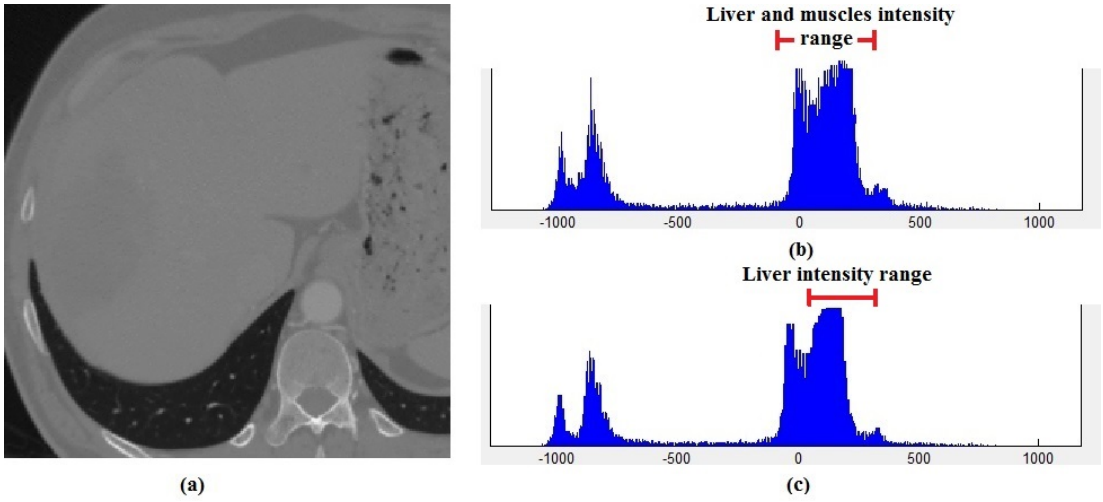


Figure 4.5: CT image and intensity range. (a) the original input CT image. (b) the histogram of CT image. (c) smoothed histogram of CT image.

After smoothing CT histogram as shown in Figure 4.5, it has two highest pitch that represents liver and muscles. Since liver intensity is higher than others (based on pre-knowledge),the first highest pitch belongs to the muscles and the second area contains the liver. Let *H(ml)* be the histogram of HU values of the liver and muscles where *m* corresponding to the first high intensity in the histogram and *l* corresponding to the second high intensity HU value. The liver likelihood HU intensity range is defined by [l_{min}, l_{max}] where l_{min} is the minimal point between the two highest pitches m and l; l_{max} is the lowest point in the tail of histogram. Then, a binary image mask for the liver area is created on the original CT image. The intensity of pixels which is in the determined range [l_{min}, l_{max}] are assigned value 1 and otherwise lower or higher than the range are assigned zero value. *M(i,j)* is binary mask for the liver defined as in Equation 4.6.

$$M(i,j) = \begin{cases} 1, & if\ \ l_{min} \leq\ f(i,j) \leq\ l_{max} \\ 0, & otherwise \end{cases} \tag{4.6}$$

Figure 4.6 displays the liver segmentation process. As the first step, the threshold of the liver intensity range is determined based on pre-knowledge and histogram-based adaptive threshold approach, as shown in Figure 4.6.a. For the adaptive threshold of the liver, a rough estimation of the intensity range [l_{min}, l_{max}] was computed from the histogram of *H(ml)*, where l_{min} and l_{max} denoted by the minimum and maximum of liver voxels intensity values, respectively. The second step, the binary liver mask is created by using the threshold l_{min} and l_{max} where the pixels having intensity in the range [l_{min}, l_{max}], as depicted in Figure 4.6.b. However, the segmented image includes pixels from the liver and the other organs as well, which have similar intensity range to the liver. In order to achieve more accurate segmentation, the post-process is applied that includes the morphological operation to refine the liver segmentation.

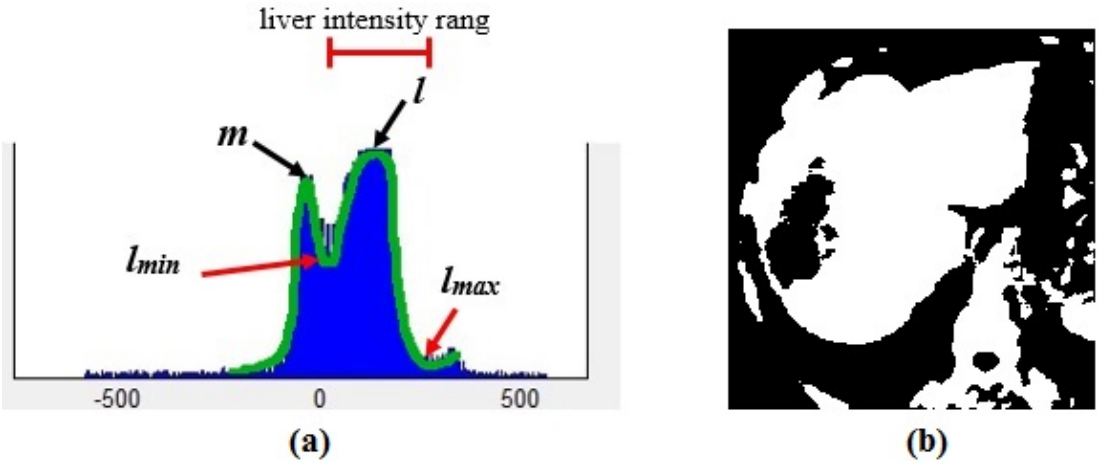


Figure 4.6: Selecting of liver intensity range from a CT image.

4.1.3 Post-processing

After segmenting the liver area from the CT image, the segmentation result usually contains not only liver tissue but also other organs, which have similar intensity to the liver. In addition, it has holes and connected pixels with the neighbour organs. Therefore, the post-processing based on morphological operation is applied on the segmented area.

The morphological process is used to delete certain unwanted region and filling the holes, which lead to detect the liver object and enhance the segmentation result. Usually a set of structure elements are used in morphological process, such as dilation, erosion, opening, and closing. The shape and size of structure element is important in selecting or extracting object to avoid of removing too many desirable objects, or keeping too many unwanted ones (Luo et al., 2009). Here a circular structure element with radius 3 is adopted which selected as a compromise between elimination of the unwanted area and the connected organs with each other and preservation of details the object. Therefore, the value of structure element was selected based on the prior-knowledge of the anatomical structure of the abdomen and the resolution of the CT image. Hence, the liver segmentation post-processing is applied as follows:

- Applied Erosion process to delete the fragments of other organs.
- Applied Connected Component Labeling algorithm (CCL) to identify connected pixel areas where the largest connect pixels is selected as liver region based on pre-knowledge (The liver is the largest internal organ in the human body).
- Applied Dilation process to fill the holes and reserve the pixels that removed from the liver by erosion process.
- Filled the remain holes inside the segmented region.
- Obtained the segmented liver by complemented and multiplied the resulting binary mask with the original CT image, as shown in Equation 4.7.

$$S(i,j) = \begin{cases} CT(i,j), & if \;\; M(i,j) = 1 \\ 0, & otherwise \end{cases} \tag{4.7}$$

  Where S(i,j) is a final segmented liver, CT(i,j) is the original CT image and M(i,j) is the liver mask.

Figure 4.7 depicts the liver segmentation process. The original CT image shown in Figure 4.7.a. The pre-processing median filter is used to smooth and remove the image noise to improve the segmentation process. The pre-knowledge and the histogram-based adaptive threshold is used to produce the initial binary mask for the liver, as

shown in Figure 4.7.b. The erosion process is used to separate the liver area from other organs, as shown in Figure 4.7.c. The CCL algorithm is applied to select the largest connected pixels area as considering the liver is the largest organ in the abdomen, as figured in Figure 4.7.d. The dilation and filling the holes is used to refine the segmented liver mask, as shown in Figure 4.7.e and Figure 4.7.f respectively. Figure 4.7.g shows the final results for liver segmentation.

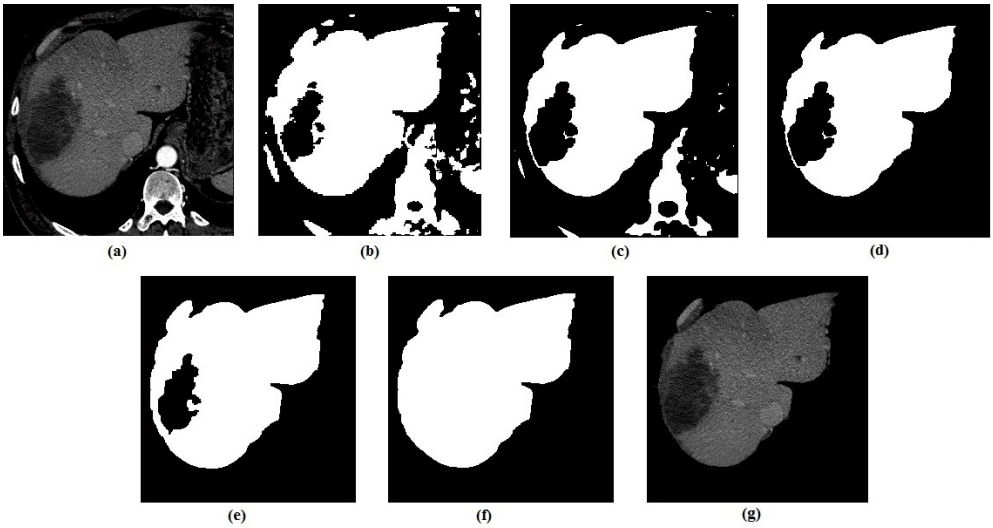


Figure 4.7: Steps of liver segmentation process. (a) The original input CT image. (b) The initial binary liver mask. (c) Operates erosion morphology. (d) Select the largest connected pixels as the liver is the largest organ that appear in CT image. (e) Operates dilation morphology. (f) The final liver mask after filling the holes. (g) The final result of the liver segmentation.

Figure 4.8 illustrates the graphical evaluation results of liver segmentation based on Dice similarity coefficient (DSC) metrics. The vertical axis (x-axis) represents the number of cases that achieved the DSC values in the horizontal axis (y-axis). The highest DSC value means the better segmentation result.

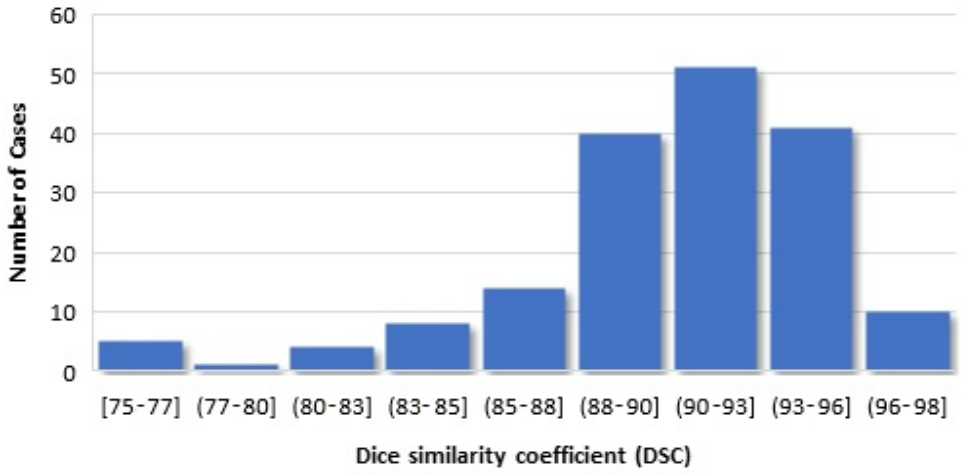


Figure 4.8: Number of cases and obtained DSC values of liver segmentation for evaluating the method.

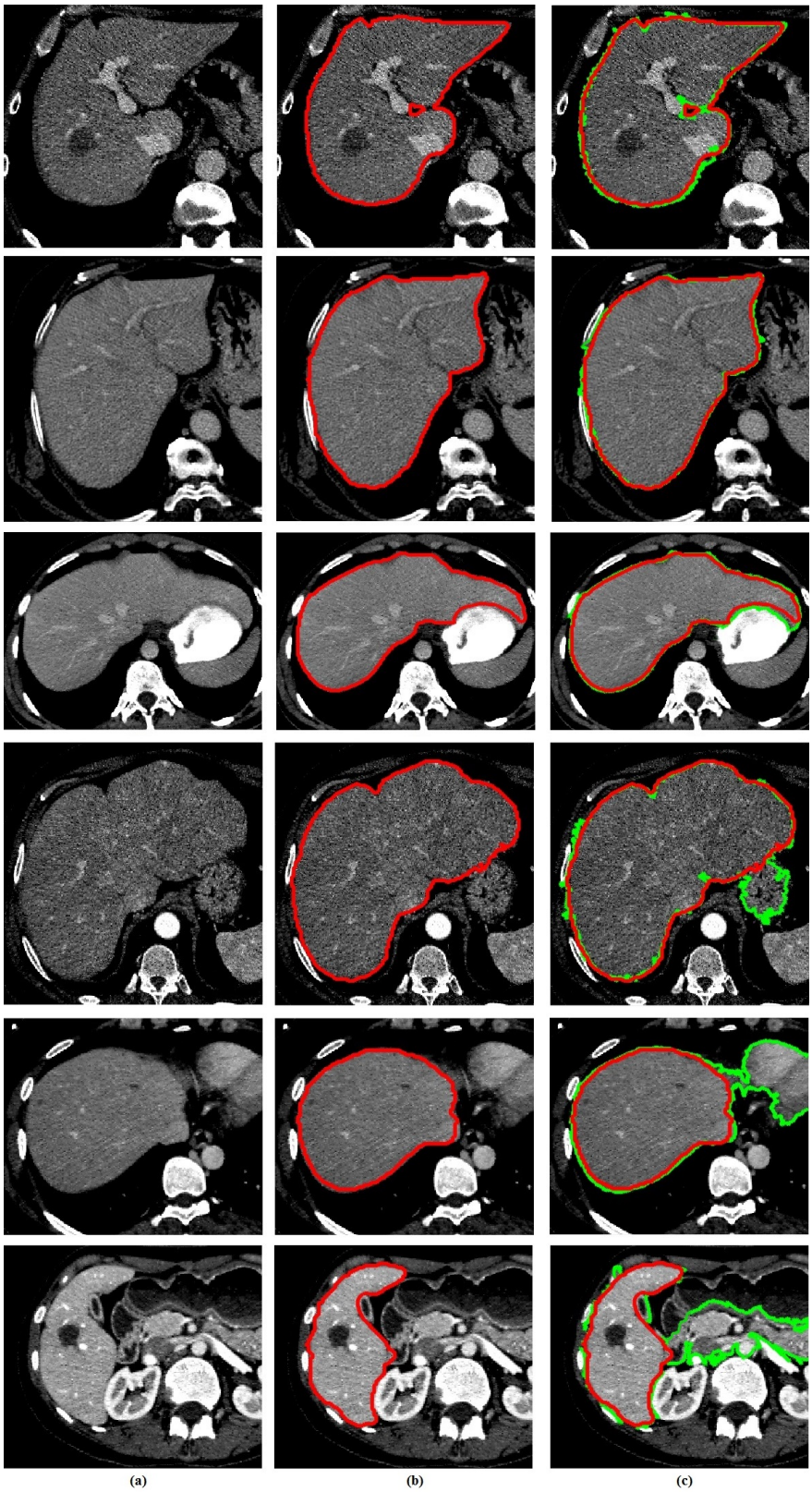


Figure 4.9: Samples of liver segmentation results (our method versus ground truth). (a) The original input CT image. (b) The ground truth of liver segmentation representing by the red line. (c) Liver segmentation result (green line) with ground truth (red line).

Figure 4.9 depicts the difference in segmentation by our automatic method compared with the ground truth obtained from the radiologist (included in the dataset). However, in some CT cases, the segmentation results are not quite satisfactory. The

increase in a false positive segmentation is due to some of the organs may surround the liver and have almost the same intensity value. In addition, the boundary between liver and contact organs may disappear and are difficult to discriminate it and this lead to incorrect segmentation by considering the other organs as a part of the liver.

By inspecting the segmentation results, we can find that some exams presented a high false positive due to the contact organs have similar intensity to the liver with no clear boundaries between of them. Hence, to overcome this problem, the segmented liver boundary was traced based on location and area across CT slices. In order to capture the significant increase in the liver area and irregular change in the boundaries over CT slices, due to overlap between the contact tissues and liver area. By comparing the liver boundary slice-by-slice, the unexpected added area will be removed, as depicted in Figure 4.10.

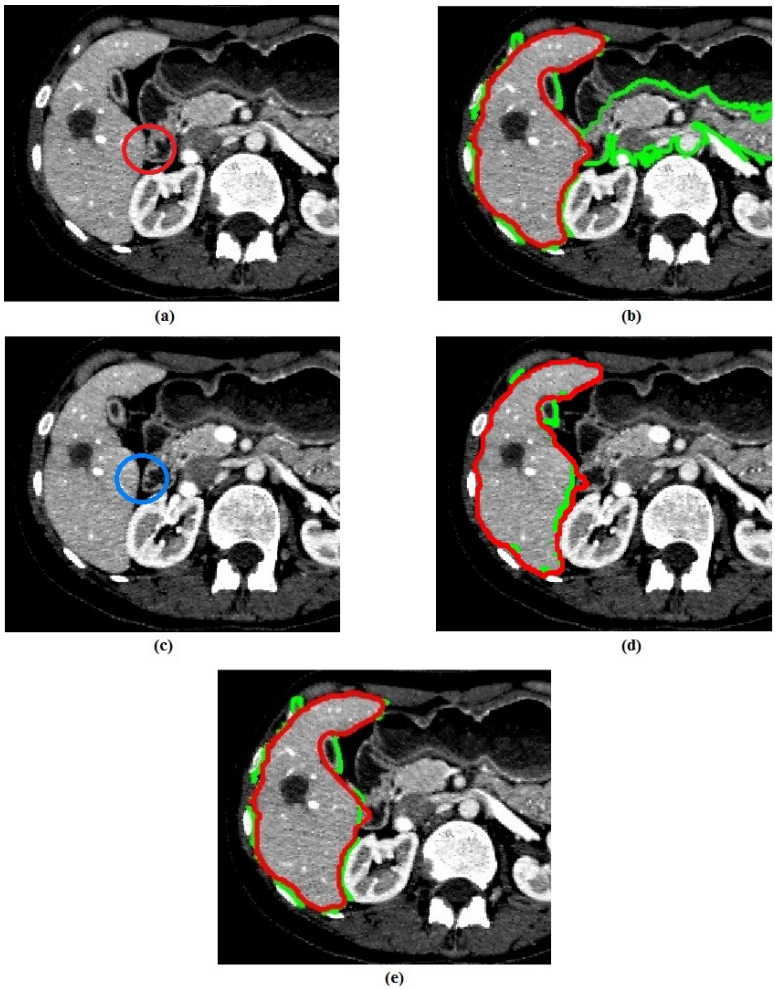


Figure 4.10: Refine segmented liver boundary. (a) The CT image in the first slice. (b) The segmentation result(green liner) compared to the ground truth (red line). (c) The CT image in the next slice. (d) The segmentation result(green liner) compared to the ground truth (red line). (e) Refine the segmentation for the first slice.

The segmentation results of the first slice, as depicted in Figure 4.10.b, shows a high false positive due to the surrounding non-liver tissues have short common boundaries with liver as shown in Figure 4.10.a (inside red circle). In the next slice, the

boundaries between liver and surrounding tissues disappeared, as shown in Figure 4.10.c (inside blue circle). By comparing the first slice to the next slice, non-liver areas will be discarded, as illustrated in Figure 4.10.e.

4.2 Liver Lesion Detection

The liver lesion segmentation is the second stage in the framework. The lesion is detected from the segmented liver, as presented in Section 4.1. The clustering method is widely used in lesion segmentation by dividing the input image into different clusters, based on changes in the intensity value. The different tissues in the medical images often comprise overlapping intensity on the gray level. Accordingly, the fuzzy clustering algorithm is particularly suitable for lesion detection from the medical image such as CT scan.

Fuzzy clustering is a popular unsupervised approach in medical image processing field, especially for liver lesion segmentation (Moghbel et al., 2016). This is because of its fuzzy nature, which allows pixels to belong to multiple clusters with varying degrees of membership. The advantages of Fuzzy clustering algorithm include a straightforward implementation and applicability to multichannel data make. Especially, the ability of Fuzzy clustering approach to model uncertainty within the data give a solution to deal with the fuzziness of the CT images. The CT images mainly suffer from inhomogeneity and uncertainty due to differences in noise and scanning geometries (e.g. starting angles of the helical scan). Furthermore, the substantial variations exist among different manufacturers when measuring CT images in phantoms, potentially owing to differences in x-ray spectra, CT detectors, and reconstruction kernels (Birnbaum et al., 2007; Fletcher et al., 2016). Because of our dataset has varied characteristics (e.g. resolution, spaces, etc.) and has produced from different manufacturers, the Fuzzy clustering approach has adopted in this work.

The fuzzy c-means (FCM) clustering algorithm assigns pixels to each group by using fuzzy memberships. Let $X = \{x_1, x_2, ..., x_n\}$ denotes an image with n pixels to be grouped into c clusters where c is a positive integer greater than one. The c clusters is a partition of X into mutually disjoint sets $\{x_1, x_2, ..., x_c\}$ where $X = x_1 \cup x_2 \cup ... \cup x_c$ or equivalently based on indicator function $\{\mu_1, \mu_2, ..., \mu_c\}$ where $\mu_i(x) = 1$ if $x \in X_i$ and $\mu_i(x) = 0$ if $x \notin X_i$ for all $i = 1, 2, ..., c$. The set $\{\mu_1, \mu_2, ..., \mu_c\}$ of indicator function is called a hard c-partition of clustering X into c classes (Wu and Yang, 2005). The following Equation 4.8 defined the hard c-means.

$$H(\mu, a) = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij} \|x_j - a_i\|^2 \tag{4.8}$$

Where $\{\mu_1, \mu_2, ..., \mu_c\}$ with $\mu_{ij} = \mu_i(x_j)$ is a fuzzy c-partition and $\{a_1, a_2, ..., a_c\}$ is the set of c cluster centres.

The hard c-means extended to allow $\mu_i(x)$ to be membership function of fuzzy sets μ_i on X assuming values in the interval $[0, 1]$, where $\sum_{i=1}^{c} \mu_i(x) = 1$ for all $x \in X$. This extension of $\{\mu_1, \mu_2, ..., \mu_c\}$ is called a fuzzy c-partition of X, as defined in Equation 4.9.

$$H_m(\mu, a) = \sum_{i=1}^{c} \sum_{j=1}^{n} \mu_{ij}^m \|x_j - a_i\|^2, m > 1 \tag{4.9}$$

where m (weighting exponent) is the degree of fuzziness, which could be used to enhance the clustering performance of FCM (Yu et al., 2004). The FCM clustering method minimises the objective function H_m in an iterative way, as depicted in Equation 4.10 and Equation 4.11.

$$\mu_{ij} = \frac{[1/\|x_j - a_i\|^2]^{1/(m-1)}}{\sum_{k=1}^{c} [1/\|x_j - a_k\|^2]^{1/(m-1)}} \tag{4.10}$$

$$a_i = \frac{\sum_{j=1}^{n} (\mu_{ij})^m x_j}{\sum_{j=1}^{n} (\mu_{ij})^m} \tag{4.11}$$

The fuzzy c-means (FCM) clustering method has been employed widely in lesion detection. However, the FCM is sensitive to noise and outliers. Therefore, the alternative fuzzy c-means (AFCM) is used for their ability to tolerate noise and outliers that often happen in medical images (Wu and Yang, 2002). The euclidean norm $E^2(x, a) = \|x - a\|^2$ used in FCM has been replaced in AFCM with new one, as presented in Equation 4.12.

$$E^2(x, x) = 1 - \exp(-\beta \|x - a\|^2) \tag{4.12}$$

The AFCM clustering method groups the similar pixels iteratively and the centres of the grouped pixels are adjusted for all iterations. The AFCM objective function measures the overall dissimilarity within the grouped pixels and the dissimilarity requires to be minimised to get the optimal partition. The objective function of AFCM is shown in Equation 4.13.

$$H_{AFCM} = \sum_{i=1}^{c} \sum_{j=1}^{n} (\mu_{ij})^m \left\{1 - \exp(-\beta \|x_j - a_i\|^2\right\} \tag{4.13}$$

Where β is a constant which can be defined in Equation 4.14.

$$\beta = \left(\frac{\sum_{j=1}^{n} \|x_j - \bar{x}\|^2}{n} \right)^{-1} \qquad with\ \bar{x} = \sum_{j=1}^{n} \frac{x_j}{n} \tag{4.14}$$

Through the objective function, the pixels are assigned to higher membership value when their intensities are close to the centroid of corresponding groups, or the lower membership values are being assigned to them when the intensities are far from the centroid. The membership value of the intensities of pixels and centre of the grouped pixels are solely dependent on the distance between each other to assign these pixels to a specific group. The Equation 4.15 and Equation 4.16 are been used to update the membership functions and the group centres.

$$\mu_{ij} = \frac{[1/(1 - \exp(-\beta\|x_j - a_i\|^2))]^{1/(m-1)}}{\sum_{k=1}^{c} [1/(1 - \exp(-\beta\|x_j - a_k\|^2))]^{1/(m-1)}} \tag{4.15}$$

and

$$a_i = \frac{\sum_{j=1}^{n} (\mu_{ij})^m \exp(-\beta\|x_j - a_i\|^2) x_j}{\sum_{j=1}^{n} (\mu_{ij})^m \exp(-\beta\|x_j - a_i\|^2)} \tag{4.16}$$

The lesion is extracted from the segmented liver using AFCM clustering method (Wu and Yang, 2002) to segregates each pixel into one of three defined classes which are liver, lesion and vessels. According to the pre-knowledge, the liver parenchyma appears less brighter intensity compared to the vessels where the lesions appear darker (Sahani and Kalva, 2004a; Oliveira et al., 2011a). The pixels in the lesion with low intensity are assigned to the first cluster, the liver pixels are assigned to the second cluster and third cluster contains the vessels pixels with the high-intensity value. The AFCM clustering algorithm assigns pixels for each class using fuzzy membership.

Figure 4.11 depicts the liver lesion segmentation process. The original CT image shown in Figure 4.11.a. The segmented liver is presented in Figure 4.11.b. Figure 4.11.c shows the image pixels clustering into one of three defined class by applying AFCM algorithm where Liver area defined as green colour, Lesion is red and vessels is blue. Figure 4.11.d. shows the final result of liver lesion segmentation.

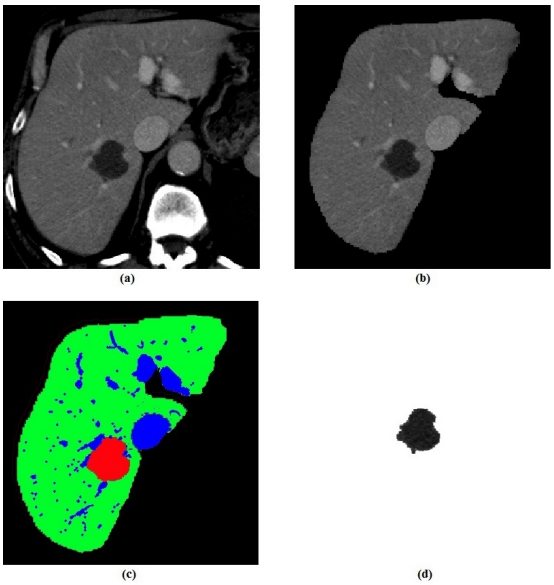


Figure 4.11: Steps of liver lesion segmentation process. (a) The original input CT image. (b) The segmented liver. (c) Operates AFCM to cluster the segmented liver into three areas where the green colour presents liver, lesion is red and vessels are blue . (d) The final result of the liver lesion segmentation.

Table 4.1 shows the liver lesion detection algorithm that analyses a segmented liver by using AFCM algorithm to classify image pixels into three classes which are liver, lesion and vessels.

ALGORITHM 1: Clustering the segmented liver into three classes (Liver, Lesion and Vessels)

INPUT:
Liver Image $X = \{x_1, x_2, ..., x_n\}$
SET number of cluster c=3;
SET degree of fuzziness m=2;
SET improvement value $\varepsilon = 0.0001$;
SET Random initial centre $a^o = \{a_1^{(o)}, ..., a_c^{(o)}\}$;

OPERATION:
Using Equation 4.14 to estimate parameter (β);
WHILE $\left\|a^{(k+1)} - a^{(k)}\right\| < \varepsilon$ **DO**
Using Equation 4.15 to calculate $\mu^{(k)}$ with $a^{(k)}$;
Using Equation 4.16 to calculate $a^{(k+1)}$ with $\mu^{(k)}$ and $a^{(k)}$;
IF $\left\|a^{(k+1)} - a^{(k)}\right\| > \varepsilon$ **THEN** k=k+1
END WHILE

OUTPUT $L = \{l_1, l_2, \ldots, l_n\}$, l_i is a label for point x_i;

Table 4.1: Our algorithm to cluster image pixels into three classes (liver, lesion and vessels) based on AFCM.

In the above algorithm, the number of cluster is three classes which are liver, lesion and vessels. The degree of fuzziness (m) selected the value two where the updated fuzzy membership value is proportional to the square of the inverse distance from a specific segment location to each cluster's centroid.

Figure 4.12 illustrates the graphical evaluation results of liver lesion segmentation based on Dice similarity coefficient (DSC) metrics. The vertical axis (x-axis) represents the number of cases that achieved the DSC values in the horizontal axis (y-axis). The highest DSC value means the better segmentation result.

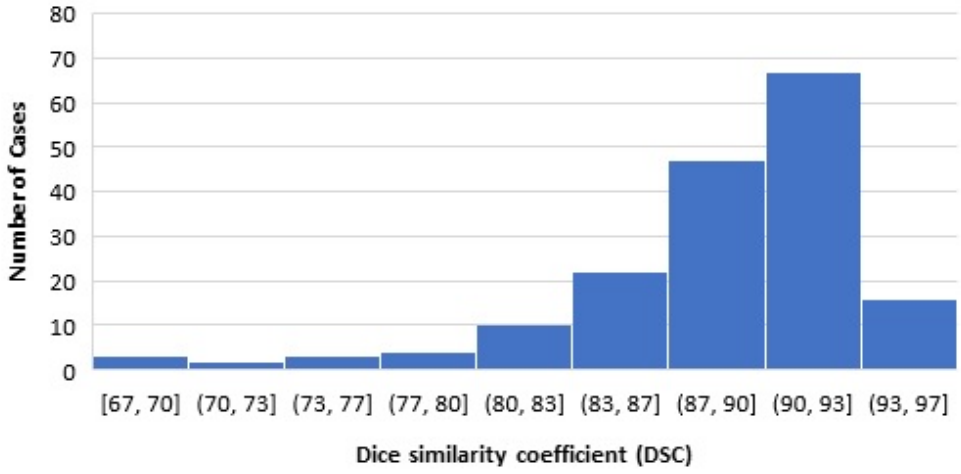


Figure 4.12: Number of cases and obtained DSC values of liver lesion segmentation for evaluating the method.

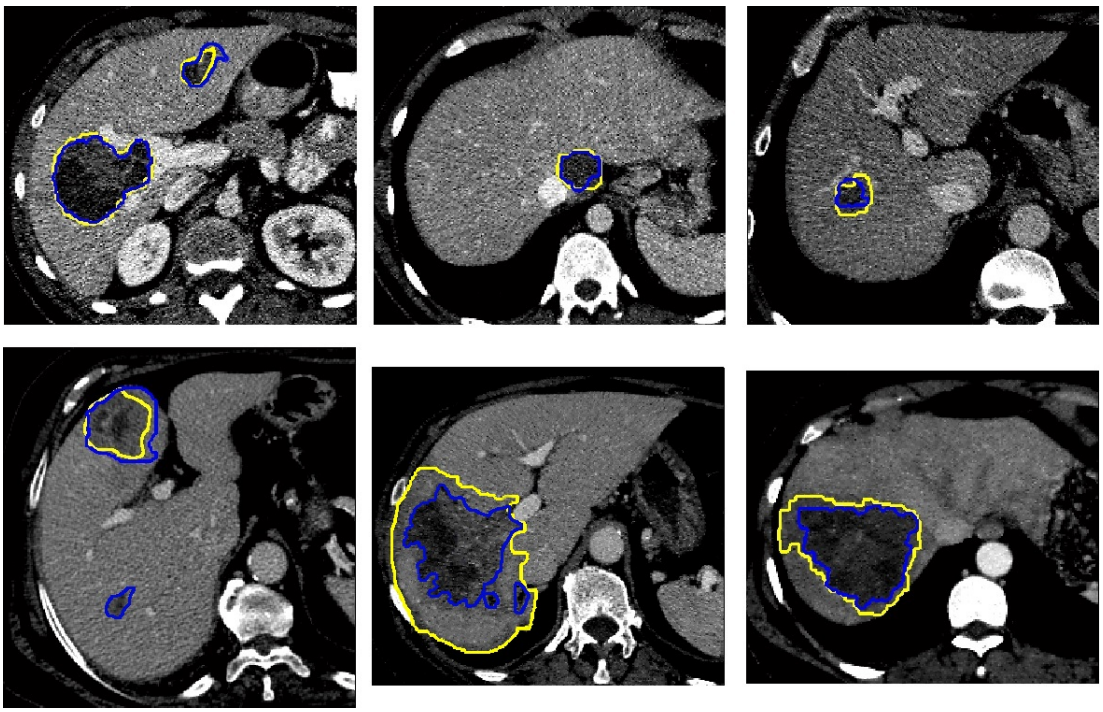


Figure 4.13: Samples of liver lesion detection results (our method versus ground truth). Our method is shown in blue and the reference detection in yellow.

Figure 4.13 illustrates the liver lesion detection accuracy of the method compared to the manual radiologist detection. The results of lesion detection shows over-segmentation and under-segmentation compared to the ground truth on some of cases especially with low intensity contrast between lesion and liver or with weak lesion boundary. This is due to the our detection method mainly relies on the gray level intensity difference between lesion and normal liver tissue.

Towards a better liver lesion segmentation, the post-processing step is applied to refine the lesion boundaries. In order to reduce the false negative/ false positive segmentation due to under/ over-segmentation of the lesion, the fast marching approach is used to enumerate the majority of the lesion voxels. Subsequently, level set approach refines lesion final shape by attaching the segmented contour to edges in the image while maintaining smoothness. The main advantages of these techniques are that arbitrarily complex shapes can be modeled and topological changes are handled implicitly. The idea behind this process is to refine the delineation of lesion boundaries that represents the abnormality area with respect to the healthy liver, as shown in Figure 4.14. Figure 4.14.a depicts the segmentation results using our method based on AFCM approach and before the refinement process. Figure 4.14. illustrates the results after the refinement process.

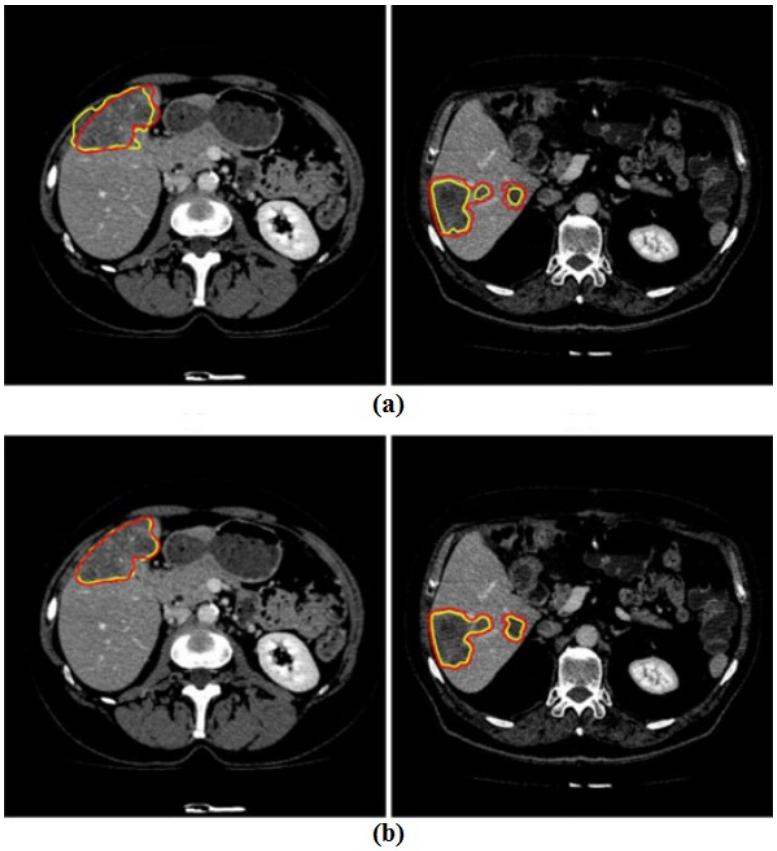


Figure 4.14: Liver lesion segmentation versus ground truth (GT) where GT is delineated as red contours. (a) The segmentation results before refinement process. (b) The segmentation results after refinement process.

4.3 Liver Vessels Extraction

The liver vessels extraction is important for anatomical liver segments which will be used later in lesion characterisation such as describe lesion location. In this step, the segmented liver is used to extract the vessels. In this work, the region growing algorithm based on the pre-processing Gaussian filter with the size 3 x 3 was adopted to extract the vessels, where several attempts were done to select the best parameter. where Gaussian is able to reduce the noise, improve the image contrast in homogenous areas and bright out the details of an image (Liu et al., 2012). The gray-level intensity distribution for the vessels is different from liver parenchyma intensity. Figure 4.15 shows the block diagram of the framework to extract liver vessels.

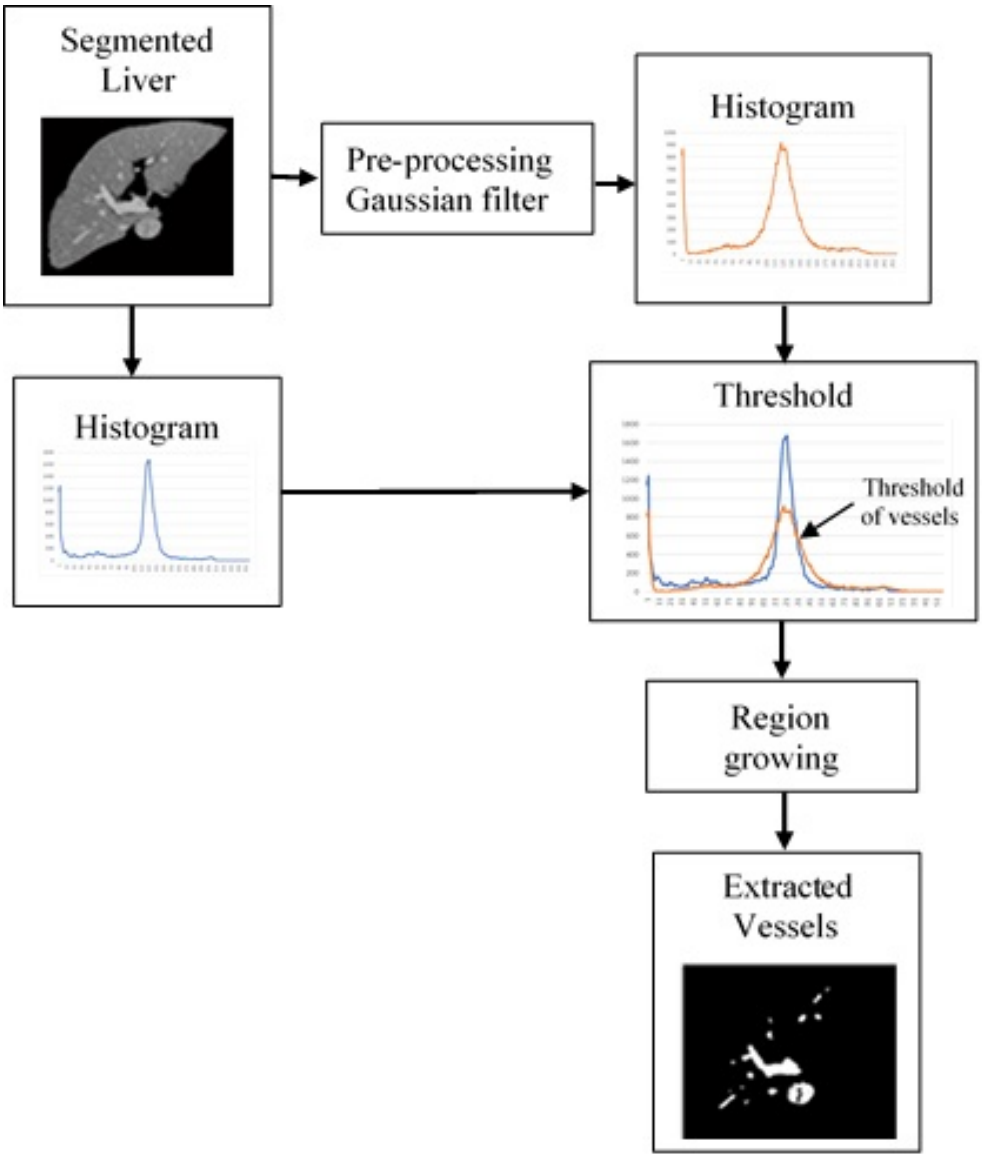


Figure 4.15: Block diagram of our framework for liver vessels extraction. The depicted output is the liver vessels from the CT image.

The main processing stages of the framework to extract liver vessels are as follows:

- At the first step, the histogram $L_i(x)$ of the segmented liver is calculated. Where

the histogram contains liver, lesion and vessels intensity $L = P + T + V$, where P corresponds to the liver parenchyma, T to the lesion and V to the vessels.

- Applied Gaussian filter $G_\sigma(x, y)$ on the segmented liver image, as presented in Equation 4.17:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{4.17}$$

Where x and y are the offset position from a discriminated pixel, σ is a scale parameter. The gradient magnitude of a Gaussian filter is presented in Equation 4.18 to enhance the boundaries.

$$\nabla G(x, y) = \sqrt{\left(\frac{\partial G(x, y)}{\partial x}\right)^2 + \left(\frac{\partial G(x, y)}{\partial y}\right)^2} \tag{4.18}$$

Where the gradient magnitude of a Gaussian filter $\nabla G(x, y)$ combines $\frac{\partial G(x,y)}{\partial x}$ (horizontal derivative) and $\frac{\partial G(x,y)}{\partial y}$ (vertical derivative) as shown in Equation 4.19 and Equation 4.20.

$$\frac{\partial G(x, y)}{\partial x} = -\frac{x}{2\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{4.19}$$

and

$$\frac{\partial G(x, y)}{\partial y} = -\frac{y}{2\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{4.20}$$

- Computed the histogram $L_G(v)$ of the new filtered image. The liver vessel is realised by intersecting both histograms resultant of: (1) directly from the CT image $L_i(x)$ and (2) after applying Gaussian filtering $L_G(v)$. After the peak point of the histogram, the brightness distribution of the image after applied Gaussian filter intersects with the input image. Thus, this point considered as vessels threshold (T_v).

- The region growing algorithm is employed to segment vessels. The intensity of the vessels inside the liver area is brighter than the intensity of liver parenchyma which have intensity distribution based on (T_v). The intensity value of the pixels greater than the threshold (T_v) are classified straightforwardly as a liver vessels and used them as seeds for the region growing algorithm that aggregates the neighbouring pixels with the intensity level value greater than threshold (T_v) to create binary mask for the vessels.

Figure 4.16 depicts the sample of liver vessels extraction process. The original input CT image shown in Figure 4.16.a. The segmented liver is depicted in Figure 4.16.b. Figure 4.16.c. shows enhanced liver after applied Gaussian filter. Figure 4.16.d shows the detected vessels where the vessels are blue. The 2D vessels mask shown in Figure 4.16.e.

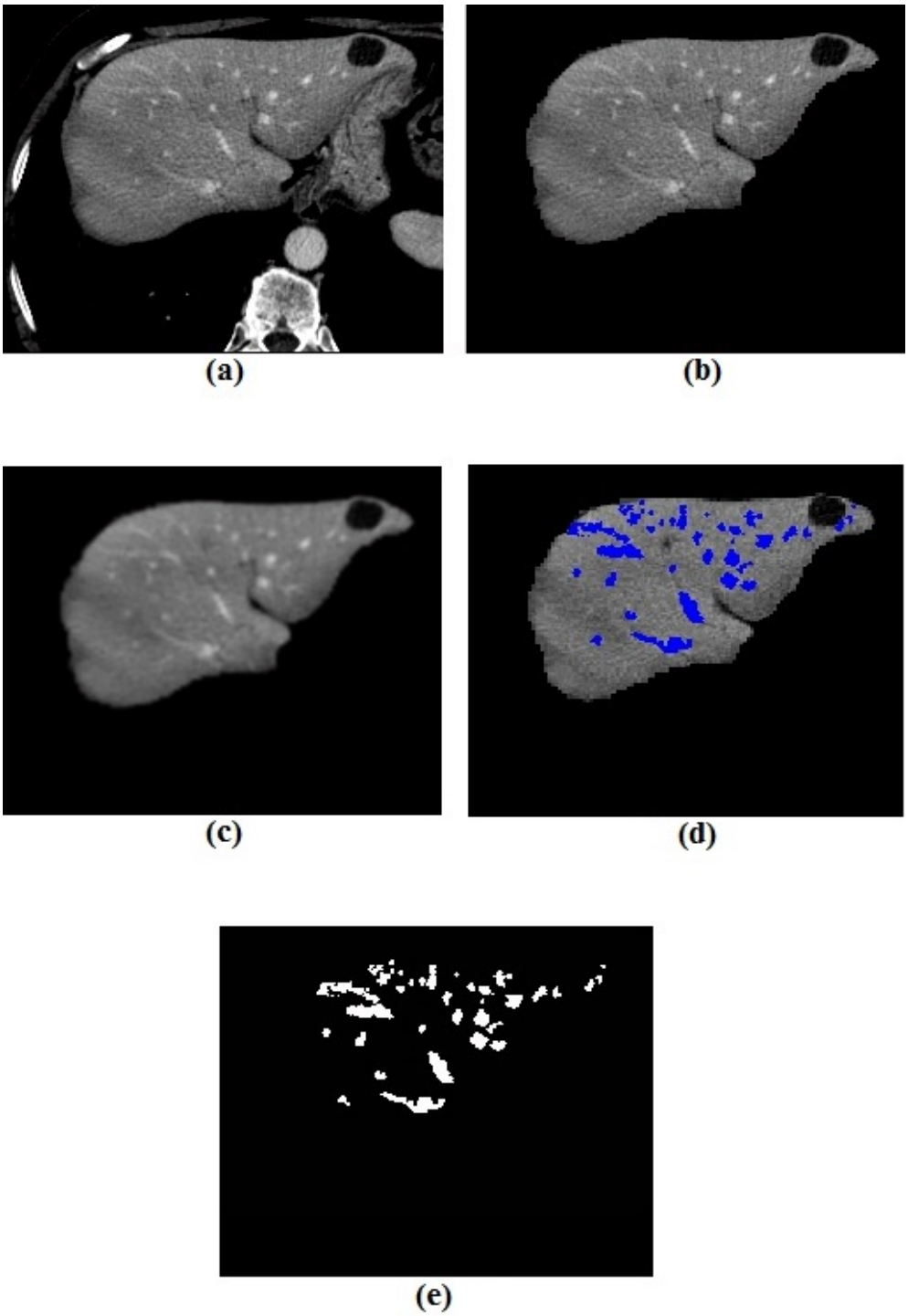


Figure 4.16: Sample of liver vessels extraction. (a The original input CT image. (b) The segmented liver. (c) Applies Gaussian filter. (d)Detected liver vessels (blue). (e) Liver vessels mask.

This part investigate the performance of our algorithm for liver vessels extraction based on Gaussian filter compared to AFCM method. In order to validate the liver vessels extraction results, the 30 CT scan of the liver were selected randomly. The calculated values of the OVE, RVD, TPVF, FPVF, JSM and DSC for our method based on 3x3 Gaussian filter is presented in Table 4.2. Table 4.3 presents the evaluation results of vessels extraction performance through AFCM clustering algorithm.

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
1	8.15	5.23	98.26	6.63	91.85	95.75
2	24.31	14.87	92.57	19.41	75.69	86.16
3	9.38	1.42	95.75	5.59	90.62	95.08
4	27.97	-10.28	79.43	11.46	72.03	83.74
5	10.29	-1.54	93.85	4.69	89.71	94.57
6	8.18	2.88	97.12	5.61	91.82	95.74
7	6.86	3.09	97.94	4.99	93.14	96.45
8	20.09	20.53	97.95	18.73	79.91	88.83
9	22.40	-8.07	83.86	8.78	77.60	87.38
10	20.63	7.75	91.93	14.68	79.37	88.50
11	16.33	13.56	97.29	14.33	83.67	91.11
12	21.45	8.34	91.66	15.40	78.55	87.99
13	9.76	-1.13	94.33	4.59	90.24	94.87
14	21.33	12.70	93.65	16.90	78.67	88.06
15	6.18	4.35	98.91	5.21	93.82	96.81
16	19.04	10.05	93.97	14.61	80.96	89.48
17	27.48	11.21	88.79	20.17	72.52	84.07
18	14.75	11.22	97.20	12.61	85.25	92.04
19	9.17	4.21	97.20	6.73	90.83	95.19
20	16.98	1.24	91.29	9.83	83.02	90.72
21	23.70	9.38	90.62	17.16	76.30	86.56
22	25.74	-9.38	81.23	10.36	74.26	85.23
23	10.91	5.06	96.61	8.04	89.09	94.23
24	15.01	9.44	96.22	12.07	84.99	91.89
25	11.79	-1.13	93.21	5.73	88.21	93.74
26	12.08	-1.16	93.03	5.88	87.92	93.57
27	15.46	1.54	92.32	9.07	84.54	91.62
28	19.09	10.07	93.96	14.64	80.91	89.45
29	19.58	7.51	92.49	13.97	80.42	89.15
30	12.89	4.70	95.30	8.98	87.11	93.11
Mean	**15.83**	**3.59**	**92.86**	**10.06**	**84.17**	**91.22**

Table 4.2: Quantitative results of vessels extraction after applying 3x3 Gaussian filter pre-processing step.

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
1	11.22	5.23	96.51	8.28	88.78	94.05
2	18.80	14.87	96.28	16.18	81.20	89.62
3	10.44	5.67	97.17	8.04	89.56	94.49
4	30.97	2.06	82.52	19.14	69.03	81.68
5	15.49	1.54	92.31	9.09	84.51	91.60
6	14.41	4.33	94.23	9.68	85.59	92.24
7	17.22	18.53	98.97	16.51	82.78	90.58
8	23.06	20.53	95.89	20.44	76.94	86.97
9	28.02	8.07	87.09	19.42	71.98	83.71
10	18.47	15.82	96.93	16.31	81.53	89.83
11	19.38	14.92	95.93	16.52	80.62	89.27
12	26.69	16.68	91.66	21.45	73.31	84.60
13	12.47	4.54	95.46	8.68	87.53	93.35
14	15.43	12.70	97.46	13.52	84.57	91.64
15	17.13	8.69	94.57	12.99	82.87	90.63
16	22.67	20.09	95.98	20.08	77.33	87.22
17	25.88	11.21	89.91	19.16	74.12	85.13
18	19.20	11.22	94.39	15.13	80.80	89.38
19	17.21	8.41	94.39	12.93	82.79	90.58
20	22.99	1.24	87.56	13.52	77.01	87.01
21	23.70	9.38	90.62	17.16	76.30	86.56
22	19.54	-9.38	84.99	6.21	80.46	89.17
23	27.04	5.06	86.50	17.66	72.96	84.37
24	11.87	9.44	98.11	10.35	88.13	93.69
25	19.13	3.40	90.94	12.05	80.87	89.42
26	17.16	10.45	95.35	13.67	82.84	90.62
27	18.39	12.28	95.39	15.04	81.61	89.88
28	25.90	30.22	97.99	24.76	74.10	85.12
29	24.51	15.01	92.49	19.58	75.49	86.03
30	17.51	8.22	94.13	13.03	82.49	90.41
Mean	**19.73**	**10.01**	**93.39**	**14.89**	**80.27**	**88.96**

Table 4.3: Quantitative results of vessels extraction using AFCM approach

The overall vessels extraction performance of the AFCM is reduced by $3.90 \pm 2.47\%$ and $2.26 \pm 1.13\%$ on the average of Jaccard Similarity Metric and Dice Similarity Coefficient respectively, compared to our algorithm based on the pre-processing Gaussian filter. The Figure 4.17 depicts the comparison of True positive Volume Fraction(TPVF) and False Positive Volume Fraction (FPVF) between our vessels extraction method and AFCM algorithm.

The results are depicted in Figure 4.17, and shows that the liver vessels extraction on our method based on medical knowledge background and Gaussian filter is slightly better by $1.53 \pm 0.67\%$ on average of True positive compared to the AFCM approach. In addition, the extracted area which wrongly indicates that as a part of liver vessels (False positive) of the AFCM approach is increased by $4.83 \pm 1.52\%$ on average, due to

the sensitivity of the AFCM to the noise and does not use any information about edge to locate vessels that will affect in the tracking procedure (Kirbas and Quek, 2004). Although the accuracy of the extraction of liver vessels is similar between the two methods, but the increase in the False Positive will affect the accuracy of the location of the lesion characterisation for the vessels.

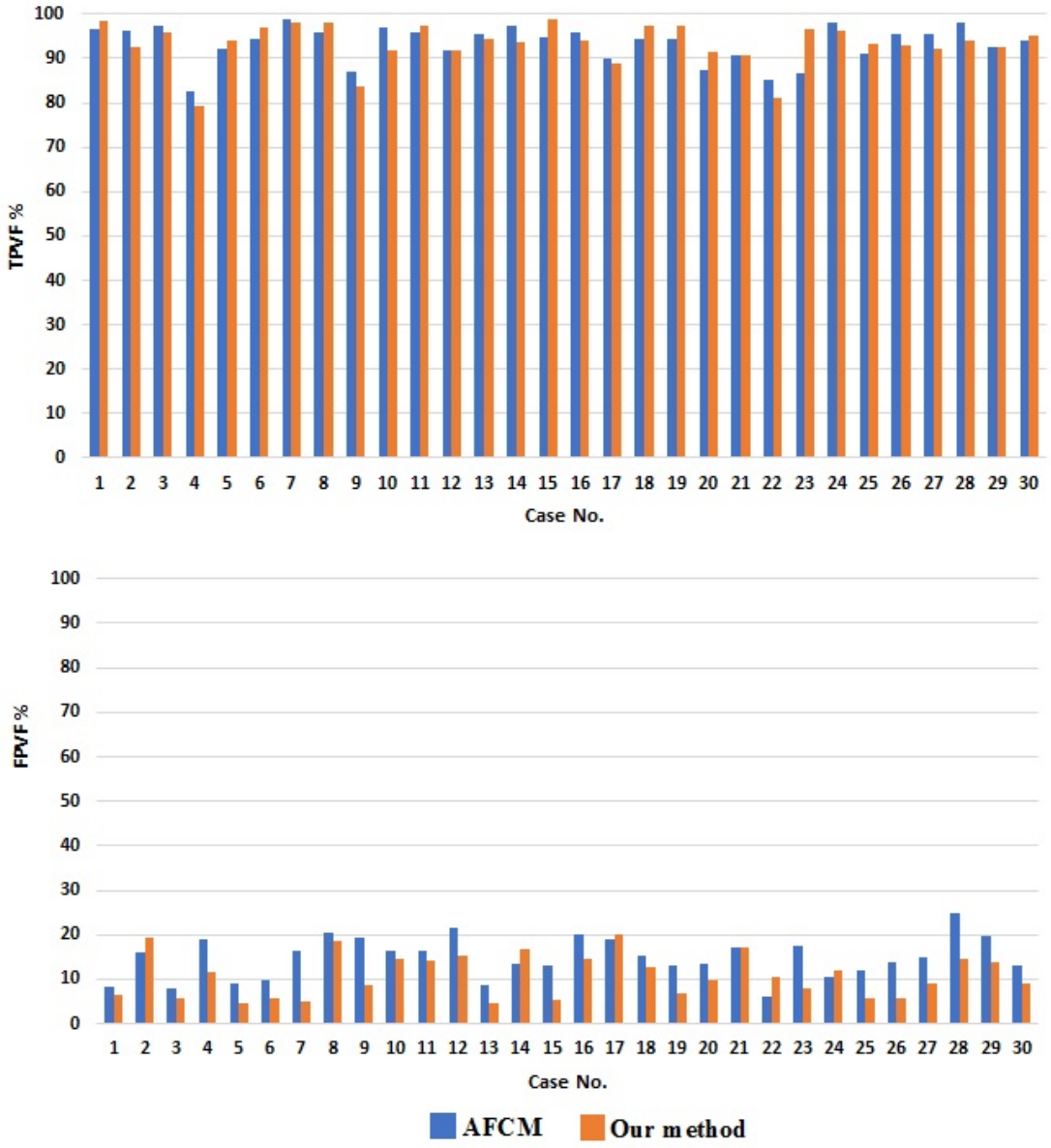


Figure 4.17: The liver vessels extraction performance based on TPVF and FPVF evaluation considering the Gaussian filter versus AFCM algorithm.

Figure 4.18 illustrates the graphical evaluation results of liver vessels segmentation based on Dice similarity coefficient (DSC) metrics. The vertical axis (x-axis) represents the number of cases that achieved the DSC values in the horizontal axis (y-axis). The highest DSC value means the better segmentation result.

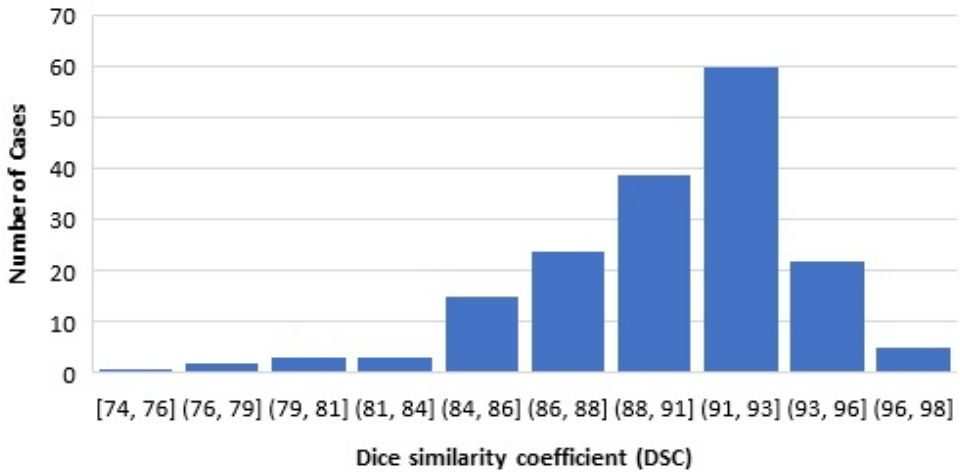


Figure 4.18: Number of cases and obtained DSC values of liver vessels segmentation for evaluating the method.

4.4 Results and Evaluation

This section presents an evaluation of liver segmentation, lesion detection and vessels extraction, coming from the dataset that was used in this thesis, as presented in details in Chapter 2. The performance of the system can be described through its accuracy and efficiency. These two factors are used to measure the success of the system to apply in clinical practice. For instance, the good accuracy of segmentation system is limited to practical use if it needs a long running time on an average case.

4.4.1 Evaluation of liver segmentation

In this thesis, an automatic liver segmentation was introduced by using anatomic medical knowledge, histogram-based adaptive threshold and morphological operations. Firstly, the HU intensity range of the liver based on medical pre-knowledge was applied to excludes air, fat and bones that having out range of liver intensity. The next step, the adaptive threshold of the liver intensity was estimated from the histogram of the image. Final step, the post-process (morphological operation) is applied to refine the liver segmentation result.

To evaluate accuracy of our liver segmentation method over the dataset compared to the ground truth; we utilised Jaccard Similarity Metric and Dice Similarity Coefficient methods which depicted in the Equation 4.21 and Equation 4.22. The average accuracy of liver segmentation was 82.85% and 90.45% respectively.

Figure 4.19 depicts a visual evaluation for automatic liver segmentation framework by using different measurement methods. Figure 4.19.a shows the evaluation results by using Volumetric Overlap Error. Figure 4.19.b illustrates the evaluation of the liver segmentation framework based on Relative Volume Difference. The evaluation results

by using True Positive Volume Fraction is shown in Figure 4.19.c. The Figure 4.19.d depicts the evaluation of over-segmentation for the framework through False Positive Volume Fraction. The full evaluation results are available in Appendix A.

The results are depicted in Figure 4.19, and shows that the best result of the liver segmentation compared to the ground truth by using JSM and DSC evaluation were 96.57% and 98.25% respectively, due to the quality of the original CT image and the sensitivity of the interpreter. On the other hand, the lowest JSM and DSC segmentation results were 59.70% and 74.77% respectively, because of the intensity similarity between liver and contact organs with the disappearance of the boundary between them. In addition, the poor quality of the CT image.

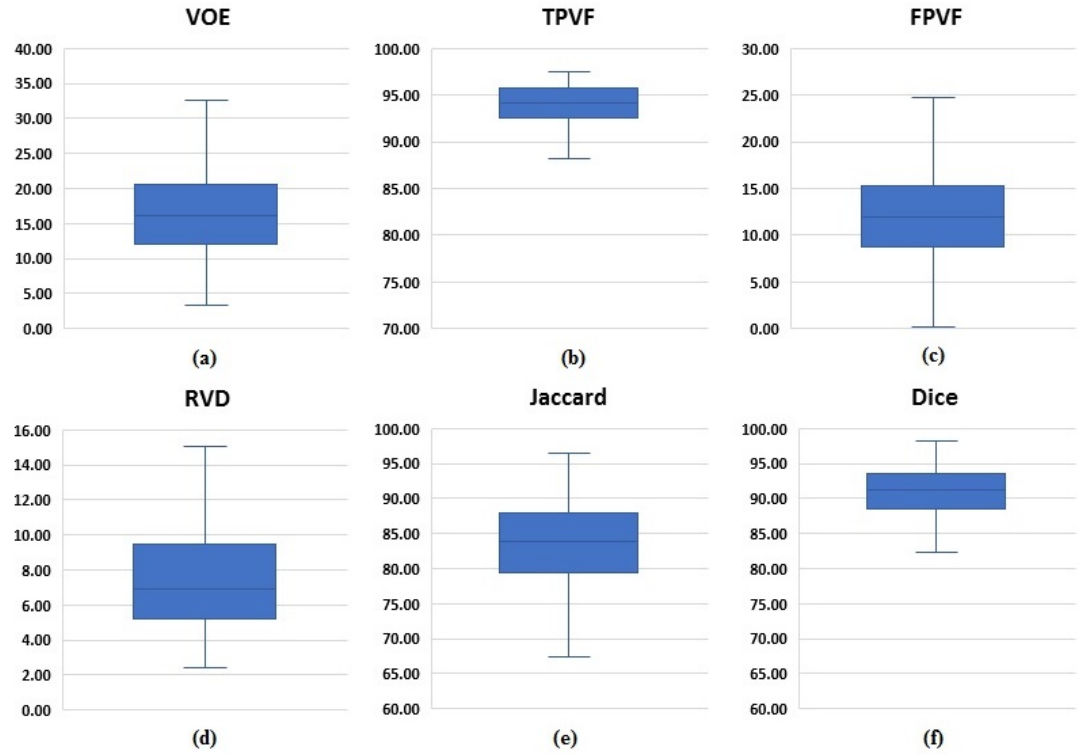


Figure 4.19: The overview of results for our liver segmentation by adopting different evaluation methods. (a) Evaluation by using VOE method. (b) TPVF method. (c) FPVF method. (d) RVD method. (e) Jaccard Similarity Metric method. (f) Dice Similarity Coefficient evaluation method.

4.4.2 Evaluation of liver lesion detection

The liver lesion is detected from the segmented liver through applying AFCM approach. The lesion segmentation process involves partitioning the segmented liver into three different clusters. Namely, liver area, lesion and vessels.

The average accuracy of liver lesion segmentation compared to the ground truth over the dataset by Jaccard Similarity Metric and Dice Similarity Coefficient evaluation methods was 79.76% and 88.51% respectively.

The Figure 4.20 depicts a visual evaluation for automatic liver lesion segmentation framework by using different measurement methods. Figure 4.20.a shows the evaluation results by using Volumetric Overlap Error. Figure 4.20.b illustrates the evaluation of our liver segmentation framework based on Relative Volume Difference. The evaluation results by using True Positive Volume Fraction is shown in Figure 4.20.c. The Figure 4.20.d depicts the evaluation of over-segmentation for our framework through False Positive Volume Fraction. The full evaluation results are available in Appendix A.

The evaluation results of the lesion detection are illustrated in Figure 4.20, and shows that the best accuracy of JSM and DSC were 91.24% and 95.42% respectively. However, the lowest lesion segmentation of JSM and DSC were 50.21% and 66.86% respectively. This is due to the AFCM lesion detection approach mainly relies on the gray level intensity difference between the clusters. Moreover, the lesion detection is affected by the CT image quality and the imaging time. However, most lesion segmentation differences between our method and ground truth occur at the boundary especially with the malignant lesion, because the characteristics of the malignant lesion are an invasion of adjacent structures with ill-defined/ poor margins (Murakami and Tsurusaki, 2014a; Minami and Kudo, 2015).

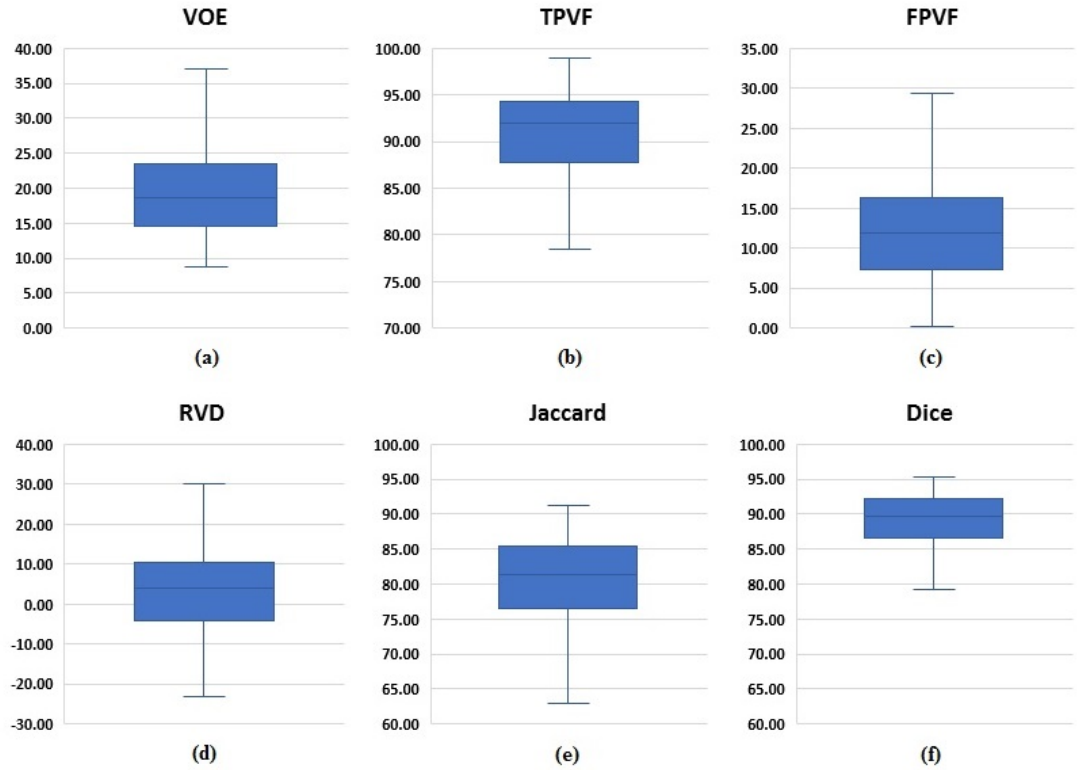


Figure 4.20: The overview of results for our liver lesion segmentation by adopting different evaluation methods. (a) Evaluation by using VOE method. (b) TPVF method. (c) FPVF method. (d) RVD method. (e) Jaccard Similarity Metric method. (f) Dice Similarity Coefficient evaluation method.

4.4.3 Evaluation of liver vessels extraction

The vessels in the segmented liver is extracted by a combination of pre-processing Gaussian filter and region growing algorithm. The adaptive threshold is obtained by intersecting the histogram of the liver before applying pre-processing process and the histogram after applying the Gaussian filter. Finally, the region growing algorithm is applied to extract the vessels. The accuracy of liver vessels extraction over the dataset by Jaccard Similarity Metric and Dice Similarity Coefficient evaluation methods was 82.06% and 90.03% respectively.

Figure 4.21 depicts a visual evaluation for automatic liver vessels segmentation our framework by using different measurement methods. Figure 4.21.a shows the evaluation results by using Volumetric Overlap Error. Figure 4.21.b illustrates the evaluation of our liver segmentation framework based on Relative Volume Difference. The evaluation results by using True Positive Volume Fraction is shown in Figure 4.21.c. The Figure 4.21.d depicts the evaluation of over-segmentation for our framework through False Positive Volume Fraction. The full evaluation results are available in Appendix A.

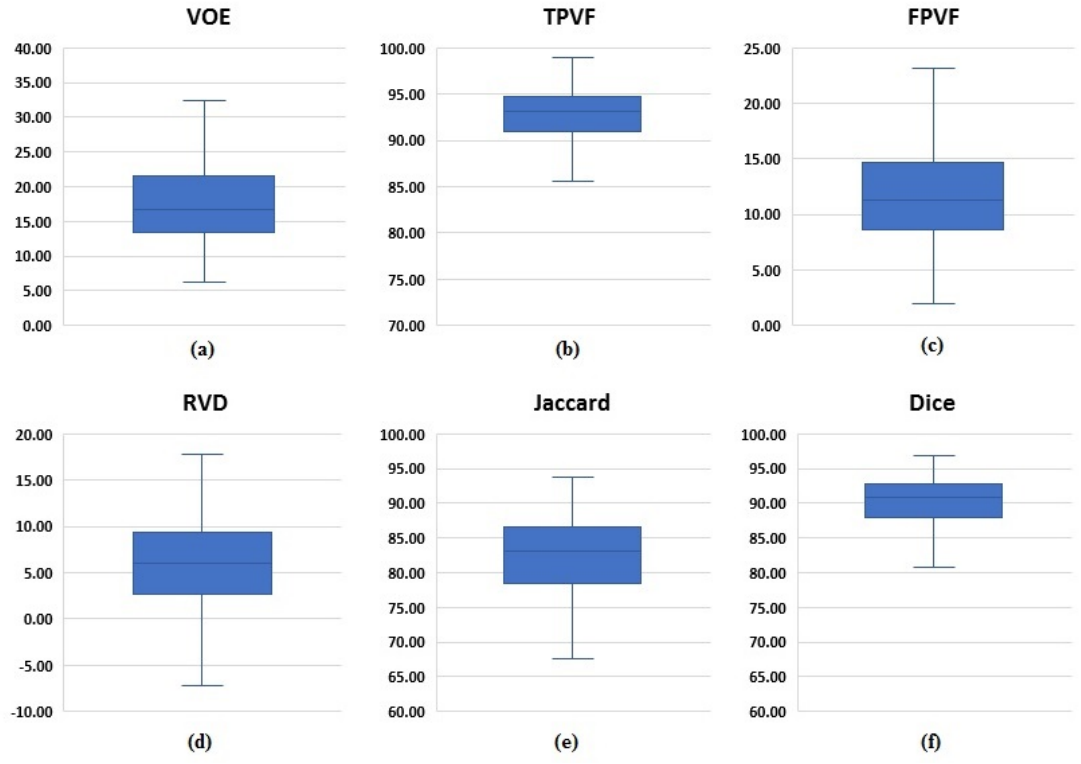


Figure 4.21: The overview of results for our liver vessels segmentation by adopting different evaluation methods. (a) Evaluation by using VOE method. (b) TPVF method. (c) FPVF method. (d) RVD method. (e) JSM method. (f) DSC evaluation method.

The evaluation results of liver vessels extraction are depicted in Figure 4.21, and shows that the best accuracy result of JSM and DSC were 93.82% and 96.81% respectively. On the other hand, the lowest liver vessels extraction of JSM and DSC were

58.76% and 74.03% respectively. This is due to the several factors that affected on the extraction performance: (1) CT image quality and noise ratio in the image. (2) Injection duration for determining CT scan timing, due to it affects the time to peak contrast enhancement in the vessels. (3) The type and size of the lesion as the large/ malignant lesions affect the structure of the vessels.

4.4.4 Comparison with Other Methods

It is always difficult to directly compare the segmentation performance with the state-of-art methods due to different datasets for evaluation, different qualities of manual segmentation, and differences in the evaluation metrics used. Hence, to overcome these challenges, different performance evaluation metrics with currently available public datasets are used for testing the performance of our framework, as a comparison would be more valuable if done using the same standard dataset with same validation criterion.

For evaluating the segmentation quality, the MICCAI criterion (Heimann et al., 2009) is used in this section, including the following measures:

Average Symmetric Surface Distance (ASD) [mm]

The ASSD is the average of all the distance from points on the framework segmentation boundary ($B_S = \{S_1, S_2, ..., S_N\}$) to the ground truth boundary ($B_{GT} = \{GT_1, GT_2, ..., GT_M\}$), as depicted in Equation 4.21.

$$ASD = \frac{\sum_{i=1}^{n} \min_{1 \le j \le m} \|S_i - GT_j\| + \sum_{j=1}^{m} \min_{1 \le i \le n} \|GT_j - S_i\|}{|B_S| + |B_{GT}|} \tag{4.21}$$

Where $\|S_i - GT_j\|$ denotes the Euclidean distance of the pixels S_i and GT_j in millimetre.

Root Mean Square Symmetric Surface Distance (RMSD) [mm]

The RMSD is highly correlated with ASD, but it is more sensitive to large deviations from the ground truth region, as defined in Equation 4.22.

$$RMSD = \sqrt{\frac{\sum_{i=1}^{n} \min_{1 \le j \le m} \|S_i - GT_j\|^2 + \sum_{j=1}^{m} \min_{1 \le i \le n} \|GT_j - S_i\|^2}{|B_S| + |B_{GT}|}} \tag{4.22}$$

Maximum symmetric surface distance (MSD) [mm]

MSD, also known as the Symmetric Hausdorff Distance, takes the maximum distance instead of the average. it is sensitive to outliers and returns the true maximum error, as depicted in Equaion 4.23.

$$MSD = \max\left\{\max_{1\leq j\leq m}\|S_i - GT_j\|, \max_{1\leq i\leq n}\|GT_j - S_i\|\right\} \tag{4.23}$$

The three surface errors ASD, RMSD and MSD are given in millimetres and equal to zero if it is a 100% accurate segmentation.

For a fair comparison, the public dataset (SLIVER07) was used to evaluate our liver segmentation framework. SLIVER07 is a liver CT dataset provided by the workshop on 3D segmentation in clinic that was held in conjunction with MICCAI 2007 conference (Heimann et al., 2007c). The dataset includes 20 CT volumes with standard liver segmentation. The CT images were acquired using different CT scanners. All the volumes have an in-plane resolution of 512×512 pixels and slice number varies from 64 to 502 (average 214). The inner-slice pixel spacing varies from 0.54 to 0.87 mm (average 0.7), and the slice thickness varies from 0.7 to 5 mm (average 1.6).

Method	VOE %	RVD %	ASD (mm)	RMSD (mm)	MSD (mm)	Time (min)
(Kainmüller et al., 2007)	**6.1±2.1**	-2.9±2.9	**0.9±0.3**	1.9±0.8	**18.7±8.5**	15
(Heimann et al., 2007b)	7.7±1.9	**1.7±3.2**	1.4±0.4	3.2±1.3	30.1±10.2	7
(Chen et al., 2012)	6.5±1.8	-2.1±2.3	1±0.4	**1.8±1**	20.5±9.3	6
(Lu et al., 2014)	7.4±1.9	4.6±2.8	1.2±0.4	2.8±1.3	38.5±18	12.4
(Yang et al., 2014)	8.9±2.2	2.3±2	1.4±0.3	2.4±1.2	24.3±9.6	2.1
(Li et al., 2015)	6.24±1.52	1.8±2.76	1.03±0.31	2.11±0.95	18.82±8.82	5.8
(Wu et al., 2016)	7.6	4.2	1	2	18.5	1.3
Our method	11.57±2.8	4.43±3.2	1.8±0.9	2.89±2.3	28.34±8.4	**0.7**

Table 4.4: Quantitative comparative results for liver segmentation over the SLIVER07 dataset.

On the other hand, the performance of our liver lesion segmentation framework is evaluated using 3Dircadb dataset from Research Institute against Digestive Cancer (IRCAD) (France, 2016), which is publicly available. This dataset is composed of 3D CT volumes of 10 men and 10 women with hepatic tumours in 75% of cases. The CT volumes correspond to 20 different patients. The expert radiologists have manually out-lined liver tumor contours for all CT volumes to determine the ground truth. All the CT images have an in-plane resolution of 512×512 pixels and slice number varies from 74 to 260 (average 141). The inner-slice pixel spacing varies from 0.56 to 0.87 mm (average 0.72), and the slice thickness varies from 1 to 4 mm (average 1.78).

Method	VOE %	RVD %	ASD (mm)	RMSD (mm)	MSD (mm)	Time (min)
(Moghbel et al., 2016)	**22.78+12.15**	8.59+18.78	NA	NA	NA	16
(Wu et al., 2017)	30.61+10.44	15.97+12.04	4.18+9.6	5.09+10.71	12.55+17.07	3.2
(Foruzan and Chen, 2016)	29.04+8.16	**2.2+15.88**	**0.72+0.33**	**1.1+0.49**	**4.25+3.03**	1.5
Our method	25.14+9.34	11.69+7.52	5.23+4.17	6.11+5.42	14.02+7.81	**0.8**

Table 4.5: Quantitative comparative results for liver lesion segmentation over the 3Dircadb dataset.

Regarding the comparison of liver vessel extraction, where there is no public database available for liver vessels segmentation based on CT scans, the quantitative comparison between the different methods is difficult. This is because the comparison would be more valuable if done using the same dataset with the same validation standard criterion. Furthermore, our method is designed to extract the major vessels that meet our target in characterising the location of the lesion for these vessels, not to extract the whole vascular network.

The liver and lesion segmentation performance achieved on SLIVER07 and 3Dircadb dataset is presented in Table 4.4 and 4.5 respectively. For comparison, we also present the segmentation results of the other methods. Results for each measure represent as the mean and standard deviations of the overall datasets. As can be seen in Table 4.4 and Table 4.5, our method achieves good segmentation results for the SLIVER07 and 3Dircadb dataset in terms of liver and lesion segmentation respectively. For all the datasets, our method outperforms the other algorithms in terms of the computational cost. The average run time of our method for liver or lesion segmentation is less than one minute. In addition, the our framework is a fully automatic that does not require any user interaction. However, the main drawback of the approach that obtained the best results in Table 4.4 is that it needed more than 100 liver shapes for the training process and besides that a semi-automatic (and manual) iteration was required in the training step. In general, a high computational cost and a hard training or initialisation was required in these methods

Among all the methods in Table 4.5, (Moghbel et al., 2016) proposed an automatic hybrid method based on FCM with cuckoo optimisation and random walkers. It achieved the highest accuracy and the average time was 16 minutes per case. (Foruzan and Chen, 2016) obtained initial tumor contours using supervised watershed, SVM, and scattered data approximation for large/small tumors and then refined the contours based on sigmoid edge model.Several seed points for the tumor, liver, and other background were marked in the middle tumor slice through user interaction. (Wu et al., 2017) proposed a semi-automatic method for liver tumor segmentation in CT volumes based on improved fuzzy C-means (FCM) and graph cuts. The performance of our method on the 3Dircadb dataset was comparable to the other methods, and the lesion segmentation procedure was performed automatically in a fast way with low computational cost. However, The segmentation errors were mainly due to the low contrast or weak boundaries of lesions. In addition, the lesions with inhomogeneous intensity

were also prone to be under-segmented.

From Table 4.4 and Table 4.5, we obtain that the liver and lesion segmentation performance of our method in terms of the five metrics are comparable to the best results. The average running time achieved by our method is shorter than that of other methods and can meet the clinical requirements. Thus, we can conclude that the our method is a good automatic liver and lesion segmentation method in terms of the segmentation accuracy and time efficiency.

4.4.5 Computational time

On the other side, The computational time for liver segmentation, lesion detection and vessels extraction is very important in the computer aided systems. Since the manual liver segmentation is time consuming. The semi-automatic and fully automatic systems have been proposed to solve the time consuming in efficient way.

For our framework, the computational time was computed to measuring the frame work performance. The Figure 4.22 shows the average computational time in milliseconds per CT slice for our system to segment the liver, detect lesion and extract vessels. The experiments were carried out using a single core of an Intel i5-3570 processor at 3.4 GHz with 8 GB of RAM.

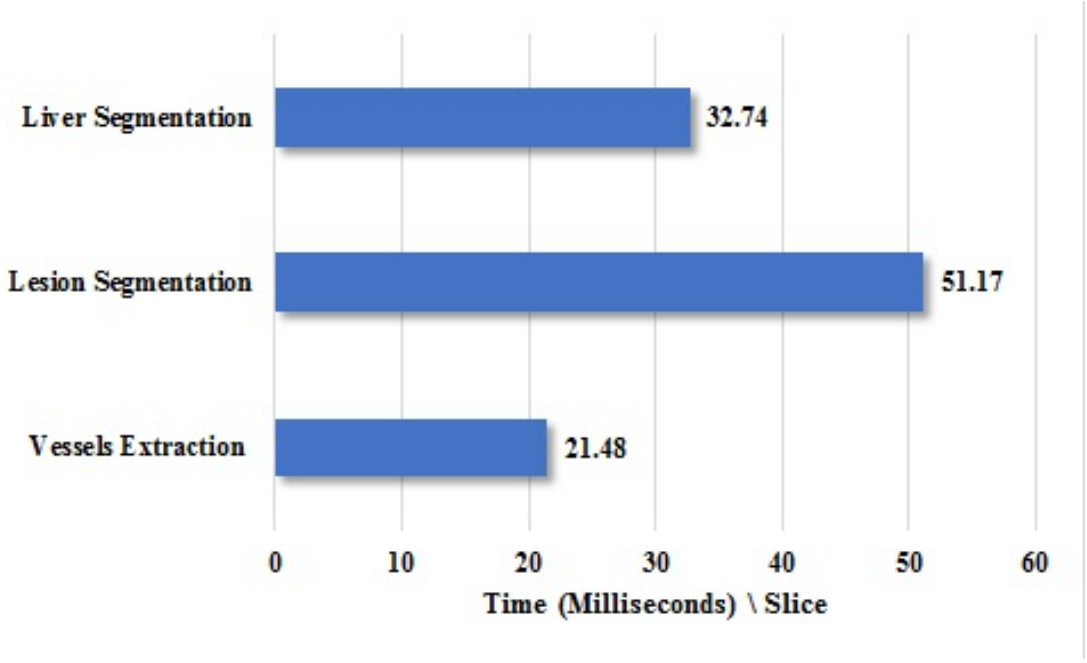


Figure 4.22: The overview of computational time for our framework in three different stages which are: (1) Liver segmentation. (2) Lesion segmentation. (3) Vessels extraction. The reported values are the time averages in milliseconds per slice, calculated over the our datasets.

4.5 Conclusion

This chapter presented a fully automatic computer aided detection framework for liver segmentation, detection the lesion inside the liver and extraction the liver vessels. This was done to utilise its potential usage benefits as a base for the proposed liver lesion characterisation and classification framework (introduced in the next chapter). To achieve these tasks, the pre-processing was used to enhance the CT image quality. The pre-knowledge such as liver location and gray-level intensity analysis was employed to extract liver area from the other organs. Finally, the post-processing was applied to refine liver segmentation. In the second stage, the liver lesion was extracted by applying the AFCM clustering algorithm. In the last stage, the pre-processing was applied on the segmented liver, and histogram analysis was performed to detect the optimal threshold to extract the liver vessels. The region growing technique was adopted to extract vessels of the liver. Several evaluation measurement methods (VOE, RVD, TPVF, FPVF, JSM and DSC) were used to assess our framework including liver segmentation, lesion detection and vessels extraction. All the results presented in this chapter were conducted using our dataset, which was already presented in Chapter 2. According to the evaluation results, the our system for liver segmentation, lesion detection and vessels extraction is showed a good results with a good performance based on computational time. Conclusively, these results highlight the possibility of utilising the automatic liver segmentation, lesion detection and vessels extraction to contribute in our target which is liver lesion characterisation and classification. The next chapter will present the structure of the proposed framework for liver lesion characterisation and classification.

Chapter 5

Liver Lesion Characterisation and Classification

This chapter presents the overall structure of the proposed framework for liver lesion classification and characterisation. The aim of this chapter is to present a novel framework and methodology for the two aspects, which are liver lesion classification and characterisation. The framework theoretical and concept basics, building constructs and its contributions to the thesis objectives are also discussed. The overview of this chapter is as follows: Section 5.1 presents the overall of the proposed framework. The liver lesion classification framework based on low-level feature is introduced in Section 5.2. The liver lesion characterisation framework is introduced in Section 5.3. The lesion classification framework based on high-level features and the fusion between high-level and low-level features are presented in Section 5.4 and 5.5 respectively. The feature selection and classifiers approaches used in this thesis, will be presented in Section 5.6 and Section 5.7 respectively. Finally the chapter is concluded in Section 5.8.

5.1 The Proposed Framework

The main objective of this research study is to develop a computer aided liver tumour characterisation/classification based on CT images. Particularly, CT scan is widely used for their high sensitivity and specificity to diagnose cancers and the ability to imaging the entire body (Sahani and Kalva, 2004b; Fass, 2008). The CT-images are considered as a rich medium of information, such information (features) which is highly beneficial for liver characterisation related tasks. The liver lesion classification and transformation the low-level features to the higher semantic level, would also be of great benefits:

- Assisting radiologists to improve and facilitate their diagnosis.
- Medical education used in training non expert radiologist.

- Medical image retrieval through searching similar cases not only by image, but also by diagnosis.

Figure 5.1 depicts block diagram of our proposed framework to diagnose the liver lesions, which has been used as a base for the investigations carried out so far.

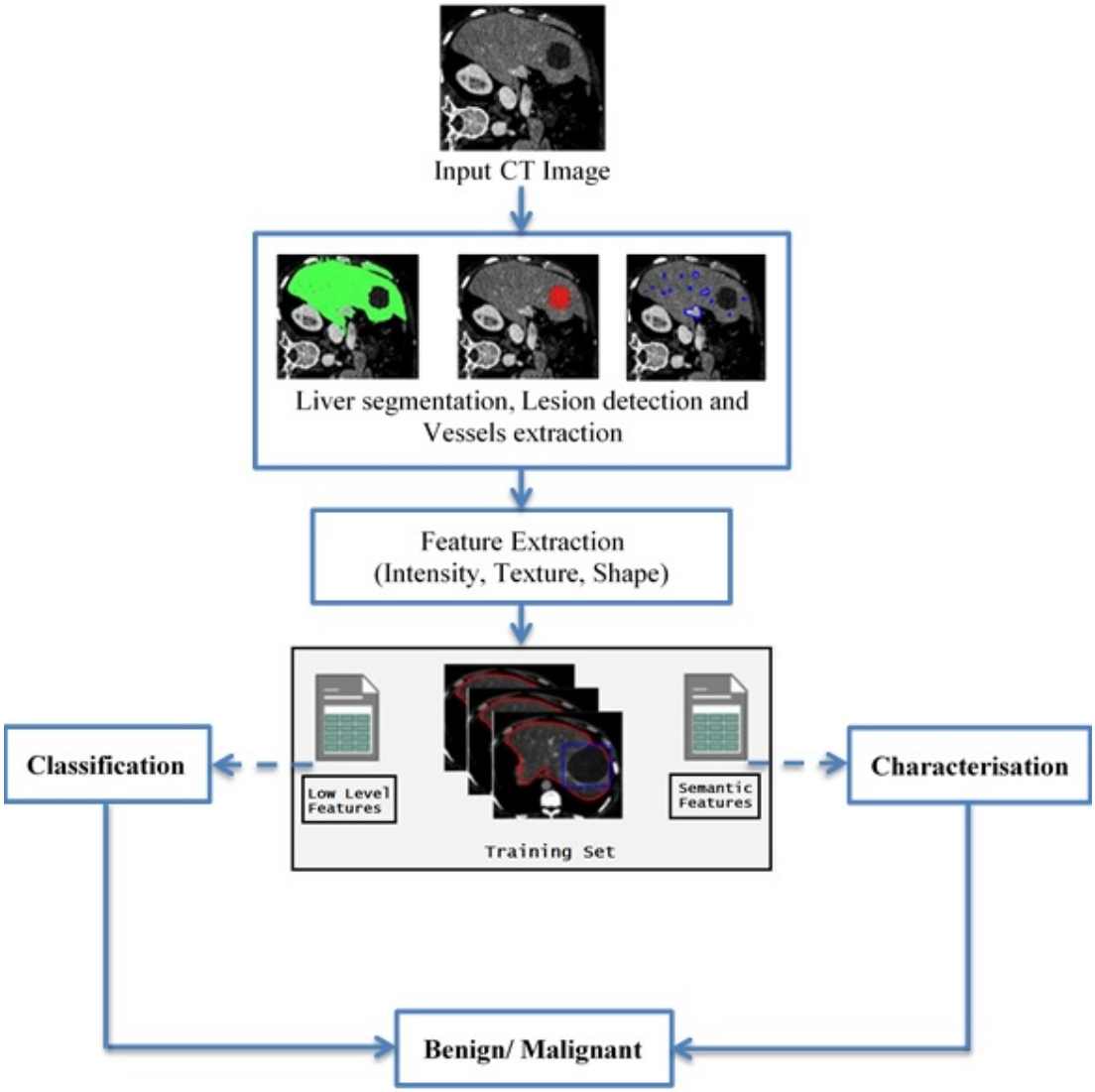


Figure 5.1: Overview block diagram of the proposed framework for liver lesion classification through low-level features and high-level features (characterisation).

The main stages of the framework system are:

- Liver segmentation, lesions detection and vessels extraction from CT image, as discussed in Chapter 4.
- Extracting features from Liver and lesions.
- Characterising lesions in high-level features (semantic features) by exploit the relationships between low-level features and high-level features.

- Classifying lesion into Benign (non-cancerous), Malignant (cancerous) in two ways, based on low-level features and through characterisation (high-level features)

The majority of the researches are being conducted to develop CAD system, which mainly focuses on improving liver segmentation, lesion detection and feature extraction to diagnose liver lesion. However, our proposed framework not only concerns on the lesion but also on the relation between the lesion and the liver (surrounding area) to extract and represent the features of the lesion, due to the characteristic differences between benign and malignant liver lesions in terms of their effect on surrounding liver tissue. Therefore, the lesion and surrounding liver tissue are used for better CAD system performance.

The framework provides two different modes for the liver lesion diagnosis. The first way is the classification-mode that uses the low-level features from the lesion and surrounding area to generate new feature vectors towards a more accurate diagnosis decision. The second way is the characterisation-mode that exploit the relationships between low-level features and high-level features to predict radiological observation in describing the liver lesions. Then, the high-level features utilise in the lesion diagnosis decision to enhance the liver lesion classification performance, by providing radiological observations / prior-knowledge through liver lesion characterisation. The high-level features and characterisation helps in interpreting and explaining the classification and is more intuitive to clinicians.

In this Chapter, we introduced two concepts. Namely, low-level feature and high-level feature. These concepts are defined as follows:

- The low-level feature is defined as a quantitative/numerical feature derived directly from raw pixel (hand-designed features).
- The high-level feature is defined as the terms (numerical, categorical, nominal, etc.) that describe the liver lesion in a way interpreted by humans.

The high-level features are based on the ontology of liver that is used by radiology, making it intuitive and understandable to radiologists, because these concepts are what radiologists used to characterise liver lesions.

5.2 Liver Lesion Classification

This section presents and discusses our proposed framework to classify liver lesions into one of the two classes: Benign or Malignant based on low-level features. Recent research efforts have been contributing considerably to enhance the performance of lesion classification (Kumar et al., 2012). According to literature, most of the literature work was focusing on feature extraction and classifiers for well-performing system

(Lin et al., 2016), and usually using the absolute value of features that extracted from lesion area, or using image patches by dividing lesion area to several regions for building dictionary BoVW based on a visual histogram. As a consequence, the performance is varied significantly under ROIs selection and different acquisition conditions such as different operators and settings. Hence, as the literature lacks this part (to the best of our knowledge), the ROI selection methods introduced as a new research area, which is importantly responsible for representing the lesion characteristics.

The next sections will present the core of the proposed framework to classify liver lesion based on low-level features. The low-level features that used in classification task are presented in Section 5.2.1. To address the challenge of the relation between lesions and liver with focusing on selected area of the interest, the novel proposed of difference-of-features technique is presented in Section 5.2.2. The novel multiple ROIs selection method based on the characteristics of the lesions is introduced in Section 5.2.3. The combination of the two proposed method (difference-of-features and multiple ROIs) is presented in Section 5.2.4.

5.2.1 Feature extraction

The next stage in the proposed framework based on low-level features is feature extraction to perform the lesions classification. As discussed in Chapter 3, there is a large diverse set of low-level features could be used for lesion classification purpose. Those features come under three categories; intensity, shape, and specifically texture feature that gained more attention as presented in Chapter 3. However, the proposed framework used the combination features (intensity, texture and shape features). These quantitative imaging features were selected upon (1) most commonly used. (2) most descriptive lesions. As discussed in Chapter 3, the texture features proved to be useful in many work. Furthermore, the malignant lesions had higher intensity and broader distribution in the CT image (more heterogeneous) compared to the benign ones. In addition, there is a difference in the shape feature between Benign and Malignant lesions in terms of regularity. Thus, the combination would yield more robust results to handle various lesion types. The images are represented in matrices to use in mathematical process. In order to extract distinctive characters of the images, statistical feature extraction methods are used.

Intensity Features

The liver lesion intensity in the CT images is important for classification task. Intensity feature derived from histogram features which describe the relative frequency of pixel intensity value in the image. Thus, five different types of intensity features are computed, the features are: (1) Mean, (2) Standard Deviation, (3) Skewness, (4) Kurtosis, (5) Entropy of gray-level histogram.

- **Mean** (μ) calculate the estimation of the average level of intensity in the ROI region.

$$\mu = \frac{1}{N} \sum_{(x,y) \in ROI}^{N} I_{(x,y)} \tag{5.1}$$

  Where I(x,y) is the gray level at pixel (x,y), and I is the total number of pixel inside the ROI, and N is the total number of pixels inside the ROI of the lesion.

- **Standard deviation** (σ) is a measure of the dispersion of intensity.

$$\sigma = \sqrt{\frac{1}{N} \sum_{j=1}^{N} (I_j - \mu)^2} \tag{5.2}$$

- **Skewness** (γ_1) is a measure of histogram symmetry.

$$\gamma_1 = \frac{1}{N \times \sigma^3} \sum_{j=1}^{N} (I_j - \mu)^3 \tag{5.3}$$

- **Kurtosis** (K) is a measure of the tail of the histogram.

$$K = \frac{1}{N \times \sigma^4} \sum_{j=1}^{N} (I_j - \mu)^4 \tag{5.4}$$

- **Entropy** (H) is a measure of uncertainty of the histogram in the segmented region.

$$H = -\sum_{j=1}^{N} I_j \log I_j \tag{5.5}$$

Texture Features

Following the literature conclusion, The texture features are widely used and considered as an important key in medical image interpretation. For the lesion classification purpose, three different texture feature types are calculated.

- **Gray-Level Co-occurrence Matrix (GLCM).**
  GLCM is a robust statistical feature and widely used for extracting the second-order texture feature to characterised the spatial distribution of the gray levels in the images (Walker et al., 1995). The GLCM shows the arrangement of the gray-level pixel (i) with the same gray scale neighbourhood (j) in predefined direction (θ) and distance (d) to generate the co-occurrence matrix $(G(i, j|d, \theta))$. The Figure 5.2 depicts the reference pixel relative to angle and (θ) and distance (d).

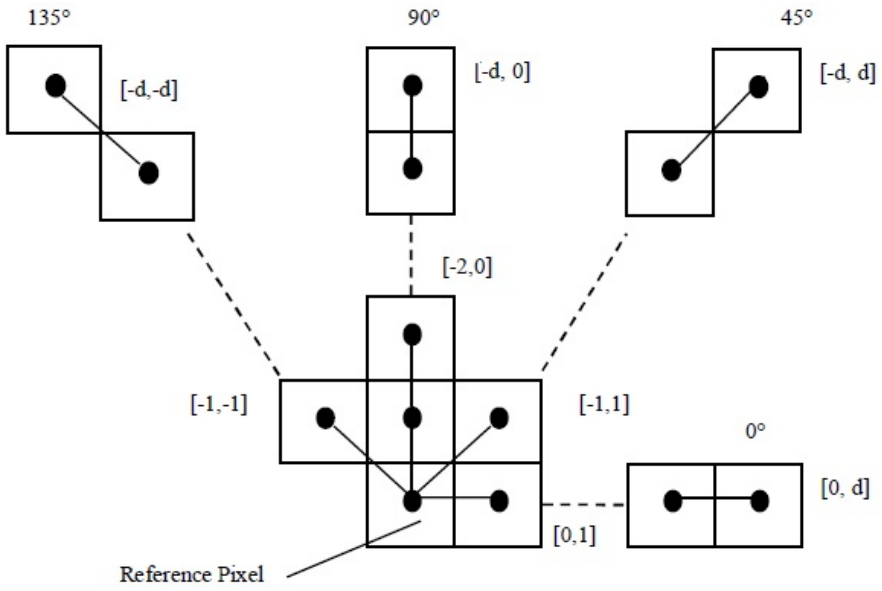


Figure 5.2: The direction of GLCM $(0^o, 45^o, 90^o, \ and \ 135^o)$ and distance (d).

In this thesis, the GLCM parameter for the distance pixels length of $d = \{1, 2, 3 \, and \, 4\}$ with orientation of $\theta = (0^o, 45^o, 90^o, \ and \, 135^o)$ are used. However, most of the research that used GLCM did not provide any justification for the selection of the distance (d) and angle (θ) (Susomboon et al., 2008). According to the analysis of variance technique and multiple pair-wise comparison (Clausi, 2002), they found that the use of "near" and "far" displacement was sufficient to capture the texture features.

According to the Haralick (Haralick et al., 1973b), the five texture coefficients derived from the GLCM matrix are utilised to capture the homogeneity and non-uniformity of the lesions. these coefficients are:
Energy quantifies the repetition of gray level pairs in an image:

$$Energy = \sum_{x=1}^{N} \sum_{y=1}^{N} \left(I_{(x,y)}\right)^2 \tag{5.6}$$

Entropy represents the randomness in the image.

$$Entropy = \sum_{x=1}^{N}\sum_{y=1}^{N} I_{(x,y)}\log_2(I_{(x,y)}) \tag{5.7}$$

Contrast is a local grey level variation in the GLCM. It can be thought of as a linear dependency of grey levels of neighbouring pixels.

$$Contrast = \sum_{x=1}^{N}\sum_{y=1}^{N} |x-y|^2 \times I_{(x,y)} \tag{5.8}$$

Homogeneity measures the uniformity of the non-zero entries in the GLCM. It weights values by the inverse of contrast weight

$$Homogeneity = \sum_{x=1}^{N}\sum_{y=1}^{N} \frac{1}{1+(x-y)^2} \times I_{(x,y)} \tag{5.9}$$

Correlation assesses the linearity of relationship between various gray level pixel pairs.

$$Correlation = \sum_{x=1}^{N}\sum_{y=1}^{N} \frac{(x-\mu_x)(y-\mu_y)I_{(x,y)}}{\sigma_x\sigma_y} \tag{5.10}$$

Where

$$\sigma_x = \sum_{x=1}^{N}(x-\mu_x)^2 \sum_{y=1}^{N} I_{(x,y)}$$

and

$$\sigma_y = \sum_{y=1}^{N}(y-\mu_y)^2 \sum_{x=1}^{N} I_{(x,y)}$$

- **Gabor Energy**
  Gabor energy reveals the localised frequency distribution pixels to describe the lesion texture homogeneity. The vector of lesion Gabor energy represented in 4 scales and 16 directions.

$$g_{(x,y)} = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(\frac{-1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi\right) \tag{5.11}$$

Shape Features

The shape features could be contributed to differentiate between Malignant lesion from Benign ones. Where Malignant lesions have different characteristic in shape from Benign. The four types of shape features are computed to define the lesion shape. These features as follows:

- **Dispersion feature** is estimating the irregularity of the lesion, which identifies the irregular shape by the equation below.

$$Dispersion = \frac{MaxRadius}{Area} \tag{5.12}$$

- **Elongation feature** differentiates the regular oval mass from the irregular. This value is given by the following equation.

$$Elongation = \frac{Area}{(2 \times MaxRadius)^2} \tag{5.13}$$

- **Circularity 1**. The circularity of the lesion is expressed by the following equation, where the result takes a value of 1 for perfect circles.

$$Circularity1 = \sqrt{\frac{Area}{(\pi \times MaxRadius^2)}} \tag{5.14}$$

- **Circularity 2**. the following formula is useful in differentiating circular/ oval lesion from irregular, where the result takes a value of 1 for perfect circles. This value measures how a lesion is similar to an ellipse.

$$Circularity2 = \sqrt{\frac{MinRadius}{MaxRadius}} \tag{5.15}$$

- **Roundness**. The roundness of the lesion is expressed by the following equation, where $R \in [0, 1]$. The bigger value of R means that the lesion shape is close to circle.

$$R = \frac{4\pi \times area}{MaxRadius^2} \tag{5.16}$$

5.2.2 Difference-of-Features (DoF)

After surveying the published papers in literature, it is observed that many researchers try to diagnose liver diseases using different techniques to increase the classification performance. However, it has been found that the previous studies on CAD systems usually used the absolute value of features, which are extracted from lesion regions. As a consequence, the performance is varied significantly under different acquisition conditions. For example, the CT machines or operators are different. In this section, the surrounding normal tissue of liver in the same image is used as reference. So for a certain feature, we calculate the difference of features between the lesion and surrounding normal liver tissue and employ it as a new feature vector in our proposed classification system based on low-level features.

Feature extraction is a crucial stage in the CAD system. Understanding the correlation between the lesion characteristics and corresponding imaging features is critical for image training, as well as for features extraction. Traditional CAD systems usually use the absolute features of the lesion area for the classification task. The major drawback of absolute features, in general, is the range of distinctive features of the same lesion type may vary, depending on the ratio of the contrast agent absorption resulting the amount of contrast agent injection and imaging time. Hence, to overcome this limitation, we proposed the DoF to calculate the contrasting (difference/relative) features between the lesion and surrounding area. This is due to the conspicuity of a liver lesion depends on the attenuation difference between the lesion and the normal liver. Contrast agents are usually needed to contrast the lesion and surrounding normal tissue. However, the contrast agent will be absorbed by the lesion, as well as by the liver parenchyma, but at a different ratio depending on the type of the lesion and the contrast-imaging phase. As a result, DoF technique uses contrast as a discriminative feature for lesions classification. Hence, the extracted lesion features are normalised by surrounding liver features based on the ratio of contrast agent absorption.

First of all, the proposed system defines two types of ROIs for extracting the features relating to intensity and texture. The first ROI is the lesion boundary (R_l), and the second ROI is the surrounding normal liver tissue (R_s) as shown in Figure 5.3. In contrast with existing works about the identification of lesions using one ROI (lesion area only), we also consider the second ROI (R_s) which surrounds the first ROI (R_l). Moreover, the second ROI will be used as well to extract low-level features. The difference of features between the first ROI and the second ROI will be employed as a new feature vector. However, there are some constrains to identify the second ROI: (1) The second ROI must be centrally surrounding the first ROI. (2) The ratio of the diameter between the (R_l) and (R_s) is heuristically chosen through exhaustive experiments to be 1:1.5, as presented in Table 5.1. (3) The first ROI is excluded from the second ROI region. (4) The liver vessels (V) are excluded from the (R_s) region. As displayed in Figure 5.3.e.

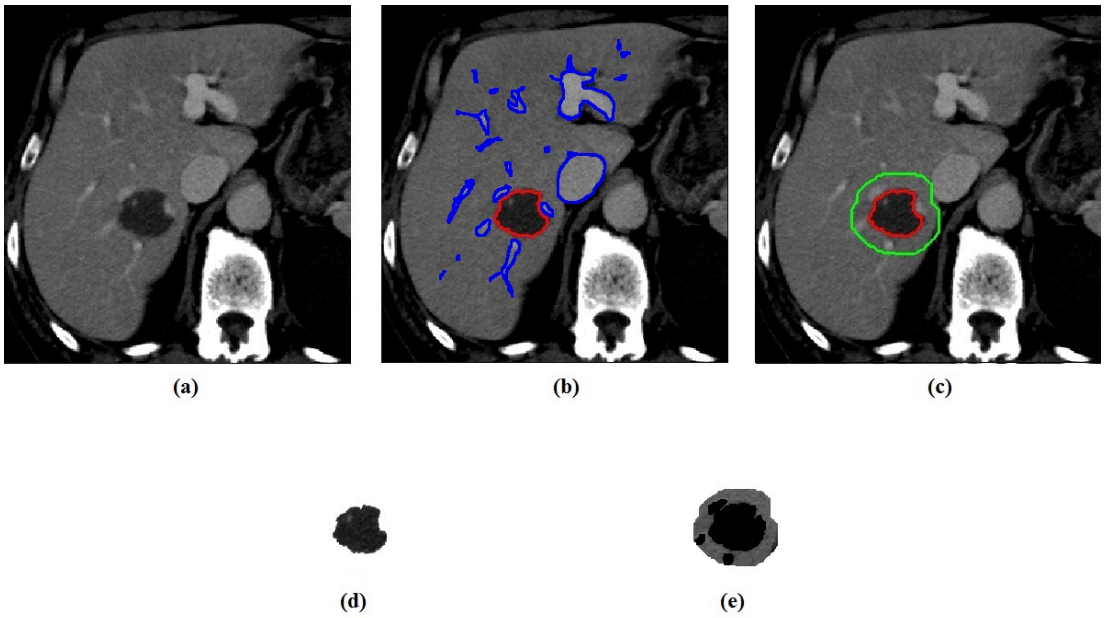


Figure 5.3: Proposed framework for Lesion and normal liver tissue ROI selection; (a) Original CT image; (b) Liver lesion and vessels detection; (c) First ROI is red border for segmented lesion and second ROI is green border for surrounding normal liver tissue; (d) Actual segmented lesion; (e) Normal liver tissue excluding lesion area and liver vessels.

Table 5.1 depicts the experiments results of DoF technique using different ratio between lesion and surrounding liver on the classification accuracy over our Dataset II.

ratio	**Sensitivity**		**Specificity**		**PPV**		**NPV**		**Average**
(Lesion:Liver)	**B**	**M**	**B**	**M**	**B**	**M**	**B**	**M**	**Accuracy**
1:0	0.863	0.745	0.745	0.863	0.829	0.792	0.792	0.829	0.815
1:0.5	0.890	0.784	0.784	0.890	0.855	0.833	0.833	0.855	0.847
1:1	0.904	0.784	0.784	0.904	0.857	0.851	0.851	0.857	0.855
1:1.5	**0.945**	**0.863**	**0.863**	**0.945**	**0.908**	**0.917**	**0.917**	**0.908**	**0.911**
1:2	0.932	0.843	0.843	0.932	0.895	0.896	0.896	0.895	0.895
1:2.5	0.880	0.765	0.765	0.880	0.859	0.796	0.796	0.859	0.836

Table 5.1: The experiments results of using different ratio between lesion and surrounding liver on the classification accuracy.

Table 5.2 presents the algorithm for extracting the normal liver tissue surrounding the lesion. Where the segmented area will be assigned as a second ROI.

ALGORITHM 1: Segmenting Surrounding Normal Liver Tissue (R_s)

INPUT:
Segmented Liver (LV);
Segmented Liver Vessels (V);
Segmented Lesion (R_l);

OPERATION:
FIND MAXDIM= maximum diameter of R_l;
CREATE Structure element, $X\_OUTSIDE = 1.5 \times MAXDIM$;
SET $\bar{R} = Dilate(R_l, X\_OUTSIDE) \in LV$;
SET $R_{sv} = \bar{R} - R_l$;
SET $R_s = R_{sv} - V \in R_{sv}$;

OUTPUT (R_s);

Table 5.2: The proposed programming algorithm to extract the second ROI from normal liver tissue that surround the lesion.

The Figure 5.4 depicts proposed framework to classify liver lesion based on low-level features by applying our proposed difference-of-feature technique.

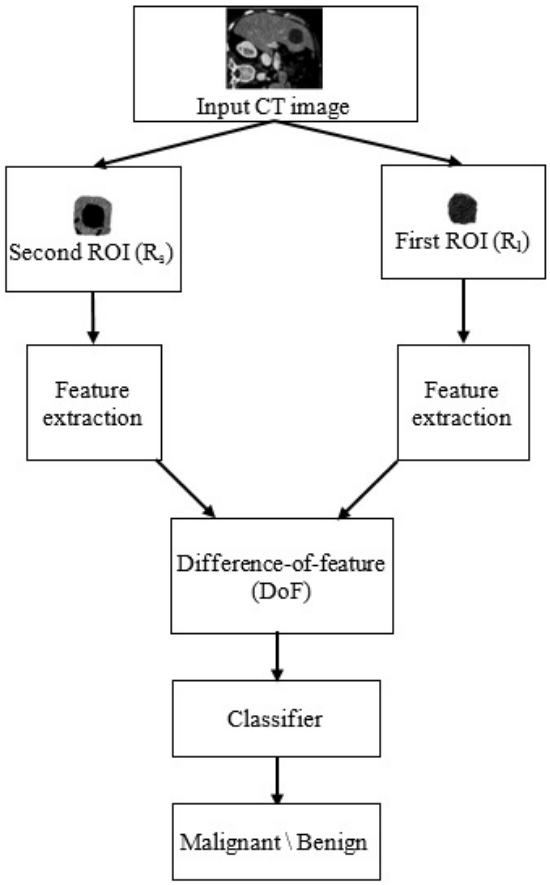


Figure 5.4: Proposed framework for Lesion classification based on difference-of-feature technique.

In our system, the low-level features, as discussed in Section 5.2.1, are extracted from the lesion and surrounding normal liver tissue area, and used the difference between them to generate a new feature vector and used in training a classifier. Equation 5.17 depicts example to calculate mean (μ) by applying difference-of-feature.

$$difference(\mu) = \frac{1}{N} \sum_{(x,y)\in R_s} I_{Normal}(x,y) - \frac{1}{M} \sum_{(x,y)\in R_l} I_{Lesion}(x,y) \tag{5.17}$$

Where $I_{Normal}(x,y)$ means the gray level at pixel (x,y) of normal surrounding liver tissue R_s, $I_{Lesion}(x,y)$ means the gray level at pixel (x,y) of lesion R_l, M is the total number of pixels inside the R_s of normal liver and N is the total number of pixels inside the R_l of lesion.

5.2.3 Multiple ROIs

This section presents a new method of the proposed multiple ROIs for liver lesion classification that efficiently utilises of ROI to build the stable classification framework with high performance. According to the literature, the majority of researches classify the liver lesion based on extracting low-level features from the lesion with no pay attention to the relation between lesion and liver where lesion classification performance depends mainly on the lesions characteristic such as surrounding area, internal structure, edge and morphology (Zhang et al., 2010b). However, the main limitation and lack of previous study (to the best of our knowledge) to represent all the lesion characteristics in a reliable way. Specifically, the lesion characteristics observation could differ based on the type of extracted ROI.

In this part, we proposed multi-region of interest from the internal lesion, border and surrounding area to capture all the lesion characteristics for better classify liver lesions. From segmented lesion (R_l), the (R_l) was divided into two areas (where $R_l = R_i \cup R_b$). (R_i) is the area internal the lesion and it considers that more than 80% of the pixels located in the central of the lesion. Otherwise, the remaining area between internal lesion and lesion boundary is considered as a lesion border (R_b). The area that surrounds the lesion from the liver tissue is taken as the third region (R_o). The Figure 5.5 shows multiple ROIs selection area which are: inside lesion, lesion border and outside lesion.

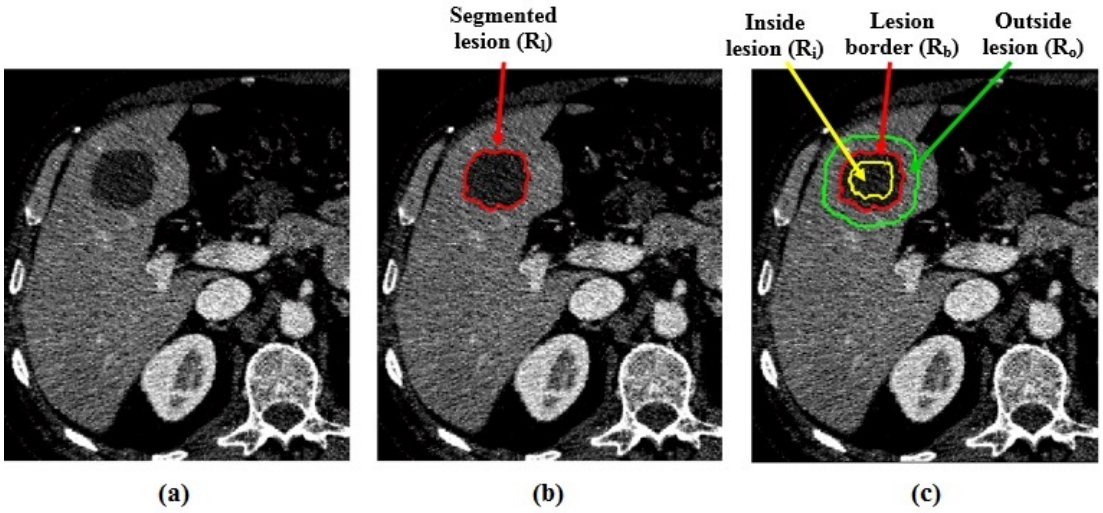


Figure 5.5: Proposed framework for multiple ROIs selection; (a) Original CT image; (b) The segmented lesion in red border; (c) Three defined ROIs which are outside lesion is green, lesion border is red and inside lesion is yellow.

Table 5.3 presents the algorithm for extracting the multiple ROIs from CT image. Where the inside lesion will be assigned as a first ROI (R_i), the second ROI for the lesion border (R_b) and the third ROI is outside lesion (R_o).

ALGORITHM 2: Extracting Multiple ROIs From CT Image

INPUT:
Segmented Liver (LV);
Segmented Lesion (R_l);

OPERATION:
FIND MINDIM= minimum diameter of R_l;
CREATE Structure element, $X\_INSIDE = MINDIM \times 0.3$;
SET $R_i = Erode(R_l, X\_INSIDE) \in R_i$;
SET $R_b = R_l - R_i$;
CREATE Structure element, $X\_OUTSIDE = 2 \times X\_INSIDE$;
SET $\bar{R} = Dilate(R_l, X\_OUTSIDE) \in LV$;
SET $R_o = \bar{R} - R_l$;

OUTPUT (R_i), (R_b), (R_o);

Table 5.3: The proposed algorithm to extract the second ROI from normal liver tissue that surround the lesion.

The low-level features within the internal, border and surrounding lesion were processed to compute intensity, texture and shape features from each ROI; as presented in Section 5.2.1. The feature-level fusion was adopted to combined all extracted features, which is considered as the most popular approach in data fusion (Zhou and Bhanu, 2008). As presented in Figure 5.6.

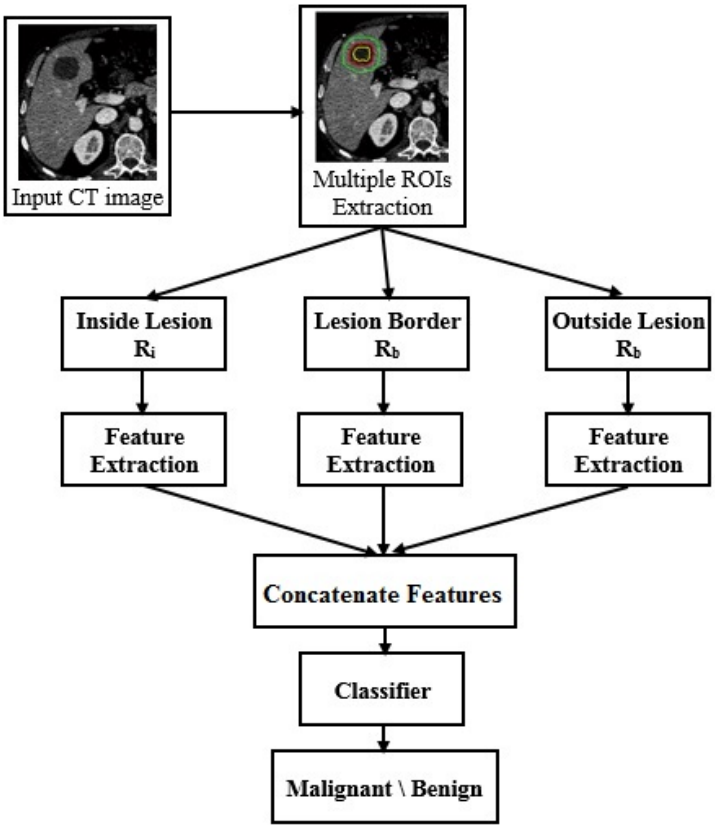


Figure 5.6: Proposed framework for lesion classification based on multiple ROIs.

5.2.4 Combined Multiple ROIs and Difference-of-features

In this section, we try to combine between the two previous proposed methods (DoF and Multiple ROIs) to generate a new feature vector that will be used in lesion classification task. From the segmented lesion R_l, we selected firstly multiple regions from inside, border and outside lesion each denoted by R_i, R_b, R_o, respectively. Multiple ROIs are chosen by interpreting the various lesion characteristics from each selected area as well as capturing the relation between the lesion and surrounding liver area, as illustrated in Figure 5.7.

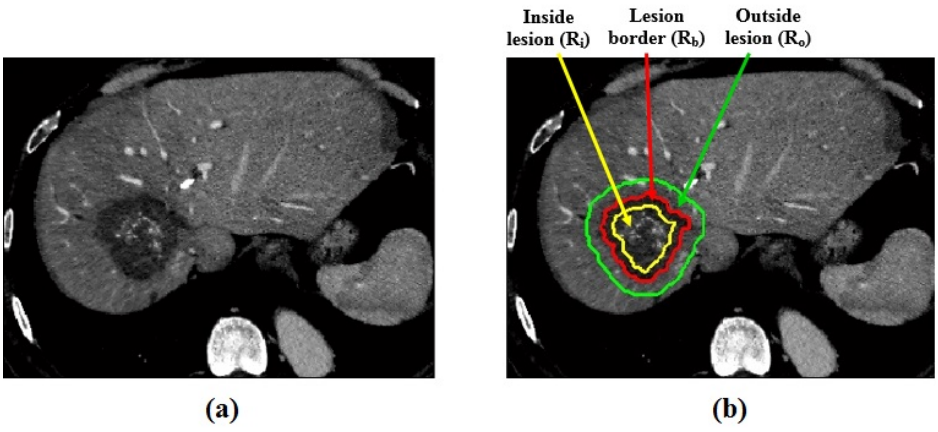


Figure 5.7: Divided lesion into Multiple ROIs; (a) Orignal CT image; (b) Multiple ROIs (inside lesion, lesion border and outside lesion).

From an input CT image $CT(x, y)$, the lesion (R_l) was detected firstly, as presented in Chapter 4. The lesion divided into two areas (inside lesion and lesion border) and denoted by $R_i(x, y)$ and $R_b(x, y)$, respectively, the liver tissue that surrounds the lesion R_l is assigned as a third ROI ($R_o(x, y)$). The multiple ROIs are selected to capture all the lesion characteristics of internal structure, border and the relation between lesion and liver. The low-level features extracted from each ROI to generate three feature vectors and each feature vector represent the features of extracted area, as formulated in Equation 5.18.

$$V_x = f(RIO_x) \qquad for\, x = inside,\, border\, and\, outside\, lesion \tag{5.18}$$

Where V_x is the feature vector, and $f(RIO_x)$ is the function that extracted the low-level features from the selected ROI. Then The difference-of-features applied between the two feature vectors (inside lesion and outside lesion) as follows:

$$V_{DoF} = V_{OutsideLesion} - V_{InsideLesion} \tag{5.19}$$

Where V_{DoF} is difference-of-features vector between outside lesion and inside lesion feature vectors. The new feature vector V_{DoF} and lesion border ($V_{LesionBorder}$) feature vector combined (fused) in one vector to feed into the classifier. The combined feature vector depicted in Equation 5.20.

$$F_v = \alpha V_{DoF} \cup \beta V_{LesionBorder} \tag{5.20}$$

Where F_v is fusion vector, α and β is weight that used in combination of the two feature vectors to generate linear fusion vector where the condition $\alpha + \beta = 1$ must be satisfied for the weighted average constraint (Weisstein and Weisstein, 2009). Through the experiments performed on the entire dataset, different values for α and β were tried, and the best results, as presented in Table 5.4, were $\alpha = 0.7$ and $\beta = 0.3$ that because the internal and surrounding lesion area are more represented for the lesion characteristics than the border area.

In this approach, a new feature vector is obtained by a linear combination of the feature vectors, obtained from lesion border features ($F_{Border} = \{l_1, l_2, l_3, ..., l_n\}$) and DoF between outside and inside lesion features ($F_{DoF} = \{d_1, d_2, d_3, ..., d_n\}$), where l, d is a feature and n is the number of features. α and β is the weighted average of a 2-vectors, with $\alpha + \beta = 1$. The weighted feature vector is $F_{DoF,Border} = \alpha F_{DoF} + \beta F_{Border}$. In the combined feature vector approach, the features vectors, from the lesion border and DoF, are concatenated to form a new feature vector $F_{DoF,Border} = F_{DoF} \cup F_{Border}$.

Weight		SN		SP		PPV		NPV		Average
α	β	B	M	B	M	B	M	B	M	Accuracy
0.3	0.7	0.851	0.813	0.813	0.851	0.842	0.823	0.823	0.842	0.833
0.4	0.6	0.894	0.863	0.863	0.894	0.884	0.873	0.873	0.884	0.879
0.5	0.5	0.947	0.913	0.913	0.947	0.927	0.936	0.936	0.927	0.931
0.6	0.4	0.957	0.925	0.925	0.957	0.938	0.949	0.949	0.938	0.943
0.7	0.3	**0.968**	**0.938**	**0.938**	**0.968**	**0.948**	**0.962**	**0.962**	**0.948**	**0.954**
0.8	0.2	0.957	0.913	0.913	0.957	0.928	0.948	0.948	0.928	0.937

Table 5.4: The experiment results of using different weight (α , β) on the classification accuracy.

The Figure 5.8 shows the overview of the proposed method to classify liver lesion based on the combination of multiple ROIs and difference-of-features (DoF). The DoF feature between inside lesion and outside lesion fused with lesion border features to generate a new feature vector. Then, the combined feature vector is feed into the classifier.

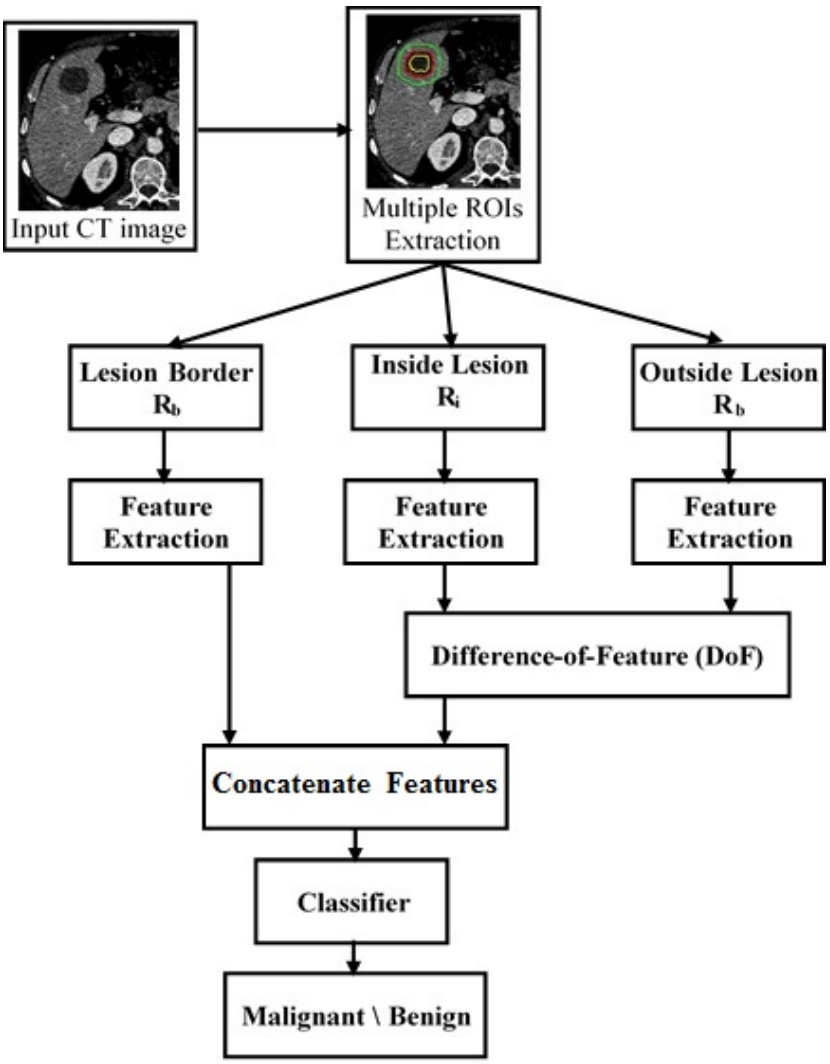


Figure 5.8: Proposed framework for lesion classification based on combination multiple ROIs and difference-of-feature technique.

5.3 Liver Lesion Characterisation

The main goal and contribution of this thesis is to characterise liver lesion automatically in CT images. In addition, exploit the relationships between low-level features and high-level features (semantic features) to enhance the liver lesion classification performance. Moreover, liver lesion characterisation provides radiological observations to interpret the classification decision. However, the importance of the explicit linkage between high-level features and low-level features has recently been emphasised (Depeursinge et al., 2014; Caputo et al., 2014). The second main contribution is enhancing lesion classification accuracy by using high-level features to classify the respective lesions and interpret the diagnosis decision. In contrast with most existing research (to the best of our knowledge), which uses low-level features only (black box) in lesion classification task, the use of high-level features and characterisation helps in interpreting and explaining the classification and is more intuitive to clinicians.

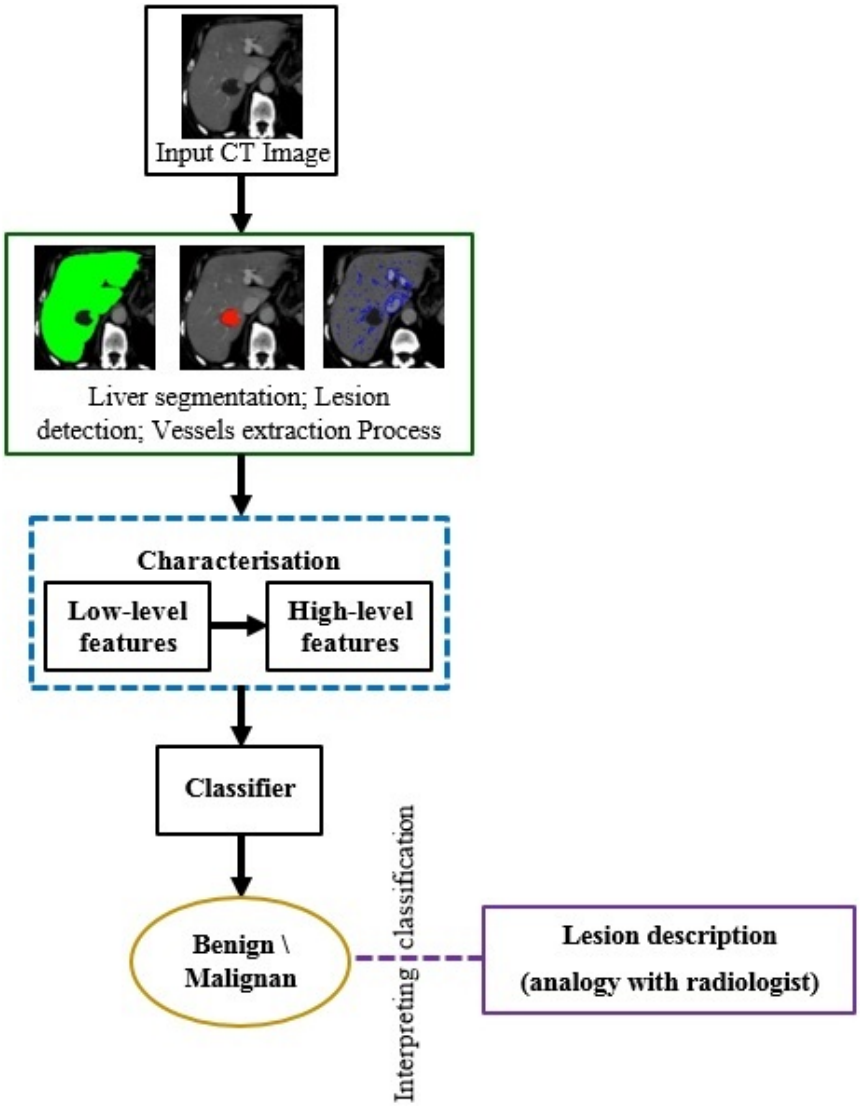


Figure 5.9: Overall proposed framework for lesion characterisation and classification based on high-level features.

An overview of the proposed framework for liver lesion characterisation and classification based on the high-level feature is shown in Figure 5.9. The first stage of the proposed system is segmenting liver, detecting lesions and extracting vessels as presented in Chapter 4. The second stage is characterising liver lesion through extracting high-level features by linking between low-level features and semantic meaning to describe the liver lesion in analogy to radiologist observation. The final stage is utilising the high-level features to classify liver lesion, the use of high-level features provides a human-interpretable explanation of the computer-based diagnostic decision to better-trusted diagnosis.

The next sections will present the core of proposed framework to characterise and classify liver lesion based on high-level features. The high-level features that calculated to characterise liver lesion are presented in Section 5.3.1. The Section 5.3.2 will present and discuss the characterisation framework in details to generate high-level features. The liver lesion classification based on high-level features is presented in Section 5.3.3.

5.3.1 High-level Features

Liver lesions originate from a variety of causes and appear in many variations on CT scan images; some are benign while others are malignant, and various diagnoses show a variety of visual manifestations. The ability to capture and characterise these lesions is important to differentiate between Benign and Malignant lesions. The malignant lesion has different characteristics compared to the benign lesion, as presented the key characteristics of benign and malignant lesion in Table 5.5.

Properties	Malignant	Benign
internal lesion structure	Wide range of changes; Heterogeneous attenuation; Invasion of adjacent structures.	Homogeneous; Water attenuation.
Lesion composition	Solid; Mix.	Cystic.
Rim	Thick enhancing rim; Continuous rim;	Thin rim; Absent rim; Discontinuous rim.
lesion shape	Irregular.	Regular; Round; Oval shape.
Lesion margin	Ill-defined margins.	Sharp margin.
Relation to surrounding tissue	Invades and destroy surrounding tissue.	Compress surrounding tissue.

Table 5.5: Liver lesion characteristics comparison between Malignant and Benign.

According to the importance of the lesion characteristics in the diagnosis process, each liver lesion has been characterised by a set of 21 high-level feature using a controlled vocabulary of 85 semantic terms, as depicted in Table 5.6. The aim of char-

acterisation stage is to analysis the imaging observations of the liver lesion domain with an emphasis on properties of the lesion such as internal structure, margin, shape and the relations between the liver and liver lesions through describing lesion rim and location.

High-level feature	Characterisation
Lesion density	Hypodense, Hyperdense, Isodense, NA.*
Lesion density type	Heterogeneous, Homogeneous, NA.
Lesion rim	Continuous Bright Rim, Discontinuous Bright Rim, Continuous Dark Rim, Discontinuous Dark Rim, NA.
Lesion rim thickness	Thick, Thin, NA.
Contrast Uptaken	Heterogeneous, Homogeneous, Dense, NA.
Enhancement Pattern	Hypoattenuation, Hyperattenuation, Isoattenuation, NA.
Lesion composition	Solid, Cystic, Mix, NA.
Lesion leveling type	Fluid fluid, Fluid gas, Fluid solid, NA.
Lesion shape	Irregular, Ovoid, Round, NA.
Lesion focality	Single lesion, Multiple lesions, NA.
Lesion margin	Smooth, ill defined, well defined, Irregular, NA.
Lesion margin definition	Defined, Diffuse, NA.
Lesion enhancement	Enhancing, Hypervascular, Nonenhancing, NA.
Lesion brightness	Hyperdense, Hypodense, Water density, NA.
Lesion surrounding	Complete, Incomplete, Absent, NA.
Calcified (inside lesion)	Yes, No, NA.
Calcified wall	Yes, No, NA.
Scar	Yes, No, NA.
Lobe	Left Lobe, Right Lobe, Caudate Lobe, NA.
Segment	Segment I, Segment II, Segment III, Segment IV, Segment V, Segment VI, Segment VII, Segment VIII, NA.
Close to vein	Left hepatic vein, Right hepatic vein, Middle hepatic vein, Left portal vein, Right portal vein, Middle portal vein, NA.

* NA is not applicable.

Table 5.6: The high-level features used in this thesis to characterise liver lesion.

Three aspects of the liver lesion characterisation were considered during proposing the framework. First, essential concepts, such as a lobe or a lesion. Second, individual properties of these concepts, such as the size or density of the lesion. Finally, the relationships between the concepts. The relations are important because they describe how different concepts relate to each other. For example, between a liver concept and a lobe concept, one can specify a LiverLobe relation to show that a liver contains lobes. Developing characterisation phase enables us to clearly specify cardinality or functionality requirements among relations. For example, a liver can have at most one

left lobe, while it can have many lesions.

Liver: The anatomical properties of the liver, such as its segments and lobes should be described. Additionally, for the referential model of segments and, regions must be defined as this is crucial in describing the location of an anomaly. Each liver is composed of three lobes which are: Right Lobe, Left Lobe and Caudate Lobe. The: LiverLobe property relates a Liver to each of its Lobes. In addition, the liver has an 8-segments. Each segment instance refers to a segment in the Liver. Hence, for example, caudate lobe can only be segmented by Segment I, whereas left lobe can be segmented by Segments II, III, or V.

Lesion: The characteristics of a lesion, such as margin, shape, internal structure, composition, calcification, and its contents must be defined. There may be abnormal areas of the liver a radiologist wishes to identify. The margin, shape and density of a lesion can be specified. For example, the shape of the lesion is described with the: LesionShape property that may take the following values: irregular, round, ovoid, and other. The density of a lesion is represented with: LesionDensity data property that takes a value of hyperdense, hypodense, or isodense. The lesion component such as calcification: LesionCalcified property indicates whether an area is calcified. Figure 5.10 shows The relations of the lesion with the characterisation concepts.

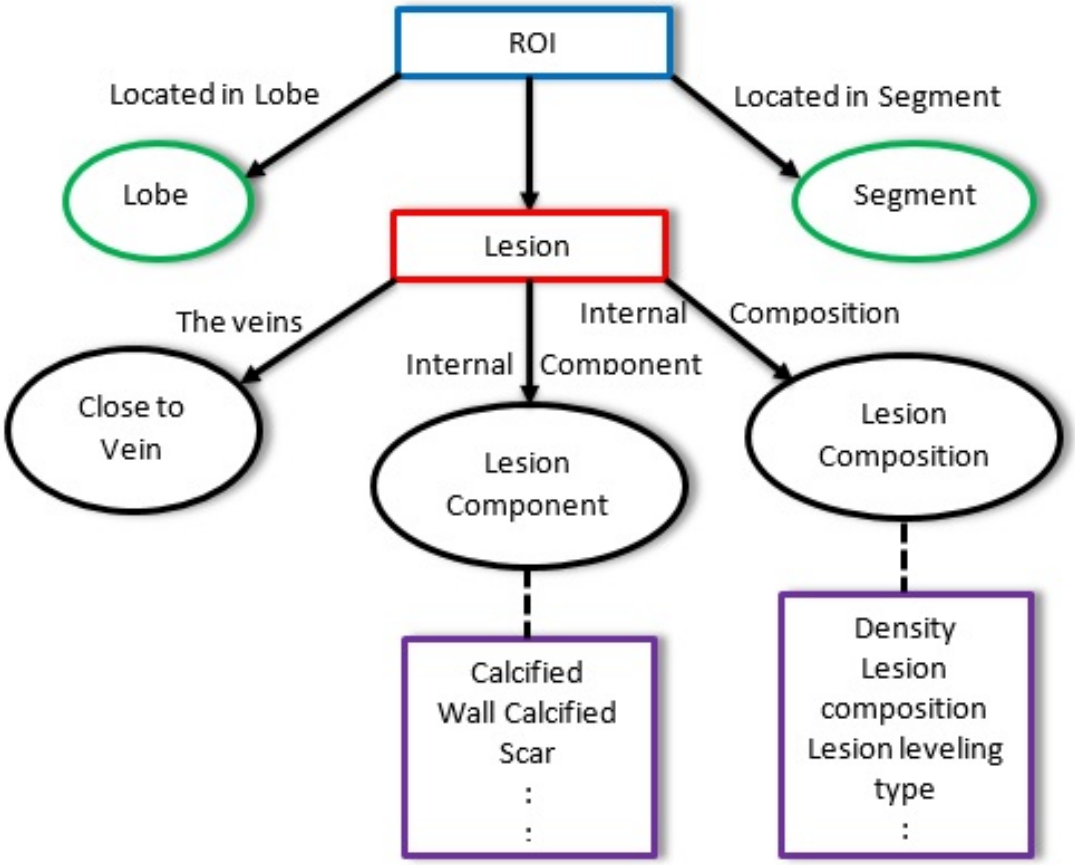


Figure 5.10: The relations of the lesion with the characterisation concepts.

The proposed liver characterisation method is composed of two main phases. The

first stage relies on machine learning process to characterise the lesion by exploiting the relationships between low-level features and high-level features. The second stage, the high-level features which are extracted from the image directly, defined as visual features from the image contents. Figure 5.11 depicts the overall proposed framework for liver lesion characterisation.

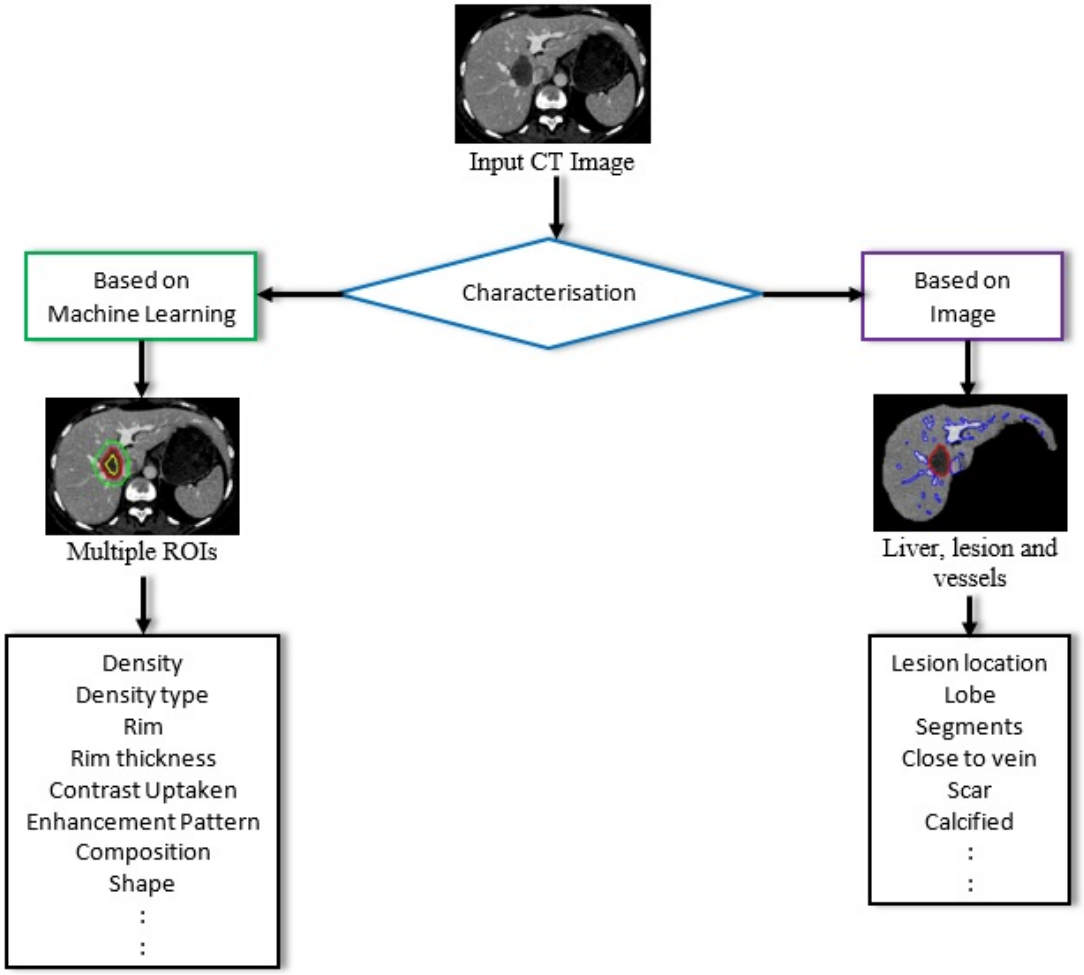


Figure 5.11: The Overall proposed framework for liver lesion characterisation.

In summary: According to the medical prior-knowledge for the characteristics of benign and malignant lesion based on the structure of the lesion, shape and their effect in the surrounding area, as reported in (Berland et al., 1988; Kondo et al., 2014; Murakami and Tsurusaki, 2014a; Minami and Kudo, 2015). In malignant lesions, the internal lesion structure shows a wide range of changes (heterogeneous attenuation) and invasion of adjacent structures; the lesion is surrounded by a thick enhancing rim; the lesion border is defined as irregular or having ill-defined margins. But in benign lesions, the internal structure is diffusely homogeneous; a thin or absent rim; the lesion border is round or oval shape with a sharp margin (Berland et al., 1988; Heiken, 2007; Murakami and Tsurusaki, 2014a). The proposed system will characterise the internal lesion structure through lesion density, lesion density type, lesion composition... The relation between a lesion and its surrounding area (i.e. rim enhancement) will be char-

acterised through lesion rim features. Furthermore, the lesion shape, size and focality add more useful information about the lesion's shape behaviour.

In this thesis, we contribute by presenting a CAD system for characterising liver lesions from CT scan images in analogy to radiologist observation. The proposed system categorises the high-level features into two groups; the first group is visual features from the image contents; the high-level features that are extracted directly from the image itself such as (Lesion location, lesion focality, Calcified, Scar, ...), as presented in Section 5.3.2; the second group is the high-level features that are calculated through machine learning to characterise the lesion such as (Lesion density, lesion rim, lesion composition, lesion shape, ...), as discussed in Section 5.3.3.

5.3.2 Visual Features From The Image Contents

This section presents the high -level features that are extracted directly from the CT image. Table 5.7 shows the high-level features which are not part of the general learning process but are rather estimated from the image itself.

High-level feature	Characterisation
Lesion focality	Single lesion, Multiple lesions, NA.*
Calcified (inside lesion)	Yes, No, NA.
Calcified wall	Yes, No, NA.
Scar	Yes, No, NA.
Lobe	Left Lobe, Right Lobe, Caudate Lobe, NA.
Segment	Segment I, Segment II, Segment III, Segment IV, Segment V, Segment VI, Segment VII, Segment VIII, NA.
Close to vein	Left hepatic vein, Right hepatic vein, Middle hepatic vein, Left portal vein, Right portal vein, Middle portal vein, NA.

* NA is not applicable.

Table 5.7: The high-level features computed directly from the CT image to characterised liver lesion.

The proposed approach is categorised the extracted high-level features from the CT image into two groups; the first group is the high-level features that characterised the liver lesion location. The second group is the high-level features that characterised the liver lesion component.

5.3.2.1 Lesion Location

The liver vessels segmentation is an important step for lesion location characterisation. In the first step, the liver area is segmented from the CT image as discussed in Chapter 4. Then the liver vessels are further extracted from the detected liver area as well presented in Chapter 4. Once the vessels are extracted in a 2D image, it is necessary to

construct the 3D vessels and assign them to either portal veins or hepatic veins for the couinaud liver segmentation.

Portal Veins and Hepatic Veins

The segmented 2D liver vessels are composed of two types of vessels which are hepatic and portal veins. At the first step the vessels skeleton is extracted from the segmented 2D vessels area in all slices to build the topology of vasculature. The importance of this step is preserving the vessels connectivity and extend through the slices in a 3D space. To achieve the goal, a 3D thinning algorithm is utilised to capture the vessels skeleton through categorising the voxels. To check the local voxel connectivity in a 3D space, a structure of 3x3x3 lattice points was built, as presented in Figure 5.12.

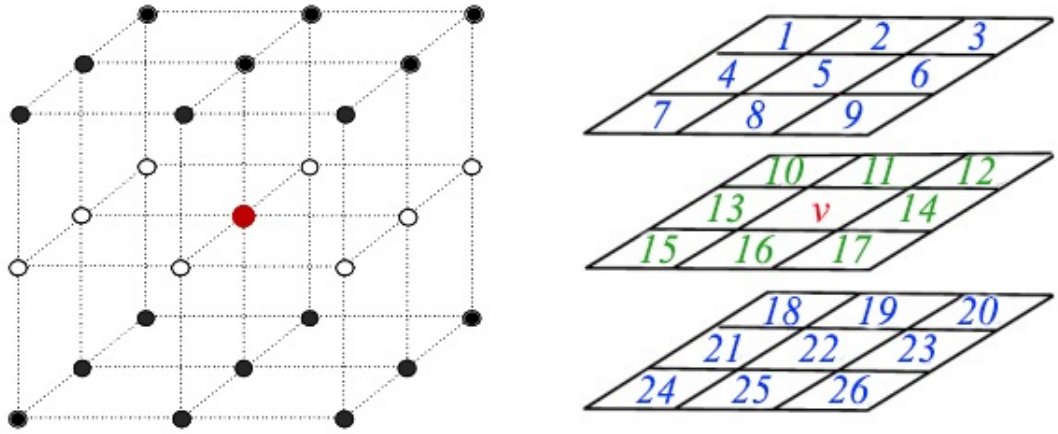


Figure 5.12: The neighbourhood connectivity of a voxel where the 8 and 18 neighbours of v are represented by white and black circles respectively.

For each check point v has two neighbours types:

- 8-connected neighbours pixels in 2D space, where $v \in V_{vessels}$ has at least one of its 8 neighbours is belonging to the vessels with a value of 1.

$$V_s = \{v | Number(N_8(v) = 1) \geq 1, v \in V_{vessels}\} \tag{5.21}$$

- 18-connected neighbours pixels in 3D space, , where $v \in V_{vessels}$ has at least one of its 18 neighbours is belonging to the vessels with a value of 1.

$$V_s = \{v | Number(N_18(v) = 1) \geq 1, v \in V_{vessels}\} \tag{5.22}$$

After tracing the vessels pixels, the connected component labeling algorithm (Chang et al., 2004; Hu et al., 2005) is applied to group the connected vessels on transverse image space and build a 3D model.

Once all vessels are extracted, it is important to classify these vessels either portal veins or hepatic veins, which are utilise later to divide liver area into eight segments

based on Couinaud segmentation method. The prior-knowledge is used in liver vessels classification process based on the following medical assumptions:

- The hepatic vein consists of three main branches (right, middle, and left hepatic vein) and they were vertically longer than 15% of the liver height, coming from upper to the bottom liver part (Oliveira et al., 2011b).
- The thickness of hepatic vessels decreases in flow direction (Selle et al., 2002c).

The first step, the hepatic veins with three main branches are identified from the binary vessels mask $V_{vessels}$ as follows:

I Choose the first slice Sc containing the vessels $V_{vessels}$ as a starting point.

II Select the largest connected pixels from the slice Sc and identified as a main hepatic vein (V_{Hvein}).

III Merge the connected pixels with (V_{Hvein}) if both conditions are met:

(a) The connected pixels is overlapping with the selected vessels in previous slice.

(b) The connected pixels thickness is equal or smaller than the selected vessels thickness in previous slice, based on second assumption.

IV Repeat step (III) until there is no overlap between (V_{Hvein}) and connected pixels.

V Compare the the height of detected vessels branch to the liver height. based on the first assumption, if the height of segmented branch compared to the liver height is equal or larger that 15%, then assigned as hepatic vein, otherwise it is rejected and restart the process.

VI Repeat the steps I to V until all the hepatic vein branches have been defined.

After finished the hepatic vein extraction process, the remain largest connected vessels pixels is defined as a portal vein vessels. Figure 5.13.a presents the 3D reconstructed liver vessels and Figure 5.13.b shows the 3D reconstructed hepatic and portal veins.

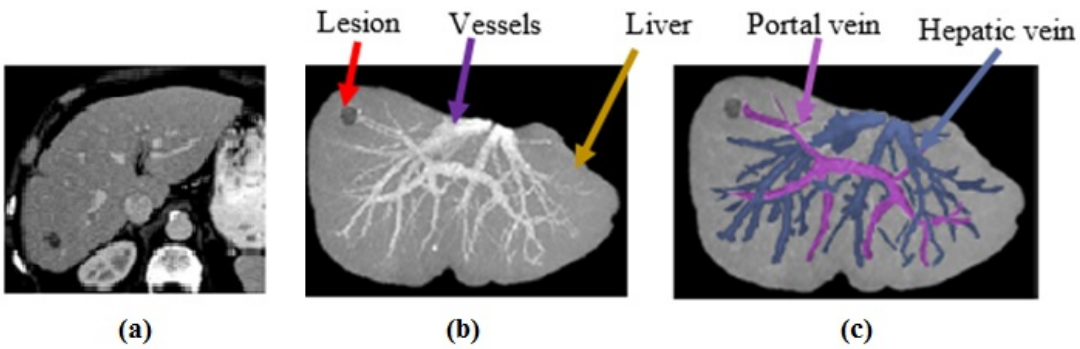


Figure 5.13: 3D Liver vessels construction and classification; (a) 2D CT scan image; (b) 3D volume rendering of liver with vessels and lesion; (c) 3D vessels classification, magenta is portal vein and blue is hepatic vein.

Anatomy of The Liver Segments

Liver anatomy can be described using two different aspects: (1) Anatomically, is not used surgically and that does not show the internal features of vessels. (2) Functionally based on Couinaud classification, is commonly used to define the anatomy of the liver segments by their relationship to vascular structures.

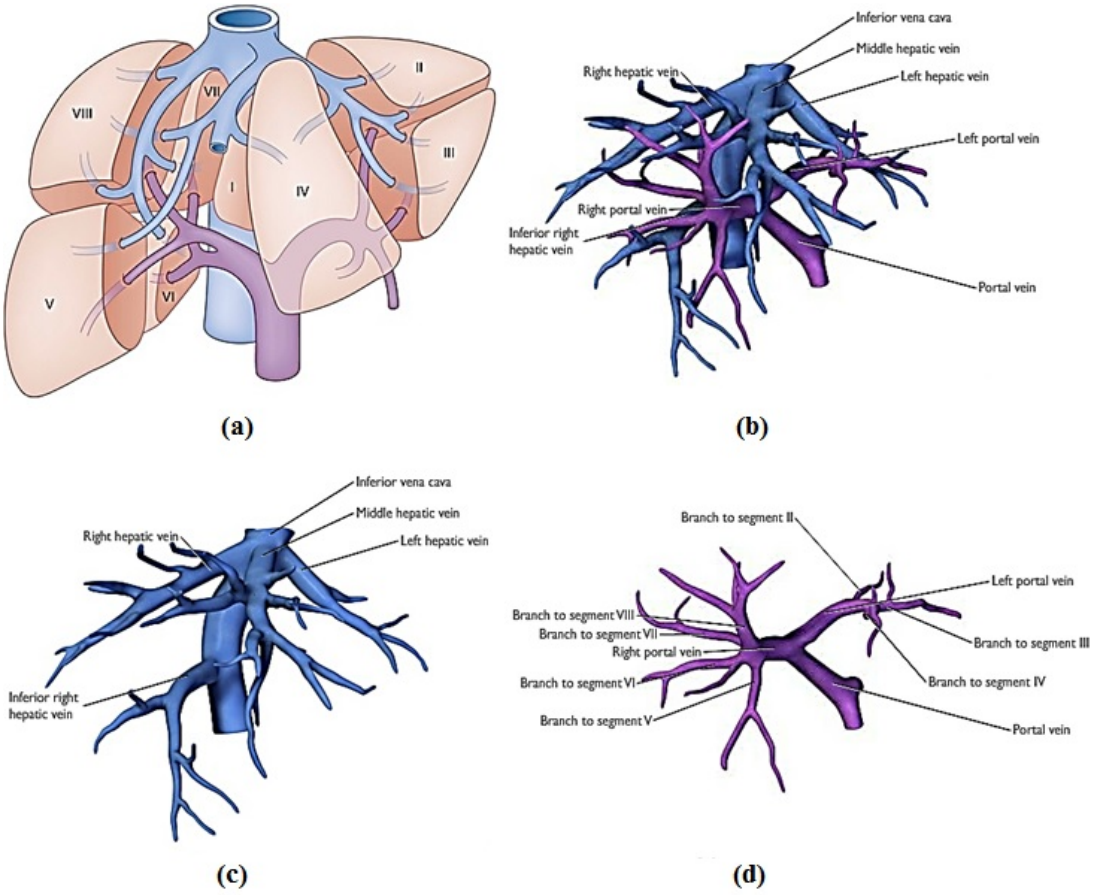


Figure 5.14: Illustrative figure of the liver Segmental anatomy with hepatic and portal vein; (a) liver segmental anatomy based on Couinaud classification system; (b) 3D of hepatic and portal vein; (c) 3D of hepatic vein; (d) 3D of portal vein.

The Couinaud classification system is based on the main branches of the portal vein. The Couinaud liver segment is the most widely used to define the anatomy of the liver since it is more convenient for surgery and more accurate for localisation and monitoring of the liver lesions (Germain et al., 2014). The liver divides into eight functionally independent segments based on the Couinaud classification where each segment has its own inflow, outflow blood vessels (Huang et al., 2007), as shown in Figure 5.14.

The liver vessels extraction is important for anatomical liver segments specifically hepatic and portal vein, as discussed in Section 5.3.2.1. According to the Couinaud classification system, each segment has centrally portal vein and peripherally hepatic vein (Bismuth, 1982; Marvasti et al., 2015). Each segment can be separated through hepatic vein where the portal vein divides the liver into upper plane that contains liver segments (II, IVa, VII, and VIII) and the lower plane that comprise liver segments (III, IVb, V, and VI). Table 5.8 depicts the relation between portal vein, hepatic vein and liver segments.

Liver Segment	**Hepatic vein** (HV)	**Portal vein** (PV)
Segment II	Left of Left HV	Superior to PV
Segment III	Left of Left HV	Inferior to PV
Segment IVa	Between the left and middle HV	Above to PV plane
Segment IVb	Between the left and middle HV	Below to PV plane
Segment V	Between the middle and right HV	Below to PV plane
Segment VI	Right of the right HV	Below to PV plane
Segment VII	Right of the right HV	Above to PV plane
Segment VIII	Between the middle and right HV	Above to PV plane

Table 5.8: Anatomy of The Liver Segments based on hepatic and portal vein.

Based on the functional segments and lobes, the proposed system used prior-knowledge to characterised liver segment and lobe that have a lesion. The liver consists of three lobes based on Couinaud classification method: (1) Caudal lobe which contains segment I. The segment I (Caudal lobe) is located between the portal vein and inferior vena cava (IVC). The IVC has been detected based on prior-knowledge where considering the the largest connected vessel in 2D from the portal vein will be seen the IVC. The IVC is located on the underside of the middle of the liver and has a circle/oval shape (Kogure et al., 2000; Hedman et al., 2016). (2) Left lobe which contains segment II, II and IV. (3) Right lobe which contains segment V, VI, VII and VIII. However, there is another method for anatomy liver segments which is considered the segment IV as a fourth lobe (quadrate lobe) (Rutkauskas et al., 2006). According to the prior-knowledge and experiments observation, the lookup table was built to utilise in liver

anatomical annotation. Table 5.9 shows the annotation results including the ratio of the appearance of liver segments in the CT slices based on transverse view point. These values of liver segments area ratio were estimated after observing more than half of all the dataset cases.

Liver Segment	Liver lobe	Slice	Appearing Ratio in Slice
VII, VIII	Right lobe	First slice	47%
IVa	Left lobe	15%	38%
II	Left lobe	20%	23%
III	Left lobe	35%	33%
IVb	Left lobe	52%	35%
V, VI	Right lobe	47%	53%

Table 5.9: Anatomy of The Liver Segments based on hepatic and portal vein.

5.3.2.2 Lesion Component

There may be abnormalities in the liver that the radiologist wants to recognise. These abnormal areas may have various components, such as lesion has scar or calcification. The calcifications are hyperdense on CT and are seen in Metastases, Cholangiocarcinomas, Fibrolamellar carcinoma and Hemangiomas (Murakami and Tsurusaki, 2014b). On the other side, The scar is visible as a hypodense structure on CT image and is seen in Fibrolamellar carcinoma, Hepatocellular carcinoma, Cholangiocarcinoma and Hemangioma (Blachar et al., 2002; Kim et al., 2009). Therefore, the characterisation of lesion components is important not only distinguish between benign and malignant but also identify the type of disease. Figure 5.15 shows two samples of liver lesion, The first Figure 5.15.a. liver lesion with scar and the second Figure 5.15.b. liver lesion with calcification.

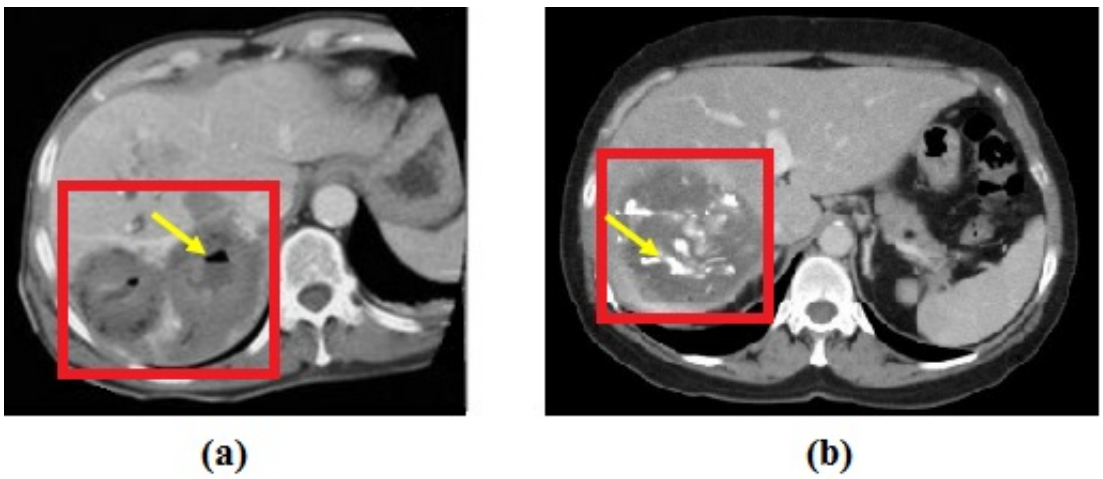


Figure 5.15: Two samples of liver lesions (red box); (a) Liver lesion with scar (arrow); (b) Liver lesion with calcification (arrow).

The proposed method relies on the following hypothesis: the lesion is darker than calcification, as calcification appears on radiographs and CT images, is easily visible as a hyper-dense or radiopaque lesion because calcification attenuates X-rays (Murakami and Tsurusaki, 2014b). The lesion is brighter than the scar where the scar is visible as a hypo-dense structure on the CT (Kim et al., 2009). From extracted lesion, the intensity analysis is performed to identify the Scar/ Calcification area. Due to the intensity variation between lesion and Scar/ Calcification area, the lesion segmentation results may have some of the holes. These holes are filled based on reconstruction step (hole-filling) which completes the lesion segmentation. The intensity of the segmented lesion before filling the holes is compared to the intensity of the holes within the lesion. The flow chart in Figure 5.16 depicts the process of detecting Scar/ Calcification region that starts with applying some morphological operations to the segmented lesion to fill the holes. Then both lesion images (after and before morphological operations) are subtracted from each other. This step is necessary to identify the Scar/ Calcification area. The intensity histogram of the segmented lesion H_l and the intensity histogram of the holes area H_a is compared to differentiate between lesion tissue and scar/calcification tissue, as defined in the following algorithm in the Table 5.10.

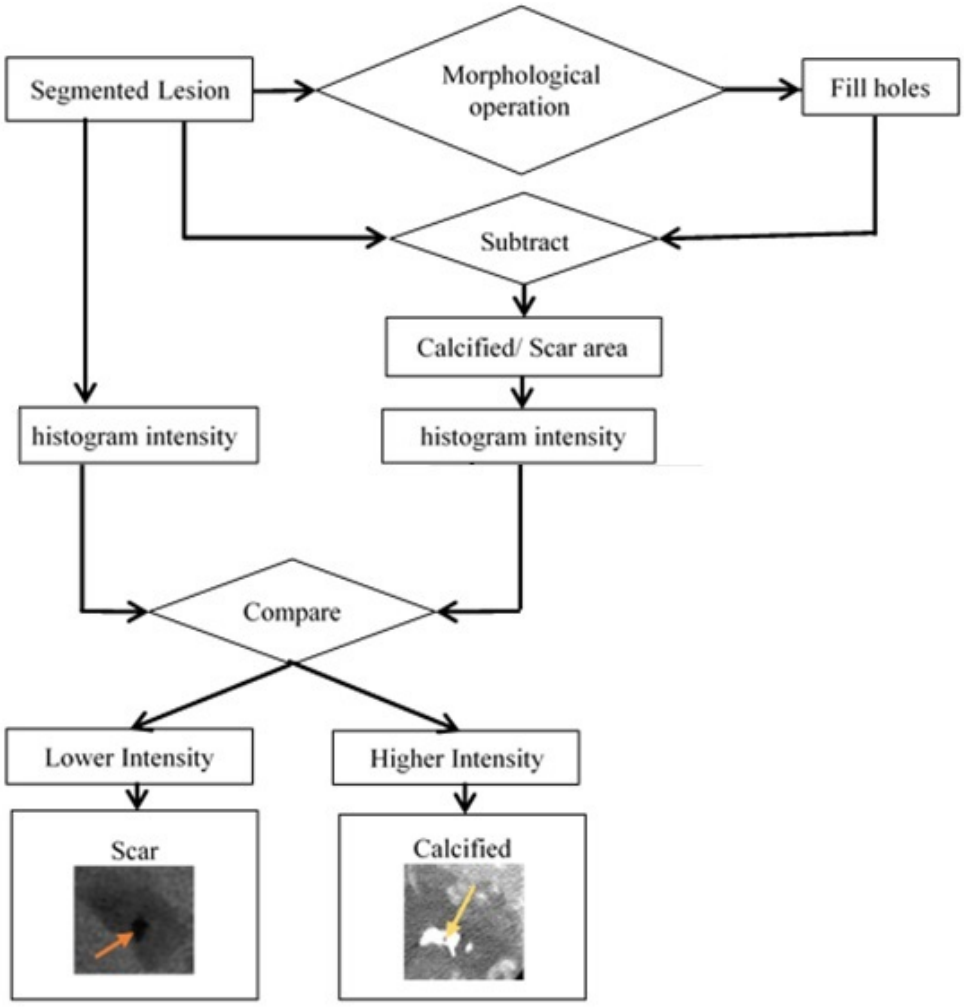


Figure 5.16: An illustration of the lesion scar/calcification detection process.

ALGORITHM 3: Detecting scar/calcification in the lesion

INPUT:
Segmented Lesion (R_l);

OPERATION:
SET R = Fill hole (R_l);
SET R_a = R - R_l;
SET H_l = Histogram((R_l);
SET H_a = Histogram((R_a);
SET P = Max Intersect Point (H_l, H_a);
IF P > Max Intensity (H_a) **OUTPUT** Calcification
ELSEIF P < Max Intensity (H_a) **OUTPUT** Calcification;
ELSE OUTPUT Lesion;

END

Table 5.10: The proposed programming algorithm to detect scar/calcification area from the lesion.

5.3.3 High-level Features Based on Machine Learning

In this section, we present the proposed liver lesion characterisation framework, based on machine learning, that emulate the human understanding of liver lesion images. Figure 5.17 displays the proposed framework overview to characterised the lesion based on machine learning. The proposed system needs to be trained through using low-level features that extracted from the lesion as an input and the high-level description as an output.

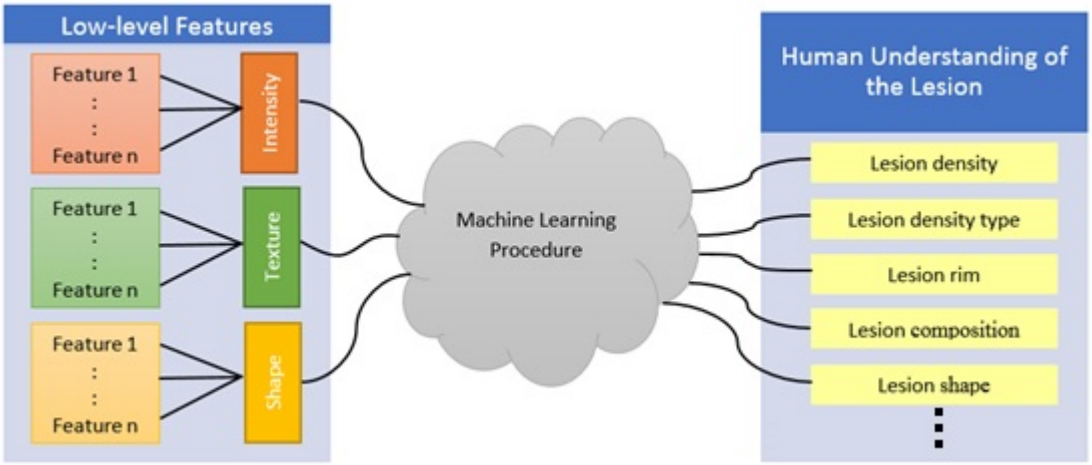


Figure 5.17: Overview of the lesion characterisation that is inferred from the low-level features through machine learning process.

Table 5.11 shows the generated high-level features that used to characterised the lesion based on machine learning. The learning process is utilised to linkage between

low-level features and high-level features through selecting the best-related region of interest (Multiple ROIs) to infer the high-level features by considering the ability of each ROI that represents a set of lesion characteristics.

High-level feature	Characterisation
Lesion density	Hypodense, Hyperdense, Isodense, NA.*
Lesion density type	Heterogeneous, Homogeneous, NA.
Lesion rim	Continuous Bright Rim, Discontinuous Bright Rim, Continuous Dark Rim, Discontinuous Dark Rim, NA.
Lesion rim thickness	Thick, Thin, NA.
Contrast Uptaken	Heterogeneous, Homogeneous, Dense, NA.
Enhancement Pattern	Hypoattenuation, Hyperattenuation, Isoattenuation, NA.
Lesion composition	Solid, Cystic, Mix, NA.
Lesion leveling type	fluid fluid, fluid gas, fluid solid, NA.
Lesion shape	Irregular, Ovoid, Round, NA.
Lesion margin	Smooth, ill defined, well defined, Irregular, NA.
Lesion margin definition	Defined, Diffuse, NA.
Lesion enhancement	Enhancing, Hypervascular, Nonenhancing, NA.
Lesion brightness	Hyperdense, Hypodense, Water density, NA.
Lesion surrounding	Complete, Incomplete, Absent, NA.

* NA is not applicable.

Table 5.11: The high-level features inferred from the low-level features based on machine learning to characterised liver lesion.

The lesion characterisation was reached based on segmenting the lesion and surrounding liver tissue with image processing methods. According to the enhancing lesion rim definition, that is an enhancing ring around the lesion (peripheral enhancement), the rim may be thin ($< 1cm$) or thick ($> 1cm$) (Elsayes et al., 2005; Martin et al., 2010; Jang et al., 2013). The area that surrounds the lesion from the liver tissue with size ($1.5cm$) is added to the segmented lesion and defined as (A_{ROI}), To ensure the capture of all lesion rim characteristics (thick / thin). The segmented area (A_{ROI}) was divided into three areas which defined as Multiple ROIs. Namely, The inner lesion (LR_{in}) and it considers that the pixels located between the central of the lesion and lesion border. The second area is lesion margin (LR_m) and denotes as the area between inner lesion and lesion edge. The area that surrounds the lesion from the liver tissue is taken as the third region (LR_{out}) to capture all the characteristics of the relation between lesion and liver.

Regarding multiple ROIs selection, the distance map is calculated for the (A_{ROI}) based on the intensity difference and the proximity distance for each voxel with respect to the normal liver tissue to generate the abnormality level map. The fast-marching method (Sethian, 1999) is adopted to generate the initial labeled regions. The speed of

fast-marching approach was empirically defined and it is equal 1 in pixels having intensity less or same the normal tissue, and 0.2 when the intensity difference is significant, due to the fast-marching approach might get stuck with the high intensity pixels such as calcified area. The abnormality map contains the zero value that represent the liver tissue and denoted by (LR_{out}), and positive values which define the lesion including the border and denoted by (LR). The further analysis of the abnormal area (LR), starting from the lesion centre and iteratively checking the abnormality neighbours in order to reduce abnormality pixels. The asymmetry and compactness features are calculated at each abnormality level to determine the inner lesion area (LR_{in}) with abnormality level is equal or above the lesion threshold, otherwise defined as lesion margin (LR_m).

Formally, consider a set of non-zero area $LR = \{l_1, l_2, l_3 ... l_n\}$ where LR is the set of (non-zero) areas in the partitioned abnormality map and n is the number of the areas in the abnormality map. For any area $l_i \in LR$ let $m(l_i)$ represent the abnormality value of the area l_i. For any subset of areas $V \subseteq LR$ let $m(V) = \min_{v \in V} m(v)$ indicates the lower abnormality value of areas included in V. let $X = LR/V$ where $LR/V = \{x \in LR \,|\, x \notin V\}$, for any $V \subseteq LR$ and any $l_i \in X$, the relation $f(V, l_i)$ considered true just only when l_i is at least neighbour for one area in V. For any $V \subseteq LR$ let $f(V) = \{d \in X | f(V, d)\}$ refers to a set of areas that are neighbour to any area in V and let $f_{-1}(V) = \{d \in f(V) | m(d) = m(V) - 1\}$ refers to the subset of the neighbourhood that contains only areas with an equal to the abnormality value of area V subtracted by 1. The area $l_i \in LR$ is considered as a maximum area when just $m(d) < m(l_i)$ for all $d \in f(\{l_i\})$, as depicted in Figure 5.18 for the abnormality level map illustration. The abnormality level b of the area embracing the maximum area l_i from the inner lesion is defined in Equation 5.23 to generate the lesion border mask and assigned as lesion margin (LR_m).

$$V_i^b = \begin{cases} \{l_i\} & , b = 0 \\ V_i^{b-1} \cup f_{-1}\left(V_i^{b-1}\right) & , 1 \leq b \leq m\left(l_i\right) - 1 \end{cases} \tag{5.23}$$

In order to build the distance map for each abnormality level $\left(V_i^b | 0 \leq b \leq m\left(l_i\right) - 1\right)$, the asymmetry and compactness features are computed for the area (l_i). These features are utilised to assign l_i (abnormality level) of V_i^l to represent a lesion area where the larger value is most likely to represent the lesion.

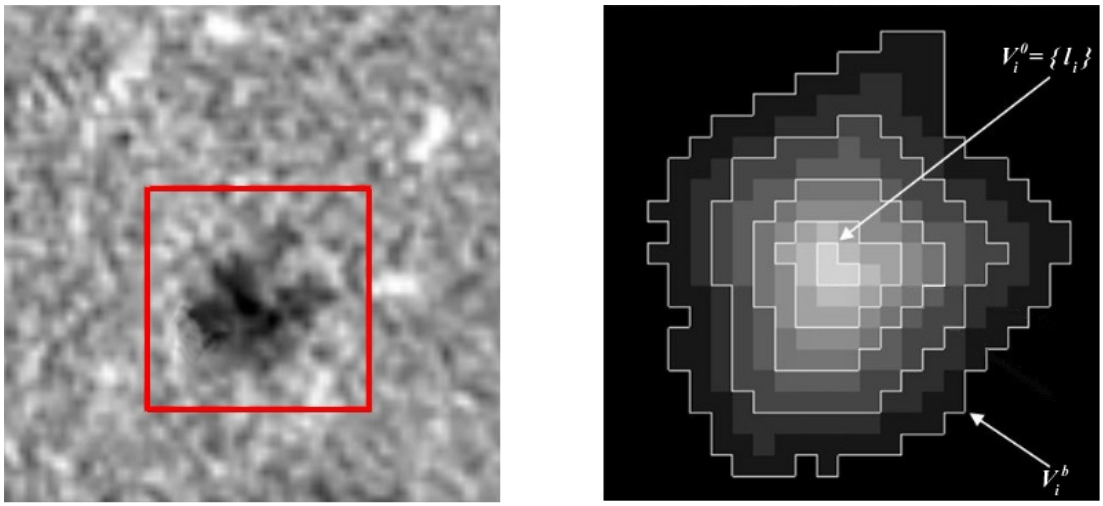


Figure 5.18: The abnormality level map for liver lesion; (a) Liver lesion CT image; (b) Liver lesion abnormality map involving in V embracing by l_i areas.

Regarding our proposed Multiple ROIs, the main idea for generating the level of abnormality map for lesion is to separate the lesion margin (LR_m) from the inner lesion (LR_{in}) area. Figure 5.18 depicts a small lesion (5.18.a) and the corresponding abnormality map (5.18.b). The brightest pixels of the abnormality map represent the maximum area l_i, where $m(l_i) = 6$, as shown in Figure 5.18.b. The first abnormality level surrounding l_i is V_i^1 where the set of V contains l_i and all areas which neighbour l_i with a total number of abnormality level equal to 5. Continuing the iteration ($b = 2, ..., 5$), each level of the abnormality area V_i is assigned as an inner lesion LR_{in} up to 80% of the lesion area. The remain abnormal level area is assigned as lesion margin LR_m. LR_{in} is the area internal the lesion and it considers that more than 80% of the pixels located in the central of the lesion. Otherwise, the remaining area between internal lesion and lesion boundary is considered as a lesion margin (LR_m).

Figure 5.19 depicts a summary of the proposed work-flow for liver lesion characterisation based on learning process. A set of expert-characterised CT images of liver lesion are utilised to train and validate the proposed framework. The validation methods and attempted experiments will introduce in next Chapter 6. The feature extraction block calculates a wide variety of intensity, texture and shape features by considering the lesion characteristics.

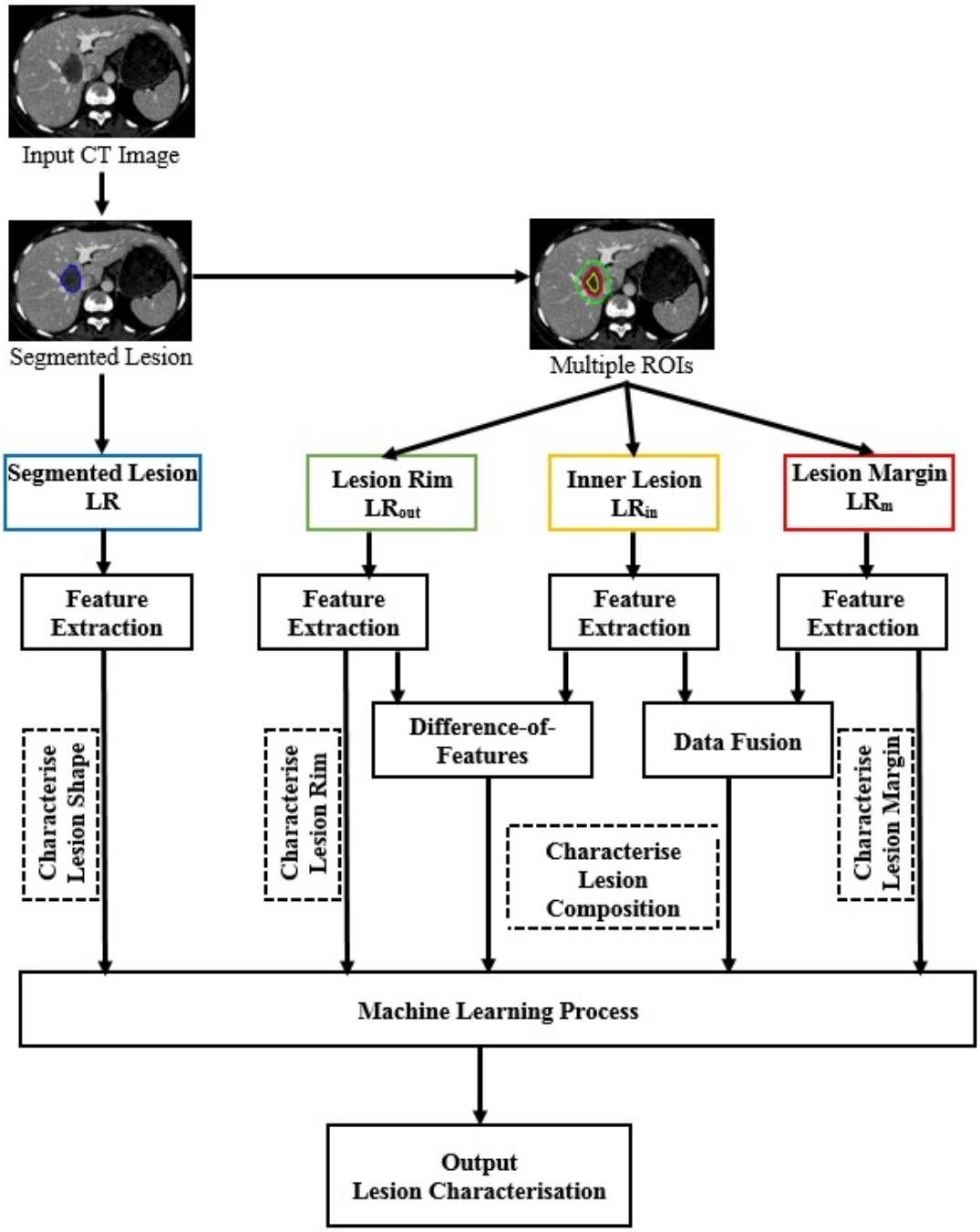


Figure 5.19: Proposed framework for liver lesion characterisation based on machine learning process.

A set of quantitative of low-level features were extracted from the segmented lesion (LR) and also from each area in multiple of ROIs that includes the inner lesion (LR_{in}), lesion margin (LR_m) and the rim area of the lesion (LR_{out}). Table 5.12 presents the low-level features that are used for lesion characterisation.

Category	Low-level Features	Represents	Dimension
Intensity	Histogram	Histogram of lesion intensity value.	32
	mean	Estimation of the average level of intensity value.	1
	Standard Deviation	Calculates dispersion of intensity value.	1
	Skewness	Measure of histogram symmetry.	1
	Kurtosis	Measure of the tail of the histogram	1
	Variance	The variation of intensity around the mean.	1
	Entropy	Measure of histogram uniformity.	1
	Energy	Measure of histogram homogeneity.	1
Texture	Gabor energy	A 27D vector of lesions Gabor energies in 3 scales and 9 directions.	27
	GLCM (Contrast)	is a local grey level variation in the GLCM (linear dependency of grey levels of neighbouring pixels).	1
	GLCM (Energy)	quantifies the repetition of gray level pairs in an image.	1
	GLCM (Correlation)	assesses the linearity of relationship between various gray level pixel pairs.	1
	GLCM (Homogeneity)	measures the uniformity of the non-zero entries in the GLCM.	1
Shape 1	Fourier descriptors	A 20D vector of the area Fourier descriptors.	20
	Smoothness	Smoothness of the lesion.	1
	Compactness	Compactness of the lesion.	1
	Sphericity	Sphericity of the lesion.	1
	Solidity	Solidity of the lesion.	1
	Roughness	Measure of boundary irregularity.	1
Shape 2	Dispersion	estimation the irregularity of the lesion.	1
	Elongation	differentiates the regular oval mass from the irregular.	1
	Circularity 1	differentiate circular/ oval lesion from irregular.	1
	Circularity 2	differentiate ellipse lesion from irregular.	1
	Roundness	differentiate circular lesion from irregular.	1

Table 5.12: Low-level features that used for lesion characterisation task.

From an input image, the lesion was segmented as a first step and denoted by (LR). A set of features (Shape 2) were extracted to generate a feature vector that used to characterised lesion shape. The (LR) was divided into two areas: inner lesion and lesion margin and denoted by (LR_{in}) and (LR_m) respectively. In addition, the surrounding lesion area from the liver selected as a third region and denoted by (LR_{out}). These three region defined as Multiple ROIs by considering the ability of each ROI that represents a set of lesion characteristics. The high-level features of lesion margin and lesion margin definition are characterised by extracting the set of features (Shape 1) from (LR_m). All the high-level features (Lesion density, density type, enhancement pattern, lesion composition, lesion leveling type and lesion brightness) are characterised by extracting the intensity and texture feature from both region (LR_{in}) and (LR_m) to generate two feature vectors. The features (LR_{in}) and (LR_m) are fused ($FV_{LR_{in}} \cup FV_{LR_m}$) to represent the mentioned high-level features. The intensity and texture feature are extracted from (LR_{out}) to characterised lesion rim, rim thickness and lesion surrounding. The difference-of-features of intensity and texture features that extracted from (LR_{in}) and (LR_{out}) are utilised to characterised the high-level features contrast uptaken and lesion enhancement.

Table 5.13 depicts the mapping between high-level, low-level features and selected ROIs that were used to generate the respective high-level features to characterise liver lesions.

High-level feature	**Low-level features**	**ROI**	**Feature vector**
Lesion density	Intensity + Texture	LR_{in}, LR_m	$FV_{LR_{in}} \cup FV_{LR_m}$
Lesion density type	Intensity + Texture	LR_{in}, LR_m	$FV_{LR_{in}} \cup FV_{LR_m}$
Lesion rim	Intensity + Texture	LR_{out}	$FV_{LR_{out}}$
Lesion rim thickness	Intensity + Texture	LR_{out}	$FV_{LR_{out}}$
Contrast Uptaken	Intensity + Texture	LR_{in}, LR_m	$FV_{DoF_{(LR_{in},LR_{out})}}$
Enhancement Pattern	Intensity + Texture	LR_{in}, LR_m	$FV_{LR_{in}} \cup FV_{LR_m}$
Lesion composition	Intensity + Texture	LR_{in}, LR_m	$FV_{LR_{in}} \cup FV_{LR_m}$
Lesion leveling type	Intensity + Texture	LR_{in}, LR_m	$FV_{LR_{in}} \cup FV_{LR_m}$
Lesion shape	Shape 2	LR	FV_{LR}
Lesion margin	Shape 1	LR_m	FV_{LR_m}
Lesion margin definition	Shape 1	LR_m	FV_{LR_m}
Lesion enhancement	Intensity + Texture	LR_{in}, LR_{out}	$FV_{DoF_{(LR_{in},LR_{out})}}$
Lesion brightness	Intensity + Texture	LR_{in}, LR_m	$FV_{LR_{in}} \cup FV_{LR_m}$
Lesion surrounding	Intensity + Texture	LR_{out}	$FV_{LR_{out}}$

Table 5.13: The high-level features inferred from the low-level features based on machine learning to characterise liver lesions.

5.4 Liver Lesion classification based on High-level Features

Classification stage of the diagnostic system (CAD) is the one in charge of making the inferences about the extracted information in the previous stages in order to be able to produce a diagnostic about the input image. The goal of the classification stage is to apply a learning-based approach considering its input feature vector(s), for the purpose of disease diagnosis. An overview of the lesion classification system through high-level features (liver lesion characterisation) is given in Figure 5.20.

Figure 5.20 depicts a summary of the proposed work-flow for liver lesion classification based on high-level features. The lesion characterisation block generates a wide variety of high-level features from each image, as presented in Table 5.6. The proposed framework categorises the high-level features into two groups, as shown in Figure 5.11; the first group is the high-level features that are extracted from the image contents such as (Lesion location, Lesion focality, Calcified, Scar, ...), as presented in Table 5.7; the second group is the high-level features that are inferred from the quantitative features through machine learning process to characterise the lesion such as (Lesion density, Lesion rim, Lesion composition, Lesion shape, etc.), as depicted in Table 5.11. The machine learning stage is used to predict the presence of each semantic term that describes liver lesions, which is inferred from the quantitative features by considering the lesion's intensity, texture and shape characteristics. Finally, we concatenate all the high-level features to create a single feature vector, and then using a new feature vector to classify the lesion. In contrast with most existing research, which uses hand-designed features, the use of high-level features (characterisation) helps in interpreting and explaining the classification decision.

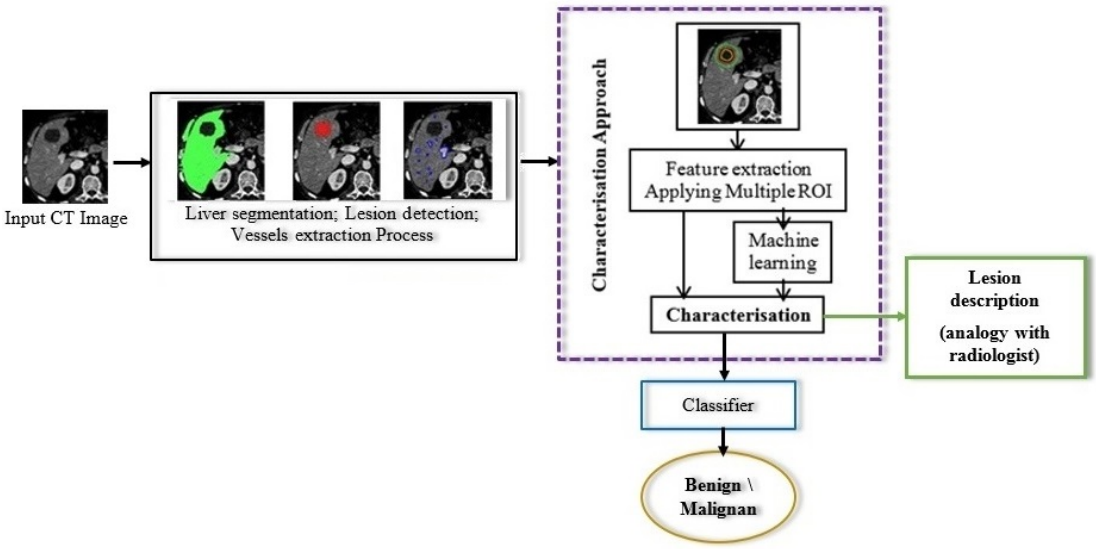


Figure 5.20: Proposed framework for liver lesion classification through high-level features (lesion characterisation).

Our proposed system consists of three main stages: (1) Automatic liver segmentation, lesion detection and vessels extraction. (2) characterised lesion by extracting high-level features from the lesion, as discussed in Section 5.3. (3) classification; the classifier is trained through using high-level features that produced via characterisation approach and the diagnosis as an output, with the benefit of interpretable characterisation that supports the diagnostic decision.

The characterisation approach provides more understanding and interpretation of the decision for the radiologists, compared to the black box low-level features approach. In contrast with most existing research, which uses low-level features only, the use of high-level features and characterisation helps in interpreting and explaining the classification and is more intuitive to clinicians. However, the classification decision based on low-level features is difficult to justify to radiologists and clinicians, who typically use subjective heuristics (lesion characterisation) to diagnose diseases.

5.5 Liver Lesion classification based on combination of High-level Features and Low-level Features

As discussed in Section 5.2, the liver lesion classification results from extracting low-level features to feed the classifier. Particularly, the novel of difference-of-features and multiple ROIs is utilised to enhance the classification accuracy. In Section 5.4, the novel feature vector from the high-level features is fed the classifier for better lesion classification performance, with the benefit of interpretable lesion characterisation in analogy to radiologist observation that supports the diagnostic decision. In this section, the lesion classification is performed through the novel feature vector that fused both, low-level feature vector and high-level feature vector.

Figure 5.21 presents the overall proposed framework for liver lesion classification based on the combination of low-level features and high-level features. Our method utilises both high-level and low-level features for lesion classification task. In classification, the input vector is created by concatenation of high-level and low-level features.

As discussed in Section 5.2.4, we extracted low-level features based on the combination of multiple ROIs and DoF techniques. The Dof features between inside and outside lesion concatenated with lesion border features to generate a new low-level feature vector. On the other hand, the lesion characterisation feature vector is based on the visual features derived from the image contents, as presented in Section 5.3.2, and the high-level features inferred through the learning-based approach, as described in Section 5.3.3. Therefore, the classification decisions made by a combination of low-level and high-level features.

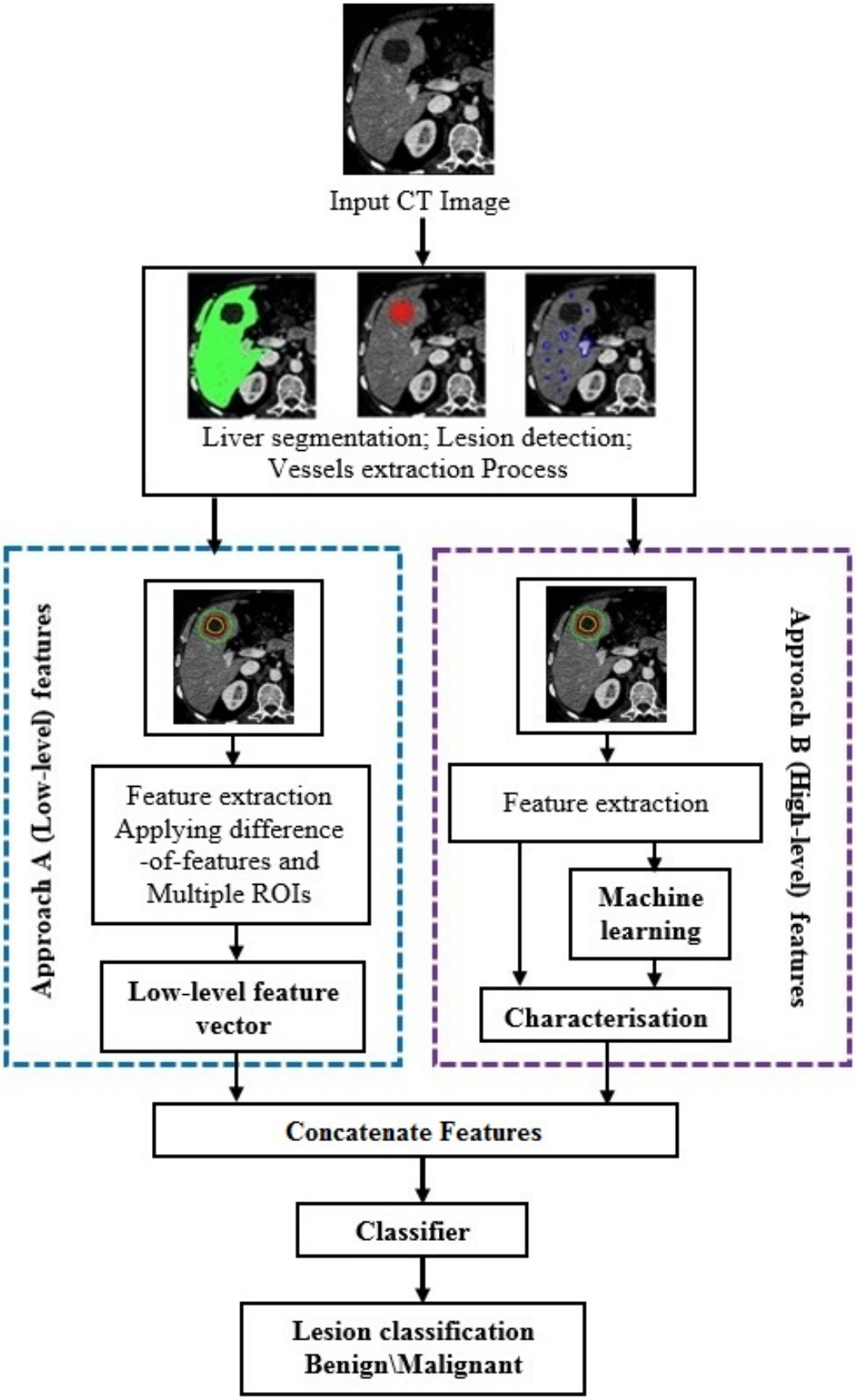


Figure 5.21: Proposed framework for liver lesion classification through combination of high-level features and low-level features.

5.6 Feature Selection

In order to increase the classification/characterisation accuracy and decrease the system costs, the feature selection approach can be used to select the most robust features from the high dimensional feature set (Jain et al., 2000). The system performance might be adversely affected by high dimensional feature set due to redundancy or lack of importance of some features (Pappu and Pardalos, 2014; Li et al., 2016). Hence, the main goal of a feature selection approach is to search for an optimal subset of relevant features and reduce the redundancy (Sun et al., 2013; Tang et al., 2014).

In this study, the Genetic Algorithm (GA) (Siedlecki and Sklansky, 1989) was adopted to measure the relevance and significance of the features and avoid the redundancy. The GA is a general adaptive optimisation procedure, which is utilising to reduce the dimensionality of the features where GAs has been successfully applied to a wide range of dimensionality reduction studies (Adams et al., 2015). Each variable is represented by a gene and the sequence of genes is called a chromosome. A number of chromosomes (population) are randomly initialised by three genetic operators: selection, crossover and mutation. The chromosomes are evaluated by a predefined fitness function to measure their quality. However, the selection operator utilises to select the high performing chromosomes to transfer it directly to the next generation. The new offspring chromosomes are created by swapping a portion of chromosomes (genes) between two chosen chromosomes and this mechanism is called a crossover operator. The mutation operator modifies one or more gene value of a selected chromosome to provide the better solution, as shown in Figure 5.22.

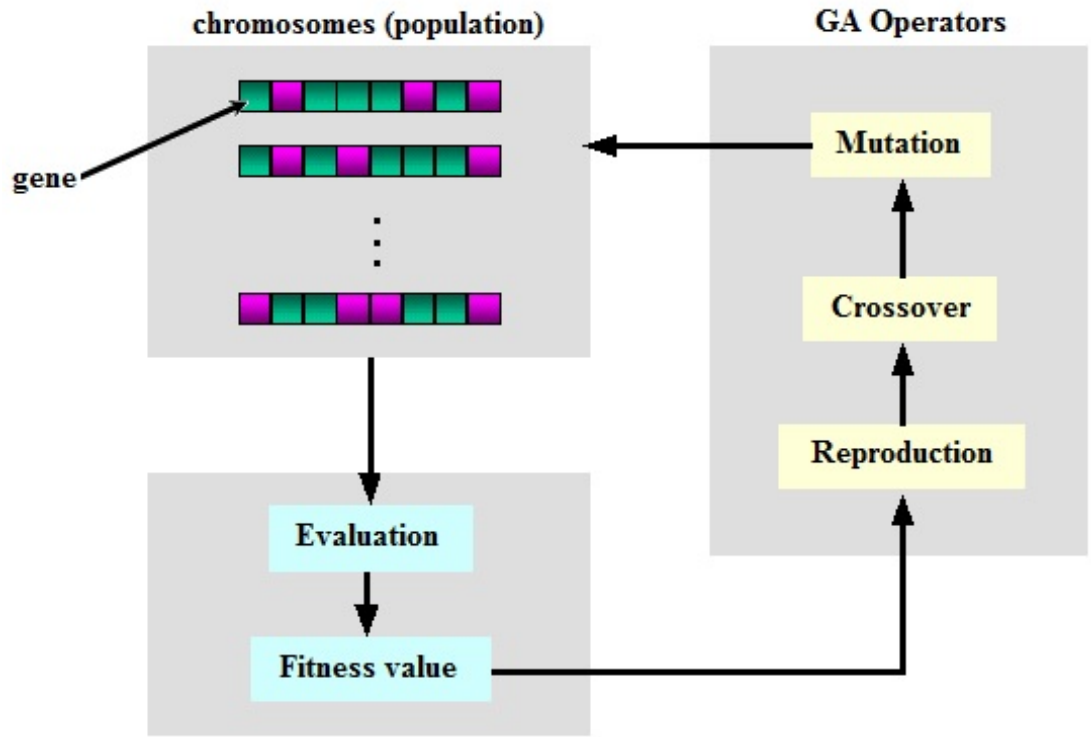


Figure 5.22: Genetic algorithm (GA) flow chart.

The schematic theory shows how the schemata featuring that appears in chromosomes with a high degree of expectation has a greater expectation of propagation through successive population as a GA evolves (McCall, 2005), as shown in equation 5.24.

$$m_H(i+1) = F_H(i)m_H(i)\left[1 - p_c\frac{l_H}{l-1}\right]\left[(1-p_m)^H\right] \tag{5.24}$$

where H is a schema features; m_H is the number of chromosomes belonging to H, $F_H(i)$ is the relative fitness of H that defined as the average fitness of all chromosomes in the population (i) belonging to H divided by the average fitness of all chromosomes in the population; l is the length of chromosome; p_m is the mutation probability, and p_c is the crossover probability.

The GAs showed better results compared to other feature selection techniques to generate a more robust feature vector, in studies related to the lesion classification/characterisation (Gletsos et al., 2003; Mougiakakou et al., 2007b; Aalaei et al., 2016). In addition, GA is considered to be an excellent choice for feature selection task due to relative insensitivity towards noisy data (Osowski et al., 2009). It differs significantly from the other existing wrapper algorithms because of the traditional methods search from a single population point unlike GA which searches from parallel population points (Akhter et al., 2016). GA is advantageous over other algorithms since it is less likely to be trapped by local minimum and provides a better global optimal solution (Garg, 2010; Ling and Liu, 2015). Hence, the GA was adopted to fulfil this task.

5.6.1 Implementation of GA

Using GA for feature selection, each feature is considered as a gene (chromosome), represented as 1 (selected) or 0 (not selected). The classification accuracy (the fitness of the chromosome) was determined as the area, A_z, under the ROC curve. The fitness function $F(c)$ for the c^{th} chromosome depicted in Equation 5.25.

$$F(c) = \left[\frac{f(c) - f_{\min}}{f_{\max} - f_{\min}}\right]^2, c = 1, 2, ..., n \tag{5.25}$$

Where $f_{\min}$ and $f_{\max}$ is the minimum and maximum $f(c)$ among the n chromosomes respectively.

The fitness function $F(c)$ based on the A_z value ranged from 0 and 1. The chromosome with the largest A_z value is assigned a fitness of 1, the chromosome with the smallest A_z value is assigned a fitness of 0. The smallest value for A_z is $A_z \leq 0.5$ and

largest value is $0.5 < A_z \leq 1$. The probability of the c^{th} chromosome being selected as a parent, $P_s(c)$ is proportional to its fitness. $P_s(c)$ calculated in Equation 5.26.

$$P_s(c) = \frac{F(c)}{\sum_{c=1}^{n} F(c)}, c = 1, 2, ..., n \tag{5.26}$$

A random sampling based on the probabilities ($P_s(c)$) allowed chromosomes with higher value of fitness to be chosen more frequently. The crossover rate determines the probability that parents will exchange genes. After crossover, another chance of introducing new features was obtained by mutation. The processes of parent selection, crossover, and mutation resulted in a new generation of n chromosomes. The best subset of features was selected to be the chromosome that provides the highest average A_z during the evolution process.

In this thesis, the follow parameters were used. The initial probability of a feature's presence (P_{init}), probability of crossover (P_c) and probability of mutation (P_m) was 0.002, 0.9 and 0.001 respectively. For better results, several studies suggest the use of a high value of crossover probability and a low value of the mutation probability (Chtioui et al., 2009; Sipper et al., 2018).

Figure 5.23 illustrates the Genetic algorithm evaluation for feature selection. Figure 5.23.a depicts the evolution of the number of selected features and Figure 5.23.b depicts the total area under the ROC curve A_z for the GA.

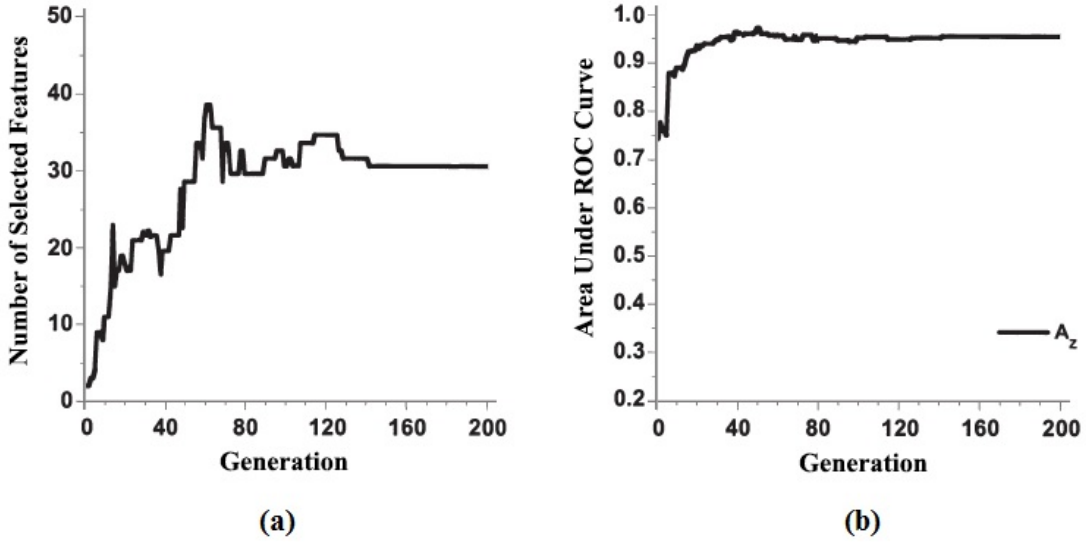


Figure 5.23: The Genetic algorithm evaluation for feature selection. (a) The evolution of the number of selected features for a GA. (b) The evolution of the area A_z under the ROC curve for the GA

The classifiers performance based on the area A_z under the ROC curve for different feature set sizes for the GA feature selection method are illustrated in Figure 5.24. It

is noted that the classification accuracies of the three classifiers have improved and the best performance of SVM was 0.97 at 39 number of features, compared to LR and LDA was 0.95 and 0.94 respectively at 78 number of features, as shown in Figure 5.24.

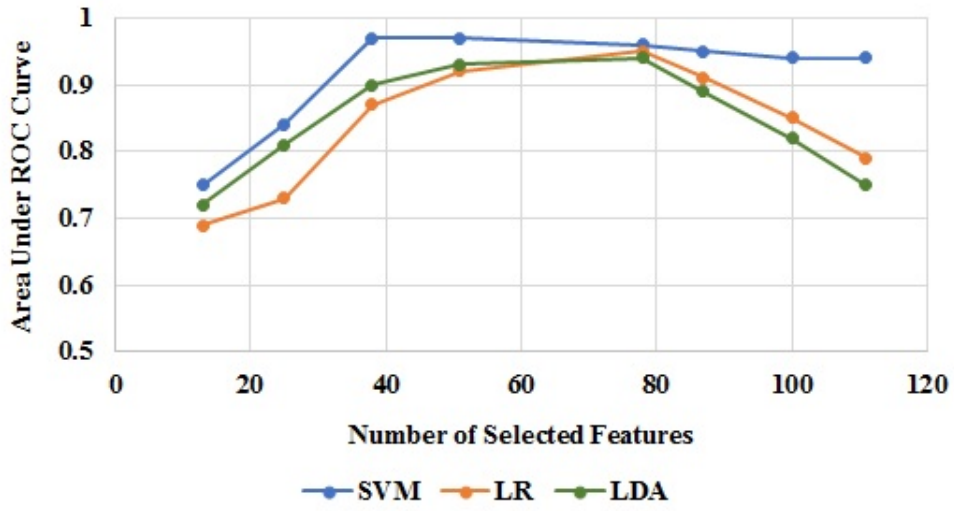


Figure 5.24: The comparisons of the area A_z under the ROC curve for each considered classifier based on the number of features.

The SVM classification performance comparison using ROC analysis is shown in Figure 5.25. The overall ROC performance of classification without using GA feature selection is $A_z = 0.94$, with comparison to $A_z = 0.97$ of classification performance after applied feature selection. For feature selection approach, less than half of the features were selected and at least one feature was selected from each category (intensity, texture and shape feature). Thus, the three categories of features complement each other to achieve better results. The results show that feature selection can improve the accuracy of the classifier.

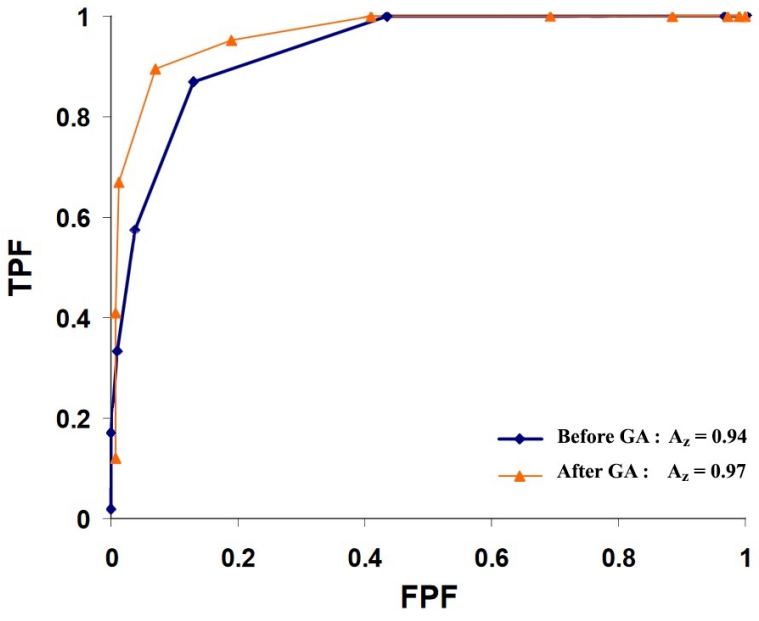


Figure 5.25: The classification performance comparison based on using GA.

We perform the GA feature selection approach to evaluate its performance. Table 5.14 summarised the classification results based on area A_z under the ROC curve, respect to the number of selected features. Using GA, the selected feature set contains only 38 of the available 111 features to achieve the best classification accuracy of 97%. The feature vector size is reduced by 65.75%. The feature set number 4 has the similar accuracy to the feature set number 3 but with an increase in the number of features by 11.71%. This theoretical result is due to the good separation between data in the selected base.

Feature set	Number of features	A_z	Percentage of reduction
1	13	0.75	88.29%
2	25	0.84	77.48%
3	**38**	**0.97**	**65.76%**
4	51	0.97	54.05%
5	78	0.96	29.73%
6	87	0.95	21.62%
7	100	0.94	9.91%
8	All features	0.94	0%

Table 5.14: The number of features and the area A_z under the ROC curve in the GA feature selection approach

The best classification accuracy 97% is achieved by using only 38 features: 9 bins of histogram, Skewness, Kurtosis, Entropy, GLCM (Contrast, Homogeneity and Correlation), 21 bins of Gabor Energy, Elongation and Roundness features. The selected features by GA approach are related with the lesion appearance and shape. By examining the CT images of the pathological area, we can see that the lesions vary in the degree of brightness, distribution and regularity of the shape. This explains the reasons for choosing these features. For example, the Elongation and Roundness feature is a descriptive characteristic of shape regularity. The malignant lesion is mostly irregular in shape compared with the benign lesion. In malignant lesions, the internal lesion structure shows a wide range of changes (heterogeneous attenuation) and invasion of adjacent structures. But in benign lesions, the internal structure is diffusely homogeneous. Therefore, these aspects also explain the choice of the homogeneity, Entropy, Gabor Energy,etc. as descriptive characteristics of the lesion. The GA approach reduces the feature size by 65.75% without compromising the accuracy of the classifier.

5.7 Classifiers

For the liver lesion classification/characterisation, the three of the most common classifiers (Support vector machine (SVM), Logistic Regression (LR) and Linear Discriminant Analysis (LDA)) are used to classify/characterise liver lesion. The LR is widely used for analysing the radiological observation (Saffari et al., 2015). Moreover, the

SVM has been widely used in the last few years for medical diagnoses. Furthermore, its capability to address problems by learning from samples and its ability to address large amounts of information simultaneously (Ulagamuthalvi et al., 2012; Jordan and Mitchell, 2015). In addition, the LDA is widely used in tissue classification stage especially with limited amount of data (Cheng et al., 2006). For classifier selection, the number of parameters to be used, especially for small-sized datasets, is critical to maximising the difference between the variances. However, for our small dataset size, these classifiers were chosen because they are controlled by fewer parameters and can be less susceptible to model violations. Hence, the use of SVM, LDA and LR are particularly worthwhile in this area.

5.7.1 Linear Discriminant Analysis (LDA)

LDA is a common classification method for predicting class membership of observation through finding a linear combination of features (variable) which best discriminate two or more classes (Ressom et al., 2008). LDA finds the optimal vector ν to discriminate the classes by maximising the ratio between class variance (Fukunaga, 2013). Consider a set of observation $\{b_1, b_2, ..., b_n\}$ where each observation belonging to the class κ. The mean μ and scatter matrix ζ for each classes is computed by the Equation 5.27 and Equation 5.28 respectively.

$$\mu_\kappa = \frac{1}{x_\kappa} \sum b_\kappa \tag{5.27}$$

$$\zeta = \sum_{i=1}^{\kappa} (b_i - \mu_i)(b_i - \mu_i)^T \tag{5.28}$$

Where x represents the number of the samples in κ. The difference between-class scatter is calculated using Equation 5.29

$$D_\zeta = \sum_{i=1}^{\kappa} (\mu_i - \mu)(\mu_i - \mu)^T \tag{5.29}$$

Where μ_i and μ are the mean of class i and entire classes respectively. The class separation in a direction of ν to maximise the distance between classes is given in Equation 5.30.

$$J(\nu) = \frac{\nu^T D_\zeta \nu}{\nu^T \zeta \nu} \tag{5.30}$$

5.7.2 Logistic Regression (LR)

LR is a discriminative machine learning classification approach used to predict the probability a class dependent variable (Hosmer Jr et al., 2013). LR is used to predict the class based on a S-shape (sigmoid) function ρ, such that $\rho(f_n) = L(c_n = 1 \,|f_n)$ for $c_n \in \{0, 1\}$ where $L(c_n = 1 \,|f_n)$ is equivalent to $1 - \rho(f_n)$ for the binary classification (Bewick et al., 2005). The sigmoid function is defined by the Equation 5.31.

$$\rho(f_n) = \frac{e^{Z(f_n)}}{1 + e^{Z(f_n)}} \tag{5.31}$$

Where $Z(f_n) = z_0 + z_1 f_{1n} + z_2 f_{2n} + \cdots + + z_x f_{xn}$, assuming x class observations and the log-likelihood function $K(z)$ of the data is defined by Equation 5.32.

$$\begin{aligned} K(z) &= \ln\left[L(f_n \,|c_n)\right] \\ &= \ln\left[\prod L(c_n = 1 \,|f_n)^{c_n} L(c_n = 0 \,|f_n)^{1-c_n}\right] \\ &= \sum_{n}^{N} \left[c_n \ln(\rho(f_n)) + (1 - c_n) \ln(1 - \rho(f_n))\right] \end{aligned} \tag{5.32}$$

The objective of LR is to select a parameter vector z with maximum likelihood $K(z)$ to estimate the probability of the class which called maximum likelihood estimator, and can be estimated through differentiating $K(z)$ with respect to z.

5.7.3 Support Vector Machine (SVM)

SVM is a supervised machine learning approach which has been successfully demonstrated for cancer classification and medical diagnosis,especially with the high dimensional feature spaces and relatively small sample size.(Chakraborty, 2011). In addition, the ability of the SVM to perform both linear and non-linear classification. The main idea of the SVM is find an optimal hyperplane by maximise the margin to separate the data into classes, as illustrated in Figure 5.26.

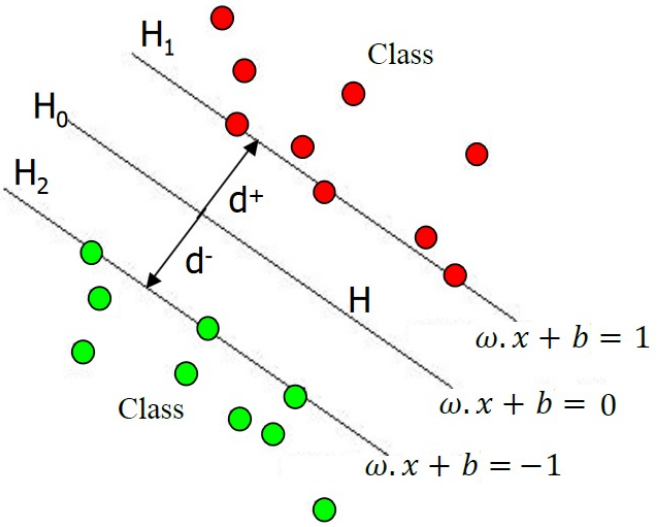


Figure 5.26: Example of two classes separated by a hyperplane H.

Regarding the Figure 5.26, the hyperplanes H_1 and H_2 is defined in Equation 5.33 where H_0 is the median in between H_1 and H_2. The margin (m) of a separating hyperplane is represented by $m = d^+ + d^-$ where d^+ and d^- is the shortest distance to the closest positive and negative point respectively, respect to H_0. The total distance between H_1 and H_2 is defined as $^2/_{\|w\|}$

$$\begin{aligned} H_1 &: w \bullet x_i + b \geq +1 \quad when \quad y_i = +1 \\ H_2 &: w \bullet x_i + b \leq -1 \quad when \quad y_i = -1 \end{aligned} \tag{5.33}$$

Where x is the input sample features, y is the output class, w is the weight vector and b represents a bias.

The performance evaluation of the proposed method for liver lesion characterisation and classification will discuss in the next Chapter 6. The experiments were performed using tenfold cross-validation and train/test spilt validation. However, The SVM classifier with Radial Basis Function kernel (SVM-RBF) was chosen because it provides the best results when compared to other classifiers that were tested. Furthermore, it can classify multi-dimensional data, unlike a linear kernel function and it has fewer parameters to set than a polynomial kernel. Two main RBF parameters applied in SVM are C and γ. Parameter C represents the cost of the penalty and parameter γ is the width of the kernel function.

5.8 Conclusion

This chapter presented the proposed thesis framework technical design details for liver lesion classification and characterisation. The framework was built to a better model of lesion classification and characterisation through linking between low-level features, high-level features and ROI. In addition, dealing with the challenge of the region of interest selection method that represent the lesion characteristics. Thus, the difference-of-features and multiple ROIs were developed for robust capturing of lesion characteristics in a reliable way. Furthermore, in contrast to the previous techniques that operate mainly over the lesion area with no pay attention to the relation between lesion and liver. The design of the liver lesion characterisation framework was inspired by an understanding of the radiologists' vision to characterise lesions as well as the utilising of prior knowledge of the medical background to support its robust performance.

This section presented a new technique for liver lesion characterisation, based on high-level features extracted automatically from the CT image, to simulate the clinician observation in describing liver lesions. In addition, the proposed framework presents three different methods to classify liver lesion. Liver lesion classification based on low-level features. Enhancing the lesion classification accuracy through utilising the

high-level features to classify the respective lesions, with the benefit of interpretable characterisation that supports the diagnostic decision. These results will be presented in the next chapter. The combination between low-level features and characterisation were used to build an extended feature vector. In summary, the chapter contributions are:

- Developing an automated technique for liver lesion characterisation in order to predict radiological observation in describing the liver lesions through using low-level features extracted from computed tomography (CT) images to infer higher level features, and simulate radiological observations for liver characterisation. In addition, overcoming the challenge of linking the image content through converting the low-level features to visual semantics by lesion characterisation. Thus, Assigning high-level descriptions to the liver lesion in analogy to radiologist observation.

- Proposing a novel Multiple ROIs for liver lesion classification/characterisation. The proposed method is based on medical knowledge and classifies the region of interest into three areas (inside lesion, lesion border and surrounding lesion) through constructing a multi-level abnormality map based on the intensity difference with respect to the normal liver. In addition, the asymmetry and compactness features are computed to define the probability of each level to represent a lesion. Thus, three regions of interest are defined and known as Multiple image ROIs. The idea behind the multiple image ROIs is to capture all the lesion appearance characteristics by considering the ability of each ROI that represents a set of lesion characteristics. This is in contrast with most existing research, which mainly relies on lesion area without considering the effect of the lesion on the surrounding area, where the performance of classification/ characterisation could be affected due to the selection of ROI methods, which represents the characteristics of the lesion.

- Proposing a difference-of-feature (DoF) technique to enhance the classification/ characterisation performance. The idea of the DoF is identifying the difference between the lesion and the surrounding normal liver tissues. The DoF emphasises the relative difference of the lesion properties, in relation to surrounding tissues of the same patient, regardless of the demographics or imaging device. Then, the two proposed techniques (DoF and Multiple ROIs) are combined towards a better and robust classification/characterisation results.

- Building a new classification framework for classifying liver lesion in three different novel ways, as follows:

    1. Classifying lesion based on low-level features that are extracted from a Multiple ROIs (inside, border and outside lesion). The multiple ROIs fused

with difference-of-features between the lesion itself and its surrounding area. This is in contrast with most existing research, which focus on extracted features from the segmented lesion only. The Multiple ROIs and difference-of-features emphasise the relative difference of the lesion properties, in relation to the surrounding tissues of the same patient. It also captures the different property/behaviour between benign and malignant lesions and how they affect the adjacent tissues.

2. Utilising the high-level features that characterised the lesion to build a novel feature vector. The new feature vector is used to classify the respective lesions, with the benefit of interpretable characterisation that supports the diagnostic decision in analogy to radiologist observation, which is in contrast with the existing works that used the black box low-level features that cannot provide an explanation in human level understanding for the diagnostic decision. However, the classification based on high-level features is more reliable for the radiologist, which provides the understanding explanation for the diagnostic decision in analogy to radiologist observation.
3. Building a novel feature vector that composed of the combination of low-level features and high-level features. The new feature vector is used to enhance the classification accuracy. In contrast with the existing works that used only low-level features. However, the classification through the fusion between high-level and low-level features will lose the advantages of the characterisation and track back to diagnosis explanation.

The next chapter will further discuss the liver lesion classification/ characterisation performance with more evaluation across the dataset. Furthermore, the proposed framework will be benchmarked against a number of state-of-art baselines in the next chapter as well.

Chapter 6

Results and evaluation

This chapter is focusing on presenting and discussing of the evaluation and results of the proposed framework and its various components in terms of their classify and characterise liver lesions. In addition, a comparative evaluation of classification accuracy, between the characterisation approach and low-level features, is presented as well as benchmarking of the entire framework against the related work baselines.

The structure of the chapter is as follows: The experimental setup and validation methods that used to evaluate the system performance is presented in Section 6.1 and Section 6.2 respectively. Section 6.3 and Section 6.4 evaluates the liver lesion characterisation and classification performance of the proposed framework. Section 6.5 presents the benchmarking of the proposed framework against a number of state-of-art baselines. Finally, the chapter is concluded in Section 6.6.

6.1 Experimental Setup

All the experiments were performed on on Intel Core I5- 3.40 GHz computer with 8 Gigabytes of RAM under windows 7 64-bit operating system. The Matlab R2015b was used to run the experiments and extract the features to achieve the thesis goal of liver lesion classification and characterisation.

6.2 Validation

In order to evaluate the performance and validate the classification/ characterisation models, a number of metrics were adopted, as previously discussed in Chapter 2. The evaluation process was performed by using a two different approaches, as follows:

- **K-fold Cross Validation** is the most widely used model validation method to assess the performance of a classifier (Geisser, 1993; Kohavi et al., 1995). Firstly, the original sample data is randomly partitioned into k equal sized sub-samples, where the $k-1$ samples are selected to training, and the rest samples perform

the validation alternately, as presented in Figure 6.1. The validation process is repeated k times (folds), with each fold k is used just once as the validation sample. The advantage of this approach is that all the samples in the dataset are used for both training and testing, and each fold is used for validation only once.

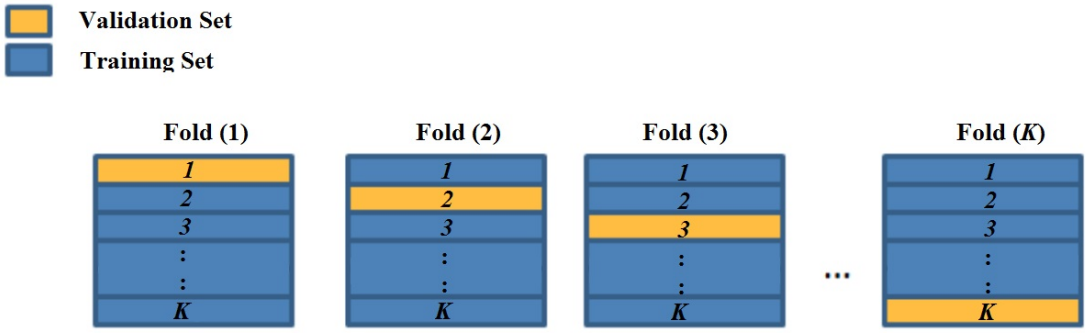


Figure 6.1: Visual representation of K-fold-cross-validation.

- **Train/Test Split Validation** is usually split the dataset into training data and test data. In our studies, the overall dataset comprises of two datasets, collected from two different institutions. The first dataset (Dataset I) was chosen to validate the proposed system, and the second dataset (Dataset II) was used for training the classifier.

6.3 Evaluation of Lesion Characterisation

This section evaluates the liver lesion characterisation performance of the proposed framework. The evaluation is done based on a number of different metrics are presented in previous Chapter 2. Figure 6.2 depicts the experiment model of liver lesion characterisation.

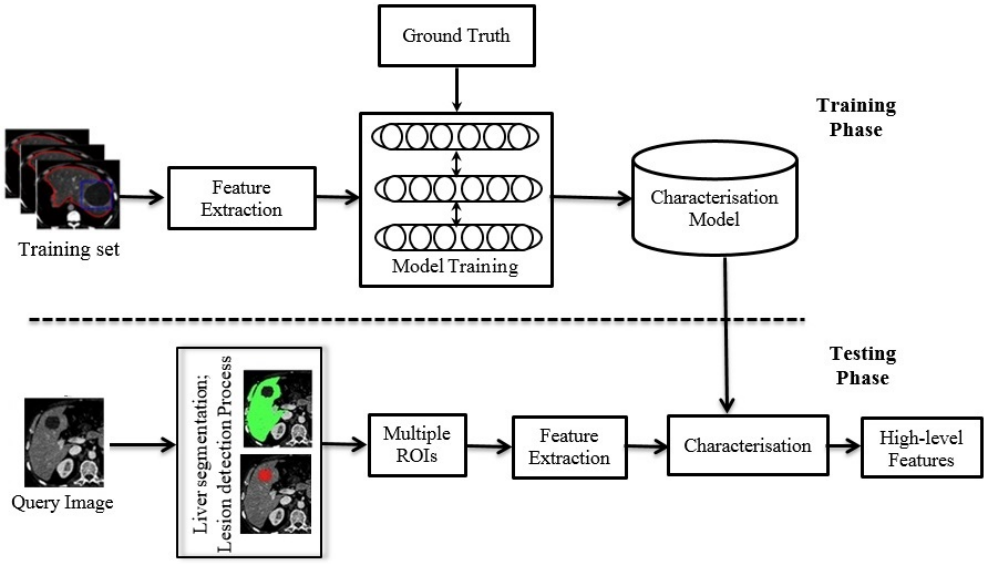


Figure 6.2: Experiment model of liver lesion characterisation.

The lesion characterisation framework includes two strategies for characterising lesion: (1) automatic lesion characterisation based on the image itself where the high-level features are estimated directly from the query image, as presented in Section 5.3.2. (2) automatic lesion characterisation using a machine learning algorithm to annotate the lesion, as discussed in Section 5.3.3.

In this thesis, we proposed a fully automatic liver lesion classification/ characterisation, which do not require any user interaction in any of the steps and that all parameters are fixed for all the images processed in beforehand. The proposed framework is designed according to certain assumptions based on medical knowledge, as discussed in Chapter 4 for liver segmentation and Chapter 5 for vessels classification and Multiple ROIs selection. For example, the liver intensity parameter is applied on CT image for a rough segmentation. This parameter is estimated according to the medical knowledge-based assumption, that the liver intensity is in the range of $[-50, 250]$ HU.

Visual Features From The Image Contents

This section evaluates the proposed framework performance for extracted the high-level features from the query image such as lesion location. The characterisation performance mainly depends on the extracted lesion and vessels, as discussed in Chapter 5. The liver anatomy mainly relies on the portal vein and the hepatic vein where these vessels are enhanced at the portal phase (Reitinger et al., 2006; Mule et al., 2015; Rajesh et al., 2015). All the experiments were done in a portal phase.

Each liver lesion was characterised by seven high-level features that extracted directly from the CT image contents. Figure 6.3 illustrates the statistics of prediction accuracy for all seven high-level features across the dataset I, dataset II and the combination of Dataset I and II. The prediction accuracy of high-level features varies over the dataset used for evaluation. This due to all seven high-level features were not equally used through the datasets since the number of cases is not equally distributed among each characteristic class.

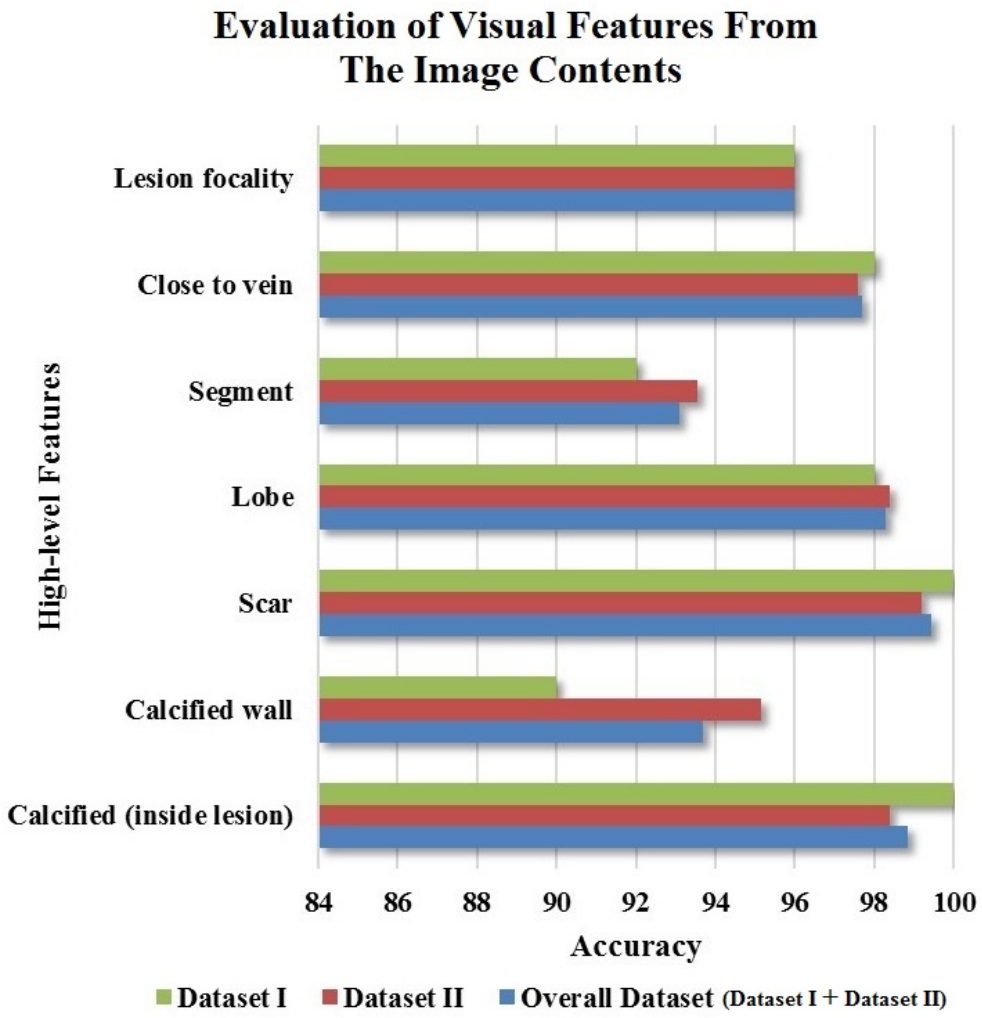
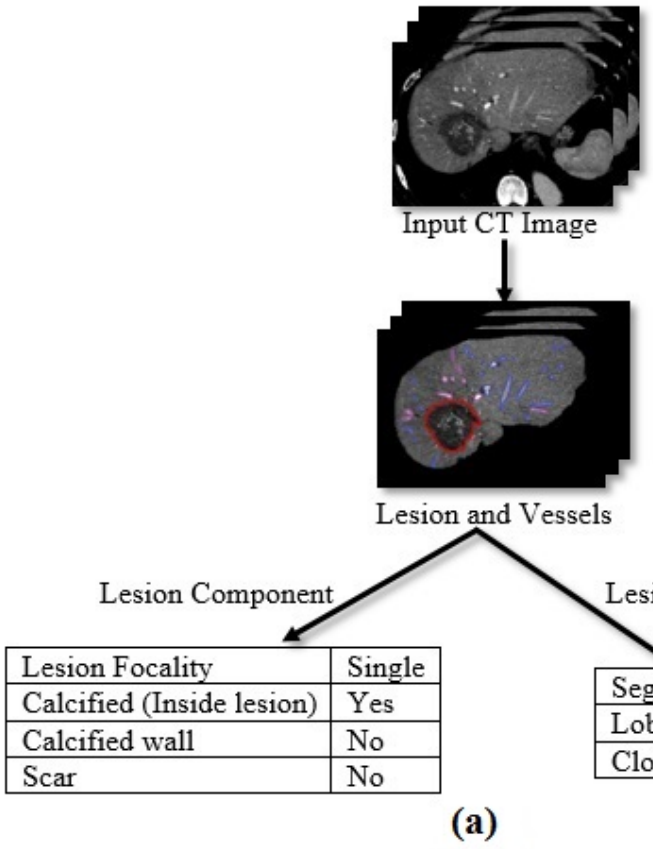


Figure 6.3: Evaluation results of lesion characterisation based on the image across the datasets.

The evaluation results of proposed approach to characterise liver lesion based on the high-level features directly extracted from the query image shown in Figure 6.3. The experiment was performed on Dataset I, Dataset II and overall dataset (Dataset I + Dataset II). The proposed framework mainly extracted high-level features from the query image based on existing medical knowledge in practice that supports lesion characterisation for better overall performance, as discussed in Section 5.3.2. The average performance of lesion characterisation for the seven high-level features that extracted from the image contents was 96.7%.

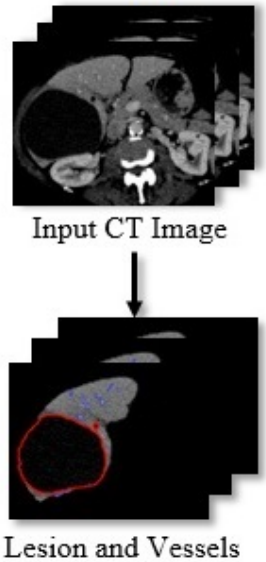
Figure 6.4 depicts samples of liver lesion characterisation results that characterise lesion location and component from the query CT image. Figure 6.4.a Sample of correct lesion characterisation. Figure 6.4.b sample of incorrect lesion characterisation of close to vein property. Only the "Right Portal Vein" was predicted wrongly as "Middle Hepatic Vein". For this high-level feature, This error is due to the effect the lesion on the vessels which lead to the loss of some vessels structure.



Lesion Focality	Single
Calcified (Inside lesion)	Yes
Calcified wall	No
Scar	No

Segments	V, VII, VIII
Lobe	Right Lobe
Close to vein	Right Portal Vein

(a)



Lesion Focality	Multiple
Calcified (Inside lesion)	No
Calcified wall	No
Scar	No

Segments	V, VI, VII
Lobe	Right Lobe
Close to vein	Middle Hepatic Vein

(b)

Figure 6.4: Sample of liver lesion characterisation results that extracted directly from query CT image; (a) Sample of correct lesion characterisation; (b) Sample of incorrect close to vein annotation.

High-level Features based on Machine Learning

In this section, the liver lesion characterisation based on learning model is evaluated through the accurate prediction of the high-level features. The accuracy is defined as the number of correct prediction for each high-level feature divided by the total number of the cases.

The Support vector machine (SVM) classifier was used to characterise liver lesion. The SVM has been widely used in the last few years for medical diagnoses. Furthermore, its capability to address problems by learning from samples and its ability to address large amounts of information simultaneously (Ulagamuthalvi et al., 2012; Jordan and Mitchell, 2015). Hence, the use of SVM is particularly worthwhile in this area.

Two validation approaches were performed to evaluate the proposed framework performance. Namely; tenfold cross-validation and train/test spilt validation. In addition, the evaluation has been done by using portal phase only and multiphase CT scan protocol (Arterial phase, Portal phase and Delayed phase). Figure 6.5 depicts the statistics of tenfold cross-validation prediction accuracy based on portal phase for all the high-level features that extracted through learning process, as discussed in Section 5.3.3. The overall dataset is randomly partitioned into ten subgroups for training and testing. For each iteration, one subgroup is left out to test the proposed framework performance. Furthermore, the liver lesion characterisation performance of our proposed framework (Multiple ROIs) is evaluated/compared with the traditional way that only relied on the lesion ROI (single ROI) to predict the lesion characterisation.

In general, the single ROI have a limited characterisation performance over the relation between the lesion and surrounding area. This is clearly reflected in the lesion surrounding, lesion rim and rim thickness high-level features where the maximum reached accuracy is $< 45\%$. This is attributed to the lesion ROI alone not represented the relationship between lesion and liver. However, The proposed multiple ROIs enhanced the characterisation framework performance for each high-level features categories, which are as follows: (1) the relationship between lesion and liver. (2) lesion margin high-level features. (3) and the lesion structure. The best overall performance of lesion characterisation based on SVM classifier were $78.28 \pm 4.63\%$ and $95.56 \pm 1.25\%$ for single ROI and Multiple ROIs respectively, compared to LR classifier was $77.22 \pm 3.71\%$ and $94.7 \pm 2.01\%$. While the accuracy achieved with LDA classifier was $74.71 \pm 3.23\%$ and $92.41 \pm 2.45\%$. However, the SVM was adopted in this section, due to the SVM achieved the best accuracy. The full results of LR and LDA classifiers are available at Appendix B.

Furthermore, the results shows that the Multiple ROIs technique is more contributive to the overall lesion characterisation. The proposed multiple ROIs records average improvement of $12.64 \pm 3.45\%$ and $55.17 \pm 5.31\%$ for characterising the internal lesion

structure and the relation between lesion and liver respectively. This is due to utilising the features of surrounding lesion from the liver tissue to interpreting the effect the lesion on the liver rather than relying on the lesion characteristics only. In addition, selecting the internal lesion is more descriptive to the lesion structure, due to reducing the false positive of the lesion segmentation and focusing on the internal structure.

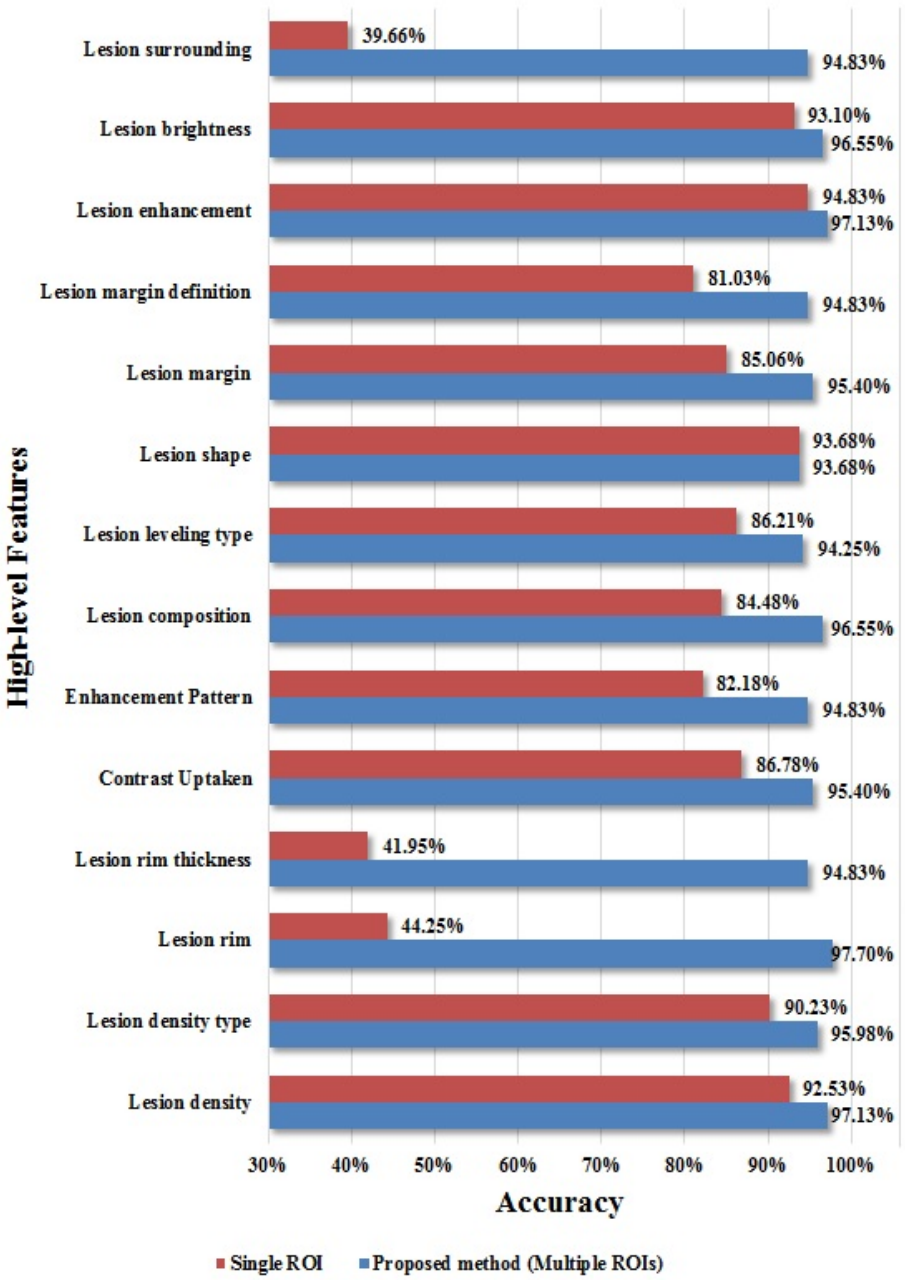


Figure 6.5: The accuracy comparison between single ROI and Multiple ROIs to predict the high-level features using the **portal phase** CT image where **tenfold cross-validation** method and SVM classifier was adopted.

Figure 6.6 shows both selection ROI methods (Single ROI and Multiple ROIs) evaluation based on multiphase CT images.A tenfold cross-validation method was performed where 90% of the dataset was used to train the classifier, while the remaining 10% was used to measure the prediction accuracy.

Both selection ROI methods (single ROI and Multiple ROIs) achieved slightly better performance in multiphase CT compared to the single phase and scored average accuracy improvement of $1.36 \pm 0.6\%$ and $0.33 \pm 0.2\%$ respectively, as shown in Figure 6.6.

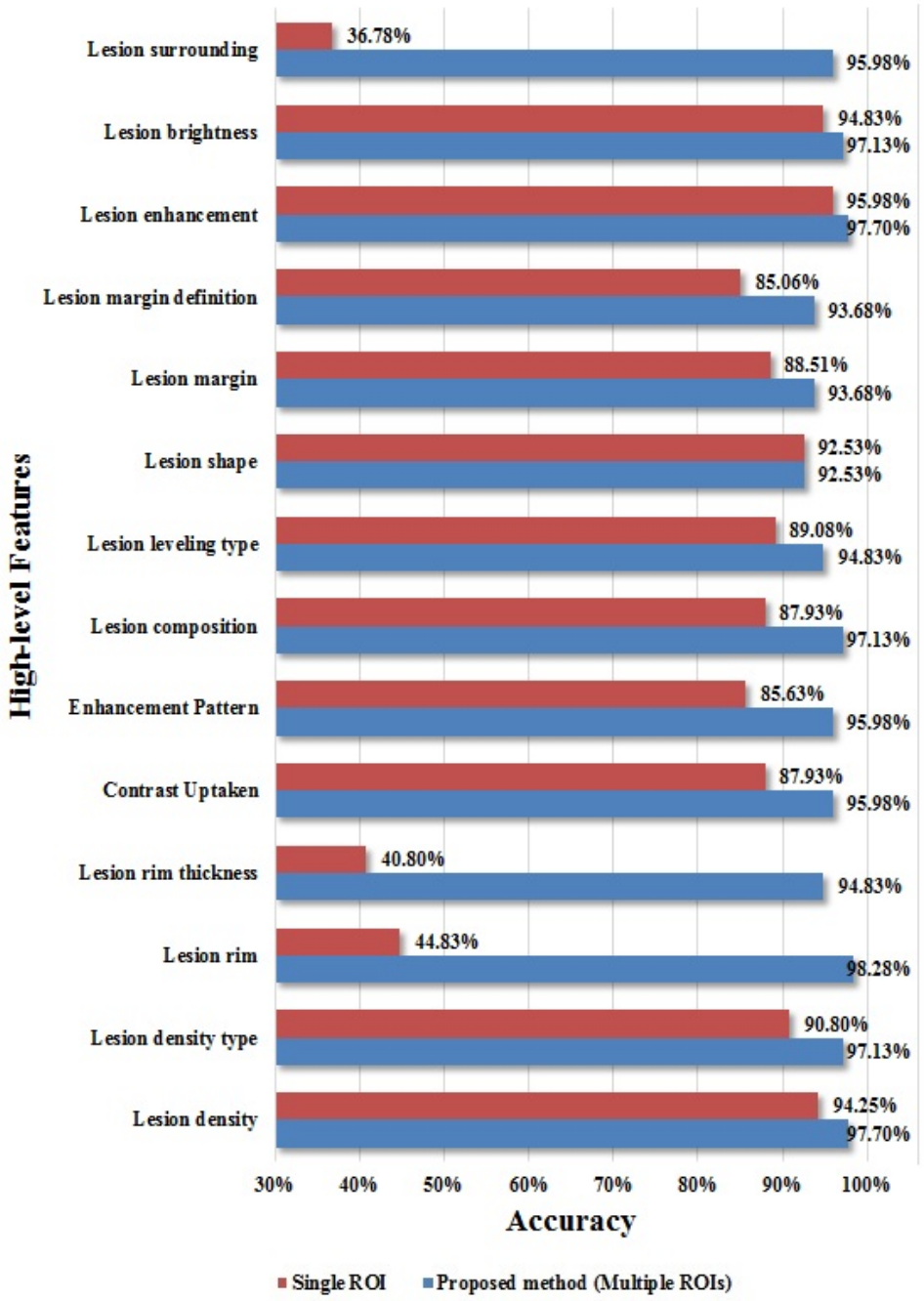


Figure 6.6: The accuracy comparison between single ROI and Multiple ROIs to predict the high-level features using the **multiphase** CT image where **tenfold cross-validation** method and SVM classifier was adopted.

Figure 6.7 depicts the evaluation of proposed liver lesion characterisation framework based on Multiple ROIs comparison by using portal phase and multiphase CT images. The SVM classifier was used to predict the high-level features and the accuracy of prediction was estimated using a ten-cross validation.

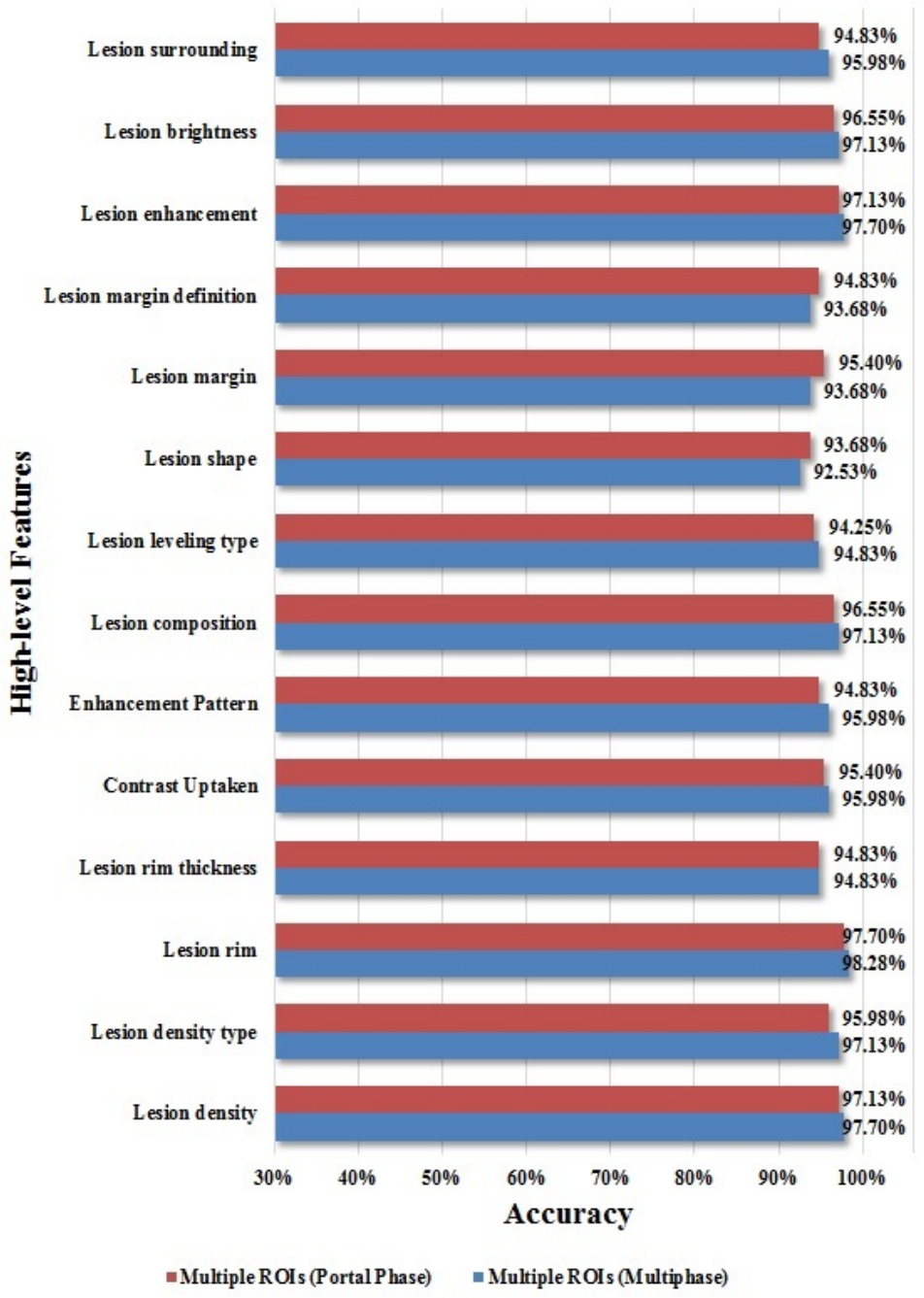


Figure 6.7: The accuracy comparison of the proposed Multiple ROIs to predict the high-level features by using **portal phase** and **multiphase** CT image where **tenfold cross-validation** method and SVM classifier was adopted.

The results are depicted in Figure 6.7, and shows that the performance of proposed framework in portal phase to characterised lesion shape and margin is better by $1.3 \pm 0.4\%$ on average compared to multiphase CT image. This is related to the lesion being affected by the rate of absorption and washout of the contrast agent during the time of screening in respect of the lesion size. On the other side, the multiple phase CT is doing better to characterise the internal structure of the lesion compared to the single phase with average improvement of $3.4 \pm 0.7\%$ for single ROI and $1.1 \pm 0.3\%$ for multiple ROIs. This can be explained as multiphase provides more useful information

about enhancement pattern variation over the phase.

To verify and confirm the performance of the proposed framework, another validation approach was implemented by splitting the dataset into training and testing sets. Figure 6.8 depicts the statistics of prediction lesion characterisation accuracy using portal phase CT image. The SVM classifier has been considered for characterisation process. the experiment has been done by using Dataset II for training the classifier and Dataset I for testing the proposed framework.

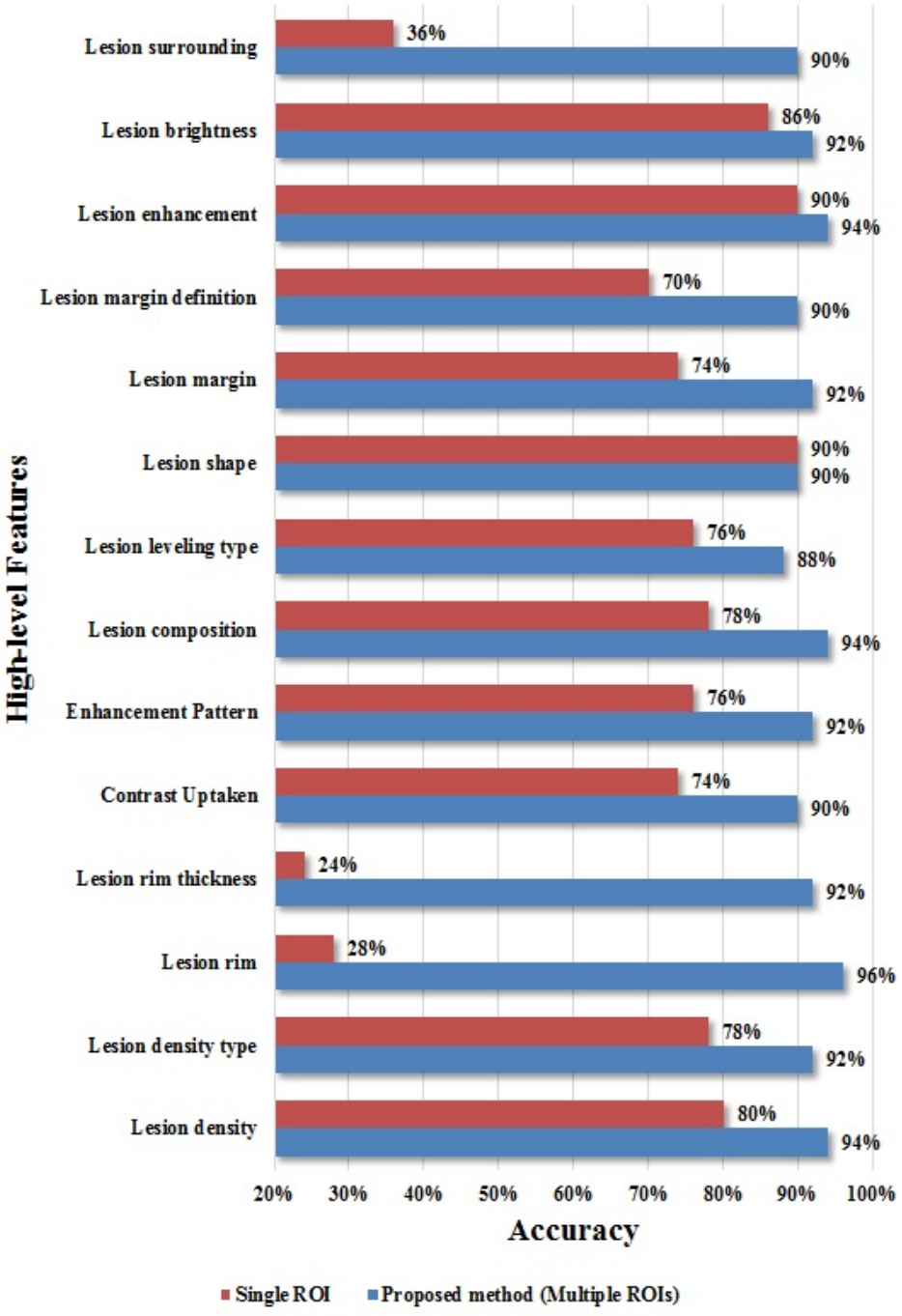


Figure 6.8: The accuracy comparison between single ROI and Multiple ROIs to predict the high-level features using the **portal phase** CT image where **Training/Testing split validation** method and SVM classifier was adopted.

Figure 6.9 shows the accuracy evaluation of prediction lesion characterisation using multiphase CT image. The dataset II was used to train the SVM classifier, while the dataset I was used to evaluate the framework prediction accuracy.

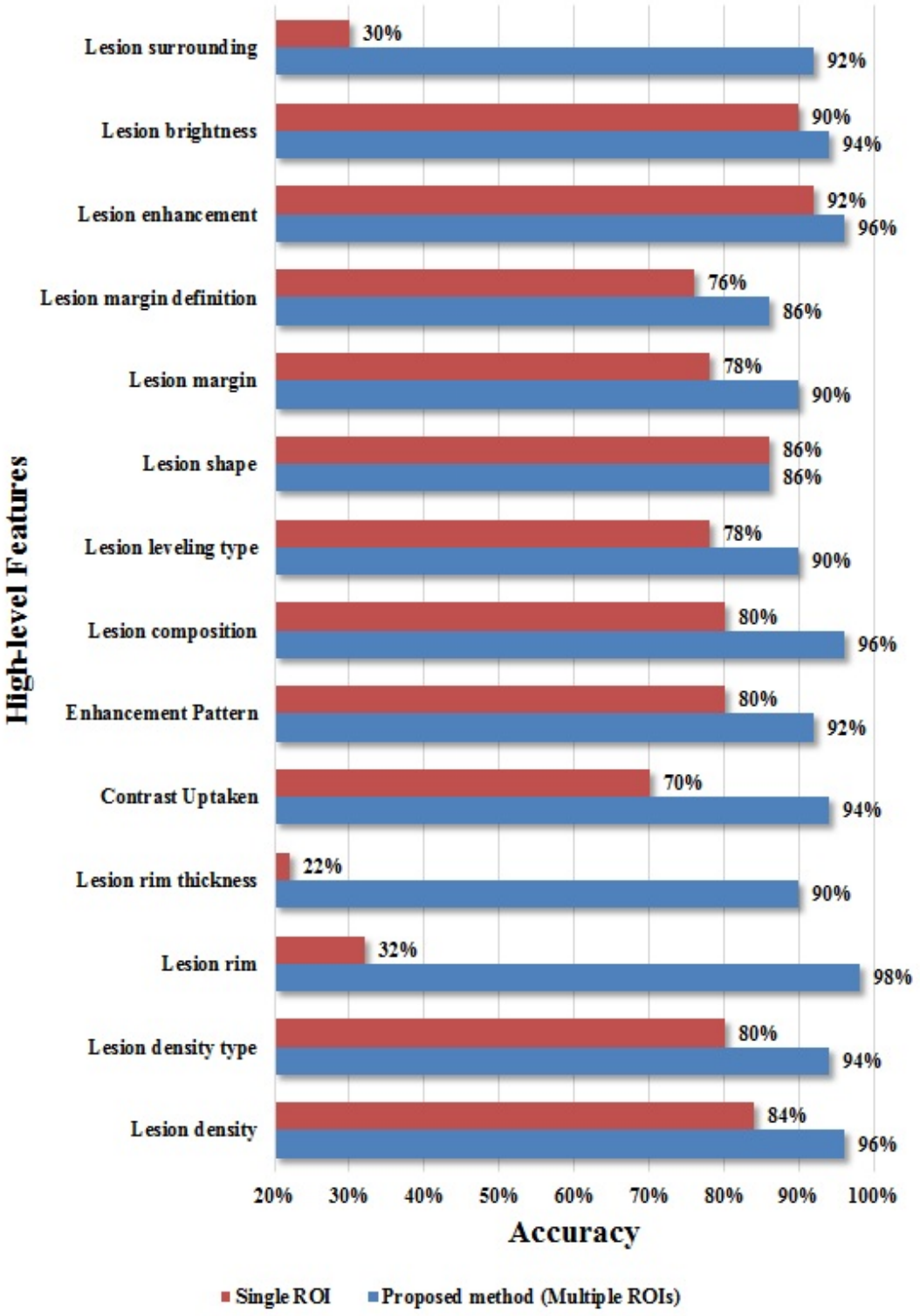


Figure 6.9: The accuracy comparison between single ROI and Multiple ROIs to predict the high-level features using **multiphase** CT image where **Training/Testing split validation** method and SVM classifier was adopted.

Figure 6.10 depicts the comparison between the multiple ROIs characterisation framework based on portal phase CT image and multiple ROIs characterisation framework using multiphase CT image. The SVM classifier was trained on dataset II and dataset I was used to evaluate the framework performance.

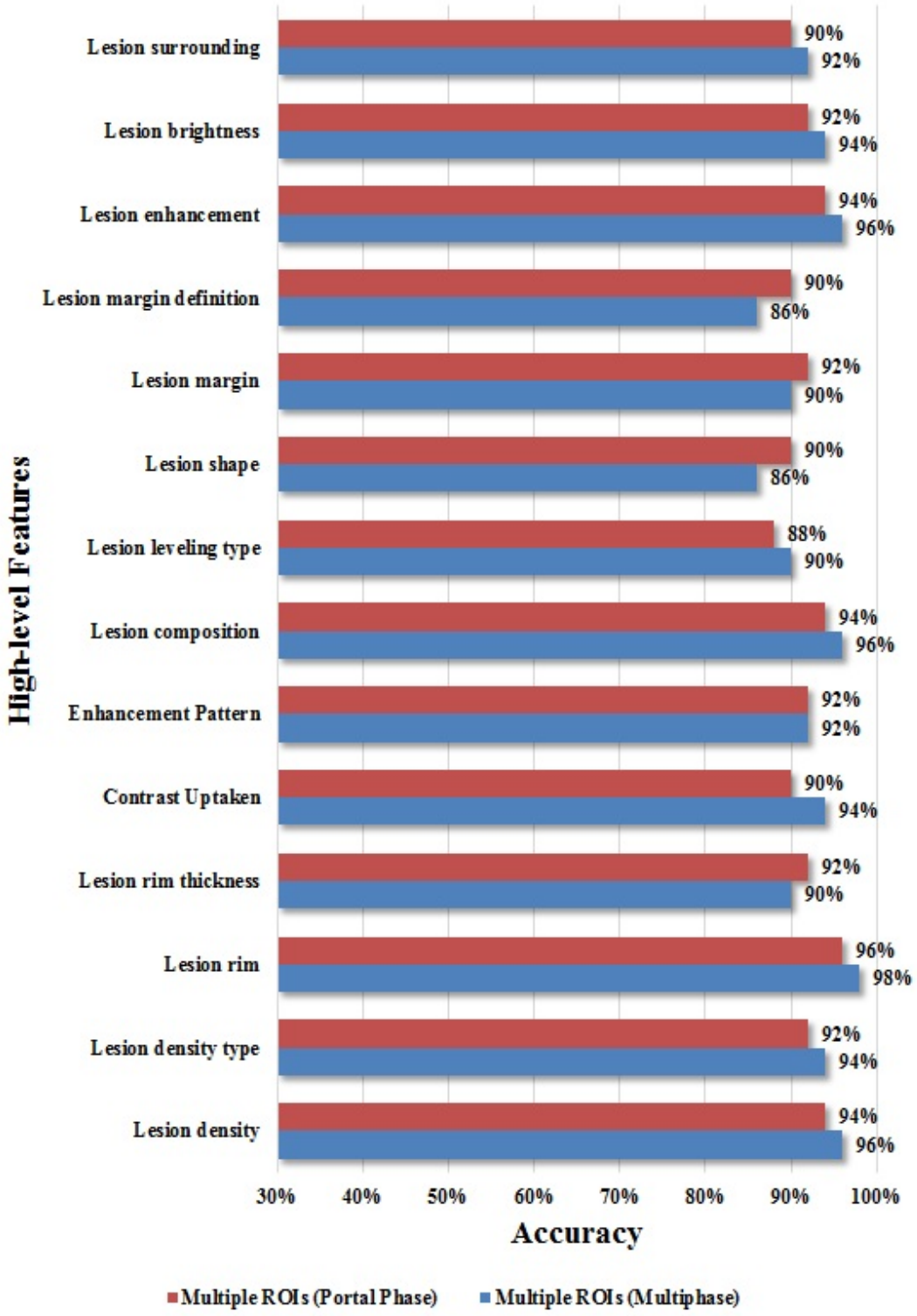


Figure 6.10: The accuracy comparison of the proposed Multiple ROIs to predict the high-level features by using **portal phase** and **multiphase** CT image where **Training/Testing split validation** method and SVM classifier was adopted.

6.4 Evaluation of Lesion Classification

This section examines the liver lesion classification performance of the proposed framework to classify liver lesions either benign or malignant. The classification accuracy (ACC), specificity (SPE), sensitivity (SEN), negative predictive value (NPV) and positive predictive value (PPV) are used in the system evaluation.

The liver lesion classification framework includes three different ways for classifying lesions: (1) automatic liver lesion classification based on low-level features (black box). (2) automatic lesion classification through high-level features. (3) lesion classification based on the feature fusion between low-level and high-level features.

For the classification task, the Support vector machine (SVM) classifier was adopted, for being the most commonly used in lesions classification. Moreover, the SVM is a well-established and very popular classifier. In addition, The SVM is less prone to over-fitting compared to others algorithms such as back-propagation neural networks.

Classification based on Low-level Features

The performance of the proposed liver lesion classification framework based on low-level features is evaluated by using three different configurations, as discussed in Section 5.2. The first way, the lesion is classified using difference-of-features (DoF) between segmented lesion and surrounding tissues. The second way, the classification task is carried out by fusing features of Multiple ROIs (internal, border and surrounding lesion). The last configuration, the classification performed through fusing lesion border features with the difference-of-features between internal and surrounding lesion. The average classification performance based on Lesion ROI, DoF, Multiple ROIs and combined Multiple ROIs and DOF technique by LR classifier were 83.33%, 89.08%, 91.38% and 93.1% respectively. While the classification accuracy achieved by LDR classifier were 81.61%, 87.93%, 89.66% and 90.23%. However, the highest accuracy was achieved by SVM as will be presented in detail in Table 6.1. Therefore, the SVM was adopted in this section. The results of LR and LDA are displayed in full at Appendix B.

Table 6.1 presents the comparison between all of the proposed configurations to classify liver lesion from portal phase CT image. The SVM classifier performance was evaluated by adopting tenfold cross-validation method.

In general, the evaluation results depicts that the combination of multiple ROIs and difference-of-features is more contributive to the overall lesion classification performance in terms of the evaluation metrics. The proposed of three different methods (DoF, Multiple ROIs and combination of Multiple ROIs and DoF) displayed the better classification accuracy over the single ROI (lesion area only) with increasing the

overall system accuracy is $7.5 \pm 1.8\%$, $8.7 \pm 2.1\%$ and $12.1 \pm 1.3\%$ respectively, as presented in Figure 6.11. This is due to the classification accuracy is affected by several factors such as the ROI selection that represent the lesion features and imaging devices/settings. However, All the lesion characteristics such as lesion its self, boundary and relation with the surrounding area were captured through our proposed multiple ROIs. Moreover, the Difference-of-Features helps to overcome the challenge issues such as the variation of intensity and texture ranges between study cases due to the imaging devices settings such as images resolution and spacing.

		SN	SP	ACC	PPV	NPV	Average ACC
Lesion ROI	Malignant	0.788	0.872	0.788	0.840	0.828	0.833
	Benign	0.872	0.788	0.872	0.828	0.840	
DoF	Malignant	0.90	0.915	0.90	0.90	0.915	0.908
	Benign	0.915	0.90	0.915	0.915	0.90	
Multiple ROIs	Malignant	0.913	0.926	0.913	0.913	0.926	0.92
	Benign	0.926	0.913	0.926	0.926	0.913	
Multiple ROIs + DoF	Malignant	0.938	0.968	0.938	0.962	0.948	0.954
	Benign	0.968	0.938	0.968	0.948	0.962	

Table 6.1: Summary of lesion classification through low-level features results obtained by **tenfold cross-validation** and SVM classifier using **portal phase** CT images. The combination of Multiple ROIs and DoF is achieving higher accuracy levels compared to the other methods.

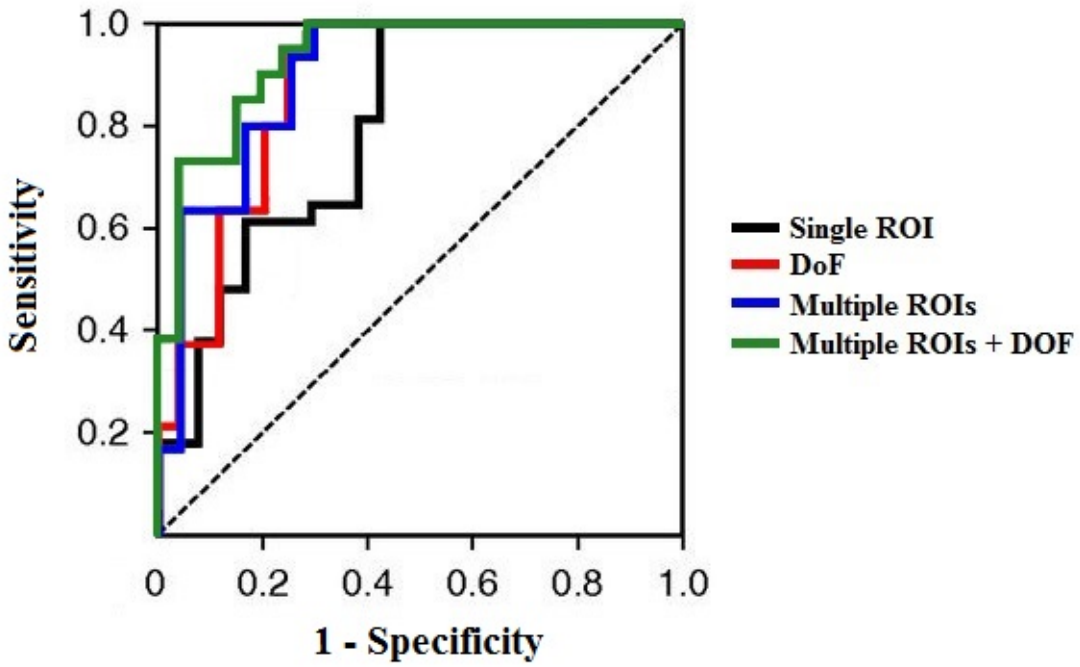


Figure 6.11: ROC-curves of different lesion classification methods using **portal phase** CT images where **tenfold cross-validation** method and SVM classifier was adopted. The combination of Multiple ROIs and DoF achieves higher accuracy over all the other methods.

The Table 6.2 and Figure 6.12 shows enhancing on overall classification performance for the proposed combination between multiple ROIs and DoF when using multiphase CT images by $1.7 \pm 0.8\%$, malignant classification accuracy by $2.5 \pm 1.1\%$ and benign classification accuracy by $1.1 \pm 0.5\%$ as compared to the single phase (portal phase) representation, due to provide more represented information, in particular on hypervascular lesions.

		SN	SP	ACC	PPV	NPV	Average ACC
Lesion ROI	Malignant	0.838	0.883	0.838	0.859	0.865	0.862
	Benign	0.883	0.838	0.883	0.865	0.859	
DoF	Malignant	0.925	0.926	0.925	0.914	0.935	0.925
	Benign	0.926	0.925	0.926	0.935	0.914	
Multiple ROIs	Malignant	0.925	0.957	0.925	0.949	0.938	0.943
	Benign	0.957	0.925	0.957	0.938	0.949	
Multiple ROIs + DoF	Malignant	0.963	0.979	0.963	0.975	0.968	0.971
	Benign	0.979	0.963	0.979	0.968	0.975	

Table 6.2: Summary of lesion classification results through low-level features obtained by **tenfold cross-validation** and SVM classifier using **multiphase** CT images. The combination of Multiple ROIs and DoF is achieving higher accuracy levels compared to the other methods.

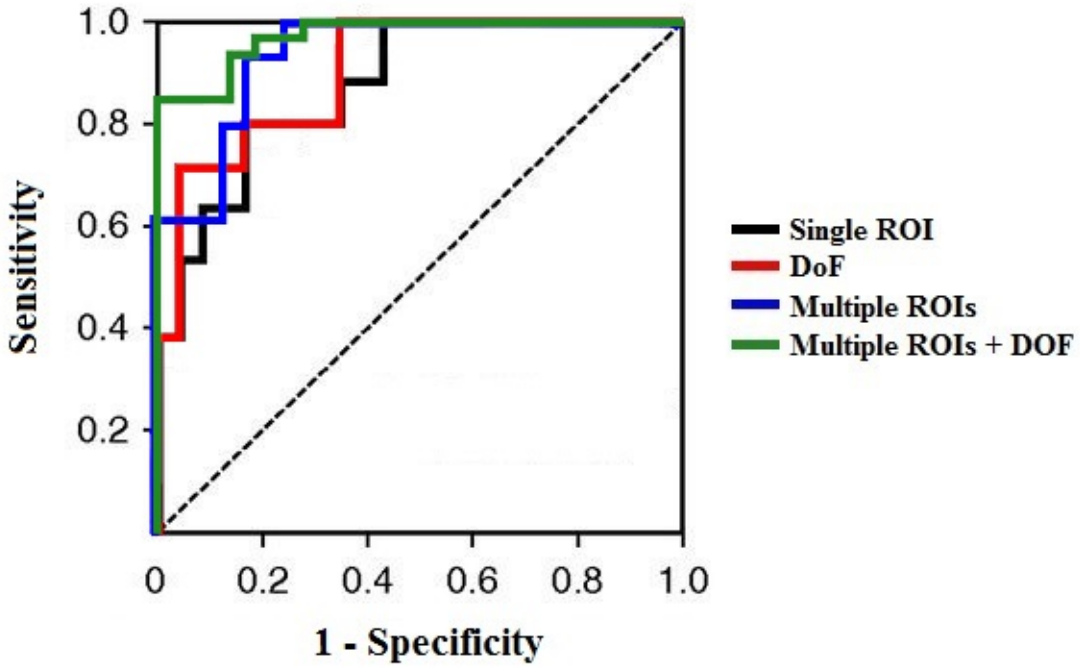


Figure 6.12: ROC-curves of different lesion classification methods using **multiphase** CT images where **tenfold cross-validation method** and SVM classifier was adopted. The combination of Multiple ROIs and DoF achieves higher accuracy over all the other methods.

Table 6.3 and Table 6.4 depicts the extended evaluation of the framework performance through using split Train/Test validation method.

		SN	SP	ACC	PPV	NPV	Average ACC
Lesion ROI	Malignant	0.690	0.714	0.69	0.769	0.625	0.70
	Benign	0.714	0.690	0.714	0.625	0.769	
DoF	Malignant	0.793	0.810	0.793	0.852	0.739	0.80
	Benign	0.810	0.793	0.810	0.739	0.852	
Multiple ROIs	Malignant	0.828	0.810	0.828	0.857	0.773	0.82
	Benign	0.810	0.828	0.810	0.773	0.857	
Multiple ROIs + DoF	Malignant	0.862	0.905	0.862	0.926	0.826	0.88
	Benign	0.905	0.862	0.905	0.826	0.926	

Table 6.3: Summary of lesion classification through low-level features results obtained by **Training/Testing split validation** and SVM classifier using **portal phase** CT images. The combination of Multiple ROIs and DoF is achieving higher accuracy levels compared to the other methods.

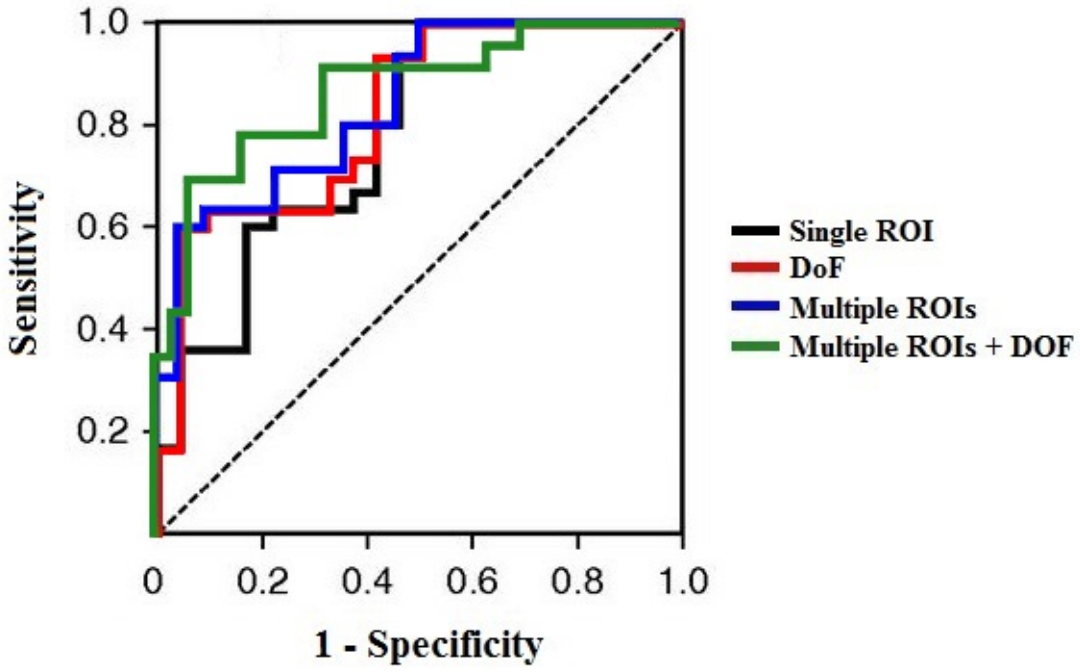


Figure 6.13: ROC-curves of different lesion classification methods using **portal phase** CT images where **Training/Testing split validation** method and SVM classifier was adopted. The combination of Multiple ROIs and DoF achieves higher accuracy over all the other methods.

		SN	SP	ACC	PPV	NPV	Average ACC
Lesion ROI	Malignant	0.759	0.762	0.759	0.815	0.696	0.76
	Benign	0.762	0.759	0.762	0.696	0.815	
DoF	Malignant	0.828	0.810	0.828	0.857	0.773	0.82
	Benign	0.810	0.828	0.810	0.773	0.857	
Multiple ROIs	Malignant	0.862	0.857	0.862	0.893	0.818	0.86
	Benign	0.857	0.862	0.857	0.818	0.893	
Multiple ROIs + DoF	Malignant	0.897	0.952	0.897	0.963	0.870	0.92
	Benign	0.952	0.897	0.952	0.870	0.963	

Table 6.4: Summary of lesion classification through low-level features results obtained by **Training/Testing split validation** and SVM classifier using **multiphase** CT images. The combination of Multiple ROIs and DoF is achieving higher accuracy levels compared to the other methods.

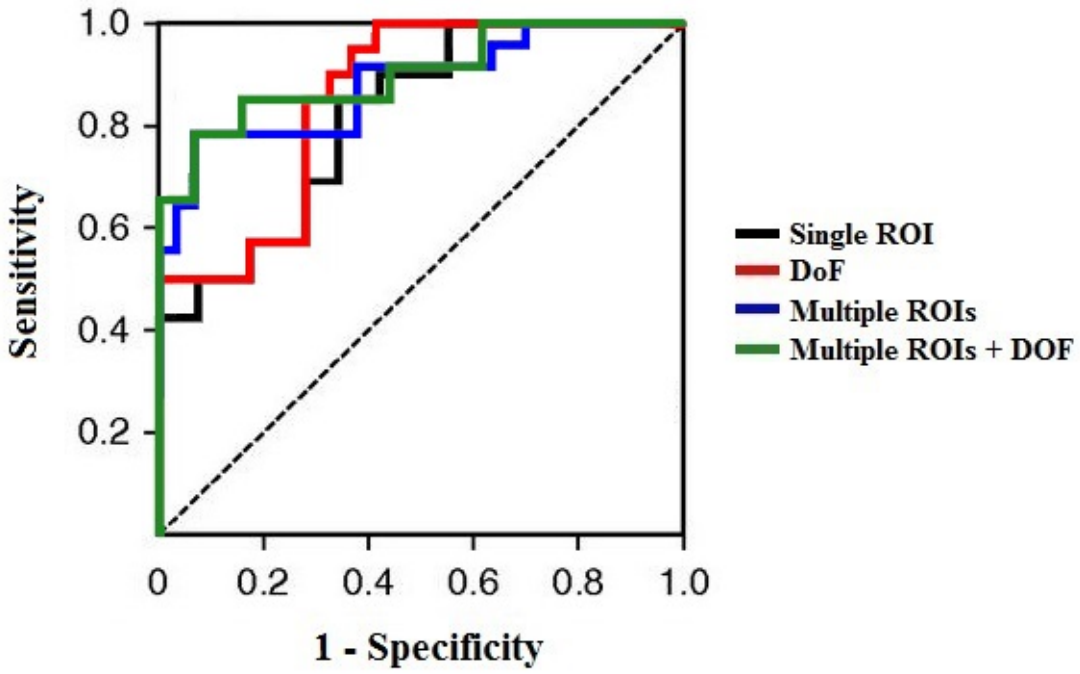


Figure 6.14: ROC-curves of different lesion classification methods using **multiphase** CT images where **Training/Testing split validation** method and SVM classifier was adopted. The combination of Multiple ROIs and DoF achieves higher accuracy over all the other methods.

In summary: In this section, we proposed three methods to classify liver lesion based on low-level features, Namely: (1) Difference-of-Feature approach (DoF). (2) Multiple ROIs. and (3) combination of multiple ROIs and DoF. These approaches were compared to the traditional way, which focuses on extracted features from the segmented lesion only. For more confident results, the McNemar test (McNemar, 1947) was used in this study to benchmark all approaches. Table 6.5 depicts the McNemar analysis p-value results in comparison of the proposed approaches based on portal phase CT images.

Lesion ROI	**Lesion ROI**		
DoF	**0.025858**	**DoF**	
Multiple ROIs	**0.009330**	0.479500	**Multiple ROIs**
Multiple ROIs + DoF	**0.000318**	**0.013328**	**0.041227**

Table 6.5: McNemar analysis p-value results in comparison all proposed approaches to classify liver lesion based on **portal phase** CT images.The p-value < 0.05 scores are in bold font.

The results show that all the proposed approaches are better than traditional ROI statistically. However, the Multiple ROIs are doing better than difference-of-features based on the classification accuracy but is not statistically significant, where the combination between Multiple ROIs and difference-of-features enhanced the accuracy results and statically significant compared to all other approaches (p-value < 0.05). Table 6.6 depicts the McNemar analysis p-value results in comparison of the proposed approaches based on multiphase CT images.

Lesion ROI	**Lesion ROI**		
DoF	**0.037056**	**DoF**	
Multiple ROIs	**0.010787**	0.546494	**Multiple ROIs**
Multiple ROIs + DoF	**0.000086**	**0.026857**	0.073638

Table 6.6: McNemar analysis p-value results in comparison all proposed approaches to classify liver lesion based on **multiphase** CT images.The p-value < 0.05 scores are in bold font.

The results based on multiphase CT images show that all the proposed approaches are better than traditional ROI statistically. However, the Multiple ROIs are doing better than difference-of-features based on the classification accuracy but is not statistically significant, where the combination between Multiple ROIs and difference-of-features enhanced the accuracy results and statically significant compared to the difference-of-features alone (p-value < 0.05). As the result, the proposed approach of combining multiple ROIs and DoF has adopted to benchmark the performance of the lesion classification through utilising high-level features and comparing it in terms of the classification performance via low-level features.

Classification based on High-level Features

In this section, the performance of the lesion classification based on high-level features is evaluated through five types of evaluation metrics and ROC curve. The semantic features are used to build the new feature vector and, fed into classifier. Table 6.7 depicts the lesion classification based on high-level features evaluation using portal phase CT images and multiphase CT images evaluation presented in Table 6.8 .The two validation approaches (tenfold cross-validation and Train/Test split validation) were performed in evaluation. The results show that the classification performance when used

portal phase CT images in characterisation achieved same accuracy levels compared to using multiphase CT images in characterisation, due to using multiphase CT images in characterisation is slightly better than portal phase CT images to characterised lesion structure where the portal phase CT images shows slightly better performance in characterised lesion shape and margin, as discussed in Section 6.3.

Validation	Lesion	SN	SP	ACC	PPV	NPV	Average ACC
cross-validation	Malignant	0.975	0.968	0.975	0.963	0.978	0.971
	Benign	0.968	0.975	0.968	0.978	0.963	
Train/Test	Malignant	0.931	0.905	0.931	0.931	0.905	0.92
	Benign	0.905	0.931	0.905	0.905	0.931	

Table 6.7: Summary of lesion classification through high-level features results obtained by **tenfold cross-validation** and **split Training/Testing validation** method. The SVM classifier and **portal phase** CT images were adopted.

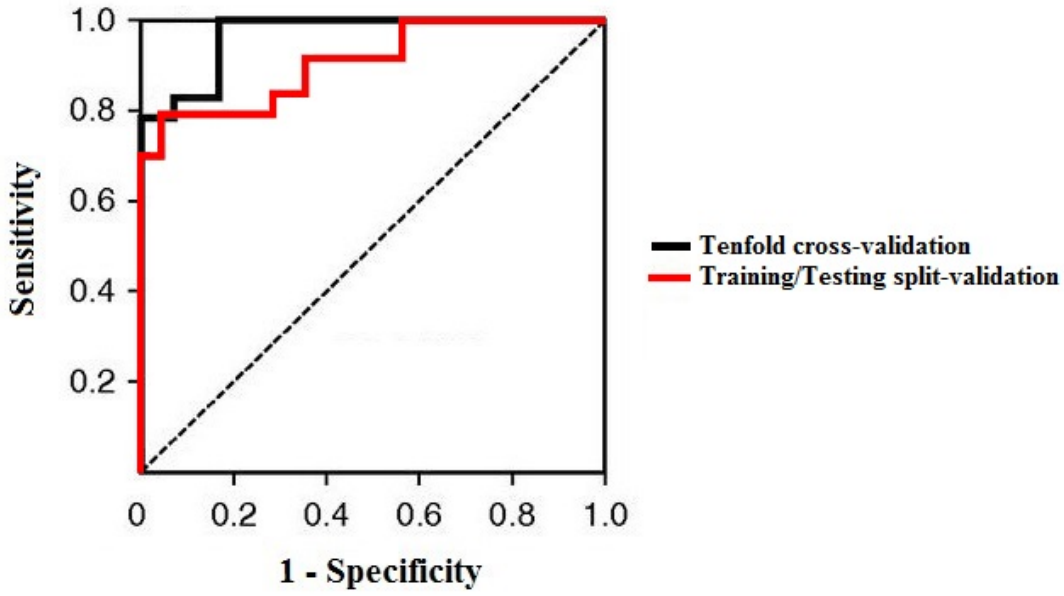


Figure 6.15: ROC-curves of lesion classification based on high-level features using **portal phase** CT images and SVM classifier where **tenfold cross-validation** and **Training/Testing split validation** approaches were used.

Validation	Lesion	SN	SP	ACC	PPV	NPV	Average ACC
cross-validation	Malignant	0.975	0.968	0.975	0.963	0.978	0.971
	Benign	0.968	0.975	0.968	0.978	0.963	
Train/Test	Malignant	0.931	0.905	0.931	0.931	0.905	0.92
	Benign	0.905	0.931	0.905	0.905	0.931	

Table 6.8: Summary of lesion classification through high-level features results obtained by **tenfold cross-validation** and **split Training/Testing validation** method. The SVM classifier and **multiphase** CT images were adopted.

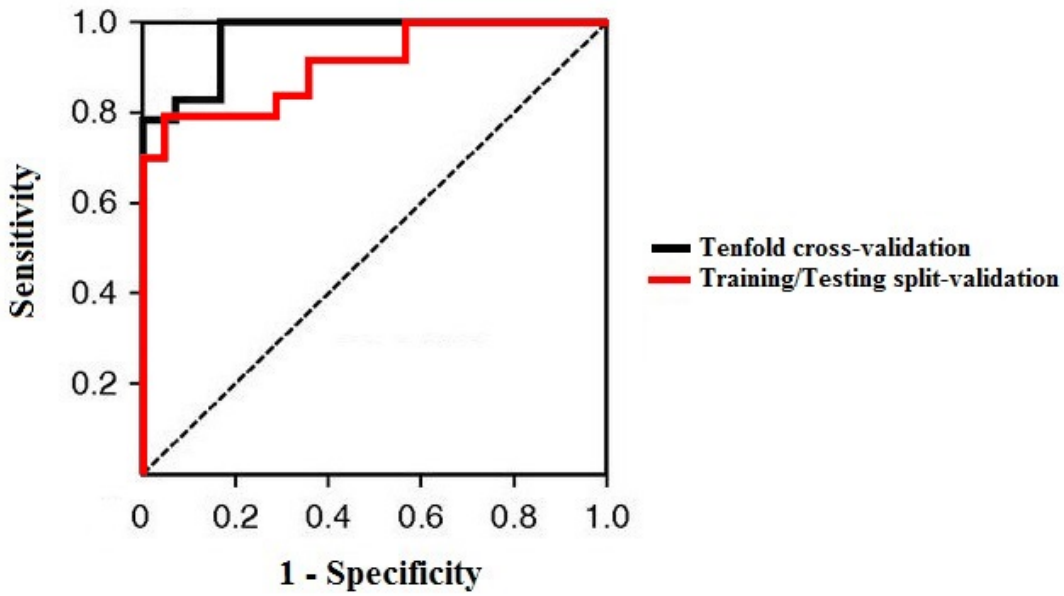


Figure 6.16: ROC-curves of lesion classification based on high-level features using **multiphase** CT images and SVM classifier where **tenfold cross-validation** and **Training/Testing split validation** approaches were used.

Furthermore, the high-level features have been shown to be more useful in lesion classification than low-level features, especially in the classification of malignant lesions. The results based on portal phase CT images show that the classification through high-level features provides up to $1.7 \pm 0.6\%$ overall accuracy increase compared to low-level features and up to $3.7 \pm 1.2\%$ accuracy increase for malignant lesion classification and the same level of accuracy for benign lesions. The use of multiphase CT images has enhanced the classification accuracy based on low-level features to achieve the same level of accuracy through utilising high-level features. However, the high-level features still offers higher accuracy in classification of malignant lesions by $1.2 \pm 0.6\%$ compared to the low-level features, which confirms the classification through high-level features higher accuracy to fulfil its job. In addition, the use of high-level features helps in interpreting and explaining the classification and is more intuitive to clinicians, which is in contrast with black box low-level features approach.

Classification based on combined High-level and Low-level Features

This section investigates the performance of the combined high-level and low-level features to further analysis their lesion classification capabilities across the datasets. Table 6.9 and Table 6.10 depicts the evaluation analysis of combined high-level and low-level features for lesion classification task across the tenfold cross-validation and Train/Test split validation approach by using portal and multiphase CT and images respectively. The results show that the merged high-level and low-level features achieved slightly higher accuracy compared to its individual components with multiphase CT images.

Validation	Lesion	SN	SP	ACC	PPV	NPV	Average ACC
cross-validation	Malignant	0.963	0.979	0.963	0.975	0.968	0.971
	Benign	0.979	0.963	0.979	0.968	0.975	
Train/Test	Malignant	0.897	0.952	0.897	0.963	0.870	0.92
	Benign	0.952	0.897	0.952	0.870	0.963	

Table 6.9: Summary of lesion classification through combined features results obtained by **tenfold cross-validation** and **split Training/Testing validation** method. The SVM classifier and **portal phase** CT images were adopted.

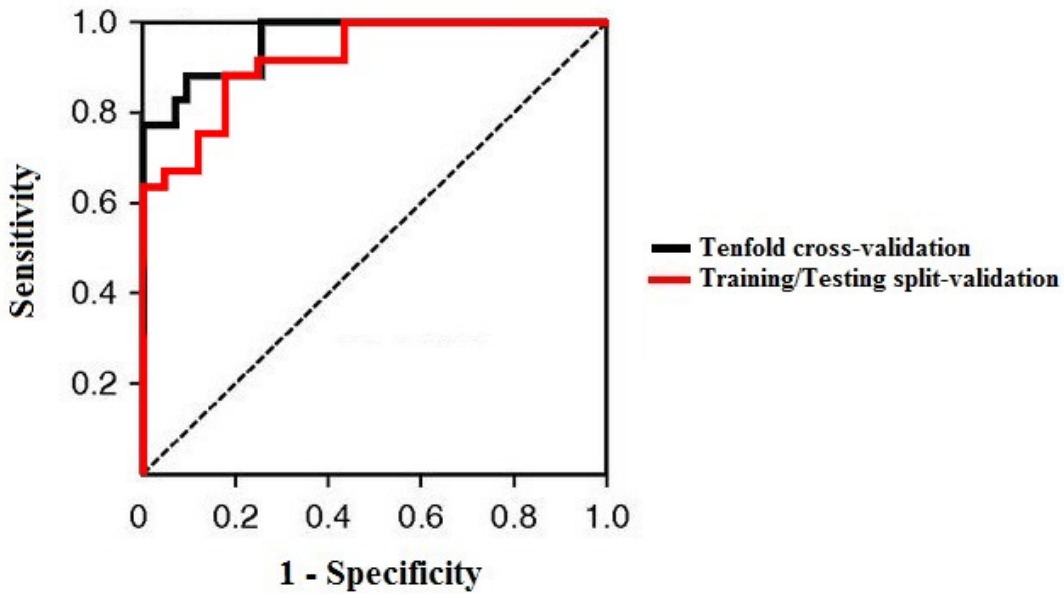


Figure 6.17: ROC-curves of lesion classification based on combined features using **portal phase** CT images and SVM classifier where **tenfold cross-validation** and **Training/Testing split validation** approaches were used.

Validation	Lesion	SN	SP	ACC	PPV	NPV	Average ACC
cross-validation	Malignant	0.975	0.979	0.975	0.975	0.979	0.977
	Benign	0.979	0.975	0.979	0.979	0.975	
Train/Test	Malignant	0.931	0.952	0.931	0.964	0.909	0.94
	Benign	0.952	0.931	0.952	0.909	0.964	

Table 6.10: Summary of lesion classification through combined features results obtained by **tenfold cross-validation** and **split Training/Testing validation** method. The SVM classifier and **multiphase** CT images were adopted.

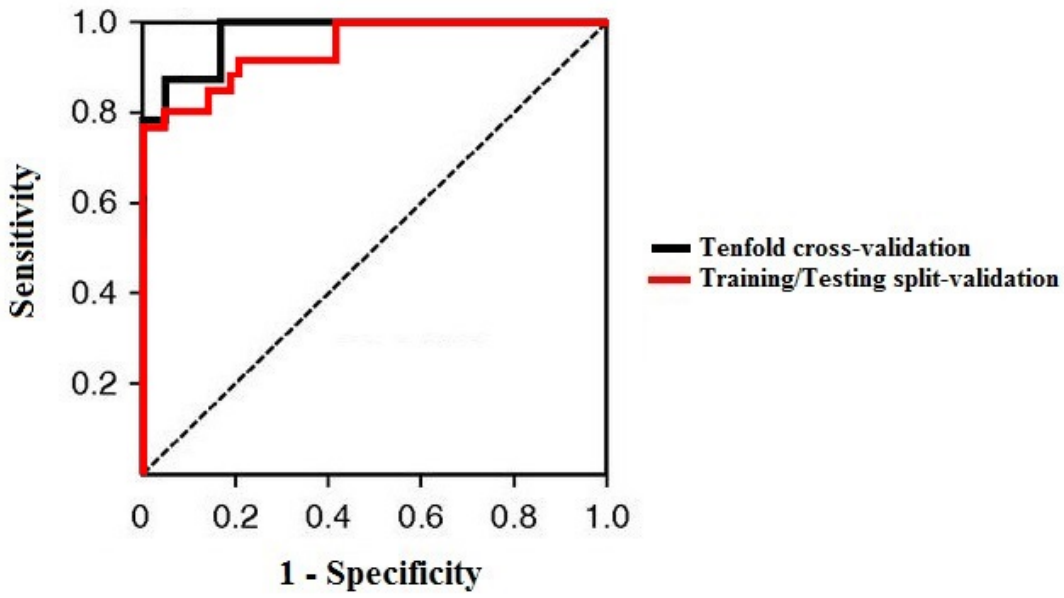


Figure 6.18: ROC-curves of lesion classification based on combined features using **multiphase** CT images and SVM classifier where **tenfold cross-validation** and **Training/Testing split validation** approaches were used.

Furthermore, since the combined features is designed to be used for enhance the classification accuracy, the results based on portal phase shows that the combined features achieved $1.7 \pm 0.8\%$ increase in overall accuracy, $2.5 \pm 0.9\%$ and $1.1 \pm 0.7\%$ increase in malignant and benign lesions classification respectively, compared to the classification through low-level features. For the comparison with the classification through high-level features, the combined features achieved $1.2 \pm 0.3\%$ less accuracy in malignant lesions classification and $1.1 \pm 0.1\%$ higher accuracy in benign lesion classification. The combined features have slightly improved the classification accuracy to 97.7% by using multiphase CT image.

6.5 Framework Benchmarking

This section shows the proposed framework for liver lesion characterisation and classification benchmarking against state-of-art baselines, in terms of system accuracy. Section 6.5.1 presents the results and evaluation of proposed liver lesion characterisation framework against the baselines. The proposed framework for liver lesion classification versus the baselines are presented in Section 6.5.2.

6.5.1 Benchmarking Characterisation and Comparisons

Table 6.11 depicts the selected list of state-of-art baselines to benchmark the proposed liver lesion characterisation framework. These baselines have been selected based on careful literature analysis, as presented and discussed in Chapter 3, in addition to cover all the used same dataset, as most of the presented works provide results for their own

dataset. Furthermore, each of the baselines uses different technique in characterising lesions. The baselines were selected by the end of February 2016, where we started the evaluation stage and prepared to write the thesis.

Generally, the selected baselines could be summarised based on the dataset as follows:

- Same our dataset (Dataset I):
  Baselines 1, 3, 4 and 6. The baseline 1 regenerated by 6.
- Using all our dataset by re-implemented the baselines:
  Baselines 1, 2, 5 and 6. The baseline 6 provided us with the code.

#	Baseline	Dataset	Approach
1	(Gimenez et al., 2012)	79 case	Machine learning. intensity, texture, shape, and edge sharpness features extracted. LASSO classifier.
2	(Depeursinge et al., 2014)	74 case	Machine learning. Lesion was divided into 12x12 patches. linear combination of texture features based on Riesz wavelets. SVM classifier.
3	(Spanier and Joskowicz, 2014)	50 case (Dataset I)	Machine learning. Intensity, texture features with Linear Discriminant Analysis (LDA) classifier.
4	(Nedjar et al., 2015)	50 case (Dataset I)	Case-based similarity. Retrieve the top five similar images then applied the majority voting.
5	(Kurtz et al., 2015)	72 case	Machine learning. Lesion was divided into 12x12 patches. A gray-level intensity histogram and multi-scale Riesz wavelets. SVM classifier.
6	(Kumar et al., 2016)	50 case (Dataset I)	Case-based similarity. weighted nearest-neighbour (WNN) with sequential feature selection.

Table 6.11: Summaries of the selected baselines approaches to benchmark the proposed liver lesion characterisation framework.

Table 6.12 depicts the evaluation results of the proposed lesion characterisation framework benchmarking against the selected baselines set. The reported proposed framework results, over the Dataset I, were obtained based on portal phase CT images, for fair comparison with the baselines.

#	High-level feature	Baseline 1	Baseline 3	Baseline 4	Baseline 6	Proposed system
1	Lesion density	0.76	0.86	0.82	0.88	**0.94**
2	Lesion density type	0.78	0.84	0.84	0.82	**0.94**
3	Lesion rim	0.62	-	-	-	**0.96**
4	Lesion rim thickness	-	-	-	-	**0.94**
5	Contrast Uptaken	0.84	0.86	0.80	0.86	**0.90**
6	Enhancement Pattern	-	0.84	0.82	0.88	**0.94**
7	Lesion composition	-	0.88	0.86	0.90	**0.92**
8	Lesion leveling type	-	0.82	0.82	0.84	**0.90**
9	Lesion shape	0.58	0.74	0.84	**0.88**	**0.88**
10	Lesion focality	0.40	0.26	0.88	0.90	**0.96**
11	Lesion margin	0.68	0.86	0.84	0.86	**0.94**
12	Lesion margin definition	0.60	0.80	0.82	0.84	**0.92**
13	Lesion enhancement	0.74	0.84	0.80	0.86	**0.92**
14	Lesion brightness	0.82	-	-	-	**0.90**
15	Lesion surrounding	-	-	-	-	**0.92**
16	Calcified (inside lesion)	-	0.82	0.90	0.94	**1**
17	Calcified wall	-	0.62	0.74	0.68	**0.90**
18	Scar	-	-	-	-	**1**
19	Lobe	-	0.62	0.52	0.48	**0.98**
20	Segment	-	0.40	0.32	0.26	**0.92**
21	Close to vein	-	0.48	0.54	0.60	**0.98**

Table 6.12: The proposed lesion characterisation framework accuracy (%) performance versus the baselines set on the same dataset (Dataset I).

The proposed framework outperformed all the baselines over the dataset I with average accuracy of 94%. In addition, the proposed system was able to characterise the liver lesions with the additional number of the high-level features compared to the baselines. The baselines accuracy to characterised lesion location such as Segment is low, due to the primitive way in which the high-level features were extracted .

Conclusively, the competitive performance of the proposed framework relates to its robust structure, which is designed to improve the lesion characterisation model through using multiple ROIs and difference-of-features. In addition, the mapping between the high-level features and the selected ROI based on the medical knowledge are the keys to capturing all the lesion characteristics efficiently.

Table 6.13 depicts the evaluation results of the proposed lesion characterisation framework benchmarking against the selected baselines set. Four baseline methods are studied to be compared with the proposed framework over the whole dataset. The baseline 1 and 6 provided us with the source code and we ran their code over the dataset. Due to the limited availability of the used dataset, we have regenerated the baseline 2 and 5 and applying it on the available dataset.

#	High-level feature	Baseline 1	Baseline 2	Baseline 5	Baseline 6	Proposed system
1	Lesion density	0.80	-	0.74	0.90	**0.97**
2	Lesion density type	0.81	0.87	0.71	0.86	**0.96**
3	Lesion rim	0.72	-	0.64	-	**0.98**
4	Lesion rim thickness	-	-	-	-	**0.95**
5	Contrast Uptaken	0.82	0.84	0.88	0.84	**0.95**
6	Enhancement Pattern	-	0.75	0.85	0.87	**0.95**
7	Lesion composition	-	-	-	0.91	**0.97**
8	Lesion leveling type	-	-	-	0.87	**0.94**
9	Lesion shape	0.67	-	0.60	0.83	**0.94**
10	Lesion focality	0.34	-	-	0.85	**0.96**
11	Lesion margin	0.76	0.82	0.86	0.84	**0.95**
12	Lesion margin definition	0.67	0.78	-	0.79	**0.95**
13	Lesion enhancement	0.72	0.76	-	0.89	**0.97**
14	Lesion brightness	0.84	0.71	0.79	-	**0.97**
15	Lesion surrounding	-	-	-	-	**0.95**
16	Calcified (inside lesion)	-	-	-	0.90	**0.99**
17	Calcified wall	-	-	-	0.74	**0.94**
18	Scar	-	-	-	-	**0.99**
19	Lobe	-	-	-	0.53	**0.98**
20	Segment	-	-	-	0.44	**0.93**
21	Close to vein	-	-	-	0.65	**0.94**

Table 6.13: The proposed lesion characterisation framework accuracy (%) performance versus re-implementation of baselines set on our dataset (Dataset I + Dataset II).

The best liver lesion characterisation performance is obtained by our proposed framework with average characterisation accuracy of 96% for the predictive high-level features. These results are consolidated by the comparison with the three other baselines methods over the dataset size of 174 cases. The proposed framework always shows the highest accuracy results with an average over of 92%.

6.5.1.1 Discussion

In this thesis, the publicly available ImageClef dataset was used since it is a popular dataset in this field. The training dataset included the ground truth used for liver lesion characterisation. A fair comparison with other recent and related work which used the same dataset or applied their methods on our dataset will be also feasible.

Liver lesion interpretation from CT image is being widely studied. In the previous studies, characterisation of the liver lesion was mainly explored in two ways: (1) Case-based retrieval approach (CBIR) and (2) Machine learning-based approach (ML). Following the ML approach, (Gimenez et al., 2012) computed a relatively large set of quantitative features from liver CT scans and fed whole feature matrix in LASSO regularisation model to predict the high-level features. (Depeursinge et al., 2014) adopted

a different modeling method and created a SVM model using only the rotation covariant Riesz wavelet features to learn the signature of each high-level feature from CT images. (Spanier and Joskowicz, 2014) trained four different classifier to characterise liver lesions using a set of quantitative image features. (Kurtz et al., 2015) adopted SVM to characterise liver lesion based on gray-level intensity histogram and multi-scale Riesz wavelet features.

Moving towards the direction of CBIR for lesion characterisation, (Nedjar et al., 2015) used CBIR approach to characterise liver lesion, by using a specific signature of the liver. The similarity metric used to retrieve the top five smilar images to the given image. The majority voting between retrieved images has been used in characterisation. (Kumar et al., 2016) Adopted the weighted nearest-neighbour search approach with sequential feature selection to find the most similar training images for characterisation task. To the best of our knowledge, no study was mapped between low-level, high-level features and represented ROI to characterise the liver lesion. our proposed framework categorises the high-level features into two groups: (1) Visual features from the image contents and (2) High-level features that predicted through learning approach. Lesion characterisation infers explicit mapping between low-level, high-level features and ROI, as presented in Chapter 5.

Table 6.12 provides the results of our proposed framework against other methods and applied on Dataset I (ImageClef). The results show that the proposed framework achieved the highest lesion characterisation accuracy, outperforming the other baseline methods. The highly significant results achieved by our proposed method was due to the way in which the high-level features were predicted. All of the high-level features number (9, 10, 16, 17, 18, 19, 20 and 21) are estimated from the image itself. which called the visual features from image contents. However, the baselines inferred these features in a primitive way based on other cases, which increases the probability of error in characterisation. For example, the lesion location (Segment, Lobe) features depend on the visual image contents where cannot be predicted correctly from other cases. However, In our proposed framework, these features calculated from the image itself by segmented the liver based on the main vessels and medical knowledge, as described in Section 5.3.2.

Regarding the two high-level features (Calcified inside lesion and Scar), the proposed framework achieves 100% accuracy. This is not surprising since the dataset contains only 3 cases with Calcified inside lesion and 4 cases with Scar which are detected correctly. These areas inside lesion are distinguishable due to the intensity difference between lesion tissue and Calcified and Scar area, located at lesion centre, as presented in Chapter 5.

The remain features accuracy of our proposed framework was higher than that of baselines. The results have experimentally confirmed the importance of mapping between low-level, high-level and ROI. We extracted a set of low-level feature form each

proposed ROIs (inside, border and outside lesion) and order the features as per their relevancy to the targeted high-level feature. For example, the lesion rim feature is inferred by considering intensity and texture feature from outside lesion ROI. The characterisation of lesions depends heavily on the characteristics of lesions including internal structure, morphology, border and rim. These characteristics are differently observed according to ROI area that may indeed significantly impact the characterisation performance. The baselines have limitation for guaranteeing robust characterisation performance for liver lesion, mainly due to the difficulties that represent all appearances of the lesion characteristics through the current single ROI. Similarly, the baseline 1 (Gimenez et al., 2012) based on LASSO regularisation model, which performs variable selection by shrinking parameter estimates (coefficients of the regression) closer to zero, has a low accuracy compared to other methods. This is due to the dataset contained a few subtle cases meaning there were a few samples from which to derive optimal features or regression coefficients. However, A more diverse training dataset would include more subtle cases and this would have a positive impact on the accuracy of methods, as illustrated in Table 6.13.

Conclusively, these results highlight the potential benefits of mapping between low-level, high-level and ROIs as a base for the proposed framework towards a more accurate characterisation performance compared with the existing characterisation approaches.

6.5.2 Benchmarking Classification and Comparisons

Table 6.14 presents the selected list of baselines to benchmark the proposed liver lesion classification framework. These baselines were selected because they are the most recent relevant work. So, they represent the state-of-art with high accuracy, as discussed in Chapter 3. It is well known that the datasets are limited in the medical field. We already contacted several researchers in the field to either share their datasets or at least to run our code on their datasets, however, this was fully refused for confidentiality purposes. Due to the limited availability of a standard benchmark dataset, we have regenerated the baselines by implementing and testing them on our dataset.

Generally, the selected baselines could be summarised based on the CT phases as follows:

- Using portal phase:
  Baselines 2, 3 and 4.
- Using multiphase:
  Baselines 1 and 5.

#	Baseline	Dataset	Approach
1	(Quatrehomme et al., 2012)	95 case	Multiphase. Law texture measures,Gaussian Markov Random Fields, Unser histograms and First order statistics features were extracted. SVM classifier.
2	(Kumar et al., 2013a)	150 case	Portal phase. Gray level, GLCM, wavelet coefficient statistics and contourlet coefficient statistics. A probabilistic neural network classifier.
3	(Doron et al., 2014b)	112 case	Portal phase. Gabor, Local Binary Patterns and grey level intensity features. SVM classifier.
4	(Diamant et al., 2017)	118 case	Portal phase. Bag-of-visual-word with patch size of 11 and visual of word size 10. SVM classifier.
5	(Chang et al., 2017)	71 case	Multiphase. 3D texture (GLCM), 3D shape (compactness, elliptic model and Margin) and Kinetic curve features were obtained from lesion . binary logistic regression classifier.

Table 6.14: Summaries of the selected baselines approaches to benchmark the proposed liver lesion classification framework.

Table 6.15 presents the performance of various proposed methods (**Proposed P** ^{1} based on low-level features, **Proposed P** ^{2} through high-level features and **Proposed P** ^{3} combined low-level and high-level features) used to classify liver lesion based on portal phase CT images, compared with three baselines using accuracy, sensitivity, specificity, positive predictive, and negative predictive value.

Baseline	Lesion	SN	SP	ACC	PPV	NPV	Average ACC
Baseline 2	Malignant	0.813	0.883	0.813	0.855	0.847	0.851
	Benign	0.883	0.813	0.883	0.847	0.855	
Baseline 3	Malignant	0.775	0.894	0.775	0.861	0.824	0.839
	Benign	0.894	0.775	0.894	0.824	0.861	
Baseline 4	Malignant	0.838	0.872	0.838	0.848	0.863	0.856
	Benign	0.872	0.838	0.872	0.863	0.848	
Proposed P ^{1}	Malignant	0.938	0.968	0.938	0.962	0.948	0.954
	Benign	0.968	0.938	0.968	0.948	0.962	
Proposed P ^{2}	Malignant	0.975	0.968	**0.975**	0.963	0.978	**0.971**
	Benign	0.968	0.975	0.968	0.978	0.963	
Proposed P ^{3}	Malignant	0.963	0.979	0.963	0.975	0.968	**0.971**
	Benign	0.979	0.963	**0.979**	0.968	0.975	

Table 6.15: Liver lesion classification based on portal phase CT image performance comparison between the proposed methods and baselines.

Regarding to the liver lesion classification, all proposed three methods achieved higher accuracy compared to each baseline in terms of the lesion type and overall system performance. The results show that $\mathbf{P}^{2}$ and $\mathbf{P}^{3}$ achieves higher overall system accuracy. Furthermore, $\mathbf{P}^{2}$ achieves the highest accuracy in the classification of the malignant lesion 97.5%. However, the $\mathbf{P}^{2}$ performed 1.1% less than $\mathbf{P}^{3}$ for the benign lesion classification. It is obvious from Figure 6.19 that $\mathbf{P}^{2}$ gained higher performance and increased the classification system accuracy 1.7% and 11.5% compared to $\mathbf{P}^{1}$ and the highest overall accuracy baseline respectively, with the advantage of interpreting the diagnosis decision in human understating and analogous to radiologist observation.

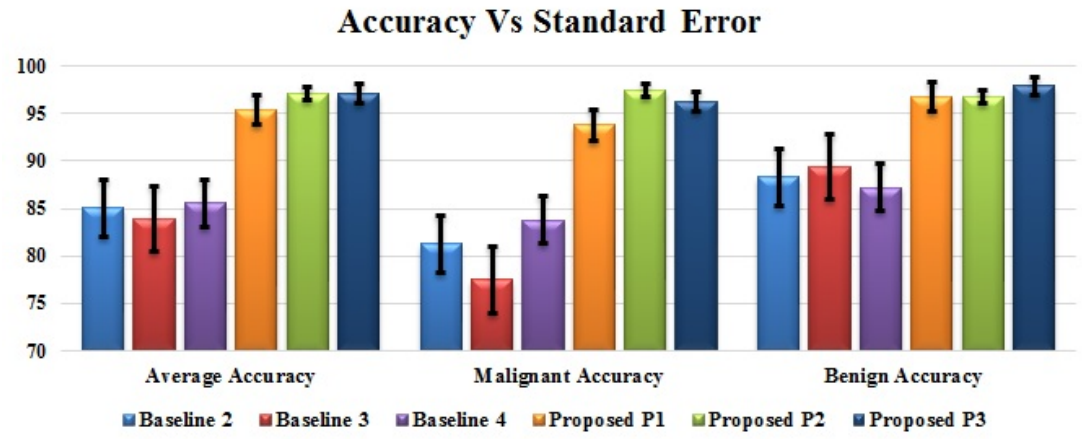


Figure 6.19: Accuracy Vs Standard error for classification performance comparison between proposed methods and baselines based on **portal phase** CT image. $\mathbf{P}^{2}$ and $\mathbf{P}^{3}$ achieves higher overall accuracy.

However, when comparing the performance of approaches, it is essential not to rely on just the statistical significant of the differences in sensitivities, specificities or accuracies; the magnitude of the differences must also be assessed (Dwyer, 1991). The two-by-two matched data array and the McNemar analysis provide a succinct format for the presentation and proper analysis of matched comparisons of binary (positive and negative) test results (Fagerland et al., 2013). For more confident results, the McNemar test was used in this study to assess our proposed system compared to the baselines. Table 6.16 Depicts the statistical analysis using McNemar test to compare proposed method over other approaches, where a p-value of <0.05 indicates a statistically significant.

Baselin 2	**Baselin 2**				
Baselin 3	0.479500	**Baselin 3**			
Baselin 4	1	0.617075	**Baselin 4**		
Proposed P1	**0.001911**	**0.001542**	**0.002076**	**Proposed P1**	
Proposed P2	**0.000204**	**0.000023**	**0.000330**	0.248213	**Proposed P2**
Proposed P3	**0.000318**	**0.000044**	**0.000523**	0.371093	0.479500

Table 6.16: McNemar analysis p-value results in comparison between all proposed approaches versus selected baselines to classify liver lesion based on **portal phase** CT images.The p-value < 0.05 scores are in bold font.

According to the results in Table 6.16, our new methods for liver lesion classification based on low-level, high-level and combination low-level and high-level features are statistically significant compared to the baselines. However, the **P** ^{2} and **P** ^{3} are doing better than **P** ^{1} but is not statistically significant, where the **P** ^{2} provides more understanding and interpretation of the decision for the radiologists, which is in contrast with black box low-level features approach.

Table 6.17 presents the performance of various proposed methods (**Proposed P** ^{1} based on low-level features, **Proposed P** ^{2} through high-level features and **Proposed P** ^{3} combined low-level and high-level features) used to classify liver lesion based on multiphase CT images, compared with two baselines using several evaluation metrics.

Baseline	**Lesion**	**SN**	**SP**	**ACC**	**PPV**	**NPV**	**Average ACC**
Baseline 1	Malignant	0.863	0.894	0.863	0.873	0.884	0.879
	Benign	0.894	0.863	0.894	0.884	0.873	
Baseline 5	Malignant	0.750	0.841	0.75	0.80	0.798	0.79.6
	Benign	0.841	0.750	0.841	0.798	0.800	
Multiple ROIs + DoF	Malignant	0.963	0.979	0.963	0.975	0.968	0.971
	Benign	0.979	0.963	**0.979**	0.968	0.975	
Through characterisation	Malignant	0.975	0.968	**0.975**	0.963	0.978	0.971
	Benign	0.968	0.975	0.968	0.978	0.963	
Combined features	Malignant	0.975	0.979	**0.975**	0.975	0.979	**0.977**
	Benign	0.979	0.975	**0.979**	0.979	0.975	

Table 6.17: Liver lesion classification based on **multiphase** CT image performance comparison between the proposed methods and baselines.

The proposed lesion classification framework, represented by the **P** ^{3}, outperformed all the baselines over the available dataset. The **P** ^{3} performed 9.8% more than the highest baseline accuracy (baseline 1) and slightly higher by only 0.8% compared to the **P** ^{1} and **P** ^{2}. However, the **P** ^{2} has an advantage of diagnosis explanation through characterisation compared to the traditional ways based on low-level features. Furthermore, **P** ^{2} provides the highest accuracy of malignant lesion classification up to 97.5%.

Accuracy and standard error measures the overall effectiveness and trueness of the proposed classification systems versus baselines as illustrated in Figure 6.20.

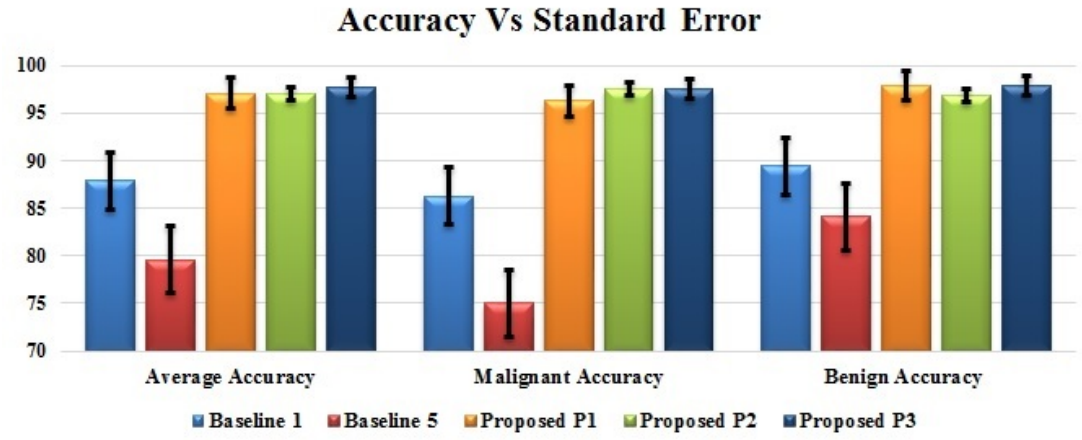


Figure 6.20: Accuracy Vs Standard error for classification performance comparison between proposed methods and baselines based on **multiphase** CT image. **P** ^{2} and **P** ^{3} achieves higher overall accuracy.

Baselin 1	**Baselin 1**			
Baselin 5	**0.002183**	**Baselin 5**		
Proposed P1	**0.001384**	**0.000005**	**Proposed P1**	
Proposed P2	**0.000796**	**0.000003**	0.479500	**Proposed P2**
Proposed P3	**0.000480**	**0.000001**	1	1

Table 6.18: McNemar analysis p-value results in comparison between all proposed approaches versus selected baselines to classify liver lesion based on **multiphase** CT images.The p-value < 0.05 scores are in bold font.

The results in Table 6.18 for the McNemar statistical analysis of the comparison between the proposed framework and baselines based on multiple CT images show that all the proposed methods is statistical significant compared to the baselines. Furthermore, the baseline 6 is better than baseline 1 statistically. Although the accuracy of **P** ^{3} is slightly higher than **P** ^{1} and **P** ^{2} but it is not statistical significant and which will lose the advantage of **P** ^{2} approach to interpret the results. Moreover, the **P** ^{2} is able to classify the liver malignant lesion more accurately compared to **P** ^{1} approach.

6.5.2.1 Discussion

In this thesis, Due to the limited availability of a standard benchmark dataset for the liver lesion classification task, we have regenerated the baselines by implementing and testing them on our dataset.

Recent studies reported classification accuracy in the range of $82 - 97\%$ in classifying liver lesion. The high variation in these reported results is partly due to different dataset size being used but also due to the fact that different studies using different validation approach, such as leave-one-out or 10-fold cross-validation. The method proposed by (Quatrehomme et al., 2012) uses SVM to separate five types of liver lesion in multiphase CT images based on texture and other statistical features. (Kumar

et al., 2013a) presented a method that extracts curvelet and texture features from the CT image, which serve as input of a probabilistic neural network. The evaluation showed the method can efficiently distinguish between two particular types of liver lesions. (Doron et al., 2014b) combines various texture, intensity features and classifiers for differential diagnosis of focal liver lesions. (Diamant et al., 2017) applied the bag-of-visual-words (BoVW) method learned from image patches. They used two dictionaries for lesion interior and boundary regions. Based on the two dictionaries they generated histograms for each lesion ROI. The final classification was made using SVM. (Chang et al., 2017) obtained three kind of features for each tumor, including texture, shape, and kinetic curve on segmented tumors. Backward elimination was used to select the best combination of features through binary logistic regression analysis to classify the tumors, as presented in Chapter 3.

Table 6.15 and Table 6.8 provide the results of our proposed framework against other methods applied on our dataset (174 CT scans) based on portal phase and multiphase CT scans respectively. The experiment results show that the classification method based on high-level features gained higher performance compared to the low-level features and recorded accuracy 97% with the advantage of interpreting the diagnosis decision in human understating and analogous to radiologist observation. The combination between low-level and high-level features was slightly higher from high-level features in multiphase CT image but will lose the advantages of characterisation and track back to diagnosis explanation, due to a diagnostic decision that not only depends on high-level feature but also on low-level features that are unable to interpret the classification result.

The baselines show lower accuracy compared to our proposed framework. This is due to the classification of liver lesions depends largely on how to represent the characteristics of lesion based on ROI selection. However, the main limitation of previous studies with traditional ROI (mainly selected lesion area) is that they may not able to represent all lesion characteristics in a reliable way. Because the ability to observe each of the lesion characteristics could differ depending on the type of extracted ROIs.

On the other side, the classification accuracy based on multiphase CT scan shows better results compared to the single phase (portal phase). This is due to the multiphase provide more represented information, in particular on hypervascular lesions.

Conclusively, these results show the classification lesion based on high-level features (characterisation) is better than all other methods, due to gained a higher accuracy. In addition, it emphasise the usefulness of the proposed framework for the lesion characterisation and classification. However, the characterisation approach provides more understanding and interpretation of the decision for the radiologists, which is in contrast with the black box low-level features approach. Another interesting result

is the ability of our proposed classification through characterisation to classify liver malignant lesion more accurately compared to low-level features.

6.6 Conclusion

This chapter provided an extensive benchmarking and evaluation of the proposed liver lesion characterisation/ classification framework. Both, lesion characterisation and classification were introduced. The evaluation sequence was as follows:

- Evaluation of the proposed liver lesion characterisation framework to assess their capabilities to extract high-level features that describe the lesions and quantify the performance obtained. The evaluation reflected the benefits of mapping between high-level feature and ROI selection based on medical prior-knowledge towards more descriptive lesion characteristics.
- Evaluation of the proposed liver lesion classification framework. The evaluation separated into three parts, as follows:
  - Evaluation of lesion classification based on low-level features. The performance of the proposed lesion classification system is evaluated by using three different configurations. The first way, the lesion is classified using difference-of-features (DoF) between segmented lesion and surrounding tissues. The second way, the classification task is carried out by fusing features of Multiple ROIs (internal, border and surrounding lesion). The last configuration, the classification performed through fusing lesion border features with the difference-of-features between internal and surrounding lesion. The evaluation showed that the fusing multiple ROIs and DoF (last configuration) performs higher than its individual components.
  - Evaluation of lesion classification based on high-level features. The evaluation related to the benefits of the utilising the high-level features that extracted through lesion characterisation framework to enhance lesion classification performance. The results showed that the characterisation approach has improved the accuracy of liver lesion classification, especially malignant lesions accuracy with the benefit of interpretable the classification decision.
  - Evaluation of lesion classification based on the fusion between low-level and high-level features. The evaluation results showed that the combined features have slightly improved the classification accuracy compared to the characterisation approach. However, the combination approach will lose the advantage of characterisation approach to interpret the results.

- Benchmarking of the liver lesion characterisation/classification framework against a related set of baselines. This evaluation quantified the proposed characterisation/classification framework power, as it demonstrated with high accuracy scores.

All the evaluations were conducted over two datasets that were discussed in detail in Chapter 2. The dataset provided a level of challenge by collecting the CT images from the different sources that were varied in the machine settings which reflected on the image resolution and spacing. This dataset added a considerable credibility to the reported results. Furthermore, the multiphase CT images results are compared to the portal phase CT (single phase) and show a slight improvement in lesion structure characterisation while the single phase better in lesion shape and margin characterisation, due to the lesion size and the fill in/washout of the contrast agent during the time of screening. On the other hand, the multiphase provides an improvement on lesion classification based on low-level features, in particular on hypervascular lesions, due to its being more descriptive for lesion.

Going a step further, a related solid and recent set of baselines were selected to measure and benchmark the performance of liver lesion characterisation/classification framework. The baselines vary in their techniques that used to enhance the lesion characterisation and classification performance which makes it more competitive to the proposed framework. The proposed framework can achieve a high characterisation accuracy with the large number of high-level features compared to the baselines and can go further, by classifying the liver lesion with the advantage of interpreting the diagnosis decision in human understating and analogous to radiologist observation, which is in contrast with black box low-level features approach. The next chapter will involve the thesis conclusion with highlights on the contributions and presents some potential future work.

Chapter 7

Conclusion

This thesis provided an integrated framework to characterise and classify liver lesions. Targeting the liver lesion characterisation was the main goal of this thesis, for the objective of characterising liver lesion with a large number of high-level features automatically. Furthermore, the proposed framework enhances liver lesion classification through utilising high-level features to classify the respective lesions, with the benefit of interpretable characterisation that supports the diagnostic decision. The liver lesion characterisation is emerging as a research issue and challenging task. The main challenge is to investigate the detection of subtle differences in low-level features of the selected ROI and to link them to the high-level features derived from a standard terminology to represent the lesion characteristics in analogy with radiology. To achieve this task, we had to face some critical challenges from the liver lesion characterisation field in general and ROI selection in particular.

Most of the available methods operate by exploiting low-level features from the lesion ROI/ patches, with no pay attention to the relation between lesion, margin and liver, to calculate the high-level features related to the lesion characteristics based on learning process or case-based retrieval, as presented in Chapter 3. The accuracy of the lesion characterisation is usually affected by selecting ROI to represent each of the lesion characteristics. This could be referred as the ROI selection problem over the lesion characterisation. In addition, there are some of the high-level features that rely on the case itself, and cannot be predicted by using case-based retrieval or machine learning approach such as lesion location.

On the other hand, Most of the existing research deal with low-level features only, for lesion classification. In addition, they mainly focused on feature extraction and classification for well-performing system. While the ROI selection methods introduced as a new research area, which is importantly responsible for representing the lesion characteristics. Through literature, there have been no attempts to classify liver lesion through high-level features (to the best of our knowledge), since the majority of relevant research attempted to classify lesions based on low-level features.

This thesis presents two main novelties, the first contribution is the automatic characterisation of the liver lesion with a large number of high-level features compared to relevant state-of-art baselines. The novelty of the proposed characterisation approach lies in:

- Assigning high-level descriptions to the liver lesion in analogy to radiologist observation.
- Proposing multiple image ROIs (inside, border and surrounding lesion) by generating abnormality level map based for the lesion and surrounding area to calculate the high-level features by considering the ability of each ROI that represents a set of lesion characteristics.
- Associating between low-level, high-level features and the selected ROI to capture all the lesion characteristics and provide a human interpretable information.

The second contribution of this thesis is enhancing liver lesion classification performance in three different novel ways, as follows:

- Classifying liver lesion based on low-level features. The novelty of the proposed approach lies in; Proposing multiple ROIs fused with difference-of-features between the internal lesion and surrounding lesion area from the normal liver tissue. All the lesion characteristics such as internal, boundary and relation with the surrounding area were captured through our proposed multiple ROIs. Moreover, the Difference-of-Features helps to overcome the challenge issues such as the variation of intensity and texture ranges between study cases due to the imaging devices/ settings such as images resolution and spacing, which makes it more generic to dealing with different datasets.
- Classifying liver lesion based on high-level features. The high-level features calculated by the lesion characterisation framework are utilised to generate a novel feature vector. In contrast with most existing research, which uses low-level features only, the use of high-level features (characterisation) helps in interpreting and explaining the classification decision and is more intuitive to clinicians. Another interesting result is the ability to use high-level features to classify liver lesion more accurately compared to the traditional ways that used low-level features in classification.
- Combining low-level and high-level features for enhanced lesion classification performance. However, using the combination approach will lose the advantages of characterisation and track back to diagnosis explanation. This is due to the lesion classification not only relies on the high-level features but also uses the low-level features in the diagnosis decision.

Conclusively, the competitive performance of the proposed framework for liver lesion characterisation relates to its underlying robust structure, which has been designed to better characterise liver lesions through using multiple ROIs and prior medical knowledge. Furthermore, the analogy between the radiological observations plus the exploit the relationships between low-level features and high-level features with respect to the ROI selection are the keys to its efficient lesion characteristics capturing. This is in contrast with majority of the baselines that attempt to characterise lesions through retrieving similar cases or using lesions its self without pay attention to the relation between lesion and liver. In addition, the advantage of the proposed framework to enhanced lesion classification performance, with the benefit of interpretable characterisation that support the diagnostic decision. In contrast with most existing research, which use black box low-level features only.

A comprehensive set of evaluation metrics were carried out over two datasets (Dataset I and overall dataset). This has been done to evaluate the overall framework for lesion characterisation performance and classification as well. A comparative evaluation of characterisation and classification accuracy is presented as well as benchmarking with the recent solid set of the related work. The experimental results and the overall benchmarking compared to the selected baselines indicated the proposed framework ability to characterised and classified liver lesions with higher accuracies. In addition, the statistical analysis confirmed that the results of the proposed approach is statistically highly-significant compared to other existing approaches for liver lesion classification with the benefit of interpretable characterisation.

However, this thesis concluded that the proposed lesion characterisation framework has many advantages; The first advantage is characterising liver lesion with high accuracy, and the larger number of high-level features compared to the other researchers. The second benefit is the ability to classify liver lesion with high accuracy with the advantage of interpreting the diagnosis decision in human understating and analogous to radiologist observation. The third advantage is the genericness aspect which makes it applicable with different datasets with different CT image settings and resolutions.

7.1 Limitations

The framework aims, in principle, to classify/ characterise the liver lesions from CT scans. However, there are a number of limitations of this proposed framework. Hence, this section summarises the proposed framework limitations.

In this thesis, the dataset was limited to four types of liver lesion. The proposed framework was developed based on the imaging characteristics of four lesions types in CT scan; two types are benign (Haemangioma and Cyst) and two types are malignant (HCC and Metastases). However, Haemangioma and Cyst are the common benign lesions, whereas HCC and Metastases are the common malignant lesions. Studies on

liver lesion and lesions in other body areas can be extended in future work to encourage continued development of relevant feature extraction methods.

The second limitation is the segmentation of vessels used in our framework. The vessels are extracted from the CT image to characterised lesion location. However, some lesions, especially closed to the vessels, have an enhanced rim with the same intensity of vessels that can be erroneously classified as a part of the vessels. This misclassification leads to a wrong characterisation of the lesion rim.

Machine learning algorithms are an important milestone that is associated with the lesion classification and characterisation filed in general. The main limitation is the performance of machine learning-based techniques depends on the training data size. A low number of training data usually results in a higher error in classification/characterisation. Therefore, we excluded any special case (not enough for training) in our dataset and focused on the most common features for characterisation and most common lesions types in terms of classification.

7.2 Future work

The work presented in this thesis can be extended in order to a better possible way to enhance the liver lesion characterisation and classification performance, the investigation can be continued in several aspects:

- **The fusion features from multiphase CT image.** In this work, the shift from portal phase to multiphase shows that the results of lesion classification are improved. However, more investigation is required about combine the extracted features from multiphase CT images, rather than the linear fusion. This will help to increase their combined performance.

- **Multimodality imaging studies.** This work is limited to developing lesion characterisation/classification methods in structural CT images. However, the consideration should be given to the use of other imaging modalities such as US/ MRI in addition to structural CT scan image as they may provide supplementary information for characterisation/classification of the lesion. In addition, explore the possible ways to fuse features from different modalities. This would fully verify the scalability aspect to achieve a better performance.

- **On semantic features.** The work presented in this thesis showed successfully enhancing on lesion classification with the advantage of interpreting the diagnosis decision in human understating and analogous to radiologist observation. In other word, the high-level features are capable of differentiating between malignant and benign lesion. However, adding more high-level features such as lesion filling/washout of contrast agent will not only improve the malignant/benign

classification but also help to classify lesion types. In addition, extend the study to diagnose other part of the body.

- **On low-level features.** We intend to explore more efficient low-level features such as 3D low-level features, which in the current work mainly use 2D low-level features. Furthermore, the future work will show whether 3D low-level features could represent lesion characteristics for building more efficient liver lesion characterisation/classification system.

- **On learning process.** We intend to investigate the use of relevance feedback to reduce the semantic gap between low-level and high-level features through allow the radiologist to evaluate the interpretation of lesion characteristics. This evaluation is used to enhance the high-level features prediction. In addition, investigate more advanced deep learning methods.

Appendix A

This appendix presents the evaluation and results of the proposed framework for liver segmentation, lesion detection and vessels extraction based on CT images. The framework performance is investigated by considering different types of evaluation measurements. Namely; Volumetric Overlap Error (**VOE**), Relative Volume Difference (**RVD**), True Positive Volume Fraction (**TPVF**), False Positive Volume Fraction (**FPVF**), Jaccard Similarity Metric (**JSM**) and Dice Similarity Coefficient (**DSC**). All the evaluations were performed over the entire dataset with a total number of 174 CT scan image.

Liver Segmentation Results

Table A.1 depicts the quantitative results of proposed liver segmentation based on preknowledge and after applying 3x3 median filter preprocessing step.

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
1	21.43	8.39	91.69	15.41	78.57	88.00
2	9.37	3.33	96.67	6.45	90.63	95.09
3	7.03	2.46	97.54	4.80	92.97	96.36
4	23.12	9.11	90.89	16.70	76.88	86.93
5	12.07	4.37	95.63	8.38	87.93	93.58
6	19.93	7.66	92.34	14.23	80.07	88.93
7	24.58	15.07	92.47	19.64	75.42	85.99
8	21.79	8.50	91.50	15.66	78.21	87.77
9	6.45	-6.45	93.55	0.00	93.55	96.67
10	6.45	-6.45	93.55	0.00	93.55	96.67
11	15.34	5.70	94.30	10.78	84.66	91.69
12	16.98	6.38	93.62	12.00	83.02	90.72
13	6.49	-6.49	93.51	0.00	93.51	96.65
14	13.15	4.80	95.20	9.17	86.85	92.96
15	4.57	-4.57	95.43	0.00	95.43	97.66
16	14.48	5.34	94.66	10.14	85.52	92.19
17	8.54	-8.54	91.46	0.00	91.46	95.54

Table A.1 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
18	3.43	-3.43	96.57	0.00	96.57	98.25
19	9.64	3.43	96.57	6.64	90.36	94.94
20	20.14	7.76	92.24	14.40	79.86	88.80
21	5.08	-5.08	94.92	0.00	94.92	97.39
22	13.83	5.08	94.92	9.67	86.17	92.57
23	9.32	3.31	96.69	6.41	90.68	95.11
24	16.94	6.37	93.63	11.97	83.06	90.74
25	6.63	-6.63	93.37	0.00	93.37	96.57
26	20.88	8.09	91.91	14.96	79.12	88.34
27	19.05	7.28	92.72	13.56	80.95	89.47
28	17.76	6.71	93.29	12.58	82.24	90.26
29	18.36	6.98	93.02	13.04	81.64	89.89
30	17.81	6.74	93.26	12.62	82.19	90.23
31	12.31	4.47	95.53	8.56	87.69	93.44
32	16.00	5.97	94.03	11.27	84.00	91.30
33	9.72	3.46	96.54	6.70	90.28	94.89
34	12.60	4.58	95.42	8.77	87.40	93.28
35	19.31	7.39	92.61	13.76	80.69	89.31
36	14.75	5.45	94.55	10.34	85.25	92.04
37	11.45	4.13	95.87	7.93	88.55	93.93
38	14.77	5.46	94.54	10.36	85.23	92.03
39	19.32	7.39	92.61	13.77	80.68	89.30
40	17.46	6.59	93.41	12.36	82.54	90.43
41	14.62	5.40	94.60	10.25	85.38	92.11
42	15.68	5.84	94.16	11.03	84.32	91.49
43	16.01	5.97	94.03	11.27	83.99	91.30
44	17.47	6.59	93.41	12.37	82.53	90.43
45	12.42	4.52	95.48	8.64	87.58	93.38
46	10.80	3.88	96.12	7.47	89.20	94.29
47	15.55	5.78	94.22	10.93	84.45	91.57
48	5.78	-5.78	94.22	0.00	94.22	97.02
49	13.70	5.02	94.98	9.57	86.30	92.65
50	17.37	6.55	93.45	12.29	82.63	90.49
51	9.56	4.38	97.06	7.01	90.44	94.98
52	18.96	9.35	93.71	14.30	81.04	89.52
53	8.87	4.04	97.29	6.49	91.13	95.36
54	8.31	3.77	97.47	6.07	91.69	95.66

Table A.1 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
55	9.11	4.16	97.21	6.67	90.89	95.23
56	9.53	4.36	97.07	6.99	90.47	95.00
57	11.57	5.38	96.38	8.53	88.43	93.86
58	15.69	7.54	94.93	11.72	84.31	91.49
59	14.54	6.92	95.35	10.82	85.46	92.16
60	16.20	7.81	94.75	12.11	83.80	91.19
61	13.29	6.26	95.79	9.86	86.71	92.88
62	19.69	9.77	93.43	14.88	80.31	89.08
63	11.19	5.19	96.51	8.25	88.81	94.07
64	12.19	5.69	96.17	9.01	87.81	93.51
65	16.02	7.71	94.81	11.98	83.98	91.29
66	11.31	5.25	96.47	8.34	88.69	94.00
67	19.31	9.55	93.58	14.58	80.69	89.31
68	15.81	7.60	94.89	11.81	84.19	91.42
69	13.54	6.39	95.70	10.05	86.46	92.74
70	16.02	7.72	94.81	11.98	83.98	91.29
71	10.31	4.75	96.81	7.58	89.69	94.56
72	19.23	9.51	93.61	14.52	80.77	89.36
73	12.68	5.95	96.00	9.39	87.32	93.23
74	11.80	5.50	96.30	8.71	88.20	93.73
75	15.49	7.43	95.01	11.56	84.51	91.61
76	13.50	6.37	95.71	10.02	86.50	92.76
77	9.76	4.47	96.99	7.16	90.24	94.87
78	8.36	3.79	97.45	6.11	91.64	95.64
79	11.24	5.21	96.49	8.29	88.76	94.04
80	16.35	7.89	94.69	12.23	83.65	91.10
81	8.54	3.88	97.39	6.24	91.46	95.54
82	11.65	5.42	96.36	8.60	88.35	93.81
83	12.30	5.75	96.13	9.10	87.70	93.44
84	12.30	5.75	96.13	9.09	87.70	93.45
85	12.02	5.61	96.23	8.88	87.98	93.60
86	21.07	10.58	92.89	16.00	78.93	88.22
87	14.22	6.75	95.46	10.57	85.78	92.35
88	19.08	9.42	93.67	14.40	80.92	89.45
89	21.67	10.93	92.65	16.48	78.33	87.85
90	24.81	12.85	91.36	19.05	75.19	85.84
91	18.17	8.90	94.01	13.67	81.83	90.01

Table A.1 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
92	11.08	5.13	96.55	8.16	88.92	94.14
93	14.06	6.67	95.52	10.45	85.94	92.44
94	12.12	5.66	96.20	8.95	87.88	93.55
95	20.46	10.22	93.13	15.51	79.54	88.60
96	12.13	5.67	96.19	8.97	87.87	93.54
97	9.77	4.48	96.99	7.17	90.23	94.86
98	10.40	4.79	96.78	7.65	89.60	94.52
99	14.50	6.90	95.36	10.79	85.50	92.18
100	9.58	4.39	97.05	7.03	90.42	94.97
101	20.12	10.02	93.26	15.23	79.88	88.81
102	22.87	11.65	92.16	17.45	77.13	87.09
103	8.91	4.06	97.27	6.52	91.09	95.34
104	18.77	9.24	93.79	14.15	81.23	89.64
105	26.12	13.69	90.79	20.14	73.88	84.98
106	10.82	5.00	96.64	7.97	89.18	94.28
107	13.36	6.30	95.77	9.91	86.64	92.84
108	15.98	7.69	94.83	11.95	84.02	91.31
109	9.11	4.16	97.21	6.67	90.89	95.23
110	18.36	9.01	93.94	13.82	81.64	89.90
111	13.10	6.16	95.86	9.71	86.90	92.99
112	8.72	3.97	97.33	6.38	91.28	95.44
113	21.65	10.92	92.66	16.47	78.35	87.86
114	17.26	8.39	94.36	12.95	82.74	90.56
115	11.09	5.14	96.55	8.17	88.91	94.13
116	16.13	7.77	94.77	12.06	83.87	91.23
117	19.95	9.92	93.33	15.09	80.05	88.92
118	21.45	10.80	92.74	16.30	78.55	87.99
119	21.59	10.89	92.68	16.42	78.41	87.90
120	20.80	9.33	92.51	15.38	79.20	88.39
121	29.42	14.17	88.62	22.38	70.58	82.75
122	32.21	15.91	87.23	24.74	67.79	80.80
123	25.41	11.83	90.50	19.07	74.59	85.45
124	20.60	9.22	92.60	15.22	79.40	88.51
125	15.02	6.43	94.83	10.90	84.98	91.88
126	28.40	13.57	89.11	21.54	71.60	83.45
127	16.57	7.18	94.23	12.08	83.43	90.97
128	38.46	20.11	83.85	30.19	61.54	76.19

Table A.1 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
129	33.53	16.76	86.55	25.88	66.47	79.85
130	15.63	6.73	94.60	11.36	84.37	91.52
131	32.63	16.18	87.01	25.10	67.37	80.50
132	28.80	13.80	88.92	21.86	71.20	83.18
133	14.98	6.42	94.85	10.87	85.02	91.90
134	29.07	13.96	88.79	22.09	70.93	83.00
135	14.86	6.36	94.90	10.77	85.14	91.97
136	19.49	8.65	93.06	14.35	80.51	89.20
137	21.52	9.70	92.21	15.95	78.48	87.94
138	25.26	11.75	90.57	18.95	74.74	85.54
139	15.53	6.68	94.64	11.28	84.47	91.58
140	19.49	8.65	93.06	14.35	80.51	89.20
141	19.59	8.70	93.02	14.42	80.41	89.14
142	15.39	6.61	94.69	11.18	84.61	91.66
143	22.81	10.39	91.66	16.97	77.19	87.13
144	20.02	8.92	92.84	14.76	79.98	88.87
145	38.08	19.84	84.07	29.85	61.92	76.48
146	15.88	6.85	94.50	11.55	84.12	91.38
147	23.00	10.50	91.57	17.13	77.00	87.01
148	16.87	7.33	94.12	12.31	83.13	90.79
149	25.94	12.13	90.26	19.50	74.06	85.10
150	28.57	13.66	89.03	21.67	71.43	83.34
151	40.30	21.44	82.78	31.83	59.70	74.77
152	17.40	7.59	93.90	12.72	82.60	90.47
153	24.25	11.18	91.02	18.13	75.75	86.20
154	30.12	14.60	88.28	22.97	69.88	82.27
155	18.78	8.28	93.35	13.79	81.22	89.64
156	18.81	8.30	93.34	13.82	81.19	89.62
157	14.90	6.37	94.88	10.80	85.10	91.95
158	18.80	8.29	93.34	13.81	81.20	89.63
159	22.08	10.00	91.97	16.39	77.92	87.59
160	38.37	20.05	83.91	30.10	61.63	76.26
161	23.77	10.92	91.24	17.74	76.23	86.51
162	39.32	20.73	83.36	30.96	60.68	75.53
163	20.88	9.36	92.48	15.44	79.12	88.35
164	16.07	6.94	94.43	11.70	83.93	91.26
165	26.30	12.34	90.09	19.80	73.70	84.86

Table A.1 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
166	21.45	9.67	92.24	15.89	78.55	87.99
167	16.30	7.05	94.34	11.87	83.70	91.13
168	24.23	11.17	91.03	18.11	75.77	86.22
169	19.50	8.65	93.05	14.36	80.50	89.19
170	32.24	15.92	87.21	24.77	67.76	80.78
171	16.76	7.28	94.16	12.23	83.24	90.85
172	22.80	10.39	91.66	16.97	77.20	87.13
173	18.82	8.31	93.33	13.83	81.18	89.61
174	16.30	7.05	94.34	11.87	83.70	91.13
Mean	**17.15**	**7.35**	**93.68**	**12.48**	**82.85**	**90.45**

Table A.1: Overview of the proposed liver segmentation results after applying preprocessing step obtained over the entire dataset

Liver Lesion Detection Results

Table A.2 presents the quantitative results of AFCM approach for liver lesion detection when applying 3x3 median filter preprocessing step.

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
1	20.63	-8.47	84.75	7.41	79.37	88.50
2	23.05	-23.05	76.95	0.00	76.95	86.97
3	17.44	10.04	94.98	13.68	82.56	90.44
4	18.91	14.96	96.26	16.27	81.09	89.56
5	15.64	11.99	97.00	13.38	84.36	91.51
6	38.29	-4.96	74.43	21.68	61.71	76.33
7	13.35	9.21	97.12	11.06	86.65	92.85
8	25.64	2.71	86.45	15.83	74.36	85.29
9	33.60	-3.39	78.45	18.79	66.40	79.81
10	13.65	8.48	96.61	10.95	86.35	92.68
11	28.11	-4.29	81.86	14.48	71.89	83.65
12	17.29	1.74	91.32	10.23	82.71	90.54
13	14.14	9.57	96.81	11.64	85.86	92.39
14	19.25	-8.11	85.72	6.71	80.75	89.35
15	12.49	10.20	98.10	10.98	87.51	93.34
16	25.22	24.27	95.96	22.78	74.78	85.57
17	24.02	16.09	93.30	19.63	75.98	86.35
18	8.76	-7.25	91.96	0.85	91.24	95.42

Table A.2 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
19	25.84	17.83	92.75	21.28	74.16	85.16
20	31.75	35.02	95.34	29.39	68.25	81.13
21	13.13	1.46	93.65	7.70	86.87	92.97
22	16.90	6.35	93.65	11.94	83.10	90.77
23	23.75	1.29	87.09	14.02	76.25	86.53
24	16.88	-0.84	90.40	8.84	83.12	90.78
25	31.19	29.33	93.48	27.72	68.81	81.53
26	14.45	13.61	98.49	13.31	85.55	92.21
27	25.23	7.78	88.89	17.52	74.77	85.57
28	20.33	17.25	96.33	17.84	79.67	88.68
29	9.60	-4.22	92.95	2.95	90.40	94.96
30	25.64	14.57	91.51	20.13	74.36	85.30
31	22.37	2.84	88.65	13.80	77.63	87.40
32	27.62	-12.05	78.92	10.27	72.38	83.98
33	13.39	-7.26	89.45	3.55	86.61	92.82
34	18.91	-2.27	88.54	9.40	81.09	89.56
35	26.77	1.35	85.12	16.02	73.23	84.55
36	16.46	5.26	93.42	11.24	83.54	91.03
37	14.48	-8.78	88.15	3.37	85.52	92.20
38	24.31	13.03	91.78	18.80	75.69	86.17
39	15.26	-7.24	88.42	4.68	84.74	91.74
40	23.60	1.17	87.13	13.88	76.40	86.62
41	11.54	-6.44	90.86	2.89	88.46	93.88
42	36.99	-3.59	75.92	21.25	63.01	77.31
43	30.02	-1.65	81.66	16.97	69.98	82.34
44	32.77	3.87	81.96	21.10	67.23	80.40
45	11.26	4.06	95.94	7.80	88.74	94.04
46	21.15	24.21	98.85	20.42	78.85	88.17
47	17.74	6.71	93.29	12.57	82.26	90.26
48	14.72	2.68	93.29	9.15	85.28	92.06
49	27.91	1.72	84.50	16.93	72.09	83.78
50	26.21	2.78	86.10	16.23	73.79	84.92
51	20.43	7.89	92.11	14.62	79.57	88.62
52	11.14	-3.85	92.29	4.01	88.86	94.10
53	18.51	-6.58	86.85	7.04	81.49	89.80
54	36.65	-20.17	69.74	12.64	63.35	77.56
55	15.63	5.82	94.18	10.99	84.37	91.52

Table A.2 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
56	14.53	8.15	95.92	11.30	85.47	92.17
57	20.41	7.87	92.13	14.60	79.59	88.64
58	34.42	2.00	80.00	21.57	65.58	79.21
59	18.67	-1.36	89.09	9.67	81.33	89.71
60	15.25	1.85	92.59	9.09	84.75	91.74
61	19.23	2.39	90.43	11.68	80.77	89.36
62	12.97	4.73	95.27	9.04	87.03	93.07
63	14.78	5.43	94.52	10.35	85.22	92.02
64	20.59	-2.16	87.57	10.50	79.41	88.53
65	16.39	-2.26	90.05	7.87	83.61	91.08
66	14.28	5.26	94.74	10.00	85.72	92.31
67	21.62	8.42	91.58	15.53	78.38	87.88
68	47.68	23.30	76.70	37.79	52.32	68.70
69	16.62	6.23	93.77	11.73	83.38	90.94
70	22.26	8.71	91.29	16.03	77.74	87.48
71	18.69	1.13	90.20	10.81	81.31	89.69
72	13.44	1.61	93.54	7.94	86.56	92.79
73	17.58	6.64	93.36	12.45	82.42	90.36
74	18.93	14.98	96.25	16.29	81.07	89.55
75	13.20	1.58	93.67	7.79	86.80	92.93
76	19.08	11.14	94.43	15.03	80.92	89.45
77	17.17	6.46	93.54	12.14	82.83	90.61
78	38.80	60.74	98.99	38.42	61.20	75.93
79	14.80	-1.28	91.42	7.40	85.20	92.01
80	14.27	5.26	94.74	9.99	85.73	92.32
81	10.03	3.58	96.42	6.92	89.97	94.72
82	20.31	7.83	92.17	14.53	79.69	88.69
83	16.87	-5.96	88.09	6.33	83.13	90.79
84	9.62	-7.59	91.34	1.15	90.38	94.95
85	15.63	5.82	94.18	10.99	84.37	91.52
86	18.94	7.23	92.77	13.48	81.06	89.54
87	12.87	-6.06	90.30	3.87	87.13	93.12
88	17.91	6.78	93.22	12.70	82.09	90.17
89	17.85	6.75	93.25	12.65	82.15	90.20
90	19.93	11.72	94.14	15.73	80.07	88.93
91	16.24	-6.29	88.29	5.78	83.76	91.16
92	16.34	-5.76	88.48	6.11	83.66	91.10

Table A.2 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
93	13.06	-9.51	88.59	2.10	86.94	93.01
94	20.22	-6.55	85.85	8.13	79.78	88.75
95	12.71	-3.70	91.49	4.99	87.29	93.22
96	12.96	3.85	94.87	8.65	87.04	93.07
97	13.74	5.04	94.96	9.60	86.26	92.62
98	22.32	-17.22	79.91	3.47	77.68	87.44
99	13.39	-10.94	87.74	1.47	86.61	92.82
100	14.57	5.00	94.45	10.05	85.43	92.14
101	22.92	1.24	87.60	13.48	77.08	87.06
102	15.80	-4.00	89.59	6.67	84.20	91.42
103	11.85	-4.11	91.77	4.29	88.15	93.70
104	45.28	-2.56	69.83	28.34	54.72	70.74
105	13.88	5.10	94.90	9.70	86.12	92.54
106	18.65	10.84	94.58	14.67	81.35	89.72
107	31.38	13.23	86.77	23.36	68.62	81.39
108	19.27	-6.86	86.27	7.37	80.73	89.34
109	13.15	-9.32	88.63	2.26	86.85	92.96
110	11.33	-2.05	93.03	5.02	88.67	93.99
111	17.50	1.29	90.99	10.16	82.50	90.41
112	12.45	-1.20	92.80	6.07	87.55	93.36
113	20.10	7.73	92.27	14.36	79.90	88.83
114	16.97	6.38	93.62	11.99	83.03	90.73
115	24.81	9.91	90.09	18.03	75.19	85.84
116	12.67	-2.83	91.92	5.41	87.33	93.24
117	39.10	-6.10	73.39	21.84	60.90	75.70
118	18.82	7.17	92.83	13.39	81.18	89.61
119	11.33	-6.46	90.96	2.76	88.67	93.99
120	11.43	0.58	94.21	6.33	88.57	93.94
121	15.08	-1.69	91.07	7.36	84.92	91.85
122	22.94	-4.99	84.87	10.67	77.06	87.04
123	17.16	6.46	93.54	12.13	82.84	90.62
124	12.27	-2.73	92.19	5.22	87.73	93.47
125	18.99	6.96	92.63	13.40	81.01	89.51
126	12.90	-2.59	91.90	5.66	87.10	93.10
127	18.55	10.77	94.61	14.59	81.45	89.78
128	26.20	16.30	91.85	21.03	73.80	84.93
129	16.20	15.86	98.41	15.05	83.80	91.19

Table A.2 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
130	49.14	-18.34	61.25	25.00	50.86	67.43
131	10.93	7.19	97.60	8.94	89.07	94.22
132	28.50	1.08	83.83	17.06	71.50	83.38
133	25.29	-6.19	82.88	11.65	74.71	85.52
134	17.67	11.01	95.28	14.17	82.33	90.31
135	15.86	3.09	92.80	9.98	84.14	91.39
136	15.29	5.68	94.32	10.74	84.71	91.72
137	12.65	10.13	97.97	11.04	87.35	93.25
138	17.86	-6.33	87.34	6.76	82.14	90.19
139	29.94	19.31	90.35	24.27	70.06	82.39
140	28.63	18.23	90.89	23.12	71.37	83.30
141	18.60	7.08	92.92	13.22	81.40	89.74
142	14.37	-8.71	88.24	3.34	85.63	92.26
143	18.32	10.62	94.69	14.40	81.68	89.92
144	28.64	18.23	90.88	23.13	71.36	83.29
145	19.49	19.78	98.02	18.16	80.51	89.20
146	20.55	12.15	93.93	16.25	79.45	88.55
147	13.13	10.57	97.89	11.47	86.87	92.98
148	18.65	10.84	94.58	14.67	81.35	89.72
149	21.54	22.37	97.76	20.11	78.46	87.93
150	19.65	-16.65	81.68	2.00	80.35	89.10
151	17.56	-12.06	84.93	3.43	82.44	90.38
152	22.84	13.78	93.11	18.17	77.16	87.10
153	13.15	10.13	97.67	11.32	86.85	92.96
154	17.32	-1.86	89.68	8.62	82.68	90.52
155	29.79	30.21	94.96	27.07	70.21	82.50
156	17.37	-15.92	83.29	0.95	82.63	90.49
157	22.27	16.26	94.58	18.64	77.73	87.47
158	18.53	10.76	94.62	14.57	81.47	89.79
159	23.79	18.44	94.47	20.24	76.21	86.50
160	17.73	6.70	93.30	12.56	82.27	90.27
161	31.73	25.51	91.50	27.10	68.27	81.14
162	20.92	12.41	93.80	16.56	79.08	88.32
163	26.70	-26.61	73.34	0.06	73.30	84.60
164	17.27	6.51	93.49	12.22	82.73	90.55
165	49.79	-19.78	60.25	24.90	50.21	66.86
166	16.83	14.96	97.61	15.10	83.17	90.81

Table A.2 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
167	23.55	10.63	91.26	17.51	76.45	86.65
168	20.77	12.30	93.85	16.43	79.23	88.41
169	10.50	-9.93	89.77	0.33	89.50	94.46
170	22.36	-19.58	78.85	1.95	77.64	87.41
171	20.47	-18.57	80.37	1.30	79.53	88.60
172	26.15	-22.28	75.49	2.87	73.85	84.96
173	23.46	17.33	94.22	19.69	76.54	86.71
174	45.33	-17.80	64.40	21.66	54.67	70.69
Mean	**20.64**	**3.50**	**89.58**	**12.87**	**79.36**	**88.07**

Table A.2: Overview of the proposed liver lesion detection results when applying preprocessing step obtained over the entire dataset

Liver Vessels Extraction Results

Table A.3 shows the quantitative results of the proposed liver vessels extraction approach based on 3x3 Gaussian filter preprocessing step.

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
1	8.15	5.23	98.26	6.63	91.85	95.75
2	34.38	14.87	85.13	25.88	65.62	79.24
3	9.38	1.42	95.75	5.59	90.62	95.08
4	27.97	-10.28	79.43	11.46	72.03	83.74
5	10.29	-1.54	93.85	4.69	89.71	94.57
6	8.18	2.88	97.12	5.61	91.82	95.74
7	6.86	3.09	97.94	4.99	93.14	96.45
8	20.09	20.53	97.95	18.73	79.91	88.83
9	22.40	-8.07	83.86	8.78	77.60	87.38
10	20.63	7.75	91.93	14.68	79.37	88.50
11	16.33	13.56	97.29	14.33	83.67	91.11
12	21.45	8.34	91.66	15.40	78.55	87.99
13	9.76	-1.13	94.33	4.59	90.24	94.87
14	21.33	12.70	93.65	16.90	78.67	88.06
15	6.18	4.35	98.91	5.21	93.82	96.81
16	19.04	10.05	93.97	14.61	80.96	89.48
17	27.48	11.21	88.79	20.17	72.52	84.07
18	14.75	11.22	97.20	12.61	85.25	92.04
19	9.17	4.21	97.20	6.73	90.83	95.19

Table A.3 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
20	16.98	1.24	91.29	9.83	83.02	90.72
21	23.70	9.38	90.62	17.16	76.30	86.56
22	25.74	-9.38	81.23	10.36	74.26	85.23
23	10.91	5.06	96.61	8.04	89.09	94.23
24	15.01	9.44	96.22	12.07	84.99	91.89
25	11.79	-1.13	93.21	5.73	88.21	93.74
26	12.08	-1.16	93.03	5.88	87.92	93.57
27	15.46	1.54	92.32	9.07	84.54	91.62
28	19.09	10.07	93.96	14.64	80.91	89.45
29	19.58	7.51	92.49	13.97	80.42	89.15
30	12.89	4.70	95.30	8.98	87.11	93.11
31	15.08	9.49	96.20	12.13	84.92	91.85
32	15.83	5.90	94.10	11.14	84.17	91.40
33	15.11	7.71	95.37	11.46	84.89	91.83
34	9.98	-1.49	94.04	4.54	90.02	94.75
35	22.37	8.76	91.24	16.11	77.63	87.41
36	11.91	1.83	94.52	7.17	88.09	93.67
37	9.77	3.48	96.52	6.73	90.23	94.86
38	13.50	2.94	94.12	8.56	86.50	92.76
39	13.70	1.14	93.17	7.87	86.30	92.65
40	13.44	10.09	97.48	11.45	86.56	92.80
41	27.35	-14.68	77.98	8.60	72.65	84.16
42	25.98	16.13	91.93	20.84	74.02	85.07
43	10.10	8.63	98.77	9.08	89.90	94.68
44	24.21	9.63	90.37	17.56	75.79	86.23
45	12.56	5.29	95.77	9.04	87.44	93.30
46	34.94	21.29	87.23	28.08	65.06	78.83
47	12.58	8.60	97.30	10.41	87.42	93.29
48	15.32	7.37	95.09	11.44	84.68	91.71
49	24.54	15.04	92.48	19.61	75.46	86.01
50	14.15	3.09	93.82	9.00	85.85	92.39
51	41.24	29.85	85.07	34.48	58.76	74.03
52	17.28	-9.03	86.45	4.96	82.72	90.54
53	16.13	9.18	95.41	12.61	83.87	91.23
54	22.89	13.82	93.09	18.21	77.11	87.07
55	13.29	-4.63	90.73	4.86	86.71	92.88
56	25.75	13.52	90.99	19.85	74.25	85.22

Table A.3 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
57	14.81	-10.13	87.34	2.82	85.19	92.00
58	27.47	17.30	91.35	22.12	72.53	84.08
59	17.86	6.76	93.24	12.66	82.14	90.19
60	22.01	8.60	91.40	15.83	77.99	87.64
61	12.46	4.53	95.47	8.67	87.54	93.35
62	19.47	7.46	92.54	13.88	80.53	89.21
63	19.05	7.27	92.73	13.56	80.95	89.47
64	12.69	-4.42	91.16	4.62	87.31	93.22
65	17.43	9.47	94.74	13.46	82.57	90.45
66	21.35	8.30	91.70	15.33	78.65	88.05
67	18.55	7.05	92.95	13.18	81.45	89.78
68	22.74	8.93	91.07	16.40	77.26	87.17
69	26.27	10.62	89.38	19.19	73.73	84.88
70	16.37	9.33	95.33	12.80	83.63	91.09
71	25.08	10.04	89.96	18.24	74.92	85.66
72	17.19	9.87	95.07	13.47	82.81	90.60
73	12.06	-1.42	92.92	5.75	87.94	93.58
74	17.33	6.53	93.47	12.26	82.67	90.51
75	29.25	16.89	89.87	23.12	70.75	82.87
76	14.69	5.43	94.57	10.30	85.31	92.07
77	14.25	5.25	94.75	9.97	85.75	92.33
78	17.38	6.55	93.45	12.30	82.62	90.48
79	13.96	5.13	94.87	9.76	86.04	92.50
80	16.36	6.12	93.88	11.53	83.64	91.09
81	22.19	10.08	91.93	16.49	77.81	87.52
82	18.25	6.93	93.07	12.95	81.75	89.96
83	12.84	4.68	95.32	8.94	87.16	93.14
84	14.39	5.31	94.69	10.08	85.61	92.25
85	16.76	6.29	93.71	11.84	83.24	90.85
86	35.68	15.60	84.40	27.00	64.32	78.29
87	15.78	5.88	94.12	11.10	84.22	91.43
88	16.17	6.04	93.96	11.39	83.83	91.21
89	19.30	7.38	92.62	13.75	80.70	89.32
90	13.19	4.82	95.18	9.19	86.81	92.94
91	36.00	15.79	84.21	27.28	64.00	78.05
92	16.52	6.19	93.81	11.65	83.48	91.00
93	21.81	8.51	91.49	15.68	78.19	87.76

Table A.3 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
94	19.87	7.63	92.37	14.18	80.13	88.97
95	18.71	7.13	92.87	13.30	81.29	89.68
96	23.71	9.39	90.61	17.16	76.29	86.55
97	13.09	4.78	95.22	9.13	86.91	92.99
98	14.68	5.42	94.58	10.29	85.32	92.08
99	16.84	6.32	93.68	11.89	83.16	90.81
100	24.78	9.90	90.10	18.01	75.22	85.86
101	30.19	12.60	87.40	22.38	69.81	82.22
102	18.27	6.94	93.06	12.97	81.73	89.94
103	13.21	4.83	95.17	9.21	86.79	92.93
104	24.10	11.12	91.10	18.02	75.90	86.30
105	21.56	11.62	93.03	16.66	78.44	87.92
106	14.41	5.31	94.69	10.09	85.59	92.24
107	19.45	10.30	93.82	14.94	80.55	89.23
108	22.15	-5.27	85.24	10.02	77.85	87.54
109	24.52	-8.90	82.20	9.77	75.48	86.02
110	32.35	17.49	87.76	25.31	67.65	80.70
111	30.80	5.34	83.98	20.28	69.20	81.79
112	16.45	-7.20	87.76	5.43	83.55	91.04
113	13.14	-1.99	92.04	6.09	86.86	92.97
114	16.20	6.05	93.95	11.42	83.80	91.19
115	19.92	7.66	92.34	14.23	80.08	88.94
116	22.54	8.84	91.16	16.25	77.46	87.30
117	15.54	5.78	94.22	10.93	84.46	91.58
118	21.63	-7.77	84.46	8.42	78.37	87.87
119	20.04	7.71	92.29	14.32	79.96	88.86
120	18.00	6.82	93.18	12.77	82.00	90.11
121	20.38	9.13	92.70	15.05	79.62	88.65
122	23.62	-4.62	84.61	11.30	76.38	86.61
123	17.63	11.91	95.71	14.48	82.37	90.33
124	12.04	4.36	95.64	8.36	87.96	93.59
125	12.79	4.66	95.34	8.90	87.21	93.17
126	13.54	4.96	95.04	9.45	86.46	92.74
127	20.10	-7.18	85.64	7.73	79.90	88.83
128	22.46	11.13	92.21	17.02	77.54	87.35
129	19.50	7.47	92.53	13.90	80.50	89.20
130	18.43	7.00	93.00	13.09	81.57	89.85

Table A.3 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
131	15.42	5.73	94.27	10.84	84.58	91.65
132	13.05	-4.55	90.91	4.76	86.95	93.02
133	16.06	6.00	94.00	11.31	83.94	91.27
134	20.99	11.26	93.24	16.19	79.01	88.27
135	13.08	-4.56	90.88	4.78	86.92	93.00
136	8.22	-4.65	93.49	1.95	91.78	95.72
137	15.99	5.97	94.03	11.26	84.01	91.31
138	21.02	-7.54	84.93	8.15	78.98	88.25
139	19.55	-6.97	86.06	7.49	80.45	89.17
140	12.43	4.52	95.48	8.65	87.57	93.37
141	8.59	-4.87	93.19	2.05	91.41	95.51
142	12.85	4.69	95.31	8.95	87.15	93.13
143	24.57	9.80	90.20	17.84	75.43	85.99
144	20.02	7.70	92.30	14.30	79.98	88.88
145	24.96	6.40	88.49	16.83	75.04	85.74
146	17.83	9.31	94.41	13.63	82.17	90.21
147	25.70	17.84	92.86	21.20	74.30	85.26
148	19.35	7.40	92.60	13.79	80.65	89.29
149	15.88	-5.59	88.82	5.92	84.12	91.37
150	19.36	10.24	93.85	14.87	80.64	89.28
151	11.93	5.00	96.00	8.58	88.07	93.66
152	19.74	11.59	94.21	15.58	80.26	89.05
153	12.24	4.44	95.56	8.51	87.76	93.48
154	22.88	-8.26	83.49	9.00	77.12	87.08
155	12.52	-4.36	91.29	4.55	87.48	93.32
156	15.69	5.84	94.16	11.04	84.31	91.49
157	13.18	4.82	95.18	9.19	86.82	92.95
158	14.50	-5.08	89.84	5.35	85.50	92.18
159	22.63	-12.86	81.64	6.32	77.37	87.24
160	13.95	5.13	94.87	9.76	86.05	92.50
161	15.10	5.60	94.40	10.60	84.90	91.83
162	16.17	2.98	92.56	10.11	83.83	91.20
163	14.38	-6.80	89.12	4.38	85.62	92.25
164	11.76	1.57	94.49	6.97	88.24	93.75
165	18.31	-11.85	84.59	4.03	81.69	89.92
166	13.38	-5.96	90.07	4.23	86.62	92.83
167	17.53	-7.70	86.92	5.84	82.47	90.40

Table A.3 continued from previous page

Case No.	VOE %	RVD %	TPVF %	FPVF %	Jaccard %	Dice %
168	17.05	9.77	95.11	13.36	82.95	90.68
169	12.77	4.65	95.35	8.89	87.23	93.18
170	15.74	5.86	94.14	11.07	84.26	91.46
171	12.91	4.71	95.29	8.99	87.09	93.10
172	12.69	-4.42	91.17	4.62	87.31	93.23
173	14.59	5.39	94.61	10.23	85.41	92.13
174	16.48	6.17	93.83	11.63	83.52	91.02
Mean	**17.94**	**5.04**	**92.24**	**11.89**	**82.06**	**90.03**

Table A.3: Overview of the proposed liver vessels extraction results based on 3x3 Gaussian filter preprocessing step obtained over the entire dataset

Appendix B

Results of Liver lesion Characterisation/Classification

This appendix displays the liver lesion characterisation/classification results for the used whole dataset (Dataset I and Dataset II), using the Logistic Regression (LR) and Linear Discriminant Analysis (LDA) Classifier.

Results of Lesion Characterisation

This section presents the evaluation results of the proposed multiple ROIs framework to characterise liver lesion using portal phase and multiphase CT images. Two types of classifiers (Logistic Regression (LR) and Linear Discriminant Analysis (LDA)) were used to characterise liver lesion. In addition, the tenfold cross-validation method was adopted to evaluate the framework.

Logistic Regression (LR) Classifier

Figure B.1 depicts the lesion characterisation accuracy comparison between single ROI and Multiple ROIs based on tenfold cross-validation method and LR classifier, using portal phase CT image.

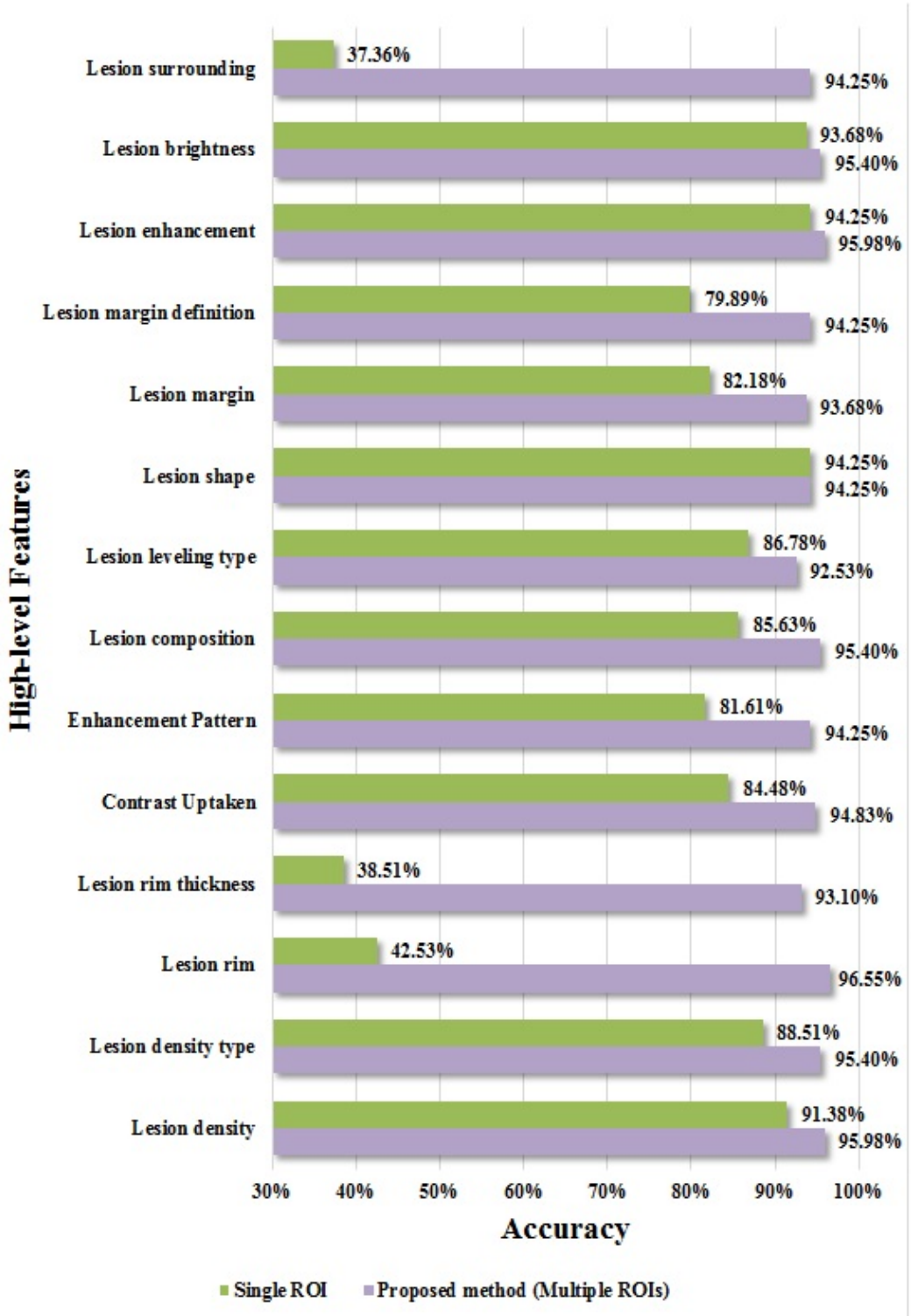


Figure B.1: The accuracy comparison between single ROI and Multiple ROIs to predict the high-level features using the **portal** phase CT image where **tenfold cross-validation** method and LR classifier was adopted.

Figure B.2 shows the lesion characterisation accuracy comparison between single ROI and Multiple ROIs based on tenfold cross-validation method and LR classifier, using Multiphase CT image.

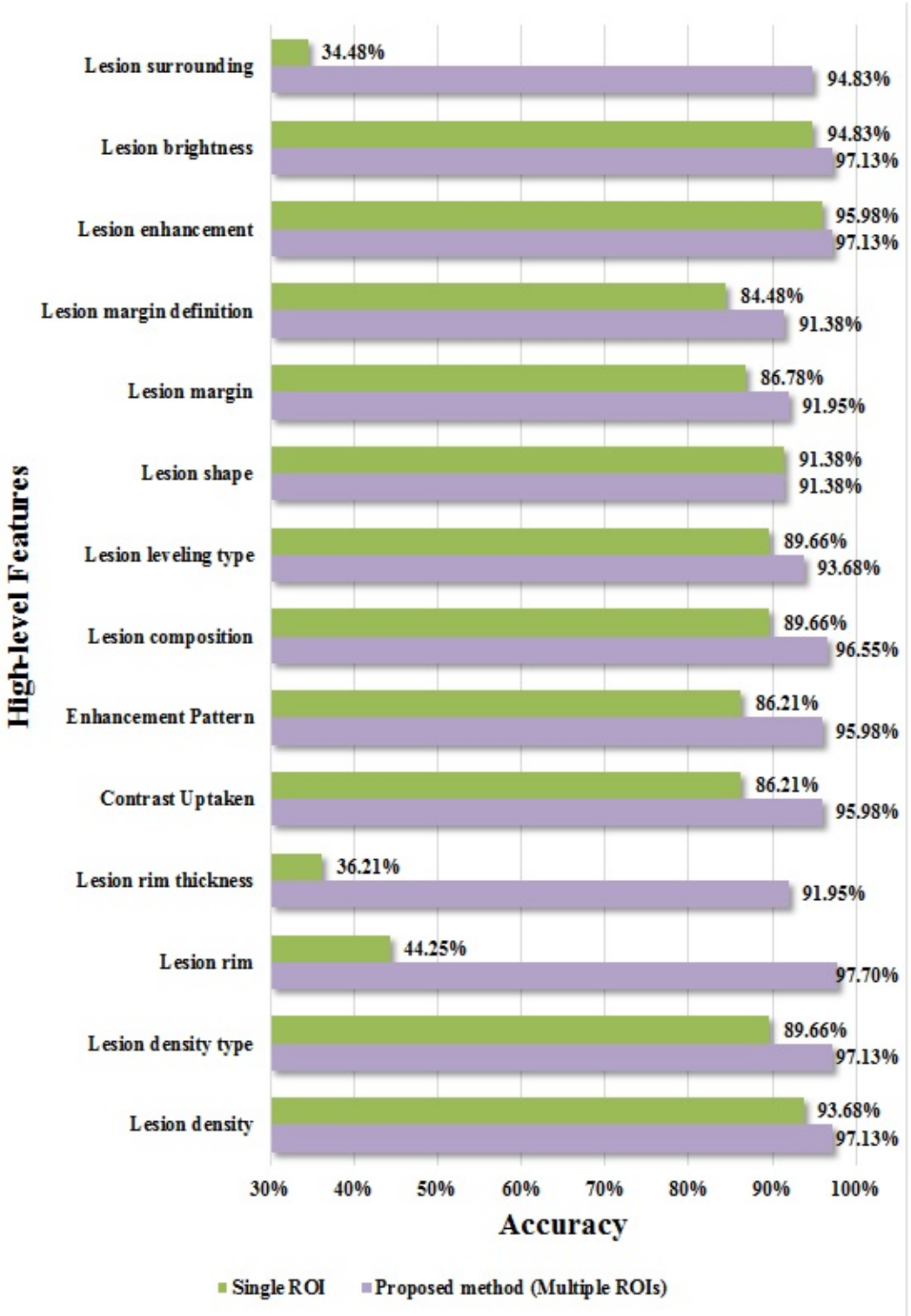


Figure B.2: The accuracy comparison between single ROI and Multiple ROIs to predict the high-level features using the **multiphase** CT image where **tenfold cross-validation** method and LR classifier was adopted.

Figure B.3 illustrates the evaluation of proposed liver lesion characterisation framework based on Multiple ROIs comparison by using portal phase and multiphase CT images. The LR classifier was used to predict the high-level features and the accuracy of prediction was estimated using a tenfold cross-validation.

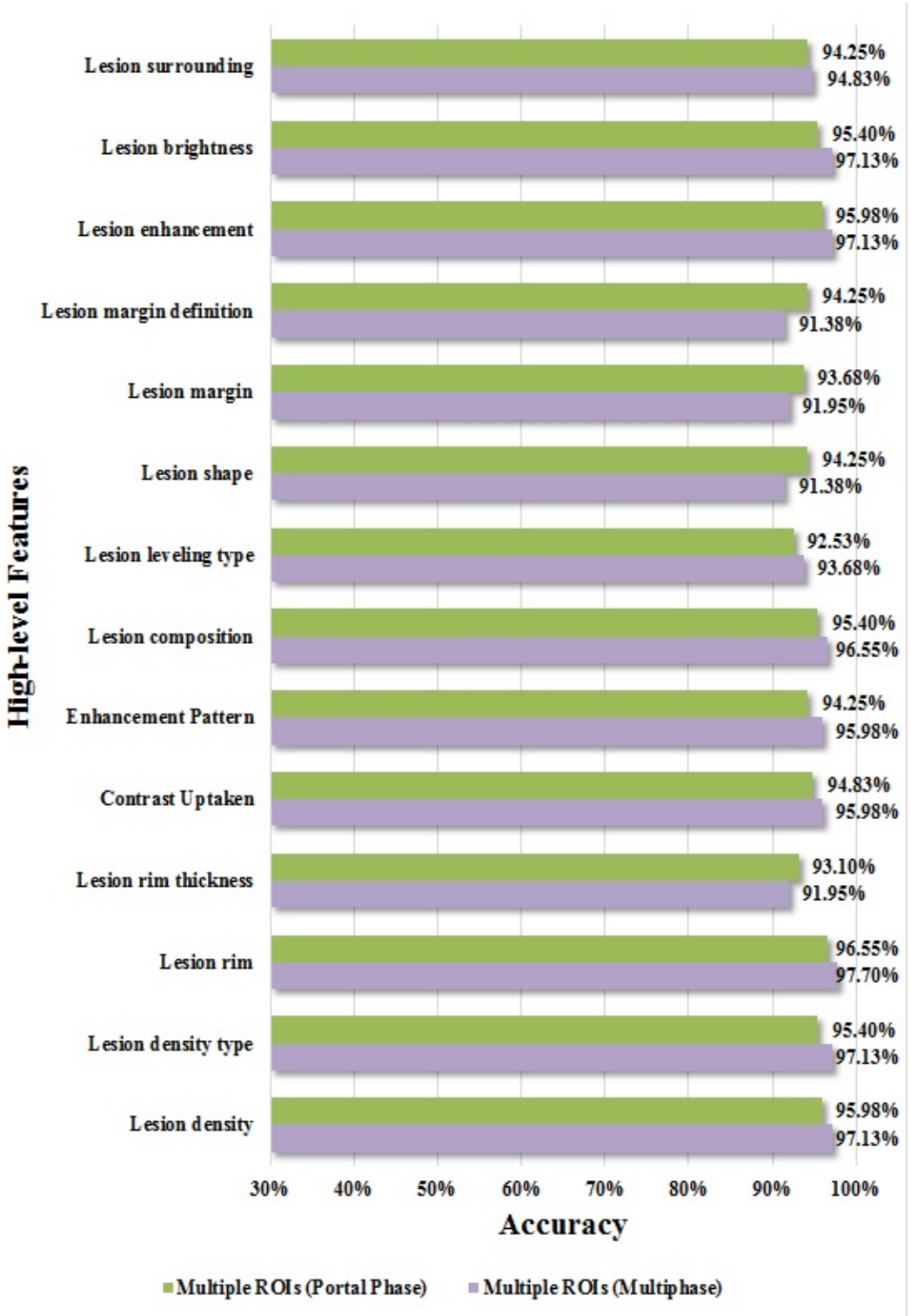


Figure B.3: The accuracy comparison of the proposed Multiple ROIs to predict the high-level features by using **portal** phase and **multiphase** CT image where **tenfold cross-validation** method and LR classifier was adopted.

Linear Discriminant Analysis (LDA) Classifier

Figure B.4 depicts the lesion characterisation accuracy comparison between single ROI and Multiple ROIs based on **tenfold cross-validation** method and LDA classifier, using **portal** phase CT image.

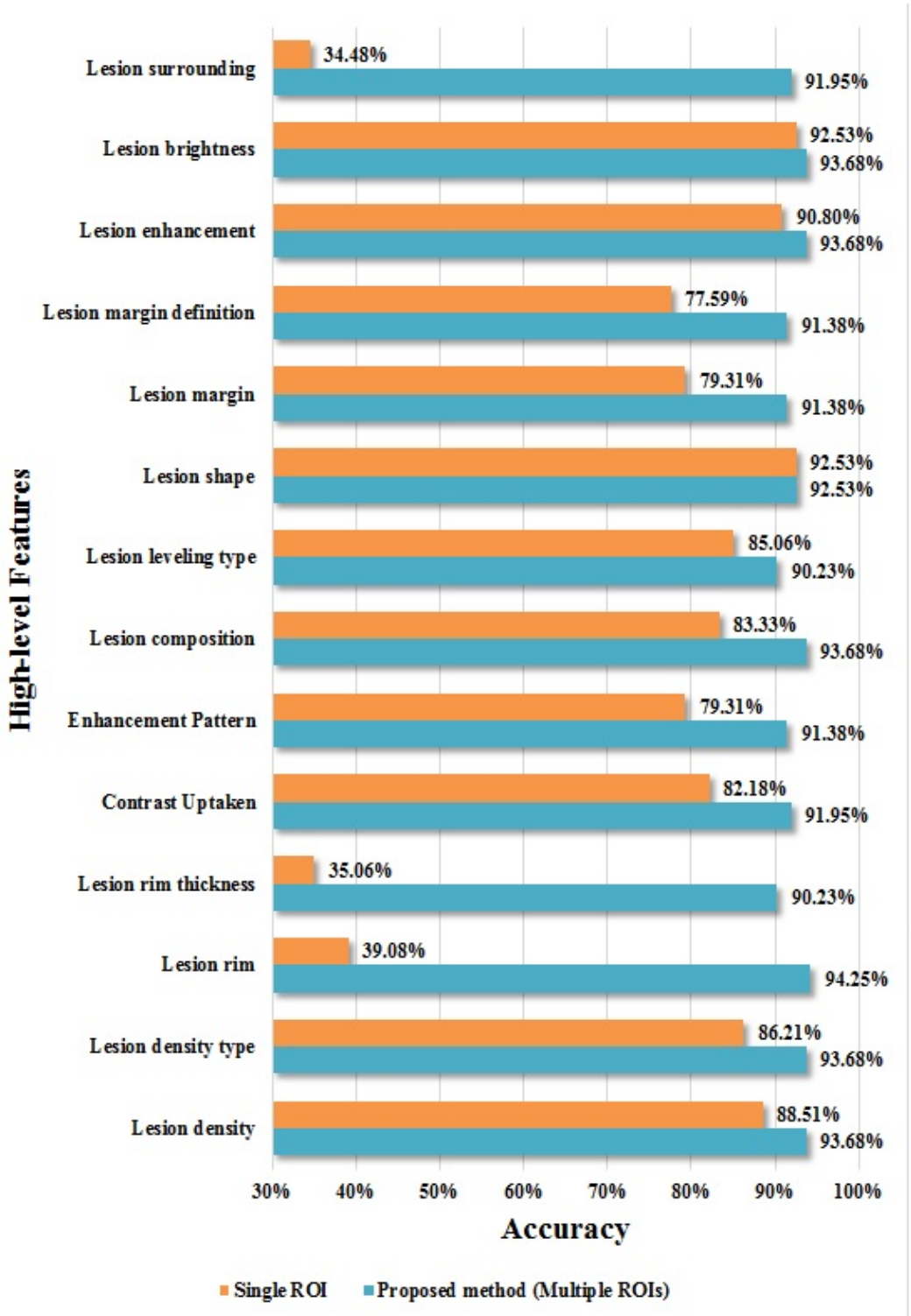


Figure B.4: The accuracy comparison between single ROI and Multiple ROIs to predict the high-level features using the **portal** phase CT image where **tenfold cross-validation** method and LDA classifier was adopted.

Figure B.5 shows the lesion characterisation accuracy comparison between single ROI and Multiple ROIs based on tenfold cross-validation method and LDA classifier, using Multiphase CT image.

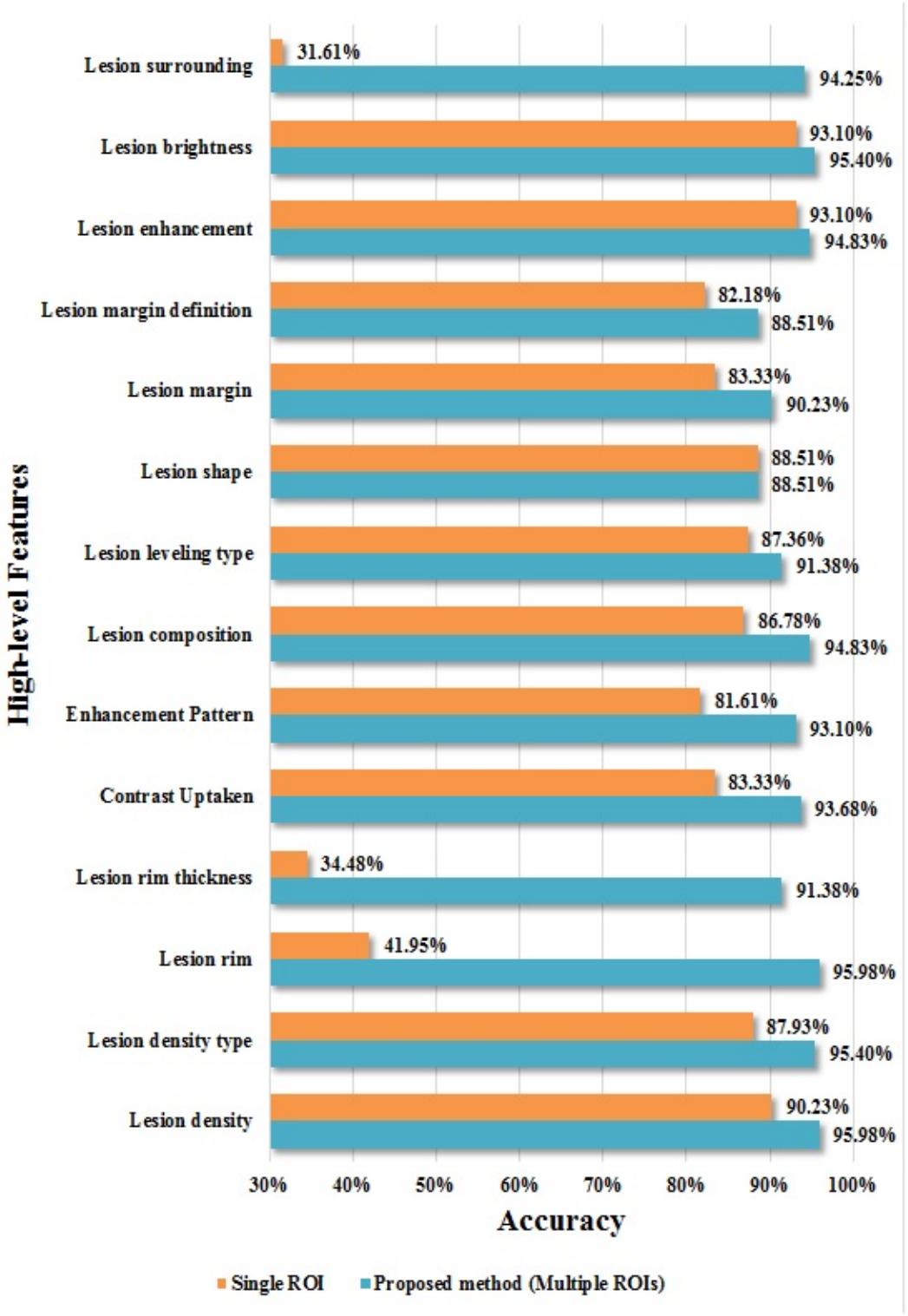


Figure B.5: The accuracy comparison between single ROI and Multiple ROIs to predict the high-level features using the **multiphase** CT image where **tenfold cross-validation** method and LDA classifier was adopted.

Figure B.6 illustrates the evaluation of proposed liver lesion characterisation framework based on Multiple ROIs comparison using portal phase and multiphase CT images. The LDA classifier was used to predict the high-level features and the accuracy of prediction was estimated using a tenfold cross-validation.

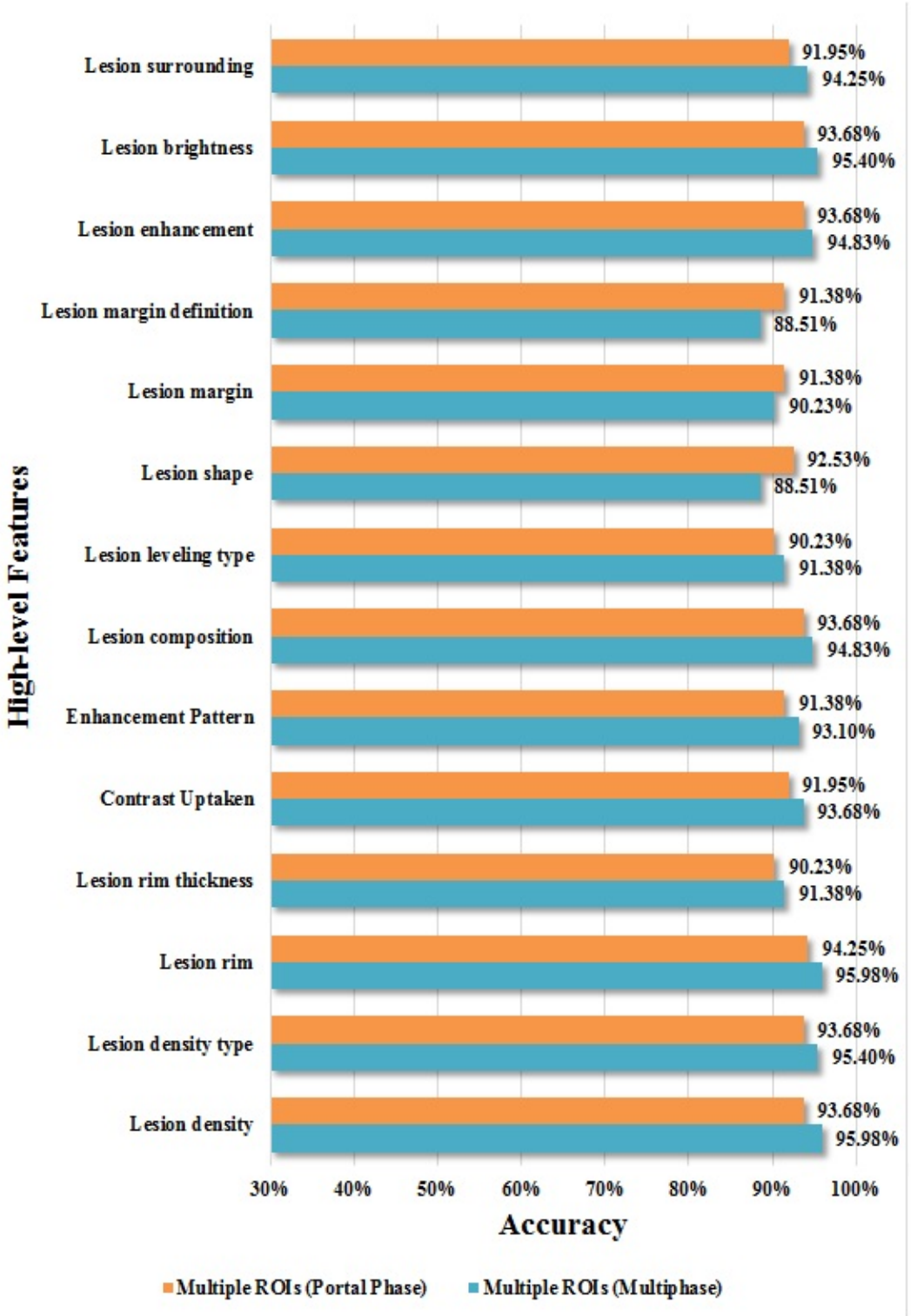


Figure B.6: The accuracy comparison of the proposed Multiple ROIs to predict the high-level features by using **portal** phase and **multiphase** CT image where **tenfold cross-validation** method and LDA classifier was adopted.

In summary

Figure B.7 depicts the evaluation comparison of the proposed Multiple ROIs for liver lesion characterisation framework based on three different types of classifiers using portal phase CT images.

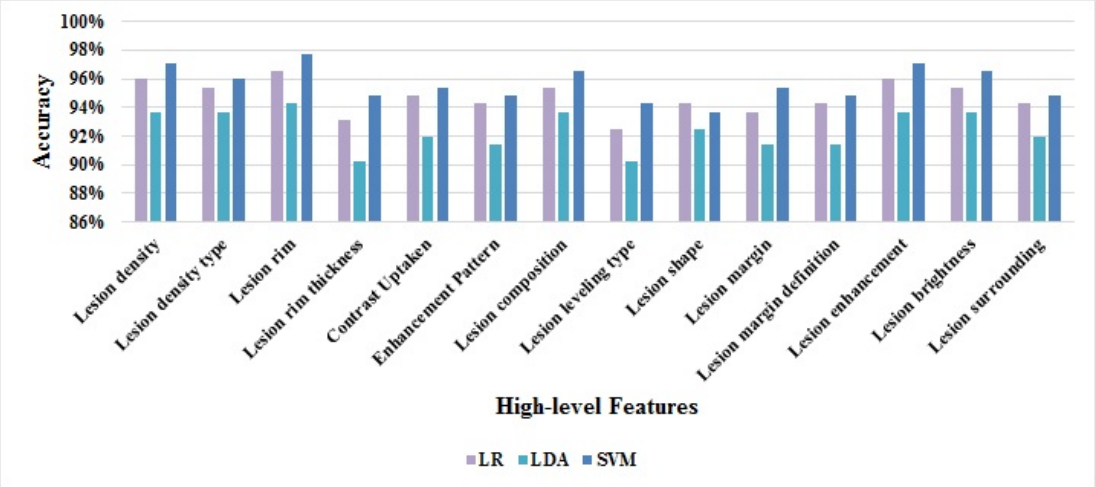


Figure B.7: The accuracy of proposed Multiple ROIs framework comparison between three different types of classifiers using **portal** phase CT image where **tenfold cross-validation** method was adopted.

Figure B.8 displays the evaluation comparison of the proposed Multiple ROIs for liver lesion characterisation framework based on three different types of classifiers using multiphase CT images.

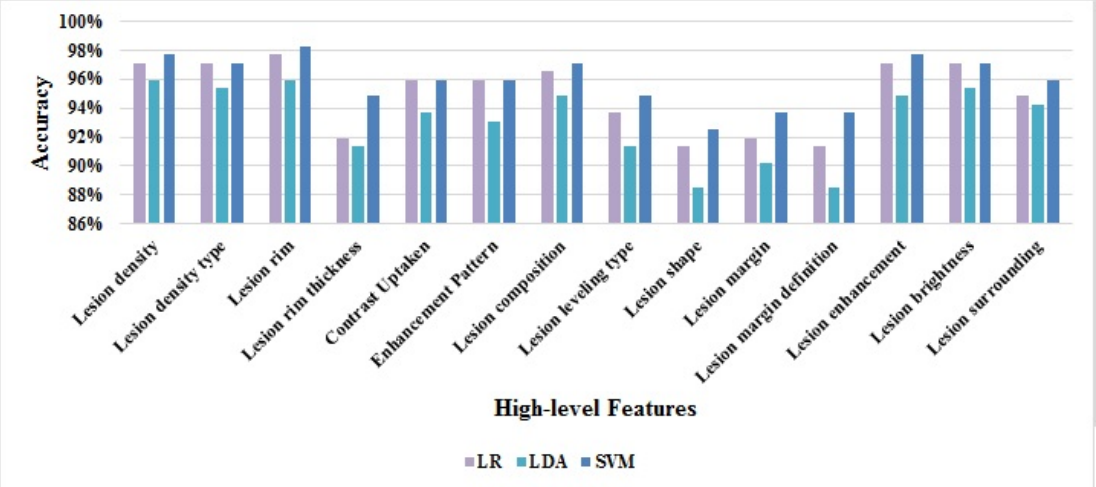


Figure B.8: The accuracy of proposed Multiple ROIs framework comparison between three different types of classifiers using **multiphase** CT image where **tenfold cross-validation** method was adopted.

Results of Lesion

This section presents the evaluation results of the proposed framework to classify liver lesion using portal phase and multiphase CT images. Two types of classifiers (Logistic Regression (LR) and Linear Discriminant Analysis (LDA)) were used to characterise liver lesion. In addition, the tenfold cross-validation method was adopted to evaluate the framework.

Logistic Regression (LR) Classifier

Table B.1 depicts comparison between all of the proposed configurations to classify liver lesion from portal phase CT image. The LR classifier performance was evaluated by adopting tenfold cross-validation method.

		SN	SP	ACC	PPV	NPV	Overall ACC
Lesion ROI	Malignant	0.775	0.883	0.775	0.849	0.822	0.833
	Benign	0.883	0.775	0.883	0.822	0.849	
DOF	Malignant	0.900	0.883	0.900	0.867	0.912	0.891
	Benign	0.883	0.900	0.883	0.912	0.867	
Multiple ROIs	Malignant	0.900	0.926	0.900	0.911	0.916	0.914
	Benign	0.926	0.900	0.926	0.916	0.911	
Multiple ROIs + DOF	Malignant	0.913	0.947	0.913	0.936	0.927	0.931
	Benign	0.947	0.913	0.947	0.927	0.936	

Table B.1: Summary of lesion classification results obtained by **tenfold cross-validation** and LR classifier using **portal** phase CT image.

Table B.2 presents comparison between all of the proposed configurations to classify liver lesion from Multiphase CT image. The LR classifier performance was evaluated by adopting tenfold cross-validation method.

		SN	SP	ACC	PPV	NPV	Overall ACC
Lesion ROI	Malignant	0.800	0.904	0.800	0.877	0.842	0.856
	Benign	0.904	0.800	0.904	0.842	0.877	
DOF	Malignant	0.913	0.915	0.913	0.901	0.925	0.914
	Benign	0.915	0.913	0.915	0.925	0.901	
Multiple ROIs	Malignant	0.900	0.936	0.900	0.923	0.917	0.920
	Benign	0.936	0.900	0.936	0.917	0.923	
Multiple ROIs + DOF	Malignant	0.938	0.957	0.938	0.949	0.947	0.948
	Benign	0.957	0.938	0.957	0.947	0.949	

Table B.2: Summary of lesion classification results obtained by **tenfold cross-validation** and LR classifier using **multiphase** CT image.

Linear Discriminant Analysis (LDA) Classifier

Table B.3 depicts comparison between all of the proposed configurations to classify liver lesion from portal phase CT image. The LDA classifier performance was evaluated by adopting tenfold cross-validation method.

		SN	SP	ACC	PPV	NPV	Overall ACC
Lesion ROI	Malignant	0.775	0.851	0.775	0.816	0.816	0.816
	Benign	0.851	0.775	0.851	0.816	0.816	
DOF	Malignant	0.875	0.883	0.875	0.864	0.892	0.879
	Benign	0.883	0.875	0.883	0.892	0.864	
Multiple ROIs	Malignant	0.888	0.904	0.888	0.888	0.904	0.897
	Benign	0.904	0.888	0.904	0.904	0.888	
Multiple ROIs + DOF	Malignant	0.900	0.904	0.900	0.889	0.914	0.902
	Benign	0.904	0.900	0.904	0.914	0.889	

Table B.3: Summary of lesion classification results obtained by **tenfold cross-validation** and LDA classifier using **portal** phase CT image.

Table B.4 presents comparison between all of the proposed configurations to classify liver lesion from Multiphase CT image. The LDA classifier performance was evaluated by adopting tenfold cross-validation method.

		SN	SP	ACC	PPV	NPV	Overall ACC
Lesion ROI	Malignant	0.800	0.872	0.800	0.842	0.837	0.839
	Benign	0.872	0.800	0.872	0.837	0.842	
DOF	Malignant	0.888	0.904	0.888	0.888	0.904	0.897
	Benign	0.904	0.888	0.904	0.904	0.888	
Multiple ROIs	Malignant	0.913	0.926	0.913	0.913	0.926	0.920
	Benign	0.926	0.913	0.926	0.926	0.913	
Multiple ROIs + DOF	Malignant	0.925	0.947	0.925	0.937	0.937	0.937
	Benign	0.947	0.925	0.947	0.937	0.937	

Table B.4: Summary of lesion classification results obtained by **tenfold cross-validation** and LDA classifier using **multiphase** CT image.

In summary

Figure B.9 depicts the **overall** accuracy comparison between all of the proposed configuration to classify liver lesion from portal phase CT image. The LR, LDA and SVM classifier were evaluated by adopting tenfold cross-validation method.

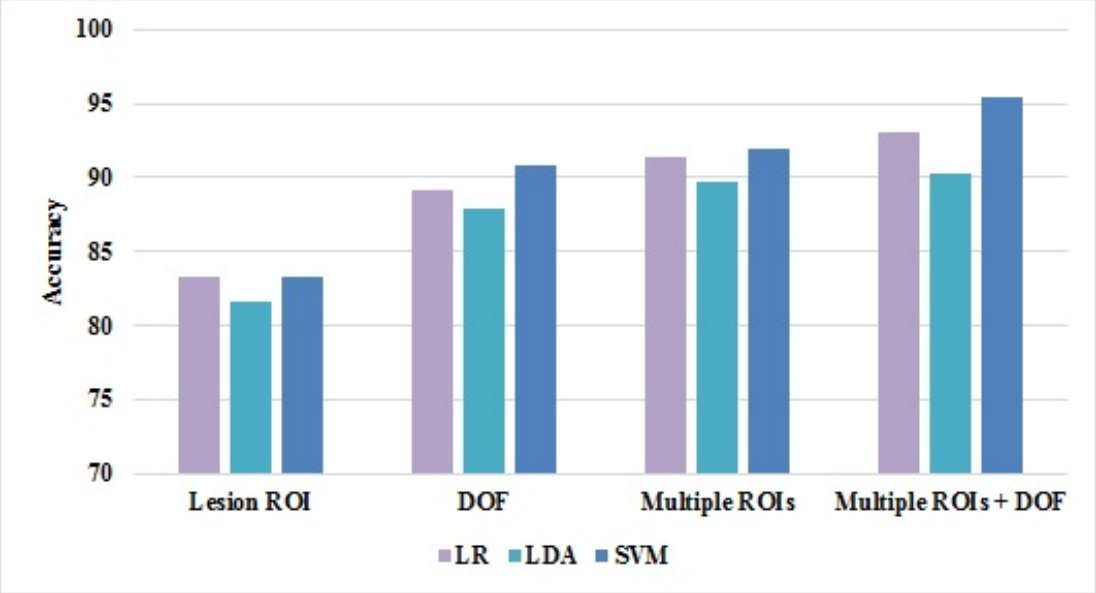


Figure B.9: The **overall** accuracy of liver lesion classification framework comparison between three different types of classifiers using **portal** CT image where **tenfold cross-validation** method was adopted.

Figure B.10 depicts the accuracy comparison between all of the proposed configuration to classify liver **malignant** lesion from portal phase CT image. The LR, LDA and SVM classifier were evaluated by adopting tenfold cross-validation method.

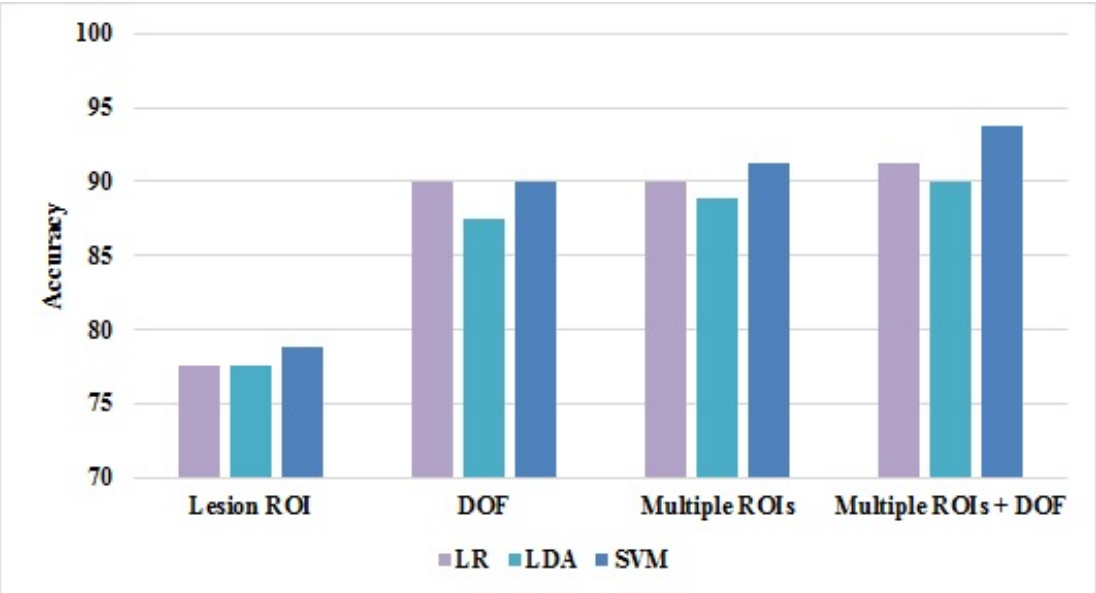


Figure B.10: The accuracy of liver **malignant** lesion classification comparison between three different types of classifiers using **portal** CT image where **tenfold cross-validation** method was adopted.

Figure B.11 depicts the accuracy comparison between all of the proposed configuration to classify liver **benign** lesion from portal phase CT image. The LR, LDA and SVM classifier were evaluated by adopting tenfold cross-validation method.

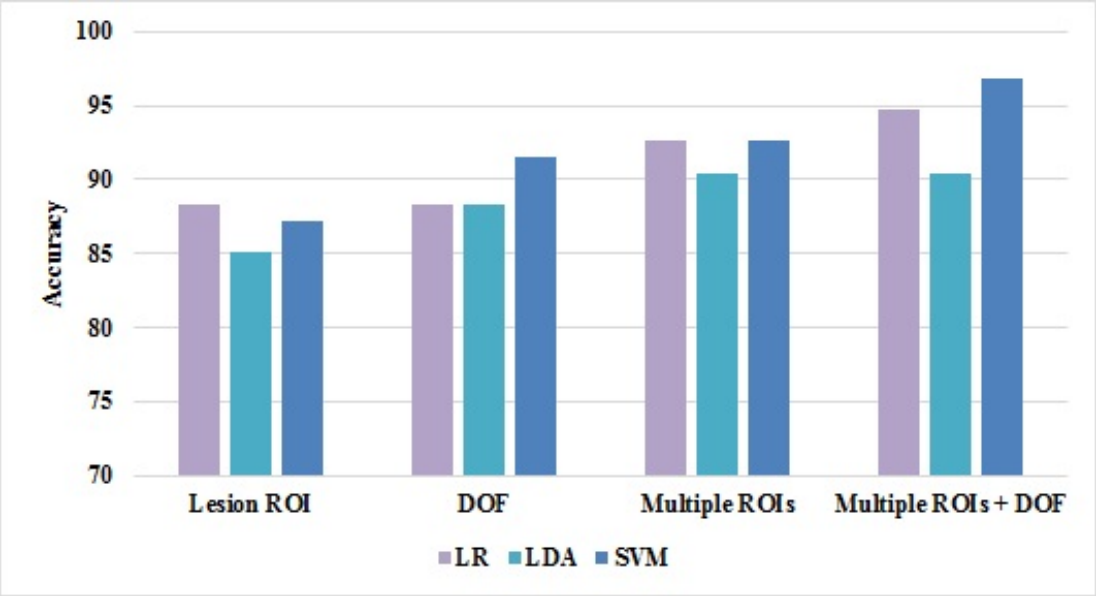


Figure B.11: The accuracy of liver **benign** lesion classification comparison between three different types of classifiers using **portal** CT image where **tenfold cross-validation** method was adopted.

Figure B.12 depicts the **overall** accuracy comparison between all of the proposed configuration to classify liver lesion from Multiphase CT image. The LR, LDA and SVM classifier were evaluated by adopting tenfold cross-validation method.

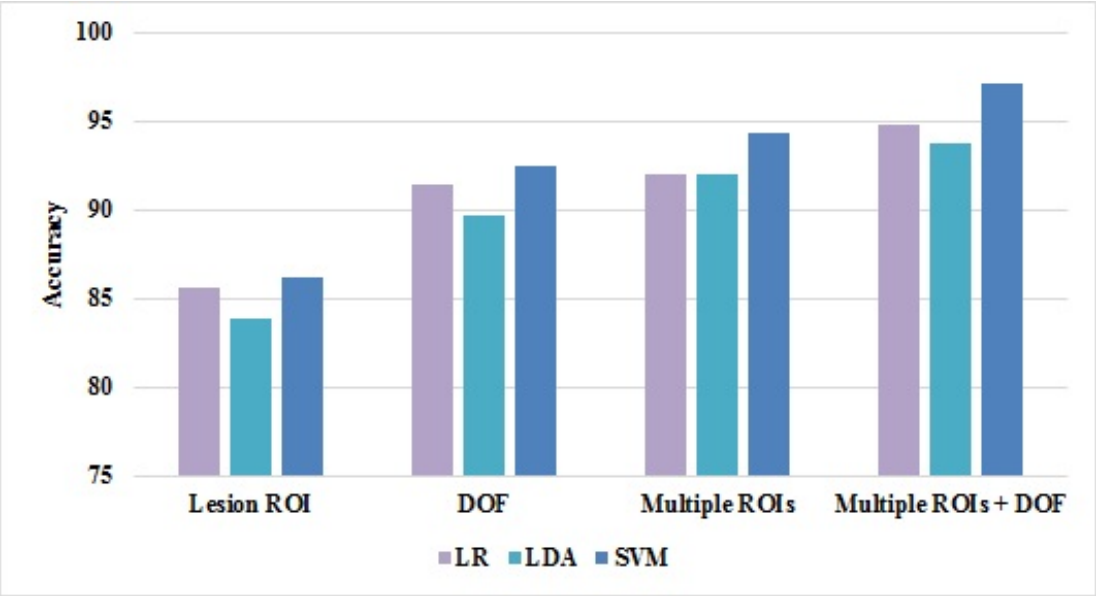


Figure B.12: The **overall** accuracy of liver lesion classification framework comparison between three different types of classifiers using **multiphase** CT image where **tenfold cross-validation** method was adopted.

Figure B.13 depicts the accuracy comparison between all of the proposed config-

uration to classify liver **malignant** lesion from Multiphase phase CT image. The LR, LDA and SVM classifier were evaluated by adopting tenfold cross-validation method.

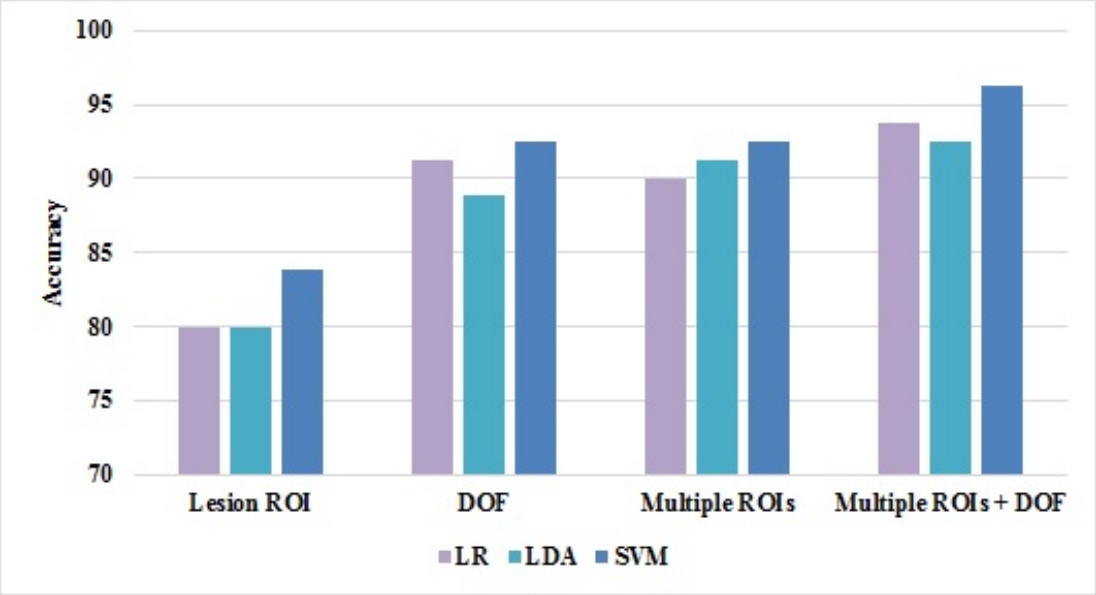


Figure B.13: The accuracy of liver **malignant** lesion classification comparison between three different types of classifiers using **multiphase** CT image where **tenfold cross-validation** method was adopted.

Figure B.14 depicts the accuracy comparison between all of the proposed configuration to classify liver **benign** lesion from Multiphase phase CT image. The LR, LDA and SVM classifier were evaluated by adopting tenfold cross-validation method.

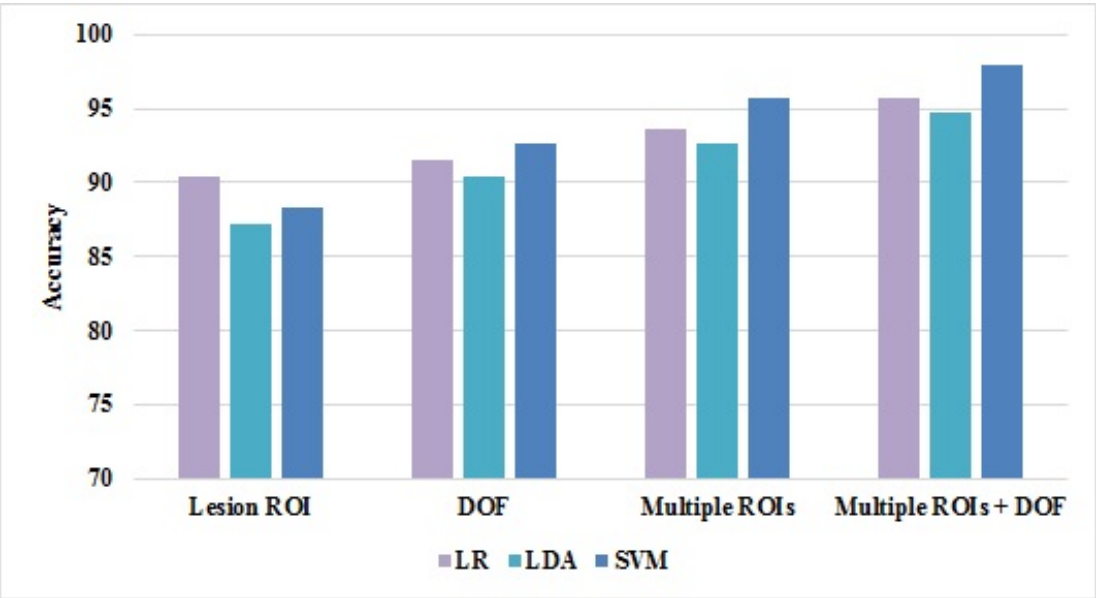


Figure B.14: The accuracy of liver **benign** lesion classification comparison between three different types of classifiers using **multiphase** CT image where **tenfold cross-validation** method was adopted.

Bibliography

Aalaei, S., Shahraki, H., Rowhanimanesh, A., and Eslami, S. (2016). Feature selection using genetic algorithm for breast cancer diagnosis: experiment on three different datasets. *Iranian journal of basic medical sciences*, 19(5):476.

Adams, L. J., Bello, G., and Dumancas, G. G. (2015). Development and application of a genetic algorithm for variable optimization and predictive modeling of five-year mortality using questionnaire data. *Bioinformatics and biology insights*, 9(Suppl 3):31.

Adams, R. and Bischof, L. (1994). Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence*, 16(6):641–647.

Afonso, P. D. (2015). Imaging techniques for the diagnosis of soft tissue tumors. *Reports in Medical Imaging*, 8:63–70.

Agarwal Vibhu, O. D. (2013). Predicting semantic features from ct images of liver lesions using deep learning.

Ahmadi, K., Karimi, A., and Fouladi Nia, B. (2016). New technique for automatic segmentation of blood vessels in ct scan images of liver based on optimized fuzzy-means method. *Computational and mathematical methods in medicine*, 2016.

Aisen, A. M., Martel, W., Braunstein, E. M., McMillin, K. I., Phillips, W. A., and Kling, T. (1986). Mri and ct evaluation of primary bone and soft-tissue tumors. *American journal of Roentgenology*, 146(4):749–756.

Akhter, N., Dabhade, S., Bansod, N., and Kale, K. (2016). Feature selection for heart rate variability based biometric recognition using genetic algorithm. In *Intelligent Systems Technologies and Applications*, pages 91–101. Springer.

Al-Salem, A. H. (2014). Pediatric liver tumors. In *An Illustrated Guide to Pediatric Surgery*, pages 469–478. Springer.

Albregtsen, F., Nielsen, B., and Danielsen, H. E. (2000). Adaptive gray level run length features from class distance matrices. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 3, pages 738–741. IEEE.

Andriole, K. P., Wolfe, J. M., Khorasani, R., Treves, S. T., Getty, D. J., Jacobson, F. L., Steigner, M. L., Pan, J. J., Sitek, A., and Seltzer, S. E. (2011). Optimizing analysis, visualization, and navigation of large image data sets: one 5000-section ct scan can ruin your whole day. *Radiology*, 259(2):346–362.

Ang, J., Heath, J., Donath, S., Khurana, S., and Auldist, A. (2007). Treatment outcomes for hepatoblastoma: an institutionâĂŹs experience over two decades. *Pediatric surgery international*, 23(2):103–109.

Anter, A. M., Azar, A. T., Hassanien, A. E., El-Bendary, N., and ElSoud, M. A. (2013). Automatic computer aided segmentation for liver and hepatic lesions using hybrid segmentations techniques. In *Computer science and information systems (FedCSIS), 2013 federated conference on*, pages 193–198. IEEE.

Arakeri, M. P. et al. (2011). Recent advances and future potential of computer aided diagnosis of liver cancer on computed tomography images. In *Computer Networks and Intelligent Computing*, pages 246–251. Springer.

Assy, N., Nasser, G., Djibre, A., Beniashvili, Z., Elias, S., and Zidan, J. (2009a). Characteristics of common solid liver lesions and recommendations for diagnostic workup. *World J Gastroenterol*, 15(26):3217–3227.

Assy, N., Nasser, G., Djibre, A., Beniashvili, Z., Elias, S., and Zidan, J. (2009b). Characteristics of common solid liver lesions and recommendations for diagnostic workup. *World J Gastroenterol*, 15(26):3217–3227.

Baaziz, N., Abahmane, O., and Missaoui, R. (2010). Texture feature extraction in the spatial-frequency domain for content-based image retrieval. *arXiv preprint arXiv:1012.5208*.

Bakoyiannis, A., Delis, S., Triantopoulou, C., and Dervenis, C. (2013). Rare cystic liver lesions: a diagnostic and managing challenge. *World Journal of Gastroenterology: WJG*, 19(43):7603.

Bandiera, S., Pfeffer, S., Baumert, T. F., and Zeisel, M. B. (2015). mir-122–a key factor and therapeutic target in liver disease. *Journal of hepatology*, 62(2):448–457.

Bankman, I. (2008). *Handbook of medical image processing and analysis*. academic press.

Bayram, U., Can, G., Duzgun, S., and Yalabik, N. (2011). Evaluation of textural features for multispectral images. In *Image and Signal Processing for Remote Sensing XVII*, volume 8180, page 81800I. International Society for Optics and Photonics.

Beichel12, R., Bauer, C., Bornik, A., Sorantin, E., and Bischof, H. (2007). Liver segmentation in ct data: A segmentation refinement approach. *Proceedings of" 3D Segmentation in The Clinic: A Grand Challenge*, pages 235–245.

Berland, L. L., Koslin, D., Kenney, P., Stanley, R., and Lee, J. (1988). Differentiation between small benign and malignant adrenal masses with dynamic incremented ct. *American Journal of Roentgenology*, 151(1):95–101.

Berlin, L. (1996). Malpractice issues in radiology. perceptual errors. *AJR. American journal of roentgenology*, 167(3):587–590.

Berlin, L. (2007). Accuracy of diagnostic procedures: has it improved over the past five decades? *American Journal of Roentgenology*, 188(5):1173–1178.

Bewick, V., Cheek, L., and Ball, J. (2005). Statistics review 14: Logistic regression. *Critical care*, 9(1):112.

Bhadauria, H. and Dewal, M. (2012). Efficient denoising technique for ct images to enhance brain hemorrhage segmentation. *Journal of digital imaging*, 25(6):782–791.

Bharti, P., Mittal, D., and Ananthasivan, R. (2017). Computer-aided characterization and diagnosis of diffuse liver diseases based on ultrasound imaging: A review. *Ultrasonic imaging*, 39(1):33–61.

Birnbaum, B. A., Hindman, N., Lee, J., and Babb, J. S. (2007). Multi–detector row ct attenuation measurements: assessment of intra-and interscanner variability with an anthropomorphic body ct phantom. *Radiology*, 242(1):109–119.

Bismuth, H. (1982). Surgical anatomy and anatomical surgery of the liver. *World journal of surgery*, 6(1):3–9.

Bismuth, H. (2013). Revisiting liver anatomy and terminology of hepatectomies. *Annals of surgery*, 257(3):383–386.

Blachar, A., Federle, M. P., Ferris, J. V., Lacomis, J. M., Waltz, J. S., Armfield, D. R., Chu, G., Almusa, O., Grazioli, L., Balzano, E., et al. (2002). RadiologistsâĂŹ performance in the diagnosis of liver tumors with central scars by using specific ct criteria. *Radiology*, 223(2):532–539.

Blechacz, B. R. and Gores, G. J. (2008). Cholangiocarcinoma. *Clinics in liver disease*, 12(1):131–150.

Bobrowski, L. and Kretowski, M. (2000). Induction of multivariate decision trees by using dipolar criteria. *Principles of Data Mining and Knowledge Discovery*, pages 335–344.

Bosch, A., Muñoz, X., and Martí, R. (2007). Which is the best way to organize/classify images by content? *Image and vision computing*, 25(6):778–791.

Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Transactions on pattern analysis and machine intelligence*, 23(11):1222–1239.

Campadelli, P., Casiraghi, E., and Esposito, A. (2009a). Liver segmentation from computed tomography scans: A survey and a new algorithm. *Artificial intelligence in medicine*, 45(2-3):185–196.

Campadelli, P., Casiraghi, E., Pratissoli, S., and Lombardi, G. (2009b). Automatic abdominal organ segmentation from ct images. *ELCVIA: electronic letters on computer vision and image analysis*, 8(1):1–14.

Caputo, B., Müller, H., Martinez-Gomez, J., Villegas, M., Acar, B., Patricia, N., Marvasti, N., Üsküdarlı, S., Paredes, R., Cazorla, M., et al. (2014). Imageclef 2014: Overview and analysis of the results. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pages 192–211. Springer.

Casciaro, S., Franchini, R., Massoptier, L., Casciaro, E., Conversano, F., Malvasi, A., and Lay-Ekuakille, A. (2012a). Fully automatic segmentations of liver and hepatic tumors from 3-d computed tomography abdominal images: comparative evaluation of two automatic methods. *IEEE Sensors journal*, 12(3):464–473.

Casciaro, S., Franchini, R., Massoptier, L., Casciaro, E., Conversano, F., Malvasi, A., and Lay-Ekuakille, A. (2012b). Fully automatic segmentations of liver and hepatic tumors from 3-d computed tomography abdominal images: comparative evaluation of two automatic methods. *IEEE Sensors journal*, 12(3):464–473.

Catheline, J.-M., Turner, R., and Champault, G. (2000). Laparoscopic ultrasound of the liver. *European journal of ultrasound*, 12(2):169–177.

Chakraborty, S. (2011). Bayesian semi-supervised learning with support vector machine. *Statistical Methodology*, 8(1):68–82.

Chang, C.-C., Chen, H.-H., Chang, Y.-C., Yang, M.-Y., Lo, C.-M., Ko, W.-C., Lee, Y.-F., Liu, K.-L., and Chang, R.-F. (2017). Computer-aided diagnosis of liver tumors on computed tomography images. *Computer Methods and Programs in Biomedicine*, 145:45–51.

Chang, F., Chen, C.-J., and Lu, C.-J. (2004). A linear-time component-labeling algorithm using contour tracing technique. *computer vision and image understanding*, 93(2):206–220.

Chang, H.-H. and Chu, W.-C. (2012). Restoration algorithm for image noise removal using double bilateral filtering. *Journal of Electronic Imaging*, 21(2):023028–1.

Chen, C.-C., DaPonte, J. S., and Fox, M. D. (1989). Fractal feature analysis and classification in medical imaging. *IEEE transactions on medical imaging*, 8(2):133–142.

Chen, E.-L., Chung, P.-C., Chen, C.-L., Tsai, H.-M., and Chang, C.-I. (1998). An automatic diagnostic system for ct liver image classification. *IEEE transactions on biomedical engineering*, 45(6):783–794.

Chen, M., Zhu, H., and Zhu, H. (2013). Segmentation of liver in ultrasonic images applying local optimal threshold method. *The Imaging Science Journal*, 61(7):579–591.

Chen, X., Udupa, J. K., Bagci, U., Zhuge, Y., and Yao, J. (2012). Medical image segmentation by combining graph cuts and oriented active appearance models. *IEEE Transactions on Image Processing*, 21(4):2035–2046.

Chen, Y., Wang, Z., Zhao, W., and Yang, X. (2009). Liver segmentation from ct images based on region growing method. In *Bioinformatics and Biomedical Engineering, 2009. ICBBE 2009. 3rd International Conference on*, pages 1–4. IEEE.

Cheng, H., Shi, X., Min, R., Hu, L., Cai, X., and Du, H. (2006). Approaches for automated detection and classification of masses in mammograms. *Pattern recognition*, 39(4):646–668.

Chi, Y., Zhou, J., Venkatesh, S. K., Huang, S., Tian, Q., Hennedige, T., and Liu, J. (2013a). Computer-aided focal liver lesion detection. *International journal of computer assisted radiology and surgery*, 8(4):511–525.

Chi, Y., Zhou, J., Venkatesh, S. K., Tian, Q., and Liu, J. (2013b). Content-based image retrieval of multiphase ct images for focal liver lesion characterization. *Medical physics*, 40(10).

Chtioui, Y., Bertrand, D., and Barba, D. (2009). Feature selection by a genetic algorithm. application to seed discrimination by artificial vision. *Journal of the Science of Food and Agriculture*, 76(1):77–86.

Ciecholewski, M. (2014). Automatic liver segmentation from 2d ct images using an approximate contour model. *Journal of Signal Processing Systems*, 74(2):151–174.

Ciecholewski, M. and Ogiela, M. R. (2007). Automatic segmentation of single and multiple neoplastic hepatic lesions in ct images. In *International Work-Conference on the Interplay Between Natural and Artificial Computation*, pages 63–71. Springer.

Clausi, D. A. (2002). An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of remote sensing*, 28(1):45–62.

Cohen, E. B. and Afdhal, N. H. (2010). Ultrasound-based hepatic elastography: origins, limitations, and applications. *Journal of clinical gastroenterology*, 44(9):637–645.

Conners, R. W. and Harlow, C. A. (1980). Toward a structural textural analyzer based on statistical methods. *Computer Graphics and Image Processing*, 12(3):224–256.

Corson, N., Sensakovic, W. F., Straus, C., Starkey, A., and Armato, S. G. (2011). Characterization of mesothelioma and tissues present in contrast-enhanced thoracic ct scans. *Medical physics*, 38(2):942–947.

Cross, G. R. and Jain, A. K. (1983). Markov random field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1):25–39.

Daněk, O., Matula, P., Maška, M., and Kozubek, M. (2012). Smooth chan–vese segmentation via graph cuts. *Pattern Recognition Letters*, 33(10):1405–1410.

Dankerl, P., Cavallaro, A., Tsymbal, A., Costa, M. J., Suehling, M., Janka, R., Uder, M., and Hammon, M. (2013). A retrieval-based computer-aided diagnosis system for the characterization of liver lesions in ct scans. *Academic radiology*, 20(12):1526–1534.

Das, A. and Sabut, S. K. (2016). Kernelized fuzzy c-means clustering with adaptive thresholding for segmenting liver tumors. *Procedia Computer Science*, 92:389–395.

Davenport, K. P., Blanco, F. C., and Sandler, A. D. (2012). Pediatric malignancies: neuroblastoma, wilm's tumor, hepatoblastoma, rhabdomyosarcoma, and sacrococcygeal teratoma. *Surgical Clinics of North America*, 92(3):745–767.

Davis, G. L., Dempster, J., Meler, J. D., Orr, D. W., Walberg, M. W., Brown, B., Berger, B. D., O'Connor, J. K., and Goldstein, R. M. (2008). Hepatocellular carcinoma: management of an increasingly common problem. In *Baylor University Medical Center. Proceedings*, volume 21, page 266. Baylor University Medical Center.

Dawant, B. M., Li, R., Lennon, B., and Li, S. (2007). Semi-automatic segmentation of the liver and its evaluation on the miccai 2007 grand challenge data set. *3D Segmentation in The Clinic: A Grand Challenge*, pages 215–221.

Depeursinge, A., Kurtz, C., Beaulieu, C., Napel, S., and Rubin, D. (2014). Predicting visual semantic descriptive terms from radiological image data: preliminary results with liver lesions in ct. *IEEE transactions on medical imaging*, 33(8):1669–1676.

Despotović, I., Goossens, B., and Philips, W. (2015). Mri segmentation of the human brain: challenges, methods, and applications. *Computational and mathematical methods in medicine*, 2015.

Dhar, V., Thomas, R. M., and Ahmad, S. A. (2016). Repeat hepatectomy for colorectal liver metastases. In *Gastrointestinal Malignancies*, pages 203–220. Springer.

Dhingra, S. and Fiel, M. I. (2014). Update on the new classification of hepatic adenomas: clinical, molecular, and pathologic characteristics. *Archives of Pathology and Laboratory Medicine*, 138(8):1090–1097.

Diamant, I., Goldberger, J., Klang, E., Amitai, M., and Greenspan, H. (2015). Multiphase liver lesions classification using relevant visual words based on mutual information. In *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, pages 407–410. IEEE.

Diamant, I., Hoogi, A., Beaulieu, C. F., Safdari, M., Klang, E., Amitai, M., Greenspan, H., and Rubin, D. L. (2016). Improved patch-based automated liver lesion classification by separate analysis of the interior and boundary regions. *IEEE journal of biomedical and health informatics*, 20(6):1585–1594.

Diamant, I., Klang, E., Amitai, M., Konen, E., Goldberger, J., and Greenspan, H. (2017). Task-driven dictionary learning based on mutual information for medical image classification. *IEEE Transactions on Biomedical Engineering*, 64(6):1380–1392.

Dilsiz, A., Aydin, T., and Gursan, N. (2009). Capillary hemangioma as a rare benign tumor of the oral cavity: a case report. *Cases journal*, 2(1):8622.

Doron, Y., Mayer-Wolf, N., Diamant, I., and Greenspan, H. (2014a). Texture feature based liver lesion classification. In *SPIE Medical Imaging*, pages 90353K–90353K. International Society for Optics and Photonics.

Doron, Y., Mayer-Wolf, N., Diamant, I., and Greenspan, H. (2014b). Texture feature based liver lesion classification. In *SPIE Medical Imaging*, pages 90353K–90353K. International Society for Optics and Photonics.

Dramiński, M., Rada-Iglesias, A., Enroth, S., Wadelius, C., Koronacki, J., and Komorowski, J. (2007). Monte carlo feature selection for supervised classification. *Bioinformatics*, 24(1):110–117.

Duda, D., Krętowski, M., and Bézy-Wendling, J. (2004). Texture-based classification of hepatic primary tumors in multiphase ct. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2004*, pages 1050–1051.

Duda, D., Kretowski, M., and Bézy-Wendling, J. (2006). Texture characterization for hepatic tumor recognition in multiphase ct. *Biocybernetics and Biomedical Engineering*, 26(4):15.

Duda, D., Krętowski, M., and Bézy-Wendling, J. (2013). Computer-aided diagnosis of liver tumors based on multi-image texture analysis of contrast-enhanced ct. selection of the most appropriate texture features. *Studies in Logic, Grammar and Rhetoric*, 35(1):49–70.

Duda, R. O., Hart, P. E., and Stork, D. G. (2012). *Pattern classification*. John Wiley & Sons.

Dwyer, A. (1991). Matchmaking and mcnemar in the comparison of diagnostic modalities. *Radiology*, 178(2):328–330.

Elsayes, K. M., Narra, V. R., Yin, Y., Mukundan, G., Lammle, M., and Brown, J. J. (2005). Focal hepatic lesions: diagnostic value of enhancement pattern approach with contrast-enhanced 3d gradient-echo mr imaging. *Radiographics*, 25(5):1299–1320.

Eltoukhy, M. M., Faye, I., and Samir, B. B. (2012). A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation. *Computers in biology and medicine*, 42(1):123–128.

Erdogan, D., Busch, O. R., Van Delden, O. M., Bennink, R. J., Ten Kate, F. J., Gouma, D. J., and Van Gulik, T. M. (2007). Management of liver hemangiomas according to size and symptoms. *Journal of gastroenterology and hepatology*, 22(11):1953–1958.

Erdt, M. and Kirschner, M. (2010). Fast automatic liver segmentation combining learned shape priors with observed shape deviation. In *Computer-Based Medical Systems (CBMS), 2010 IEEE 23rd International Symposium on*, pages 249–254. IEEE.

Fagerland, M. W., Lydersen, S., and Laake, P. (2013). The mcnemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. *BMC medical research methodology*, 13(1):91.

Faraj, W., Dar, F., Marangoni, G., Bartlett, A., Melendez, H. V., Hadzic, D., Dhawan, A., Mieli-Vergani, G., Rela, M., and Heaton, N. (2008). Liver transplantation for hepatoblastoma. *Liver Transplantation*, 14(11):1614–1619.

Farid, S. G., Prasad, K. R., and Morris-Stiff, G. (2013). Operative terminology and post-operative management approaches applied to hepatic surgery: Trainee perspectives. *World journal of gastrointestinal surgery*, 5(5):146.

Fass, L. (2008). Imaging and cancer: a review. *Molecular oncology*, 2(2):115–152.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern recognition letters*, 27(8):861–874.

Feinstein, S. B., Coll, B., Staub, D., Adam, D., Schinkel, A. F., Folkert, J., and Thomenius, K. (2010). Contrast enhanced ultrasound imaging. *Journal of nuclear cardiology*, 17(1):106–115.

Fergusson, J. (2012a). Investigation and management of hepatic incidentalomas. *Journal of gastroenterology and hepatology*, 27(12):1772–1782.

Fergusson, J. (2012b). Investigation and management of hepatic incidentalomas. *Journal of gastroenterology and hepatology*, 27(12):1772–1782.

Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D., and Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012. *International journal of cancer*, 136(5):E359–E386.

Flach, P., Blockeel, H., Ferri, C., Hernández-Orallo, J., and Struyf, J. (2003). Decision support for data mining. In *Data Mining and Decision Support*, pages 81–90. Springer.

Fletcher, J. G., Leng, S., Yu, L., and McCollough, C. H. (2016). Dealing with uncertainty in ct images.

Foruzan, A. H. and Chen, Y.-W. (2016). Improved segmentation of low-contrast lesions using sigmoid edge model. *International journal of computer assisted radiology and surgery*, 11(7):1267–1283.

France, I. (2016). 3dircadb, 3d image reconstruction for comparison of algorithm database. https://www.ircad.fr/research/3dircadb/.

Freund, Y. and Schapire, R. E. (1995). A desicion-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pages 23–37. Springer.

Fukunaga, K. (2013). *Introduction to statistical pattern recognition*. Academic press.

Furlan, A., van der Windt, D. J., Nalesnik, M. A., Sholosh, B., Ngan, K.-K., Pealer, K. M., Ijzermans, J. N., and Federle, M. P. (2008). Multiple hepatic adenomas associated with liver steatosis at ct and mri: a case-control study. *American Journal of Roentgenology*, 191(5):1430–1435.

Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Computer graphics and image processing*, 4(2):172–179.

Ganeshan, B., Miles, K. A., Young, R. C., and Chatwin, C. R. (2009). Texture analysis in non-contrast enhanced ct: impact of malignancy on texture in apparently disease-free areas of the liver. *European journal of radiology*, 70(1):101–110.

Garg, P. (2010). A comparison between memetic algorithm and genetic algorithm for the cryptanalysis of simplified data encryption standard algorithm. *arXiv preprint arXiv:1004.0574*.

Garrean, S., Hering, J., Saied, A., Helton, W. S., and Espat, N. J. (2008). Radiofrequency ablation of primary and metastatic liver tumors: a critical review of the literature. *The American Journal of Surgery*, 195(4):508–520.

Geisser, S. (1993). *Predictive inference*, volume 55. CRC press.

Ger, R. (1989). Surgical anatomy of the liver. *Surgical Clinics of North America*, 69(2):179–192.

Germain, T., Favelier, S., Cercueil, J.-P., Denys, A., Krausé, D., and Guiu, B. (2014). Liver segmentation: practical tips. *Diagnostic and interventional imaging*, 95(11):1003–1016.

Gimenez, F., Xu, J., Liu, Y., Liu, T. T., Beaulieu, C. F., Rubin, D. L., and Napel, S. (2012). Automatic annotation of radiological observations in liver ct images. In *AMIA*.

Gletsos, M., Mougiakakou, S. G., Matsopoulos, G. K., Nikita, K. S., Nikita, A. S., and Kelekis, D. (2003). A computer-aided diagnostic system to characterize ct focal liver lesions: design and optimization of a neural network classifier. *IEEE transactions on information technology in biomedicine*, 7(3):153–162.

Glinkova, V., Shevah, O., Boaz, M., Levine, A., and Shirin, H. (2004). Hepatic haemangiomas: possible association with female sex hormones. *Gut*, 53(9):1352–1355.

Glisson, C. L., Altamar, H. O., Herrell, S. D., Clark, P., and Galloway, R. L. (2011). Comparison and assessment of semi-automatic image segmentation in computed tomography scans for image-guided kidney surgery. *Medical physics*, 38(11):6265–6274.

Gloger, O., Kühn, J., Stanski, A., Völzke, H., and Puls, R. (2010). A fully automatic three-step liver segmentation method on lda-based probability maps for multiple contrast mr images. *Magnetic Resonance Imaging*, 28(6):882–897.

Glover, C., Douse, P., Kane, P., Karani, J., Meire, H., Mohammadtaghi, S., and Allen-Mersh, T. (2002). Accuracy of investigations for asymptomatic colorectal liver metastases. *Diseases of the colon & rectum*, 45(4):476–484.

Gonzales, R. and Woods, E. (2007). Digital image processing, 3, d edition.

Gravel, P., Beaudoin, G., and De Guise, J. A. (2004). A method for modeling noise in medical images. *IEEE Transactions on medical imaging*, 23(10):1221–1232.

Grazioli, L., Bondioni, M. P., Faccioli, N., Gambarini, S., Tinti, R., Schneider, G., and Kirchin, M. (2010). Solid focal liver lesions: dynamic and late enhancement patterns with the dual phase contrast agent gadobenate dimeglumine. *Journal of gastrointestinal cancer*, 41(4):221–232.

Grazioli, L., Bondioni, M. P., Haradome, H., Motosugi, U., Tinti, R., Frittoli, B., Gambarini, S., Donato, F., and Colagrande, S. (2012). Hepatocellular adenoma and focal nodular hyperplasia: value of gadoxetic acid–enhanced mr imaging in differential diagnosis. *Radiology*, 262(2):520–529.

Grazioli, L., Federle, M. P., Brancatelli, G., Ichikawa, T., Olivetti, L., and Blachar, A. (2001). Hepatic adenomas: Imaging and pathologic findings 1. *Radiographics*, 21(4):877–892.

Gunasundari, S. and Ananthi, M. S. (2012). Comparison and evaluation of methods for liver tumor classification from ct datasets. *International journal of computer applications*, 39(18):46–51.

Guo, Z., Zhang, L., and Zhang, D. (2010). Rotation invariant texture classification using lbp variance (lbpv) with global matching. *Pattern recognition*, 43(3):706–719.

Halankar, J. A., Kim, T. K., Jang, H.-J., Khalili, K., Masoom, H. A., et al. (2012). Understanding the natural history of focal nodular hyperplasia in the liver with mri. *Indian Journal of Radiology and Imaging*, 22(2):116.

Häme, Y. (2008). Liver tumor segmentation using implicit surface evolution. *The Midas Journal*, pages 1–10.

Han, K. and Talavage, T. M. (2011). Effects of combining field strengths on auditory functional mri group analysis: 1.5 t and 3t. *Journal of Magnetic Resonance Imaging*, 34(6):1480–1488.

Han, N. Y., Park, B. J., Kim, M. J., Sung, D. J., and Cho, S. B. (2015). Hepatic parenchymal heterogeneity on contrast-enhanced ct scans following oxaliplatin-based chemotherapy: natural history and association with clinical evidence of sinusoidal obstruction syndrome. *Radiology*, 276(3):766–774.

Haralick, R. M., Shanmugam, K., et al. (1973a). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, 3(6):610–621.

Haralick, R. M., Shanmugam, K., et al. (1973b). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621.

Hedman, K., Nylander, E., Henriksson, J., Bjarnegård, N., Brudin, L., and Tamás, É. (2016). Echocardiographic characterization of the inferior vena cava in trained and untrained females. *Ultrasound in medicine & biology*, 42(12):2794–2802.

Heiken, J. P. (2007). Distinguishing benign from malignant liver tumours. *Cancer Imaging*, 7(Special issue A):S1.

Heimann, T., Meinzer, H., and Wolf, I. (2007a). A statistical deformable model for the segmentation of liver ct volumes using extended training data. *Proc. MICCAI Work*, pages 161–166.

Heimann, T., Münzing, S., Meinzer, H.-P., and Wolf, I. (2007b). A shape-guided deformable model with evolutionary algorithm initialization for 3d soft tissue segmentation. In *Biennial International Conference on Information Processing in Medical Imaging*, pages 1–12. Springer.

Heimann, T., Van Ginneken, B., and Styner, M. (2007c). 3d segmentation in the clinic: A grand challenge. *3D segmentation in the clinic: a grand challenge*, pages 7–15.

Heimann, T., Van Ginneken, B., Styner, M. A., Arzhaeva, Y., Aurich, V., Bauer, C., Beck, A., Becker, C., Beichel, R., Bekes, G., et al. (2009). Comparison and evaluation of methods for liver segmentation from ct datasets. *IEEE transactions on medical imaging*, 28(8):1251–1265.

Heimann, T., Wolf, I., and Meinzer, H.-P. (2006). Active shape models for a fully automated 3d segmentation of the liver–an evaluation on clinical data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 41–48. Springer.

Hibet-Allah, O., Hajer, J., and Kamel, H. (2016). Vascular tree segmentation in mra images using hessian-based multiscale filtering and local entropy thresholding. In *Advanced Technologies for Signal and Image Processing (ATSIP), 2016 2nd International Conference on*, pages 325–329. IEEE.

Hong, J.-S., Kaneko, T., Sekiguchi, R., and Park, K.-H. (2001). Automatic liver tumor detection from ct. *IEICE Transactions on information and systems*, 84(6):741–748.

Horng, M.-H., Sun, Y.-N., and Lin, X.-Z. (2002). Texture feature coding method for classification of liver sonography. *Computerized medical imaging and graphics*, 26(1):33–42.

Hosmer Jr, D. W., Lemeshow, S., and Sturdivant, R. X. (2013). *Applied logistic regression*, volume 398. John Wiley & Sons.

Hsieh, J. (2003a). *Computed tomography: principles, design, artifacts, and recent advances*, volume 114. SPIE press.

Hsieh, J. (2003b). *Computed tomography: principles, design, artifacts, and recent advances*, volume 114. SPIE press.

Hsu, C.-W. and Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, 13(2):415–425.

Hu, H. (1999). Multi-slice helical ct: Scan and reconstruction. *Medical physics*, 26(1):5–18.

Hu, Q., Qian, G., and Nowinski, W. L. (2005). Fast connected-component labelling in three-dimensional binary images based on iterative recursion. *Computer Vision and Image Understanding*, 99(3):414–434.

Huang, J., Meng, L., Qu, W., and Wang, C. (2011). Based on statistical analysis and 3d region growing segmentation method of liver. In *Advanced Computer Control (ICACC), 2011 3rd International Conference on*, pages 478–482. IEEE.

Huang, S., Wang, B., Cheng, M., Huang, X., and Ju, Y. (2007). A simplified method to segment liver according to couinaud's classification. In *Bioinformatics and Biomedical Engineering, 2007. ICBBE 2007. The 1st International Conference on*, pages 523–526. IEEE.

Huang, Y.-L., Chen, J.-H., and Shen, W.-C. (2006). Diagnosis of hepatic tumors with texture analysis in nonenhanced computed tomography images. *Academic radiology*, 13(6):713–720.

Iroju, O. G. and Olaleke, J. O. (2015). A systematic review of natural language processing in healthcare. *International Journal of Information Technology and Computer Science*, 8:44–50.

Iwatsuki, S., Todo, S., Marsh, J. W., Madariaga, J. R., Lee, R. G., Dvorchik, I., Fung, J. J., and Starzl, T. E. (1998). Treatment of hilar cholangiocarcinoma (klatskin tumors) with hepatic resection or transplantation. *Journal of the American College of Surgeons*, 187(4):358–364.

Jain, A. K., Duin, R. P. W., and Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on pattern analysis and machine intelligence*, 22(1):4–37.

Jang, J. Y., Kim, M. Y., Jeong, S. W., Kim, T. Y., Kim, S. U., Lee, S. H., Suk, K. T., Park, S. Y., Woo, H. Y., Kim, S. G., et al. (2013). Current consensus and guidelines of contrast enhanced ultrasound for the characterization of focal liver lesions. *Clinical and molecular hepatology*, 19(1):1.

Javed, S. G., Majid, A., Mirza, A. M., and Khan, A. (2016). Multi-denoising based impulse noise removal from images using robust statistical features and genetic programming. *Multimedia Tools and Applications*, 75(10):5887–5916.

Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA: a cancer journal for clinicians*, 61(2):69–90.

Ji, H., He, J., Yang, X., Deklerck, R., and Cornelis, J. (2013). Acm-based automatic liver segmentation from 3-d ct images by combining multiple atlases and improved mean-shift techniques. *IEEE journal of biomedical and health informatics*, 17(3):690–698.

Jiang, H., He, B., Fang, D., Ma, Z., Yang, B., and Zhang, L. (2013a). A region growing vessel segmentation algorithm based on spectrum information. *Computational and mathematical methods in medicine*, 2013.

Jiang, H., He, B., Fang, D., Ma, Z., Yang, B., and Zhang, L. (2013b). A region growing vessel segmentation algorithm based on spectrum information. *Computational and mathematical methods in medicine*, 2013.

Jimenez-Carretero, D., Fernandez-de Manuel, L., Pascau, J., Tellado, J. M., Ramon, E., Desco, M., Santos, A., and Ledesma-Carbayo, M. J. (2011). Optimal multiresolution 3d level-set method for liver segmentation incorporating local curvature constraints. In *Engineering in medicine and biology society, EMBC, 2011 annual international conference of the IEEE*, pages 3419–3422. IEEE.

Jin, J., Yang, L., Zhang, X., and Ding, M. (2013). Vascular tree segmentation in medical images using hessian-based multiscale filtering and level set method. *Computational and mathematical methods in medicine*, 2013.

Jordan, M. I. and Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260.

Jung, E., Schreyer, A., Schacherer, D., Menzel, C., Farkas, S., Loss, M., Feuerbach, S., Zorger, N., and Fellner, C. (2009). New real-time image fusion technique for characterization of tumor vascularisation and tumor perfusion of liver tumors with contrast-enhanced ultrasound, spiral ct or mri: first results. *Clinical hemorheology and microcirculation*, 43(1-2):57–69.

Jurie, F. and Triggs, B. (2005). Creating efficient codebooks for visual recognition. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 1, pages 604–610. IEEE.

Kainmüller, D., Lange, T., and Lamecker, H. (2007). Shape constrained automatic segmentation of the liver based on a heuristic intensity model. In *Proc. MICCAI Workshop 3D Segmentation in the Clinic: A Grand Challenge*, pages 109–116.

Kan, Z. and Madoff, D. C. (2008). Liver anatomy: microcirculation of the liver. In *Seminars in interventional radiology*, volume 25, pages 077–085. © Thieme Medical Publishers.

Kang, X., Zhao, Q., Sharma, K., Shekhar, R., Wood, B. J., and Linguraru, M. G. (2014a). Automatic labeling of liver veins in ct by probabilistic backward tracing. In *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on*, pages 1115–1118. IEEE.

Kang, X., Zhao, Q., Sharma, K., Shekhar, R., Wood, B. J., and Linguraru, M. G. (2014b). Automatic labeling of liver veins in ct by probabilistic backward tracing. In *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on*, pages 1115–1118. IEEE.

Kaur, R. and Juneja, M. (2018a). Comparison of different renal imaging modalities: an overview. In *Progress in intelligent computing techniques: theory, practice, and applications*, pages 47–57. Springer.

Kaur, R. and Juneja, M. (2018b). Comparison of different renal imaging modalities: an overview. In *Progress in intelligent computing techniques: theory, practice, and applications*, pages 47–57. Springer.

Keevil, S. F. (2001). Magnetic resonance imaging in medicine. *Physics Education*, 36(6):476.

Khan, S. A., Thomas, H. C., Davidson, B. R., and Taylor-Robinson, S. D. (2005). Cholangiocarcinoma. *The Lancet*, 366(9493):1303–1314.

Kim, T., Hori, M., and Onishi, H. (2009). Liver masses with central or eccentric scar. In *Seminars in Ultrasound, CT and MRI*, volume 30, pages 418–425. Elsevier.

Kingsbury, N. (2001). Complex wavelets for shift invariant analysis and filtering of signals. *Applied and computational harmonic analysis*, 10(3):234–253.

Kirbas, C. and Quek, F. (2004). A review of vessel extraction techniques and algorithms. *ACM Computing Surveys (CSUR)*, 36(2):81–121.

Kirbas, C. and Quek, F. K. (2003). Vessel extraction techniques and algorithms: a survey. In *Bioinformatics and bioengineering, 2003. Proceedings. Third IEEE symposium on*, pages 238–245. IEEE.

Kitasaka, T., Mori, K., Hasegawa, J.-i., Suenaga, Y., and Toriwaki, J.-i. (2003). Extraction of bronchus regions from 3d chest x-ray ct images by using structural features of bronchus. In *International Congress Series*, volume 1256, pages 240–245. Elsevier.

Kogure, K., Kuwano, H., Fujimaki, N., and Makuuchi, M. (2000). Relation among portal segmentation, proper hepatic vein, and external notch of the caudate lobe in the human liver. *Annals of surgery*, 231(2):223.

Kohavi, R. et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Stanford, CA.

Kokciyan, N., Turkay, R., Uskudarli, S., Yolum, P., Bakir, B., and Acar, B. (2014). Semantic description of liver ct images: An ontological approach. *IEEE journal of biomedical and health informatics*, 18(4):1363–1369.

Kondo, F., Fukusato, T., and Kudo, M. (2014). Pathological diagnosis of benign hepatocellular nodular lesions based on the new world health organization classification. *Oncology*, 87(Suppl. 1):37–49.

Kretowski, M. (2002). *Tissue modeling and classification in biomedical imaging*. PhD thesis, University of Rennes.

Krishan, A. and Mittal, D. (2015). Detection and classification of liver cancer using ct images. *International Journal on Recent Technologies in Mechanical and Electrical Engineering*, 2:93–98.

Krupinski, E. A. (2011). The role of perception in imaging: past and future. In *Seminars in nuclear medicine*, volume 41, pages 392–400. Elsevier.

Kumar, A., Dyer, S., Kim, J., Li, C., Leong, P. H., Fulham, M., and Feng, D. (2016). Adapting content-based image retrieval techniques for the semantic annotation of medical images. *Computerized Medical Imaging and Graphics*, 49:37–45.

Kumar, A., Dyer, S., Li, C., Leong, P. H. W., and Kim, J. (2014). Automatic annotation of liver ct images: the submission of the bmet group to imageclefmed 2014. In *CLEF (Working Notes)*, pages 428–437.

Kumar, A., Tripathi, A., and Jain, S. (2011). Extracorporeal bioartificial liver for treating acute liver diseases. *The journal of extra-corporeal technology*, 43(4):195–206.

Kumar, S., Moni, R., and Rajeesh, J. (2012). Liver tumor diagnosis by gray level and contourlet coefficients texture analysis. In *Computing, Electronics and Electrical Technologies (ICCEET), 2012 International Conference on*, pages 557–562. IEEE.

Kumar, S., Moni, R., and Rajeesh, J. (2013a). An automatic computer-aided diagnosis system for liver tumours on computed tomography images. *Computers & Electrical Engineering*, 39(5):1516–1526.

Kumar, S., Moni, R., and Rajeesh, J. (2013b). Automatic liver and lesion segmentation: a primary step in diagnosis of liver diseases. *Signal, Image and Video Processing*, 7(1):163–172.

Kurtz, C., Idoux, P.-A., Thangali, A., Cloppet, F., Beaulieu, C. F., and Rubin, D. L. (2015). Semantic retrieval of radiological images with relevance feedback. In *Multimodal Retrieval in the Medical Domain*, pages 11–25. Springer.

Laws, K. I. (1980). Textured image segmentation. Technical report, DTIC Document.

Lee, C.-C., Chiang, Y.-C., Tsai, C.-L., and Chen, S.-H. (2007). Distinction of liver disease from ct images using kernel-based classifiers. *International Journal of Intelligent Computing in Medical Sciences & Image Processing*, 1(2):113–120.

Lee, C.-C., Chung, P.-C., and Chen, Y.-J. (2005). Classification of liver diseases from ct images using bp-cmac neural network. In *Cellular Neural Networks and Their Applications, 2005 9th International Workshop on*, pages 118–121. IEEE.

Lee, J.-S. (1980). Digital image enhancement and noise filtering by use of local statistics. *IEEE transactions on pattern analysis and machine intelligence*, (2):165–168.

Lee, M., Kim, J. H., Park, M. H., Kim, Y.-H., Seong, Y. K., Cho, B. H., and Woo, K.-G. (2014). Computer-aided classification of liver tumors with combined deformable model segmentation and support vector machine. In *SPIE Medical Imaging*, pages 90341N–90341N. International Society for Optics and Photonics.

Lerski, R., Straughan, K., Schad, L., Boyce, D., Blüml, S., and Zuna, I. (1993). Viii. mr image texture analysisâĂŤan approach to tissue characterization. *Magnetic resonance imaging*, 11(6):873–887.

Leventon, M. E., Grimson, W. E. L., and Faugeras, O. (2000). Statistical shape influence in geodesic active contours. In *Computer vision and pattern recognition, 2000. Proceedings. IEEE conference on*, volume 1, pages 316–323. IEEE.

Li, C., Wang, X., Eberl, S., Fulham, M., Yin, Y., and Feng, D. (2010). Fully automated liver segmentation for low-and high-contrast ct volumes based on probabilistic atlases. In *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pages 1733–1736. IEEE.

Li, G., Chen, X., Shi, F., Zhu, W., Tian, J., and Xiang, D. (2015). Automatic liver segmentation based on shape constraints and deformable graph cut in ct images. *IEEE Transactions on Image Processing*, 24(12):5315–5329.

Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., and Liu, H. (2016). Feature selection: A data perspective. *arXiv preprint arXiv:1601.07996*.

Li, J., Du, Q., and Sun, C. (2009). An improved box-counting method for image fractal dimension estimation. *Pattern Recognition*, 42(11):2460–2469.

Li, J., Yin, J., Sun, J., and Guo, Z. (2013). A weighted dual porous medium equation applied to image restoration. *Mathematical Methods in the Applied Sciences*, 36(16):2117–2127.

Liao, M., Zhao, Y.-q., Liu, X.-y., Zeng, Y.-z., Zou, B.-j., Wang, X.-f., and Shih, F. Y. (2017). Automatic liver segmentation from abdominal ct volumes using graph cuts and border marching. *Computer methods and programs in biomedicine*, 143:1–12.

Lin, D.-T., Lei, C.-C., and Hung, S.-W. (2006). Computer-aided kidney segmentation on abdominal ct images. *IEEE transactions on information technology in biomedicine*, 10(1):59–65.

Lin, Y.-L., Lin, C.-L., Cheng, S.-C., Yang, C.-K., and Wu, Y.-H. (2016). A roi-bag approach for automatic liver cirrhosis diagnosis using ultrasound images. *Journal of Marine Science and Technology*, 24(3):548–561.

Ling, J.-M. and Liu, P.-H. (2015). Economic analysis of lagrangian and genetic algorithm for the optimal capacity planning of photovoltaic generation. *Mathematical Problems in Engineering*, 2015.

Linguraru, M. G., Wang, S., Shah, F., Gautam, R., Peterson, J., Linehan, W. M., and Summers, R. M. (2011). Automated noninvasive classification of renal cancer on multiphase ct. *Medical physics*, 38(10):5738–5746.

Liu, B., Zhang, H., Hua, S., Jiang, Q., Huang, R., Liu, W., Zhang, S., Zhang, B., and Yue, Z. (2016). An automatic segmentation system of acetabulum in sequential ct images for the personalized artificial femoral head design. *Computer methods and programs in biomedicine*, 127:318–335.

Liu, W., Wu, J., Zuo, L., Yuan, H., and Zhao, H. (2012). Rapid image segmentation using color, texture and syntactic visual features. *Artificial Intelligence and Computational Intelligence*, pages 424–434.

Liu, Y. and Zheng, Y. F. (2005). One-against-all multi-class svm classification using reliability measures. In *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*, volume 2, pages 849–854. IEEE.

López-Mir, F., González, P., Naranjo, V., Pareja, E., Morales, S., and Solaz-Mínguez, J. (2015). A method for liver segmentation on computed tomography images in venous phase suitable for real environments. *Journal of Medical Imaging and Health Informatics*, 5(6):1208–1216.

López-Mir, F., González, P., Naranjo, V., Pareja, E., Raya, M. A., and Solaz, J. (2013). Liver segmentation on ct images. a fast computational method based on 3d morphology and a statistical filter. In *IWBBIO*, pages 483–490.

Lu, S., Huang, H., Liang, P., Chen, G., and Xiao, L. (2017). Hepatic vessel segmentation using variational level set combined with non-local robust statistics. *Magnetic resonance imaging*, 36:180–186.

Lu, X., Wu, J., Ren, X., Zhang, B., and Li, Y. (2014). The study and application of the improved region growing algorithm for liver segmentation. *Optik-International Journal for Light and Electron Optics*, 125(9):2142–2147.

Luo, S., Hu, Q., He, X., Li, J., Jin, J. S., and Park, M. (2009). Automatic liver parenchyma segmentation from abdominal ct images using support vector machines. In *Complex Medical Engineering, 2009. CME. ICME International Conference on*, pages 1–5. IEEE.

Majno, P., Mentha, G., Toso, C., Morel, P., Peitgen, H. O., and Fasel, J. H. (2014). Anatomy of the liver: an outline with three levels of complexity–a further step towards tailored territorial liver resections. *Journal of hepatology*, 60(3):654–662.

Mala, K. and Sadasivam, V. (2010). Classification of fatty and cirrhosis liver using wavelet-based statistical texture features and neural network classifier. *International Journal of Software and Informatics*, 4(2):151–163.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7):674–693.

Mandelbrot, B. B. and Pignoni, R. (1983). *The fractal geometry of nature*, volume 1. WH freeman New York.

Manjunath, B. S., Salembier, P., and Sikora, T. (2002). *Introduction to MPEG-7: multimedia content description interface*, volume 1. John Wiley & Sons.

Martin, D. R., Kalb, B., Sarmiento, J. M., Heffron, T. G., Coban, I., and Adsay, N. V. (2010). Giant and complicated variants of cystic bile duct hamartomas of the liver: Mri findings and pathological correlations. *Journal of Magnetic Resonance Imaging*, 31(4):903–911.

Maruyama, M., Isokawa, O., Hoshiyama, K., Hoshiyama, A., Hoshiyama, M., and Hoshiyama, Y. (2013). Diagnosis and management of giant hepatic hemangioma: the usefulness of contrast-enhanced ultrasonography. *International journal of hepatology*, 2013.

Marvasti, N. B., García, M. d. M. R., Üsküdarli, S., Montes, J. F. A., and Acar, B. (2015). Overview of the imageclef 2015 liver ct annotation task. In *CLEF (Working Notes)*.

Marvasti, N. B., Yoruk, E., and Acar, B. (2017). Computer-aided medical image annotation: Preliminary results with liver lesions in ct. *IEEE Journal of Biomedical and Health Informatics*.

Massoptier, L. and Casciaro, S. (2007). Fully automatic liver segmentation through graph-cut technique. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 5243–5246. IEEE.

Massoptier, L. and Casciaro, S. (2008). A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from ct scans. *European radiology*, 18(8):1658.

McCall, J. (2005). Genetic algorithms for modelling and optimisation. *Journal of Computational and Applied Mathematics*, 184(1):205–222.

McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157.

Mejino Jr, J. L., Rubin, D. L., and Brinkley, J. F. (2008). Fma-radlex: An application ontology of radiological anatomy derived from the foundational model of anatomy reference ontology. In *AMIA Annual Symposium Proceedings*, volume 2008, page 465. American Medical Informatics Association.

Memon, N. A., Mirza, A. M., and Gilani, S. (2006). Segmentation of lungs from ct scan images for early diagnosis of lung cancer. In *Proceedings of world academy of science, engineering and technology*, volume 14.

Miliadis, L., Giannakopoulos, T., Boutsikos, G., Terzis, I., and Kyriazanos, I. D. (2010). Spontaneous rupture of a large non-parasitic liver cyst: a case report. *Journal of medical case reports*, 4(1):2.

Militzer, A., Hager, T., Jager, F., Tietjen, C., and Hornegger, J. (2010). Automatic detection and segmentation of focal liver lesions in contrast enhanced ct images. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 2524–2527. IEEE.

Minami, Y. and Kudo, M. (2015). Imaging modalities for assessment of treatment response to nonsurgical hepatocellular carcinoma therapy: contrast-enhanced us, ct, and mri. *Liver cancer*, 4(2):106–114.

Mir, A., Hanmandlu, M., and Tandon, S. (1995). Texture analysis of ct images. *IEEE Engineering in Medicine and Biology Magazine*, 14(6):781–786.

Mitra, V. and Metcalf, J. (2012). Functional anatomy and blood supply of the liver. *Anaesthesia & intensive care medicine*, 13(2):52–53.

Mitrea, D., Nedevschi, S., Lupsor, M., Socaciu, M., and Badea, R. (2009). Classification of the hepatocellular carcinoma in ultrasound images based on the imagistic textural model of this tumor. In *International Conference on Advancements of Medicine and Health Care through Technology*, pages 267–272. Springer.

Moghbel, M., Mashohor, S., Mahmud, R., and Saripan, M. I. B. (2016). Automatic liver tumor segmentation on computed tomography for patient treatment planning and monitoring. *EXCLI journal*, 15:406.

Moreira, R. K. (2015). Malignant pediatric liver tumors. In *Surgical Pathology of Liver Tumors*, pages 403–433. Springer.

Mougiakakou, S. G., Valavanis, I. K., Nikita, A., and Nikita, K. S. (2007a). Differential diagnosis of ct focal liver lesions using texture features, feature selection and ensemble driven classifiers. *Artificial Intelligence in Medicine*, 41(1):25–37.

Mougiakakou, S. G., Valavanis, I. K., Nikita, A., and Nikita, K. S. (2007b). Differential diagnosis of ct focal liver lesions using texture features, feature selection and ensemble driven classifiers. *Artificial Intelligence in Medicine*, 41(1):25–37.

Mule, S., Colosio, A., Cazejust, J., Kianmanesh, R., Soyer, P., and Hoeffel, C. (2015). Imaging of the postoperative liver: review of normal appearances and common complications. *Abdominal imaging*, 40(7):2761–2776.

Müller, H., Michoux, N., Bandon, D., and Geissbuhler, A. (2004). A review of content-based image retrieval systems in medical applicationsâĂŤclinical benefits and future directions. *International journal of medical informatics*, 73(1):1–23.

Murakami, T. and Tsurusaki, M. (2014a). Hypervascular benign and malignant liver tumors that require differentiation from hepatocellular carcinoma: key points of imaging diagnosis. *Liver Cancer*, 3(2):85–96.

Murakami, T. and Tsurusaki, M. (2014b). Hypervascular benign and malignant liver tumors that require differentiation from hepatocellular carcinoma: key points of imaging diagnosis. *Liver Cancer*, 3(2):85.

Nam, C. Y., Chaudhari, V., Raman, S. S., Lassman, C., Tong, M. J., Busuttil, R. W., and Lu, D. S. (2011). Ct and mri improve detection of hepatocellular carcinoma, compared with ultrasound alone, in patients with cirrhosis. *Clinical Gastroenterology and Hepatology*, 9(2):161–167.

Navarro, A. P., Gomez, D., Lamb, C. M., Brooks, A., and Cameron, I. C. (2014). Focal nodular hyperplasia: a review of current indications for and outcomes of hepatic resection. *HPB*, 16(6):503–511.

Nedjar, I., Mahmoudi, S., Chikh, M. A., Abi-Yad, K., and Bouafia, Z. (2015). Automatic annotation of liver ct image: Imageclefmed 2015. In *CLEF (Working Notes)*.

Nicolau, C., Vilana, R., Catalá, V., Bianchi, L., Gilabert, R., García, A., and Brú, C. (2006). Importance of evaluating all vascular phases on contrast-enhanced sonography in the differentiation of benign from malignant focal liver lesions. *American Journal of Roentgenology*, 186(1):158–167.

Ning, J., Zhang, L., Zhang, D., and Wu, C. (2010). Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition*, 43(2):445–456.

Obayya, M. I., Areed, N. F., and Abdulhadi, A. O. (2016). Liver cancer identification using adaptive neuro-fuzzy inference system. *International Journal of Computer Applications*, 140(8).

Oberkampf, H., Zillner, S., Overton, J. A., Bauer, B., Cavallaro, A., Uder, M., and Hammon, M. (2015). Semantic representation of reported measurements in radiology. *BMC medical informatics and decision making*, 16(1):5.

Oliveira, D. A., Feitosa, R. Q., and Correia, M. M. (2011a). Segmentation of liver, its vessels and lesions from ct images for surgical planning. *Biomedical engineering online*, 10(1):30.

Oliveira, D. A., Feitosa, R. Q., and Correia, M. M. (2011b). Segmentation of liver, its vessels and lesions from ct images for surgical planning. *Biomedical engineering online*, 10(1):30.

Osher, S. and Fedkiw, R. (2006). Level set methods and dynamic implicit surfaces. *Implicit Surfaces*, 153.

Osowski, S., Siroic, R., Markiewicz, T., and Siwek, K. (2009). Application of support vector machine and genetic algorithm for improved blood cell recognition. *IEEE Transactions on Instrumentation and Measurement*, 58(7):2159–2168.

Pamulapati, V., Venkatesan, A., Wood, B. J., and Linguraru, M. G. (2011). Liver segmental anatomy and analysis from vessel and tumor segmentation via optimized graph cuts. In *International MICCAI workshop on computational and clinical challenges in abdominal imaging*, pages 189–197. Springer.

Pappu, V. and Pardalos, P. M. (2014). High-dimensional data classification. In *Clusters, Orders, and Trees: Methods and Applications*, pages 119–150. Springer.

Park, H., Bland, P. H., and Meyer, C. R. (2003a). Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Transactions on medical imaging*, 22(4):483–492.

Park, H., Bland, P. H., and Meyer, C. R. (2003b). Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Transactions on medical imaging*, 22(4):483–492.

Park, S.-J., Seo, K.-S., and Park, J.-A. (2005). Automatic hepatic tumor segmentation using statistical optimal threshold. In *International Conference on Computational Science*, pages 934–940. Springer.

Paulson, E. K. (2001). Evaluation of the liver for metastatic disease. In *Seminars in liver disease*, volume 21, pages 225–236. Copyright© 2001 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA. Tel.:+ 1 (212) 584-4662.

Peng, J., Wang, Y., and Kong, D. (2014). Liver segmentation with constrained convex variational model. *Pattern Recognition Letters*, 43:81–88.

Peris, A., Zagli, G., Bonizzoli, M., Cianchi, G., Ciapetti, M., Spina, R., Anichini, V., Lapi, F., and Batacchi, S. (2010). Implantation of 3951 long-term central venous catheters: performances, risk analysis, and patient comfort after ultrasound-guidance introduction. *Anesthesia & Analgesia*, 111(5):1194–1201.

Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, 12(7):629–639.

Pham, D. L., Xu, C., and Prince, J. L. (2000). Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2(1):315–337.

Pichler, B. J., Judenhofer, M. S., and Pfannenberg, C. (2008). Multimodal imaging approaches: Pet/ct and pet/mri. In *Molecular Imaging I*, pages 109–132. Springer.

Platero, C., Tobar, M. C., Sanguino, J., Poncela, J. M., and Velasco, O. (2011). Level set segmentation with shape and appearance models using affine moment descriptors. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 109–116. Springer.

Pourghassem, H. and Ghassemian, H. (2008). Content-based medical image classification using a new hierarchical merging scheme. *Computerized Medical Imaging and Graphics*, 32(8):651–661.

QIAN, Y.-j. (2017). A method for quickly and exactly extracting hepatic vessel. *DEStech Transactions on Computer Science and Engineering*, (cece).

Quatrehomme, A., Millet, I., Hoa, D., Subsol, G., and Puech, W. (2012). Assessing the classification of liver focal lesions by using multi-phase computer tomography scans. In *MICCAI International Workshop on Medical Content-Based Retrieval for Clinical Decision Support*, pages 80–91. Springer.

Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.

Rademaker, J., Widjaja, A., and Galanski, M. (2000a). Hepatic hemangiosarcoma: imaging findings and differential diagnosis. *European radiology*, 10(1):129–133.

Rademaker, J., Widjaja, A., and Galanski, M. (2000b). Hepatic hemangiosarcoma: imaging findings and differential diagnosis. *European radiology*, 10(1):129–133.

Rajesh, S., Mukund, A., and Arora, A. (2015). Imaging diagnosis of splanchnic venous thrombosis. *Gastroenterology research and practice*, 2015.

Ramamoorthy, S., Kirubakaran, R., and Subramanian, R. S. (2015). Texture feature extraction using mgrlbp method for medical image classification. In *Artificial Intelligence and Evolutionary Algorithms in Engineering Systems*, pages 747–753. Springer.

Rao, S.-X., Lambregts, D. M., Schnerr, R. S., van Ommen, W., van Nijnatten, T. J., Martens, M. H., Heijnen, L. A., Backes, W. H., Verhoef, C., Zeng, M.-S., et al. (2014). Whole-liver ct texture analysis in colorectal cancer: Does the presence of liver metastases affect the texture of the remaining liver? *United European gastroenterology journal*, 2(6):530–538.

Razi, T., Niknami, M., and Ghazani, F. A. (2014). Relationship between hounsfield unit in ct scan and gray scale in cbct. *Journal of dental research, dental clinics, dental prospects*, 8(2):107.

Reddy, K. R. and Faust, T. (2005). *The clinician's guide to liver disease*. Slack Incorporated.

Reddy, K. R. and Faust, T. (2006). *The clinician's guide to liver disease*. Slack Incorporated.

Reitinger, B., Bornik, A., Beichel, R., and Schmalstieg, D. (2006). Liver surgery planning using virtual reality. *IEEE Computer Graphics and Applications*, 26(6).

Ressom, H. W., Varghese, R. S., Zhang, Z., Xuan, J., and Clarke, R. (2008). Classification algorithms for phenotype prediction in genomics and proteomics. *Frontiers in bioscience: a journal and virtual library*, 13:691.

Reynolds, M. (2001). Pediatric liver tumors. In *Liver-Directed Therapy for Primary and Metastatic Liver Tumors*, pages 299–312. Springer.

Robinson, P. (1997). Radiology's achilles' heel: error and variation in the interpretation of the röntgen image. *The British Journal of Radiology*, 70(839):1085–1098.

Robinson, P. (2000). Imaging liver metastases: current limitations and future prospects. *The British journal of radiology*, 73(867):234–241.

Rockey, D. C., Caldwell, S. H., Goodman, Z. D., Nelson, R. C., and Smith, A. D. (2009a). Liver biopsy. *Hepatology*, 49(3):1017–1044.

Rockey, D. C., Caldwell, S. H., Goodman, Z. D., Nelson, R. C., and Smith, A. D. (2009b). Liver biopsy. *Hepatology*, 49(3):1017–1044.

Rogowska, J. (2000). Overview and fundamentals of medical image segmentation. *Handbook of medical imaging, processing and analysis*, pages 69–85.

Roy, S., Chi, Y., Liu, J., Venkatesh, S. K., and Brown, M. S. (2014a). Three-dimensional spatiotemporal features for fast content-based retrieval of focal liver lesions. *IEEE Transactions on Biomedical Engineering*, 61(11):2768–2778.

Roy, S., Chi, Y., Liu, J., Venkatesh, S. K., and Brown, M. S. (2014b). Three-dimensional spatiotemporal features for fast content-based retrieval of focal liver lesions. *IEEE Transactions on Biomedical Engineering*, 61(11):2768–2778.

Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268.

Ruskó, L., Bekes, G., and Fidrich, M. (2009a). Automatic segmentation of the liver from multi-and single-phase contrast-enhanced ct images. *Medical Image Analysis*, 13(6):871–882.

Ruskó, L., Bekes, G., and Fidrich, M. (2009b). Automatic segmentation of the liver from multi-and single-phase contrast-enhanced ct images. *Medical Image Analysis*, 13(6):871–882.

Rusko, L., Bekes, G., Nemeth, G., and Fidrich, M. (2007a). Fully automatic liver segmentation for contrast-enhanced ct images. *MICCAI Wshp. 3D Segmentation in the Clinic: A Grand Challenge*, 2(7).

Rusko, L., Bekes, G., Nemeth, G., and Fidrich, M. (2007b). Fully automatic liver segmentation for contrast-enhanced ct images. *MICCAI Wshp. 3D Segmentation in the Clinic: A Grand Challenge*, 2(7).

Rutkauskas, S., Gedrimas, L. V., Pundzius, J., Barauskas, G., and Basevičius, A. (2006). Clinical and anatomical basis for the classification of the structural parts of liver. *Medicina*, 42(2):98–106.

Sabih, A., Sabih, Q., Khan, A. N., et al. (2011). Image perception and interpretation of abnormalities; can we believe our eyes? can we do something about it? *Insights into imaging*, 2(1):47–55.

Saddi, K. A., Rousson, M., ChefdâĂŹhotel, C., and Cheriet, F. (2007). Global-to-local shape matching for liver segmentation in ct imaging. In *Proceedings of MICCAI workshop on 3D segmentation in the clinic: a grand challenge*, pages 207–214.

Safdari, M., Pasari, R., Rubin, D., and Greenspan, H. (2013). Image patch-based method for automated classification and detection of focal liver lesions on ct. In *SPIE Medical Imaging*, pages 86700Y–86700Y. International Society for Optics and Photonics.

Saffari, S. E., Löve, Á., Fredrikson, M., and Smedby, Ö. (2015). Regression models for analyzing radiological visual grading studies–an empirical comparison. *BMC medical imaging*, 15(1):49.

Sahani, D. V. and Kalva, S. P. (2004a). Imaging the liver. *The oncologist*, 9(4):385–397.

Sahani, D. V. and Kalva, S. P. (2004b). Imaging the liver. *The oncologist*, 9(4):385–397.

Saito, A., Yamamoto, S., Nawano, S., and Shimizu, A. (2017). Automated liver segmentation from a postmortem ct scan based on a statistical shape model. *International journal of computer assisted radiology and surgery*, 12(2):205–221.

Sanches, J. M., Nascimento, J. C., and Marques, J. S. (2008). Medical image noise reduction using the sylvester–lyapunov equation. *IEEE transactions on image processing*, 17(9):1522–1539.

Sankar, D. and Thomas, T. (2010). Fractal features based on differential box counting method for the categorization of digital mammograms. *Journal of Computer Information Systems and Industrial Management Applications*, 2:11–19.

Schenck, J. F. (2000). Safety of strong, static magnetic fields. *Journal of magnetic resonance imaging*, 12(1):2–19.

Schiff, E. R. (2009). Advances in hepatology: Current developments in the treatment of hepatitis and hepatobiliary disease. *Gastroenterology & hepatology*, 5(6):414.

Schmidt, J., Strotzer, M., Fraunhofer, S., Boedeker, H., and Zirngibl, H. (2000). Intraoperative ultrasonography versus helical computed tomography and computed tomography with arterioportography in diagnosing colorectal liver metastases: lesion-by-lesion analysis. *World journal of surgery*, 24(1):43–48.

Schnater, J. M., Köhler, S. E., Lamers, W. H., von Schweinitz, D., and Aronson, D. C. (2003). Where do we stand with hepatoblastoma? *Cancer*, 98(4):668–678.

School, H. M. (2014). Liver cancer. http://www.health.harvard.edu/cancer/liver-cancer.

Seghers, D., Slagmolen, P., Lambelin, Y., Hermans, J., Loeckx, D., Maes, F., and Suetens, P. (2007a). Landmark based liver segmentation using local shape and local intensity models. In *Proc. Workshop of the 10th Int. Conf. on MICCAI, Workshop on 3D Segmentation in the Clinic: A Grand Challenge*, pages 135–142.

Seghers, D., Slagmolen, P., Lambelin, Y., Hermans, J., Loeckx, D., Maes, F., and Suetens, P. (2007b). Landmark based liver segmentation using local shape and local intensity models. In *Proc. Workshop of the 10th Int. Conf. on MICCAI, Workshop on 3D Segmentation in the Clinic: A Grand Challenge*, pages 135–142.

Seghers, D., Slagmolen, P., Lambelin, Y., Hermans, J., Loeckx, D., Maes, F., and Suetens, P. (2007c). Landmark based liver segmentation using local shape and local intensity models. In *Proc. Workshop of the 10th Int. Conf. on MICCAI, Workshop on 3D Segmentation in the Clinic: A Grand Challenge*, pages 135–142.

Selle, D., Preim, B., Schenk, A., and Peitgen, H.-O. (2002a). Analysis of vasculature for liver surgical planning. *IEEE transactions on medical imaging*, 21(11):1344–1357.

Selle, D., Preim, B., Schenk, A., and Peitgen, H.-O. (2002b). Analysis of vasculature for liver surgical planning. *IEEE transactions on medical imaging*, 21(11):1344–1357.

Selle, D., Preim, B., Schenk, A., and Peitgen, H.-O. (2002c). Analysis of vasculature for liver surgical planning. *IEEE transactions on medical imaging*, 21(11):1344–1357.

Selvalakshmi, V. et al. (2017). A novel region based segmentation of hepatic tumors and hepatic vein in low contrast cta images using bernstein polynomials. *Biomedical Research*.

Seo, K.-S., Kim, H.-B., Park, T., Kim, P.-K., and Park, J.-A. (2005). Automatic liver segmentation of contrast enhanced ct images based on histogram processing. *Advances in Natural Computation*, pages 421–421.

Seo, K.-S. and Park, J.-A. (2005). Improved automatic liver segmentation of a contrast enhanced ct image. In *Pacific-Rim Conference on Multimedia*, pages 899–909. Springer.

Sethian, J. A. (1999). *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, volume 3. Cambridge university press.

Setia, L., Teynor, A., Halawani, A., and Burkhardt, H. (2008). Grayscale medical image annotation using local relational features. *Pattern Recognition Letters*, 29(15):2039–2045.

Shang, Q., Clements, L., Galloway, R. L., Chapman, W. C., and Dawant, B. M. (2008). Adaptive directional region growing segmentation of the hepatic vasculature. In *Medical Imaging 2008: Image Processing*, volume 6914, page 69141F. International Society for Optics and Photonics.

Shimizu, A. (2006). Multi-organ segmentation in three dimensional abdominal ct images. *Proc of Computer Assisted Radiology and Surgery, 2006*.

Shimizu, A., Nakagomi, K., Narihira, T., Kobatake, H., Nawano, S., Shinozaki, K., Ishizu, K., and Togashi, K. (2010). Automated segmentation of 3d ct images based on statistical atlas and graph cuts. In *International MICCAI Workshop on Medical Computer Vision*, pages 214–223. Springer.

Shoemaker, M., Barrie, M., Holman, H., Wolk, K. E., Stromberg, P. C., and Aeffner, F. (2016). Pathology in practice. *Journal of the American Veterinary Medical Association*, 248(2):153–155.

Sibulesky, L. (2013). Normal liver anatomy. *Clinical Liver Disease*, 2(S1).

Siedlecki, W. and Sklansky, J. (1989). A note on genetic algorithms for large-scale feature selection. *Pattern recognition letters*, 10(5):335–347.

Šimundić, A.-M. (2008). Measures of diagnostic accuracy: basic definitions. *Medical and biological sciences*, 22(4):61–65.

Singh, M., Singh, S., and Gupta, S. (2014). Investigations on roi selection for liver classification. In *Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on*, pages 1–6. IEEE.

Sipper, M., Fu, W., Ahuja, K., and Moore, J. H. (2018). Investigating the parameter space of evolutionary algorithms. *BioData Mining*, 11(1):2.

Slagmolen, P., Elen, A., Seghers, D., Loeckx, D., Maes, F., and Haustermans, K. (2007). Atlas based liver segmentation using nonrigid registration with a b-spline transformation model. In *Proceedings of MICCAI workshop on 3D segmentation in the clinic: a grand challenge*, pages 197–206. Citeseer.

Smeets, D., Stijnen, B., Loeckx, D., De Dobbelaer, B., and Suetens, P. (2008). Segmentation of liver metastases using a level set method with spiral-scanning technique and supervised fuzzy pixel classification. In *MICCAI workshop*, volume 42, page 43.

Smeulders, A. W., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on pattern analysis and machine intelligence*, 22(12):1349–1380.

Socas, L., Zumbado, M., Perez-Luzardo, O., Ramos, A., Perez, C., Hernandez, J., and Boada, L. (2005). Hepatocellular adenomas associated with anabolic androgenic steroid abuse in bodybuilders: a report of two cases and a review of the literature. *British journal of sports medicine*, 39(5):e27–e27.

Soler, L., Delingette, H., Malandain, G., Montagnat, J., Ayache, N., Koehl, C., Dourthe, O., Malassagne, B., Smith, M., Mutter, D., et al. (2001a). Fully automatic anatomical, pathological, and functional segmentation from ct scans for hepatic surgery. *Computer Aided Surgery*, 6(3):131–142.

Soler, L., Delingette, H., Malandain, G., Montagnat, J., Ayache, N., Koehl, C., Dourthe, O., Malassagne, B., Smith, M., Mutter, D., et al. (2001b). Fully automatic anatomical, pathological, and functional segmentation from ct scans for hepatic surgery. *Computer Aided Surgery*, 6(3):131–142.

Spanier, A. B. and Joskowicz, L. (2014). Towards content-based image retrieval: From computer generated features to semantic descriptions of liver ct scans. In *CLEF (Working Notes)*, pages 438–447.

Stawiaski, J., Decenciere, E., and Bidault, F. (2008). Interactive liver tumor segmentation using graph-cuts and watershed. In *Workshop on 3D segmentation in the clinic: a grand challenge II. Liver tumor segmentation challenge. MICCAI, New York, USA*.

Stoitsis, J., Valavanis, I., Mougiakakou, S. G., Golemati, S., Nikita, A., and Nikita, K. S. (2006). Computer aided diagnosis based on medical image processing and artificial intelligence methods. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 569(2):591–595.

Stringer, M. D. (2014). Anatomy and development of the liver. In *Diseases of the Liver in Children*, pages 3–21. Springer.

Suh, Y. J., Kim, M.-J., Choi, J.-Y., Park, Y. N., Park, M.-S., and Kim, K. W. (2011). Differentiation of hepatic hyperintense lesions seen on gadoxetic acid–enhanced hepatobiliary phase mri. *American Journal of Roentgenology*, 197(1):W44–W52.

Sun, X., Liu, Y., Wei, D., Xu, M., Chen, H., and Han, J. (2013). Selection of interdependent genes via dynamic relevance analysis for cancer diagnosis. *Journal of biomedical informatics*, 46(2):252–258.

Susomboon, R., Raicu, D., Furst, J., and Johnson, T. B. (2008). A co-occurrence texture semi-invariance to direction, distance and patient size. In *Proc. SPIE Medical Imaging*, volume 6914.

Susomboon, R., Raicu, D. S., and Furst, J. (2007). A hybrid approach for liver segmentation. In *Proceedings of MICCAI workshop on 3D segmentation in the clinic: a grand challenge*, pages 151–160.

Suzuki, K. (2011). Computerized segmentation of organs by means of geodesic active-contour level-set algorithm. In *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, pages 103–128. Springer.

Suzuki, K., Huynh, H. T., Liu, Y., Calabrese, D., Zhou, K., Oto, A., and Hori, M. (2013). Computerized segmentation of liver in hepatic ct and mri by means of level-set geodesic active contouring. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 2984–2987. IEEE.

Tang, J., Alelyani, S., and Liu, H. (2014). Feature selection for classification: A review. *Data Classification: Algorithms and Applications*, page 37.

Taylor, G. A. (2017). Perceptual errors in pediatric radiology. *Diagnosis*, 4(3):141–147.

Toledo, R., Orriols, X., Radeva, P., Binefa, X., Vitria, J., Canero, C., and Villanuev, J. (2000). Eigensnakes for vessel segmentation in angiography. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 4, pages 340–343. IEEE.

Tommasi, T., Orabona, F., and Caputo, B. (2008). Discriminative cue integration for medical image annotation. *Pattern Recognition Letters*, 29(15):1996–2002.

Toriyabe, Y., Nishimura, T., Kita, S., Saito, Y., and Miyokawa, N. (1997). Differentiation between benign and metastatic cervical lymph nodes with ultrasound. *Clinical radiology*, 52(12):927–932.

Tourassi, G., Voisin, S., Paquit, V., and Krupinski, E. (2013). Investigating the link between radiologists' gaze, diagnostic decision, and image content. *Journal of the American Medical Informatics Association*, 20(6):1067–1075.

Tsagaan, B., Shimizu, A., Kobatake, H., Miyakawa, K., and Hanzawa, Y. (2001). Segmentation of kidney by using a deformable model. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 3, pages 1059–1062. IEEE.

Tsao, J., Boesiger, P., and Pruessmann, K. P. (2003). k-t blast and k-t sense: Dynamic mri with high frame rate exploiting spatiotemporal correlations. *Magnetic Resonance in Medicine*, 50(5):1031–1042.

Ulagamuthalvi, V., Kulanthaive, G., and Sridharan, D. (2012). a novel approach for diagnosing liver lesion images in telemedicine mode. *Journal of Health Science*, 2(4):50–53.

Unser, M. (1986). Sum and difference histograms for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1):118–125.

Van Gemert, J. C., Veenman, C. J., Smeulders, A. W., and Geusebroek, J.-M. (2010). Visual word ambiguity. *IEEE transactions on pattern analysis and machine intelligence*, 32(7):1271–1283.

van Herk, M. and Kooy, H. M. (1994). Automatic three-dimensional correlation of ct-ct, ct-mri, and ct-spect using chamfer matching. *Medical physics*, 21(7):1163–1178.

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.

Venturi, A., Piscaglia, F., Vidili, G., Flori, S., Righini, R., Golfieri, R., and Bolondi, L. (2007). Diagnosis and management of hepatic focal nodular hyperplasia. *Journal of ultrasound*, 10(3):116–127.

Vilgrain, V., Boulos, L., Vullierme, M.-P., Denys, A., Terris, B., and Menu, Y. (2000). Imaging of atypical hemangiomas of the liver with pathologic correlation 1. *Radiographics*, 20(2):379–397.

Walker, R. F., Jackway, P., and Longstaff, I. (1995). Improving co-occurrence matrix feature discrimination. In *DICTA'95, 3rd Conference on Digital Image Computing: Techniques and Application*, pages 643–648.

Wang, G., Zhang, S., Li, F., and Gu, L. (2013a). A new segmentation framework based on sparse shape composition in liver surgery planning system. *Medical physics*, 40(5).

Wang, G., Zhang, S., Li, F., and Gu, L. (2013b). A new segmentation framework based on sparse shape composition in liver surgery planning system. *Medical physics*, 40(5).

Wang, J., Cheng, Y., Guo, C., Wang, Y., and Tamura, S. (2016). Shape–intensity prior level set combining probabilistic atlas and probability map constrains for automatic liver segmentation from abdominal ct images. *International journal of computer assisted radiology and surgery*, 11(5):817–826.

Wang, J., Han, X.-H., Xu, Y., Lin, L., Hu, H., Jin, C., and Chen, Y.-W. (2017). Sparse codebook model of local structures for retrieval of focal liver lesions using multi-phase medical images. *International journal of biomedical imaging*, 2017.

Wang, J., Yang, J., Yu, K., Lv, F., Huang, T., and Gong, Y. (2010). Locality-constrained linear coding for image classification. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 3360–3367. IEEE.

Wang, L., Zhang, Z., Liu, J., Jiang, B., Duan, X., Xie, Q., Hu, D., and Li, Z. (2009). Classification of hepatic tissues from ct images based on texture features and multiclass support vector machines. *Advances in Neural Networks–ISNN 2009*, pages 374–381.

Wang, X., Yang, J., Ai, D., Zheng, Y., Tang, S., and Wang, Y. (2015a). Adaptive mesh expansion model (amem) for liver segmentation from ct image. *PloS one*, 10(3):e0118064.

Wang, Z., Yang, J., Zheng, Y., Ai, D., Xia, L., and Wang, Y. (2015b). A novel image representation method for liver tumor classification.

Weisstein, E. W. and Weisstein, E. W. (2009). *CRC encyclopedia of mathematics*. CRC Press Boca Raton, FL.

Wennerberg, P., Schulz, K., and Buitelaar, P. (2011). Ontology modularization to improve semantic medical image annotation. *Journal of biomedical informatics*, 44(1):155–162.

Weszka, J. S., Dyer, C. R., and Rosenfeld, A. (1976). A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man, and Cybernetics*, (4):269–285.

Wimmer, A., Soza, G., and Hornegger, J. (2007). Two-stage semi-automatic organ segmentation framework using radial basis functions and level sets. *3D segmentation in the clinic: a grand challenge*, pages 179–188.

Winn, J., Criminisi, A., and Minka, T. (2005). Object categorization by learned universal visual dictionary. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1800–1807. IEEE.

Wintermark, M., Meuli, R., Browaeys, P., Reichhart, M., Bogousslavsky, J., Schnyder, P., and Michel, P. (2007). Comparison of ct perfusion and angiography and mri in selecting stroke patients for acute treatment. *Neurology*, 68(9):694–697.

Wippold, F. J. (2007). Head and neck imaging: the role of ct and mri. *Journal of magnetic resonance imaging*, 25(3):453–465.

Wolf, D. C. (1990). Evaluation of the size, shape, and consistency of the liver.

Wu, K.-L. and Yang, M.-S. (2002). Alternative c-means clustering algorithms. *Pattern recognition*, 35(10):2267–2278.

Wu, K.-L. and Yang, M.-S. (2005). A cluster validity index for fuzzy clustering. *Pattern Recognition Letters*, 26(9):1275–1291.

Wu, W., Wu, S., Zhou, Z., Zhang, R., and Zhang, Y. (2017). 3d liver tumor segmentation in ct images using improved fuzzy c-means and graph cuts. *BioMed research international*, 2017.

Wu, W., Zhou, Z., Wu, S., and Zhang, Y. (2016). Automatic liver segmentation on volumetric ct images using supervoxel-based graph cuts. *Computational and mathematical methods in medicine*, 2016.

Yang, W., Feng, Q., Huang, M., Lu, Z., and Chen, W. (2013a). A non-parametric method based on nbnn for automatic detection of liver lesion in ct images. In *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*, pages 366–369. IEEE.

Yang, W., Feng, Q., Huang, M., Lu, Z., and Chen, W. (2013b). A non-parametric method based on nbnn for automatic detection of liver lesion in ct images. In *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*, pages 366–369. IEEE.

Yang, X., Yu, H. C., Choi, Y., Lee, W., Wang, B., Yang, J., Hwang, H., Kim, J. H., Song, J., Cho, B. H., et al. (2014). A hybrid semi-automatic method for liver segmentation based on level-set methods using multiple seed points. *Computer methods and programs in biomedicine*, 113(1):69–79.

Ye, J., Sun, Y., Wang, S., Gu, L., Qian, L., and Xu, J. (2009). Multi-phase ct image based hepatic lesion diagnosis by svm. In *Biomedical Engineering and Informatics, 2009. BMEI'09. 2nd International Conference on*, pages 1–5. IEEE.

Yu, J., Cheng, Q., and Huang, H. (2004). Analysis of the weighting exponent in the fcm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(1):634–639.

Yushkevich, P. A., Piven, J., Hazlett, H. C., Smith, R. G., Ho, S., Gee, J. C., and Gerig, G. (2006). User-guided 3d active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128.

Zhang, L., Chen, J., Zhu, Y., and Luo, J. (2009a). Comparisons of several new denoising methods for medical images. In *Bioinformatics and Biomedical Engineering, 2009. ICBBE 2009. 3rd International Conference on*, pages 1–4. IEEE.

Zhang, X., Tian, J., Deng, K., Wu, Y., and Li, X. (2010a). Automatic liver segmentation using a statistical shape model with optimal surface detection. *IEEE Transactions on Biomedical Engineering*, 57(10):2622–2626.

Zhang, X.-Y., Diao, X.-F., Wang, T.-F., and Chen, S.-p. (2010b). Study on feature extraction for ultrasonic differentiation of liver space-occupying lesions. In *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on*, pages 1–4. IEEE.

Zhang, Y.-L., Yuan, L., Shen, F., and Wang, Y. (2009b). Hemorrhagic hepatic cysts mimicking biliary cystadenoma. *World J Gastroenterol*, 15(36):4601–3.

Zheng, S., Fang, B., Li, L., Gao, M., and Wang, Y. (2018). A variational approach to liver segmentation using statistics from multiple sources. *Physics in Medicine & Biology*, 63(2):025024.

Zhou, X. and Bhanu, B. (2008). Feature fusion of side face and gait for video-based human identification. *Pattern Recognition*, 41(3):778–795.

Zhou, X., Kitagawa, T., Hara, T., Fujita, H., Zhang, X., Yokoyama, R., Kondo, H., Kanematsu, M., and Hoshi, H. (2006). Constructing a probabilistic model for automated liver region segmentation using non-contrast x-ray torso ct images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 856–863. Springer.

Zhu, F., Carpenter, T., Gonzalez, D. R., Atkinson, M., and Wardlaw, J. (2012). Computed tomography perfusion imaging denoising using gaussian process regression. *Physics in Medicine & Biology*, 57(12):N183.

Zviniene, K. (2012). *Differential Diagnosis of Hepatocellular Carcinoma on Computed Tomography*. INTECH Open Access Publisher.